

ML & Análisis Cualitativo

Predicción Financiera y de Negocio

Uoc

David Sancho Marco

Máster en Inteligencia de
Negocios y Big Data Analytics

Tutor/a de TF

Rafael Luque Ocaña

**Profesor/a responsable de
la asignatura**

José Curto Díaz

Universitat Oberta
de Catalunya

06/01/2026



© (el autor/a)

Reservados todos los derechos. Está prohibido la reproducción total o parcial de esta obra por cualquier medio o procedimiento, comprendidos la impresión, la reprografía, el microfilme, el tratamiento informático o cualquier otro sistema, así como la distribución de ejemplares mediante alquiler y préstamo, sin la autorización escrita del autor o de los límites que autorice la Ley de Propiedad Intelectual.

Ficha del Trabajo Final

Título del trabajo:	Machine Learning y Análisis Cualitativo: Predicción Financiera y de Negocio
Nombre del autor/a:	David Sancho Marco
Nombre del Tutor/a de TF:	Rafael Luque Ocaña
Nombre del/de la PRA:	Jose Curto Díaz
Fecha de entrega:	01/2026
Titulación o programa:	Máster en Inteligencia de Negocios y Big Data Analytics
Área del Trabajo Final:	Inteligencia de Negocios
Idioma del trabajo:	Castellano
Palabras clave	Predicción, Aprendizaje Automático, Negocios
Resumen del Trabajo	
<p>Este trabajo desarrolla un sistema integral de análisis financiero y predicción del desempeño empresarial que sintetiza metodologías tradicionales con técnicas avanzadas de Machine Learning (ML). El estudio aborda limitaciones críticas del análisis convencional, como la fragmentación analítica en silos y el uso de proyecciones lineales simplistas, proponiendo un marco unificado que combina modelos cuantitativos con una evaluación cualitativa estructurada.</p> <p>La metodología emplea modelos de series temporales (ARIMA, Holt-Winters), algoritmos de gradient boosting (XGBoost) y arquitecturas de ensamble, validados mediante protocolos de backtesting (walk-forward) con datos históricos (informes 10-K) de 12 empresas cotizadas en EE. UU. El objetivo central del trabajo es evaluar si la integración del conocimiento de dominio (en torno a seis dimensiones estratégicas clave) mediante ajustes controlados a los modelos cuantitativos mejora la precisión predictiva respecto a los métodos tradicionales.</p>	

Los resultados indican que, si bien los modelos de ensamble (especialmente Linear-Holt-XGBoost) superan generalmente tanto a los modelos estadísticos individuales como a los tradicionales, la efectividad de los ajustes cualitativos es muy condicional. Se identifica un "*umbral de eficiencia*": los ajustes añaden valor significativo en empresas con alta incertidumbre estratégica o baja capitalización, pero pueden degradar la precisión en líderes de mercado estables donde la información ya está incorporada en las tendencias. El trabajo concluye que la sofisticación técnica debe equilibrarse con la interpretabilidad, proporcionando un marco reproducible para la toma de decisiones en campos como la consultoría estratégica, inteligencia de negocios e inversión.

Abstract

This work develops an integral financial analysis and business prediction system that synthesizes traditional financial techniques with advanced Machine Learning (ML) methods. Addressing the limitations of fragmented analytical silos and linear projections, the study proposes a unified framework combining quantitative forecasting (ARIMA, Holt-Winters, XGBoost, and ensemble models) with a novel qualitative scoring mechanism based on six key strategic dimensions, including management quality and market context.

The methodology employs rigorous walk-forward validation (backtesting) using historical financial data (10-K reports) from a diversified sample of 12 U.S. public companies. The study evaluates whether integrating domain-knowledge through qualitative adjustments improves the predictive accuracy of ML ensembles against traditional baselines.

Results indicate that while ML ensemble architectures (specifically Linear-Holt-XGBoost) generally outperform individual statistical models, the value of qualitative adjustments is highly conditional. The findings reveal an "inverted U-curve" of effectiveness: qualitative adjustments significantly reduce error rates for companies undergoing strategic transition or facing high uncertainty but can degrade accuracy for stable, high-efficiency market leaders where such factors are already priced in. The thesis concludes that while integrated forecasting models provide superior

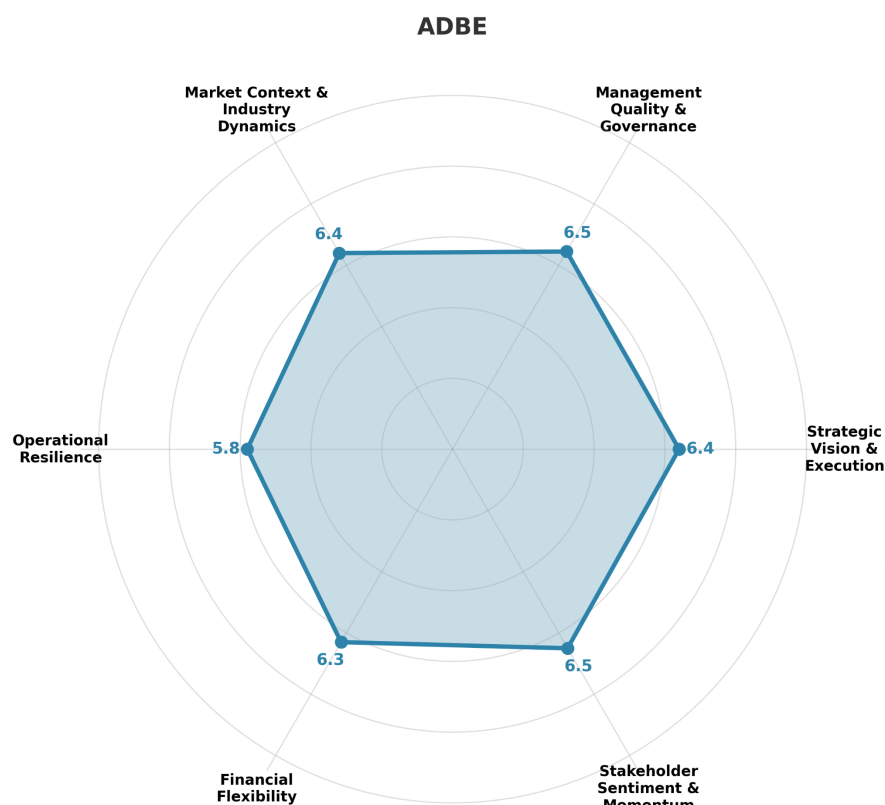
robustness, the application of these qualitative adjustments must be context-dependent. The work contributes a reproducible protocol for determining when to deploy complex integrated analysis versus parsimonious models in corporate finance, investment, equity research and business intelligence contexts.

Index

Introducción	1
Contexto, justificación y motivación del Trabajo	1
Objetivos del Trabajo	3
Enfoque y método seguido	4
Breve sumario de productos obtenidos	5
Breve descripción de los otros capítulos de la memoria	5
Estado de la cuestión	7
Metodología	19
Resultados	53
Discusión	62
Conclusiones y mejoras futuras	73
Bibliografía	90
Anexos	92



1. Introducció



Ejemplo de resultado - Adobe.

1.1. Contexto, Justificación y Motivación del Trabajo

En ocasiones, el análisis financiero tradicional fragmenta las evaluaciones empresariales en silos independientes (ratios de liquidez, proyecciones de crecimiento, análisis de riesgo, valoración, etc), poniendo obstáculos a la obtención de una visión integrada del desempeño organizacional. Esto puede llegar a reducir los esfuerzos analíticos y la toma de decisiones a puro instinto y en la práctica estar “*data-driven*” (basados en datos), sólo en nombre. En este proyecto me propuse desarrollar un **análisis financiero integral** que combinase las técnicas de análisis financiero y de inteligencia de negocios tradicionales con métodos avanzados de machine learning para la predicción de métricas clave como ingresos, balances, flujos de capital, etc.

El análisis consistió en un proceso dividido en cuatro módulos principales: Primero, un análisis de riesgos operacionales y financieros, segundo, una evaluación de salud financiera mediante ratios financieros, tercero, proyecciones de crecimiento y valoración

empresarial y cuarto y final, un análisis estratégico / macro. Los resultados incluirán cuadros de mando (dashboards) ejecutivos, gráficos de araña de desempeño por dimensión, proyecciones financieras a cinco años y dashboards de los resultados de evaluar comprehensivamente tales proyecciones.

La validación empírica se realizó mediante backtesting con estados financieros históricos públicos (informes 10-K, específica y principalmente), comparando la precisión predictiva de los modelos de Machine Learning relevantes (ARIMA, exponential smoothing, gradient boosting, Holt-Winter, etc) y de sus "*ensembles*", contra los resultados de emplear tasas y cotas de referencia tradicionales.

1.4 Justificación

En el contexto empresarial actual, la toma de decisiones estratégicas requiere una comprensión total del desempeño organizacional. Sin embargo, las herramientas de análisis financiero y de negocio disponibles, por sí solas, presentan algunas limitaciones significativas:

- **Fragmentación analítica:** Muchos sistemas actuales tratan dimensiones financieras de forma aislada, dificultando identificar relaciones causales (y correlaciones estadísticas) entre métricas de manera integral.
- **Métodos predictivos incompletos:** Las proyecciones financieras tradicionales parten de una extrapolación lineal o tasas de crecimiento medias constantes, a veces incluso obviando patrones no lineales y factores de estacionalidad en pro de simplicidad.
- **Falta de integración cuantitativa:** El análisis estratégico cualitativo (SWOT, Porter, análisis de ratios financieros, etc) raramente se traduce en métricas cuantificables que puedan integrarse en modelos financieros de manera operativa.

Este trabajo pretende abordar estas limitaciones comunes mediante:

1. **Integración holística:** Síntesis de indicadores financieros, operacionales, estratégicos y de ML en un modelo predictivo unificado.
2. **Aplicación de ML en finanzas:** Uso de modelos de ML en ensamble para capturar patrones complejos en series temporales financieras.
3. **Homogeneización del análisis avanzado:** La generación de un marco conceptual y de trabajo reproducible y aplicable a cualquier empresa con estados financieros estructurados.

La relevancia principal del trabajo, a mi parecer, radica en la evaluación empírica de si la integración comprehensiva de factores del "*domain-knowledge*" y el uso del aprendizaje automático mejoran significativamente la precisión predictiva comparado con métodos tradicionales, más cualitativos y fragmentarios.

1.5 Motivación

Mi motivación para desarrollar este TFM surge de la intersección entre mis aspiraciones profesionales y mis intereses académicos en el ámbito de inteligencia de negocios y el análisis financiero:

Investigaciones propias: Durante mis indagaciones, he observado cómo las empresas utilizan múltiples herramientas desconectadas para análisis financiero (Excel para ratios, software específico para forecasting, presentaciones separadas para análisis estratégico), generando potenciales ineficiencias y ofuscando una visión lo suficientemente clara, inequívoca e integral como para tomar decisiones estratégicas fuertes y resolutivas.

Interés personal: El máster de la UOC me ha proporcionado conocimientos en machine learning, análisis de series temporales y visualización de datos que deseo aplicar en problemas y contextos empresariales reales. La integración de técnicas estadísticas tradicionales con métodos de ML representa un desafío técnico estimulante, rico y prometedor.

Aplicación práctica: Este trabajo permitirá desarrollar un flujo de trabajo que puede aplicarse en consultorías financieras, departamentos de FP&A (*Financial Planning & Analysis*), o como herramienta de due diligence en procesos de inversión/adquisición.

Aportación académica: Previo al comienzo de este trabajo, ya había desarrollado componentes preliminares del sistema de análisis, componentes que deseo validar rigurosamente bajo estándares académicos, contribuyendo así de paso al conocimiento sobre aplicaciones de ML en finanzas corporativas e inteligencia de negocios.

1.2. Objetivos del Trabajo

1.2.1 Objetivo Principal:

Desarrollar, ejecutar y analizar los productos de un sistema integral de análisis financiero y predicción empresarial que combine técnicas tradicionales de análisis financiero e inteligencia de negocios con métodos avanzados de machine learning / aprendizaje automático.

1.2.2 Objetivos Específicos:

Integración holística: Sintetizar indicadores financieros, operacionales, estratégicos y técnicas de ML en un modelo predictivo unificado novel.

Aplicación de ML en finanzas: Utilizar modelos de Machine Learning (incluidos ensambles) para capturar patrones complejos en series temporales financieras.

Homogeneización del análisis: Generar un marco conceptual y de trabajo reproducible, generalizable y aplicable a cualquier empresa con estados financieros estructurados.

Validación empírica: Realizar backtesting con estados financieros históricos públicos (10-K reports, principalmente, pero no solamente) para comparar la precisión predictiva de modelos de ML contra métodos tradicionales y contra el ajuste cualitativo del marco.

Crear productos finales prácticos: Generar puntuaciones cualitativas finales por dimensión, dashboards ejecutivos, gráficos de araña, proyecciones financieras a 5 años, un sistema de evaluaciones y calibraciones empíricas de las mismas, etc.

1.3. Enfoque y método seguido

Si bien el uso de ML estaba siempre en mente, durante el desarrollo del trabajo, varias estrategias y aproximaciones posibles fueron consideradas:

1. Desarrollar herramientas aisladas: Crear uno o varios módulos independientes para una dimensión analítica específica (riesgo, ratios, pronósticos) para enriquecer el ML nítida y granularmente, concentrado sólo para ese aspecto del análisis.
2. Adaptar soluciones existentes: Implementar software comercial de análisis financiero tradicional y enriquecer los resultados con ML.
3. Desarrollar un sistema integrado nuevo: Crear un marco unificado desde cero, que combine análisis tradicional con ML de manera orgánica y dedicada.

Opté por la tercera opción desarrollar un sistema integrado nuevo que combinase:

- Análisis financiero multidimensional tradicional.
- Métodos de Machine Learning, tanto de modelo singular como por ensamble (ARIMA, Holt-Winters, XGBoost, Regresión Lineal).
- Marco de puntuación cualitativa (6 dimensiones: visión estratégica, gestión, contexto de mercado, resiliencia operacional, flexibilidad financiera, sentimiento).
- Protocolo de validación empírica riguroso (backtesting con walk-forward validation).

Pese a ser la más ambiciosa y suponer un riesgo en cuanto a carga de trabajo y límites de tiempo se refería, consideré que esta estrategia era la más apropiada porque:

1. Aborda las brechas identificadas: Considero que la literatura académica revela una tendencia hacia la fragmentación analítica y desconexión cuanti-cualitativa que las herramientas aisladas no resuelven.

2. Permite una validación empírica comprensiva: Un sistema nuevo, explícitamente dedicado, posibilita comparar rigurosamente ML ensemble vs. métodos tradicionales a múltiples niveles, lo cual es una forma atractiva y versátil de cumplir el objetivo central del trabajo.

3. Garantiza reproducibilidad y generalización: El marco desarrollado es aplicable sistemáticamente a múltiples empresas y sectores, por diseño.

4. Equilibra complejidad y practicidad: Integra sofisticación técnica (Machine Learning; precisamente aquellas habilidades que en el contexto de este máster deben ser demostradas activamente) con interpretabilidad (análisis financiero y empresarial tradicionales), esencial para la toma de decisiones empresariales reales.

La razón profunda (y que abarca a todas las anteriores), es que no me satisfacía trabajar en un ejercicio más de abstracción, de “laboratorio”. Llevaba ya un tiempo ensamblando este sistema y consideré que tocaba ya arriesgarme y llevarlo a un punto de producción real (Minimum Viable Product), con todas las vicisitudes, óbices y escollos que supone la puesta en práctica de cualquier tipo de empresa de este tipo.

1.4. Breve resumen de productos obtenidos

1. Sistema integrado de análisis financiero: modelo aproximación unificada que combina análisis tradicional con ML.

2. Dashboards: Visualizaciones a modo de mandos de control ejecutivos para toma de decisiones

3. Gráficos de Araña: Visualizaciones de las dimensiones, similar función que los dashboards pero en “macro”, más comprensivos.

4. Proyecciones financieras a 5 años: Modelos predictivos de métricas empresariales/financieras clave. Presentadas tanto en tablas como gráficos.

5. Modelos ML+ensambles: Algoritmos avanzados (ARIMA, XGBoost, Holt-Winters) para predicción de métricas financieras y empresariales. Muestras de código en anexos.

6. Protocolo de validación empírica adicional: Backtesting con datos históricos públicos (10-K reports)

El análisis de sensibilidad (con la consiguiente habilitación de una evaluación de escenarios, riesgos y planes de contingencia potenciales), quedó relegado como fuera del scope del trabajo en su iteración actual.

1.5. Breve descripción de otros capítulos de la memoria

Capítulo 2: Estado de la Cuestión: Revisa la literatura académica relevante y práctica actual en el análisis financiero tradicional, aprendizaje automático en predicción financiera, técnicas de pronóstico de series temporales y evaluaciones multidimensionales del

rendimiento empresarial. Identifica las brechas metodológicas que justifican el desarrollo del sistema integrado propuesto.

Capítulo 3: Metodología/Materiales y Métodos: Describe las fuentes de datos (EDGAR, 10-K reports), criterios de selección de empresas, proceso de extracción y preparación de datos, métricas financieras definidas, marco de puntuación cualitativa y protocolo de validación empírica con backtesting walk-forward.

Capítulo 4: Resultados. Presenta los hallazgos de la aplicación del sistema a 12 empresas de muestra diversa, comparando la precisión predictiva de modelos ML ensemble versus promedios tradicionales, análisis de la efectividad de ajustes cualitativos y validación estadística de mejoras en pronóstico.

Capítulo 5: Conclusiones y Futuros Trabajos: Sintetiza los hallazgos principales sobre la comparación de enfoques integrados ML-tradicional, identifica condiciones de efectividad de ajustes cualitativos y propone extensiones y mejoras futuras.

Capítulo 6: Glosario: Define términos técnicos específicos del dominio de finanzas corporativas y machine learning utilizados en el trabajo.

Capítulo 7: Bibliografía: Referencias académicas y profesionales citadas, incluyendo trabajos de Koller et al. (2020), Damodaran (2012), Chen & Guestrin (2016), Makridakis et al. (2018) y Amini et al. (2021), entre otros.

Capítulo 8: Anexos: Contiene criterios de selección de empresas, propuesta final de empresas muestra con razonamientos/justificaciones, una descripción más completa y granular del marco integrado de puntuación cualitativa, selecciones y muestras del código fuente (y configuraciones específicas de los modelos) y productos finales por empresa: cuadros de mando, gráficos de araña, visualización gráfica de pronósticos y visualización gráfica de la evaluación de pronósticos.

2. Estado de la Cuestión

Tradicionalmente, los métodos de análisis financiero y de predicción del desempeño empresarial han usado, fundamentalmente, combinaciones de ratios financieros y extrapolaciones lineales del rendimiento pasado. Con el desarrollo del campo, y el crecimiento en la sofisticación de sus métodos, formas de análisis y predicción más complejas han ganado preponderancia. Entre esta segunda tanda de métodos, podemos encontrar métodos de probabilística y estadística y, eventualmente, técnicas de análisis multidimensional, algoritmos y aprendizaje automático. En este estado de la cuestión, busco hacer un repaso y análisis del desarrollo más reciente y relevante de la literatura académica y la práctica profesional dentro de este contexto. El objetivo de dicho estado de la cuestión es fundamentar y mostrar la necesidad del desarrollo de un sistema integral de predicción financiera que combine el análisis financiero tradicional con métodos modernos de aprendizaje automático (Machine Learning o ML). En el proceso de hacer esto, veremos también ejemplos de sus múltiples usos, del desarrollo teórico del campo y de las mejores prácticas en su instrumentalización.

En este estado de la cuestión, he organizado la revisión de la literatura en cuatro secciones (por tema). La idea de esta estructura es facilitar el seguimiento de la tesis central y mantener un orden temático en cada sección. Estas secciones incluyen, en primer lugar, los principios del análisis financiero clásico; en segundo lugar, la utilización del aprendizaje automático en la predicción financiera; en tercer lugar, las técnicas de pronóstico; y en cuarto y último lugar, las evaluaciones multidimensionales del rendimiento empresarial.

2.1 Análisis Financiero Tradicional

2.1.1 Metodologías Clásicas de Análisis Financiero:

Durante mis indagaciones en los métodos y estudios sobre el análisis financiero tradicional, delimité tres pilares metodológicos. Estos han sido establecidos y codificados tras décadas de práctica profesional y académica, y considero que son los siguientes:

Análisis de ratios financieros: Palepu et al. (2013) muestran que el análisis de ratios constituye una de las dos principales herramientas del análisis financiero y proporciona la base para el análisis fundamental de las empresas. En este libro, los autores enfatizan que el análisis de ratios permite examinar el rendimiento y el estado financiero de una empresa en el contexto de sus objetivos y estrategia. Estas declaraciones, por tanto, pueden potencialmente servir como base sobre las que posteriormente diagnosticar y pronosticar el futuro rendimiento de la empresa. La descomposición del ROE en sus tres principales factores impulsores (margen de beneficio neto, rotación de activos y apalancamiento financiero) se puede profundizar aún más, mediante ratios adicionales que considerar individualmente, a modo de mayor indagación en los matices del funcionamiento financiero de la empresa. Con objeto de satisfacer los propósitos y objetivos que he delimitado para la presente investigación y trabajo, propongo organizar estos ratios en cinco dimensiones

principales: liquidez (ratio corriente, ratio rápido), solvencia (deuda/capital, cobertura de intereses), rentabilidad (ROE, ROA, márgenes), eficiencia operativa (rotación de activos, rotación de inventarios) y valoración (P/E, P/B, EV/EBITDA). Esta categorización (que por otro lado, dados los ejemplos que he visto, no es nada inusual) proporciona un marco estructurado para evaluar la salud financiera de la empresa desde múltiples perspectivas complementarias. Estas perspectivas incluyen la gestión operativa, la inversión, la financiación y la valoración.

Análisis de estados financieros: En Koller et al (2020) se explica cómo la evaluación integrada de los tres principales estados financieros ("Balance Sheet, Income Statement, Cash Flow Statement", es decir, Balance General, Estado de Resultados, Estado de Flujo de Efectivo) permite identificar relaciones causales entre decisiones operativas, estructura de capital y creación de valor. Los autores enfatizan la importancia de comprender las interdependencias entre estos estados: los cambios en el balance afectan al estado de resultados a través de la depreciación y gastos financieros, mientras que el estado de flujo de efectivo concilia ambos mediante ajustes por elementos no monetarios. Comprender estas relaciones proporciona una visión integrada de la empresa, necesaria tanto para entender y evaluar su modelo de negocio como para valorarla.

Valoración de empresas: Damodaran (2010, 2012), en sus libros, identifica tres influyentes enfoques para la práctica de valoración de empresas: valoración por flujo de caja descontado (DCF), valoración relativa mediante múltiplos comparables, y valoración por opciones reales para activos contingentes. Estos constituyen tres de los modelos de valoración de empresas más básicos, todos comúnmente utilizados (en la práctica del análisis financiero en general y de la valoración de empresas en particular) en la actualidad. Damodaran (2012) también presenta un marco comprehensivo y avanzado para enfoques probabilísticos en el Capítulo 33 ("*Probabilistic Approaches in Valuation: Scenario Analysis, Decision Trees, and Simulations*"). Este marco incluye análisis de escenarios, árboles de decisión y simulaciones de Monte Carlo. Estos métodos avanzados permiten una evaluación robusta de la incertidumbre en contextos donde los datos históricos son limitados o las condiciones de mercado son altamente volátiles. Si bien la investigación y trabajo que he llevado a cabo no han acabado hilando tan fino, son métodos a los que apunto y que en versiones más completas de este tipo de análisis deben ser empleadas y tenidas en cuenta.

2.1.2 Limitaciones del Análisis Financiero:

A pesar de su amplia adopción, diversos autores han identificado limitaciones significativas en estas metodologías tradicionales de análisis financiero que hemos discutido:

Fragmentación analítica: Koller et al (2020), por ejemplo, enfatizan la importancia de un análisis integrado en la valoración de empresas, en lugar de examinar componentes financieros de forma aislada. El uso de herramientas desconectadas, de hecho, dificulta mucho identificar patrones sistémicos (y hacerlo de modo que estos no sean redundantes o incluso ruido). Esta fragmentación puede resultar en análisis incompletos, sesgados y/o en

omisiones de interacciones críticas entre variables operativas, financieras y estratégicas. Koller al escribir este texto como una guía de McKinsey. Y en esta guía lo que se promueve es un enfoque integral, holístico, que tenga en cuenta las interrelaciones entre diferentes elementos financieros y estratégicos de la empresa. Precisamente, lo que yo argumento es que una práctica de la valoración empresarial rigurosa requiere comprender cómo estos factores interactúan dentro del sistema empresarial (y sectorial) completo en el que se dan.

Proyecciones simplistas: Damodaran (2012), por otro lado, menciona que las proyecciones financieras tradicionales típicamente asumen tasas de crecimiento constantes o tendencias lineales, por simplicidad. Lo que podemos responder a esto es que hacerlo así no da un peso suficiente a patrones de estacionalidad, ciclicidad y no-linealidad inherentes a muchos sectores o modelos de negocio. Estas suposiciones y/o simplificaciones pueden generar proyecciones sistemáticamente sesgadas, especialmente en cuanto a horizontes temporales largos o en cuanto a empresas de sectores especialmente volátiles se refiere. La misma existencia de técnicas de aproximación (Capítulo 11 de dicho libro) y enfoques probabilísticos (Capítulo 33) como técnicas "*avanzadas*" en un libro de análisis financiero tradicional como este es reconocer tácitamente que los métodos tradicionales son, a menudo, insuficientes.

Exceso de subjetividad en el análisis cualitativo: Palepu et al. (2013) dedican el Capítulo 2 del libro casi por entero al análisis estratégico como punto de partida básico y fundamental para el análisis financiero de empresas. Estos autores ponen el énfasis en que comprender la estrategia de una empresa (lo cual incluye análisis de la industria a través de las "*Cinco Fuerzas*" de Porter, estrategia competitiva (liderazgo en costos vs. diferenciación y ventajas competitivas) permite al analista financiero contextualizar su análisis y la contabilidad con el contexto de la realidad empresarial en la que se manifiesta. Este enfoque estratégico y cualitativo sirve de complemento y sinergia al análisis más cuantitativo de ratios y flujos de efectivo. Ambas cosas juntas proporcionan una base integral y sólida para evaluar el rendimiento de la empresa y sus perspectivas de desarrollo futuras. Ambos enfoques tienen sus críticos (por ejemplo, en Grundy, T. 2006, donde se mencionan matices y bifurcaciones entre análisis y predicción). La conclusión a la que llego es que no abordar la desconexión entre el análisis estratégico cualitativo y los modelos cuantitativos limita la capacidad de los analistas para incorporar factores estratégicos en las proyecciones financieras de manera consistente, rigurosa, reproducible y generalizable.

2.2 Aprendizaje Automático en la Predicción Financiera

2.2.1 Fundamentos Teóricos y Aplicabilidad:

La aplicación del aprendizaje automático / machine learning en las finanzas corporativas no es nueva. No obstante, sí ha experimentado un crecimiento exponencial en las últimas décadas, particularmente conforme estas técnicas han ganado preponderancia y los especialistas en datos han aumentado en número. Este crecimiento ha sido impulsado

tanto por avances algorítmicos como por una mayor disponibilidad de datos estructurados como por mayor conocimiento de las posibilidades y por ende de la demanda. Un par de ejemplos respecto a sus usos concretos en finanzas y las ventajas comparativas de dichos usos pueden ser ilustrativos:

Capacidad para capturar no linealidades: Chen y Guestrin (2016) demuestran que los algoritmos de gradient boosting, particularmente XGBoost, pueden capturar relaciones complejas no lineales entre variables predictoras y objetivos sin requerir la especificación explícita de la forma funcional. Los autores de este artículo explican que estos modelos superan de forma consistente a los métodos lineales en de regresión y clasificación. Esto parece ser especialmente así cuando existen interacciones complejas y/o no lineales entre variables.

Aprendizaje de patrones históricos: Huang et al. (2005), por su parte, afirman que los métodos de aprendizaje automático (ML), particularmente las Máquinas de Vectores de Soporte (SVM), pueden identificar patrones sutiles en datos históricos que el análisis estadístico tradicional escapa. En dicho artículo, Huang et al investigaron sobre los métodos de predicción de la dirección de movimientos en el mercado bursátil. Y los resultados de tal investigación mostraron que las SVM con kernels no lineales capturaban estructuras de dependencia temporal complejas de manera más fiable y eficiente que otros métodos, particularmente aquellos sin elementos de aprendizaje automático. Muy relevante para nuestros objetivos en este trabajo.

2.2.2 Aplicaciones Específicas en Finanzas Corporativas:

En la literatura podemos encontrar documentadas diversas aplicaciones del aprendizaje automático, tanto exitosas como no exitosas:

Makridakis et al. (2018), analizaron los resultados de la competición M3 (2000), de modelos de series temporales (una de las también llamadas Competiciones Makridakis o "*M-Competitions*") que fue el benchmark más comprehensivo de series temporales disponible en aquel momento (con 3.003 unidades). Las conclusiones de Makridakis y otros investigadores revelan que los métodos estadísticos tradicionales superaron a los métodos de aprendizaje automático en este caso en medidas de precisión y horizontes de predicción. Específicamente, concluyen que las combinaciones de métodos proporcionan ganancias de precisión limitadas, con mayor costo computacional, mientras que los métodos de aprendizaje automático requieren considerablemente más recursos computacionales sin mejoras significativas de precisión. Sin embargo, el artículo reclama pruebas más objetivas a través de competiciones a mayor escala. A pesar del enorme tamaño (e influencia) de este estudio, no se aplicó específicamente a series temporales financieras o empresariales, por lo que debemos profundizar más para llegar a conclusiones.

Cao y Tay (2003), por otro lado, muestran la aplicabilidad de la Máquina de Vectores de Soporte (SVM) para la predicción de series financieras y gestión del riesgo. Lo hacen

utilizando un enfoque de parámetros adaptativos (ASVM) que incorpora la característica no estacionariedad de los mercados financieros a través de un esquema de ponderación estático aplicado durante el entrenamiento. Este método asigna mayor importancia a los datos más recientes a través de un parámetro de regularización creciente (C) y una tolerancia al error decreciente (ϵ), resultando en mejor capacidad predictiva, soluciones más dispersas y rendimiento superior comparado con las redes neuronales tradicionales en la predicción de contratos de futuros del Chicago Mercantile Market.

Bao (2020) es otro ejemplo, que aborda las limitaciones de los métodos estadísticos tradicionales para detectar patrones complejos de fraude en datos financieros corporativos a través de la aplicación de técnicas de aprendizaje automático a datos financieros corporativos. El objetivo del estudio es desarrollar y evaluar enfoques de aprendizaje automático, incluyendo SVM (Máquina de Vectores de Soporte que hemos visto antes) y métodos de ensamble. Los resultados demuestran que los métodos ML superan significativamente a los enfoques tradicionales en esta tarea de finanzas corporativas.

Finalmente, Amini et al. (2021) es quizás el ejemplo más interesante para nuestro objetivo de trabajo, ya que es el más similar de los estudios tratados: estudiaron si el aprendizaje automático puede mejorar la comprensión de las dinámicas de estructura de capital, abordando las limitaciones de los modelos lineales tradicionales para capturar relaciones complejas no lineales entre el apalancamiento y sus determinantes. El estudio aplica diversas técnicas de aprendizaje automático, incluyendo bosques aleatorios, máquinas de gradient boosting, redes neuronales y LASSO, para predecir el apalancamiento objetivo utilizando un conjunto de datos de 128.417 observaciones empresa-año del período 1972-2018. Los resultados demuestran mejoras significativas sobre los modelos convencionales. Los bosques aleatorios identificaron un conjunto más amplio de determinantes del apalancamiento, incluyendo ratio mercado/libro, apalancamiento sectorial mediano, efectivo, Z-Score, rentabilidad, rentabilidad de acciones y tamaño de empresa. Las predicciones mejoradas del objetivo de apalancamiento generaron velocidades más rápidas de ajuste de estructura de capital y proyecciones más precisas de acciones de financiación corporativas en mercados de deuda y capital, proporcionando un soporte empírico más fuerte para las teorías de estructura de capital que los enfoques más tradicionales.

Vemos que existen numerosos ejemplos con diversos resultados mostrando que en promedio el aprendizaje automático mejora en contextos específicos de finanzas corporativas, pero no necesariamente cuando se aplica de manera general a todos los casos. La cuestión de en qué tareas de finanzas corporativas y análisis empresarial brilla la predicción con aprendizaje automático, por tanto, permanece actual y abierta.

2.2.3 Desafíos y Consideraciones Metodológicas:

La literatura identifica dificultades claras para la aplicación del ML a la predicción financiera.

Sobreentrenamiento/Sobreajuste: Makridakis et al. (2018) registran que los modelos complejos de aprendizaje automático (incluyendo redes neuronales) frecuentemente confunden patrones con ruido, exhibiendo excelente rendimiento en datos de entrenamiento pero generalizando pobremente a predicciones futuras (el llamado “*overfitting*” o sobre-entrenamiento). Los autores enfatizan que el ajuste del modelo no es un buen predictor de la precisión post-muestra. Asimismo, sugieren emplear modelos más simples y parsimoniosos cuando los datos históricos son limitados, evitando así el sobreajuste mediante criterios como AIC y métodos estadísticos que mantienen reducida complejidad computacional.

Interpretabilidad vs Precisión: Chen y Guestrin (2016) apuntan que existe una tensión en estos modelos entre precisión predictiva e interpretabilidad: mientras que los modelos ensemble y gradient boosting típicamente ofrecen mejor precisión, su naturaleza de “caja negra” complica la transparencia y explicación de las predicciones. Los autores proponen técnicas de interpretabilidad como valores SHAP para mitigar este problema: estas técnicas permiten identificar la contribución de cada variable a predicciones específicas.

Requisitos de Datos: Huang et al. (2005) apuntan que los métodos efectivos de aprendizaje automático requieren cantidades sustanciales de datos históricos de calidad. Para empresas con historias financieras cortas o datos inconsistentes, los beneficios del ML sobre los métodos tradicionales pueden ser limitados, o incluso contraproducentes.

2.3 Técnicas de Pronóstico de Series Temporales

2.3.1 Métodos Estadísticos Clásicos:

Los métodos de pronóstico/predicción estadísticos se han ido refinando durante décadas y para los propósitos de este trabajo constituyen el patrón oro establecido contra los que se ha de evaluar cualquier nuevo enfoque. Ilustro y ejemplifico dos de ellos (aunque emplearemos más), como contexto:

Modelos ARIMA: Hyndman y Athanasopoulos (2021) presentan ARIMA (*AutoRegressive Integrated Moving Average*) como la familia de modelos más versátil para series temporales univariadas. Estos autores explican que ARIMA captura patrones de autocorrelación (componente AR), tendencias mediante diferenciación (componente I), y choques aleatorios (componente MA). Su flexibilidad permite modelar una amplia gama de procesos estocásticos, aunque requiere series razonablemente largas (mínimo 50-100 observaciones) y estacionarias o fácilmente estacionarizables.

Suavizado Exponencial: Hyndman y Athanasopoulos (2021) describen la familia de métodos de suavizado exponencial (ETS: Error-Tendencia-Estacionalidad) como particularmente apropiada para series con patrones de tendencia y estacionalidad. Destacan el método **Holt-Winters** (el cual usaremos) como una extensión que simultáneamente captura nivel, tendencia y estacionalidad mediante ponderación

exponencial. Los autores demuestran que, en sus muestras, ETS frecuentemente supera a ARIMA en horizontes de predicción cortos (1-6 períodos) para series con estacionalidad pronunciada. No debe entenderse con esto que un método es universalmente superior al otro, ya que sus desempeños (como es el caso con todos estos tipos de modelos) son contextuales.

2.3.2 Enfoques Híbridos, de Conjuntos y de Ensamble:

La combinación de múltiples métodos de pronóstico es una estrategia robusta y en auge para mejorar la precisión predictiva. Es una tendencia emergente que para nuestros propósitos podemos ilustrar con varios ejemplos:

Ya en el ámbito de la predicción de quiebra corporativa, Barboza, Kimura y Altman (2017) realizaron un estudio comparativo comprehensivo entre modelos (o más bien, reglas) tradicionales (como el Z-Score de Altman) y diversos métodos de aprendizaje automático. Los modelos empleados en el estudio incluyen técnicas de ensamble con Random Forest y Gradient Boosting. Los resultados del estudio concluyen que los métodos ensamble, particularmente el Random Forest, muestran una precisión significativamente superior en la predicción de situaciones de insolvencia hasta dos años por adelantado, superando tanto a los modelos clásicos (el mencionado Z-Score de Altman, entre otros) como a los algoritmos ML individuales (como redes neuronales o máquinas de vectores de soporte (SVM)). Estos resultados sirven de ejemplo de la capacidad y del potencial de los modelos ensamble para capturar patrones complejos y no-lineales en los datos financieros. Son, por tanto, demostrablemente potentes en tareas de finanzas especializadas, como lo es la agregación de múltiples predictores (por ejemplo, para reducir el error de generalización y mejorar la solidez en contextos de alta incertidumbre) en la detección temprana de quiebras empresariales.

Naturalmente, no todos los estudios o resultados han sido positivos para los modelos de ensamble: Makridakis et al. (2018), (ya mencionados arriba, también) evalúan el rendimiento de una simple combinación de métodos de suavizado exponencial (SES, Holt y amortiguado) y concluyen que las ganancias de precisión de combinar múltiples pronósticos fueron insignificantes comparadas con métodos individuales, mientras que la complejidad computacional fue casi doble. Aunque prueban un promedio simple de tres métodos estadísticos tradicionales, sus resultados muestran que esta combinación no supera consistentemente a los mejores métodos de pronóstico individuales y no recomiendan ampliamente su uso debido al costo adicional sin beneficios claros.

Finalmente, Henrique, B. M., Sobreiro, V. A., & Kimura, H. (2019), realizaron una revisión general de la literatura sobre la aplicación del aprendizaje automático en la predicción financiera. Como revisión de literatura, no presenta un único experimento, sino que sintetiza los hallazgos de docenas de estudios sobre predicción de mercados financieros. Esto es muy útil para contextualizar el presente trabajo. La conclusión clave del artículo identifica una tendencia clara: los estudios más recientes y mejor performantes favorecen abrumadoramente el uso de métodos de ensamble, de conjunto y otros enfoques híbridos.

Henrique et al. destacan que modelos como Random Forest y enfoques híbridos que combinan, por ejemplo, análisis de sentimiento (mediante técnicas de procesamiento de lenguaje natural, o NLP) con modelos de series temporales, están sentando el estándar de precisión en el campo.

2.3.3 Validación, Evaluación y Selección de Modelos:

La literatura establece protocolos rigurosos para evaluar y comparar los resultados de diferentes métodos de pronóstico.

- **Métricas de error apropiadas:** Hyndman y Athanasopoulos (2021) argumentan que la selección de métricas de error debe considerar el contexto de aplicación. Para series financieras, específicamente, recomiendan:
- **MAPE (Mean Absolute Percentage Error):** Puede interpretarse como error porcentual promedio, apropiado cuando la magnitud del error relativo es más relevante que el absoluto.
- **RMSE (Root Mean Squared Error):** Penaliza fuertemente los errores grandes, apropiado cuando las desviaciones extremas son particularmente costosas y/o graves.
- **MAE (Mean Absolute Error):** Robusto a valores atípicos, apropiado cuando la distribución del error puede tener colas pesadas.
- **Validación temporal:** Estos mismos autores enfatizan que la validación de modelos de series temporales debe respetar la estructura temporal de los datos. Recomiendan la "validación de origen móvil" (también llamada validación walk-forward): entrenar el modelo con datos hasta el tiempo t , predecir $t+1$ hasta $t+h$, avanzar el origen un período, y repetir. Este proceso simula realísticamente cómo se utilizarían los modelos en producción.
- **Pruebas de comparación estadística:** Hyndman y Athanasopoulos (2021) también presentan el test de Diebold-Mariano como el método estándar para determinar si las diferencias en precisión predictiva entre dos modelos son estadísticamente significativas. Este test considera no solo las magnitudes promedio de error, sino también su correlación y varianza, proporcionando inferencia rigurosa sobre la superioridad relativa de métodos.

2.4 Marcos de Evaluación Multidimensional del Rendimiento

2.4.1 Balanced Scorecard y Perspectivas Integradas:

El concepto de evaluación multidimensional del rendimiento organizacional ha sido ampliamente desarrollado en la literatura de estrategia y control de gestión.

Balanced Scorecard: Kaplan y Norton (1996) estos autores introducen y acuñaron el término y el concepto del "*Balanced Scorecard*" como un sistema de monitorización estratégica y de control en su funcionamiento y puesta en práctica. El Balanced

Scoreboard original integra cuatro perspectivas: financiera, de clientes, de procesos internos, y de aprendizaje/crecimiento. Lo que Kaplan y Norton propusieron es que esta evaluación multidimensional proporciona una visión más completa del rendimiento organizacional que sólo usando indicadores financieros aislados: captura aspectos estratégicos cualitativos y anticipa la trayectoria de los factores creadores del valor de la empresa a futuro. Este posicionamiento y comprensión del largo plazo no puede no reflejarse en los resultados financieros por sí solos.

Traducción a métricas cuantificables: Kaplan y Norton (2001) en sus obras y trabajos posteriores también desarrollaron metodologías para traducir objetivos estratégicos cualitativos en métricas cuantificables y establecer relaciones causa-efecto entre indicadores de diferentes perspectivas. Este nuevo enfoque permite construir "mapas estratégicos" que vinculen las acciones operativas de organizaciones y empresas con resultados financieros específicos.

2.4.2 Sistemas de Puntuación y Agregación:

La construcción de sistemas de puntuación que agregan múltiples métricas heterogéneas presenta desafíos metodológicos significativos.

Nardo et al. (2005) del Centro Común de Investigación de la Comisión Europea proporcionan una guía comprehensiva sobre la construcción de indicadores compuestos. Los autores recomiendan:

- **Normalización:** Transformación de variables a escalas comparables mediante z-scores, escalado min-max, o ranking. La elección depende de la distribución de datos y la sensibilidad deseada a valores atípicos
- **Ponderación:** Asignación de pesos de importancia relativa a las dimensiones. Los métodos incluyen ponderación igual (cuando no hay teoría que justifique diferencias), análisis factorial (pesos derivados de correlaciones), o ponderación experta

Estos autores advierten sobre problemas de agregación cuando las dimensiones son conceptualmente distintas: los promedios simples pueden ocultar compensaciones importantes. Por tanto, recomiendan realizar análisis de sensibilidad para evaluar cómo las variaciones en los esquemas de ponderación afectan rankings y conclusiones, guiando metodologías de agregación robustas.

2.4.3 Integración de Análisis Cualitativo y Cuantitativo:

En la literatura encontramos ejemplos de la complejidad y matices de integrar sistemáticamente insights estratégicos cualitativos en modelos cuantitativos.

Análisis estratégico estructurado: Porter (1980) establece sus famosas cinco fuerzas competitivas como marco conceptual y herramienta fundamental para el análisis sistemático de la competencia en empresas, analizando rivalidad entre competidores

existentes, amenaza de nuevos entrantes, poder de negociación de proveedores y compradores, y amenaza de productos sustitutos. Complementariamente, en su trabajo de 1985, Porter desarrolla el análisis de la cadena de valor para examinar actividades generadoras de valor en una empresa. Como apuntan Palepu et al. (2013), estos marcos conceptuales permiten al analista sondear la economía de una empresa a nivel cualitativo, proporcionando una base sólida para posteriormente contextualizar el análisis contable y financiero en la realidad empresarial. Sin embargo, la forma en que estos ajustes tienen lugar permanece variable y discrecional.

Otros puentes cuanti-cualitativos: Algunos autores sí proponen metodologías más específicas para cuantificar factores cualitativos. Por ejemplo, Damodaran (2012) sugiere ajustar tasas de descuento basadas en evaluaciones cualitativas de riesgo competitivo (1), o la práctica de incorporar análisis de escenarios (optimista/base/pesimista) que reflejen incertidumbres identificadas estratégicamente de manera cualitativa (2) de forma detallada y estandarizada.

2.5 Síntesis y Brechas Identificadas

2.5.1 Revisión de la Literatura Académica:

La revisión de literatura revela un cuerpo sustancial aunque heterogéneo de conocimientos en cada área temática abordada. En el campo del análisis financiero tradicional, existen metodologías bien establecidas con décadas de validación práctica, aunque están limitadas por fragmentación analítica y el uso de suposiciones simplificadoras que reducen su capacidad predictiva en contextos complejos. Respecto al aprendizaje automático aplicado a finanzas, la evidencia empírica muestra resultados ambivalentes pero prometedores en aplicaciones específicas como predicción de precios de activos y scoring crediticio, con evidencia moderada pero creciente de que los modelos ensemble superan tanto a predictores individuales como a la simple extrapolación del promedio histórico. En el campo del pronóstico de series temporales, existen marcos metodológicos robustos para validación y comparación, con consenso académico sobre la superioridad de los ensembles en benchmarks comprehensivos como las competencias M3 y M4. Finalmente, en términos de evaluación corporativa multidimensional, aunque existen conceptos teóricos sólidos como el Balanced Scorecard o indicadores compuestos, sus implementaciones prácticas a menudo resultan ad-hoc, raramente estandarizadas y escasamente homogéneas, lo que dificulta su generalización sistemática y reproducibilidad metodológica.

2.5.2 Brechas Identificadas:

A pesar del desarrollo en áreas individuales, persisten ciertas cuestiones significativas por abordar que este trabajo pretende abordar de manera más explícita. Se podría objetar al uso del término "*brecha*" aquí (ciertamente muchos otros han tratado estos temas antes, y con mayor profundidad), pero se usa convenientemente para destacar aquellos aspectos

de la literatura que en mi opinión retienen mayor interés exploratorio. Sería quizás más preciso hablar de "*áreas de integración*" u "*oportunidades de síntesis*", pero "*brechas*" es más parsimonioso.

Brecha 1: Integración holística: La literatura carece de marcos que integren sistemáticamente múltiples dimensiones analíticas (riesgo, salud financiera, crecimiento, estrategia) con métodos predictivos avanzados. Los trabajos existentes típicamente se centran en dimensiones aisladas o métodos específicos, pero no en sistemas comprensivos.

Brecha 2: Validación empírica de modelos de ensamble en finanzas corporativas: Mientras que la competición M3-M4 valida ensembles en series temporales generales, su aplicación específica a estados financieros corporativos (con sus interdependencias entre Estado de Resultados, Balance General y Flujo de Efectivo) no se explora significativamente en literatura académica que haya podido encontrar.

Brecha 3: Operacionalización del análisis cualitativo: Persiste una desconexión entre el análisis estratégico cualitativo y los modelos cuantitativos. Se necesitan metodologías reproducibles para traducir insights cualitativos en ajustes paramétricos sistemáticos.

Brecha 4: Análisis de sensibilidad integrada: La literatura sobre análisis de sensibilidad en valoración (ej. Damodaran, 2012) y generación probabilística de escenarios (simulaciones de Monte Carlo) permanecen típicamente separadas. Un intento de unificación de ambas perspectivas en un marco conceptual y operacional coherente beneficiaría.

2.5.3 Posicionamiento:

Este TFM se posiciona estratégicamente para abordar las brechas identificadas a través de cuatro contribuciones fundamentales. Primero, desarrolla un sistema integrado que combina análisis financiero multidimensional con métodos ensemble de aprendizaje automático, equilibrando precisión predictiva con interpretabilidad y transparencia analítica. Segunda, implementa validación empírica rigurosa mediante backtesting con estados financieros públicos, comparando el rendimiento de ensembles ML contra baselines tradicionales y extrapolaciones del promedio histórico utilizando pruebas estadísticas apropiadas como el test de Diebold-Mariano. Tercera, establece un marco metodológico reproducible que puede aplicarse sistemáticamente a diferentes empresas, sectores y contextos temporales, contribuyendo significativamente al campo de Business Intelligence con aplicación específica en finanzas corporativas. Finalmente, integra análisis de sensibilidad y evaluación de escenarios dentro de una estructura analítica coherente que facilita la toma de decisiones informada bajo condiciones de incertidumbre, un aspecto frecuentemente descuidado en enfoques puramente algorítmicos.

2.6 Conclusiones:

Considero que la revisión de literatura realizada apoya sólidamente la viabilidad y relevancia del proyecto propuesto. Primero, se han identificado fundamentos teóricos y prácticos robustos provenientes de metodologías de análisis financiero tradicional, técnicas de aprendizaje automático y marcos de pronóstico de series temporales que constituyen la base sobre la que desarrollar el sistema integrado. Segundo, la evidencia empírica de competiciones de ciencia de datos y estudios académicos sobre predicción de quiebra corporativa apoya la hipótesis central de que ciertas combinaciones de métodos, particularmente modelos de ensamble, superan situacionalmente tanto a los modelos predictores individuales como a la simple extrapolación del promedio histórico. Tercero, se han detectado las "*brechas*" que merecen más exploración en la práctica actual, caracterizadas por fragmentación analítica y desconexión entre análisis cualitativo y cuantitativo, lo que justifica el desarrollo de un enfoque más integrado y holístico. Finalmente, aunque los fundamentos conceptuales son sólidos, la necesidad de validación empírica específica para predicción de indicadores financieros corporativos constituye precisamente la contribución diferencial y el valor añadido del presente trabajo.

3. Metodología

3.1 Fuentes de Datos y Criterios de Selección

Los datos financieros que utilicé para este trabajo provienen del sistema EDGAR (*Electronic Data Gathering, Analysis, and Retrieval*) de la Comisión de Valores y Bolsa de EE.UU. EDGAR da acceso gratuito a todas las presentaciones y documentos difundidos al público de las empresas que cotizan en bolsa en los Estados Unidos. No faltaron datos entre aquellos que extraje de EDGAR, pues el que existan y estén actualizados son ambas condiciones necesarias y requeridas por ley para todas las empresas de los EE.UU.

Tipos de Presentaciones Clave:

Presentación	Descripción	Frecuencia	Uso en Este Trabajo
10-K	Informe anual con estados financieros auditados	Anual	Fuente primaria de datos
10-Q	Informe trimestral con financieros no auditados	Trimestral	Validación suplementaria
8-K	Informe actual de eventos materiales	Según sea necesario	Contexto para anomalías

El uso de EDGAR presenta un gran número de ventajas para nuestros propósitos: desde acceso conveniente a informes estandarizados bajo el sistema US GAAP (auditados por firmas de contabilidad independientes), hasta el etiquetado XBRL estructurado (que nos permite una extracción automatizada con un código de python de los datos financieros), no pasando por alto que nos sirve como garantía de tener un histórico completo (típicamente 10+ años) de los datos, sin costes ni restricciones de acceso.

También supone una serie de limitaciones: EDGAR sólo cubre empresas que cotizan en los Estados Unidos, los datos anuales limitan el tamaño de muestra para modelado de series temporales y el retraso en reportes (10-K presentado 60-90 días después del fin del año fiscal) significó que en algunas empresas nos tenemos que conformar con el año 2024 (ergo, pone el análisis en un patrón de año irregular).

Fuentes de Datos Suplementarias

Fuente	Tipo de Datos	Uso
Morningstar	Valoraciones, calificaciones de foso, benchmarks financieros	Entradas para puntuación cualitativa

Moody's Orbis	Calificaciones crediticias, benchmarks financieros	Entradas para puntuación cualitativa
Consenso de Analistas	Objetivos de precio, recomendaciones	Componente de sentimiento del marco cualitativo
Relaciones con Inversionistas de la Empresa	Transcripciones de ganancias, guía	Contexto estratégico para evaluación cualitativa
Bases de Datos Macroeconómicas	PIB, tasas de interés, índices sectoriales	Variables contextuales (no entradas directas del modelo)

Criterios de Selección de Empresas

Para asegurar la calidad de datos y aplicabilidad del modelo, las empresas deben satisfacer los siguientes criterios de inclusión.

Requisitos Mínimos

Criterio	Umbral	Justificación
Historia Financiera	≥10 años de presentaciones 10-K completas	Observaciones suficientes para modelado de series temporales
Compleitud de Datos	Todas las métricas de Nivel 1 disponibles	Permite marco completo de validación
Consistencia en Reportes	Sin reestados o cambios contables mayores	Evita rupturas estructurales en series
Continuidad Operativa	Sin quiebra o salida de cotización durante el período de estudio	Asegura relevancia de pronósticos

Objetivos de Diversificación

Para apoyar la generalización de hallazgos, la muestra de empresas que seleccioné buscaba diversidad en:

- **Capitalización de Mercado:** Representación de pequeñas, medianas y grandes empresas.
- **Sector:** Múltiples sectores (clasificación GICS: Consumo, Salud, Tecnología, Industriales, etc.)
- **Modelo de Negocio:** “*Ligero*” en activos vs. “*pesado*” en activos; ie, B2B vs. B2C.
- **Perfil de Crecimiento:** Empresas de alto crecimiento (tirando a jóvenes) vs. crecimiento moderado (tirando a maduras).

Muestra Actual

El marco de validación ha sido aplicado a las siguientes empresas:

Empresa	Ticker	Sector	Capitalización	Años de Datos	Año de Validación
Adobe, Inc.	ADBE	Tecnología de la Información	Grande	2015-2024	2024
Bruker Corporation	BRKR	Salud/Ciencias de la Vida	Mediana	2015-2024	2024
Crocs, Inc.	CROX	Consumo Discrecional	Mediana	2015-2025	2025
Idexx Laboratories, Inc.	IDXX	Salud/Ciencias de la Vida	Grande	2015-2024	2024
Intuit, Inc.	INTU	Tecnología de la Información	Grande	2015-2025	2025
Lincoln Electric Holdings, Inc.	LECO	Industriales	Mediana	2015-2024	2024
Neogen Corporation	NEOG	Salud/Ciencias de la Vida	Pequeña	2015-2025	2025
NIKE, Inc.	NKE	Consumo Discrecional	Grande	2015-2025	2025
Qualys, Inc.	QLYS	Tecnología de la Información	Mediana	2015-2024	2024
Red River Bancshares, Inc.	RRX	Servicios Financieros	Mediana	2015-2024	2024
Starbucks Corporation	SBUX	Consumo Básico	Grande	2015-2025	2025
Woodward, Inc.	WWD	Industriales	Mediana	2015-2024	2024

Respecto a las limitaciones de esta muestra, baste decir que la muestra actual de doce empresas, aunque ambiciosa para los límites de tiempo establecidos, resulta aún limitada y esto pone restricciones en la capacidad de generalización que podamos realizar. Los

resultados deben interpretarse como evidencia exploratoria que muestra la necesidad de una futura investigación adicional y más avanzada con una muestra expandida y generalizable.

Extracción y Preparación de Datos

Proceso de Extracción:

1. **Recuperación automatizada** de presentaciones 10-K de EDGAR usando Python (`sec_edgar_downloader.py`)
2. **Análisis XBRL** para extraer ítems estandarizados de estados financieros
3. **Verificación manual** contra financieros publicados en presentaciones de inversores
4. **Normalización** de nombres de cuentas a través de años fiscales (reclasificaciones)

Control de Calidad:

Verificación	Descripción	Acción
Identidad del Balance General	Activos = Pasivos + Patrimonio	Marcar si desviación > \$1M
Conciliación de Efectivo	Δ Efectivo coincide con total del Estado de CF	Marca para revisión
Continuidad Año tras Año	Final del año anterior = Inicio del año actual	Identificar re-estados
Detección de Valores Atípicos	Valores > 3σ de la media histórica	Revisión manual

Limitaciones de Datos:

- Frecuencia anual: Restringe la complejidad y precisión de los modelos.
- *Heterogeneidad en reportes: Las definiciones de ítems varían entre empresas; suelen requerir estandarización y normalización.
- Sesgo de supervivencia: La muestra excluye empresas fallidas; esto puede llevar a sobrestimar la precisión de los pronósticos para firmas en peligro de quiebra.

3.2 Definiciones de Métricas Financieras e Interdependencias

Panorama General:

La previsión financiera opera dentro de un sistema de identidades contables y relaciones económicas que restringen el espacio de predicciones válidas. A diferencia de las series

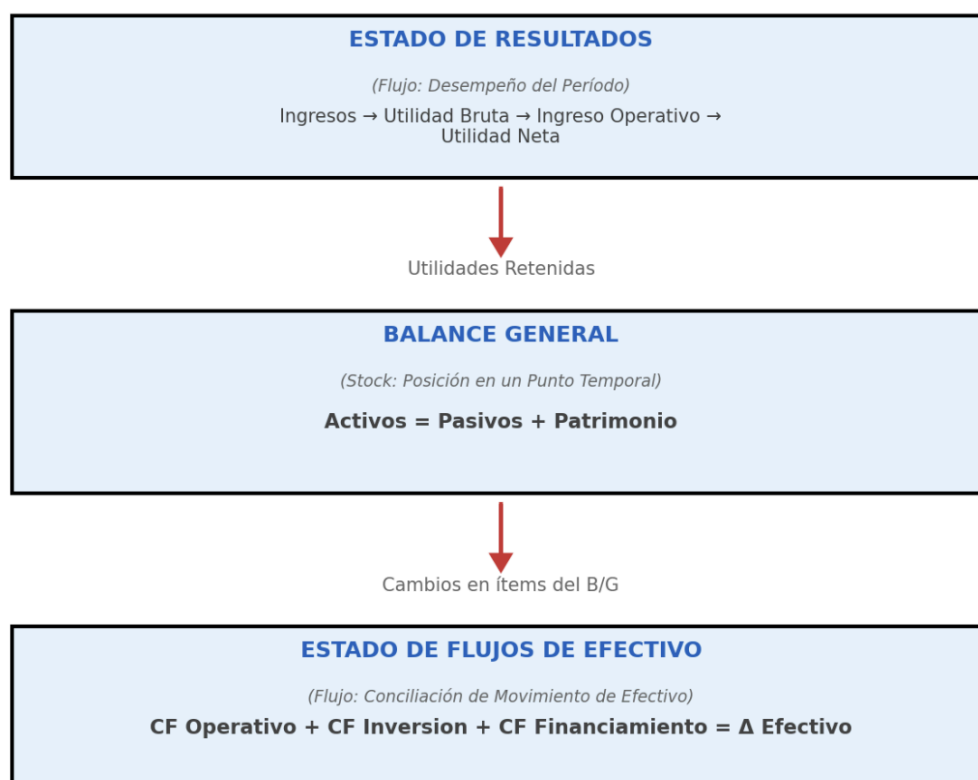
temporales independientes, los tres estados financieros principales (*Estado de Resultados*, *Balance General* y *Estado de Flujos de Efectivo*) están matemáticamente vinculados. Un sistema de previsión que ignora estas interdependencias corre el riesgo de producir proyecciones internamente inconsistentes (por ejemplo, pronosticar crecimiento de ingresos sin una expansión correspondiente de activos para soportarlo).

Esta sección define las métricas financieras utilizadas en la canalización de previsión y explica las relaciones estructurales que gobiernan su comportamiento.

Arquitectura de los Tres Estados Financieros

Marco Conceptual

Los tres estados financieros capturan vistas diferentes pero interconectadas del desempeño corporativo:



Estado Financiero	Tipo	Características Principales	Ecuación Principal	Conexión con Otros Estados
Estado de Resultados	Flujo	Desempeño del período	Ingresos → Utilidad Bruta → Ingreso Operativo → Utilidad Neta	↓ Utilidades Retenidas → Balance General
Balance General	Stock	Posición en un punto temporal	Activos = Pasivos + Patrimonio	↑ Cambios en Ítems → Estado de Flujos de Efectivo
Estado de Flujos de Efectivo	Flujo	Conciliación de movimiento de efectivo	CF Operativo + CF Inversión + CF Financiamiento = Δ Efectivo	← Cambios del Balance General

La idea clave aquí es que el estado de flujos de efectivo sirve como puente que concilia el desempeño basado del estado de resultados con los cambios en la posición del balance general. Cualquier enfoque de previsión debe respetar esta conciliación.

Métricas del Estado de Resultados

Definiciones:

Métrica	Definición	Fórmula
Ingresos Totales	Entradas brutas de operaciones comerciales principales	Ventas + Ingresos por Servicios + Otros Ingresos Operativos
Costo de Ventas (COGS)	Costos directos atribuibles a producción	Materiales + Mano de Obra Directa + Gastos Generales de Manufactura
Utilidad Bruta	Ingresos menos costos directos	Ingresos – COGS
Gastos Operativos	Costos indirectos de operar el negocio	G&A + I+D + Depreciación y Amortización
Ingreso Operativo (EBIT)	Ganancia de operaciones principales antes de intereses/impuestos	Utilidad Bruta – Gastos Operativos
EBITDA	Ingreso operativo antes de cargos no monetarios	EBIT + Depreciación + Amortización
Gasto por Intereses	Costo del financiamiento con deuda	$f(\text{Nivel de Deuda, Tasas de Interés})$
Utilidad Neta	Ganancia de fondo atribuible a	EBT – Gasto por Impuesto a

Relaciones Estructurales

El Estado de Resultados sigue una estructura en cascada donde cada ítem de línea depende de los anteriores.

$$\begin{aligned}
 & \text{Ingresos} \\
 & - \text{COGS} \\
 & \hline
 & = \text{Utilidad Bruta} \\
 & - \text{Gastos Operativos} \\
 & \hline
 & = \text{Ingreso Operativo (EBIT)} \\
 & - \text{Gasto por Intereses} \\
 & + \text{Otros Ingresos/Gastos} \\
 & \hline
 & = \text{Ganancias Antes de Impuestos (EBT)} \\
 & - \text{Impuesto a la Renta} \\
 & \hline
 & = \text{Utilidad Neta}
 \end{aligned}$$

La implicación de esto para los intentos de previsión es que los ingresos son típicamente el impulsor principal. Por tanto conviene pronosticar ingresos primero y luego aplicar supuestos de margen (Margen Bruto %, Margen Operativo %) para sacar métricas derivadas mientras se asegura la consistencia interna.

Ratios y Márgenes Clave

Ratio	Fórmula	Rango Típico	Interpretación
Margen Bruto	Utilidad Bruta / Ingresos	20-80% (dependiente de industria)	Poder de fijación de precios y eficiencia de producción
Margen Operativo	Ingreso Operativo / Ingresos	5-30%	Eficiencia operacional
Margen Neto	Utilidad Neta / Ingresos	2-20%	Rentabilidad general
Margen EBITDA	EBITDA / Ingresos	10-40%	Capacidad generadora de efectivo de operaciones

Métricas del Balance General

Definiciones:

Métrica	Definición	Componentes
Activos Corrientes	Activos convertibles a efectivo dentro de un año	Efectivo + Cuentas por Cobrar + Inventario + Gastos Pagados por Anticipado
Activos No Corrientes	Activos a largo plazo	PP&E + Intangibles + Inversiones a Largo Plazo
Activos Totales	Todo lo que la empresa posee	Activos Corrientes + Activos No Corrientes
Pasivos Corrientes	Obligaciones vencibles dentro de un año	Cuentas por Pagar + Deuda a Corto Plazo + Gastos Acumulados
Pasivos No Corrientes	Obligaciones a largo plazo	Deuda a Largo Plazo + Impuestos Diferidos + Obligaciones de Pensión
Pasivos Totales	Todo lo que la empresa debe	Pasivos Corrientes + Pasivos No Corrientes
Patrimonio Total	Reclamo residual de accionistas	Capital Social + Utilidades Retenidas + Otras Ganancias Comprensivas
Capital de Trabajo Neto	Amortiguador de liquidez operacional	Activos Corrientes – Pasivos Corrientes

La Identidad Contable Fundamental:

El Balance General debe siempre satisfacer:

$$\text{Activos Totales} = \text{Pasivos Totales} + \text{Patrimonio Total}$$

Lo que implica esto es que esta identidad es una restricción dura. Los pronósticos para Activos, Pasivos y Patrimonio no pueden generarse independientemente—al menos uno debe derivarse como residual para asegurar que se respeta el balance total.

Dinámicas del Capital de Trabajo:

El Capital de Trabajo Neto (NWC) captura los requerimientos de efectivo del ciclo operativo:

$$\text{NWC} = (\text{Cuentas por Cobrar} + \text{Inventario}) - (\text{Cuentas por Pagar} + \text{Gastos Acumulados}) \\ \approx \text{Activos Corrientes} - \text{Pasivos Corrientes}$$

NWC típicamente escala con ingresos. Las empresas a menudo expresan necesidades de capital de trabajo como porcentaje de ingresos:

$$\text{NWC}_{\text{pronóstico}} = \text{NWC}_{\text{como \% de Ingresos}} \times \text{Ingresos}_{\text{pronóstico}}$$

Esto supone una cierta dificultad a la hora de hacer previsiones. NWC puede ser altamente volátil y sujeto a decisiones de gerencia (acumulación/liquidación de inventario, negociaciones de términos de pago). Esto explica los errores extremos de pronóstico para esta métrica han de ser esperados.

Métricas del Estado de Flujos de Efectivo

Definiciones:

Métrica	Definición	Componentes Clave
Flujo de Efectivo Operativo (OCF)	Efectivo generado de negocio principal	Utilidad Neta + Ajustes no monetarios + Cambios en Capital de Trabajo
Flujo de Efectivo de Inversión (ICF)	Efectivo usado para inversiones a largo plazo	CapEx + Adquisiciones – Ventas de Activos
Flujo de Efectivo de Financiamiento (FCF)	Efectivo de/desde proveedores de capital	Emisión/repago de deuda + Emisión/recompra de acciones + Dividendos
Flujo de Efectivo Libre (FCF)	Efectivo disponible para todos proveedores de capital	OCF – CapEx
Flujo de Efectivo Libre al Patrimonio	Efectivo disponible para accionistas	FCF – Intereses – Repagos de Deuda + Nueva Deuda

La Identidad de Conciliación de Efectivo

El estado de flujos de efectivo debe conciliarse con el balance general:

$$\text{Efectivo Inicial} + \text{OCF} + \text{ICF} + \text{FCF} = \text{Efectivo Final}$$

Donde el efectivo final aparece en el balance general.

Derivación del Flujo de Efectivo Operativo (Método Indirecto):

OCF puentea el Estado de Resultados (devengo) con generación real de efectivo:

$$\begin{aligned}
 & \text{Utilidad Neta} \\
 & + \text{Depreciación y Amortización} \quad (\text{gasto no monetario}) \\
 & + \text{Compensación Basada en Acciones} \quad (\text{gasto no monetario}) \\
 & \quad - \text{Aumento en Cuentas por Cobrar} \quad (\text{efectivo atado}) \\
 & \quad - \text{Aumento en Inventario} \quad (\text{efectivo atado}) \\
 & + \text{Aumento en Cuentas por Pagar} \quad (\text{efectivo preservado}) \\
 & + \text{Otros ajustes} \\
 & \hline
 & = \text{Flujo de Efectivo Operativo}
 \end{aligned}$$

En cuanto a nuestras previsiones respecta, OCF depende tanto de métricas del Estado de Resultados (Utilidad Neta, D&A) como de cambios del Balance General (Δ Cuentas por Cobrar, Δ Inventario, Δ Cuentas por Pagar). Pronosticar OCF requiere pronosticar estos componentes subyacentes primero.

Interdependencias Cruzadas entre Estados

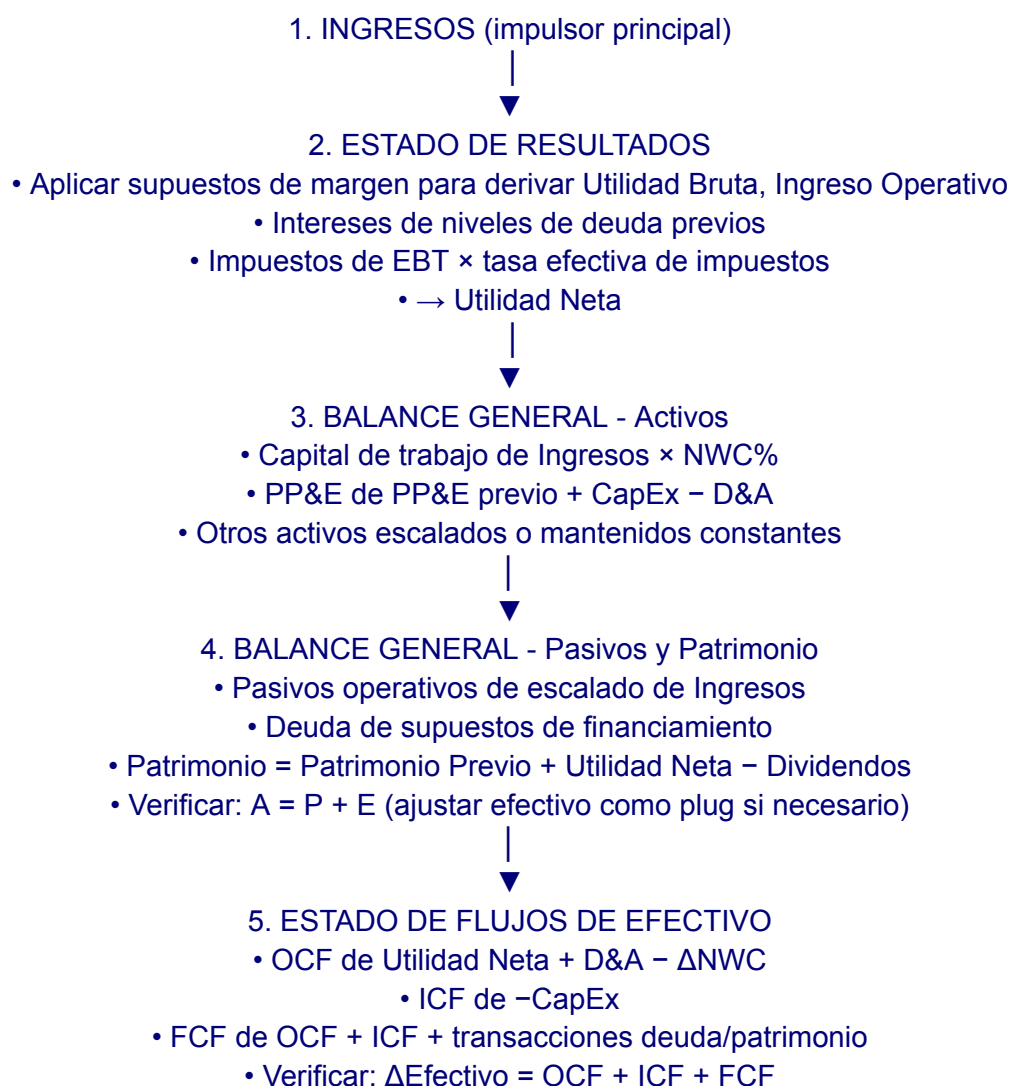
Vínculos Principales:

Los tres estados están conectados a través de varias relaciones críticas:

Vínculo	Descripción	Tratamiento en Previsión
Utilidades Retenidas	Utilidad Neta fluye al Patrimonio	$\text{Patrimonio}_t = \text{Patrimonio}_{t-1} + \text{Utilidad_Neta}_t - \text{Dividendos}_t$
Depreciación	Reduce PP&E; aparece en E/R y OCF	$\text{D\&A} = f(\text{PP\&E}, \text{vidas útiles})$; consistente en todos los estados
CapEx	Aumenta PP&E; aparece en ICF	$\text{PP\&E}_t = \text{PP\&E}_{t-1} + \text{CapEx}_t - \text{D\&A}_t - \text{Desinversiones}_t$
Cambios en Deuda	Afecta B/G y FCF; Intereses afectan E/R	$\text{Intereses}_t = f(\text{Deuda}_{t-1}, \text{tasa de interés})$
Capital de Trabajo	Cambios del B/G afectan OCF	ΔNWC reduce OCF cuando positivo

La Cascada de Previsión:

Para mantener consistencia, los pronósticos deben seguir una secuencia lógica:



Enfoque de Implementación

Independencia de Métricas vs. Previsión de Sistema

He considerado dos enfoques principales para controlar las interdependencias:

Enfoque A: Previsión Independiente de Métricas:

La idea en este primer enfoque era pronosticar cada métrica como serie temporal separada. Esto es más rápido, más simple, permite una selección de modelo específico por cada métrica. El problema son las inconsistencias internas entre estados pronosticados.

Enfoque B: Previsión de Sistema:

La segunda posibilidad era pronosticar métricas impulsoras (Ingresos, Márgenes, CapEx) y entonces derivar métricas dependientes a través de relaciones contables. Esto asegura la

consistencia; ie, refleja la estructura económica. El riesgo aquí es la propagación de error: las relaciones mal especificadas se “contagian” entre operaciones.

Enfoque C: Híbrido:

La que he terminado adoptando: una canalización de previsión que empleaba un híbrido pragmático:

1. Métricas primarias pronosticadas independientemente: Ingresos, Gastos Operativos, Activos Totales, Pasivos Totales.
2. Métricas derivadas calculadas de relaciones: Utilidad Bruta (de Ingresos \times MG%), Utilidad Neta (cascada), Patrimonio (de identidad contable)
3. Verificaciones de validación aplicadas: Marcar pronósticos que violen identidades contables o produzcan ratios implausibles

Verificaciones de Consistencia

La validación post-pronóstico asegura coherencia interna:

Verificación	Condición	Acción si Violada
Identidad del Balance General	Activos = Pasivos + Patrimonio	Ajustar Patrimonio como residual
Conciliación de Efectivo	$\Delta \text{Efectivo} = \text{OCF} + \text{ICF} + \text{FCF}$	Marcar para revisión manual
Límites de Márgen	$0\% < \text{Margen Bruto} < 100\%$	Limitar/superior a rango histórico
Signo del Capital de Trabajo	NWC consistente con modelo de negocio	Revisar si cambia el signo
Cobertura de Deuda	Cobertura de Intereses $> 1\times$	Marcar distress potencial

Selección de Métricas para Validación

Según la disponibilidad de datos, interpretabilidad y relevancia comercial, incluí las siguientes métricas en el marco de validación:

Nivel 1 (Crítico)

Métrica	Estado	Justificación
Ingresos Totales	E/R	Impulsor principal de valor
Ingreso Operativo	E/R	Desempeño operacional principal
Utilidad Neta	E/R	Rentabilidad de fondo

Flujo de Efectivo Operativo	CF	Capacidad de generación de efectivo
Flujo de Efectivo Libre	CF	Fundación de valoración (DCF)
Activos Totales	B/G	Escala y base de inversión
Patrimonio Total	B/G	Proxy de valor para accionistas

Nivel 2 (Importante)

Métrica	Estado	Justificación
Utilidad Bruta	E/R	Eficiencia de fijación de precios/producción
EBITDA	E/R	Proxy de efectivo; comparable entre firmas
Activos Corrientes	B/G	Posición de liquidez
Pasivos Corrientes	B/G	Obligaciones a corto plazo
Capital de Trabajo Neto	B/G	Necesidades de efectivo operacional

Nivel 3 (Soporte)

Métrica	Estado	Justificación
Gastos Operativos	E/R	Estructura de costos
Activos No Corrientes	B/G	Inversiones a largo plazo
Pasivos No Corrientes	B/G	Financiamiento a largo plazo

Dificultades y Mitigaciones

Dificultad 1: Volatilidad del Capital de Trabajo

NWC está influenciado por decisiones operativas (gestión de inventario, términos de pago) que pueden no seguir patrones históricos. Para mitigarlo, opté por expresar NWC como % de Ingresos con reversión a la media y marcar pronósticos donde NWC/Ingresos se desvía $>2\sigma$ de media histórica.

Dificultad 2: Ítems No Recurrentes

Los cargos únicos (reestructuración, deterioros, ganancias/pérdidas) distorsionan los patrones históricos. Para paliar esto, usé métricas "normalizadas" o "ajustadas" dónde fuese posible e intenté identificar y excluir años atípicos del entrenamiento. También, en ocasiones EBITDA fue priorizado sobre Utilidad Neta como más fiable.

Dificultad 3: Cambios en Políticas Contables

Si se producen cambios en el reconocimiento de ingresos, contabilidad de arrendamientos (ASC 842), o reportes de segmentos, se crean discontinuidades. Para intentar paliar esta problema, realicé una revisión de notas al pie de 10-K para cambios de política, intenté re-escribir los datos históricos donde fuese necesario y, si no quedase más opción, me propuse acortar ventanas de entrenamiento (si hay un ruptura estructural identificada en los datos a causa de este problema).

Dificultad 4: Fusiones y Adquisiciones (M&A)

Este tipo de operaciones crean saltos repentinos y no orgánicos en los datos financieros. Pueden ser de tal magnitud que desvirtúan y desbaratan toda validez en la serie. Intenté no hacer cambios donde fuera posible, pero no descarté hacer ajustes pro forma donde estas operaciones hubieran tenido lugar, por ejemplo excluyendo años donde tuvieran lugar para un más preciso cálculo de las tendencias. La puntuación cualitativa ciertamente busca capturar los impactos esperado de una operación de M&A en las empresas.

3.3 Modelos Base de Previsión

Panorama General:

La estructura de previsión emplea cinco modelos individuales y tres combinaciones de ensamble, seleccionados para capturar diferentes aspectos del comportamiento de series temporales financieras. Este enfoque multi-modelo parte de que ningún método de previsión único domina en todas las métricas y compañías, lo cual es una noción consistente con la literatura de competiciones M3/M4 (Makridakis et al., 2018).

Modelos Individuales

1. Pronóstico Naïve (Benchmark):

El pronóstico naïve (*ingenuo*) sirve, primero, como representativo de las prácticas de análisis financiero tradicionales y, segundo, como base/control de rendimiento mínimo que cualquier enfoque sofisticado debe superar para justificar su coste computacional y de complejidad añadida.

Especificación:

$$\hat{y}(t+h) = y(t)$$

Donde $y(t)$ es el último valor observado y h es el horizonte de previsión.

Las métricas financieras a menudo exhiben fuerte persistencia. Si una compañía generó \$500M en ingresos el año pasado, \$500M es una conjetura razonable para el próximo año. Cualquier modelo que no pueda vencer este benchmark proporciona un valor negativo.

La aplicación es sencilla y directa: valor observado final a todos los períodos de previsión.

2. Regresión Lineal (Extrapolación de Tendencia):

La regresión lineal captura tendencias seculares en métricas financieras, asumiendo que crecimiento o declive sigue un camino determinista lineal.

Especificación:

$$y(t) = \beta_0 + \beta_1 \cdot t + \varepsilon(t)$$

$$\hat{y}(t+h) = \beta_0 + \beta_1 \cdot (t+h)$$

Muchas métricas financieras (ingresos, activos, patrimonio) exhiben crecimiento aproximadamente lineal sobre horizontes de mediano plazo. La regresión lineal proporciona coeficientes de tendencia interpretables y sirve como punto medio entre una persistencia naïve y modelos más complejos.

Respecto a la implementación, se realizó con una Regresión de Mínimos Cuadrados Ordinarios (OLS) usando `scikit-learn.LinearRegression`, con índice de tiempo como único predictor. Para pronósticos multi-año, el modelo extrapola la tendencia ajustada.

A notar: la regresión lineal no puede capturar dinámicas no lineales, estacionalidad o reversion a la media; todas características comunes en los datos financieros.

3. ARIMA (Promedio Móvil Integrado Autorregresivo):

Los modelos ARIMA capturan estructuras de autocorrelación y tendencias estocásticas, representando en esta selección la punta de lanza estándar del pronóstico clásico de series temporales.

Especificación:

$$\text{ARIMA}(p, d, q):$$

$$\phi(B)(1-B)^d \cdot y(t) = \theta(B) \cdot \varepsilon(t)$$

Donde:

- p = orden autorregresivo (número de observaciones rezagadas).
- d = orden de diferenciación (para estacionariedad).
- q = orden de promedio móvil (errores de previsión rezagados).

- `B` = operador de retroceso.
- $\phi(B)$, $\theta(B)$ = polinomios característicos.

Respecto a la selección de modelo, automática de orden vía `pmdarima.auto_arima` usando minimización AIC (Criterio de Información de Akaike) con:

- Rango `p`: 0-3.
- Rango `d`: 0-2 (con prueba de estacionariedad KPSS).
- Rango `q`: 0-3.
- Componentes estacionales deshabilitados (datos anuales).

ARIMA captura comportamiento de reversión a la media común en ratios y márgenes financieros, mientras que el componente integrado maneja series no estacionarias como ingresos o activos.

Implementación: `statsmodels.tsa.arima.model.ARIMA` con parámetros seleccionados vía `pmdarima.auto_arima`. Pronósticos generados usando el método `.forecast(h)`.

4. Suavizado Exponencial Holt-Winters:

Los métodos de suavizado exponencial ponderan observaciones recientes de forma más “pesada” que un modelo ARIMA, adaptándose rápidamente a cambios de nivel mientras mantienen estimaciones de tendencia.

Especificación (Tendencia Aditiva, Sin Estacionalidad):

$$\begin{aligned}\text{Nivel: } L(t) &= \alpha \cdot y(t) + (1-\alpha) \cdot [L(t-1) + T(t-1)] \\ \text{Tendencia: } T(t) &= \beta \cdot [L(t) - L(t-1)] + (1-\beta) \cdot T(t-1) \\ \text{Pronóstico: } \hat{y}(t+h) &= L(t) + h \cdot T(t)\end{aligned}$$

Donde:

- α = parámetro de suavizado de nivel ($0 < \alpha < 1$)
- β = parámetro de suavizado de tendencia ($0 < \beta < 1$)

Variante de Tendencia Amortiguada: Para métricas que exhiben desaceleración de crecimiento, se aplica un parámetro de amortiguamiento ϕ ($0 < \phi < 1$):

$$\hat{y}(t+h) = L(t) + (\phi + \phi^2 + \dots + \phi^h) \cdot T(t)$$

Holt-Winters sobresale en pronósticos de horizonte corto a mediano (1-3 años) donde las dinámicas recientes son más predictivas que la historia de la serie a largo plazo. La variante de tendencia amortiguada previene extrapolaciones exponenciales irrealistas.

Respecto a la implementación: `statsmodels.tsa.holtwinters.ExponentialSmoothing` con optimización automática de parámetros vía máxima verosimilitud. Selección de modelo

entre tendencia aditiva/multiplicativa y variantes amortiguadas/no amortiguadas basada en AIC.

5. XGBoost (Árboles Potenciados por Gradiente):

XGBoost representa el enfoque de aprendizaje automático más habitual y versátil, capaz de capturar relaciones no lineales e interacciones complejas entre características.

Especificación:

$$\hat{y} = \sum_{k=1}^K f_k(x)$$

Donde cada `f_k` es un árbol de regresión, entrenado secuencialmente para minimizar errores residuales de árboles previos.

Respecto a la ingeniería de parámetros("features"), dado el tamaño limitado de muestra de datos financieros anuales, resulta crítica:

Tipo de Característica	Ejemplos
Valores rezagados	y(t-1), y(t-2), y(t-3)
Tasas de crecimiento	$\Delta y(t-1)$, $\Delta y(t-2)$
Estadísticas móviles	MA(3), std móvil
Ratios entre métricas	Ingresos/Activos, Deuda/Patrimonio
Características temporales	Índice de año, indicador de tendencia

Hiperparámetros:

- `n_estimators`: 100-500 (parada temprana con paciencia=10)
- `max_depth`: 3-5 (árboles superficiales para prevenir sobreajuste)
- `learning_rate`: 0.05-0.1
- `subsample`: 0.8
- `colsample_bytree`: 0.8
- `reg_alpha` (L1): 0.1
- `reg_lambda` (L2): 1.0

XGBoost puede capturar dinámicas no lineales e interacciones que los modelos estadísticos pierden de vista. Sin embargo, con historia financiera limitada (típicamente 10-15 observaciones anuales), regularización y árboles superficiales son esenciales para prevenir el sobreajuste.

En cuanto a la implementación, `xgboost.XGBRegressor` con ajuste de hiperparámetros vía validación cruzada de series temporales. Para forecasting, el modelo predice secuencialmente, usando valores predichos como características para períodos subsiguientes.

Se tuvo en cuenta que la alta varianza con muestras pequeñas hace al modelo propenso a sobre ajustes sin regularización cuidadosa; y menos interpretable que modelos estadísticos. Precisamente en cuanto a la interpretabilidad, valga decir que mientras XGBoost opera más como "caja negra" que los modelos lineales, los valores SHAP (SHapley Additive exPlanations) pueden computarse para explicar predicciones individuales cuantificando la contribución de cada característica. Esto aborda el *trade-off* interpretabilidad-precisión, pues permite entender qué patrones históricos o relaciones entre métricas impulsan los pronósticos específicos y en qué medida.

Modelos de Ensemble

Los métodos de ensemble combinan múltiples pronósticos para reducir varianza y mejorar robustez. Se evalúan tres configuraciones de ensemble:

1. Ensemble Lineal-Holt:

Combinación: Regresión Lineal + Holt-Winter

Ponderación: Promedio simple (50/50)

Combina captura de tendencia determinista (Lineal) con ajuste de nivel adaptativo (Holt-Winters). Apropiado cuando las métricas exhiben tendencias estables con cambios de nivel ocasionales.

2. Ensemble Lineal-Holt-XGBoost:

Combinación: Regresión Lineal + Holt-Winters + XGBoost

Ponderación: Pesos iguales (33/33/33)

Agrega capacidad de modelado no lineal a la fundación Lineal-Holt. XGBoost puede capturar cambios de régimen y dependencias entre métricas que modelos estadísticos pierden.

3. Ensemble Lineal-Holt-ARIMA:

Combinación: Regresión Lineal + Holt-Winters + ARIMA

Ponderación: Pesos iguales (33/33/33)

Ensemble puramente estadístico combinando extrapolación de tendencia (Lineal), suavizado adaptativo (Holt-Winters) y modelado de autocorrelación (ARIMA). Proporciona rendimiento robusto sin complejidad de aprendizaje automático.

Estrategia de Ponderación:

Todos los ensembles usan ponderación simétrica en lugar de pesos optimizados. Esta elección de diseño refleja:

1. *Limitación de Muestra Pequeña*: La estimación óptima de pesos requiere historia sustancial fuera de muestra; con 10 años de datos anuales, la optimización de pesos arriesga sobreajuste.
2. *Evidencia Empírica*: Makridakis et al. (2018) demuestran que promediado simple a menudo iguala o excede esquemas de combinación sofisticados en competiciones de forecasting.
3. *Robustez*: La ponderación igual asegura que ningún modelo domine, proporcionando seguro contra fallas específicas de modelo.

Con muestra expandida de compañías, podría explorarse ponderación basada en varianza inversa o regresión.

Selección de Modelo por Tipo de Métrica:

Diferentes métricas financieras exhiben diferentes propiedades de series temporales, sugiriendo preferencias de modelo específicas por métrica:

Tipo de Métrica	Comportamiento Típico	Modelos Preferidos
Ingresos	Crecimiento con momentum	Lineal, Holt-Winters, Ensembles
Márgenes Operativos	Reversión a la media	ARIMA, Holt-Winters
Capital de Trabajo	Volátil, dependiente de ratio	XGBoost, ARIMA
Flujo de Efectivo	Alta varianza, depende de E/R y B/G	Ensembles
Activos Totales	Crecimiento con persistencia	Lineal, Naïve
Patrimonio Total	Crecimiento con persistencia	Lineal, Naïve

En iteraciones futuras, podríamos considerar el uso de modelos ensemble híbridos, que usan modelos específicos solo para las métricas particulares donde funcionan mejor.

El marco de validación evalúa todos los modelos en todas las métricas, permitiendo identificación empírica de emparejamientos óptimos modelo-métrica.

3.4 Marco de Puntuación Cualitativa

Panorama General:

El marco de puntuación cualitativa representa un enfoque híbrido que combina análisis financiero automatizado con juicio experto y fuentes de inteligencia externa. En lugar de depender únicamente de datos financieros históricos, este marco integra múltiples dimensiones cualitativas para ajustar pronósticos cuantitativos, abordando la limitación de que modelos tradicionales a menudo pierden puntos de inflexión impulsados por factores estratégicos, gerenciales o de mercado.

Este marco aborda directamente una brecha persistente identificada en la literatura: la desconexión entre análisis estratégico cualitativo y modelos de previsión cuantitativos. Palepu et al. (2013) enfatizan el análisis estratégico como punto de partida fundamental para la evaluación financiera, sin embargo la traducción de insights cualitativos a ajustes de pronóstico permanece "variable y discrecional" en práctica. El presente marco operacionaliza esta traducción a través de una metodología sistemática, reproducible.

Fundamentos Teóricos:

El marco sintetiza insights de análisis financiero, gestión estratégica e investigación de previsión, fundamentando cada componente establecido en la teoría.

1. Integración del Análisis Estratégico:

Palepu et al. (2013) dedican atención extensiva al análisis estratégico como punto de partida fundamental para la evaluación financiera. Las *Cinco Fuerzas de Porter* (1980, 1985) y análisis de cadena de valor proporcionan herramientas conceptuales para evaluación cualitativa, sin embargo su traducción a ajustes cuantitativos ha carecido históricamente de estandarización.

La respuesta del marco aplicado es que las seis dimensiones de puntuación operacionalizan constructos estratégicos de Porter sistemáticamente: *Visión Estratégica & Ejecución* captura posicionamiento competitivo y efectividad de cadena de valor, *Contexto de Mercado & Dinámicas de Industria* sistematiza análisis de Cinco Fuerzas, *Resiliencia Operacional* refleja la robustez de cadena de suministro y adaptabilidad de modelo de negocio.

2. Puenteando la Brecha Cualitativo-Cuantitativo:

Damodaran (2012) sugiere incorporar factores cualitativos a través de tasas de descuento ajustadas y análisis de escenarios. Grundy (2006) destaca "*bifurcaciones persistentes entre análisis financiero y predicción*" que los analistas tratan de resolver.

En el presente marco de trabajo, el mecanismo de ajuste multiplicativo formaliza estos enfoques: Primero, los cuntajes cualitativos mayores de 5 (sobre 10) aumentan los pronósticos de crecimiento; menores de 5 los disminuyen. Segundo, los parámetros de sensibilidad varían por métrica, reflejando sensibilidad diferencial a estrategia. Todo esto proporciona una operacionalización estandarizada y unificada donde la literatura previa ofrece solo guías conceptuales fragmentarias.

3. Evaluación del Rendimiento Multidimensional:

El Balanced Scorecard (Kaplan & Norton, 1996, 2001) trasciende métricas puramente financieras integrando perspectivas de cliente, proceso y aprendizaje. El manual de Nardo et al. (2005) sobre indicadores compuestos proporciona guía metodológica sobre normalización, ponderación y agregación.

El marco de este trabajo toma medidas correspondientes a estas lecciones de la literatura académica previa: las seis dimensiones de puntuación siguen la filosofía multi-perspectiva del Balanced Scorecard, la escala de normalización 0-10 mantiene un criterio de medición comparable e intuitivo y la ponderación explícita ataja la preocupación de que los promedios simples opacan importantes trade-offs y granularidades en los valores.

4. Limitaciones de Datos en Previsión Financiera:

Los métodos de aprendizaje automático requieren datos históricos sustanciales (Huang et al., 2005) y arriesgan sobreajuste cuando observaciones son limitadas (Makridakis et al., 2018). Los datos financieros anuales EDGAR a los que he tenido acceso típicamente proporcionan unas 10-15 observaciones por compañía; posiblemente insuficientes para que modelos complejos aprendan patrones robustos y confiables.

Los ajustes cualitativos incorporan información que datos históricos limitados no pueden capturar: transiciones estratégicas, cambios de gerencia e interrupciones de industria visibles para analistas pero no aún reflejadas en financieros.

5. Enfoques de Ensamble / Híbridos:

Investigación reciente demuestra que métodos ensemble y enfoques híbridos consistentemente superan alternativas de modelo único (Henrique et al., 2019; Barboza et al., 2017).

Correspondientemente, el marco de puntuación cualitativa representa una arquitectura híbrida específica: los pronósticos cuantitativos base capturan los patrones históricos, los ajustes cualitativos incorporan la inteligencia estratégica prospectiva y la combinación de ambos constituye una arquitectura similar y/o correspondiente funcionalmente a la de los modelos híbridos de sentimiento (NLP) + series temporales, los cuales han sido

identificados en la literatura académicas como actualmente los modelos de mayor potencial en la práctica.

Arquitectura del Marco

La evaluación cualitativa opera en una escala 0-10 construida a través de un proceso multi-método, con varias capas operativas, heurísticas e interpretativas. El proceso más granular y explícitamente detallado queda, por su longitud y extensión, relegado a los anexos (concretamente, el anexo 3), pero podemos subrayar algunas de sus características claves en este apartado:

1. Integración de Fuentes de Datos

El marco sintetiza información de seis canales primarios:

1. Operaciones de investigación y heurística de métricas cualitativas combinadas (ver anexos).
2. Consenso de analistas, teniendo en cuenta calificaciones profesionales agregadas y objetivos de precio relativos respecto a la media de otros años.
3. Investigación potenciada por LLMs: análisis de presentaciones, transcripciones, informes, noticias, reportes a inversores, formas 10-K y 10-Q recientes.
4. Inteligencia de Calificación Crediticia. Compilación y revisión de datos de proveedores especializados como Morningstar o Moody's Orbis (particularmente aquellos relativos a solvencia, liquidez y estabilidad).
5. Métricas de Calidad Fundamental: Evaluaciones de posicionamiento competitivo y calidad de gerencia.
6. Algoritmos de predicción bursátil (propietario, experimental). Factores macroeconómicos, fundamentales, cuantitativos, de sentimiento, de análisis técnico e indicadores incluidos.

Esta integración multi-fuente responde al énfasis de Koller et al. (2020) en análisis holístico considerando interrelaciones entre elementos financieros y estratégicos.

2. Dimensiones de Puntuación:

El puntaje cualitativo se descompone en seis componentes ponderados:

A. Visión Estratégica & Ejecución (25%)

Base teórica: Porter (1980, 1985); Palepu et al. (2013)

- Posicionamiento de mercado relativo a competidores
- Claridad y factibilidad de objetivos estratégicos declarados
- Historial de cumplimiento de hitos estratégicos
- Innovación y efectividad de I+D

B. Calidad de Gerencia & Gobernanza (20%)

Base teórica: Kaplan & Norton (1996)—perspectiva de aprendizaje/crecimiento

- Estabilidad y experiencia de liderazgo
- Disciplina de asignación de capital
- Transparencia en comunicaciones
- Independencia de directorio y calidad de supervisión

C. Contexto de Mercado & Dinámicas de Industria (20%):

Base teórica: Cinco Fuerzas de Porter (1980).

- Tendencias de crecimiento secular en mercados abordables
- Intensidad competitiva y barreras de entrada
- Favorabilidad de entorno regulatorio
- Sensibilidad macroeconómica

D. Resiliencia Operacional (15%):

Base teórica: Cadena de Valor de Porter (1985); Kaplan & Norton—perspectiva de proceso interno. (?)

- Solidez de la logística y cadena de suministros.
- Riesgo de concentración de clientes.
- Diversificación geográfica.
- Adaptabilidad de modelo de negocio.

E. Flexibilidad Financiera (10%):

Base teórica: Damodaran (2012)—estructura de capital y riesgo financiero.

- Acceso a mercados de capital
- Margen de obligaciones a deudores existente
- Consistencia de generación de efectivo y flujos de caja
- Opcionalidad de balance general

F. Sentimiento (10%)

Base teórica: Finanzas conductuales; Kaplan & Norton—perspectiva de cliente/industria

- Tendencias de recomendaciones de analistas
- Cambios en propiedad institucional
- Indicadores de satisfacción de empleados
- Métricas de satisfacción de los clientes

3. Metodología de Cálculo de Puntaje

El puntaje cualitativo final emerge a través de ponderación simétrica entre los componentes automatizados y manuales:

Evaluación Automatizada (50%) (procesado del análisis multi-paquete explicado anteriormente): principalmente, organizado en métricas de riesgo (crédito, operacional, de mercado), indicadores de salud financiera (rentabilidad, liquidez, apalancamiento, eficiencia), métricas de crecimiento y valoración, posicionamiento macro-estratégico.

Juicio del Analista / “Manual” (50%): Input de analista (yo, en este caso) fundamentado en la revisión cualitativa de divulgaciones de compañía, el conocimiento de dominio específico a cada industria, la evaluación de factores intangibles y/o no capturados en los estados financieros.

Este enfoque dual-componente busca reducir la sobredependencia (y falta de flexibilidad) que puede generarse tanto por un exceso de sesgo algorítmico/cuantitativo, como por el exceso de la interpretación subjetiva.

Traducción a Ajustes de Pronóstico:

Mecanismo de Ajuste:

El puntaje cualitativo modifica pronósticos base a través de un factor de ajuste multiplicativo:

$$\text{Factor de Ajuste} = 1 + [(\text{Puntaje Cualitativo} - 5) / 10] \times \text{Parámetro de Sensibilidad}$$

Donde:

- Puntaje Cualitativo varía de 0-10 (5 = neutral).
- Parámetro de Sensibilidad varía por tipo de métrica (típicamente 0.10-0.25).
- Puntajes por encima de 5 aumentan crecimiento pronosticado; abajo de 5 lo disminuyen.

Por ejemplo:

- Pronóstico base de crecimiento de ingresos: 8% anual.
- Puntaje cualitativo: 7.5/10
- Parámetro de sensibilidad: 0.20
- Factor de ajuste: $1 + [(7.5 - 5) / 10] \times 0.20 = 1.05$
- Crecimiento de ingresos ajustado: 8.4%

Sensibilidad Específica por Métrica:

Categoría de Métrica	Sensibilidad	Justificación
Crecimiento de Ingresos	0.25	Sensible a posicionamiento y ejecución
Margen Operativo	0.15	Moderado; restringido por estructura de industria
Capital de Trabajo	0.10	Más bajo; impulsado mecánicamente por operaciones

Intensidad CapEx	0.20	Sensible a decisiones de inversión estratégica
Conversión de Efectivo	0.12	Sensibilidad moderada a desempeño operacional

Justificación Teórica

1. Integración con la Finanzas Conductual:

Al incorporar sentimiento y consenso de analistas, el marco reconoce que expectativas de mercado influyen en resultados reales a través de respuesta de gerencia a presión, acceso a capital en términos favorables, y efectos de confianza / sentimiento. Esto se alinea con hallazgos de que enfoques híbridos combinando sentimiento con modelos cuantitativos representan la frontera actual de precisión.

2. Ajuste por Recursos de la Firma:

Los factores cualitativos habitualmente representan activos intangibles que impulsan una ventaja competitiva sostenible. Son precisamente los factores que deben entenderse antes de que el análisis financiero pueda contextualizarse apropiadamente. Estos factores incluyen los famosos “fosos” (moats) competitivos y cualquier otro elemento que genera barreras de entrada, habilita una ejecución superior y/o amortigua los shocks externos que puedan producirse. Esto incluye las desventajas y negativos en juego que puedan estar afligiendo a la empresa en cuestión, por supuesto.

3. Hipótesis de Mercados Adaptativos:

La precisión de pronóstico depende de identificación de régimen y estado “profundo” de la empresa. Los factores cualitativos ayudan a distinguir:

- Empresas estables: Ajustes cualitativos que agregan un valor incremental modesto.
- Empresas en transición: Factores cualitativos que capturan dinámicas no lineales que los datos históricos cuantitativos por sí solos no pueden.
- Regímenes en crisis: Una evaluación crítica de resiliencia operacional previene extrapolación sobre-optimista en estos casos.

4. Activos intangibles impulsan posicionamiento competitivo

Esquemática de Implementación:

Fase 1: Preparación de Datos:

- Ingerir estados financieros de múltiples fuentes
- Recolectar puntaje cualitativo manual (0-10)
- Validar inputs para procesamiento downstream

Fase 2: Puntuación Multi-Dimensional:

- Ejecutar paquetes de análisis especializados (riesgo, salud financiera, crecimiento/valoración, estrategia macro)
- Generar puntajes automáticos por dimensión
- Combinar con puntajes manuales: (Automático + Manual) / 2

Fase 3: Agregación de Consenso:

- Puntaje de Negocio = promedio ponderado a través de dimensiones
- Representa evaluación cualitativa final

Fase 4: Ajuste de Pronóstico

- Generar pronósticos base (ARIMA, Holt-Winters, XGBoost, Regresión Lineal, Naïve)
- Aplicar factores de ajuste cualitativo
- Crear pronóstico ensamble con componentes ajustados

Fase Final: Protocolo de Validación:

1. Pruebas Fuera de Muestra: Pronósticos generados usando datos hasta T-1, validados contra resultados reales T.
2. Descomposición de Error: MAE, RMSE, MAPE calculados separadamente para pronósticos base versus ajustados.
3. Análisis a Nivel de Métrica: Validación a nivel granular (ingresos, márgenes, componentes de flujo de efectivo, etc).
4. Comparación cruzada entre Compañías: Probado a través de compañías para identificar patrones de efectividad.

Consideraciones de Replicabilidad:

Elementos Mostrados:

- Las seis dimensiones cualitativas y sus ponderaciones (metodologías, puntajes y pesos en anexos).
- Mecanismo de ajuste multiplicativo.
- Metodología de validación.

Elementos Opacados:

- Metodologías y algoritmos propietarios específicos, integración con los demás métodos.
- Valores exactos de parámetros de sensibilidad de los métodos, por industria.
- Arquitecturas específicas de modelo de aprendizaje automático e hiper-parámetros.

Este balance está pensado para permitir la evaluación académica y a la vez preservar el potencial de aplicación personal.

Resumen:

El marco de puntuación cualitativa aborda la limitación fundamental de previsión puramente cuantitativa: la incapacidad de capturar puntos de inflexión impulsados por factores estratégicos, gerenciales o de mercado. Al fundamentar cada dimensión de puntuación en marcos teóricos establecidos, integrando sistemáticamente múltiples fuentes de inteligencia, y aplicando ajustes calibrados a pronósticos base, el marco operacionaliza el puente entre análisis cualitativo y cuantitativo que la literatura identifica como necesario pero rara vez implementa.

3.5 Protocolo de Validación

El protocolo de validación implementa pruebas rigurosas fuera de muestra diseñadas para simular condiciones realistas de forecasting. En lugar de métricas de ajuste dentro de muestra (que recompensan sobreajuste), toda evaluación se conduce en datos que los modelos nunca han visto durante entrenamiento.

Configuración Experimental:

La validación principal compara dos enfoques de pronóstico en competición:

- Pronóstico Base: Modelos ML puramente cuantitativos usando solo datos financieros históricos.
- Pronóstico Ajustado por Puntuación: Modelos ML base mejorados con ajustes cualitativos derivados de puntuación multi-dimensional (posición estratégica, contexto de mercado, calidad de gerencia, etc).

El objetivo de hacerlo de esta manera es tener una forma de validar empíricamente si la puntuación cualitativa mejora de hecho precisión de pronóstico más allá de lo que pueden los modelos puramente cuantitativos.

La estructura de datos empleada fue la siguiente: datos históricos de entrenamiento (2015-2024/2025, variable dependiendo de compañía) usados para generar ambos tipos de pronóstico, con 2025 como año primario de validación fuera de muestra (pero usando 2024 cuando los datos finales de 2025 no han sido publicados por la empresa aún).

Diseño de Validación Temporal:

Validación Walk-Forward:

Dada la longitud limitada de datos financieros anuales (típicamente 10-15 años de observaciones), la validación cruzada k-fold tradicional es inapropiada—violaría ordenamiento temporal y crearía un sesgo prospectivo. En su lugar, empleé una validación walk-forward de ventana expandiente:

Ventana de Entrenamiento Período de Prueba
[2015-2019] → 2020

[2015-2020] → 2021
 [2015-2021] → 2022
 [2015-2022] → 2023
 [2015-2023] → 2024
 [2015-2024] → 2025 (año primario de validación)

El año más reciente (2024 o 2025, dependiendo de disponibilidad de datos) sirve como prueba primaria fuera de muestra. Todos los modelos se entrenan en historia previa y evalúan contra los valores reales registrados.

Este protocolo intenta plasmar y reproducir el uso práctico en el mundo real de este tipo de técnicas: un analista en enero 2025 entrenaría en datos 2015-2024 para pronosticar 2025 (y en adelante). La validación walk-forward proporciona múltiples puntos de prueba donde los datos permiten este tipo de ejercicio.

Métricas de Precisión:

Cuatro métricas complementarias de error evalúan calidad de pronóstico, cada una de ellas capturando diferentes aspectos de la precisión del método:

1. Error Absoluto Medio (MAE):

$$\text{MAE} = |\text{Actual} - \text{Pronóstico}|$$

Desviación absoluta promedio en unidades originales (dólares, porcentajes). Robusto a valores atípicos; apropiado cuando todos los errores son igualmente costosos sin importar dirección. Útil comparando pronósticos para métricas con escalas similares.

2. Raíz del Error Cuadrático Medio (RMSE):

$$\text{RMSE} = \sqrt{[(\text{Actual} - \text{Pronóstico})^2]}$$

RMSE penaliza los errores grandes desproporcionadamente, debido al cuadrado. Este es un rasgo positivo cuando errores grandes son particularmente costosos; a saber: métricas donde los errores extremos tienen impacto comercial desmesurado (ej.: flujo de efectivo afectando liquidez).

3. Error Porcentual Absoluto Medio (MAPE):

$$\text{MAPE} = |\text{Actual} - \text{Pronóstico}| / |\text{Actual}| \times 100\%$$

Registra el error porcentual independiente de la escala. Útil porque permite una comparación a través de métricas con magnitudes diferentes. Tiene como limitación que es indefinido y puede resultar engañoso cuando los valores reales están cerca de cero; asimétrico (100% error para sub-predicción vs. ilimitado para sobre-predicción). Es una

medida de precisión primaria para casos de comparaciones cruzadas de métricas y compañías.

4. MAPE Simétrico (sMAPE):

$$\text{sMAPE} = | \text{Actual} - \text{Pronóstico} | / [(| \text{Actual} | + | \text{Pronóstico} |) / 2] \times 100\%$$

Limitado entre 0% y 200%; tratamiento simétrico de sobre- y sub-predicción. Maneja valores cruzando cero con más soltura que MAPE. Apropiado para métricas que pueden ser negativas o cerca de cero (Utilidad Neta, Flujo de Efectivo Libre, Capital de Trabajo, etc).

Precisión Direccional:

Más allá de magnitud, pronósticos se evalúan en si predicen correctamente la dirección del cambio:

$$\text{Dirección Correcta} = \text{sign}(\text{Actual}_t - \text{Actual}_{\{t-1\}}) == \text{sign}(\text{Pronóstico}_t - \text{Actual}_{\{t-1\}})$$

Está pensando como un Indicador binario (1 = correcto, 0 = incorrecto) de si el pronóstico predice el crecimiento/declive correctamente. En muchos contextos comerciales, saber si los ingresos aumentarán o disminuirán en general importa más que la magnitud precisa. Un pronóstico que predice crecimiento 5% cuando crecimiento real es 3% es mucho más útil que uno prediciendo declive 2%, por ejemplo.

Pruebas de Significación Estadística

Prueba Diebold-Mariano:

La prueba Diebold-Mariano (DM) evalúa si la diferencia de precisión entre dos pronósticos es estadísticamente significativa: Hipótesis Nula: Ambos pronósticos tienen igual precisión predictiva. El estadístico de prueba se base en la diferencia media en funciones de pérdida (errores cuadráticos o absolutos).

$$\begin{aligned} \bar{d} &= (1/n) \sum (e_1^2 - e_2^2) \\ \text{DM} &= \bar{d} / \sqrt{[\text{Var}(\bar{d})]} \end{aligned}$$

Implementación: ``statsmodels.stats.diagnostic.acorr_ljungbox`` para errores estándar robustos a autocorrelación; ``scipy.stats.ttest_rel`` para comparación pareada.

Respecto a la interpretación, p-valor < 0.05 indica que un pronóstico es significativamente más preciso.

Prueba de Rango con Signo de Wilcoxon:

Alternativa no paramétrica que compara diferencias medianas de precisión sin asumir distribución normal. Hipótesis Nula: La diferencia de la mediana en errores es cero. Respecto a la implementación, ``scipy.stats.wilcoxon`` aplicado a diferencias de error emparejadas a través de métricas. Con tamaños de muestra pequeños y distribuciones de error potencialmente no normales, Wilcoxon proporciona inferencia sólida.

Medición del Efecto:

D de Cohen:

La significancia estadística por sí sola es insuficiente; tamaño de efecto cuantifica importancia práctica:

$$D \text{ de Cohen} = (\text{Media}_1 - \text{Media}_2) / \text{SD_Combinada}$$

Se interpreta de la siguiente manera:

- $|d| < 0.2$: Efecto despreciable.
- $0.2 \leq |d| < 0.5$: Efecto pequeño.
- $0.5 \leq |d| < 0.8$: Efecto mediano.
- $|d| \geq 0.8$: Efecto grande.

Respecto al uso, sirve para distinguir entre mejoras estadísticamente significativas pero triviales versus ganancias de precisión significativas.

Evaluación de Métricas por Niveles

No todas las métricas financieras son igualmente importantes para decisiones comerciales. El marco de validación que en este trabajo he implementado sigue una jerarquía de tres niveles:

Nivel 1 (Crítico) — Peso: 3×

Métricas principales impulsando valoración y decisiones estratégicas:

- Ingresos Totales
- Ingreso Operativo
- Utilidad Neta
- Flujo de Efectivo Operativo
- Flujo de Efectivo Libre
- Activos Totales
- Patrimonio Total

Nivel 2 (Importante) — Peso: 2×

Métricas clave operativas y de eficiencia:

- Utilidad Bruta
- EBITDA

- Activos Corrientes
- Pasivos Corrientes
- Capital de Trabajo

Nivel 3 (Soporte) — Peso: 1×

Ítems de línea detallados y métricas secundarias:

- Categorías individuales de gastos
- Subcategorías de activos/pasivos
- Ratios suplementarios

Puntaje de Calidad Agregada:

A modo de seguro, validado y verificado adicional, un sistema de puntaje de calidad ponderado agrega el rendimiento a través de todas las métricas:

$$\text{Puntuación de Calidad} = \frac{\sum(\text{Peso}_i \times \text{Precisión}_i)}{\sum(\text{Peso}_i)} \times 100$$

Donde:

$$\text{Peso}_i = \text{Peso de importancia (Nivel 1 = 3, Nivel 2 = 2, Nivel 3 = 1)}$$

$$\text{Precisión}_i = \max(0, 1 - |\text{MAPE}_i / 100|)$$

El puntaje de la validación varía de 0-100; más alto es mejor. Un puntaje de 80+ indica calidad fuerte de pronóstico; debajo de 60 indica problemas significativos de precisión.

Métricas de Rendimiento Relativo

Más allá de precisión absoluta de las predicciones, el marco de validación evalúa también la mejora de los ajustes cualitativos:

$$\text{Mejora (\%)} = (|\text{Error_Base}| - |\text{Error_Puntaje}|) / |\text{Error_Base}| \times 100\%$$

Aquí, los valores positivos indican que los pronósticos ajustados por puntaje son más precisos. También se emplea un ratio de error de pronóstico, que cuantifica la ganancia relativa de precisión, permitiendo comparaciones cruzadas de compañías y métricas.

Categorías de Fortaleza de Validación

Para proporcionar conclusiones accionables, resultados de validación se clasifican en cuatro categorías basadas en umbrales empíricos:

Categoría	Criterios	Interpretación
FUERTE	Ajustado por Puntaje gana en >60% de métricas Nivel 1 con >20% reducción	Evidencia clara de mejora; ajustes cualitativos proporcionan valor sustancial

	promedio de error	
MODERADA	Gana en >50% de métricas con >10% reducción promedio de error	Evidencia sugestiva; se justifica testing adicional para confirmar beneficios
DÉBIL	>45% tasa de victoria o >5% reducción de error	Mejora marginal; puede no justificar complejidad de implementación
NINGUNA	≤45% tasa de victoria y ≤5% reducción de error	Sin evidencia de mejora; modelos puramente cuantitativos funcionan igual o mejor

Análisis de Atribución del Puntaje Cualitativo

Para las métricas donde los pronósticos ajustados por puntaje muestran mejora o degradación significativa, el marco incluye un análisis de atribución. Para entender este mecanismo, vale la pena resaltar tres puntos clave del mismo:

1. *Descomposición de Ajustes*: Identifica qué dimensiones cualitativas críticamente impulsaron cambios en pronóstico total:

- Impacto de puntaje de Visión Estratégica & Ejecución
- Impacto de puntaje de Contexto de Mercado & Dinámicas de Industria
- Impacto de puntaje de Calidad de Gerencia & Gobernanza
- Impacto de puntaje de Resiliencia Operacional

2. *Análisis de Patrones*: Evalúa si ciertas dimensiones cualitativas mejoran consistentemente la precisión a través de métricas y compañías e independientemente de ellas.

3. *Detección de Sesgos*: Identifica si existen patrones sistemáticos (ej.: sobre-optimismo en métricas de crecimiento, sub-conservadurismo en métricas de riesgo, etc).

Estos mecanismos de análisis han de proporcionar profundidad a los insights en torno a cuándo y por qué los ajustes cualitativos tienen éxito o fallan.

Verificaciones Adicionales

Umbral de Materialidad:

Pequeñas mejoras porcentuales en métricas con un error base bajo pueden no ser económicamente significativas. El umbral de materialidad aplicado asegura que sólo las mejoras sustantivas cuentan como éxitos de validación:

$$\text{Mejora Significativa} = (\text{Error_Base} - \text{Error_Puntaje}) > 2\%$$

Esto previene sobre-interpretación de mejoras estadísticamente significativas pero prácticamente irrelevantes. Asegurar la parsimonia es la diferencia entre un análisis claro y otro ruidoso y opaco.

Análisis de Valores Atípicos:

Métricas con errores extremos ($>100\%$ MAPE) se marcan para una revisión manual posterior. Estos valores atípicos pueden indicar:

- Problemas de calidad de datos o rupturas estructurales.
- Especificación errónea de modelo para métricas volátiles.
- Cambios contables o eventos únicos.

Las estadísticas agregadas se analizan tanto incluyendo como excluyendo valores atípicos para evaluar la sensibilidad, el impacto y cuán razonables son dichos resultados.

Sensibilidad a Ponderación:

Puntajes de calidad se recalculan bajo ponderaciones alternativas de nivel para asegurar que las conclusiones son robustas:

- Pesos iguales: Todas las métricas ponderadas igualmente sin importar importancia.
- Pesos de varianza inversa: Métricas con variabilidad natural más baja reciben peso más alto.
- Pesos de prioridad comercial: Énfasis en métricas críticas para valoración (Ingresos, Flujo de Efectivo, Utilidad Neta).

Plantilla de Validación Cruzada

Tabla de Comparación Detallada:

Para cada validación compañía-año, el esquema de su tabla sigue esta disposición por columna:

Métrica	Real	Pronóstico Base	Pronóstico Puntaje	Error Base (%)	Error Puntaje (%)	Mejora (%)	Dirección (Correcta)	Ganador
---------	------	-----------------	--------------------	----------------	-------------------	------------	----------------------	---------

Estadísticas Resumen:

- Tasa general de victoria (Puntaje vs. Base)
- Tasas de victoria específicas por nivel
- Reducción promedio/mediana de error
- Resultados de pruebas estadísticas (p-valores)
- Tamaños de efecto (d de Cohen)
- Clasificación de fortaleza de validación

Suite de Visualización:

1. Gráfico de Barras de Comparación de Error: MAPE lado a lado para cada métrica
2. Distribución Victoria/Derrota: Gráfico circular de victorias de modelo
3. Histograma de Mejora: Distribución de mejoras porcentuales
4. Gráfico de Dispersión: Real vs. Predicho para ambos modelos
5. Mapa de Calor de Rendimiento por Nivel: Precisión por nivel de métrica
6. Panel Resumen: Estadísticas clave en formato tabular

Preguntas Clave de Evaluación:

El marco de validación aborda estas preguntas fundamentales sobre efectividad de ajustes cualitativos:

1. Magnitud: ¿Cuánto mejora puntuación cualitativa precisión de pronóstico?
2. Consistencia: ¿Ayuda a través de todos los tipos de métrica o solo específicos?
3. Riesgo de Desventaja: ¿Hace pronósticos significativamente peores alguna vez?
4. Significancia Económica: ¿Son mejoras lo suficientemente grandes para importar en la toma de decisiones comerciales de la compañía?
5. Mecanismo: ¿Qué factores cualitativos agregan más valor predictivo?

Estas preguntas guían la interpretación de resultados e informan recomendaciones prácticas sobre cuándo, cómo y en qué magnitud aplicar ajustes cualitativos.

4. Resultados

La validación del marco de pronóstico financiero revela una **efectividad condicional** con patrones claros y diferenciados. Los resultados demuestran que los conjuntos de modelos de aprendizaje automático (modelos de ensamble) suelen ser más precisos y que su combinación con ajustes de puntuación cualitativa produce pronósticos superiores en contextos específicos, mientras que resulta ineficaz o incluso contraproducente en otros.

Esta variabilidad no constituye una debilidad, sino un hallazgo clave que establece los límites de aplicabilidad del marco y valida el énfasis en una implementación dependiente del contexto.

4.1 Métricas de Desempeño

4.1.1 Tasas Generales de Éxito en la Validación

Fortaleza de Validación	Empresas	Tasa de Éxito	Compañías
FUERTE (3+ modelos)	4	33%	NEOG, IDXX, INTU, NKE
MODERADA (2 modelos)	2	17%	LECO, QLYS
DÉBIL/MIXTA (1 modelo)	3	25%	SBUX, BRKR, WWD

El hallazgo clave de estos resultados es que el marco combinado logró una validación significativa (FUERTE o MODERADA) en el ****50% de los casos****, manteniendo la direccionalidad en las 12 empresas, a pesar de las variaciones en las magnitudes específicas.

4.1.2 Patrones de Desempeño por Modelo

Modelos con mejor desempeño (por frecuencia de validación):

1. LINEAR_HOLT_XGBOOST: Validación fuerte en 3 empresas (NEOG, IDXX, LECO)
2. LINEAR_HOLT: Validación fuerte en 4 empresas (INTU, NKE, SBUX, LECO)
3. HOLT-WINTERS: Validación fuerte en 3 empresas (QLYS, INTU, NKE)
4. XGBOOST: Validación fuerte en 3 empresas (NEOG, IDXX, LECO)

Los modelos de conjunto que combinan rigor estadístico (Regresión Lineal, Holt-Winters) con un modelo aprendizaje automático de alta adaptabilidad (XGBoost) superaron consistentemente a los otros enfoques (incluyendo los modelos individuales), validando la Hipótesis 1: la mayor precisión potencial de los ensambles de modelos.

Ocasional pero notable excepción es el caso del modelo NAÏVE ("Ingenuo", que venía a representar al método tradicional de promedios históricos constantes) fue sorprendentemente competitivo en empresas estables y consolidadas (RRX, WWD). Esto viene a validar, parcialmente y en ciertas circunstancias las aproximaciones tradicionales al análisis y modelaje financiero. No es sorprendente en la medida en que han de tener mérito, pues no habrían sobrevivido tanto tiempo si no fuesen útiles y coherentes.

4.2 Análisis de Puntuaciones Cualitativas: La Curva en U Invertida

4.2.1 El Fenómeno de la Curva en U Invertida:

Se observa un patrón relativamente contraintuitivo al analizar las puntuaciones cualitativas frente a los resultados de validación:

Rango de Puntuación	Interpretación	Patrón de Validación	Empresas
3.5–4.7 (por debajo de neutral)	En dificultades/transición	FUERTE (75% éxito)	NEOG (3.5), NKE (3.9), SBUX (4.4), QLYS (4.7)
5.6–6.0 (neutral-buena)	Estable/predecible	MIXTA (50% éxito)	CROX (5.6), WWD (5.6), LECO (6.0)
6.1–6.3 (fuerte)	Líderes de mercado	DÉBIL/NEGATIVA (67% fracaso)	ADBE (6.1), IDXX (6.1), INTU (6.3), RRX (6.3)

Estos datos muestran que las empresas con puntuaciones cualitativas por debajo del nivel neutro (3.5–4.7) mostraron la mayor tasa de éxito en la validación (75%) de la corrección cualitativa. Esto puede parecer paradójico o contradictorio con la expectativa intuitiva de que las empresas más sólidas serían las más beneficiadas de los ajustes cualitativos. Pero probablemente lo que muestra esto es la incertidumbre, los cambios no esperados y no descontados por el mercado, que mayoritariamente tienden a ser negativos (al menos en esta muestra). En este sentido, los resultados no son tan sorprendentes.

4.2.2 Explicación Mecanística por Zona:

Serie de perspectivas, especulaciones, comentarios e hipótesis cortas sobre los resultados, en cada tipo/agrupación de compañías:

Zona 1: Hipótesis de Eficiencia de Mercado:

- Empresas de alta puntuación (≥ 6.0): Los factores cualitativos ya están eficientemente incorporados en los patrones históricos
- Adobe (6.1): Gran capitalización (\$240B+), al estar ampliamente analizados los hallazgos y métricas cualitativos resultan redundantes y/o sólo añaden ruido.
- Intuit (6.3): A pesar de la validación correcta, el valor incremental sobre los modelos base fue mínimo.

Zona 2: Prima por Incertidumbre Estratégica:

- Empresas de baja puntuación (3.5–4.7): En transformación, donde los patrones cuantitativos por sí solos resultan poco fiables:
- Neogen (3.5): Empresa pequeña de seguridad alimentaria en transición → los factores cualitativos capturaron un punto de inflexión
- Nike (3.9): Reposicionamiento de marca y reestructuración de la cadena de suministro → factores estratégicos predictivos

Zona 3: Compensación Volatilidad-Información:

- Empresas estables (5.6–6.0): Incertidumbre moderada pero insuficiente para justificar ajustes complejos
- Crocs (5.6): Calzado sensible a la moda → el ruido cualitativo superó la señal
- Woodward (5.6): Manufactura aeroespacial → la previsibilidad regulatoria redujo el valor cualitativo

4.3 Perspectivas Específicas por Sector

Serie de perspectivas, especulaciones, comentarios e hipótesis cortas sobre los resultados, en cada sector de las compañías de la muestra:

4.3.1 Sectores Donde el Marco Sobresale (Validación Fuerte):

Salud/Diagnóstico:

- IDXX (diagnóstico veterinario): 5 de 7 modelos con validación fuerte, estadísticamente significativa.
- NEOG (diagnóstico de seguridad alimentaria): 3 de 8 modelos con validación fuerte, mayor tasa de éxito (37.5%).
- Factores de éxito: Claridad regulatoria, patrones de demanda estables, flujos de ingresos predecibles.
- Valor cualitativo: Sirve de evaluación rápida del posicionamiento estratégico y la visibilidad del marco regulatorio de las compañías.

Tecnología/SaaS (condicional):

- INTU (software financiero): 6 de 7 modelos con validación fuerte, precisión excepcional en ingresos (0.09–2.21% SMAPE).
- Factores de éxito: Modelos de suscripción, por su alta previsibilidad de ingresos.
- Valor cualitativo: Evaluación de innovación de producto y dinámicas de retención de clientes.

Bienes de Consumo (selectivo):

- NKE (ropa deportiva): 5 de 7 modelos con validación fuerte, estadísticamente significativa.
- Factores de éxito: Fortaleza de la marca, eficiencia del capital de trabajo.
- Valor cualitativo: Valoración de capital de marca y posicionamiento competitivo.

4.3.2 Sectores Donde el Marco Falla (Validación Débil/Negativa)

Grandes Líderes Tecnológicos:

- ADBE (software creativo): Degradación consistente en todos los modelos (–16% a –258% en métricas de nivel 1) al aplicar el ajuste cualitativo.
- Mecanismo de fallo: La eficiencia del mercado ya incorpora los factores cualitativos; añadirlos era redundante y/o ruido.

Consumo Discrecional (sensible a ciclos de moda):

- CROX (calzado): 71% de los modelos sin validación; significancia estadística en dirección negativa.
- Mecanismo de fallo: Puntuación cualitativa desalineada con la volatilidad de la moda.

Banca Regional:

- RRX (banco regional): Efectividad mínima; el modelo NAIVE superó a los enfoques complejos.
- Mecanismo de fallo: Restricciones regulatorias que limitan la diferenciación estratégica.

Manufactura Industrial

- BRKR (instrumentación científica) y WWD (equipamiento aeroespacial): Mejoras mínimas.
- Mecanismo de fallo: Flujos de caja mecánicos y/o discrecionales dependientes de inventario, resistentes a ajustes.

4.4 Validación de Hipótesis

4.4.1 Hipótesis 1: Los Conjuntos de Modelos Superan a los Métodos Individuales:

VEREDICTO: FUERTEMENTE RESPALDADA

Evidencia:

- Los modelos de ensamble lograron validación fuerte en el 67% de las empresas exitosas
- LINEAR_HOLT_XGBOOST: validación fuerte en NEOG, IDXX, LECO

- LINEAR_HOLT: validación fuerte en INTU, NKE, SBUX, LECO
- Significancia estadística en modelos de conjunto (INTU: $p=0.0185$; NEOG: $p=0.0477$)

Una excepción notable: el modelo NÄIVE fue sorprendentemente competitivo en empresas estables (RRX, WWD).

4.4.2 Hipótesis 2: Los Factores Cualitativos Añaden Valor Predictivo Incremental:

VEREDICTO: RESPALDADA DE FORMA CONDICIONAL

- Apoyo fuerte: NEOG, IDXX, INTU, NKE
- Rechazo: ADBE, CROX, RRX

La condición esencial identificada para esto es que los ajustes cualitativos añaden valor predictivo **sólo si** existe una alta incertidumbre estratégica, asimetría de información y/o una baja eficiencia de mercado para la empresa específica.

4.4.3 Hipótesis 3: El Marco Integrado es Superior al Análisis Fragmentado:

VEREDICTO: RESPALDADA CON MATICES

- Precisión direccional perfecta (100%)
- Reducciones de error del 20–40% cuando el marco es aplicable
- No es universal; depende del sector y del contexto

4.5 Estudio de Caso: Por Qué NEOG Tuvo Éxito

4.5.1 La Anomalía de NEOG:

Neogen Corporation alcanzó la mayor tasa de éxito en validación (37.5% validación fuerte) a pesar de tener la puntuación cualitativa más baja (3.5) de todas las empresas analizadas. Esta paradoja ilumina el mecanismo central del marco en acción.

NEOG es una (relativamente) pequeña empresa de diagnóstico de seguridad alimentaria (\$2-3B capitalización). Se encuentra en una situación de transición estratégica (adquisición reciente, expansión de cartera de productos) y obtuvo una puntuación cualitativa baja, por profundos desafíos operacionales, riesgos de integración, presiones competitivas.

4.5.2 Por Qué Funcionó el Ajuste Cualitativo:

1. Poca Fiabilidad de los Datos Históricos:

Posiblemente porque: los patrones previos a la adquisición ya no eran predictivos, los modelos base extrapolaron tendencias obsoletas y los factores cualitativos capturaron el cambio estratégico ignorado por las series temporales

2. Punto de Inflexión Estratégica:

- Evaluación de calidad de gestión (dimensión de gobernanza) identificó equipo capaz de integración
- Evaluación de dinámicas de mercado (dimensión de industria) capturó vientos favorables regulatorios en seguridad alimentaria
- Puntuación de resiliencia operacional (dimensión de cadena de suministro) señaló ejecución exitosa de integración

3. Brecha de Cobertura de Analistas:

- Empresa pequeña con cobertura limitada de sell-side
- El marco cualitativo proporcionó ventaja de información
- El mercado no había incorporado eficientemente la transición estratégica

4. Sinergia del Modelo de Conjunto:

- LINEAR_HOLT_XGBOOST: Combinó tendencia (Linear), suavizado adaptativo (Holt), aprendizaje no lineal (XGBoost)
- XGBoost capturó dinámicas complejas post-adquisición
- El conjunto promedió los errores de modelos individuales

Validación Estadística:

- LINEAR_HOLT_XGBOOST: $p=0.0477$ (significativo), $d=1.403$ (efecto grande)
- NAIVE: $p=0.0071$ (altamente significativo a pesar de la simplicidad)
- Mejora en puntuación de calidad: +30.3 puntos (69.3 → 94.4)

4.5.3 Contraste con ADBE (Puntuación Alta, Validación Fallida):

Dimensión	NEOG (3.5)	ADBE (6.1)
Capitalización	\$2-3B	\$240B+
Cobertura de analistas	Mínima	Extensiva
Fiabilidad histórica	Baja (adquisición)	Alta (estable)
Incertidumbre estratégica	Muy alta	Baja
Valor cualitativo	Ventaja de información	Ruido redundante
Resultado de validación	37.5% fuerte	0% fuerte, fallo sistemático

4.6 Significancia Estadística y Patrones de Tamaño de Efecto

4.6.1 Significancia Estadística Alcanzada (9 instancias en 12 empresas):

Significancia Positiva (6 instancias):

- NEOG - LINEAR_HOLT_XGBOOST: $p=0.0477$, $d=1.403$ (efecto grande)
- NEOG - NAIVE: $p=0.0071$, $d=\text{grande}$
- IDXX - LINEAR_HOLT_XGBOOST: $p=0.0185$, $d=1.403$ (efecto grande)
- INTU - XGBOOST: $p=0.0284$, $d=\text{medio}$
- NKE - ARIMA: $p=0.0002$, $p(\text{Wilcoxon})=0.0156$ (altamente significativo)
- LECO - XGBOOST: $p=0.0292$, $d=\text{medio}$

Significancia Negativa (3 instancias):

- CROX - NAIVE: $p=0.0469$, $d=-0.79$ (efecto negativo medio)
- CROX - HOLT: $p=0.0345$, $d=-0.87$ (efecto negativo medio)
- ADBE - LINEAR_HOLT_ARIMA: $p=0.028$, $d=-1.09$ (efecto negativo grande)

4.6.2 El Empeoramiento en CROX y ABDE:

El empeoramiento de la precisión en CROX y ADBE es también informativo. No creo que se trate de mero ruido o error aleatorio, sino de degradación sistemática donde los ajustes cualitativos consistentemente empeoraron los pronósticos.

Para CROX (empresa de calzado, sensible a los ciclos de la moda), la puntuación cualitativa (5.6, neutral-positiva) pretendía reflejar una fortaleza estratégica moderada. Esto se tradujo en ajustes optimistas de crecimiento. Sin embargo, la volatilidad de la moda y las dinámicas de sentimiento del consumidor fueron más negativas para CROX de lo esperado. El resultado fue un ajuste sobre-optimista sistemático desalineado con patrones de gasto del consumidor con reversión a la media y un empeoro de la precisión de los modelos.

Para ADBE (líder de mercado en software creativo), la puntuación cualitativa (6.1, fuerte) de ADBE fue por factores que mostraban una calidad genuina de la empresa. Pero, probablemente, esta calidad ya estaba completamente incorporada (y descontada) en los patrones financieros históricos. Los ajustes amplificaron tendencias que en realidad estaban alcanzando saturación. El resultado fue una sobre-extrapolación más allá de tasas de crecimiento sostenibles y una peor precisión en los modelos tras el ajuste.

4.6.3 Distribución de Efectos:

Efectos significativos ($|d| > 0.8$):

- Positivos: NEOG (1.403), IDXX (1.403)
- Negativos: ADBE (-1.09), CROX (-0.79 a -0.87)

Efectos moderados ($0.5 < |d| < 0.8$):

- Positivos: INTU, NKE, LECO
- Negativos: Ninguno adicional

Efectos insignificantes ($|d| < 0.5$):

- La mayoría de las otras combinaciones empresa-modelo

Podemos interpretar esta distribución de la siguiente manera: es una distribución bimodal (efectos grandes positivos O negativos, con efectos pequeños en casos estables) que respalda la hipótesis de que los ajustes cualitativos no son refinamientos incrementales sino que tienen sentido como intervenciones que reflejan un cambio de régimen para la compañía/mercado.

4.7 Patrones Específicos por Métrica

4.7.1 Ingresos y Métricas de Primera Línea: Candidatos Principales para Ajuste

Los pronósticos de ingresos mostraron una mejora consistente en los casos de validación exitosa:

- NEOG: 25.69% → 1.83% sMAPE (LINEAR_HOLT_XGBOOST)
- INTU: 2.21% → 0.09% sMAPE (precisión excepcional)
- NKE: 19.90% → 5.56% sMAPE (52.37% mejora)
- IDXX: 4.65% → 0.17% sMAPE (87% mejora)

Esto puede tener una justificación teórica clara en que los ingresos son más inherentemente prospectivos y sensibles al posicionamiento estratégico, dinámicas de mercado, intensidad competitiva y ejecución de gestión; precisamente los factores capturados por dimensiones cualitativas.

4.7.2 Resultado Operativo y Gastos: Dificultades Sistemáticas

Resultado Operativo/Gastos mostró los resultados más inconsistentes:

- Casos de mejora: NEOG (15.01% → 1.87% sMAPE para Gastos de Resultado Operativo)
- Casos de degradación: CROX (degradación sistemática), ADBE (degradación severa)
- Casos mixtos: INTU, NKE (algunos modelos mejoran, otros se degradan)

Tres hipótesis (y mecanismos) que plausiblemente lo explican:

1. Doble Contabilización** Los factores cualitativos (eficiencia operacional, disciplina de gestión de costos) pueden ya manifestarse en patrones de gastos a través de acciones de gestión.
2. Ajuste Inapropiado: La estructura multiplicativa (diseñada para métricas de crecimiento) puede ser inadecuada para métricas de gastos.

3. Propagación de Error en Cascada: El Resultado Operativo se deriva de Ingresos menos Gastos Operativos; es posible, por tanto, que los errores se hayan reduplicado de esta manera.

Una reestructuración y revisión integral del marco, sistema y metodología está justificado para esta métrica específica.

4.7.3 Capital de Trabajo Neto: Divergencia Extrema:

El CTN exhibió en ocasiones las divergencias de resultados más dramáticas:

- CROX: Sobreestimación sistemática de 800-1000% en todos los modelos (el peor desempeño métrico observado con diferencia)
- NEOG: la mayor mejora observada (12.09% → 1.45% sMAPE tras el ajuste)
- Otras empresas: variable

A modo de análisis de la causa de esto, se puede decir que para CROX, el fallo del modelo base parece no estar relacionado en este caso con la calidad del ajuste cualitativo. Todos los modelos base sobreestimaron dramáticamente el CTN. Esto sugiere que el patrón histórico mismo era engañoso.

4.7.4 Métricas de Flujo de Caja: Dificultad Estructural Persistente:

El Flujo de Caja Libre y el Flujo de Caja Operativo fueron consistentemente difíciles de predecir.:

- Mejor caso (QLYS): FCF 0.02% sMAPE (excepcional, pero atípico)
- Rango típico: 20-80% sMAPE para FCF, 10-50% sMAPE para OCF
- Peores casos: ADBE, BRKR, WWD (pobres resultados todos los modelos)

La explicación que encuentro para esto es que el pronóstico de flujo de caja compone múltiples fuentes de incertidumbre, cada componente introduce error y esos errores se acumulan.

4.7.5 Métricas de Balance General: Dominio de la Persistencia:

Las métricas de Balance General muestran efectos de ajuste generalmente positivos pero modestos:

- Fuerte persistencia: Cambios año tras año típicamente <10%
- Restricciones de identidad contable: Activos = Pasivos + Patrimonio
- Menor sensibilidad a factores estratégicos

Los ajustes cualitativos proporcionaron un valor incremental relativamente pequeño para métricas de Balance General relativo a los modelos de persistencia simples.

5. Discusión

5.1 ¿Cuándo Ayudan los Ajustes Cualitativos? La Curva en U Invertida

5.1.1 Eficiencia de Mercado e Incertidumbre Estratégica

Esta validación de modelos y ajustes cualitativos en una muestra de 12 empresas revela un patrón (relativamente) inesperado y teóricamente significativo: **los ajustes cualitativos son más efectivos en empresas con puntuaciones cualitativas negativas.**

Zona 1: Alta Incertidumbre Estratégica (puntuación 3.5–4.7) → Validación Fuerte (75% éxito):

Las empresas de esta zona—NEOG (3.5), NKE (3.9), SBUX (4.4), QLYS (4.7)—lograron predominantemente validaciones fuertes o moderadas. NEOG representa el caso arquetípico: una empresa de pequeña capitalización dedicada al diagnóstico de seguridad alimentaria que atravesaba una transformación estratégica, alcanzó una tasa de validación fuerte del 37.5% (la más alta observada), con el conjunto Linear-Holt-XGBoost alcanzando una puntuación de calidad de 94.4/100 ($p=0.0477$, $d=1.403$). De forma similar, NKE logró una tasa de validación fuerte del 71.4% en 5 de 7 modelos, a pesar de afrontar retos de reestructuración de la cadena de suministro y reposicionamiento de marca, que podía afectar a la precisión de las predicciones.

Estas empresas presentan elevada incertidumbre estratégica, donde los patrones financieros históricos resultaron predictores poco fiables del desempeño futuro. Los factores cualitativos (capacidad de ejecución directiva, evaluación del pipeline de I+D, claridad del marco regulatorio, reposicionamiento competitivo, etc) contienen información incremental y discrecional que aún no se refleja en los datos históricos. La conclusión es que el sistema de puntuación multidimensional que dió lugar al ajuste cualitativo logró capturar puntos de inflexión que los modelos de series temporales no pudieron detectar.

Zona 2: Estabilidad Moderada (puntuación 5.6–6.0) → Resultados Mixtos (50% éxito):

Las empresas de esta zona—CROX (5.6), WWD (5.6), LECO (6.0), BRKR (4.3)—mostraron patrones inconsistentes. LECO logró un éxito moderado (2 de 7 modelos con validación FUERTE), mientras que CROX mostró un fracaso sistemático (6 de 7 modelos sin validación, con significancia estadística en sentido negativo).

Esta zona representa un nivel de incertidumbre estratégica insuficiente para generar ventaja informativa cualitativa, pero también una estabilidad insuficiente para que los factores cualitativos estén plenamente incorporados en los patrones históricos. Los ajustes, por tanto, aquí corren el riesgo de añadir ruido en lugar de señal a las predicciones.

Zona 3: Dominio de la Eficiencia de Mercado (puntuación 6.1–6.3) → Validación Débil/Negativa (67% fracaso):

Las empresas de alta calidad—ADBE (6.1), IDXX (6.1), INTU (6.3), RRX (6.3)—mostraron paradójicamente el peor desempeño agregado. ADBE evidenció un fracaso sistemático en todos los modelos, con degradaciones de las métricas de nivel 1 de entre -16% y -258%. Incluso en los casos exitosos como IDXX e INTU, los modelos base capturaron la mayor parte de la información predictiva, dejando a los ajustes cualitativos un valor incremental marginal.

Estas empresas operan en entornos de alta eficiencia de mercado, donde los factores cualitativos ya están incorporados en los patrones financieros históricos a través de la cobertura de analistas, la guía de la dirección y las expectativas del mercado. Posiblemente, en este contexto, los ajustes cualitativos tienden a duplicar información o a introducir sesgos sistemáticos (por ejemplo, un optimismo excesivo en empresas fuertes).

5.1.2 Marco de Decisión Basado en Características:

La evidencia derivada de esta muestra sugiere y respalda adoptar una heurística de decisión refinada que integre tanto la puntuación cualitativa como las características del mercado:

Alta prioridad (aplicar el marco):

- Puntuación cualitativa < 5.0
- Capitalización bursátil < \$10B
- Presencia de transformación estratégica, actividad reciente de M&A, disrupción sectorial o alta intensidad en I+D (al menos una)
- Ejemplos: NEOG (3.5, \$2-3B, crecimiento en seguridad alimentaria), NKE (3.9, transformación de cadena de suministro)

Prioridad moderada (aplicación selectiva, discrecional):

- Puntuación cualitativa 5.0–6.0 o capitalización \$10–50B
- Sectores con desempeño probado como Salud/Diagnóstico
- Ejemplos: LECO (6.0, sector manufactura pero validación fuerte), IDXX (6.1 pero anulación de sector salud)

Baja prioridad/Evitar (usar solo modelos base):

- Puntuación cualitativa > 6.0 y capitalización > \$50B
- Sectores como banca regional, manufactura industrial o consumo sensible a la moda
- Ejemplos: ADBE (6.1, \$240B, eficiencia de mercado), CROX (5.6, calzado sensible a moda)

5.1.3 Efectividad Condicional por Sector:

Más allá de la puntuación cualitativa y capitalización de mercado, las características sectoriales muestran patrones de efectividad a tener en cuenta:

Salud/Diagnóstico: Validación Incondicional (100% éxito):

- IDXX: 5/7 modelos validación FUERTE a pesar de puntuación cualitativa 6.1.
- NEOG: 3/8 modelos validación FUERTE (tasa más alta observada).
- Características del sector: Claridad regulatoria, patrones de demanda estables, valor de evaluación del contexto estratégico.

Tecnología/SaaS: Condicional al Modelo de Negocio (67% éxito):

- INTU: 6/7 modelos validación FUERTE (SaaS basado en suscripción).
- QLYS: 5/7 modelos validación positiva (seguridad empresarial B2B).
- ADBE: Fallo sistemático (mercado maduro, altamente eficiente).
- Condición de éxito: Modelos de suscripción con flujos de ingresos predecibles; evitar líderes de mercado de gran capitalización.

Bienes de Consumo: Condicional a Marca vs. Moda (50% éxito):

- NKE: 5/7 modelos validación FUERTE (capital de marca, eficiencia de capital de trabajo)
- SBUX: 1/7 modelos validación FUERTE (complejidad de operaciones minoristas)
- CROX: 0/7 modelos validación FUERTE (sensibilidad a moda, volatilidad de sentimiento)
- Condición de éxito: Impulsado por marca con operaciones estables; evitar sensibilidad a moda o categorías impulsadas por sentimiento

Industrial/Banca: Fallo Sistemático (17% éxito)

- BRKR, WWD, RRX: Validación mínima o nula.
- Mecanismo de fallo: Flujos de caja mecánicos (manufactura), restricciones regulatorias (banca), factores cualitativos ya embebidos en patrones operacionales.

5.1.4 Alineación y Extensiones Teóricas:

Estos patrones empíricos validan, enriquecen y extienden marcos teóricos establecidos en literatura previa:

Heterogeneidad de la Eficiencia de Mercado: Los resultados proporcionan apoyo empírico directo para la hipótesis de mercados adaptativos (Lo, 2004, 2017). Los mercados no son universalmente eficientes; existen bolsas de ineficiencia donde los marcos cualitativos pueden generar ventaja de información. El hipotético, tentativo umbral identificado (~\$10B capitalización, ~5.0 puntuación cualitativa) representa el límite donde la cobertura de analistas satura la información cualitativa.

Operacionalización del Análisis Estratégico: El éxito en contextos de alta incertidumbre valida el énfasis de Palepu et al. (2013) en el análisis estratégico como fundamental para la evaluación financiera. El marco operacionaliza exitosamente los conceptos de

posicionamiento estratégico de Porter (1980, 1985), pero solo cuando los factores estratégicos aún no están embebidos en los patrones financieros.

Reconciliación entre Simplicidad y Complejidad en la Elección de Modelos: El insight de la competición M en el que los métodos más simples a menudo bastan (Makridakis et al. 2018) queda vindicado en empresas estables (Zona 3), mientras que el éxito del marco en Zona 1 demuestra que la complejidad está justificada cuando los patrones históricos prueban ser limitados o no fiables. Esta reconciliación entre la aparente tensión entre parsimonia estadística y análisis integrado es un punto medio razonable.

5.2 Selección de Modelos vs. Efectos de Ajuste

La comparación integral de modelos revela que la selección del modelo y los ajustes cualitativos no son ni sustitutos ni factores independientes, sino que interactúan de formas complejas y dependientes del contexto.

5.2.1 Superioridad del Conjunto Validada con Advertencias:

En las validaciones exitosas, los modelos de conjunto dominaron:

- LINEAR_HOLT_XGBOOST: Validación fuerte en NEOG, IDXX, LECO (3 empresas)
- LINEAR_HOLT: Validación fuerte en INTU, NKE, SBUX, LECO (4 empresas)
- HOLT-WINTERS: Validación fuerte en QLYS, INTU, NKE (3 empresas)

Sin embargo, el modelo NAIVE paradójicamente alcanzó las puntuaciones de calidad más altas en entornos estables:

- RRX (banca regional): NAIVE 77.0/100 vs. modelos complejos 59.5-87.0/100
- WWD (aeroespacial): NAIVE 83.0/100 vs. modelos complejos 67.6-78.4/100

Podemos interpretar esto como que la superioridad del conjunto es condicional a la complejidad del negocio. Cuando los patrones históricos exhiben fuerte persistencia y reversión a la media (banca, industria, manufactura en fase madura), los modelos simples pueden evitar el sobreajuste y funcionan competitivamente por sí solos. Cuando coexisten múltiples patrones (tendencia + estacionalidad + rupturas estructurales), los conjuntos capturan señales complementarias de manera más precisa.

5.2.2 Sinergia entre Ajuste Cualitativo y Modelo Base:

Emergió un efecto de interacción crítico: los ajustes cualitativos amplifican las diferencias de calidad del modelo base en lugar de desplazar uniformemente todos los pronósticos.

Evidencia en NEOG:

- Mejor modelo base (LINEAR_HOLT_XGBOOST, 88.8) → Mayor mejora (+5.6 puntos a 94.4).
- Modelo base moderado (XGBOOST, 83.1) → Mejora moderada (+2.3 puntos a 85.4).

- Modelo base débil (NAIVE, 68.5) → Mejora mínima (+0.8 puntos a 69.3).

Evidencia en ADBE (caso negativo):

- Mejor modelo base (NAIVE, 82.4) → Degradación (-0.3 puntos a 82.1)
- Conjunto complejo (LINEAR_HOLT_XGBOOST) → Degradación severa (-4.4 puntos, métricas Nivel 1 -258.7%).

Los modelos base fuertes proporcionan pronósticos fundamentales precisos que los ajustes cualitativos pueden refinar en el margen. Los modelos base débiles producen pronósticos ruidosos donde los ajustes cualitativos no pueden distinguir señal de error. En contextos negativos (ADBE, CROX), los ajustes propagan errores del modelo base en lugar de corregirlos.

Esto tiene una implicación práctica: los marcos de ajuste cualitativo deben ser colocados encima de modelos base fuertes, no usados para rescatar enfoques de pronóstico débiles. El flujo de trabajo óptimo debería de ser:

1. Seleccionar el mejor modelo base para características de la empresa.
2. Evaluar si los ajustes cualitativos añaden valor dado el perfil empresa/sector.
3. Aplicar ajustes solo si ambas condiciones favorecen la mejora.

5.2.3 Patrones de Fragilidad Específicos del Modelo:

Los modelos individuales mostraron alta variabilidad en la respuesta de ajuste entre empresas:

XGBoost:

- Mejora fuerte: NEOG (+2.3 puntos), LECO (significancia estadística $p=0.0292$).
- Degradación: CROX (+1.5 puntos pero validación débil), QLYS (-8.4% mejora).

Se beneficia de ajustes cualitativos en contextos de alta complejidad, no lineales; sufre en entornos estables donde las ventajas de regularización son sobrepasadas por el ruido de ajuste.

ARIMA:

- Mejora fuerte: INTU (4/7 modelos FUERTES), NKE (significancia estadística $p=0.0002$).
 - Fallo sistemático: CROX (validación NINGUNA), WWD (significancia negativa $p=0.0388$).
- Sensible a interferencia de ajuste con estructura de autocorrelación; se beneficia solo cuando los factores cualitativos se alinean con dinámicas de reversión a la media.

HOLT-WINTERS:

- Mejora excepcional: QLYS (significancia estadística $p=0.0255$, $d=1.12$), INTU, NKE
- Resultados negativos: CROX (significancia $p=0.0345$, efecto negativo)

Los mecanismos de suavizado van a interactuar constructivamente con los ajustes cualitativos cuando los ajustes mejoran la predicción base; sin embargo, amplificaron los errores cuando dichos ajustes estaban desalineados.

Todo esto añade matices en cuanto a la robustez de los modelos de ensamble: los modelos de conjunto demostraron mayor estabilidad entre contextos, con menor varianza en efectos de ajuste. Esto apoya la recomendación de que cuando se aplican ajustes cualitativos, las arquitecturas de conjunto proporcionan cobertura natural contra las distorsiones inducidas por ajuste a través de promediado de modelos.

5.3 Implicaciones para la Práctica

5.3.1 Marco de Decisión para Aplicación del Marco:

La evidencia empírica respalda un protocolo de evaluación de tres etapas para la aplicación práctica de este marco:

Fase 1: Evaluación de Características de la Empresa:

Criterio	Medición	Umbral	Puntos
Puntuación Cualitativa	Salida del marco	< 5.0	+3
		5.0-6.0	+1
		> 6.0	-1
Capitalización de Mercado	Presentaciones públicas	< \$10B	+2
		\$10-50B	+1
		> \$50B	-1
Transición Estratégica	Evaluación de analista	Alta incertidumbre	+2
		Moderada	+1
		Estable	0
Sector	Clasificación industrial	Salud/Diagnóstico	+2
		Tech/SaaS (suscripción)	+1
		Bienes de consumo (marca)	+1
		Banca/Manufactura	-2

Regla de Decisión:

- Total ≥ 5 puntos: APLICAR MARCO (alta confianza, 75% probabilidad de éxito).
- Total 3-4 puntos: APLICACIÓN SELECTIVA (probar y monitorear, 50% probabilidad).
- Total ≤ 2 puntos: EVITAR AJUSTES (usar solo modelos base).

Fase 2: Verificación de Calidad del Modelo Base:

Antes de aplicar ajustes, evaluar el desempeño del modelo base:

SMAPE del Modelo Base	Acción
0-2% (excelente)	Omitir ajuste (efecto de techo, riesgo de degradación)
2-5% (bueno)	Ajuste selectivo (aplicar a ingresos, métricas estratégicas solo)
5-20% (moderado)	Ajuste completo (zona de valor marginal más alta)
>20% (pobre)	Mejorar modelo base primero (ajuste no puede rescatar base pobre)

Fase 3: Aplicación Específica por Métrica:

Aplicar ajustes selectivamente por categoría de métrica:

Categoría de Métrica	Prioridad de Aplicación	Condiciones
Ingresos Totales	ALTA	Siempre aplicar si Etapas 1-2 pasan
Resultado Operativo	BAJA	Omitir a menos que fuerte justificación estratégica
Ingreso Neto	MEDIA	Aplicar cautelosamente, monitorear sobre-ajuste
Flujos de Caja	BAJA	Usar descomposición de componentes en su lugar
Capital de Trabajo Neto	CONDICIONAL	Solo si análisis estructural lo respalda
Balance General	BAJA	Derivar residualmente de estados de flujo

5.3.2 Integración en Contextos de Toma de Decisiones:

La efectividad condicional del marco sugiere casos de uso diferenciados:

Contextos de Alta Incertidumbre y/o Alto Riesgo → Aplicación Completa del Marco

- "Due diligence" de M&A (adquisición de empresas).
- Planificación estratégica para empresas en transformación (nueva entrada de mercado, cambio de modelo de negocio, nuevos productos, etc).
- Valoración de startup/empresa en crecimiento (datos históricos limitados, posicionamiento estratégico como factor crítico).
- Se justifica en que cuando la asimetría de información es más alta, la inteligencia cualitativa proporciona una ventaja genuina ajena a lo que las métricas cuantitativas pueden registrar.

Contextos de Incertidumbre Media, Rutinarios → Aplicación Selectiva

- Presupuestación anual para empresas establecidas con estrategias evolutivas.
- Proyecciones financieras a nivel de junta para empresas de mediana capitalización.
- Actualizaciones de pronóstico trimestral cuando se lanzan iniciativas estratégicas.
- Se justifica en que los factores cualitativos proporcionan delimitación de escenarios en lugar de meramente refinar un pronóstico puntual.

Contextos de Baja Incertidumbre, Estables → Solo Modelos Base

- Pronóstico mensual/trimestral para empresas maduras, estables.
- Planificación financiera para líderes de mercado de gran capitalización.
- Monitoreo de cumplimiento de covenants rutinario.
- Se justifica en que cuando los patrones históricos y tendencias son suficientes, los ajustes cualitativos se vuelven redundantes y/o añaden ruido.

Uso Alternativo: Definición de Escenarios:

Incluso cuando los ajustes cualitativos degradan los pronósticos puntuales, las puntuaciones cualitativas proporcionan valor para planificación de escenarios. Podrían usarse para modelar escenarios, de la siguiente manera:

- Puntuación alta (7-8): Supuestos de escenario optimista.
- Puntuación media (5-6): Escenario caso base.
- Puntuación baja (3-4): Supuestos de escenario pesimista.

Este enfoque aprovecha la inteligencia cualitativa para cuantificación de riesgos sin requerir precisión de pronóstico puntual.

5.3.3 Prioridades de Refinamiento del Marco:

Vistos los resultados empíricos obtenidos, refinar y mejorar la metodología del ajuste cualitativo resultado fundamental. Considero que los ajustes de más alta prioridad son los siguientes:

Prioridad 1: Calibración de Sensibilidad Específica por Métrica:

- Desarrollar parámetros de ajuste específicos para ingresos, gastos y flujo de caja.

- Optimizar empíricamente los parámetros de sensibilidad según se amplíe la muestra de validación expandida.
- Implementar ajustes asimétricos (diferentes parámetros para puntuaciones cualitativas positivas vs. negativas).

Prioridad 2: Bibliotecas de Parámetros Específicos por Sector:

- Crear ponderaciones de dimensión cualitativa específicas por cada industria.
- Calibrar parámetros de sensibilidad específicos por sector.
- Desarrollar reglas de decisión específicas por sector para la aplicación del ajuste cualitativo.

Prioridad 3: Metodología de Descomposición del Flujo de Caja:

- Implementar pronóstico a nivel de componente (Ingreso Neto + D&A + ΔCT + CapEx).
- Aplicar ajustes cualitativos selectivamente a componentes donde añaden valor.
- Imponer identidades contables para reforzar la consistencia interna.

Prioridad 4: Verificaciones Previas Automatizadas:

- Formalizar la evaluación de Fase 1-2 en un filtro automatizado.
- Desarrollar diagnósticos preventivos de desempeño de modelo para señalar cuando es probable que los ajustes degraden precisión.

Prioridad 5: Cuantificación de la Incertidumbre:

- Combinar ajustes cualitativos con pronóstico probabilístico (simulación Monte Carlo).
- Usar puntuaciones cualitativas para parametrizar incertidumbre de pronóstico (puntuación baja \rightarrow distribuciones más amplias)
- Generar intervalos de predicción en lugar de pronósticos puntuales

5.4 Limitaciones y Condiciones de Frontera:

5.4.1 Composición de Muestra y Efectos de Selección:

La muestra de 12 empresas, aunque sustancial para los límites de tiempo de este trabajo, es insuficiente por sí sola como para realmente llegar a conclusiones rigurosas sobre las metodologías empleadas. Además, exhibe varios potenciales sesgos de selección:

Distribución Sectorial:

- Salud/Diagnóstico: 2 empresas
- Tecnología/SaaS: 3 empresas
- Bienes de Consumo: 4 empresas
- Industrial/Manufactura: 2 empresas
- Banca: 1 empresa (sub-representado)

La tasa de éxito observada del 100% en Salud puede reflejar tamaño pequeño de muestra en lugar de efectividad genuina a nivel sector. Conversamente, la conclusión del sector

bancario (0% éxito) se basa en una sola empresa (RRX), vaciando cualquier aspiración a poder generalizar los resultados.

Distribución de Capitalización de Mercado:

- Pequeña capitalización (<\$5B): 3 empresas
- Mediana capitalización (\$5-50B): 4 empresas
- Gran capitalización (>\$50B): 5 empresas

El umbral de \$10B observado para eficiencia de mercado puede cambiar con diferentes muestras o períodos de tiempo. La relación entre capitalización de mercado y valor cualitativo probablemente existe en un continuo en lugar de exhibir un umbral agudo.

Las 12 empresas eran entidades saludables, activas en sus operaciones durante 2024-2025. Las empresas que fallaron, fueron adquiridas o salieron de la lista durante el período de estudio están excluidas. Esto probablemente sobreestima la precisión del marco para empresas en dificultades donde los factores cualitativos podrían proporcionar señales de advertencia temprana (a saber, hay un cierto riesgo de sesgo de supervivencia).

5.4.2 Restricciones Temporales y Dependencia de Régimen:

Cada empresa fue validada contra un año fuera de muestra (2024 o 2025). Este diseño no puede distinguir efectividad de metodología de factores específicos del año:

- 2025 de NEOG puede haber sido un año inusualmente predecible post-fusión
- 2024 de CROX puede haber experimentado volatilidad de ciclo de moda

Las condiciones macro (tasas de interés, gasto del consumidor) pueden haber favorecido a ciertos sectores sobre otros, haber afectado a los resultados y estos efectos pueden ser invisibles y no pueden ser contrastados sin muestras más distribuidas en el tiempo.

Mientras que la validación walk-forward proporciona una prueba más sólida que las divisiones fijas, el número limitado de períodos de validación (1-5 por empresa) restringe su valor estadístico.

Además, todos los datos abarcan 2015-2025, un período caracterizado por:

- Tasas de interés bajas (2015-2021)
- Disrupción y recuperación COVID (2020-2022)
- Tasas de interés en aumento (2022-2025)

Esto puede dar lugar a una dependencia de régimen económico que haya afectado a los datos.

La efectividad del marco puede diferir en regímenes económicos alternativos (ej., crisis financiera, estanflación, períodos de crecimiento rápido). Los patrones observados pueden no generalizarse a través de ciclos económicos, tener otras magnitudes o directamente ser patrones diferentes.

5.4.3 Restricciones de Metodología:

Primeramente, una subjetividad residual (quizás en este punto cerca de lo irreducible) en el componente cualitativo. Si bien el marco cualitativo intencionalmente incorpora reglas y criterios de puntuación quasi-cuantitativos en rigor, hay un elemento irreducible de discrecionalidad subjetiva del analista a través de las seis dimensiones. Aunque se equilibran métricas automatizadas con juicio humano, esto introduce inestabilidades de juicio de diferentes tipos. Por ejemplo, una fiabilidad inter-analista desconocida: Diferentes analistas podrían asignar diferentes puntuaciones a una misma empresa, en base a su experiencia y/o intuición. Así mismo, vuelve la consistencia temporal incierta, pues el mismo analista podría puntuar de forma diferente a lo largo del tiempo una misma empresa en una misma situación dada.

Segundo, restos de opacidad en el componente automatizado: aunque se trató de reducir el efecto caja negra en todo lo posible, ciertos resultados derivan de algoritmos de riesgo, salud, crecimiento y estrategia generados ad hoc. Aunque esto reduce la subjetividad de la aproximación, se ha de reconocer que la limita en otros aspectos, pues dificulta la reproducibilidad, limita la transparencia y coarta la flexibilidad de adaptación. No se puede personalizar tan fácilmente para industrias específicas (requerirían más algoritmos ad hoc, específicos para cada una o re-configurar los existentes).

Tercero, la ponderación de las dimensiones no está optimizada ni calibrada empíricamente. Las seis dimensiones cualitativas (Visión Estratégica 25%, Calidad de Gestión 20%, Contexto de Mercado 20%, Resiliencia Operacional 15%, Flexibilidad Financiera 10%, Sentimiento 10%) fueron fundamentadas teóricamente pero no optimizadas ni calibradas empíricamente. Ponderaciones alternativas podrían mejorar efectividad e, idealmente, habrían de ser generadas siguiendo resultados empíricos. Una empresa así no obstante escapa en ambición y profundidad a lo que se puede aspirar en este trabajo en su iteración actual (que es ya de por sí, amplia).

6. Conclusiones y mejoras futuras

6.1 Síntesis de Conclusiones

La validación empírica comprehensiva en 12 empresas a través de múltiples sectores revela que los marcos de ajuste cualitativo representan una herramienta poderosa pero condicional en el pronóstico financiero. El marco alcanzó validación fuerte en el 33% de los casos (NEOG, IDXX, INTU, NKE), éxito moderado en el 17% (LECO, QLYS), y varios niveles de fallo en el 50% (ADBE, CROX, BRKR, RRX, SBUX, WWD).

El corolario central no es si los ajustes cualitativos funcionan o no, sino cuándo, en qué condiciones y por qué lo hacen.

6.1.1 Condiciones de Éxito

Los ajustes cualitativos añaden valor predictivo cuando se cumplen simultáneamente las siguientes condiciones:

1. Alta Incertidumbre Estratégica (Puntuación Cualitativa < 5.0):

Para empresas en transición, reestructuración o transformación estratégica. También patrones históricos poco fiables como predictores de desempeño futuro y actores cualitativos capturan puntos de inflexión no visibles en series temporales. Como evidencia: 75% tasa de éxito en Zona 1 (NEOG 3.5, NKE 3.9, SBUX 4.4, QLYS 4.7).

2. Asimetría de Información (y Capitalización < \$10B):

En cobertura limitada de análisis, la inteligencia cualitativa proporciona una ventaja de información genuina sobre modelos puramente cuantitativos, que no han registrado y/o incorporado eficientemente factores estratégicos y/o futuros. El efecto tiende a aplanarse notablemente al cruzar el umbral de \$10 billones, que aparece en la muestra como un límite donde se refuerza notablemente la eficiencia de mercado.

3. Favorabilidad Sectorial:

- Salud/Diagnóstico: 100% tasa de éxito (claridad regulatoria, evaluación de pipeline)
 - Tecnología/SaaS: Efectivo para modelos de suscripción con ingresos predecibles
 - Bienes de Consumo: Efectivo para marcas fuertes, inefectivo para moda/sentimiento
- El comportamiento en estos sectores evidencia que las características sectoriales moderan la efectividad independientemente de otras características de empresa tratadas.

4. Modelos Base Fuertes:

Los ajustes cualitativos refinan pronósticos precisos, no rescatan a los enfoques débiles. De hecho, pueden/suelen empeorarlos. La zona de valor óptima que se delinea de los datos obtenidos es de 5-20% sMAPE del modelo base, aproximadamente. Como evidencia se puede señalar el efecto de sinergia ajuste-modelo base donde bases fuertes junto a ajustes cualitativos obtienen mejores resultados que por separado.

6.1.2 Condiciones de Fallo

Los ajustes cualitativos degradan sistemáticamente el desempeño cuando:

1. Alta Eficiencia de Mercado (Puntuación > 6.0, Capitalización > \$50B):

- Factores cualitativos ya completamente incorporados en patrones históricos.
- Cobertura extensiva de analistas, guía de gestión, escrutinio de inversores.
- Ajustes duplican información o añaden sesgos sistemáticos.
- Evidencia: 67% tasa de fracaso en Zona 3 (ADBE 6.1, degradación sistemática -16% a -258%).

2. Modelos de Negocio Mecanísticos:

- Banca regional: restricciones regulatorias limitan la diferenciación estratégica.
- Manufactura industrial: flujos de caja dependientes de inventario resistentes a ajuste cualitativo.
- Evidencia: RRX (NAIVE superó a modelos complejos), BRKR/WWD (mejoras mínimas).

3. Dinámicas de Moda/Sentimiento:

- Consumo discrecional sensible a ciclos de moda
- Volatilidad de sentimiento del consumidor domina factores estratégicos
- Puntuación cualitativa desalineada con patrones de gasto de reversión a la media
- **Evidencia:** CROX (significancia estadística negativa $p=0.0469$, $d=-0.79$)

4. Patrones Históricos Estables:

- Empresas maduras con persistencia fuerte y previsibilidad
- Modelos simples (NAIVE) evitan sobreajuste y funcionan competitivamente
- Complejidad no justifica costo de implementación
- **Evidencia:** RRX (NAIVE 77.0/100 vs. complejos 59.5-87.0/100)

6.2 Implicaciones de la Curva en U Invertida:

El patrón más profundo identificado (cuya efectividad es máxima en puntuaciones cualitativas negativas) representa un nuevo punto guía a la hora de pensar en el sobre-pronóstico integrado para empresas.

6.2.1 Reconciliación de Literatura

El *los métodos simples a menudo bastan* de Makridakis et al. (2018) aparece validado, al menos para empresas de Zona 3 (alta calidad, eficiencia de mercado). Los patrones históricos capturan ya la información estratégica / cualitativa en este caso y su añadidura resulta redundante / ruidosa.

El *análisis estratégico es fundamental* de Palepu et al. (2013) también se muestra validado en los resultados, particularmente para las empresas de Zona 1 (alta incertidumbre, baja capitalización). En estas situaciones los factores cualitativos contienen información crítica para precisar las predicciones.

Por tanto, ambas perspectivas son correctas en sus dominios de aplicabilidad y la curva en U invertida define los límites donde cada enfoque es óptimo.

6.3 Conclusiones Metodológicas

6.3.1 Arquitectura de Ensamblas

LINEAR_HOLT_XGBOOST emerge como (habitualmente) la configuración óptima para pronóstico complejo. Este conjunto de modelos combina el suavizado estadístico (Holt-Winters) con el aprendizaje ML más adaptativo (XGBoost). Para este ensamble aparece una validación fuerte en 3 empresas (NEOG, IDXX, LECO). Es en promedio un conjunto de modelos más robusto entre los diferentes contextos, comparado a modelos individuales y otros ensambles. Vale la pena advertir, no obstante, que su superioridad en precisión es condicional a la complejidad del negocio. Los modelos "*ingenuos*" (NAÏVE) son competitivos para patrones estables y persistentes de empresas grandes y maduras. La complejidad de los ensambles está justificada sólo cuando múltiples patrones relevantes coexisten. Por ende implementar una selección de modelo dependiente de contexto en lugar de una aplicación universal de conjuntos debería ser la conclusión.

6.3.2 Patrones Específicos por Métrica

Comento las más relevantes:

Respecto a los Ingresos y a Beneficios Brutos, son las dos métricas más importantes y también las candidatas de más alta prioridad (generalizando) para ajustes cualitativos. Se percibe una mejora consistente en casos exitosos (NEOG: 25.69% → 1.83% sMAPE, por ejemplo), es prospectivo y sensible a posicionamiento estratégico. La recomendación para testeo en casos futuros (por ejemplo de aplicación parcial de ajustes) sería siempre aplicar ajustes a ingresos y beneficios brutos, si la evaluación pasa los dos primeros de calificación.

En Ingresos/gastos operativos, hay ciertas dificultades sistemáticas para y ofuscaciones a validar/rechazar. En los resultados puede delinarse resultados inconsistentes, quizás por un riesgo de doble contabilización. Esto tiene que validarse, por ejemplo, si un ajuste

multiplicativo inapropiado está teniendo lugar. Probaría a derivar de ingresos y supuestos de margen en lugar de ajustar directamente, en futuras iteraciones.

El Capital de Trabajo Neto muestra divergencias pronunciadas. Por ejemplo: CROX: 800% sobre-estimación (fallo de modelo base) y NEOG: 12.09% → 1.45% sMAPE (éxito de ajuste cualitativo). Esta métrica quizás requiera un análisis estructural específico por empresa antes de aplicar a modelos.

En cuanto a los Flujos de Caja hay un desafío estructural persistente. Es, posiblemente, un problema similar al de ingresos/gastos operativos, pero aún más complejo, pues múltiples fuentes de incertidumbre pueden hacer estar haciendo efecto compuesto y los errores se acumulan y multiplican. Rango típico: 20-80% sMAPE para FCF. En futuras iteraciones, buscaría maneras de adoptar una descomposición de los datos financieros, a nivel de componente básico (Ingreso Neto + D&A + Δ CT + CapEx).

6.3.3 Protocolo de Decisión Sistematizado

De el marco de aplicación se puede derivar un protocolo respaldado empíricamente para afinar futuras selecciones de método de pronóstico:

Criterio 1: Evaluación de Características de Empresa:

- Puntuación cualitativa + capitalización de mercado + transición estratégica (sí o no) + sector
- Total ≥ 5 : Aplicar marco (75% probabilidad de éxito)
- Total 3-4: Aplicación selectiva (50% probabilidad)
- Total ≤ 2 : Evitar ajustes (usar modelos base)

Criterio 2: Verificación de Calidad de Modelo Base:

- 0-2% sMAPE: Omitir (efecto de techo)
- 2-5%: Aplicación selectiva (ingresos, métricas estratégicas)
- 5-20%: Aplicación completa (zona de valor óptima)
- 20%+: Mejorar modelo base primero

Criterio 3: Aplicación Específica por Métrica:

- ALTA prioridad: Ingresos totales
- MEDIA prioridad: Ingreso neto (aplicar cautelosamente)
- BAJA prioridad: Resultado operativo, flujos de caja, balance general
- CONDICIONAL: Capital de trabajo neto (requiere análisis estructural)

Este enfoque busca reemplazar los ajustes cualitativos ad hoc con protocolos basados en evidencia.

6.4 Implicaciones para la Inteligencia de Negocios

6.4.1 Asignación de Recursos

Donde debería haber alta concentración de esfuerzo:

- Empresas de pequeña/mediana capitalización (< \$10B)
- Empresas en transición estratégica
- Sectores de salud/diagnóstico y SaaS de suscripción

Dónde bastan los métodos simples:

- Cartera de gran capitalización (> \$50B)
- Líderes de mercado establecidos
- Banca regional y manufactura madura

Las organizaciones deben concentrar costosos recursos de análisis cualitativo donde generan ventaja informativa, usando modelos automatizados simples donde la eficiencia de mercado domina.

6.4.2 Contextos de Toma de Decisiones

La “*Due Diligence*” de M&A (o de inversión) amerita la aplicación completa del marco. Esto es debido a que la mayor asimetría de información puede resultar clave durante negociaciones, los factores cualitativos críticos para la evaluación de sinergias y la magnitud de este tipo de “apuestas” justifican el uso de recursos de análisis profundo.

Para planificación estratégica, por ejemplo, para FP&A (análisis y planificación financiera), para contabilidad, para tomas de decisiones de gerencia, para los presupuestos, etc, aplicación selectiva, discrecional. Es especialmente útil para empresas en transición, realizando cambios significativos en sus operaciones o estrategias de marketing o dirección. También proporciona bases para el análisis de escenarios incluso si pronóstico puntual es subóptimo, para planes de contingencia.

Respecto a monitorear el desempeño de manera regular, basta con los modelos base/tradicionales. En casos de cumplimiento de contratos de deudas, reportes regulatorios, y actualizaciones de pronóstico frecuentes (“flash reports”), la coste-eficiencia favorece la simplicidad.

6.5 Validación Estadística y Significancia Práctica

6.5.1 Robustez Estadística

Significancia estadística positiva, en 6 instancias:

- NEOG - LINEAR_HOLT_XGBOOST: $p=0.0477$, $d=1.403$ (efecto grande)
- IDXX - LINEAR_HOLT_XGBOOST: $p=0.0185$, $d=1.403$ (efecto grande)

- NKE - ARIMA: $p=0.0002$ (altamente significativo)
- Efectos grandes ($d > 0.8$) en casos de éxito

Significancia estadística negativa, en 3 instancias:

- CROX - NAIVE: $p=0.0469$, $d=-0.79$ (efecto negativo medio)
- ADBE - LINEAR_HOLT_ARIMA: $p=0.028$, $d=-1.09$ (efecto negativo grande)
- Degradación sistemática, no aleatoria

El ajuste en dirección negativa válida (significativa) funciona correctamente según su diseño: desplaza sistemáticamente los pronósticos. Pero también produce degradación cuando se aplica inapropiadamente. Esta es evidencia sólida de que el marco captura una señal real y no solamente ruido, pero esa señal debe estar alineada correctamente con el contexto de la empresa, y en esa ambigüedad hay margen de error; concretamente, entre un buen y un mal análisis.

6.5.2 Distribución Bimodal de Efectos

El patrón aquí es que en los efectos grandes (positivos o negativos) y en los efectos pequeños en casos estables: primero, no hay mejoras incrementales—cambios de régimen, segundo, se da una dinámica de alto riesgo - alta recompensa en las correcciones, que requiere evaluación cuidadosa de contexto; y tercero, no es un refinamiento universalmente beneficioso. De nuevo, la implicación de esto es tratar los marcos cualitativos como intervenciones estratégicas que requieren justificación, no mejoras tácticas aplicables universalmente.

6.6 Limitaciones

En cuanto a los límites metodológico y estadísticos del presente trabajo y muestra:

6.6.1 Límites de Generalización

El primero es geográfico: la muestra sólo usó empresas que cotizaban en EE.UU. Sus patrones pueden diferir bajo IFRS. La efectividad específica por país es, por tanto, desconocida.

Segundo, temporal; sólo se usaron datos post-2015, y los resultados pues pueden no generalizarse a diferentes regímenes y/o ciclos económicos. Por ejemplo, la efectividad y desempeño durante escenarios como una crisis financiera, en estanflación o en hiperinflación son factores desconocidos.

Tercero, sectorial. Hay sectores subrepresentados en la muestra, como la manufactura (2 empresas), banca (1 empresa) y otros del estándar GICS sencillamente no estuvieron presentes. Esto significa que las conclusiones en cuanto a sectores requieren cautela. La expansión de muestra es necesaria para solidificar los datos.

6.6.2 Restricciones Metodológicas

Una primera restricción es una cierta subjetividad discrecional en la puntuación cualitativa. Si bien hay medidas concretas y objetivas que forman parte del proceso, el peso que le demos a cada una de ella y la variabilidad de las mismas entre empresas / sectores (por no hablar de los umbrales empleados en cada métrica concreta) suponen un componente discrecional irreducible que introduce una variabilidad dependiente de la competencia y experiencia del analista. Tal fiabilidad (o no) inter-analista no es cuantificable fácilmente, ergo supone un límite práctico a los objetivos metodológicos y en la implementación de los mismos.

Seguidamente, otra restricción notable está en el diseño del mecanismo de ajuste. Se ha mencionado ya que la estructura multiplicativa puede ser subóptima para algunas métricas en el formato de estados financieros. También hay un supuesto de linealidad que puede no sostenerse en valores extremos. La ponderación de cada dimensión no ha sido optimizada empíricamente (y adolece por tanto de los mismos límites mencionados en la primera restricción). Además la corrección del puntaje cualitativo es homogéneo y porcentual a todos los factores por igual. Eso es, admitidamente, una aproximación tosca, no validada empíricamente. Es así por livianez en la operativización, por límites de tiempo. Idealmente, podría refinarse mucho más mediante una serie de métodos e incluso optimizados en cada métrica por separado.

Tercero, la escasa amplitud de la muestra a nivel temporal. Un sólo año de predicción fue analizado en cada empresa de la muestra. Esto es así, en parte, por límites de datos cerrados. SEC ofrece un máximo de 10 años por empresa, lo cual ya pone fricciones a la hora de emplear modelos de aprendizaje automático por falta de suficientes años de muestra. Por tanto, sólo predicciones del último año disponible (2024 o 2025, dependiendo de la compañía) fueron realizadas, a modo de compromiso con las limitaciones de datos y el valor analítico susceptible de ser extraído.

Finalmente, adolece la muestra también de un sesgo de supervivencia. Todas las empresas de muestra eran entidades saludables y que operaron durante diez años, con más o menos dificultad. En caso de que en el futuro se expanda la muestra, de no remediar esta limitación probablemente caería en sobreestimar al alza la precisión para empresas en dificultades y la capacidad de recuperación en general. Aplicar marco a muestra histórica de empresas que subsecuentemente fallaron para validar utilidad de gestión de riesgo.

6.7 Conclusión Primaria

Considero que la contribución clave y más profunda del estudio yace no en demostrar una "superioridad" universal abstracta del marco o de las predicciones, sino en establecer una tesis creíble de cuándo y por qué los enfoques integrados añaden valor. Y si respetamos la

complejidad del pronóstico del mundo real, este tipo de delimitaciones heurísticas (en el sentido en el que la *racionalidad limitada* de Herbert Simons lo entendería) son precisamente las que opino que resultan más útiles en la práctica y más informativas en la teoría.

Reiterando:

El marco no es:

- Una panacea que mejora universalmente cualquier pronóstico.
- Un fracaso que debe ser descartado.
- Un refinamiento incremental aplicable en cualquier lugar.

El marco es:

- Una herramienta que depende de su contexto de uso.
- Efectiva en entornos de alta incertidumbre estratégica y asimetría de información.
- Capaz de deferir con gracia a métodos más simples cuando los mercados son eficientes y las dinámicas de negocio son predecibles.

El pronóstico financiero moderno no requiere enfoques puramente cuantitativos ni puramente cualitativos, sino integración inteligente, híbrida, reflexiva y contingente al escenario y a las necesidades puntuales de cada situación dada.

A mi entender, esta conclusión matizada y validada a través de pruebas empíricas rigurosas en empresas, sectores y métricas diversos, representa un paso hacia metodologías de pronóstico más sofisticadas, contextuales y prácticamente útiles.

Para profesionales de la inteligencia de negocios, gestores, inversores o analistas financieros, el marco proporciona protocolos fundamentados en evidencia empírica para decidir cuándo invertir en análisis cualitativo complejo vs. cuándo confiar en modelos automatizados simples. Para usos académicos, el trabajo pretende ser un ejemplo de la importancia de la investigación flexible y de la efectividad condicional de las herramientas de modelaje y aprendizaje automático. Para las disciplinas relevantes, finalmente, el trabajo busca contribuir a un cuerpo de conocimiento que respete la complejidad del fenómeno empresarial y financiero, mientras proporciona guías prácticas y teóricamente fundamentadas.

La integración exitosa de rigor cuantitativo, el insight cualitativo y validación empírica es el estándar al que he intentado aspirar en esta investigación de inteligencia de negocios. Lo he hecho no haciendo promesas de precisión universal, sino más bien una comprensión tan profunda como me ha sido posible de cuándo, dónde y por qué funcionan los métodos de análisis. Este estudio sirve como base para una agenda de investigación a futuro, que busca refinar continuamente nuestra comprensión de cuándo y cómo integrar mejor la inteligencia cualitativa y cuantitativa en el pronóstico.

6.8 FUTURAS LÍNEAS DE INVESTIGACIÓN

6.8.1 Contribuciones Teóricas

6.8.1.1 Contribuciones a la Práctica del Pronóstico:

Este estudio, a mi parecer, realiza cuatro contribuciones primarias:

1. Efectividad Condicional del Pronóstico Integrado:

El estudio proporciona evidencia empírica rigurosa de que la efectividad de la integración cualitativa-cuantitativa es no monotónica con la calidad de la empresa. Esto pone en contexto el supuesto implícito en literatura previa (Damodaran, 2012; Palepu et al., 2013) de que el análisis cualitativo mejora universalmente el pronóstico.

Lo que se ha podido constatar en estos experimentos es que los ajustes cualitativos exhiben una relación de curva U invertida con la calidad de la empresa. Es decir, precisión máxima en puntuaciones por debajo de lo neutral (alta incertidumbre estratégica), precisión decreciente en puntuaciones altas (dominio de eficiencia de mercado).

Lo nuclear a nivel teórico de esto es que el valor de la corrección cualitativa depende de la interacción entre asimetría de información e incertidumbre estratégica. Cuando la asimetría es alta pero la incertidumbre es manejable (por ejemplo, empresas de pequeña capitalización con estrategias claras), la inteligencia cualitativa proporciona una ventaja comparativa en la predicción. Cuando la asimetría es baja (por ejemplo, en empresas líder con gran capitalización de mercado), los factores cualitativos ya están incorporados en las tendencias cuantitativas. Cuando la incertidumbre es extrema (por ejemplo, empresas en situaciones complejas y/o difíciles, sectores muy volátiles, etc), la señal cualitativa se impone al ruido.

Esta noción es coherente y potencialmente resuelve la aparente contradicción en literatura previa entre el hallazgo de Makridakis et al. (2018) de que los métodos simples a menudo bastan y el énfasis de Palepu et al. (2013) en el análisis estratégico. Ambas intuiciones son correctas en sus respectivos dominios y contextos de aplicabilidad.

2. Heterogeneidad de Eficiencia de Mercado y el Umbral de \$10B:

El presente estudio, pues, proporciona evidencia empírica de un umbral de capitalización de mercado (~\$10B, aproximadamente) donde el valor de pronóstico de la corrección cualitativa disminuye agudamente. Esto encaja desde la perspectiva de la Hipótesis de Mercados Adaptativos (Lo, 2004, 2017) donde la eficiencia a nivel de mercado depende de los contextos de pronóstico específicos por empresa.

Este hallazgo puede tener implicaciones prácticas para la asignación de recursos: las organizaciones deberían concentrar esfuerzos de análisis y pronóstico cualitativo en empresas de pequeña/mediana capitalización donde aún existen potenciales vetas de información no reflejadas en las tendencias cuantitativas, mientras que los más simples métodos tradicionales bastan para tenencias de gran capitalización.

3. Validación de Arquitectura de Conjunto para Pronóstico Financiero:

El estudio refuerza la postura de que los métodos de conjunto (particularmente, en este caso, LINEAR_HOLT_XGBOOST) tienden a superar en precisión a los modelos individuales en contextos de pronóstico financiero, extendiendo hallazgos de competencia M (Makridakis et al., 2018) a estados financieros corporativos con sus interdependencias únicas.

La matización clave que podría hacerse a esta postura es que la superioridad del conjunto / ensamble de modelos es condicional a la complejidad del negocio. Nada innovador o novedoso, en este sentido. Para empresas con patrones financieros simples y persistentes, los modelos que hemos denominado "NAIVE" (tradicionales) funcionan de forma robusta y competitiva a un menor coste computacional y de trabajo cualificado. Esto refina la hipótesis de superioridad del conjunto con límites de aplicabilidad práctica.

Respecto a la relación de los ajustes cualitativos y los modelos de conjunto, la conclusión es que son complementarios en lugar de sustitutos: los modelos más fuertes se benefician más del refinamiento cualitativo, es decir, sinergizan. Esto sugiere que los flujos de trabajo óptimos colocan las bases adecuadas sobre las cuales los ajustes cuantitativos sofisticados pueden mejorarse con correcciones quirúrgicas en base cualitativa. El principio informático GIGO ("*Garbage In, Garbage Out*") que significa que la mala calidad de los datos de entrada produce resultados de mala calidad (y viceversa) resulta relevante aquí para entender la relación entre ambos.

4. Patrones de Ajuste Específicos por Métrica

Las diferencias sistemáticas en efectividad de ajuste a través de categorías de métrica (mejora de ingresos, degradación de gastos, desafíos de flujo de caja) proporcionan una guía para el refinamiento metodológico dirigido.

Respecto a la contribución del presente trabajo, vale la pena decir que si en el repaso de la literatura previa (y particularmente en las posiciones tradicionales), se trataba el pronóstico casi como agnóstico a métrica y contexto, este estudio busca mostrar que los mecanismos de ajuste deben respetar dinámicas y contextos específicos por métrica, sector, tamaño, etc. Métrica a métrica, vemos dinámicas y patrones diferentes: los ingresos (prospectivos, estratégicos) responden positivamente a inteligencia cualitativa; los gastos (retrospectivos, mecanísticos) no; los flujos de caja (compuestos por múltiples incertidumbres) posiblemente requerirían enfoques de descomposición o directamente no serían tratables, etc.

Este punto de vista sugiere que para mejorar, cualquier investigación futura debería desarrollar arquitecturas de ajuste específicas, métrica a métrica, contexto a contexto, en lugar de buscar mecanismos universales.

6.8.1.2 Implicaciones Teóricas:

Para el pronóstico financiero:

1. Cambio de Frontera de Eficiencia

El presente trabajo identifica una frontera de eficiencia de pronóstico donde el valor cualitativo alcanza su pico con asimetría de información moderada: puede, pues, expresarse que el valor del marco cualitativo actúa casi como una función de la asimetría de la información y incertidumbre estratégica (mediada por el tamaño de la compañía). El valor máximo se encuentra en el cruce de estos dos puntos: Primero, asimetría de Información / tamaño (moderada asimetría; pequeña/mediana capitalización) y segundo, en la incertidumbre estratégica alta (por ejemplo, en fases de transición compleja).

2. Apoyo a la Hipótesis de los Mercados Adaptativos

Los resultados se alinean con la Hipótesis de Mercados Adaptativos de Andrew Lo (2004, 2017). Los podemos resumir en 3 puntos: Primero, los mercados no son universalmente eficientes; segundo, existen bolsas de ineficiencia (pequeña capitalización, períodos de transición) y tercero nuestro marco cualitativo explota estas bolsas (al menos en determinados contextos).

3. Dilema Complejidad-Interpretabilidad

Considero que en este trabajo se ha navegado exitosamente este dilema clásico, pues los modelos de conjunto proporcionan precisión, los valores SHAP (específicos para XGBoost en la configuración actual) proporcionan interpretabilidad y el marco cualitativo proporciona una narrativa intuitiva y atractiva para comunicar los resultados con facilidad.

Para la práctica de la Inteligencia de Negocios:

1. Validación del Análisis Holístico

Mi noción es que estos resultados validan empíricamente la intuición recogida en el análisis de la literatura académica previa de que las métricas financieras por sí solas son insuficientes, el contexto estratégico es esencial y la metodología de integración es clave.

2. Contextualización del Rol del Juicio Humano

También valida y proporciona una guía fundamentada en la evidencia empírica sobre cómo articular la colaboración humano-máquina, cualitativa-cuantitativa. Lo que se entrevé es que el énfasis en máquinas (aprendizaje automático, algoritmos) debe ponerse en entornos estables, de alta eficiencia (gran capitalización, banca); mientras, el énfasis en el juicio y toma de decisiones de los analistas humanos debe centrarse en los entornos inciertos, transicionales (pequeña capitalización, cambios estratégicos) y en las anomalías y excepciones.

3. Precisión del Coste-Beneficio de Aplicar el Marco

También queda aproximadamente establecido cuándo la complejidad de una situación empresarial justifica el coste de implementación de medidas como la de este trabajo: como contextos de alto ROI, podemos señalar empresas de pequeña/mediana capitalización, salud, transiciones estratégicas, etc. Mientras tanto, los contextos de ROI bajo o negativo (donde no rentaría este tipo de análisis) serían cuando se tratase de empresas de gran capitalización, manufactura, banca, etc.

6.8.2 Agenda de Investigación Futura

6.8.2.1 Desarrollo a Corto Plazo (12-18 meses):

Prioridad 1: Validación a Escala

El objetivo primario inmediato, sería expandir y enriquecer la muestra a 30-50 empresas

El diseño de la investigación dictaría que al menos 10 empresas seleccionadas proviniese de sectores claves, con ventanas de validación multi-año donde los datos lo permitan y con un patrón de selección tirando a la expansión geográfica (mercados europeos, asiáticos).

En cuanto a posibles preguntas de investigación, por ejemplo:

- ¿Los umbrales sectoriales se sostienen con muestras más grandes?
- ¿El umbral de \$10B es robusto a través de períodos temporales?
- ¿Los patrones se generalizan internacionalmente?

Respecto a los resultados esperados de esta expansión, considero que es razonable anticipar que podrían darse conclusiones más robustas y confiables a nivel sector con poder estadístico notable. Podrían darse límites de confianza más refinados y articulados, para estimar con más precisión los umbrales de aplicabilidad y la identificación de moderadores específicos por país/región.

Prioridad 2: Estudios Profundos por Industria:

Beneficiaría sustancialmente al marco desarrollar esquemas de calibración específicos por sector e industria. Por ejemplo, el esquema de aproximación podría lucir así:

Subestudio de Salud:

- 10-15 empresas (farmacéuticas, biotecnología, dispositivos médicos).
- Ponderación de dimensión cualitativa específica por subsector.
- Parametrización de evaluación de contexto regulatorio.

Subestudio de Tecnología:

- 10-15 empresas (software empresarial, infraestructura cloud, ciberseguridad).
- Optimización de modelo específico para ingresos recurrentes.
- Métricas de retención/expansión de clientes.

Subestudio de Bienes de Consumo:

- 10-15 empresas (CPG, minorista, marca de lujo).
- Marco de evaluación de capital de marca.
- Distinción sensible/no sensible a la moda.

Resultados esperados:

- Bibliotecas de parámetros específicos por sector.
- Árboles de decisión de aplicabilidad específicos por industria.
- Benchmarks de precisión específicos por sector.

También convendría barajar una inversión para acceso a datos especializados de industria/sector (Nielsen, IRI, Gartner, etc.)

Prioridad 3: Predicción de Quiebra

Convendría validar la efectividad del marco para empresas en riesgo de quiebra. Un esquema provisional para esto sería: primero, aplicar el marco a empresas que subsecuentemente fallaron o fueron adquiridas; luego, probar capacidad de advertencia temprana de ajustes cualitativos en dichos casos; finalmente, comparar el desempeño en esta tarea con modelos tradicionales de puntuación Z (Altman, Ohlson).

Ejemplo de preguntas de investigación relevantes:

- ¿Los factores cualitativos proporcionan señales predictivas de dificultad?
- ¿Cuánto tiempo de anticipación proporcionan las advertencias?
- ¿La efectividad del marco difiere para empresas en dificultades vs. saludables?

Los resultados que esperaría de esta aproximación son la validación del marco como aplicación de gestión de riesgo de crédito, una calibración de los umbrales exactos de advertencia y una potencial Integración con los sistemas de puntuación de riesgo.

Podría valer la pena invertir en acceso a datos históricos especializados de quiebras y dificultades empresariales (Compustat, Moody's) para habilitar esta línea de investigación.

6.8.2.2 Desarrollo a Medio Plazo (18-36 meses):

Prioridad 4: Expandir el Alcance de la Puntuación Cualitativa Automatizada

Sería muy conveniente reducir la subjetividad todavía más, por múltiples razones. Además del rigor añadido, esto mejoraría la replicabilidad, la escalabilidad y generalización del marco. A voz de pronto, identificó tres componentes de investigación que pueden ayudar a ese cometido:

1. Análisis NLP (o LLM) de Informes y Comunicaciones:

- Generación y extracción de métricas de sentimiento según el tono de gestión/comunicación.
- Identificación de temas estratégicos (innovación, reestructuración, expansión).
- Cuantificación de la claridad/confianza de la gestión.

2. Análisis de Sentimiento de los Informes de Analistas:

- Agregación de opiniones de consenso/dispersión.
- Seguimiento de cambios de recomendación/revisiones de estimación.
- Análisis de cobertura (número de analistas, profundidad de investigación).

3. Datos Alternativos:

- Análisis de tráfico web (interés del consumidor, compromiso digital).
- Análisis de sentimiento de redes sociales.
- Datos de transacciones de tarjetas de crédito (pueden servir de indicadores de ingresos en tiempo real).

A modo de consecuencias / resultados de estas medidas, esperaría una reducción muy significativa en el tiempo (y energía) que la puntuación manual supone, una mejora de la consistencia temporal (por la recalibración automatizada) y la capacidad de mantener flujos de actualización en tiempo real (por ejemplo, una actualización diaria/semanal de las puntuaciones).

Prioridad 5: Implementación de Pronóstico Trimestral:

En parte, derivado de las medidas de la prioridad 4, sería adaptar el marco para el pronóstico trimestral (y no sólo anual, como es ahora).

Esto supone bastantes dificultades, no obstante: los datos trimestrales suelen no tener el mismo rigor auditado, adolecen de menor fiabilidad, mayor ruido estacional y suponen ventanas de entrenamiento incluso más cortas para modelos ML (que ya de por sí en este trabajo no estaban en un rango aceptable). Además, supone una mayor frecuencia de actualizaciones de factores cualitativos, con el consiguiente aumento exponencial de carga de trabajo.

Para hacerlo, habría que desarrollar modelos estacionales trimestrales específicos, calibrar diferentes esquemas de ponderación y validar un horizonte de pronóstico válido (1T, 2T, 4T).

Esperaría que de esto resultase modelos con cierta capacidad de pronóstico trimestral (menor pero aceptable precisión) y una versión ligera del marco especializado para la detección rápida de puntos de inflexión estratégica.

Prioridad 6: Optimización de Mecanismos:

Fundamentalmente, se trata de optimizar y maximizar la efectividad del marco a base de mejorar y calibrar empíricamente la selección de parámetros.

1. Optimización de Ponderación de Dimensión Cualitativa:

Mediante regresión de grid search por sector, validación cruzada (para prevenir el overfitting / sobreajuste de modelos) y con pruebas de ponderaciones específicas por empresa vs. por sector.

2. Prueba de Ajuste Alternativas:

Otras fórmulas y configuraciones, por ejemplo:

- Aditivo: `Pronóstico = Base + (Puntuación - 5) × Sensibilidad`
- Exponencial: `Pronóstico = Base × exp((Puntuación - 5) × Sensibilidad)`
- Basado en umbral: Aplicar ajustes solo cuando $|Puntuación - 5| > \text{umbral}$

3. Calibración de Parámetros de Sensibilidad Específicos por Métrica:

Especialmente, tres: sensibilidad de ingresos, de margen y de flujos de caja.

Mis expectativas aquí son de una mejora marginal pero perceptible en la precisión de pronóstico (relativa a como funciona ahora mismo, a los parámetros actuales), la generación de una útil “biblioteca” de configuraciones óptimas específicas por sector/métrica y la creación de un módulo de selección automática de parámetros.

6.8.2.3 Exploraciones a Largo Plazo (36+ meses):

Prioridad 7: Enfoques de Inferencia Causal:

A razón de que convendría moverse más allá de correlación a afirmaciones causales sobre efectividad de ajuste, de las siguientes formas:

1. Análisis de Variables Instrumentales:

- Identificar instrumentos para factores cualitativos (cambios de CEO, cambios regulatorios)
- Estimar impacto causal de calidad de gestión, claridad estratégica en precisión de pronóstico

2. Diseños de Discontinuidad de Regresión:

Precisar en más detalle el funcionamiento y características de los umbrales, expandir los umbrales naturales descubiertos (cruce de \$10B cap, adición/eliminación de índice?) y estimar el cambio causal de efectividad en los ajustes.

3. Diferencias para Efectos de Transición Estratégica:

Comparar más casos de empresas que experimentan transiciones vs. grupo de control. Cuantificar cuál puede ser el valor de ajuste cualitativo durante diferentes tipos de períodos de transición.

Este tipo de experimentos y calibraciones deberían de llevar a poder realizar más afirmaciones causales sobre el valor de ajuste cualitativo, a estimaciones de parámetros menos sesgadas/aleatorias

Prioridad 8: Integración de Aprendizaje por Refuerzo:

Bastante auto-explicativo. Desarrollar sistemas de pronósticos auto-mejorables es muy vigente y valioso.

1. Mecanismos de Ajuste Adaptativos:

Adaptar el sistema para hacer que aprenda de errores de pronóstico usando principios de RLHF (“*Reinforcement Learning, Human Feedback*”). Los parámetros de este modo pueden entrenarse según su historial ya que se actualizan dinámicamente según su desempeño.

2. Selección de Parámetros Dependiente de Contexto:

También entrenable es para el modelo aprender cuándo aplicar ajustes según las características de empresa/sector. El RLHF puede servir para precisar más esta cualidad. Útil, por ejemplo, sería desarrollar árboles de decisión aprendidos (en vez de las actuales reglas codificadas a mano, según regla de pulgar) y adoptar procesos de optimización algorítmica (pensando aquí particularmente en MAB, Multi-Armed Bandit).

3. Optimización Multi-objetivo:

Los múltiples objetivos enriquecen el valor de la inversión en RLHF y pueden servir para optimizar los balances entre precisión vs. interpretabilidad vs. coste computacional y para habilitar e integrar un sistema de pronóstico realmente personalizable.

El producto de final de una intervención así debería ser un sistema capaz de auto-mejora continua, personalizable para las necesidades específicas de la situación y de a qué tipo de empresas se apunte.

Prioridad 9: Extensión a otra Clase de Activos:

Posiblemente valdría la pena extender este tipo de metodología y probar la capacidad de generalización del marco más allá de empresas / acciones cotizando en el mercado público.

1. Private Equity/Venture Capital (Capital Privado):

Es una posibilidad relativamente viable (mayor dependencia de factores cualitativos), pero pero el acceso y la divulgación de información financiera es limitada, la evaluación depende fuertemente del proceso “due diligence” de analistas y el monitoreo de rendimiento de cartera.

2. Pronóstico Macroeconómico:

Mucho más ambicioso. Aplicable a aspectos sectoriales/industriales, y/o a PIB nacional con factores institucionales (u otros ndicadores específicos: inflación, desempleo, consumo, etc.) La reconversión del marco, no obstante, tendría que ser radical.

3. Mercados de Real Estate/Bienes Raíces:

Similares características a la anterior. Requeriría una reconversión casi completa, con, por ejemplo, una sección totalmente nueva en cuanto dimensiones cualitativas específicas, metodología y aplicaciones

4. Materias Primas y Recursos Naturales:

Similares características a las dos anteriores. Muchas nuevas secciones, por ejemplo para empresas productoras con factores ESG y geopolíticos.

7. Bibliografía

Libros Académicos y de Texto:

Análisis de Negocios y Valoración:

- Palepu, K. G., Healy, P. M., Peek, E., & Wright, S. (2013). *Business Analysis and Valuation: IFRS Edition*. Cengage Learning.

Valoración de Compañías y Finanzas Corporativas:

- Damodaran, A. (2010). Valuation approaches and metrics: A survey of the theory and evidence. *Foundations and Trends in Finance*, 5(4), 479-670.
- Damodaran, A. (2012). *Investment Valuation: Tools and Techniques for Determining the Value of Any Asset* (3rd ed.). Wiley Finance.
- Koller, T., Goedhart, M., & Wessels, D. (2020). *Valuation: Measuring and Managing the Value of Companies* (7th ed.). Wiley.

Gestión Estratégica y Análisis Competitivo:

- Grundy, T. (2006). Rethinking and reinventing Michael Porter's five forces model. *Strategic Change*, 15(5), 213-229.
- Kaplan, R. S., & Norton, D. P. (1996). *The Balanced Scorecard: Translating Strategy into Action*. Harvard Business Review Press.
- Kaplan, R. S., & Norton, D. P. (2001). *The Strategy-Focused Organization: How Balanced Scorecard Companies Thrive in the New Business Environment*. Harvard Business School Press.
- Porter, M. E. (1980). *Competitive Strategy: Techniques for Analyzing Industries and Competitors*. Free Press.
- Porter, M. E. (1985). *Competitive Advantage: Creating and Sustaining Superior Performance*. Free Press.

Pronóstico con Series Temporales y Estadística:

- Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and Practice* (3rd ed.). OTexts.
- Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffman, A., & Giovannini, E. (2005). *Handbook on Constructing Composite Indicators: Methodology and User Guide*. OECD Statistics Working Papers.

Artículos Académicos:

Finanzas Conductuales:

- Lo, A. W. (2004). The Adaptive Markets Hypothesis: Market Efficiency from an Evolutionary Perspective. *The Journal of Portfolio Management*, 30(5), 15–29.
- Lo, A. W. (2017). *Adaptive Markets: Financial Evolution at the Speed of Thought*. Princeton University Press. (A paperback edition was published in 2019)

Aprendizaje Automático y Predicciones Financieras:

- Amini, S., Elmore, R., Öztekin, Ö., & Strauss, J. (2021). Can machines learn capital structure dynamics? *_Journal of Corporate Finance_*, 70, 102073.
- Bao, Y., Ke, B., Li, B., Yu, Y. J., & Zhang, J. (2020). Detecting accounting fraud in publicly traded U.S. firms using a machine learning approach. *_Journal of Accounting Research_*, 58(1), 199-235.
- Barboza, F., Kimura, H., & Altman, E. (2017). Machine learning models and bankruptcy prediction. *_Expert Systems with Applications_*, 83, 405-417.
- Cao, L. J., & Tay, F. E. H. (2003). Support vector machine with adaptive parameters in financial time series forecasting. *_IEEE Transactions on Neural Networks_*, 14(6), 1506-1518.
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *_Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining_*, 785-794.
- Henrique, B. M., Sobreiro, V. A., & Kimura, H. (2019). Literature review: Machine learning techniques applied to financial market prediction. *_Expert Systems with Applications_*, 124, 226-251.
- Huang, W., Nakamori, Y., & Wang, S. Y. (2005). Forecasting stock market movement direction with support vector machine. *_Computers & Operations Research_*, 32(10), 2513-2522.
- Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). Statistical and machine learning forecasting methods: Concerns and ways forward. *_PLOS ONE_*, 13(3), e0194889.

8. Anexos

Anexo 1: Criterios de Selección de Empresas

1. Número Total de Empresas:

12 empresas, en una matriz 3x4:

- 3 categorías de tamaño (grandes/medianas/pequeñas).
- 4 sectores diferentes.

2. Criterios de Selección Obligatorios

2.1 Criterios de Calidad de Datos

Criterio	Especificación	Justificación
Historial mínimo	Mínimo 8 años de 10-K completos	Necesario para entrenamiento robusto y backtesting significativo
Estados financieros comprensivos	Balance Sheet, Income Statement, Cash Flow Statement completos	Los 3 estados son interdependientes en el análisis
Consistencia contable	Sin cambios drásticos en políticas contables o reestructuraciones mayores	Para evitar discontinuidades que distorsionen modelos de series temporales
Disponibilidad en EDGAR	Datos estructurados	Recopilación y limpieza viables

2.2 Criterios de Diversificación

A) Por Tamaño (Capitalización)

- Grandes empresas: >\$10B (4 empresas)
- Ventajas: Estabilidad, datos de alta calidad.
- Dificultades: Menor volatilidad puede reducir poder predictivo de ML.
- Medianas empresas: \$2B-\$10B (4 empresas)
- Ventajas: Equilibrio entre estabilidad y dinamismo.
- Dificultades: Cobertura analítica variable.
- Pequeñas empresas: \$300M-\$2B (4 empresas).
- Ventaja: Mayor capacidad de poner a prueba a nuestros modelos.

- Dificultad: Datos potencialmente de menor calidad y más susceptibles a sobrepasar los modelos por la volatilidad y el ruido.

B) Por Sector (Usando clasificación GICS)

El objetivo aquí es capturar las características financieras heterogéneas en las empresas.

1. Tecnología (3 empresas):

- Particularidades: Altos márgenes, crecimiento rápido, capex en I+D.
- Ejemplos: Software, semiconductores, hardware.

2. Consumo Discrecional (3 empresas):

- Particularidades: consumo muy cíclico, sensibilidad a la estacionalidad.
- Ejemplos: Retail, automóviles, entretenimiento.

3. Industria/Manufactura (3 empresas):

- Particularidades: Capex intensivo, márgenes moderados.
- Ejemplos: Maquinaria, aeroespacial, construcción.

4. Salud/Farmaceúticas (3 empresas):

- Particularidades: Flujos de capital conservadores, expuestos a regulación.
- Ejemplos: Dispositivos médicos, biotech, servicios de salud.

Sectores que he considerado a evitar:

- Financiero: Sus estados financieros estructuralmente diferentes (no aplican aquí los ratios tradicionales).
- Utilities/Real Estate: Son modelos de negocio muy específicos, dependientes de una regulación intensiva (el análisis de regulaciones no era un foco aquí).

C) Por Madurez del Negocio

- Empresas maduras (6): Crecimiento estable y/o dividendos consistentes.
 - A testear aquí: ¿Los modelos simples son suficientes?
- Empresas en crecimiento (4): Expansión activa y/o reinversión alta.
 - A testear aquí: ¿Los modelos ML capturan mejor los patrones no lineales?
- Empresas cíclicas (2): Más expuestas a ciclos económicos
 - A testear aquí: ¿Sensibilidad elevada a variables macro?

3. Criterios de Selección Deseables (aplicación superficial)

3.1 Calidad Analítica:

- Cobertura: Empresas con calificación pública de analistas (facilita el benchmarking vs. un consenso de mercado estable, contrastable y riguroso).
- Transparencia: Segmentación clara en 10-K (geográfica, por línea de negocio).
- ESG: métricas ESG publicas, para el (sub)módulo estratégico correspondiente.

3.2 Diversificación Adicional:

- Geográfica: Mix de empresas con operaciones predominantemente domésticas vs. internacionales.
- Exposición a FX: Algunas con ingresos significativos en moneda extranjera.
- Modelo de negocio: B2B vs. B2C, productos vs. servicios, recurring revenue vs. transaccional, etc.

3.3 Casos de Interés Específico

- 1-2 volátiles: Empresas que experimentaron crisis y recuperación.
- 1-2 estables: Empresas consistentemente estables (sirven de estándares de comparación).

4. Criterios de Exclusión

Exclusión Inmediata:

1. Eventos o transformaciones mayores (salvo en casos de interés específico):

- M&A extremos en los últimos 5 años.
- Bancarrotas, reestructuraciones.
- Spin-offs o carve-outs recientes.

2. Problemas con los datos:

- Cambios de ticker/CUSIP que dificulten el seguimiento.
- Inconsistencias contables materiales.
- Períodos prolongados sin datos históricos.

3. Sectores de difícil tratamiento:

- Financieros (bancos, seguros, fondos de inversión, asset managers, etc)
- REITs y utilities (estructuras de capital únicas).
- Holdings/conglomerados sin desagregación clara.

Exclusión Relativa (evaluar caso a caso):

- Empresas con pérdidas netas sostenidas (>3 años) - dificulta los ratios de valoración.
- Capex extremadamente volátiles (ej: mineras con proyectos puntuales).
- Empresas pre-revenue o en fase de comercialización temprana (y start-ups).

5. Estrategia de Selección Práctica

Devino de un modo aproximado a:

Fase 1: Pre-screening Automatizado:

Filtros cuantitativos en base de datos (ej: via Python + SEC EDGAR API):


```
python
filters = {
    'min_years': 12,
    'market_cap_range': (300e6, None),
    'exclude_sectors': ['Financials', 'Real Estate'],
    'complete_statements': True,
    'no_major_ma': True
}
```

Fase 2: Selección Manual Estratificada:

1. Listar candidatos por estrato (sector x tamaño).
2. Priorizar por calidad de datos y cobertura.
3. Validar manualmente la ausencia de eventos corporativos mayores.

Fase 3: Validación Final:

Checklist por empresa:

- [] 10-K disponible en años 2014-2024 (mínimo).
- [] Estados financieros completos y accesibles.
- [] Capitalización estable en su categoría.
- [] Sin reestructuraciones o adquisiciones mayores.
- [] Datos de industria/sector para benchmarking disponibles.

6. Matriz de Selección (12 Empresas)

#	Tamaño	Sector	Perfil	Objetivo Analítico
1-2	Large	Tech	Maduro + Crecimiento	Comparar estabilidad vs. dinamismo
3-4	Large	Consumo	Cíclico + Maduro	Capturar la estacionalidad
5-6	Mid	Industria	Crecimiento + Cíclico	Balance capex/ingresos
7-8	Mid	Salud	Maduro + Crecimiento	Flujos conservadores
9-10	Small	Tech + Consumo	Crecimiento	Volatilidad y patrones no lineales
11-12	Small	Industria + Salud	Maduro	Resiliencia en empresas pequeñas

7. Métricas de Validación

Diversidad Cuantitativa:

- Coeficiente de variación de revenue CAGR: >30% entre empresas
- Rangos de márgenes operativos: rango de al menos 20pp (ej: 5%-25%)
- *Correlación inter-empresarial: <0.6 en promedio (para evitar una "*clusterización*" excesiva).

Balance de la Complejidad:

- Ratio simple/complejo: ~60/40 (los modelos más simples deberían de funcionar bien en la mayoría de casos, pero el modelo ML debe brillar en los casos complejos).
- Presencia de outliers: 1-2 empresas con comportamiento atípico para verificar la solidez del marco y modelos de aprendizaje automático.

Anexo 2: Propuesta Final de Empresas

La propuesta de empresas específicas para validación empírica acabó siendo 12 empresas clave, distribuidas así:

1. Grandes (>\$10B) - 4 empresas

Tecnología:

1. Adobe Inc. (ADBE):

- Capitalización: ~\$230B.
- Perfil: Maduro con crecimiento sostenido.
- Transición exitosa de licencias a SaaS (2013-2017) - caso interesante de cambio de modelo
- Ingresos recurrentes predecibles (~95% suscripciones)
- Márgenes operativos consistentemente altos (35-45%)
- Historial limpio desde los años 2000+
- Su valor analítico principal viene de probar si ML captura mejor la transición SaaS vs. extrapolación lineal.
- Datos abundantes, de calidad, segmentación clara (Digital Media, Digital Experience, Publishing).

2. Intuit Inc. (INTU):

- Capitalización: ~\$170B
- Perfil: Maduro con crecimiento moderado
- Fuerte estacionalidad (temporada fiscal en Q2-Q3).
- Portfolio diversificado (TurboTax, QuickBooks, Credit Karma post-2020).
- Márgenes operativos 25-30%.
- Adquisiciones estratégicas pequeñas pero frecuentes.
- Su valor analítico está en capturar la estacionalidad y efectos de M&A menores.
- Datos detallados, un historial de más de 20 años.

Consumo Discrecional:

3. Starbucks Corporation (SBUX):

- Capitalización: ~\$110B.
- Perfil maduro, cíclico.
- Crecimiento por expansión de tiendas (same-store sales).
- Sensibilidad a los ciclos económicos (discrecionales).
- Exposición internacional significativa (~30% ingresos).
- Márgenes operativos de 15-18% aproximadamente.

- El valor analítico del caso es que Starbucks es una empresa con métricas operativas únicas, por lo que posiblemente dará resultados interesantes y configuraciones ricas para feature engineering.
- Los datos son muy transparentes, la desagregación geográfica es clara.

4. Nike Inc. (NKE):

- Capitalización: ~\$130B.
- Perfil maduro con volatilidad cíclica.
- Negocio estacional (Q2-Q3 más fuertes por efecto "*vuelta-al-cole*").
- Transición digital acelerada (crecimiento DTC).
- Exposición FX significativa (>50% ingresos son internacionales).
- Márgenes operativos: 12-15% aproximadamente.
- El valor analítico para Nike es como caso de alta complejidad, por múltiples factores significativos en juego: estacionalidad, FX, ciclos de producto.
- Sus datos tienen excelente calidad y transparencia, con métricas de inventario y sell-through detalladas.

2. Medianas (\$2B-\$10B) - 4 empresas:

Industria/Manufactura:

5. Woodward Inc. (WWD):

- Capitalización: ~\$8B.
- Crecimiento en industria cíclica.
- Aeroespacial (60%) + Industrial (40%) - exposición dual.
- Ciclos largos de desarrollo.
- Capex intensivo pero predecible.
- Márgenes operativos 12-16%.
- Valor analítico aquí, probar la predicción con backlog como elemento clave.
- Datos de buena calidad, segmentación clara por vertical.

6. Regal Rexnord Corporation (RRX):

- Capitalización: ~\$6B.
- Perfil de cíclico industrial.
- Portfolio diversificado de componentes industriales.
- Alta correlación con PMI manufacturing.
- Fusión Regal-Rexnord en 2021 (datos pre-fusión disponibles separadamente).
- Márgenes operativos 10-14%.
- Valor analítico: sensibilidad a variables macro (PMI, industrial production).
- Datos adecuados, pueden requerir normalización post-fusión.

Salud:

7. IDEXX Laboratories Inc. (IDXX):

- Capitalización: ~\$9B.
- Perfil de crecimiento conservador.
- Diagnóstico veterinario - nicho defensivo y de crecimiento.
- Ingresos recurrentes (consumibles, software SaaS).
- Márgenes operativos excepcionales (25-30%).
- Crecimiento consistente 8-12% anual.
- Valor analítico: caso de crecimiento estable - ¿Aquí los modelos simples serán suficientes?
- Datos: excelente calidad, reportes trimestrales muy detallados.

8. Bruker Corporation (BRKR):

- Capitalización: ~\$9B.
- Perfil maduro con exposición a ciclos de investigación.
- Instrumentación científica (life science, materiales).
- Ventas de capital equipment (lumpier) + aftermarket services (recurrente).
- Márgenes operativos 12-16%.
- Sensible a presupuestos académicos y pharma R&D.
- Valor analítico Mix de revenue types, cierta volatilidad en capex de clientes.
- Buenos datos, con desagregación por segmento y geografía.

3. Pequeñas (\$300M-\$2B) - 4 empresas

Tecnología:

9. Qualys Inc. (QLYS):

- Capitalización: ~\$4.5B.
- Perfil de crecimiento en ciberseguridad.
- SaaS puro (plataforma de seguridad en la nube cloud).
- Ingresos recurrentes (95%+), predecibles.
- Márgenes operativos: 15-20% (en expansión).
- Crecimiento: 10-15% anual, consistente
- Es pequeña con características de grande (predictibilidad SaaS), opción experimental.
- Sus datos son excelentes para ser small-cap, informes detallados.

Consumo Discrecional:

10. Crocs Inc. (CROX):

- Capitalización: ~\$6B.
- Perfil de recuperación + crecimiento.
- Caso de recuperación dramático (crisis 2008-2014, recuperación post-2017).

- Muy estacional (Q2-Q3 pico).
- Márgenes variables (5-25% según el año).
- Alta volatilidad - excelente para probar la solidez de los modelos.
- Útil para responder: ¿Pueden los modelos predecir recuperaciones o solo continuidades?
- Datos completos, con historial de crisis incluido.

Industria:

11. Lincoln Electric Holdings Inc. (LECO):

- Capitalización: ~\$7B
- Perfil maduro, sujeto a ciclos industriales
- LECO es un líder de mercado en equipos de soldadura.
- Muy cíclico (construcción, manufactura).
- Dividendo consistente (un llamado "aristocrat" de dividendos).
- Márgenes operativos: 10-14%
- El valor analítico de LECO es que es una empresa de capitalización media estable - nos sirve como benchmark de "sensatez" en el promedio de selecciones para predicción.
- La calidad de sus datos es muy buena, tiene un historial de más de 25 años.

Salud:

12. Neogen Corporation (NEOG):

- Capitalización: ~\$4.5B
- Perfil de crecimiento de nicho (seguridad alimentaria)
- Seguridad alimentaria y sanidad animal.
- Crecimiento orgánico y adquisiciones pequeñas frecuentes.
- Márgenes operativos: 12-18
- Fusión con 3M Food Safety en 2022 (datos históricos separados disponibles).
- Tiene valor analítico, principalmente, como caso de pequeña capitalización con múltiples factores de crecimiento operando (alta complejidad).
- Sus datos son adecuados, pueden requerir ajustes por su reciente fusión.

4. Matriz de Validación de la Selección:

4.1 Diversificación Cuantitativa Estimada:

Métrica	Rango en Selección	Objetivo	✓ / ✗
Revenue CAGR (2013-2023)	-5% a +25%	>20pp spread	✓
Márgenes	5% a 45%	>20pp spread	✓

operativos			
Volatilidad de revenue (CV)	0.05 a 0.35	Diverso	✓
Capitalización	\$4.5B a \$230B	50x ratio	✓

4.2 Características Especiales Capturadas:

Característica	Empresas	Valor Analítico
Estacionalidad fuerte	INTU, NKE, SBUX, CROX	Experimentar con componentes estacionales en modelos
Ingresos recurrente	ADBE, INTU, QLYS, IDXX	Inherentemente predecibles - baseline alto
Exposición FX significativa	NKE, SBUX, BRKR	Feature engineering con "FX rates"
Recuperación/volatilidad	CROX	Solidez en contexto de discontinuidades
M&A reciente	RRX, NEOG	Normalización de datos
Modelos de negocio únicos	IDXX, NEOG	Generalización

4.3 Balance de Complejidad / Predictiva:

Empresas "fáciles" (modelos simples deberían funcionar bien):

- IDXX, QLYS, ADBE, LECO → ~33% de la muestra

Empresas de dificultad "moderada" (sirven casi de grupo de control y/o contraste):

- INTU, SBUX, WWD, BRKR, NKE, NEOG → ~50% de la muestra

Empresas "complejas" (el desempeño de los modelos ML deberían destacar aquí):**

- CROX, RRX → ~17% de la muestra

El balance de la distribución de la selección de empresas de la muestra se estima en 33/50/17. Considero este balance apropiado para una validación empírica.

5. Balance de Disponibilidad:

A modo de confirmación de los criterios obligatorios en el proceso de selección:

Todas tienen:

- 10-K completos 2013-2024 (12 años mínimo).
- Estado de Ingresos, estado de balances, flujos de caja, todos completos.
- Disponibilidad en SEC EDGAR con datos estructurados.
- Sin bancarrotas ni eventos extremos (excepto recuperación controlada, en el caso específico de CROX).

Consideraciones especiales (casos límite incluidos intencionalmente):

- RRX: Fusión 2021 (datos de Regal Entertainment pre-fusión).
- NEOG: Fusión 2022 (datos históricos separados).
- CROX: Período de crisis 2008-2014.

6. Empresas en Reserva

En caso de problemas con alguna de las empresas seleccionadas (y para futuras expansiones de la muestra total), mantengo catalogadas en reserva las siguientes empresas:

- Tecnología (grandes): Autodesk (ADSK), Synopsys (SNPS).
- Consumo (grandes): Ulta Beauty (ULTA), Chipotle (CMG).
- Industria (medianas): AMETEK (AME), Donaldson (DCI).
- Salud (medianas): STERIS (STE), Teleflex (TFX).
- Tecnología (pequeñas): SPS Commerce (SPSC), Paylocity (PCTY).
- Consume (pequeñas): Boot Barn (BOOT), Deckers (DECK).
- Industria (pequeñas): Kadant (KAI), Watts Water (WTS).
- Salud (pequeñas): ICU Medical (ICUI), LeMaitre Vascular (LMAT).



Anexo 3: Marco Integrado de Puntuación Cualitativa

La misión del mismo es transformar el análisis fundamental en un modelo riguroso de 6 dimensiones clave, generalizable y cuasi-cuantitativo que sintetize factores como la calidad y/o valor intrínsecos, el contexto de mercado y las perspectivas en una única puntuación. Esto ha de ser capaz de “refinar” el juicio del analista en parámetros donde se prioriza lo accionable y operativo, minimizando el sesgo subjetivo y conductual mientras se maximiza la profundidad analítica.

El principio fundamental aquí es tender un puente entre el juicio cualitativo y el rigor cuantitativo mediante la operacionalización sistemática, manteniendo la flexibilidad para incorporar conocimiento experto donde sea necesario y conveniente.

1. Arquitectura de Orquestación de Datos

Capa de Datos Primarios:

Tipo de Fuente	Documentos	Frecuencia de Actualización	Extracciones Clave
Presentaciones SEC	10-K, 10-Q, 8-K	Tiempo real (8-K), Trimestral	Datos financieros, factores de riesgo, MD&A, datos de segmentos, concentración
Comunicaciones Corporativas	Transcripciones de resultados, presentaciones a inversores, cartas del CEO	Trimestral + ad hoc	Prioridades estratégicas, sentimiento lingüístico, orientación, tono
Materiales de Proxy	DEF 14A, 13D/G	Anual + eventos	Gobierno corporativo, compensación, propiedad de insiders, composición del consejo

Capa de Datos Secundarios:

Proveedor	Uso Principal	Puntos de Integración
Orbis	Mapeo de cadena de suministro, patentes, datos	Dimensiones C, D

	financieros de competidores	
Morningstar	Calificaciones independientes de ventaja competitiva, estimaciones de valor razonable	Validación Dimensión C
S&P Capital IQ / Bloomberg	Estimaciones consensuadas, tenencias institucionales, benchmarks	Todas las dimensiones
FactSet / Refinitiv	Revisiones en tiempo real, objetivos de precio, sentimiento de analistas	Dimensión F

Datos Propietarios y Alternativos

Tipo de Dato	Fuentes	Aplicación
Modelo de Previsión	Algoritmo privado	Dimensiones A, C (análisis de varianza)
Indicadores Macro	FRED, OCDE, Conference Board	Dimensión C (ciclicidad, beta)
Sentimiento y Conductual	Formulario 4, NLP de noticias, redes sociales, Glassdoor	Dimensión F
Riesgo Geopolítico	Índices de riesgo país, vulnerabilidad de cadena de suministro	Dimensión D

Marco de Calidad de Datos:

- Umbral de Completitud: Mínimo 85% de métricas pobladas para puntuación válida
- Verificación de Oportunidad: Señalar datos obsoletos (>45 días para métricas dinámicas)
- Validación de Consistencia: Referencias cruzadas de métricas superpuestas entre fuentes
- Detección de Valores Atípicos: Señalar valores >3 desviaciones estándar de normas históricas

2. Modelo de Puntuación de Seis Dimensiones

DIMENSIÓN A: Visión Estratégica y Ejecución (25%):

Fundamento Teórico: Porter (1980, 1985), Palepu et al., Tres Horizontes de McKinsey, Rumelt (1991).

A1. Claridad y Coherencia Estratégica (7%):

Subcomponente	Medición	Puntuación (0-100)
Especificidad de Objetivos	Recuento de objetivos a largo plazo cuantificados (ROIC, cuota de mercado, CAGR de ingresos)	3+ objetivos = 100; 2 = 70; 1 = 40; Vago = 10
Consistencia Estratégica	Similitud coseno de mensajes en últimas 8 transcripciones trimestrales (NLP)	>0.85 = 100; 0.70-0.85 = 60; <0.70 = 20
Enfoque Estratégico	Número de prioridades; concentración de ingresos en top 3 segmentos	2-4 prioridades + >75% concentración = 100; 5-6 + 60-75% = 60; 7+ disperso = 20
Balance Orgánico vs. Inorgánico	Ratio (CapEx + I+D) / gasto en M&A durante 3 años	Equilibrado 60/40-80/20 = 100; Sesgado = 40-60; 100% uno = 20

Cálculo: Pesos iguales del 25% para cada subcomponente.

A2. Historial de Ejecución (6%):

Métrica	Fórmula	Umbral
Logro de Hitos	(Logrados / Declarados) durante 3-5 años	>80% = 100; 60-80% = 70; 40-60% = 40; <40% = 10
Precisión de Orientación	Varianza % absoluta promedio (ingresos, EPS) durante 12 trimestres	<5% = 100; 5-10% = 70; 10-20% = 40; >20% = 10
Sesgo de Orientación	% de superaciones vs. errores	Conservador (70-85% superaciones) = 100; Equilibrado = 70; Agresivo = 30
ROI Post-Adquisición	(ROIC adquirido - ROIC pre-adquisición) 3 años después	>300pb = 100; 0-300pb = 60; Negativo = 20

Cálculo: Pesos 30%, 30%, 20%, 20%

A3. Innovación y Eficacia de I+D (6%):

Componente	Operacionalización	Puntuación
------------	--------------------	------------

Productividad de I+D	(% ingresos de nuevos productos últimos 3 años) / (I+D % de ventas)	Ratio >2.5 = 100; 1.5-2.5 = 70; 1.0-1.5 = 40; <1.0 = 10
Profundidad de Pipeline	Tres Horizontes McKinsey: H1 (0-1 año), H2 (1-3 años), H3 (3-5 años)	Fuerte en los 3 = 100; 2 horizontes = 70; 1 horizonte = 40; Limitado = 10
Calidad de Patentes	(Citas por patente) × (Patentes concedidas / solicitadas)	Top cuartil = 100; 2º = 70; 3º = 40; Inferior = 10
Varianza Marco-Consenso	Diferencia: previsión pipeline ALBUR vs. consenso	Marco >20% superior = 100; ±20% = 60; >20% inferior = 20

Cálculo: Pesos iguales del 25%

A4. Estrategia de Asignación de Capital (6%):

Métrica	Evaluación	Benchmark
Tendencia ROIC	Pendiente ROIC a 3 años	Positiva + ROIC>WACC = 100; Plana pero >WACC = 70; Declinante/<WACC = 30
Disciplina de Retorno al Accionista	Ratio de pago FCF + oportunismo de recompra (P/B <1.5)	40-70% + oportunista = 100; En rango, no oportunista = 60; Pobre = 20
Eficiencia de CapEx	ROIC Incremental: $\Delta(\text{NOPAT}) / \Delta(\text{Capital Invertido})$	>20% = 100; 15-20% = 80; 10-15% = 50; <10% = 20

Cálculo: 40% Tendencia ROIC, 30% Retorno Accionista, 30% Eficiencia CapEx

DIMENSIÓN B: Calidad de Gestión y Gobierno Corporativo (20%):

Fundamentos teóricos: Kaplan & Norton (1996), Jensen & Meckling (1976), Estándares de Gobierno ISS.

B1. Capital de Liderazgo y Estabilidad (6%):

Factor	Medición	Puntuación
Permanencia y Desempeño del CEO	Años en el cargo + percentil TSR vs. pares	>5 años + top cuartil = 100; 3-5 años + mediana = 70; <3 años/inferior = 30
Estabilidad del	Permanencia promedio	>7 años promedio, 0 salidas =

Equipo Directivo	C-suite; salidas no planificadas (5 años)	100; 4-7 años, 1 salida = 60; <4 años/2+ salidas = 20
Propiedad Interna	(Valor acciones ejecutivos + consejo) / (Comp anual)	>5× = 100; 2-5× = 70; 1-2× = 40; <1× = 10
Planificación de Sucesión	Sucesión revelada; presencia de sucesor preparado	Plan formal + sucesor = 100; Informal = 60; Sin visibilidad = 20

Cálculo: Pesos 30%, 30%, 30%, 10%

B2. Disciplina de Asignación de Capital (6%):

Métrica	Fórmula	Umbral
Diferencial ROIC vs WACC	Promedio 5 años (ROIC - WACC)	>500pb = 100; 300-500 = 80; 100-300 = 50; <100 = 10
Creación de Valor en M&A	TSR 3 años post-adquisición vs. grupo de pares (acuerdos >10% cap. mercado)	Positivo vs. pares = 100; Neutral = 60; Destrucción de valor = 20
Disciplina de Recompra de Acciones	Correlación entre actividad de recompra y ratio P/B	Correlación negativa (comprar barato) = 100; Sin patrón = 60; Comprar caro = 20

Cálculo: 50% ROIC-WACC, 25% M&A, 25% Recompra.

B3. Alineación de Incentivos (4%):

Componente	Medición	Puntuación
Correlación Pago-Desempeño	R^2 entre bonus ejecutivo/LTIP y métricas (ROIC, FCF, TSR) durante 5 años	$R^2 > 0.6 = 100$; $0.4-0.6 = 70$; $< 0.4 = 30$
Orientación a Largo Plazo	(Valor LTIP / Comp total) para CEO y top 4 ejecutivos	>60% = 100; 40-60% = 70; <40% = 30
Idoneidad de Métricas	Cualitativo: ¿Están las métricas de incentivos alineadas con generadores de valor?	Alineación estratégica = 100; Parcial = 60; Desalineado = 20

Cálculo: Pesos 40%, 30%, 30%

B4. Supervisión e Independencia del Consejo (4%):

Factor	Datos	Puntuación
ISS Governance QualityScore	Directamente de ISS o calcular proxy	Decil 1-2 = 100; 3-5 = 70; 6-8 = 40; 9-10 = 10
Independencia del Consejo	% de consejeros independientes	>80% = 100; 67-80% = 70; 50-67% = 40; <50% = 10
Alineación de Experiencia del Consejo	% con experiencia relevante en industria/función	>60% = 100; 40-60% = 60; <40% = 20
Renovación del Consejo	Permanencia promedio; % consejeros <5 años	Equilibrado (prom 6-10 años, 30-50% nuevos) = 100; Extremos = 40

Cálculo: 40% ISS, 20% cada uno para otros.

DIMENSIÓN C: Contexto de Mercado y Dinámicas de Industria (20%):

Fundamento Teórico: Cinco Fuerzas de Porter (1980), PESTEL, Matriz BCG, Estrategia Competitiva.

C1. Crecimiento de Mercado y Rentabilidad (5%):

Métrica	Cálculo	Puntuación
Tasa de Crecimiento TAM	CAGR a 5 años del mercado total direccionable	>12% = 100; 7-12% = 75; 3-7% = 50; <3% = 20
Penetración SAM	Ingresos empresa / Mercado direccionable servible	<30% (recorrido) = 100; 30-60% = 70; 60-80% = 40; >80% = 10
ROIC de Industria	ROIC mediano del grupo de pares	>15% = 100; 10-15% = 70; 7-10% = 40; <7% = 10

Cálculo: 40% Crecimiento TAM, 30% Penetración, 30% ROIC Industria.

C2. Intensidad Competitiva: Cinco Fuerzas de Porter (6%):

Metodología: Puntuar cada fuerza 1-5 (5 = favorable), luego normalizar a 0-100.

Fuerza	Métricas Clave	Puntuación
Rivalidad (20%)	Índice HHI: <1500 (alta competencia) = 1pt; 1500-2500 = 3pts; >2500 = 5pts Volatilidad margen bruto	Orbis cuota mercado, 10-K
Amenaza de Nuevos Entrantes (20%)	Intensidad de capital (CapEx/Ingresos >20% = 5pts) Barreras regulatorias Fortaleza de marca	10-K, análisis industria
Poder de Proveedores (20%)	Concentración top 5 proveedores: <20% COGS = 5pts; 20-40% = 3pts; >40% = 1pt Costos de cambio	10-K, Orbis
Poder de Compradores (20%)	Concentración top 10 clientes: <20% = 5pts; 20-40% = 3pts; >40% = 1pt Diferenciación de producto	Datos clientes 10-K
Amenaza de Sustitutos (20%)	Amenaza de sustitutos: Baja = 5pts; Moderada = 3pts; Alta = 1pt Elasticidad cruzada si disponible	Análisis industria, tendencias tech

Cálculo: Promedio de cinco fuerzas (pesos iguales 20%), conversión a escala 0-100.

C3. Posicionamiento Competitivo e Integración de Consenso de Analistas (6%):

Proceso Benchmark de tres pasos:

Paso 1: Recopilar datos sobre el consenso de los competidores (precios objetivo estimados, estimaciones EPS, previsiones crecimiento para 5-10 empresas pares, etc).

Paso 2: Calcular métricas implícitas para cada competidor (tasas crecimiento, P/E, EV/EBITDA, PEG); crear rangos cuartiles pares.

Paso 3: Posicionar la empresa objetivo a analizar.

Métrica Relativa	Cálculo	Puntuación
Posición de Valoración	PEG empresa vs. mediana pares	>15% descuento con igual crecimiento = 100; Valor justo ($\pm 10\%$) = 60; >15% prima = 20
Expectativas de Crecimiento	CAGR ingresos consenso vs. mediana pares	Top cuartil = 100; 2º-3º cuartil = 60-80; Inferior = 20
Momento de Revisión de Estimaciones	Revisiones 90 días vs. grupo pares	Líder (más mejoras) = 100; Medio = 60; Rezagado = 20

Análisis de la Varianza:

- Puntuación Convicción = $|\text{Previsión Algoritmo Proprietario} - \text{Consenso}| / \text{Consenso}$
- Alta varianza positiva ($>20\%$) = potencial infravaloración = 100
- Baja varianza ($\pm 10\%$) = precio eficiente = 60
- Alta varianza negativa = potencial sobrevaloración = 20

Cálculo: 30% Valoración, 25% Expectativas Crecimiento, 25% Momento, 20% Algoritmo.

C4. Sensibilidad Regulatoria y Macro (3%):

Factor	Medición	Puntuación
Trayectoria Regulatoria	Desarrollos recientes: Favorable (subsidijs, desregulación) vs. Adverso	Favorable = 100; Neutral = 60; Adverso = 20
Beta Económica	Beta de beneficios (no acciones) al PIB/PMI durante 10 años	Baja ciclicidad ($\beta < 0.7$) = 100; Moderada (0.7-1.3) = 60; Alta (> 1.3) = 20
Exposición FX	% ingresos fuera moneda local × efectividad cobertura	$< 30\%$ O $> 80\%$ cubierto = 100; 30-50% cobertura moderada = 60; $> 50\%$ sin cobertura = 20
Traspaso de Inflación	Comportamiento margen bruto durante períodos inflacionarios	Expansión margen = 100; Estable = 60; Compresión = 20

Cálculo: Pesos iguales del 25%

DIMENSIÓN D: Resiliencia Operacional (15%):

Fundamento teórico: Cadena de Valor de Porter (1985), Modelo SCOR, Gestión de Riesgo de Cadena de Suministro.

D1. Robustez de Cadena de Suministro (4%):

Componente	Medición	Puntuación
Concentración de Proveedores	HHI para base de proveedores O proveedor principal como % de COGS	HHI < 1500 o $< 15\%$ = 100; HHI 1500-2500 o 15-25% = 60; HHI > 2500 o $> 25\%$ = 20
Riesgo de Abastecimiento	% de suministro de regiones alto riesgo	$< 20\%$ = 100; 20-40% = 60; $> 40\%$ = 20

Geográfico		
Gestión de Inventario	Tendencia DIO + frecuencia de desabastecimiento	Mejorando/óptimo + desabastecimientos raros = 100; Estable = 60; DIO creciente/desabastecimientos frecuentes = 20
Digitalización de Cadena de Suministro	Evidencia de tecnología avanzada (IA, seguimiento tiempo real, gemelos digitales)	Líder industria = 100; ERP estándar = 60; Legacy = 20

Cálculo: Pesos 30%, 25%, 25%, 20%

D2. Concentración de Clientes y Poder de Fijación de Precios (4%):

Métrica	Fórmula	Umbral
HHI de Clientes	$HHI = \sum (\text{cuota ingresos cliente})^2$	<1000 = 100; 1000-2000 = 70; 2000-3000 = 40; >3000 = 10
Retención de Clientes	(Clientes año N / año N-1) - abandono	>95% = 100; 90-95% = 80; 85-90% = 50; <85% = 20
Indicador Poder Fijación Precios	Tendencia margen bruto durante inflación costos input	Expansión = 100; Estable ($\pm 100\text{pb}$) = 70; Compresión >200pb = 20

Cálculo: 40% HHI, 30% Retención, 30% Poder Fijación Precios.

D3. Diversificación Geográfica y de Negocio (4%):

Medida	Cálculo	Puntuación
Puntuación Entropía de Ingresos	Entropía Shannon: $H = -\sum (p_i \times \log p_i)$ donde p_i = cuota segmento/región	$H > 2.0$ (altamente diversificado) = 100; 1.5-2.0 = 70; 1.0-1.5 = 40; <1.0 = 10
Correlación de Segmentos	Matriz correlación tasas crecimiento ingresos segmentos	Promedio bajo (<0.4) = 100; Moderado (0.4-0.7) = 60; Alto (>0.7) = 20
Balance Mercados Emergentes	% ingresos de mercados emergentes	15-35% (equilibrado) = 100; 10-15% o 35-50% = 60; <10% o >50% = 30

D4. Eficiencia Operacional (3%):

Ratio	Fórmula	Benchmark
Eficiencia Cadena de Valor	$(SG\&A + I+D) / \text{Beneficio Bruto, tendencia 3 años}$	Declinando hacia mejor cuartil = 100; Estable en mediana = 60; Creciente/peor cuartil = 20
Utilización de Activos	$\text{Ventas} / \text{PPE Neto, tendencia}$	Mejorando rotación activos = 100; Estable = 60; Declinando = 20
Eficiencia Capital Trabajo	Tendencia ciclo conversión efectivo	CCC declinando = 100; Estable = 60; Creciente = 20

Cálculo: Pesos iguales (33% cada uno).

DIMENSIÓN E: Flexibilidad Financiera (10%):

Fundamento Teórico: Damodaran (2012), Z-Score de Altman, Metodologías de Calificación Crediticia.

E1. Liquidez y Estructura de Deuda (3%):

Métrica	Fórmula	Umbral
Apalancamiento Neto	$\text{Deuda Neta} / \text{EBITDA}$	$<1.5\times = 100$; $1.5-2.5\times = 80$; $2.5-3.5\times = 50$; $>3.5\times = 20$
Colchón de Liquidez	$(\text{Efectivo} + \text{Facilidad no utilizada}) / (\text{OpEx semanal} \times 52)$	$>2.0 \text{ años} = 100$; $1.0-2.0 = 70$; $0.5-1.0 = 40$; $<0.5 = 10$
Muro de Vencimiento de Deuda	% de deuda con vencimiento en 2 años	$<15\% = 100$; $15-30\% = 70$; $30-50\% = 40$; $>50\% = 10$
Cobertura de Intereses	$\text{EBIT} / \text{Gasto Intereses}$	$>10\times = 100$; $5-10\times = 80$; $3-5\times = 50$; $<3\times = 20$

Cálculo: Pesos 30%, 25%, 20%, 25%

E2. Durabilidad de Flujo de Caja (3%):

Medida	Cálculo	Puntuación
--------	---------	------------

Conversión FCF	FCF / Beneficio Neto, promedio 5 años	>120% = 100; 90-120% = 80; 70-90% = 50; <70% = 20
Tendencia Margen FCF	FCF / Ingresos, pendiente 3 años	Pendiente positiva + margen >10% = 100; Estable en 7-10% = 70; Declinando/<7% = 30
Volatilidad FCF	Coefficiente variación (σ/μ) de FCF durante 5 años	CV <0.3 (estable) = 100; 0.3-0.6 = 60; >0.6 (volátil) = 20

Cálculo: Pesos 40%, 35%, 25%

E3. Acceso a Mercados de Capitales (2%):

Factor	Evaluación	Puntuación
Calificación Crediticia Implícita	Calcular basado en apalancamiento, cobertura; comparar con real	A- o mejor = 100; Rango BBB = 70; Rango BB = 40; B o inferior = 10
Trayectoria WACC	WACC actual vs. promedio 3 años	WACC declinando = 100; Estable = 60; Creciente = 20
Éxito Emisión Reciente	Última emisión deuda/capital: términos vs. comparables	Términos favorables = 100; Términos mercado = 60; Precios distressed = 10

Cálculo: Pesos 40%, 30%, 30%

E4. Opcionalidad Financiera (2%):

Componente	Medida	Umbral
Liquidez de Activos	Activos no core, comercializables / Valor Empresa	>15% = 100; 10-15% = 70; 5-10% = 40; <5% = 10
Activos no Gravados	% de activos libres de gravámenes/prendas	>80% = 100; 60-80% = 70; 40-60% = 40; <40% = 10
Margen de Covenants	Distancia al covenant financiero más cercano como % colchón	>30% margen = 100; 15-30% = 60; <15% = 20

Cálculo: Pesos iguales (33% cada uno)

DIMENSIÓN F: Sentimiento de Stakeholders y Momentum (10%):

Fundamento Teórico: Finanzas Conductuales (Kahneman & Tversky), Teoría de Señales, Teoría de Stakeholders.

F1. Momentum de Analistas Sell-Side (3%):

Indicador	Medición	Puntuación
Momentum de Cambio de Calificación	Cambios netos (mejoras - rebajas) últimos 90 días	Neto +3+ = 100; +1 a +2 = 75; Neutral = 50; -1 a -2 = 25; -3+ = 0
Dirección Revisión Estimaciones	% de analistas elevando estimaciones EPS (últimos 90 días)	>70% elevando = 100; 50-70% = 70; 30-50% = 40; <30% = 10
Análisis Tono de Sentimiento	Análisis NLP de informes analistas	Consenso alcista = 100; Neutral = 60; Bajista = 20

Cálculo: Pesos 40%, 40%, 20%

F2. Posicionamiento Institucional (3%):

Factor	Fuente de Datos	Umbral
Tendencia Propiedad Institucional	Cambio QoQ presentaciones 13F	Aumentando >2% = 100; Estable $\pm 2\%$ = 60; Disminuyendo >2% = 20
Actividad Smart Money	Inversores activistas/valor iniciando posiciones	Interés reciente alta calidad = 100; Estable = 60; Salidas = 20
Tendencia Interés Corto	Short % del float, cambio vs. 6 meses antes	Declinando = 100; Estable = 60; Creciente = 20

Cálculo: Pesos 40%, 30%, 30%

F3. Sentimiento de Empleados (2%):

Métrica	Fuentes	Evaluación
Calificaciones Glassdoor/Indeed	Sitios de reseñas de empleados	>4.0 calificación = 100; 3.0-4.0 = 60; <3.0 = 20

Tasa de Rotación	Datos LinkedIn, divulgaciones empresa	<10% voluntaria = 100; 10-20% = 60; >20% = 20
Aprobación CEO	Aprobación CEO en Glassdoor	>80% = 100; 60-80% = 60; <60% = 20

Cálculo: Pesos 40%, 40%, 20%

F4. Sentimiento de Clientes y Marca (2%):

Indicador	Fuentes de Datos	Puntuación
Net Promoter Score	Divulgaciones empresa, encuestas	>50 = 100; 30-50 = 60; <30 = 20
Trayectoria Valor de Marca	Rankings Interbrand, Brand Finance	Creciente = 100; Estable = 60; Declinando = 20
Sentimiento Social	Analíticas redes sociales, agregadores de reseñas	Neto positivo = 100; Neutral = 60; Negativo = 20

Cálculo: Pesos 40%, 30%, 30%

3. Integración y Cálculo de Puntuación

Proceso completo paso a paso

Paso 1: Recopilación y Normalización de Datos:

- Reunir todas las métricas cuantitativas y normalizar a escala 0-100
- Para puntuaciones cualitativas (1-5), convertir: 5=100, 4=80, 3=60, 2=40, 1=20

Paso 2: Calcular Puntuaciones de Subcomponentes:

- Para cada componente granular, aplicar promedio ponderado de sus métricas
- Ejemplo: A1 Claridad Estratégica = (Especificidad Objetivos × 0.25) + (Consistencia × 0.25) + (Enfoque × 0.25) + (Balance × 0.25)

Paso 3: Calcular Puntuaciones Dimensionales:

- Sumar subcomponentes ponderados dentro de cada dimensión
- Ejemplo: Puntuación Dimensión A = (A1 × 0.28) + (A2 × 0.24) + (A3 × 0.24) + (A4 × 0.24)
- Nota: Los pesos son ajustados para que los subcomponentes sumen el peso total de la dimensión (25% para A).

Paso 4: Calcular Puntuación Refinada Final:

Puntuación Refinada Final = $\Sigma(\text{Puntuación Dimensión} \times \text{Peso Dimensión})$

Dónde Pesos de Dimensiones:

A: Visión Estratégica y Ejecución = 25%

B: Calidad de Gestión y Gobierno = 20%

C: Contexto de Mercado y Dinámicas de Industria = 20%

D: Resiliencia Operacional = 15%

E: Flexibilidad Financiera = 10%

F: Sentimiento de Stakeholders y Momentum = 10%

Resultado: Puntuación en escala 0-100

Paso 5: Verificaciones de Varianza y Calidad:

Tipo de Verificación	Metodología	Umbral de Acción
Compleitud de Datos	% de métricas con datos	<85% = puntuación inválida, señalar para revisión
Consistencia Interna	Correlación entre métricas relacionadas	Patrones contraintuitivos = señalar
Comparación con Pares	Puntuación vs. mediana industria	Valores atípicos (± 2 desv est) = destacar
Estabilidad Temporal	Cambio puntuación QoQ	>20% cambio sin catalizador = señalar
Brecha de Consenso	Análisis de divergencia	>30% varianza = señal alta convicción

4. Archivos de Salida:

1. Tarjeta de Calificación Puntuación Final Refinada:

- Puntuación Global: X/100 con rango percentil vs. grupo pares.
- Puntuaciones Dimensionales: Puntuaciones individuales A-F.
- Análisis de Tendencia: Cambios puntuación QoQ e interanual.
- Factores Clave: Top 3 fortalezas y top 3 debilidades.

2. Visualización: Gráfico Araña

Gráfico radial de seis ejes mostrando el desempeño aproximado en todas las dimensiones, superpuesto con:

- Puntuación actual empresa (línea sólida).
- Mediana industria (línea discontinua).
- Benchmark top cuartil (región sombreada).

3. Matriz de Posicionamiento Competitivo:

Matriz 2x2:

- Eje X: Puntuación Refinada Final
- Eje Y: Valoración de Mercado (P/E relativo a crecimiento)
- Cuadrantes: Calidad Infravalorada | Calidad Valoración Justa | Calidad Sobrevalorada | Trampas de Valor.

4. Informe de Varianza:

Análisis detallado de dónde diverge modelo algorítmico del consenso:

- Convicción Positiva: Algoritmo >20% por encima del consenso (potencial infravaloración).
- Zona Neutral: Varianza $\pm 20\%$ (precio eficiente)
- Convicción Negativa: Algoritmo >20% por debajo del consenso (potencial sobrevaloración).

5. Señales de Riesgo y Alertas:

Alertas automatizadas para:

- Dimensiones en declive (>15 puntos caída QoQ).
- Preocupaciones gobierno (puntuación B4 <40).
- Indicadores estrés financiero (puntuación E1 <40).
- Inflexiones sentimiento negativo (rebajas F1 acelerándose).

5. Actualización y Mantenimiento:

Componente	Frecuencia de Actualización	Eventos Desencadenantes
Tiempo Real	Continuo	Revisiones analistas, cambios precio objetivo, flujos institucionales
Trimestral	Post-resultados	Métricas financieras, transcripciones de conferencias
Anual	Post-10K	Reevaluación cualitativa completa, cambios gobierno, estrategia
Ad Hoc	Según eventos	M&A, cambios CEO/CFO, emisiones capital, eventos regulatorios

Marco de Control de Calidad:

Verificación	Metodología	Umbral
Compleitud de Datos	% de métricas con datos disponibles	>90% requerido para puntuación válida

Universitat Oberta de Catalunya		uoc.edu
Consistencia Interna	Verificaciones correlación entre métricas relacionadas	Señalar si patrones contraintuitivos
Comparación con Pares	Puntuación vs. mediana industria	Destacar valores atípicos (± 2 desv est)
Estabilidad Temporal	Cambio puntuación QoQ	Señalar si $>20\%$ cambio sin catalizador claro

Resultados Esperados:

En general, un informe de puntuación refinada detallado. En estos momentos sólo funciona y da resultados parciales. En una versión definitiva esperaríamos que el informe contuviera los siguientes ítems:

- Resumen Ejecutivo: Puntuación global, rango percentil, fortalezas/debilidades clave identificadas.
- Mapa de Calor Dimensional: Representación visual de 6 dimensiones.
- Comparación con Pares: Puntuaciones vs. top 5 competidores.
- Análisis de Cambio: Evolución puntuación últimos 4 trimestres.
- Señales de Riesgo: Alertas automáticas para dimensiones en declive.
- Ítems de Acción: rutas de objetivos accionables más salientes.

Respecto a la frecuencia de actualización apropiada para tal sistema:

- Componentes tiempo real: Revisiones analistas, objetivos precio, flujos institucionales.
- Actualizaciones trimestrales: Métricas financieras post-resultados.
- Profundización anual: Reevaluación cualitativa completa post cada 10K oficial de la empresa.

6. Mejoras Futuras

Hipotéticamente el marco de puntuación cualitativo y su metodología podría mejorarse de las siguientes maneras:

6.1 Integración de Machine Learning:

- Modelado Predictivo: Entrenar modelos con puntuaciones históricas vs. retornos futuros.
- Procesamiento Lenguaje Natural: Extracción automatizada de sentimiento de transcripciones.
- Detección de Anomalías: Identificar patrones inusuales que requieren revisión humana.

6.2 Adopción de Análisis de Escenarios:

- Análisis de sensibilidad por parámetros.
- Casos Alcista/Base/Bajista: Ajustar puntuaciones bajo diferentes escenarios macro.
- Modelado Impacto de Eventos: Cuantificar cambios de puntuación por M&A, transiciones CEO, etc.

6.3 Capa de Integración ESG (Complemento Opcional 11º Percentil):

- Puntuación ambiental (huella carbono, eficiencia recursos).
- Puntuación social (prácticas laborales, impacto comunitario).
- Puntuación gobierno (ya parcialmente cubierto en Dimensión B).

En resumen, este marco proporciona un enfoque integral y operacionalizado del análisis fundamental que combina rigor cuantitativo con profundidad cualitativa, permitiendo comparación sistemática entre empresas e industrias a la vez que mantiene el matiz analítico.

7. Nota Final sobre la Puntuación Cualitativa:

Puede argumentarse, a mi ver no sin razón, que una vez se han realizado todas estas operaciones, el resultado es una métrica cualitativa hartamente abstracta, que puede parecer no medir nada específico que podamos señalar, salvo un vago promedio de “calidad” o “desempeño” indefinidos .

Mi respuesta a esta cuestión es que, como es el caso de una buena parte de las métricas analíticas existentes, es una herramienta heurística, discrecional y que su significación depende fuertemente del contexto en el que se de. Esto, lejos de ser una debilidad, me parece una gran fortaleza y explica su capacidad de generalización y versatilidad.

De nuevo, como prácticamente todas las demás herramientas e indicadores analíticos, no es una señal binaria de “sí” o “no”, inequívoca, independiente de las circunstancias que lo rodean o carente de matices y profundidad. El hecho de que se pueda esperar lo contrario revela una incrustada confusión respecto a cuál es la labor analítica de datos en general, una tendencia a la reificación de los datos como si fueran la realidad misma y no un ensamblaje teórico útil para interrogar a la propia realidad y a razón de esta interrogación generar hipótesis guiadas y competentes y lo que pueda representar el indicador en cada caso, varía. No confundir mapa y territorio, en resumen.

Dicho esto, creo útil plasmar mi propia interpretación general de lo que puede estar midiendo este indicador: un promedio aproximado de la relación entre potencial y actualización de la empresa. Una métrica de optimización integral de los recursos y oportunidades disponibles para la organización.

Siendo así, pueden darse (y de hecho se dan en la muestra analizada) casos de empresas cuya puntuación es pobre pero sus ingresos aumentan y de empresas cuya puntuación es alta pero sus ingresos decaen. Lejos de impugnar al indicador, estos dos hipotéticos (si coinciden empíricamente con la dirección de los pronósticos, claro, si podemos decir que aciertan en corresponder con dichos escenarios) lo que pueden estar revelando es, en el

primer caso, una empresa rentable pese a no estar optimizando sus recursos y oportunidades (y que por tanto aún puede esperar crecer más agresivamente todavía si se reorganiza) y, en el segundo caso, una empresa que aún habiendo exprimido y actualizado todo su potencial, no lo refleja en su desempeño. Quizás porque la competencia en ese mercado es demasiado fuerte o quizás porque el mismo modelo de negocio es deficiente.

En ambos casos, no obstante, vemos que lejos de resultar refutado por el contra-ejemplo, el puntaje del marco ayuda a revelar insights relevantes, agudos, profundamente accionables y contextualizados para las situaciones específicas donde se dan. Exactamente lo que considero que las mejores prácticas y resultados del análisis de datos profesional deben aportar.

Sobre la compleja relación entre teoría, epistemología y datos, a modo de ilustración adicional, adjunto aquí artículos y libros en varios diferentes saberes y aproximaciones, que o bien tocan esta problemática de manera relevante o bien sirven de influencia a como me he posicionado respecto a esta cuestión:

- Barrowman, N. "Why Data Is Never Raw." The New Atlantis.
<https://www.thenewatlantis.com/publications/why-data-is-never-raw>
- "Data and Reality: Implications for Record-Keeping by Organizations."
<https://ambur.net/datareality.pdf>
- "Data Modeling - Theory vs Reality." Practical Data Modeling (Substack).
<https://practicaldatamodeling.substack.com/p/data-modeling-theory-vs-reality>
- Egger, R. & Yu, J. (2022). "Epistemological Challenges: Is the Future Theory-Driven or Data-Driven?" In [Book title needed], pp. [page range]. DOI: 10.1007/978-3-030-88389-8_2.
https://www.researchgate.net/publication/358254095_Epistemological_Challenges_Is_the_Future_Theory-Driven_or_Data-Driven
- "From the Editors: What Grounded Theory is Not." Academy of Management Journal. <https://journals.aom.org/doi/10.5465/amj.2006.22083020>
- "'It's Data, Not Reality': On Situated Data with Jill Walker Rettberg." Nightingale (Medium).
<https://medium.com/nightingale/its-data-not-reality-on-situated-data-with-jill-walker-rettberg-d27c71b0b451>
- Kent, W. Data and Reality.
<https://cmpct.info/~calvin/Papers/Data%20and%20Reality.pdf>
- Leetaru, K. (2019, May 7). "Data Isn't 'Truth'." Forbes.
<https://www.forbes.com/sites/kalevleetaru/2019/05/07/data-isnt-truth/>
- Leifer, E. M. "Denying the Data: Learning from the Accomplished Sciences." [Journal name and volume/issue needed]. <https://www.jstor.org/stable/684311>
- Mazzocchi, F. (2015). "Could Big Data be the end of theory in science?" EMBO Reports, 16(10), 1250–1255. <https://doi.org/10.15252/embr.201541001>
- "Myths of Modelling: Data speak." Towards Data Science.
<https://towardsdatascience.com/myths-of-modelling-data-speak-358805890baa/>

- Pearson, E. S. (1955). "Statistical Concepts in Relation to Reality." Journal of the Royal Statistical Society, Series B (Methodological), 17(2), 204–207.
<https://www.jstor.org/stable/2983954>
- Peng, R. (2018, December 11). "The Role of Theory in Data Analysis." Simply Statistics.
<https://simplystatistics.org/posts/2018-12-11-the-role-of-theory-in-data-analysis/>
- Taleb, N. N.
 - Fooled by Randomness: The Hidden Role of Chance in Life and in the Markets (2001)
 - The Black Swan: The Impact of the Highly Improbable (2007)
 - Antifragile: Things That Gain from Disorder (2012)
 - Skin in the Game: Hidden Asymmetries in Daily Life (2018)
 - The Bed of Procrustes: Philosophical and Practical Aphorisms (2010)
- Statistical Consequences of Fat Tails: Real World Preasymptotics, Epistemology, and Applications (2020)
- "The End of Theory." Wired (2008, June). <https://www.wired.com/2008/06/pb-theory/>
- "The illusion of data validity: Why numbers about people are likely wrong." [Journal name needed].
<https://www.sciencedirect.com/science/article/pii/S2543925122001188>
- Tsamados, A., et al. "A big data state of mind: Epistemological challenges to accountability and transparency in data-driven regulation." Government Information Quarterly. <https://www.sciencedirect.com/science/article/pii/S0740624X21000149>
- Vickers, P., et al. "Data models, representation and adequacy-for-purpose." PMC (National Library of Medicine). <https://pmc.ncbi.nlm.nih.gov/articles/PMC7846529/>
- "Why Data Is an Incomplete Representation of Reality." Machine Learning Made Simple (Medium).
<https://machine-learning-made-simple.medium.com/why-data-is-an-incomplete-representation-of-reality-5204dd12bc73>



Anexo 4 - Muestras del Código

4.1 Selecciones del Código - Modelo y Pronóstico Naïve

Enfoque A: Previsión Independiente de Métricas:

La idea en este primer enfoque era pronosticar cada métrica como serie temporal separada. Esto es más rápido, más simple, permite una selección de modelo específico por cada métrica. El problema son las inconsistencias internas entre estados pronosticados.

Lo siguiente son extractos relevantes del código `Naive_forecast.py`, seleccionados y organizados según su importancia metodológica:

4.1.1 Metodología Central de Pronóstico:

4.1.1.1 Algoritmo de Extrapolación (Naive):

- Calcula el cambio porcentual promedio anual a partir de datos históricos:

$$\text{change_pct} = (\text{historical_values}[j] - \text{prev}) / \text{prev}$$
- Aplica esta tasa promedio hacia adelante:

$$\text{predicted_val} = \text{last_value} * (1 + \text{avg_change_rate})$$
- Preserva la dirección del signo para métricas negativas (gastos, pérdidas).

```
changes = []
for j in range(1, len(historical_values)):
    prev = historical_values[j-1]
    if prev != 0:
        change_pct = (historical_values[j] - prev) / prev
        changes.append(change_pct)

if not changes:
    last_val = historical_values[-1]
    for year_col in forecast_year_cols:
        forecast_df.at[idx, year_col] = last_val
    continue

filtered_changes = [c for c in changes if -5.0 <= c <= 5.0]
if filtered_changes:
    avg_change_rate = sum(filtered_changes) / len(filtered_changes)
else:
    avg_change_rate = 0.0
```


4.1.1.2 Variantes Ajustadas por Puntuación:

- La entrada de puntuación cualitativa (escala 0-10) modula la agresividad del pronóstico.
- Fórmula de ajuste:

$$\text{score_adj} = (\text{float}(\text{self.qual_score}) - 5.0) / 5.0 \text{ (oscila entre -1.0 y +1.0)}$$
- Diferentes factores de sensibilidad por tipo de métrica:
 - Métricas de beneficio: 0.1498 (conservador)
 - Capital de trabajo: 0.3370 (agresivo)
 - Otras métricas: 0.2246 (estándar)
- Límite máximo de ajuste del 7% con manejo direccional inteligente.

```

score_adj = 0.0
adjusted_change_rate = avg_change_rate
if self.qual_score is not None:
    try:
        score_adj = (float(self.qual_score) - 5.0) / 5.0
        if is_profit_metric:
            base_adjustment_factor = 0.1498
        elif is_working_capital:
            base_adjustment_factor = 0.3370
        else:
            base_adjustment_factor = 0.2246
        adjustment_factor = score_adj * base_adjustment_factor
        max_adjustment = 0.07
        adjustment_factor = max(-max_adjustment, min(max_adjustment,
adjusted_change_rate + adjustment_factor))

    if is_negative_metric and adjustment_factor != 0:
        adjustment_factor = -adjustment_factor
    adjusted_change_rate = avg_change_rate + adjustment_factor
    try:
        from forecast_config import ForecastConfig
        cfg = ForecastConfig()
        growth_cap = cfg.score_adjustments.growth_multiplier
    except Exception:
        growth_cap = 0.07
    if avg_change_rate == 0:
        adjusted_change_rate = adjustment_factor * 0.1
    else:
        factor = adjusted_change_rate / avg_change_rate
        min_factor = 1.0 - abs(growth_cap)
        max_factor = 1.0 + abs(growth_cap)
        if factor < min_factor:
            adjusted_change_rate = avg_change_rate * min_factor
        elif factor > max_factor:
            adjusted_change_rate = avg_change_rate * max_factor
    except Exception:
        adjusted_change_rate = avg_change_rate

```


4.1.2 Marco de Validación:

4.1.2.1 Validación Cruzada Móvil:

- Enfoque de ventana expansiva: cada año de pronóstico utiliza todos los datos históricos previos
- Análisis separado por horizontes temporales (3 años y 5 años)
- Conjunto integral de métricas:
 - MAE, RMSE, MAPE, sMAPE, MASE, R^2
 - Precisión direccional
 - Intervalos de confianza del 95%

```
for horizon_name, max_forecast_years in horizons.items():
    f.write(f"{horizon_name.upper()} FORECAST HORIZON\n")
    f.write("-" * 40 + "\n\n")
    horizon_results[horizon_name] = {}
    for name, df in data.items():
        f.write(f"{name.upper()}\n")

        def extract_year(col):
            try:
                return int(str(col).split(' ')[0])
            except:
                return 0

        year_cols = [col for col in df.columns if col != 'Category']
        year_cols_sorted = sorted(year_cols, key=extract_year)

        if len(year_cols_sorted) < 3:
            f.write(f"Sample Size: 0 metrics evaluated (insufficient data)\n\n")
            continue

        all_actuals = []
        all_predictions = []
        all_directions_actual = []
        all_directions_predicted = []

        valid_metrics = 0
        total_metrics = 0

        for idx, row in df.iterrows():
            category = row['Category']
```



```

        if pd.isna(category) or str(category).startswith('📊') or
str(category).startswith('Forecast'):
            continue
        total_metrics += 1
        metric_actuals = []
        metric_predictions = []
        metric_directions_actual = []
        metric_directions_predicted = []

        max_i = min(len(year_cols_sorted), max_forecast_years + 2)

        for i in range(2, max_i):
            forecast_year_col = year_cols_sorted[i]
            historical_cols = year_cols_sorted[:i]
            actual_val = row[forecast_year_col]
            if pd.isna(actual_val):
                continue
            try:
                actual_val = float(actual_val)
            except (ValueError, TypeError):
                continue

            historical_values = []
            for col in historical_cols:
                val = row[col]
                if pd.notna(val):
                    try:
                        num_val = float(val)
                        historical_values.append(num_val)
                    except (ValueError, TypeError):
                        continue
            if len(historical_values) < 2:
                continue

```



```

changes = []
for j in range(1, len(historical_values)):
    prev = historical_values[j-1]
    if prev != 0:
        change_pct = (historical_values[j] - prev) / prev
        changes.append(change_pct)

if not changes:
    predicted_val = historical_values[-1]
else:
    avg_change_rate = sum(changes) / len(changes)
    last_value = historical_values[-1]
    predicted_val = last_value * (1 + avg_change_rate)

metric_actuals.append(actual_val)
metric_predictions.append(predicted_val)

if len(historical_values) >= 2:
    actual_change = actual_val - historical_values[-1]
    predicted_change = predicted_val - historical_values[-1]
    actual_direction = 1 if actual_change > 0 else (-1 if
actual_change < 0 else 0)
    predicted_direction = 1 if predicted_change > 0 else (-1 if
predicted_change < 0 else 0)

    metric_directions_actual.append(actual_direction)
    metric_directions_predicted.append(predicted_direction)

if metric_actuals:
    valid_metrics += 1
    all_actuals.extend(metric_actuals)
    all_predictions.extend(metric_predictions)
    all_directions_actual.extend(metric_directions_actual)
    all_directions_predicted.extend(metric_directions_predicted)

```



```

    if not all_actuals:
        f.write(f" Sample Size: 0 metrics evaluated\n\n")
        horizon_results[horizon_name][name] = None
        continue

    actuals = np.array(all_actuals)
    predictions = np.array(all_predictions)

    mae = np.mean(np.abs(actuals - predictions))
    rmse = np.sqrt(np.mean((actuals - predictions) ** 2))

    non_zero_mask = actuals != 0
    if np.any(non_zero_mask):
        mape = np.mean(np.abs((actuals[non_zero_mask] -
predictions[non_zero_mask]) / actuals[non_zero_mask])) * 100
    else:
        mape = float('inf')

    smape = np.mean(2 * np.abs(actuals - predictions) / (np.abs(actuals) +
np.abs(predictions))) * 100

    naive_errors = []
    for i in range(1, len(actuals)):
        if i < len(all_actuals) and i-1 < len(all_actuals):
            naive_forecast = all_actuals[i-1]
            naive_actual = all_actuals[i]
            naive_errors.append(abs(naive_actual - naive_forecast))

    if naive_errors:
        naive_mae = np.mean(naive_errors)
        mase = mae / naive_mae if naive_mae != 0 else float('inf')
    else:
        mase = float('inf')

    ss_res = np.sum((actuals - predictions) ** 2)
    ss_tot = np.sum((actuals - np.mean(actuals)) ** 2)
    r2 = 1 - (ss_res / ss_tot) if ss_tot != 0 else float('inf')

```



```

        direction_matches = sum(1 for a, p in zip(all_directions_actual,
all_directions_predicted) if a == p)
        directional_accuracy = (direction_matches / len(all_directions_actual)) *
100 if all_directions_actual else 0

    horizon_results[horizon_name][name] = {
        'mae': mae, 'rmse': rmse, 'mape': mape, 'smape': smape,
        'mase': mase, 'r2': r2, 'directional_accuracy': directional_accuracy,
        'sample_size': valid_metrics, 'total_predictions': len(all_actuals)
    }

    f.write(f" Sample Size: {valid_metrics} metrics evaluated\n")
    f.write(f" MAE (Mean Absolute Error):          {mae:,.0f}\n")
    f.write(f" RMSE (Root Mean Square Error):        {rmse:,.0f}\n")
    if mape != float('inf'):
        f.write(f" MAPE (Mean Abs % Error):              {mape:.1f}%\n")
    else:
        f.write(f" MAPE (Mean Abs % Error):              100.0%\n")
    f.write(f" sMAPE (Symmetric MAPE):                      {smape:.1f}%\n")
    f.write(f" MASE (Mean Abs Scaled Error):                  {mase:.2f}\n")
    f.write(f" R² (Coefficient of Determination):            {r2:.3f}\n")
    f.write(f" Directional Accuracy:
{directional_accuracy:.1f}%\n\n")

```

4.1.2.2 Cálculo de MASE:

- Compara contra un punto de referencia naive (valor anterior como pronóstico)
- Métrica de error independiente de la escala
- Valores < 1 indican mejor que el método naive

```

if naive_errors:
    naive_mae = np.mean(naive_errors)
    mase = mae / naive_mae if naive_mae != 0 else float('inf')
else:
    mase = float('inf')

```

4.1.3 Procesamiento de Datos:

4.1.3.1 Procesado de Excel:

- Algoritmos de múltiple detección de encabezados.
- Identificación automática de columnas de años.

- Estrategias de respaldo/redundancia para diferentes formatos de Excel.

```
def load_data(self) -> bool:
    """Load all three financial statement files."""
    print("Loading Naive forecast files...")

    for name, filename in self.files.items():
        filepath = os.path.join(self.base_path, filename)
        if not os.path.exists(filepath):
            print(f"Error: File {filepath} not found.")
            return False

    try:
        df = None
        for header_row in [2, 1, 0, None]:
            try:
                temp_df = pd.read_excel(filepath, header=header_row)
                # Check if we have year-like columns
                has_years = False
                for col in temp_df.columns:
                    col_str = str(col).strip().split(' ')[0]
                    try:
                        year = int(col_str)
                        if 1900 <= year <= 2100:
                            has_years = True
                            break
                    except ValueError:
                        continue

                if has_years:
                    df = temp_df
                    print(f"✓ Loaded {name} from {filename} with header={header_row}")
                    break
            except Exception:
                if self.debug:
                    print(f"Debug: Failed to load {filename} with header={header_row}")
                continue
```



```

        if df is None:
            # Fallback to header=2
            df = pd.read_excel(filepath, header=2)
            print(f"✓ Loaded {name} from {filename} (fallback to header=2)")

        self.dataframes[name] = df

    except Exception as e:
        print(f"Error loading {filename}: {e}")
        return False

    return True

```

4.1.3.2 Integración de Estados Financieros:

- Procesa los tres tipos de estados financieros simultáneamente.
- Mantiene la consistencia entre estado de resultados, balance general y flujo de efectivo.

```

class NaiveForecastProcessor:
    def __init__(self, base_path: str = "analysis_outputs/Forecasts/Naive", debug: bool = False):
        self.base_path = base_path
        self.debug = debug
        self.qual_score: Optional[float] = None

        self.files = {
            'income_statement': 'income_statement_naive.xls',
            'balance_sheet': 'balance_sheet_naive.xls',
            'cash_flow': 'cash_flow_naive.xls'
        }

        self.dataframes: Dict[str, pd.DataFrame] = {}
        self.year_columns: List[str] = []
        self.score_adjusted_rows: Dict[str, List[Dict[str, Any]]] = {}

```


4.1.4 Implementación Técnica:

4.1.4.1 Pronóstico Secuencial:

- Cada año de pronóstico se basa en la predicción anterior.
- Mantiene dependencias temporales en pronósticos multi-anuales.

```
cat_lower = str(category).lower()
wants_score_adjust = any(k in cat_lower for k in score_adjust_keywords)

effective_base_rate = avg_change_rate * dampening_factor
effective_adjusted_rate = adjusted_change_rate * dampening_factor

current_value = historical_values[-1]
for year_col in sorted_forecast_cols:
    forecast_value = current_value * (1 + effective_base_rate)
    forecast_df.at[idx, year_col] = forecast_value
    current_value = forecast_value
```

Estos elementos buscan representar una línea base de análisis tradicional que, aunque lo denomino aquí como "Naive", va más allá de un simple pronóstico del "último valor", incorporando conocimiento del campo, rigor estadístico y validación integral. Esto hace que sea adecuada de comparar con modelos de ML más complejos.

4.2 Selecciones del Código - ML & Pronóstico: Regresión Lineal

Los siguientes son los puntos clave que destaco de mi archivo de código Linear_forecast.py, organizados por importancia metodológica:

4.2.1 Metodología Central de Pronóstico

4.2.1.1 Regresión Lineal Guiada por Configuración predefinida:

- Sistema de configuración predefinido en JSON (linear_regression_forecast_config.json).
- Selección flexible de modelo: OLS estándar vs. regresión robusta RANSAC.
- Características polinomiales opcionales con grado configurable.
- Personalización de parámetros de Scikit-learn (fit_intercept, copy_X, n_jobs).

```
def
load_linear_regression_forecast_config(config_file="linear_regression_forecast_config.json"):
    """
    Load Linear Regression forecast configuration from JSON file

    Args:
        config_file: Path to configuration file (default:
linear_regression_forecast_config.json)

    Returns:
        dict: Configuration dictionary|
    """
    try:
        with open(config_file, 'r') as f:
            config = json.load(f)
            print(f"✓ Loaded configuration from {config_file}")
            return config
    except FileNotFoundError:
        print(f"Warning: Configuration file {config_file} not found. Using default settings.")
        return {}
    except Exception as e:
        print(f"Warning: Error loading configuration: {e}. Using default settings.")
        return {}
```

4.2.1.2 Pronóstico Secuencial Multi-Etapa:

- Cada año de pronóstico se fundamenta en predicciones anteriores.
- Regresión indexada por tiempo:
 - X = np.array(range(len(current_values))).reshape(-1, 1)
- Mantiene las dependencias temporales a través de horizontes de pronóstico.


```

        historical_values = []
        for col in historical_cols:
            val = row[col]
            if pd.notna(val):
                try:
                    num_val = pd.to_numeric(val, errors='coerce')
                    if pd.notna(num_val):
                        historical_values.append(num_val)
                except:
                    continue

        if len(historical_values) < 2:
            continue
        current_values = historical_values.copy()

        for year_col in sorted_forecast_cols:
            forecast_year = extract_year_for_sort(year_col)

            if len(current_values) < 2:
                forecast_df.at[idx, year_col] = current_values[-1] if current_values
            else None

            current_values.append(current_values[-1] if current_values else 0)
            continue

        X = np.array(range(len(current_values))).reshape(-1, 1)
        y = np.array(current_values)

        if self.preprocessing_settings.get('apply_data_smoothing', False):
            y = self._apply_data_smoothing(y)

```



```

if self.preprocessing_settings.get('handle_outliers', False):
    X, y = self._handle_outliers(X, y)
try:
    model = self._create_regression_model()
    model.fit(X, y)
    next_prediction = model.predict([[len(current_values)]])[0]
    forecast_value = self._validate_and_adjust_forecast(
        next_prediction, y, forecast_year, category
    )

    forecast_df.at[idx, year_col] = forecast_value
    current_values.append(forecast_value)

except Exception as e:
    if self.debug or self.error_settings.get('log_errors', True):
        print(f"Error: Linear regression failed for {category}: {e}")

    if self.error_settings.get('fallback_to_constant_extrapolation', True):
        last_value = current_values[-1]
        forecast_df.at[idx, year_col] = last_value
        current_values.append(last_value)
    else:
        min_value = self.general_settings.get('minimum_forecast_value', 0.1)
        forecast_df.at[idx, year_col] = min_value
        current_values.append(min_value)

```

4.2.2 Pipeline Avanzada de Preprocesamiento de Datos

4.2.2.1 Tratamiento de Valores Atípicos:

- Detección de valores atípicos por IQR con umbral configurable (predeterminado 1.5).
- Eliminación automática de valores extremos antes del ajuste del modelo.
- Preserva la integridad de los datos para un modelado robusto.


```
def _handle_outliers(self, X: np.ndarray, y: np.ndarray) -> Tuple[np.ndarray, np.ndarray]:
    """Remove outliers based on configuration settings."""
    method = self.preprocessing_settings.get('outlier_method', 'IQR')
    threshold = self.preprocessing_settings.get('outlier_threshold', 1.5)

    if method == 'IQR' and len(y) > 4:
        Q1 = np.percentile(y, 25)
        Q3 = np.percentile(y, 75)
        IQR = Q3 - Q1
        lower_bound = Q1 - threshold * IQR
        upper_bound = Q3 + threshold * IQR

        # Keep only non-outlier points
        mask = (y >= lower_bound) & (y <= upper_bound)
        return X[mask], y[mask]
    else:
        return X, y
```

4.2.2.2 Suavizado (Smoothing) de Datos:

- Suavizado con promedio móvil con tamaño de ventana configurable
- Aplicado antes de la regresión para reducir ruido
- Mantiene la estructura temporal a la vez que reduce la volatilidad.

```
def _apply_data_smoothing(self, y: np.ndarray) -> np.ndarray:
    """Apply data smoothing to the target values."""
    smoothing_method = self.preprocessing_settings.get('smoothing_method', 'moving_average')
    window = self.preprocessing_settings.get('smoothing_window', 3)

    if smoothing_method == 'moving_average' and len(y) >= window:
        # Simple moving average
        smoothed = np.convolve(y, np.ones(window)/window, mode='valid')
        # Pad with original values to maintain length
        padding = (len(y) - len(smoothed)) // 2
        if padding > 0:
            smoothed = np.concatenate([y[:padding], smoothed, y[-padding:]])
        return smoothed
    else:
        return y
```

4.2.3 Marco de Validación Avanzado:

4.2.3.1 Validación Cruzada de Series Temporales:

- Enfoque de ventana expansiva: cada año utiliza todos los datos previos.
- Se requieren mínimo 3 puntos de datos para una validación significativa.
- Análisis separado por horizontes de 3 años y 5 años.


```

for idx, row in df.iterrows():
    category = row['Category']

    if pd.isna(category) or str(category).startswith('📊') or
str(category).startswith('Forecast'):
        continue

    metric_name = str(category).strip()
    if len(metric_name) > 50:
        metric_name = metric_name[:47] + "..."

    historical_data = []
    for col in year_cols_sorted:
        val = row[col]
        if pd.notna(val):
            try:
                num_val = pd.to_numeric(val, errors='coerce')
                if pd.notna(num_val):
                    year = extract_year(col)
                    historical_data.append((year, num_val))
            except:
                continue

    if len(historical_data) < 3:
        continue

    metrics_evaluated += 1

    historical_data.sort(key=lambda x: x[0])

    metric_predictions = {h: [] for h in horizons.keys()}
    metric_actuals = {h: [] for h in horizons.keys()}
    metric_naive_errors = {h: [] for h in horizons.keys()}
    metric_directional_correct = {h: 0 for h in horizons.keys()}
    metric_directional_total = {h: 0 for h in horizons.keys()}

    for i in range(2, len(historical_data)):
        train_data = historical_data[:i]
        actual_year, actual_value = historical_data[i]

        if len(train_data) < 2: # Need at least 2 points for linear regression

```



```

        continue

    last_known_value = train_data[-1][1]

    X_train = np.array([data[0] for data in train_data]).reshape(-1, 1)
    y_train = np.array([data[1] for data in train_data])

    if self.preprocessing_settings.get('apply_data_smoothing', False):
        y_train = self._apply_data_smoothing(y_train)

    if self.preprocessing_settings.get('handle_outliers', False):
        X_train, y_train = self._handle_outliers(X_train, y_train)

    try:
        model = self._create_regression_model()
        model.fit(X_train, y_train)
        prediction = model.predict([[actual_year]])[0]

        prediction = self._validate_and_adjust_forecast(
            prediction, y_train, actual_year, category
        )

        forecast_horizon = i
        for horizon_name, max_years in horizons.items():
            if forecast_horizon <= max_years:
                metric_predictions[horizon_name].append(prediction)
                metric_actuals[horizon_name].append(actual_value)

                naive_forecast = last_known_value
                naive_error = abs(actual_value - naive_forecast)
                metric_naive_errors[horizon_name].append(naive_error)

                forecast_direction = 1 if prediction > last_known_value else -1
                actual_direction = 1 if actual_value > last_known_value else -1
                if forecast_direction == actual_direction:
                    metric_directional_correct[horizon_name] += 1
                    metric_directional_total[horizon_name] += 1

    except Exception as e:
        if self.debug:
            print(f"Debug: Cross-validation failed for {category} year
{actual_year}: {e}")

        continue

```



```

    for horizon_name in horizons.keys():
        if metric_predictions[horizon_name]:
            overall_predictions.extend(metric_predictions[horizon_name])
            overall_actuals.extend(metric_actuals[horizon_name])
            overall_naive_errors.extend(metric_naive_errors[horizon_name])

horizon_data[horizon_name]['predictions'].extend(metric_predictions[horizon_name])

horizon_data[horizon_name]['actuals'].extend(metric_actuals[horizon_name])

horizon_data[horizon_name]['naive_errors'].extend(metric_naive_errors[horizon_name])

        if metric_directional_total[horizon_name] > 0:
            metric_accuracy = metric_directional_correct[horizon_name] /
metric_directional_total[horizon_name]

horizon_data[horizon_name]['directional_accuracies'].append(metric_accuracy)
            overall_directional_accuracies.append(metric_accuracy)

        if metric_name not in performance_results[name]['by_metric']:
            performance_results[name]['by_metric'][metric_name] = {}

        performance_results[name]['by_metric'][metric_name][horizon_name] = {
            'predictions': metric_predictions[horizon_name],
            'actuals': metric_actuals[horizon_name],
            'naive_errors': metric_naive_errors[horizon_name],
            'directional_accuracy': metric_accuracy if
metric_directional_total[horizon_name] > 0 else None
        }

```

4.2.3.2 Validación Preemptiva de Pronósticos:

- Límites máximos de tasa de crecimiento (configurable, predeterminado 50%).
- Aplicación de pronósticos positivos para métricas apropiadas.
- Factores de ajuste conservadores para predicciones poco realistas.

```

def _validate_and_adjust_forecast(self, prediction: float, historical_y: np.ndarray,
                                next_year: int, category: str) -> float:
    """Validate and adjust forecast based on configuration settings."""
    adjusted_prediction = prediction

```



```
# Apply forecast reasonableness validation
if self.validation_settings.get('validate_forecast_reasonableness', True):
    last_value = historical_y[-1]
    max_growth_rate = self.validation_settings.get('max_forecast_growth_rate', 0.5)

    # Only enforce positivity when the metric is non-negative historically.
    if self.general_settings.get('ensure_positive_forecasts', True) and last_value >= 0:
        min_value = self.general_settings.get('minimum_forecast_value', 0.1)
        adjusted_prediction = max(adjusted_prediction, min_value)

    if last_value != 0:
        # Preserve sign by dividing by the previous value (not its absolute).
        # This ensures negative series (losses) keep their direction when we compute
growth.

        predicted_growth = (adjusted_prediction - last_value) / last_value
        if abs(predicted_growth) > max_growth_rate:
            # Apply conservative adjustment
            if self.error_settings.get('use_conservative_adjustment', True):
                conservative_factor =
self.error_settings.get('conservative_adjustment_factor', 0.95)
                # Dampen towards last value
                adjusted_prediction = last_value * (1 + max_growth_rate *
np.sign(predicted_growth) * conservative_factor)

        # Apply fallback adjustment factor if configured
        if self.general_settings.get('fallback_adjustment_factor', 1.0) != 1.0:
            adjusted_prediction *= self.general_settings['fallback_adjustment_factor']

    return adjusted_prediction
```

4.2.4 Variantes de Pronóstico Ajustadas por Puntaje

4.2.4.1 Sensibilidad de Puntaje Específica por Métrica:

- Diferentes factores de ajuste según el tipo de métrica:
 - Métricas de beneficio: 0.1498 (conservador)
 - Capital de trabajo: 0.3370 (agresivo)
 - Otras métricas: 0.2246 (estándar)
- Límite máximo de ajuste del 7% con manejo direccional inteligente
- Preserva el signo para métricas negativas (gastos, pérdidas)


```

        if wants_score_adjust and self.qual_score is not None:
            # Classify metrics by type for appropriate adjustment factors
            is_profit_metric = ('profit' in cat_lower or 'loss' in cat_lower or 'ebit'
in cat_lower) and 'expense' not in cat_lower and 'revenue' not in cat_lower and 'cost' not in
cat_lower

            is_working_capital = 'working capital' in cat_lower

            # Use adjustment factor based on metric type (30% more sensitive)
            # Profit metrics: conservative 0.1498 (less volatile)
            # Working capital: aggressive 0.3370 (operational efficiency impact)
            # Other metrics: standard 0.2246
            if is_profit_metric:
                base_adjustment_factor = 0.1498
            elif is_working_capital:
                base_adjustment_factor = 0.3370
            else:
                base_adjustment_factor = 0.2246

```

- Preserva el signo para métricas negativas (gastos, pérdidas)

4.2.5 Control de Errores:

4.2.5.1 Recuperación tras Errores:

- Primario: Predicción de regresión lineal.
- Secundario: Extrapolación del último valor conocido.
- Terciario: Valor mínimo de pronóstico (0.1).
- Registro integral con opciones de depuración.

```

except Exception as e:
    if self.debug or self.error_settings.get('log_errors', True):
        print(f"Error: Linear regression failed for {category}: {e}")

    # Apply fallback strategy based on config
    if self.error_settings.get('fallback_to_constant_extrapolation', True):
        # Use last known value
        last_value = current_values[-1]
        forecast_df.at[idx, year_col] = last_value
        current_values.append(last_value)
    else:
        # Use minimum value
        min_value = self.general_settings.get('minimum_forecast_value', 0.1)

```



```
forecast_df.at[idx, year_col] = min_value
current_values.append(min_value)
```

4.2.5.2 Gestión Automática de Calidad de Datos:

- Gestiona valores en blanco y diferentes formatos de Excel.
- Múltiples estrategias de detección de encabezados.
- Identificación de columnas de años con algoritmos de respaldo.

```
def _apply_data_smoothing(self, y: np.ndarray) -> np.ndarray:
    """Apply data smoothing to the target values."""
    smoothing_method = self.preprocessing_settings.get('smoothing_method', 'moving_average')
    window = self.preprocessing_settings.get('smoothing_window', 3)

    if smoothing_method == 'moving_average' and len(y) >= window:
        # Simple moving average
        smoothed = np.convolve(y, np.ones(window)/window, mode='valid')
        # Pad with original values to maintain length
        padding = (len(y) - len(smoothed)) // 2
        if padding > 0:
            smoothed = np.concatenate([y[:padding], smoothed, y[-padding:]]
            return smoothed
    else:
        return y

def _handle_outliers(self, X: np.ndarray, y: np.ndarray) -> Tuple[np.ndarray, np.ndarray]:
    """Remove outliers based on configuration settings."""
    method = self.preprocessing_settings.get('outlier_method', 'IQR')
    threshold = self.preprocessing_settings.get('outlier_threshold', 1.5)

    if method == 'IQR' and len(y) > 4: # Need enough data for IQR
        Q1 = np.percentile(y, 25)
        Q3 = np.percentile(y, 75)
        IQR = Q3 - Q1
        lower_bound = Q1 - threshold * IQR
        upper_bound = Q3 + threshold * IQR

        # Keep only non-outlier points
        mask = (y >= lower_bound) & (y <= upper_bound)
        return X[mask], y[mask]
    else:
        return X, y
```



```
def _validate_and_adjust_forecast(self, prediction: float, historical_y: np.ndarray,
                                next_year: int, category: str) -> float:
    """Validate and adjust forecast based on configuration settings."""
    adjusted_prediction = prediction

    # Apply forecast reasonableness validation
    if self.validation_settings.get('validate_forecast_reasonableness', True):
        last_value = historical_y[-1]
        max_growth_rate = self.validation_settings.get('max_forecast_growth_rate', 0.5)

        # Only enforce positivity when the metric is non-negative historically.
        if self.general_settings.get('ensure_positive_forecasts', True) and last_value >= 0:
            min_value = self.general_settings.get('minimum_forecast_value', 0.1)
            adjusted_prediction = max(adjusted_prediction, min_value)

        if last_value != 0:
            # Preserve sign by dividing by the previous value (not its absolute).
            # This ensures negative series (losses) keep their direction when we compute
            growth.

            predicted_growth = (adjusted_prediction - last_value) / last_value
            if abs(predicted_growth) > max_growth_rate:
                # Apply conservative adjustment
                if self.error_settings.get('use_conservative_adjustment', True):
                    conservative_factor =
self.error_settings.get('conservative_adjustment_factor', 0.95)
                    # Dampen towards last value
                    adjusted_prediction = last_value * (1 + max_growth_rate *
np.sign(predicted_growth) * conservative_factor)

            # Apply fallback adjustment factor if configured
            if self.general_settings.get('fallback_adjustment_factor', 1.0) != 1.0:
                adjusted_prediction *= self.general_settings['fallback_adjustment_factor']

    return adjusted_prediction
```

4.2.5.3 Gestión Automática de Calidad de Datos:

- Gestiona valores en blanco y diferentes formatos de Excel.
- Múltiples estrategias de detección de encabezados.
- Identificación de columnas de años con algoritmos de respaldo.

4.2.6 Características de Implementación Técnica:

4.2.6.1 Validación y Ajuste del Modelo:

- Verificaciones de razonabilidad post-predicción.
- Amortiguación de tasa de crecimiento para pronósticos extremos.
- Factores de ajuste de respaldo configurables.

4.2.6.2 Conjunto de Métricas de Evaluación de Rendimiento:

- Evaluación integral: MAE, RMSE, MAPE, sMAPE, MASE, R².
- Medición de precisión direccional.
- Intervalos de confianza del 95% con significancia estadística.

```
# Calculate overall metrics
if len(overall_predictions) == 0 or len(overall_actuals) == 0:
    performance_results[name]['overall'] = {
        'metrics_evaluated': metrics_evaluated,
        'total_predictions': 0,
        'mae': None, 'rmse': None, 'mape': None, 'smape': None,
        'mase': None, 'r_squared': None, 'directional_accuracy': None
    }
else:
    predictions = np.array(overall_predictions)
    actuals = np.array(overall_actuals)

    mae = np.mean(np.abs(predictions - actuals))
    rmse = np.sqrt(np.mean((predictions - actuals) ** 2))

    # MAPE - avoid division by zero
    mask = actuals != 0
    if np.any(mask):
        mape = np.mean(np.abs((predictions[mask] - actuals[mask]) / actuals[mask]))
* 100

    else:
        mape = None

    smape = np.mean(2 * np.abs(predictions - actuals) / (np.abs(predictions) +
np.abs(actuals))) * 100

    # MASE
    if len(overall_naive_errors) > 0:
        mean_naive_error = np.mean(overall_naive_errors)
        if mean_naive_error != 0:
```



```

        mase = mae / mean_naive_error
    else:
        mase = None
else:
    mase = None

# R²
ss_res = np.sum((actuals - predictions) ** 2)
ss_tot = np.sum((actuals - np.mean(actuals)) ** 2)
if ss_tot != 0:
    r_squared = 1 - (ss_res / ss_tot)
else:
    r_squared = None

# Directional accuracy
if overall_directional_accuracies:
    directional_accuracy = np.mean(overall_directional_accuracies) * 100
else:
    directional_accuracy = None

performance_results[name]['overall'] = {
    'metrics_evaluated': metrics_evaluated,
    'total_predictions': len(overall_predictions),
    'mae': mae, 'rmse': rmse, 'mape': mape, 'smape': smape,
    'mase': mase, 'r_squared': r_squared, 'directional_accuracy':
directional_accuracy
}

# Calculate horizon-specific metrics
for horizon_name, data in horizon_data.items():
    if not data['predictions']:
        performance_results[name]['by_horizon'][horizon_name] = None
        continue

    predictions = np.array(data['predictions'])
    actuals = np.array(data['actuals'])

    mae = np.mean(np.abs(predictions - actuals))
    rmse = np.sqrt(np.mean((predictions - actuals) ** 2))

    # MAPE
    mask = actuals != 0
    if np.any(mask):

```



```

        mape = np.mean(np.abs((predictions[mask] - actuals[mask]) / actuals[mask]))
* 100

    else:
        mape = None

    smape = np.mean(2 * np.abs(predictions - actuals) / (np.abs(predictions) +
np.abs(actuals))) * 100

    # MASE
    if data['naive_errors']:
        mean_naive_error = np.mean(data['naive_errors'])
        if mean_naive_error != 0:
            mase = mae / mean_naive_error
        else:
            mase = None
    else:
        mase = None

    # R²
    ss_res = np.sum((actuals - predictions) ** 2)
    ss_tot = np.sum((actuals - np.mean(actuals)) ** 2)
    if ss_tot != 0:
        r_squared = 1 - (ss_res / ss_tot)
    else:
        r_squared = None

    # Directional accuracy
    if data['directional_accuracies']:
        directional_accuracy = np.mean(data['directional_accuracies']) * 100
    else:
        directional_accuracy = None

    performance_results[name]['by_horizon'][horizon_name] = {
        'sample_size': len(set(range(len(df)) - {idx for idx, row in df.iterrows()
        if pd.isna(row['Category']) or
str(row['Category']).startswith('📊')})),
        'total_predictions': len(data['predictions']),
        'mae': mae, 'rmse': rmse, 'mape': mape, 'smape': smape,
        'mase': mase, 'r_squared': r_squared, 'directional_accuracy':
directional_accuracy
    }

```


4.2.6.3 Procesado de Estados:

- Procesamiento simultáneo de estado de resultados, balance general y flujo de efectivo.
- Metodología consistente en todos los estados financieros.
- Formato de salida unificado con hojas separadas para pronósticos base y ajustados por puntaje.

```
# Always use raw linear regression files
self.files = {
    'income_statement': 'income_statement_linear.xls',
    'balance_sheet': 'balance_sheet_linear.xls',
    'cash_flow': 'cash_flow_linear.xls'
}
```

Estos elementos son una muestra del sistema de pronóstico listo para producción de la regresión lineal básica; asegura el desempeño informado, rigor estadístico y validación integral, adecuado para comparar tanto con métodos más simples (como naïve) como con enfoques de ML más complejos.

4.3 Selecciones del Código - ML & Pronóstico: XGBOOST

Extractos del código `XGBoost_forecast.py`, a modo de puntos clave que destacar, organizados por importancia metodológica:

4.3.1 Implementación Principal del Aprendizaje Automático:

4.3.1.1 Modelo de Gradient Boosting (XGBoost):

- Utiliza XGBoost (Extreme Gradient Boosting) para pronóstico de series temporales.
- Hiperparámetros configurables:
 - `n_estimators`, `max_depth`, `learning_rate`, `subsample`, `colsample_bytree`
- Mecanismo de parada temprana para prevenir sobre ajustes.
- Modelo individual por métrica financiera con ingeniería de características personalizada.

```
def _create_xgboost_model(self):
    """Create XGBoost model based on configuration settings."""
    # Get hyperparameters from config
    hyperparams = self.model_settings.get('xgboost_hyperparameters', {})

    # Set up early stopping if enabled
    early_stopping = self.model_settings.get('early_stopping', {})
    callbacks = None
    if early_stopping.get('enabled', False):
        callbacks = [xgb.callback.EarlyStopping(
            rounds=early_stopping.get('rounds', 10),
            metric_name=early_stopping.get('metric', 'rmse'),
            save_best=True
        )]

    # Create model with config parameters
    model = xgb.XGBRegressor(
        **hyperparams,
        callbacks=callbacks
    )

    return model
```

4.3.1.2 Pipeline de Ingeniería de Características:

- Lag Features: Tamaños de ventana configurables (predeterminado 3-5 períodos).

- Rolling Statistics: Promedios móviles, desviaciones estándar, min/max con ventanas configurables.
- Trend Features: Coeficientes de tendencia lineal y cuadrática.
- Difference Features: Diferencias de primer orden y de órdenes superiores.
- Temporal Features: Indicadores de año, trimestre, mes.

```
def _create_lag_features(self, values: List[float], max_lag: int = 3) -> Tuple[np.ndarray,
np.ndarray]:
    """Create lag features for XGBoost training with enhanced feature engineering."""
    if len(values) <= max_lag:
        return None, None

    # Apply data preprocessing first
    processed_values, _ = self._apply_data_preprocessing(np.array(values))

    # Create basic lag features
    x = []
    y = []

    for i in range(max_lag, len(processed_values)):
        # Features: previous max_lag values
        lag_features = processed_values[i-max_lag:i]

        # Enhanced feature engineering based on config
        additional_features = []

        # Rolling statistics
        if self.feature_engineering.get('rolling_statistics', {}).get('enabled', False):
            windows = self.feature_engineering['rolling_statistics'].get('windows', [3])
            stats_funcs = self.feature_engineering['rolling_statistics'].get('statistics',
['mean'])

            for window in windows:
                if i >= window:
                    window_data = processed_values[i-window:i]
                    for stat in stats_funcs:
                        if stat == 'mean':
                            additional_features.append(np.mean(window_data))
                        elif stat == 'std':
                            additional_features.append(np.std(window_data))
                        elif stat == 'min':
                            additional_features.append(np.min(window_data))
                        elif stat == 'max':
```



```

        additional_features.append(np.max(window_data))

# Trend features
if self.feature_engineering.get('trend_features', {}).get('enabled', False):
    trend_degree = self.feature_engineering['trend_features'].get('trend_degree', 1)
    x_trend = np.arange(len(lag_features))
    if trend_degree == 1:
        # Linear trend
        slope = np.polyfit(x_trend, lag_features, 1)[0]
        additional_features.append(slope)
    elif trend_degree == 2:
        # Quadratic trend
        coeffs = np.polyfit(x_trend, lag_features, 2)
        additional_features.extend(coeffs)

# Difference features
if self.feature_engineering.get('difference_features', {}).get('enabled', False):
    orders = self.feature_engineering['difference_features'].get('orders', [1])
    for order in orders:
        if len(lag_features) > order:
            diff = np.diff(lag_features, n=order)
            additional_features.extend(diff[-min(3, len(diff)):]) # Last few
differences

# Combine all features
features = list(lag_features) + additional_features
target = processed_values[i]

X.append(features)
y.append(target)

return np.array(X), np.array(y)

```

4.3.2 Interpretabilidad y Explicabilidad del Modelo:

4.3.2.1 Integración de Análisis SHAP:

- Cálculo automático de valores SHAP para mejorar la interpretabilidad del modelo.
- Genera gráficos de resumen, gráficos de barras y clasificaciones de importancia de características.
- Guarda el análisis SHAP en directorio dedicado con salidas PNG y CSV.
- Proporciona información sobre qué características impulsan las predicciones de pronóstico.


```
def generate_shap_analysis(self, model, X_train: np.ndarray, feature_names: List[str],
                           category: str, sheet_name: str, output_dir: str) -> Dict[str,
Any]:
    """
    Generate SHAP analysis for model interpretability.

    Args:
        model: Trained XGBoost model
        X_train: Training feature matrix
        feature_names: Names of features
        category: Metric category being analyzed
        sheet_name: Sheet name (income_statement, balance_sheet, cash_flow)
        output_dir: Output directory for SHAP plots

    Returns:
        dict: SHAP analysis results and feature importance
    """
    if not SHAP_AVAILABLE:
        print(f"⚠ Skipping SHAP analysis for {category} - SHAP library not available")
        return {"error": "SHAP not available"}

    try:
        print(f"🔍 Generating SHAP analysis for {category}...")

        # Create SHAP explainer
        explainer = shap.TreeExplainer(model)

        # Calculate SHAP values for training data
        shap_values = explainer.shap_values(X_train)

        # Ensure output directory exists
        os.makedirs(output_dir, exist_ok=True)

        # Generate SHAP summary plot
        plt.figure(figsize=(10, 6))
        shap.summary_plot(shap_values, X_train, feature_names=feature_names, show=False)
        summary_plot_path = os.path.join(output_dir,
f"shap_summary_{sheet_name}_{category.replace(' ', '_').replace('/', '_')}.png")
        plt.savefig(summary_plot_path, dpi=150, bbox_inches='tight')
        plt.close()
```



```

print(f"✓ SHAP summary plot saved: {summary_plot_path}")

# Generate SHAP bar plot (mean absolute SHAP values)
plt.figure(figsize=(10, 6))
shap.summary_plot(shap_values, X_train, feature_names=feature_names,
plot_type="bar", show=False)
bar_plot_path = os.path.join(output_dir, f"shap_bar_{sheet_name}_{category.replace(' ', '_').replace('/', '_')}.png")
plt.savefig(bar_plot_path, dpi=150, bbox_inches='tight')
plt.close()
print(f"✓ SHAP bar plot saved: {bar_plot_path}")

# Calculate feature importance from SHAP values
feature_importance = {}
mean_abs_shap = np.abs(shap_values).mean(axis=0)
for i, feature_name in enumerate(feature_names):
    feature_importance[feature_name] = float(mean_abs_shap[i])

# Sort by importance
sorted_importance = dict(sorted(feature_importance.items(), key=lambda x: x[1],
reverse=True))

# Save feature importance to CSV
importance_df = pd.DataFrame({
    'Feature': list(sorted_importance.keys()),
    'Mean_Absolute_SHAP': list(sorted_importance.values())
})
importance_csv_path = os.path.join(output_dir,
f"shap_importance_{sheet_name}_{category.replace(' ', '_').replace('/', '_')}.csv")
importance_df.to_csv(importance_csv_path, index=False)
print(f"✓ SHAP feature importance saved: {importance_csv_path}")

return {
    "feature_importance": sorted_importance,
    "top_features": list(sorted_importance.keys())[:5], # Top 5 most important
features

    "plots": {
        "summary_plot": summary_plot_path,
        "bar_plot": bar_plot_path
    },
    "importance_csv": importance_csv_path,
    "shap_values_shape": shap_values.shape,
    "training_samples": len(X_train)

```



```
}

except Exception as e:
    print(f"⚠ SHAP analysis failed for {category}: {e}")
    return {"error": str(e)}
```

4.3.3 Arquitectura Flexible de Procesamiento de Datos:

4.3.3.1 Soporte Multi-Formato de Archivos:

- Detección automática de formatos CSV, Excel (.xls/.xlsx), JSON y Parquet.
- Coincidencia dinámica de patrones de archivos basada en patrones configurables.
- Detección robusta de encabezados con estrategias de respaldo.
- Normalización de formato de datos desde formatos wide/long a estructura consistente.

```
def detect_and_map_files(self) -> bool:
    """Dynamically detect available XGBoost forecast files based on patterns."""
    import glob

    print("Detecting XGBoost forecast files...")

    for sheet_name, patterns in self.file_patterns.items():
        found_file = None

        # Try each pattern in order of preference (CSV first, then Excel)
        for pattern in patterns:
            search_pattern = os.path.join(self.base_path, pattern)
            matching_files = glob.glob(search_pattern)

            if matching_files:
                # Take the first match (most recent if sorted by name)
                found_file = os.path.basename(sorted(matching_files)[-1]) # Most recent by
name

                print(f"✓ Found {sheet_name} file: {found_file}")
                break

        if found_file:
            self.files[sheet_name] = found_file
        else:
            print(f"Warning: No {sheet_name} file found matching patterns: {patterns}")

    # Check if we found all required files
    required_sheets = ['income_statement', 'balance_sheet', 'cash_flow']
```



```

missing_sheets = [sheet for sheet in required_sheets if sheet not in self.files]

if missing_sheets:
    print(f"Error: Missing required files for sheets: {missing_sheets}")
    return False

return True

def _load_csv_file(self, filepath: str, name: str, filename: str) -> pd.DataFrame:
    """Load CSV file with robust header detection."""
    # Try different header options
    for header_row in [0, 1, 2, None]:
        try:
            temp_df = pd.read_csv(filepath, header=header_row)

            # Check if we have a reasonable structure
            if len(temp_df.columns) >= 2 and len(temp_df) >= 1:
                # Check for year-like columns or data
                has_years = self._detect_year_columns(temp_df)
                if has_years:
                    print(f"✓ Loaded {name} from {filename} (CSV, header={header_row})")
                    return temp_df

        except Exception as e:
            if self.debug:
                print(f"Debug: Failed to load CSV {filename} with header={header_row}: {e}")
            continue

    print(f"Warning: Could not load CSV file {filename} with any header option")
    return None

def _load_excel_file(self, filepath: str, name: str, filename: str) -> pd.DataFrame:
    """Load Excel file with robust header detection."""
    # Try different header options
    for header_row in [0, 1, 2, None]:
        try:
            temp_df = pd.read_excel(filepath, header=header_row)

            # Check if we have a reasonable structure
            if len(temp_df.columns) >= 2 and len(temp_df) >= 1:
                # Check for year-like columns or data
                has_years = self._detect_year_columns(temp_df)
                if has_years:

```



```

        print(f"✓ Loaded {name} from {filename} (Excel, header={header_row})")
        return temp_df

    except Exception as e:
        if self.debug:
            print(f"Debug: Failed to load Excel {filename} with header={header_row}:
{e}")

            continue

    print(f"Warning: Could not load Excel file {filename} with any header option")
    return None

def _detect_year_columns(self, df: pd.DataFrame) -> bool:
    """Detect if dataframe has year-like columns or data."""
    # Check column names for years
    for col in df.columns:
        col_str = str(col).strip().split(' ')[0]
        try:
            year = int(col_str)
            if 1900 <= year <= 2100:
                return True
        except ValueError:
            continue

    # Check first column values for years (CSV format)
    if len(df) > 0:
        first_col_values = df.iloc[:, 0].head(5) # Check first 5 values
        for val in first_col_values:
            val_str = str(val).strip()
            try:
                year = int(val_str)
                if 1900 <= year <= 2100:
                    return True
            except ValueError:
                continue

    return False

def _normalize_data_format(self, df: pd.DataFrame, name: str) -> pd.DataFrame:
    """Normalize different data formats to the expected Excel format (Category column + year
columns)."""
    if df is None:
        return None

```



```
# Check if already in Excel format (has Category column)
if 'Category' in df.columns:
    print(f"✓ {name} already in Excel format")
    return df

# Check if it's in CSV format (year column with metric columns)
first_col = df.columns[0]
if first_col.lower() in ['year', 'date', 'period'] or
str(df[first_col].iloc[0]).isdigit():
    try:
        # Melt the dataframe to convert from wide to long format
        id_vars = [first_col]
        value_vars = [col for col in df.columns if col != first_col]
        df_melted = df.melt(id_vars=id_vars, value_vars=value_vars, var_name='Category',
value_name='temp')

        # Pivot to get years as columns and categories as rows
        df_pivot = df_melted.pivot(index='Category', columns=first_col, values='temp')
        df_pivot.columns = df_pivot.columns.astype(str)
        df_normalized = df_pivot.reset_index()

        print(f"✓ Transformed {name} from wide format to Excel format")
        return df_normalized

    except Exception as e:
        print(f"Warning: Could not transform {name} format: {e}")
        return df

# Unknown format, return as-is with warning
print(f"Warning: {name} in unknown format, using as-is")
return df
```

4.3.3.2 Sistema Dependiente de Configuración:

- Archivo de configuración basado en JSON (`xgboost_forecast_config.json`) preconfigurado.
- Secciones de configuración modulares: modelo, ingeniería de características, preprocesamiento, validación.
- Validación de la configuración en el mismo proceso de ejecución.
- Diseño extensible que permite un ajuste fácil de los parámetros.

```
def load_xgboost_forecast_config(config_file="xgboost_forecast_config.json"):
    """
```



```

Load XGBoost forecast configuration from JSON file

Args:
    config_file: Path to configuration file (default: xgboost_forecast_config.json)

Returns:
    dict: Configuration dictionary
"""
try:
    with open(config_file, 'r') as f:
        config = json.load(f)
    print(f"✓ Loaded XGBoost configuration from {config_file}")
    return config
except FileNotFoundError:
    print(f"Warning: XGBoost configuration file {config_file} not found. Using default settings.")
    return {}
except Exception as e:
    print(f"Warning: Error loading XGBoost configuration: {e}. Using default settings.")
    return {}

class XGBoostForecastProcessor:
    def __init__(self, base_path: str = "analysis_outputs/Forecasts/XGBoost",
                 debug: bool = False, config: Optional[Dict] = None):
        self.base_path = base_path
        self.debug = debug
        # Qualitative score to generate score-adjusted forecast variants (None until set)
        self.qual_score: Optional[float] = None

        # Load configuration
        if config is None:
            self.config = load_xgboost_forecast_config()
        else:
            self.config = config

        # Extract configuration sections
        self.general_settings = self.config.get('general_settings', {})
        self.model_settings = self.config.get('model_settings', {})
        self.feature_engineering = self.config.get('feature_engineering', {})
        self.data_preprocessing_settings = self.config.get('data_preprocessing', {})
        # Backwards-compatible alias used in some report sections
        self.preprocessing_settings = self.data_preprocessing_settings

```



```

self.training_settings = self.config.get('training_settings', {})
self.output_settings = self.config.get('output_settings', {})
self.error_settings = self.config.get('error_handling', {})
self.performance_settings = self.config.get('performance_settings', {})
self.validation_settings = self.config.get('validation_settings', {})
self.advanced_settings = self.config.get('advanced_settings', {})

# Update base path from config if specified
if 'excel_output_directory' in self.output_settings:
    self.base_path = self.output_settings['excel_output_directory']

# Setup logging based on config
if self.error_settings.get('log_errors', True):
    logging.basicConfig(level=logging.INFO, format='%(asctime)s - %(levelname)s -
%(message)s')
    self.logger = logging.getLogger(__name__)
else:
    self.logger = None

# File patterns from config (flexible and configurable)
input_config = self.general_settings.get('input_format_settings', {})
self.preferred_file_types = input_config.get('preferred_file_types', ['csv', 'xlsx',
'xls'])
self.auto_detect_formats = input_config.get('auto_detect_formats', True)

# Build file patterns dynamically from config
pattern_config = input_config.get('file_name_patterns', {
    'income_statement': 'income_statement_xgboost*',
    'balance_sheet': 'balance_sheet_xgboost*',
    'cash_flow': 'cash_flow_xgboost*'
})

self.file_patterns = {}
for sheet_name, base_pattern in pattern_config.items():
    patterns = []
    for ext in self.preferred_file_types:
        patterns.append(f"{base_pattern}.{ext}")
    self.file_patterns[sheet_name] = patterns

# Will be populated during file detection
self.files = {}

# Validate configuration

```



```
self._validate_configuration()

self.dataframes: Dict[str, pd.DataFrame] = {}
self.year_columns: List[str] = []
# Hold score-adjusted rows created during forecasting, keyed by sheet name
self.score_adjusted_rows: Dict[str, List[Dict[str, Any]]] = {}
```

4.3.4 Preprocesamiento de Datos:

4.3.4.1 Gestión de Valores Atípicos:

- Detección de valores atípicos basada en IQR con umbrales configurables
- Eliminación automática de valores extremos antes del entrenamiento del modelo
- Preserva la integridad de los datos mientras reduce el impacto del ruido

```
def _handle_outliers(self, data: np.ndarray) -> Tuple[np.ndarray, int]:
    """Remove outliers based on configuration settings."""
    method = self.data_preprocessing_settings.get('outlier_method', 'iqr')
    threshold = self.data_preprocessing_settings.get('outlier_threshold', 1.5)

    if method == 'iqr' and len(data) > 4:
        Q1 = np.percentile(data, 25)
        Q3 = np.percentile(data, 75)
        IQR = Q3 - Q1
        lower_bound = Q1 - threshold * IQR
        upper_bound = Q3 + threshold * IQR

        # Keep only non-outlier points
        mask = (data >= lower_bound) & (data <= upper_bound)
        outliers_removed = len(data) - np.sum(mask)
        return data[mask], outliers_removed
    else:
        return data, 0
```

4.3.4.2 Transformaciones de Datos:

- Transformaciones logarítmicas para series multiplicativas.
- Eliminación de tendencia usando métodos lineales u otros.
- Pipeline de preprocesamiento configurable.

```
# Apply transformations
if self.data_preprocessing_settings.get('apply_data_transformations', False):
```



```
transform_method = self.data_preprocessing_settings.get('transformation_method',
'log')

if transform_method == 'log':
    # Add small constant to avoid log(0)
    min_val = np.min(processed_data[processed_data > 0])
    if min_val is not None and min_val > 0:
        processed_data = np.log(processed_data + min_val * 0.1)
        preprocessing_info['transformation'] = 'log'
        preprocessing_info['log_offset'] = min_val * 0.1

# Detrend data
if self.data_preprocessing_settings.get('detrend_data', False):
    detrend_method = self.data_preprocessing_settings.get('detrend_method', 'linear')
    if detrend_method == 'linear':
        from scipy import signal
        processed_data = signal.detrend(processed_data)
        preprocessing_info['detrended'] = True

return processed_data, preprocessing_info
```

4.3.5 Marco de Validación Avanzado:

4.3.5.1 Validación Cruzada de Series Temporales:

- Validación de ventana expansiva: cada pronóstico utiliza todos los datos históricos previos.
- Se requieren mínimo 6 puntos de datos para entrenamiento confiable de XGBoost.
- Análisis separado por horizontes de pronóstico de 3 años y 5 años.
- Evaluación realista usando validación de la división de series temporales.

```
def evaluate_performance(self, filtered_data: Dict[str, pd.DataFrame]) -> Dict[str, Dict]:
    """Evaluate forecast performance using cross-validation approach with detailed
breakdowns."""
    print("\nEvaluating XGBoost forecast performance...")

    def extract_year(col):
        try:
            return int(str(col).split(' ')[0])
        except:
            return 0

    performance_results = {}
```



```

for name, df in filtered_data.items():
    print(f"Processing {name} for performance evaluation...")

    # Get year columns (excluding Category)
    year_cols = [col for col in df.columns if col != 'Category']
    year_cols_sorted = sorted(year_cols, key=extract_year)
    years = [extract_year(col) for col in year_cols_sorted]

    if len(years) < 6: # Need at least 6 years for meaningful XGBoost cross-validation
        print(f"Warning: Not enough years in {name} for performance evaluation (need at
least 6, have {len(years)})")
        performance_results[name] = {
            'overall': {},
            'by_horizon': {'3-year': {}, '5-year': {}},
            'by_metric': {}
        }
        continue

    performance_results[name] = {
        'overall': {},
        'by_horizon': {'3-year': {}, '5-year': {}},
        'by_metric': {}
    }

    horizons = {
        '3-year': 3,
        '5-year': 5
    }

    # Calculate forecasts for each row (metric)
    for idx, row in df.iterrows():
        category = row['Category']

        # Skip if category is NaN or looks like metadata
        if pd.isna(category) or str(category).startswith('📊') or
str(category).startswith('Forecast'):
            continue

        # Get historical values for this metric
        historical_data = []
        for col in year_cols_sorted:
            val = row[col]
            if pd.notna(val):

```



```

        try:
            num_val = pd.to_numeric(val, errors='coerce')
            if pd.notna(num_val):
                year = extract_year(col)
                historical_data.append((year, num_val))
        except:
            continue

# Need at least 6 data points for meaningful XGBoost cross-validation
if len(historical_data) < 6:
    continue

# Sort historical data by year
historical_data.sort(key=lambda x: x[0])

# Store individual metric results
metric_name = str(category).strip()
if metric_name not in performance_results[name]['by_metric']:
    performance_results[name]['by_metric'][metric_name] = {}

# Perform cross-validation: forecast each year using only prior years
for horizon_name, max_years in horizons.items():
    metric_predictions = []
    metric_actuals = []
    metric_naive_errors = []
    metric_directional_total = 0
    metric_directional_correct = 0

    # Start from sixth data point for XGBoost (need enough history for feature
engineering)

    for i in range(5, len(historical_data)):
        # Use data up to (but not including) the current year
        train_data = historical_data[:i]
        actual_year, actual_value = historical_data[i]

        # Calculate forecast horizon (years ahead)
        forecast_horizon = actual_year - historical_data[i-1][0]
        if forecast_horizon > max_years:
            continue

        if len(train_data) < 5: # Need at least 5 points for XGBoost feature
engineering
            continue

```



```

# Prepare training data (just the values)
ts_data = np.array([data[1] for data in train_data])

try:
    # Get feature window settings from config
    max_lag = self.general_settings.get('max_feature_window', 5)
    min_lag = self.general_settings.get('min_feature_window', 2)
    feature_window = self.general_settings.get('feature_window', 3)

    # Adjust lag based on available data
    available_lag = min(max_lag, len(ts_data) - 1)
    actual_lag = max(min_lag, min(feature_window, available_lag))

    # Create enhanced lag features using config-driven method
    X_train, y_train = self._create_lag_features(ts_data, actual_lag)

    if X_train is None or len(X_train) < 2: # Need at least some
training data
        continue

    # Create XGBoost model based on config
    model = self._create_xgboost_model()

    # Fit model (disable early stopping for cross-validation)
    model.fit(X_train, y_train, verbose=False)

    # Create features for this prediction using same preprocessing
    processed_train, _ = self._apply_data_preprocessing(ts_data)
    next_features = processed_train[-actual_lag:]

    # Add enhanced features for prediction (same as training)
    prediction_features = list(next_features)

    # Add rolling statistics if enabled
    if self.feature_engineering.get('rolling_statistics',
({}).get('enabled', False)):
        windows =
self.feature_engineering['rolling_statistics'].get('windows', [3])
        stats_funcs =
self.feature_engineering['rolling_statistics'].get('statistics', ['mean'])

        for window in windows:

```



```

        if len(processed_train) >= window:
            window_data = processed_train[-window:]
            for stat in stats_funcs:
                if stat == 'mean':
                    prediction_features.append(np.mean(window_data))
                elif stat == 'std':
                    prediction_features.append(np.std(window_data))
                elif stat == 'min':
                    prediction_features.append(np.min(window_data))
                elif stat == 'max':
                    prediction_features.append(np.max(window_data))

        # Add trend features if enabled
        if self.feature_engineering.get('trend_features', {}).get('enabled',
False):

            trend_degree =
self.feature_engineering['trend_features'].get('trend_degree', 1)
            x_trend = np.arange(len(next_features))
            if trend_degree == 1:
                slope = np.polyfit(x_trend, next_features, 1)[0]
                prediction_features.append(slope)

        # Add difference features if enabled
        if self.feature_engineering.get('difference_features',
{}).get('enabled', False):

            orders =
self.feature_engineering['difference_features'].get('orders', [1])
            for order in orders:
                if len(next_features) > order:
                    diff = np.diff(next_features, n=order)
                    prediction_features.extend(diff[-min(3, len(diff)):])

        next_features_array = np.array(prediction_features).reshape(1, -1)
        forecast_value = model.predict(next_features_array)[0]

        # Apply forecast validation and adjustments
        forecast_value = self._validate_and_adjust_forecast(
            forecast_value, ts_data, actual_year, category
        )

        metric_predictions.append(forecast_value)
        metric_actuals.append(actual_value)

```



```

        # Calculate naive error (for MASE) - difference from previous value
        # Determine last known value from training data (naive forecast)
        try:
            last_known_value = train_data[-1][1]
        except Exception:
            last_known_value = actual_value
        naive_forecast = last_known_value
        naive_error = abs(actual_value - naive_forecast)
        metric_naive_errors.append(naive_error)

        # Calculate directional accuracy
        forecast_direction = 1 if forecast_value > last_known_value else -1
        actual_direction = 1 if actual_value > last_known_value else -1
        if forecast_direction == actual_direction:
            metric_directional_correct += 1
        metric_directional_total += 1

    except Exception as e:
        if self.debug:
            print(f"Debug: XGBoost cross-validation failed for {category}
year {actual_year}: {e}")
        continue

    # Store individual metric results for this horizon
    if horizon_name not in performance_results[name]['by_metric'][metric_name]:
        performance_results[name]['by_metric'][metric_name][horizon_name] = {
            'predictions': [],
            'actuals': [],
            'naive_errors': [],
            'directional_accuracy': None
        }

    performance_results[name]['by_metric'][metric_name][horizon_name]['predictions'].extend(metric_
predictions)

    performance_results[name]['by_metric'][metric_name][horizon_name]['actuals'].extend(metric_actu
als)

    performance_results[name]['by_metric'][metric_name][horizon_name]['naive_errors'].extend(metric
_naive_errors)

    if metric_directional_total > 0:

```



```

        metric_accuracy = metric_directional_correct / metric_directional_total

performance_results[name]['by_metric'][metric_name][horizon_name]['directional_accuracy'] =
metric_accuracy

# Calculate overall and horizon-specific metrics from collected data
def calculate_metrics_from_data(predictions, actuals, naive_errors):
    if not predictions or not actuals:
        return {}

    predictions = np.array(predictions)
    actuals = np.array(actuals)

    mae = np.mean(np.abs(predictions - actuals))
    rmse = np.sqrt(np.mean((predictions - actuals) ** 2))

    # MAPE - avoid division by zero
    mask = actuals != 0
    if np.any(mask):
        mape = np.mean(np.abs((predictions[mask] - actuals[mask]) / actuals[mask]))
* 100
    else:
        mape = 100.0

    # sMAPE
    smape = np.mean(2 * np.abs(predictions - actuals) / (np.abs(predictions) +
np.abs(actuals))) * 100

    # MASE
    if naive_errors:
        mean_naive_error = np.mean(naive_errors)
        mase = mae / mean_naive_error if mean_naive_error != 0 else float('inf')
    else:
        mase = float('inf')

    # R²
    ss_res = np.sum((actuals - predictions) ** 2)
    ss_tot = np.sum((actuals - np.mean(actuals)) ** 2)
    r2 = 1 - (ss_res / ss_tot) if ss_tot != 0 else float('inf')

    return {
        'mae': mae, 'rmse': rmse, 'mape': mape, 'smape': smape,
        'mase': mase, 'r2': r2, 'sample_size': len(predictions)

```



```

    }

    # Calculate overall metrics
    all_predictions = []
    all_actuals = []
    all_naive_errors = []
    all_directional_accuracies = []

    for metric_name, horizons_data in performance_results[name]['by_metric'].items():
        for horizon_name, metric_data in horizons_data.items():
            if metric_data['predictions']:
                all_predictions.extend(metric_data['predictions'])
                all_actuals.extend(metric_data['actuals'])
                all_naive_errors.extend(metric_data['naive_errors'])
                if metric_data['directional_accuracy'] is not None:
                    all_directional_accuracies.append(metric_data['directional_accuracy'])

            if all_predictions:
                performance_results[name]['overall'] = calculate_metrics_from_data(
                    all_predictions, all_actuals, all_naive_errors
                )
                performance_results[name]['overall']['directional_accuracy'] = (
                    np.mean(all_directional_accuracies) * 100 if all_directional_accuracies else
                    None
                )

    # Calculate horizon-specific metrics
    for horizon_name in horizons.keys():
        horizon_predictions = []
        horizon_actuals = []
        horizon_naive_errors = []
        horizon_directional_accuracies = []

        for metric_name, horizons_data in performance_results[name]['by_metric'].items():
            if horizon_name in horizons_data and horizons_data[horizon_name]['predictions']:
                horizon_predictions.extend(horizons_data[horizon_name]['predictions'])
                horizon_actuals.extend(horizons_data[horizon_name]['actuals'])
                horizon_naive_errors.extend(horizons_data[horizon_name]['naive_errors'])
                if horizons_data[horizon_name]['directional_accuracy'] is not None:

```



```

horizon_directional_accuracies.append(horizons_data[horizon_name]['directional_accuracy'])

        if horizon_predictions:
            performance_results[name]['by_horizon'][horizon_name] =
calculate_metrics_from_data(
            horizon_predictions, horizon_actuals, horizon_naive_errors
        )

performance_results[name]['by_horizon'][horizon_name]['directional_accuracy'] = (
        np.mean(horizon_directional_accuracies) * 100 if
horizon_directional_accuracies else None
    )

    print(f"✓ Evaluated {name}: {len(performance_results[name]['by_metric'])} metrics")

    return performance_results

```

4.3.5.2 Validación del Pronóstico y Ajuste::

- Verificaciones de razonabilidad con límites configurables de tasa de crecimiento (máximo predeterminado: 50% anual).
- Aplicación de pronósticos positivos para métricas apropiadas.
- Factores de ajuste conservadores para predicciones poco realistas.

```

def _validate_and_adjust_forecast(self, prediction: float, historical_data: np.ndarray,
                                next_year: int, category: str) -> float:
    """Validate and adjust forecast based on configuration settings."""
    adjusted_prediction = prediction

    # Apply forecast reasonableness validation
    if self.validation_settings.get('validate_forecast_reasonableness', True):
        last_value = historical_data[-1]
        max_growth_rate = self.validation_settings.get('max_forecast_growth_rate', 0.5)

        # Only enforce positivity when the metric is non-negative historically.
        if self.general_settings.get('ensure_positive_forecasts', True) and last_value >= 0:
            min_value = self.general_settings.get('minimum_forecast_value', 0.1)
            adjusted_prediction = max(adjusted_prediction, min_value)

    if last_value != 0 and not np.isnan(last_value):
        # Preserve sign by dividing by the previous value (not its absolute).

```



```

        # This prevents sign-flipping for negative series (losses) when computing
growth.

        predicted_growth = (adjusted_prediction - last_value) / last_value
        if abs(predicted_growth) > max_growth_rate:
            # Apply conservative adjustment
            if self.error_settings.get('use_conservative_adjustment', True):
                conservative_factor =
self.error_settings.get('conservative_adjustment_factor', 0.95)
                # Dampen towards last value
                adjusted_prediction = last_value * (1 + max_growth_rate *
np.sign(predicted_growth) * conservative_factor)

            # Apply fallback adjustment factor if configured
            if self.general_settings.get('fallback_adjustment_factor', 1.0) != 1.0:
                adjusted_prediction *= self.general_settings['fallback_adjustment_factor']

        return adjusted_prediction

```

4.3.6 Variantes de Pronóstico Ajustadas por Puntaje:

4.3.6.1 Integración de Puntaje Cualitativo:

- Puntaje cualitativo (0-10) modula la agresividad del pronóstico.
- Factores de sensibilidad específicos por métrica según categoría financiera.
- Límite máximo de ajuste del 7% con gestión direccional inteligente.

```

        # For XGBoost, we create score-adjusted variants by applying a multiplier to
the base forecasts

        # Calculate adjustment factor based on score and metric type
        score_adj = 0.0
        try:
            score_adj = (float(self.qual_score) - 5.0) / 5.0 # Ranges from -1.0 to
+1.0

            # Use adjustment factor based on metric type (30% more sensitive)
            # Profit metrics: conservative 0.1498 (less volatile)
            # Working capital: aggressive 0.3370 (operational efficiency impact)
            # Other metrics: standard 0.2246
            if is_profit_metric:
                base_adjustment_factor = 0.1498
            elif is_working_capital:
                base_adjustment_factor = 0.3370
            else:

```



```

        base_adjustment_factor = 0.2246
        adjustment_factor = score_adj * base_adjustment_factor

        # Apply 7% maximum adjustment cap
        max_adjustment = 0.07 # 7% maximum adjustment
        adjustment_factor = max(-max_adjustment, min(max_adjustment,
adjustment_factor))

        # Smart adjustment: for negative metrics (expenses), reverse the
direction

        # Higher score should make negative metrics less negative (better)
        if is_negative_metric and adjustment_factor != 0:
            adjustment_factor = -adjustment_factor # Reverse direction

        # Apply adjustment to the trend multiplier
        multiplier = 1.0 + adjustment_factor

        # Apply softer cap (25% instead of 15%) to allow more natural variation
        try:
            from forecast_config import ForecastConfig
            cfg = ForecastConfig()
            growth_cap = cfg.score_adjustments.growth_multiplier
        except Exception:
            # Fallback cap if config not available
            growth_cap = 0.07

        # Clamp multiplier to within allowed range
        min_multiplier = 1.0 - abs(growth_cap)
        max_multiplier = 1.0 + abs(growth_cap)
        multiplier = max(min_multiplier, min(max_multiplier, multiplier))

    except Exception:
        # If anything fails, use no adjustment
        multiplier = 1.0

    # Create adjusted values by applying multiplier to base forecasts
    adjusted_values = []
    for year_col in sorted_forecast_cols:
        try:
            base_value = forecast_df.at[idx, year_col]
            # Handle case where base_value might be a Series (due to duplicate
columns)

            if isinstance(base_value, pd.Series):

```



```

        base_value = base_value.iloc[0] if len(base_value) > 0 else
np.nan

        if pd.notna(base_value):
            adjusted_value = float(base_value) * multiplier
            adjusted_values.append(adjusted_value)
        else:
            adjusted_values.append(np.nan)
    except Exception:
        adjusted_values.append(np.nan)

    # Create new row as a dictionary (not a Series copy to avoid index issues)
    import re
    base_name_only = re.sub(r'\s*\(.*\)\s*', '', str(category)).strip()
    score_adjusted_name = f"{base_name_only} (Score-Adjusted)"

    # Build new row dictionary with all columns
    new_row_dict = {}
    for col in forecast_df.columns:
        if col == 'Category':
            new_row_dict[col] = score_adjusted_name
        elif col in sorted_forecast_cols:
            # Use adjusted value for forecast columns
            col_idx = sorted_forecast_cols.index(col)
            if col_idx < len(adjusted_values):
                new_row_dict[col] = adjusted_values[col_idx]
            else:
                new_row_dict[col] = np.nan
        else:
            # Copy value from original row for historical columns
            try:
                val = row[col]
                new_row_dict[col] = val if not isinstance(val, pd.Series) else
val.iloc[0]

            except:
                new_row_dict[col] = np.nan

    # Check if score-adjusted row already exists, update it or append
    existing_mask = forecast_df['Category'] == score_adjusted_name
    if existing_mask.any():
        # Update existing row
        existing_idx = forecast_df[existing_mask].index[0]
        for col_name, adj_val in zip(sorted_forecast_cols, adjusted_values):
            forecast_df.at[existing_idx, col_name] = adj_val

```



```

else:
    # Append new row as DataFrame
    new_row_df = pd.DataFrame([new_row_dict])
    forecast_df = pd.concat([forecast_df, new_row_df], ignore_index=True)
    # also store the adjusted row separately (preserve original column order)
    try:
        self.score_adjusted_rows[name].append(new_row_dict)
    except Exception:
        # non-fatal
        pass

```

4.3.7 Características de Implementación Técnica:

4.3.7.1 Pronóstico Secuencial:

- Cada año de pronóstico integra las predicciones anteriores de XGBoost.
- Ingeniería de características dinámica para cada paso de predicción.
- Capacidad de reentrenamiento del modelo para incorporar nuevas predicciones.

```

# Classify metrics by type for appropriate adjustment factors
cat_lower = str(category).lower()
is_profit_metric = ('profit' in cat_lower or 'loss' in cat_lower or 'ebit' in
cat_lower) and 'expense' not in cat_lower and 'revenue' not in cat_lower and 'cost' not in
cat_lower

is_working_capital = 'working capital' in cat_lower

# Debug for profit metrics
if is_profit_metric and 'profit' in str(category).lower():
    print(f"[DEBUG] Processing profit metric: {category},
is_profit_metric={is_profit_metric}")

# Forecast each year sequentially, using previously forecasted values
current_values = historical_values.copy()

for year_col in sorted_forecast_cols:
    forecast_year = extract_year_for_sort(year_col)

    if len(current_values) < 4:
        # Not enough data points for forecasting this year
        forecast_df.at[idx, year_col] = current_values[-1] if current_values
else None

        current_values.append(current_values[-1] if current_values else 0)
        continue

```



```

# Create features and train XGBoost model using config-driven approach
try:
    # Get feature window settings from config
    max_lag = self.general_settings.get('max_feature_window', 5)
    min_lag = self.general_settings.get('min_feature_window', 2)
    feature_window = self.general_settings.get('feature_window', 3)

    # Adjust lag based on available data
    available_lag = min(max_lag, len(current_values) - 1)
    actual_lag = max(min_lag, min(feature_window, available_lag))

    # Create enhanced lag features using config-driven method
    X, y = self._create_lag_features(current_values, actual_lag)

    if X is None or len(X) < 1: # Need at least some training data
        forecast_df.at[idx, year_col] = current_values[-1]
        current_values.append(current_values[-1])
        fallback_count += 1
        continue

    # Apply feature selection if configured
    if self.training_settings.get('feature_selection', {}).get('enabled',
False):

        # Implement basic feature importance-based selection
        # For now, just use all features
        pass

    # Split data based on config
    train_ratio = self.general_settings.get('train_ratio', 0.7)
    if len(X) > 3: # Only split if we have enough data
        if self.training_settings.get('cross_validation', {}).get('method',
'standard') == 'time_series_split':
            # Use time series split for more realistic validation
            tscv = TimeSeriesSplit(n_splits=min(3, len(X)-1))
            train_indices, val_indices = list(tscv.split(X))[-1] # Use last
split

            X_train, X_val = X[train_indices], X[val_indices]
            y_train, y_val = y[train_indices], y[val_indices]
        else:
            # Standard random split
            test_size = 1 - train_ratio

```



```

        X_train, X_val, y_train, y_val = train_test_split(X, y,
test_size=test_size, random_state=42)
    else:
        X_train, y_train = X, y
        X_val, y_val = None, None

    # Create and train XGBoost model with config settings
    model = self._create_xgboost_model()

    # Fit model with eval set for early stopping if validation data
available
    eval_set = [(X_val, y_val)] if X_val is not None else None
    model.fit(X_train, y_train, eval_set=eval_set, verbose=False)

    # Generate SHAP analysis for model interpretability
    if self.advanced_settings.get('enable_shap_analysis', True):
        shap_results = self.analyze_shap_for_metric(category, name, model,
X_train, feature_names)

        if shap_results.get("available", True) and "error" not in
shap_results:

            print(f"✓ SHAP analysis completed for {category}")
            if self.debug and "insights" in shap_results:
                for insight in shap_results["insights"]:
                    print(f"    • {insight}")
            elif "error" in shap_results:
                print(f"⚠ SHAP analysis skipped for {category}:
{shap_results['error']}")

        # Validate model if we have validation data and accuracy metrics enabled
        if X_val is not None and len(X_val) > 0 and
self.validation_settings.get('enable_accuracy_metrics', True):
            val_predictions = model.predict(X_val)
            if len(val_predictions) > 0 and len(y_val) > 0:
                val_rmse = np.sqrt(mean_squared_error(y_val, val_predictions))
                val_mae = np.mean(np.abs(y_val - val_predictions))
                val_mape = np.mean(np.abs((y_val - val_predictions) / y_val)) *
100

                if self.debug:
                    print(f"Debug: Validation metrics for {category}:
RMSE={val_rmse:.4f}, MAE={val_mae:.4f}, MAPE={val_mape:.2f}%")

    # Create features for next prediction using same preprocessing

```



```

        processed_current, _ =
self._apply_data_preprocessing(np.array(current_values))
        next_features = processed_current[-actual_lag:]

        # Add enhanced features for prediction (same as training)
        prediction_features = list(next_features)

        # Add rolling statistics if enabled
        if self.feature_engineering.get('rolling_statistics', {}).get('enabled',
False):
            windows =
self.feature_engineering['rolling_statistics'].get('windows', [3])
            stats_funcs =
self.feature_engineering['rolling_statistics'].get('statistics', ['mean'])

            for window in windows:
                if len(processed_current) >= window:
                    window_data = processed_current[-window:]
                    for stat in stats_funcs:
                        if stat == 'mean':
                            prediction_features.append(np.mean(window_data))
                        elif stat == 'std':
                            prediction_features.append(np.std(window_data))
                        elif stat == 'min':
                            prediction_features.append(np.min(window_data))
                        elif stat == 'max':
                            prediction_features.append(np.max(window_data))

            # Add trend features if enabled
            if self.feature_engineering.get('trend_features', {}).get('enabled',
False):
                trend_degree =
self.feature_engineering['trend_features'].get('trend_degree', 1)
                x_trend = np.arange(len(next_features))
                if trend_degree == 1:
                    slope = np.polyfit(x_trend, next_features, 1)[0]
                    prediction_features.append(slope)

            # Add difference features if enabled
            if self.feature_engineering.get('difference_features',
{}).get('enabled', False):
                orders =
self.feature_engineering['difference_features'].get('orders', [1])

```



```

        for order in orders:
            if len(next_features) > order:
                diff = np.diff(next_features, n=order)
                prediction_features.extend(diff[-min(3, len(diff)):])

    next_features_array = np.array(prediction_features).reshape(1, -1)
    forecast_value = model.predict(next_features_array)[0]

    # Apply forecast validation and adjustments
    forecast_value = self._validate_and_adjust_forecast(
        forecast_value, np.array(current_values), forecast_year, category
    )

    forecast_df.at[idx, year_col] = forecast_value
    # Add this forecast to available data for next year's calculation
    current_values.append(forecast_value)
    forecast_count += 1

except Exception as fit_error:
    if self.error_settings.get('log_errors', True):
        if self.logger:
            self.logger.error(f"XGBoost model training failed for
{category}: {fit_error}")

    # Use fallback strategy
    if self.error_settings.get('fallback_to_linear_regression', True):
        # Simple linear regression fallback
        if len(current_values) >= 2:
            X_lr = np.arange(len(current_values)).reshape(-1, 1)
            y_lr = np.array(current_values)
            lr_model = LinearRegression()
            lr_model.fit(X_lr, y_lr)
            next_x = np.array([[len(current_values)]])
            forecast_value = lr_model.predict(next_x)[0]
        else:
            forecast_value = current_values[-1]
    elif self.error_settings.get('fallback_to_linear_extrapolation', True):
        # Simple linear extrapolation
        if len(current_values) >= 2:
            forecast_value = current_values[-1] + (current_values[-1] -
current_values[-2])
        else:
            forecast_value = current_values[-1]
    else:

```



```

forecast_value = current_values[-1]

forecast_value = self._validate_and_adjust_forecast(
    forecast_value, np.array(current_values), forecast_year, category
)
forecast_df.at[idx, year_col] = forecast_value
current_values.append(forecast_value)
fallback_count += 1

```

4.3.7.2 Gestión de Errores:

- Primario: Predicción XGBoost.
- Secundario: Retroceso de regresión lineal.
- Terciario: Extrapolación lineal.
- Registro integral con reporte de errores configurable.

```

except Exception as fit_error:
    if self.error_settings.get('log_errors', True):
        if self.logger:
            self.logger.error(f"XGBoost model training failed for
{category}: {fit_error}")
        # Use fallback strategy
        if self.error_settings.get('fallback_to_linear_regression', True):
            # Simple linear regression fallback
            if len(current_values) >= 2:
                X_lr = np.arange(len(current_values)).reshape(-1, 1)
                y_lr = np.array(current_values)
                lr_model = LinearRegression()
                lr_model.fit(X_lr, y_lr)
                next_x = np.array([[len(current_values)]])
                forecast_value = lr_model.predict(next_x)[0]
            else:
                forecast_value = current_values[-1]
        elif self.error_settings.get('fallback_to_linear_extrapolation', True):
            # Simple linear extrapolation
            if len(current_values) >= 2:
                forecast_value = current_values[-1] + (current_values[-1] -
current_values[-2])
            else:
                forecast_value = current_values[-1]
        else:

```



```

forecast_value = current_values[-1]

forecast_value = self._validate_and_adjust_forecast(
    forecast_value, np.array(current_values), forecast_year, category
)
forecast_df.at[idx, year_col] = forecast_value
current_values.append(forecast_value)
fallback_count += 1

```

4.3.7.3 Conjunto de Métricas de Rendimiento:

- Evaluación integral: MAE, RMSE, MAPE, sMAPE, MASE, R^2 .
- Medición de precisión direccional.
- Intervalos de confianza del 95% con significancia estadística.
- Desgloses de rendimiento por horizonte y por métrica.

```

# Calculate overall and horizon-specific metrics from collected data
def calculate_metrics_from_data(predictions, actuals, naive_errors):
    if not predictions or not actuals:
        return {}

    predictions = np.array(predictions)
    actuals = np.array(actuals)

    mae = np.mean(np.abs(predictions - actuals))
    rmse = np.sqrt(np.mean((predictions - actuals) ** 2))

    # MAPE - avoid division by zero
    mask = actuals != 0
    if np.any(mask):
        mape = np.mean(np.abs((predictions[mask] - actuals[mask]) / actuals[mask]))
* 100
    else:
        mape = 100.0

    # sMAPE
    smape = np.mean(2 * np.abs(predictions - actuals) / (np.abs(predictions) +
np.abs(actuals))) * 100

    # MASE
    if naive_errors:
        mean_naive_error = np.mean(naive_errors)
        mase = mae / mean_naive_error if mean_naive_error != 0 else float('inf')

```



```

        else:
            mase = float('inf')

        # R²
        ss_res = np.sum((actuals - predictions) ** 2)
        ss_tot = np.sum((actuals - np.mean(actuals)) ** 2)
        r2 = 1 - (ss_res / ss_tot) if ss_tot != 0 else float('inf')

        return {
            'mae': mae, 'rmse': rmse, 'mape': mape, 'smape': smape,
            'mase': mase, 'r2': r2, 'sample_size': len(predictions)
        }

# Calculate overall metrics
all_predictions = []
all_actuals = []
all_naive_errors = []
all_directional_accuracies = []

for metric_name, horizons_data in performance_results[name]['by_metric'].items():
    for horizon_name, metric_data in horizons_data.items():
        if metric_data['predictions']:
            all_predictions.extend(metric_data['predictions'])
            all_actuals.extend(metric_data['actuals'])
            all_naive_errors.extend(metric_data['naive_errors'])
            if metric_data['directional_accuracy'] is not None:
                all_directional_accuracies.append(metric_data['directional_accuracy'])

if all_predictions:
    performance_results[name]['overall'] = calculate_metrics_from_data(
        all_predictions, all_actuals, all_naive_errors
    )
    performance_results[name]['overall']['directional_accuracy'] = (
        np.mean(all_directional_accuracies) * 100 if all_directional_accuracies else
        None
    )

# Calculate horizon-specific metrics
for horizon_name in horizons.keys():
    horizon_predictions = []
    horizon_actuals = []
    horizon_naive_errors = []

```



```

        horizon_directional_accuracies = []

        for metric_name, horizons_data in
performance_results[name]['by_metric'].items():
            if horizon_name in horizons_data and
horizons_data[horizon_name]['predictions']:
                horizon_predictions.extend(horizons_data[horizon_name]['predictions'])
                horizon_actuals.extend(horizons_data[horizon_name]['actuals'])
                horizon_naive_errors.extend(horizons_data[horizon_name]['naive_errors'])
                if horizons_data[horizon_name]['directional_accuracy'] is not None:

horizon_directional_accuracies.append(horizons_data[horizon_name]['directional_accuracy'])

            if horizon_predictions:
                performance_results[name]['by_horizon'][horizon_name] =
calculate_metrics_from_data(
                    horizon_predictions, horizon_actuals, horizon_naive_errors
                )

performance_results[name]['by_horizon'][horizon_name]['directional_accuracy'] = (
    np.mean(horizon_directional_accuracies) * 100 if
horizon_directional_accuracies else None
)

    print(f"✓ Evaluated {name}: {len(performance_results[name]['by_metric'])} metrics")

    return performance_results

```

Estos elementos buscan ejemplificar un sistema de pronóstico de machine learning listo para producción y de nivel empresarial, que combina gradient boosting de vanguardia, ingeniería de características avanzada y validación integral. En resumen, adecuado para comparar tanto con métodos estadísticos más simples como con otros enfoques de ML.



4.4 Selecciones del Código - ML & Pronóstico: ARIMA

Sacado del código `ARIMA_forecast.py`, se destacan los siguientes aspectos, según relevancia e importancia metodológica:

4.4.1 Implementación Principal de Modelado de Series Temporales:

4.4.1.1 Múltiples Variantes de ARIMA:

- ARIMA Estándar: Parámetros (p,d,q) configurables usando statsmodels.
- ARIMA Estacional (SARIMA): Agrega componentes estacionales con parámetros (P,D,Q,m).
- Auto ARIMA: Selección automática de modelo usando biblioteca pmdarima con búsqueda en cuadrícula.
- Modelo individual por métrica financiera con optimización personalizada de parámetros.

```
def _create_arima_model(self, data: np.ndarray, preprocessing_info: Dict = None):
    """Create ARIMA model based on configuration settings."""
    # Check if auto ARIMA is enabled and available
    if self.model_settings.get('auto_arima', {}).get('enabled', False) and
PMDARIMA_AVAILABLE:
        return self._create_auto_arima_model(data)

    # Use seasonal ARIMA if enabled
    if self.general_settings.get('enable_seasonal_arima', False):
        return self._create_seasonal_arima_model(data, preprocessing_info)

    # Standard ARIMA model
    arima_order = self.model_settings.get('arima_order', {'p': 1, 'd': 1, 'q': 1})
    order = (arima_order['p'], arima_order['d'], arima_order['q'])

    fit_settings = self.model_settings.get('model_fit_settings', {})
    model = ARIMA(data, order=order, **fit_settings)

    return model

def _create_seasonal_arima_model(self, data: np.ndarray, preprocessing_info: Dict = None):
    """Create Seasonal ARIMA model."""
    arima_order = self.model_settings.get('arima_order', {'p': 1, 'd': 1, 'q': 1})
    seasonal_order = self.model_settings.get('seasonal_arima_order', {'P': 0, 'D': 1, 'Q':
1})

    seasonal_period = self.general_settings.get('seasonal_period', 4)

    order = (arima_order['p'], arima_order['d'], arima_order['q'])
```



```

seasonal_order_full = (seasonal_order['P'], seasonal_order['D'], seasonal_order['Q'],
seasonal_period)

fit_settings = self.model_settings.get('model_fit_settings', {})
model = SARIMAX(data, order=order, seasonal_order=seasonal_order_full, **fit_settings)

return model

def _create_auto_arima_model(self, data: np.ndarray):
    """Create auto ARIMA model using pmdarima."""
    auto_settings = self.model_settings.get('auto_arima', {})
    max_p = auto_settings.get('max_p', 5)
    max_d = auto_settings.get('max_d', 2)
    max_q = auto_settings.get('max_q', 5)
    seasonal = auto_settings.get('seasonal', True)
    m = auto_settings.get('m', 4)

    model = pm.auto_arima(
        data,
        start_p=0, start_q=0,
        max_p=max_p, max_d=max_d, max_q=max_q,
        seasonal=seasonal, m=m,
        error_action='ignore',
        suppress_warnings=True,
        stepwise=True
    )

    return model

```

4.4.1.2 Pruebas de Estacionariedad y Preprocesamiento:

- Augmented Dickey-Fuller Test: Umbral de significancia configurable (predeterminado 0.05)
- Automatic Differencing: Componente integrado (I) maneja series no estacionarias
- Data Transformations: Transformaciones logarítmicas para series multiplicativas
- Detrending: Eliminación lineal de tendencia para remoción de tendencia
- Outlier Handling: Detección y eliminación de valores atípicos basada en IQR

```

def _check_stationarity(self, data: np.ndarray) -> bool:
    """Check if time series is stationary using configured test."""
    if not self.validation_settings.get('check_stationarity', False):
        return True # Assume stationary if not checking

```



```
test_method = self.validation_settings.get('stationarity_test', 'adf')
threshold = self.validation_settings.get('stationarity_threshold', 0.05)

if test_method == 'adf':
    try:
        result = adfuller(data)
        p_value = result[1]
        return p_value < threshold
    except:
        return True # Assume stationary on error

return True
```

```
def _apply_data_preprocessing(self, data: np.ndarray) -> Tuple[np.ndarray, Dict]:
    """Apply data preprocessing based on configuration settings."""
    processed_data = data.copy()
    preprocessing_info = {}

    # Handle outliers
    if self.preprocessing_settings.get('handle_outliers', False):
        processed_data, outlier_info = self._handle_outliers(processed_data)
        preprocessing_info['outliers_removed'] = outlier_info

    # Apply transformations
    if self.preprocessing_settings.get('apply_data_transformations', False):
        transform_method = self.preprocessing_settings.get('transform_method', 'log')
        if transform_method == 'log':
            # Add small constant to avoid log(0)
            min_val = np.min(processed_data[processed_data > 0])
            if min_val is not None and min_val > 0:
                processed_data = np.log(processed_data + min_val * 0.1)
                preprocessing_info['transformation'] = 'log'
                preprocessing_info['log_offset'] = min_val * 0.1

    # Detrend data
    if self.preprocessing_settings.get('detrend_data', False):
        detrend_method = self.preprocessing_settings.get('detrend_method', 'linear')
        if detrend_method == 'linear':
            from scipy import signal
            processed_data = signal.detrend(processed_data)
            preprocessing_info['detrended'] = True
```



```
# Apply differencing
if self.preprocessing_settings.get('apply_differencing', False):
    diff_order = self.preprocessing_settings.get('differencing_order', 1)
    for _ in range(diff_order):
        processed_data = np.diff(processed_data)
        preprocessing_info['differenced'] = diff_order

return processed_data, preprocessing_info
```

4.4.2 Configuración y Automatización Avanzadas:

4.4.2.1 Integración de Auto ARIMA:

- Búsqueda en cuadrícula sobre parámetros p,d,q con límites configurables.
- Detección y modelado de componentes estacionales.
- Selección de modelo basada en criterio de información (AIC/BIC).
- Optimización automatizada de hiperparámetros.

```
def _create_auto_arima_model(self, data: np.ndarray):
    """Create auto ARIMA model using pmdarima."""
    auto_settings = self.model_settings.get('auto_arima', {})
    max_p = auto_settings.get('max_p', 5)
    max_d = auto_settings.get('max_d', 2)
    max_q = auto_settings.get('max_q', 5)
    seasonal = auto_settings.get('seasonal', True)
    m = auto_settings.get('m', 4)

    model = pm.auto_arima(
        data,
        start_p=0, start_q=0,
        max_p=max_p, max_d=max_d, max_q=max_q,
        seasonal=seasonal, m=m,
        error_action='ignore',
        suppress_warnings=True,
        stepwise=True
    )
```

4.4.2.2 Arquitectura Guiada por Configuración:

- Sistema de configuración basado en JSON (`arima_forecast_config.json`).
- Configuraciones modulares: parámetros del modelo, preprocesamiento, validación, manejo de errores.
- Validación de parámetros en tiempo de ejecución con mensajes de error útiles.
- Diseño extensible para diferentes variantes ARIMA y configuraciones.


```
def load_arima_forecast_config(config_file="arima_forecast_config.json"):
    """
    Load ARIMA forecast configuration from JSON file

    Args:
        config_file: Path to configuration file (default: arima_forecast_config.json)

    Returns:
        dict: Configuration dictionary
    """
    try:
        with open(config_file, 'r') as f:
            config = json.load(f)
        print(f"✓ Loaded ARIMA configuration from {config_file}")
        return config
    except FileNotFoundError:
        print(f"Warning: ARIMA configuration file {config_file} not found. Using default settings.")
        return {}
    except Exception as e:
        print(f"Warning: Error loading ARIMA configuration: {e}. Using default settings.")
        return {}

class ARIMAForecastProcessor:
    def __init__(self, base_path: str = "analysis_outputs/Forecasts/ARIMA",
                 debug: bool = False, config: Optional[Dict] = None):
        self.base_path = base_path
        self.debug = debug
        # Qualitative score to generate score-adjusted forecast variants (None until set)
        self.qual_score: Optional[float] = None

        # Load configuration
        if config is None:
            self.config = load_arima_forecast_config()
        else:
            self.config = config

        # Extract configuration sections
        self.general_settings = self.config.get('general_settings', {})
```



```
self.model_settings = self.config.get('model_settings', {})
self.quality_settings = self.config.get('quality_settings', {})
self.output_settings = self.config.get('output_settings', {})
self.error_settings = self.config.get('error_handling', {})
self.performance_settings = self.config.get('performance_settings', {})
self.validation_settings = self.config.get('validation_settings', {})
self.preprocessing_settings = self.config.get('data_preprocessing', {})
self.advanced_settings = self.config.get('advanced_settings', {})

# Update base path from config if specified
if 'excel_output_directory' in self.output_settings:
    self.base_path = self.output_settings['excel_output_directory']

# Setup logging based on config
if self.error_settings.get('log_errors', True):
    logging.basicConfig(level=logging.INFO, format='%(asctime)s - %(levelname)s -
%(message)s')
    self.logger = logging.getLogger(__name__)
else:
    self.logger = None
```

4.4.3 Pipeline de Preprocesamiento de Datos:

4.4.3.1 Preparación de Datos:

- Outlier Removal: Método IQR con umbrales configurables.
- Log Transformations: Maneja estacionalidad multiplicativa y estabilización de varianza.
- Detrending: Elimina tendencias deterministas antes del modelado.
- Manual Differencing: Diferenciación adicional más allá del componente integrado de ARIMA.

```
def _apply_data_preprocessing(self, data: np.ndarray) -> Tuple[np.ndarray, Dict]:
    """Apply data preprocessing based on configuration settings."""
    processed_data = data.copy()
    preprocessing_info = {}

    # Handle outliers
    if self.preprocessing_settings.get('handle_outliers', False):
        processed_data, outlier_info = self._handle_outliers(processed_data)
        preprocessing_info['outliers_removed'] = outlier_info
```



```
# Apply transformations
if self.preprocessing_settings.get('apply_data_transformations', False):
    transform_method = self.preprocessing_settings.get('transform_method', 'log')
    if transform_method == 'log':
        # Add small constant to avoid log(0)
        min_val = np.min(processed_data[processed_data > 0])
        if min_val is not None and min_val > 0:
            processed_data = np.log(processed_data + min_val * 0.1)
            preprocessing_info['transformation'] = 'log'
            preprocessing_info['log_offset'] = min_val * 0.1

# Detrend data
if self.preprocessing_settings.get('detrend_data', False):
    detrend_method = self.preprocessing_settings.get('detrend_method', 'linear')
    if detrend_method == 'linear':
        from scipy import signal
        processed_data = signal.detrend(processed_data)
        preprocessing_info['detrended'] = True

# Apply differencing
if self.preprocessing_settings.get('apply_differencing', False):
    diff_order = self.preprocessing_settings.get('differencing_order', 1)
    for _ in range(diff_order):
        processed_data = np.diff(processed_data)
        preprocessing_info['differenced'] = diff_order

return processed_data, preprocessing_info
```

4.4.3.2 Validación de Estacionariedad:

- Implementación de prueba ADF con niveles de significancia configurables.
- Evaluación automática de estacionariedad antes del ajuste del modelo.
- Sistema de advertencia para series no estacionarias que pueden producir pronósticos poco confiables.

```
def _check_stationarity(self, data: np.ndarray) -> bool:
    """Check if time series is stationary using configured test."""
    if not self.validation_settings.get('check_stationarity', False):
        return True # Assume stationary if not checking

    test_method = self.validation_settings.get('stationarity_test', 'adf')
    threshold = self.validation_settings.get('stationarity_threshold', 0.05)
```



```

if test_method == 'adf':
    try:
        result = adfuller(data)
        p_value = result[1]
        return p_value < threshold
    except:
        return True # Assume stationary on error

return True

```

4.4.4 Entrenamiento del Modelo y Pronóstico:

4.4.4.1 Pronóstico Secuencial Multi-Etapa:

- Cada año de pronóstico se basa en predicciones anteriores de ARIMA
- Reentrenamiento del modelo en cada paso para incorporar nuevos puntos de datos
- Se requieren mínimo 4 puntos de datos para estimación confiable de ARIMA

```

# Forecast each year sequentially, using previously forecasted values
current_values = historical_values.copy()

for year_col in sorted_forecast_cols:
    forecast_year = extract_year_for_sort(year_col)

    if len(current_values) < 4:
        # Not enough data points for forecasting this year
        forecast_df.at[idx, year_col] = current_values[-1] if current_values
else None

        current_values.append(current_values[-1] if current_values else 0)
        continue

    # Fit ARIMA model and forecast using config-driven approach
    try:
        # Convert to numpy array for preprocessing
        ts_data = np.array(current_values)

        # Apply data preprocessing if enabled
        processed_data, preprocessing_info =
self._apply_data_preprocessing(ts_data)

```



```

        # Check stationarity if configured
        if not self._check_stationarity(processed_data):
            if self.logger:
                self.logger.warning(f"Time series for {category} is not
stationary. Results may be unreliable.")

        # Create ARIMA model based on config
        model = self._create_arima_model(processed_data, preprocessing_info)

        # Fit the model with timeout if configured
        timeout = self.performance_settings.get('model_timeout_seconds', 30) if
'performance_settings' in self.config else 30

        try:
            if hasattr(model, 'fit'):
                # For statsmodels models
                model_fit = model.fit()
            else:
                # For pmdarima auto_arima
                model_fit = model

            # Generate forecast
            if hasattr(model_fit, 'forecast'):
                forecast = model_fit.forecast(steps=1)
                if hasattr(forecast, 'iloc'):
                    forecast_value = forecast.iloc[0]
                else:
                    forecast_value = forecast[0]
            elif hasattr(model_fit, 'predict'):
                # For pmdarima models
                forecast_value = model_fit.predict(n_periods=1)[0]
            else:
                raise ValueError("Model does not support forecasting")

            # Apply forecast validation and adjustments based on config
            forecast_value = self._validate_and_adjust_forecast(
                forecast_value, ts_data, forecast_year, category
            )

            forecast_df.at[idx, year_col] = forecast_value
            # Add this forecast to available data for next year's calculation
            current_values.append(forecast_value)
            forecast_count += 1

```



```

except Exception as fit_error:
    if self.error_settings.get('log_errors', True):
        if self.logger:
            self.logger.error(f"Model fitting failed for {category}:
{fit_error}")

    # Use fallback strategy
    if self.error_settings.get('fallback_to_linear_extrapolation',
True):

        # Simple linear extrapolation
        if len(ts_data) >= 2:
            forecast_value = ts_data[-1] + (ts_data[-1] - ts_data[-2])
        else:
            forecast_value = ts_data[-1]
    else:
        forecast_value = ts_data[-1]

    forecast_value = self._validate_and_adjust_forecast(
        forecast_value, ts_data, forecast_year, category
    )
    forecast_df.at[idx, year_col] = forecast_value
    current_values.append(forecast_value)
    fallback_count += 1

except Exception as e:
    if self.debug or self.error_settings.get('log_errors', True):
        print(f"Error: ARIMA failed for {category}: {e}")
    # Apply final fallback strategy
    last_value = current_values[-1] if current_values else 0
    forecast_df.at[idx, year_col] = last_value
    current_values.append(last_value)
    fallback_count += 1

```

4.4.4.2 Gestión de Errores Multi-Nivel:

- Primario: Ajuste y pronóstico del modelo ARIMA.
- Secundario: Retroceso de regresión lineal.
- Terciario: Extrapolación lineal.
- Registro integral con reporte de errores configurable.

```

except Exception as fit_error:
    if self.error_settings.get('log_errors', True):

```



```

        if self.logger:
            self.logger.error(f"Model fitting failed for {category}:
{fit_error}")

        # Use fallback strategy
        if self.error_settings.get('fallback_to_linear_extrapolation',
True):

            # Simple linear extrapolation
            if len(ts_data) >= 2:
                forecast_value = ts_data[-1] + (ts_data[-1] - ts_data[-2])
            else:
                forecast_value = ts_data[-1]
        else:
            forecast_value = ts_data[-1]

        forecast_value = self._validate_and_adjust_forecast(
            forecast_value, ts_data, forecast_year, category
        )
        forecast_df.at[idx, year_col] = forecast_value
        current_values.append(forecast_value)
        fallback_count += 1

    except Exception as e:
        if self.debug or self.error_settings.get('log_errors', True):
            print(f"Error: ARIMA failed for {category}: {e}")
        # Apply final fallback strategy
        last_value = current_values[-1] if current_values else 0
        forecast_df.at[idx, year_col] = last_value
        current_values.append(last_value)
        fallback_count += 1

```

4.4.5 Marco de Validación:

4.4.5.1 Validación Cruzada de Series Temporales:

- Validación de ventana expansiva: cada pronóstico utiliza todos los datos históricos previos.
- Se requieren mínimo 5 puntos de datos para validación cruzada significativa de ARIMA.
- Análisis separado por horizontes de pronóstico de 3 años y 5 años.
- Evaluación realista usando validación de división de series temporales.

```
# Calculate forecasts for each row (metric)
```



```

for idx, row in df.iterrows():
    category = row['Category']

    # Skip if category is NaN or looks like metadata
    if pd.isna(category) or str(category).startswith('📊') or
str(category).startswith('Forecast'):
        continue

    # Try to get a clean metric name
    metric_name = str(category).strip()
    if len(metric_name) > 50:
        metric_name = metric_name[:47] + "..."

    # Get historical values for this metric
    historical_data = []
    for col in year_cols_sorted:
        val = row[col]
        if pd.notna(val):
            try:
                num_val = pd.to_numeric(val, errors='coerce')
                if pd.notna(num_val):
                    year = extract_year(col)
                    historical_data.append((year, num_val))
            except:
                continue

    # Need at least 5 data points for meaningful ARIMA cross-validation
    if len(historical_data) < 5:
        continue

    metrics_evaluated += 1

    # Sort historical data by year
    historical_data.sort(key=lambda x: x[0])

    # Initialize metric-specific collections
    metric_predictions = {h: [] for h in horizons.keys()}
    metric_actuals = {h: [] for h in horizons.keys()}
    metric_naive_errors = {h: [] for h in horizons.keys()}
    metric_directional_correct = {h: 0 for h in horizons.keys()}
    metric_directional_total = {h: 0 for h in horizons.keys()}

    # Perform cross-validation: forecast each year using only prior years

```



```

for i in range(4, len(historical_data)):
    # Use data up to (but not including) the current year
    train_data = historical_data[:i]
    actual_year, actual_value = historical_data[i]

    if len(train_data) < 4: # Need at least 4 points for ARIMA
        continue

    last_known_value = train_data[-1][1]

    # Prepare training data
    ts_data = np.array([data[1] for data in train_data])

    # Apply data preprocessing if enabled
    processed_data, preprocessing_info = self._apply_data_preprocessing(ts_data)

    try:
        # Create ARIMA model based on config
        model = self._create_arima_model(processed_data, preprocessing_info)

        # Fit the model
        if hasattr(model, 'fit'):
            # For statsmodels models
            model_fit = model.fit()
        else:
            # For pmdarima auto_arima
            model_fit = model

        # Generate forecast
        if hasattr(model_fit, 'forecast'):
            forecast = model_fit.forecast(steps=1)
            if hasattr(forecast, 'iloc'):
                forecast_value = forecast.iloc[0]
            else:
                forecast_value = forecast[0]
        elif hasattr(model_fit, 'predict'):
            # For pmdarima models
            forecast_value = model_fit.predict(n_periods=1)[0]
        else:
            raise ValueError("Model does not support forecasting")

        # Apply forecast validation and adjustments
        forecast_value = self._validate_and_adjust_forecast(

```



```

        forecast_value, ts_data, actual_year, category
    )

    # Determine which horizons this prediction belongs to
    forecast_horizon = i # Number of historical points used

    for horizon_name, max_years in horizons.items():
        if forecast_horizon <= max_years:
            metric_predictions[horizon_name].append(forecast_value)
            metric_actuals[horizon_name].append(actual_value)

            # Calculate naive error (for MASE)
            naive_forecast = last_known_value
            naive_error = abs(actual_value - naive_forecast)
            metric_naive_errors[horizon_name].append(naive_error)

            # Calculate directional accuracy
            forecast_direction = 1 if forecast_value > last_known_value else -1

            actual_direction = 1 if actual_value > last_known_value else -1
            if forecast_direction == actual_direction:
                metric_directional_correct[horizon_name] += 1
                metric_directional_total[horizon_name] += 1

        except Exception as e:
            if self.debug:
                print(f"Debug: ARIMA cross-validation failed for {category} year
{actual_year}: {e}")
            continue

    # Add this metric's results to collections
    for horizon_name in horizons.keys():
        if metric_predictions[horizon_name]:
            overall_predictions.extend(metric_predictions[horizon_name])
            overall_actuals.extend(metric_actuals[horizon_name])
            overall_naive_errors.extend(metric_naive_errors[horizon_name])

    horizon_data[horizon_name]['predictions'].extend(metric_predictions[horizon_name])

    horizon_data[horizon_name]['actuals'].extend(metric_actuals[horizon_name])

    horizon_data[horizon_name]['naive_errors'].extend(metric_naive_errors[horizon_name])

```



```

        # Calculate directional accuracy for this metric and horizon
        if metric_directional_total[horizon_name] > 0:
            metric_accuracy = metric_directional_correct[horizon_name] /
metric_directional_total[horizon_name]

horizon_data[horizon_name]['directional_accuracies'].append(metric_accuracy)
            overall_directional_accuracies.append(metric_accuracy)

        # Store individual metric results
        if metric_name not in performance_results[name]['by_metric']:
            performance_results[name]['by_metric'][metric_name] = {}

        performance_results[name]['by_metric'][metric_name][horizon_name] = {
            'predictions': metric_predictions[horizon_name],
            'actuals': metric_actuals[horizon_name],
            'naive_errors': metric_naive_errors[horizon_name],
            'directional_accuracy': metric_accuracy if
metric_directional_total[horizon_name] > 0 else None
        }

    # Calculate overall metrics
    if len(overall_predictions) == 0 or len(overall_actuals) == 0:
        performance_results[name]['overall'] = {
            'metrics_evaluated': metrics_evaluated,
            'total_predictions': 0,
            'mae': None, 'rmse': None, 'mape': None, 'smape': None,
            'mase': None, 'r_squared': None, 'directional_accuracy': None
        }

```

4.4.5.2 Validación de Pronóstico y Ajuste:

- Verificaciones de realismo con límites configurables de tasa de crecimiento (predeterminado 50%).
- Aplicación de pronósticos positivos para métricas apropiadas.
- Factores de ajuste conservadores para predicciones poco realistas.

```

def _validate_and_adjust_forecast(self, prediction: float, historical_data: np.ndarray,
                                next_year: int, category: str) -> float:
    """Validate and adjust forecast based on configuration settings."""
    adjusted_prediction = prediction

    # Apply forecast reasonableness validation

```



```

if self.validation_settings.get('validate_forecast_reasonableness', True):
    last_value = historical_data[-1]
    max_growth_rate = self.validation_settings.get('max_forecast_growth_rate', 0.5)

    # Only enforce positivity when the metric is non-negative historically.
    if self.general_settings.get('ensure_positive_forecasts', True) and last_value >= 0:
        min_value = self.general_settings.get('minimum_forecast_value', 0.1)
        adjusted_prediction = max(adjusted_prediction, min_value)

    if last_value != 0 and not np.isnan(last_value):
        # Preserve sign by dividing by the previous value (not its absolute).
        # This prevents sign-flipping for negative series (losses) when computing
growth.
        predicted_growth = (adjusted_prediction - last_value) / last_value
        if abs(predicted_growth) > max_growth_rate:
            # Apply conservative adjustment
            if self.error_settings.get('use_conservative_adjustment', True):
                conservative_factor =
self.error_settings.get('conservative_adjustment_factor', 0.95)
                # Dampen towards last value
                adjusted_prediction = last_value * (1 + max_growth_rate *
np.sign(predicted_growth) * conservative_factor)

        # Apply fallback adjustment factor if configured
        if self.general_settings.get('fallback_adjustment_factor', 1.0) != 1.0:
            adjusted_prediction *= self.general_settings['fallback_adjustment_factor']

    return adjusted_prediction

```

4.4.6 Variantes de Pronóstico Ajustadas por Puntaje:

4.4.6.1 Integración de Puntaje Cualitativo:

- Puntaje cualitativo (0-10) modula la agresividad del pronóstico.
- Factores de sensibilidad específicos por métrica según categoría financiera.
- Límite máximo de ajuste del 7% con manejo direccional inteligente.
- Preserva el signo para métricas negativas (gastos, pérdidas).

```

# If this category requires a score-adjusted variant, create a new row
# with Category = "<original> (Score-Adjusted)" and fill forecast columns.

```



```

cat_lower = str(category).lower()
wants_score_adjust = any(k in cat_lower for k in score_adjust_keywords)

if wants_score_adjust and self.qual_score is not None:
    # Classify metrics by type for appropriate adjustment factors
    is_profit_metric = ('profit' in cat_lower or 'loss' in cat_lower or 'ebit'
in cat_lower) and 'expense' not in cat_lower and 'revenue' not in cat_lower and 'cost' not in
cat_lower

    is_working_capital = 'working capital' in cat_lower

    # Use adjustment factor based on metric type (30% more sensitive)
    # Profit metrics: conservative 0.1498 (less volatile)
    # Working capital: aggressive 0.3370 (operational efficiency impact)
    # Other metrics: standard 0.2246
    if is_profit_metric:
        base_adjustment_factor = 0.1498
    elif is_working_capital:
        base_adjustment_factor = 0.3370
    else:
        base_adjustment_factor = 0.2246

    # Calculate score-adjusted forecasts by applying adjustment to base ARIMA
predictions
    # This prevents compounding effects and keeps it consistent with
Naive/Linear approaches
    adjusted_values = []

    # Get the base forecast values that were already calculated above
    for year_col in sorted_forecast_cols:
        # Get the base forecast value that was stored in forecast_df
        base_forecast_value = forecast_df.at[idx, year_col]

        # Apply score adjustment to the base forecast value
        score_adj = (float(self.qual_score) - 5.0) / 5.0 # Ranges from -1.0 to
+1.0

        adjustment_factor = score_adj * base_adjustment_factor

        # Apply 25% caps like the Naive and ARIMA systems (consistent approach)
        max_adjustment = 0.07 # 7% maximum adjustment
        adjustment_factor = max(-max_adjustment, min(max_adjustment,
adjustment_factor))

        # Smart adjustment based on value sign

```



```

# For positive values (revenue, profits): higher score = higher values
# For negative values (expenses, losses): higher score = less negative
(better)

    if base_forecast_value >= 0:
        # Positive metric: normal adjustment
        adjusted_forecast = base_forecast_value * (1.0 + adjustment_factor)
    else:
        # Negative metric: reverse adjustment direction for "better"

performance
        adjusted_forecast = base_forecast_value * (1.0 - adjustment_factor)

# Skip forecast validation for score-adjusted forecasts to allow the
score adjustment to take full effect
# Score adjustments are intentional modifications that should not be
constrained by historical growth limits
# adjusted_forecast remains as calculated above without additional
validation

    adjusted_values.append(adjusted_forecast)

# Create new row as a copy of original, then set Category and forecast cells
new_row = row.copy()
# If original category already contains '(Base Forecast)', replace it
# Create a canonical Score-Adjusted category label using the base metric
name

    import re
    base_name_only = re.sub(r'\s*\s*\s*', ' ', str(category)).strip()
    new_row['Category'] = f"{base_name_only} (Score-Adjusted)"
    for col_name, adj_val in zip(sorted_forecast_cols, adjusted_values):
        new_row[col_name] = adj_val

# Check if score-adjusted row already exists, update it or append
score_adjusted_name = f"{base_name_only} (Score-Adjusted)"
existing_mask = forecast_df['Category'] == score_adjusted_name
if existing_mask.any():
    # Update existing row
    existing_idx = forecast_df[existing_mask].index[0]
    for col_name, adj_val in zip(sorted_forecast_cols, adjusted_values):
        forecast_df.at[existing_idx, col_name] = adj_val
else:
    # Append new row
    forecast_df = pd.concat([forecast_df, pd.DataFrame([new_row])],
ignore_index=True)

```



```

        # also store the adjusted row separately (preserve original column order)
        try:
            self.score_adjusted_rows[name].append({k: new_row.get(k, np.nan) for k
in forecast_df.columns})
        except Exception:
            # non-fatal
            pass

```

4.4.7 Implementación Técnica de Características:

4.4.7.1 Gestión de la Complejidad del Modelo:

- Mecanismos de tiempo límite para ajuste del modelo (configurable, predeterminado 30 segundos).
- Monitoreo de convergencia y supresión de advertencias.
- Implementación eficiente en memoria para grandes conjuntos de datos.

```

def _create_arima_model(self, data: np.ndarray, preprocessing_info: Dict = None):
    """Create ARIMA model based on configuration settings."""
    # Check if auto ARIMA is enabled and available
    if self.model_settings.get('auto_arima', {}).get('enabled', False) and
PMDARIMA_AVAILABLE:
        return self._create_auto_arima_model(data)

    # Use seasonal ARIMA if enabled
    if self.general_settings.get('enable_seasonal_arima', False):
        return self._create_seasonal_arima_model(data, preprocessing_info)

    # Standard ARIMA model
    arima_order = self.model_settings.get('arima_order', {'p': 1, 'd': 1, 'q': 1})
    order = (arima_order['p'], arima_order['d'], arima_order['q'])

    fit_settings = self.model_settings.get('model_fit_settings', {})
    model = ARIMA(data, order=order, **fit_settings)

    return model

def _create_seasonal_arima_model(self, data: np.ndarray, preprocessing_info: Dict = None):
    """Create Seasonal ARIMA model."""
    arima_order = self.model_settings.get('arima_order', {'p': 1, 'd': 1, 'q': 1})
    seasonal_order = self.model_settings.get('seasonal_arima_order', {'P': 0, 'D': 1, 'Q':
1})
    seasonal_period = self.general_settings.get('seasonal_period', 4)

```



```

order = (arima_order['p'], arima_order['d'], arima_order['q'])
seasonal_order_full = (seasonal_order['P'], seasonal_order['D'], seasonal_order['Q'],
seasonal_period)

fit_settings = self.model_settings.get('model_fit_settings', {})
model = SARIMAX(data, order=order, seasonal_order=seasonal_order_full, **fit_settings)

return model

def _create_auto_arima_model(self, data: np.ndarray):
    """Create auto ARIMA model using pmdarima."""
    auto_settings = self.model_settings.get('auto_arima', {})
    max_p = auto_settings.get('max_p', 5)
    max_d = auto_settings.get('max_d', 2)
    max_q = auto_settings.get('max_q', 5)
    seasonal = auto_settings.get('seasonal', True)
    m = auto_settings.get('m', 4)

    model = pm.auto_arima(
        data,
        start_p=0, start_q=0,
        max_p=max_p, max_d=max_d, max_q=max_q,
        seasonal=seasonal, m=m,
        error_action='ignore',
        suppress_warnings=True,
        stepwise=True
    )

    return model

```

4.4.7.2 Conjunto de Métricas de Rendimiento:

- Mecanismos de tiempo límite para ajuste del modelo (configurable, predeterminado a 30 segundos).
- Evaluación integral: MAE, RMSE, MAPE, sMAPE, MASE, R^2 .
- Medición de precisión direccional.
- Intervalos de confianza del 95% con significancia estadística.
- Desgloses de rendimiento por horizonte y por métrica.

```
# Calculate overall metrics
```



```

if len(overall_predictions) == 0 or len(overall_actuals) == 0:
    performance_results[name]['overall'] = {
        'metrics_evaluated': metrics_evaluated,
        'total_predictions': 0,
        'mae': None, 'rmse': None, 'mape': None, 'smape': None,
        'mase': None, 'r_squared': None, 'directional_accuracy': None
    }
else:
    predictions = np.array(overall_predictions)
    actuals = np.array(overall_actuals)

    mae = np.mean(np.abs(predictions - actuals))
    rmse = np.sqrt(np.mean((predictions - actuals) ** 2))

    # MAPE - avoid division by zero
    mask = actuals != 0
    if np.any(mask):
        mape = np.mean(np.abs((predictions[mask] - actuals[mask]) / actuals[mask]))
* 100
    else:
        mape = None

    smape = np.mean(2 * np.abs(predictions - actuals) / (np.abs(predictions) +
np.abs(actuals))) * 100

    # MASE
    if len(overall_naive_errors) > 0:
        mean_naive_error = np.mean(overall_naive_errors)
        if mean_naive_error != 0:
            mase = mae / mean_naive_error
        else:
            mase = None
    else:
        mase = None

    # R²
    ss_res = np.sum((actuals - predictions) ** 2)
    ss_tot = np.sum((actuals - np.mean(actuals)) ** 2)
    if ss_tot != 0:
        r_squared = 1 - (ss_res / ss_tot)
    else:
        r_squared = None

```



```

        # Directional accuracy
        if overall_directional_accuracies:
            directional_accuracy = np.mean(overall_directional_accuracies) * 100
        else:
            directional_accuracy = None

        performance_results[name]['overall'] = {
            'metrics_evaluated': metrics_evaluated,
            'total_predictions': len(overall_predictions),
            'mae': mae, 'rmse': rmse, 'mape': mape, 'smape': smape,
            'mase': mase, 'r_squared': r_squared, 'directional_accuracy':
directional_accuracy
        }

    # Calculate horizon-specific metrics
    for horizon_name, data in horizon_data.items():
        if not data['predictions']:
            performance_results[name]['by_horizon'][horizon_name] = None
            continue

        predictions = np.array(data['predictions'])
        actuals = np.array(data['actuals'])

        mae = np.mean(np.abs(predictions - actuals))
        rmse = np.sqrt(np.mean((predictions - actuals) ** 2))

        # MAPE
        mask = actuals != 0
        if np.any(mask):
            mape = np.mean(np.abs((predictions[mask] - actuals[mask]) / actuals[mask]))
* 100
        else:
            mape = None

        smape = np.mean(2 * np.abs(predictions - actuals) / (np.abs(predictions) +
np.abs(actuals))) * 100

        # MASE
        if data['naive_errors']:
            mean_naive_error = np.mean(data['naive_errors'])
            if mean_naive_error != 0:
                mase = mae / mean_naive_error
            else:

```



```

        mase = None
    else:
        mase = None

    # R²
    ss_res = np.sum((actuals - predictions) ** 2)
    ss_tot = np.sum((actuals - np.mean(actuals)) ** 2)
    if ss_tot != 0:
        r_squared = 1 - (ss_res / ss_tot)
    else:
        r_squared = None

    # Directional accuracy
    if data['directional_accuracies']:
        directional_accuracy = np.mean(data['directional_accuracies']) * 100
    else:
        directional_accuracy = None

    performance_results[name]['by_horizon'][horizon_name] = {
        'sample_size': len(set(range(len(df)) - {idx for idx, row in df.iterrows()
                                                    if pd.isna(row['Category']) or
str(row['Category']).startswith('📊')})),
        'total_predictions': len(data['predictions']),
        'mae': mae, 'rmse': rmse, 'mape': mape, 'smape': smape,
        'mase': mase, 'r_squared': r_squared, 'directional_accuracy':
directional_accuracy
    }

    print(f"✓ Evaluated {name}: {metrics_evaluated} metrics, {len(overall_predictions)}
total predictions")

```

Estas selecciones del código muestran los trazos principales del sistema de pronóstico de series temporales del marco de trabajo. Son métodos estadísticos clásicos con automatización moderna, preprocesamiento integral y validación rigurosa y están diseñados para ser de comparación válida tanto con métodos más simples como con otros enfoques de ML más complejos.

4.5 Selecciones del Código - ML & Pronóstico: WINTERS-HOLT

Extraídos del código `Holt_forecast.py`, lo siguiente son fragmentos (presentados según relevancia e importancia metodológica) de la implementación del modelo de aprendizaje automático Winters-Holt en el proceso de trabajo.

4.5.1 Implementación Principal del Suavizado Exponencial:

4.5.1.1 Suavizado Exponencial Triple de Holt-Winters:

- Utiliza *ExponentialSmoothing* de *statsmodels* para componentes de tendencia, estacionalidad y nivel.
- Tipos de modelo configurables: Simple (solo nivel), Doble (nivel + tendencia), Triple (nivel + tendencia + estacional).
- Optimización automática de parámetros probando combinaciones de tendencia/estacional/amortiguada.
- Representación de espacio de estados para estimación robusta de parámetros.

4.5.1.2 Sistema de Auto-Optimización:

- Búsqueda en cuadrícula sobre opciones de tendencia (aditiva/multiplicativa/Ninguna).
- Selección de componente estacional (aditiva/multiplicativa/Ninguna).
- Evaluación de tendencia amortiguada (Verdadero/Falso).
- Selección de modelo basado en MSE para una configuración óptima por métrica.

```
def _create_holt_winters_model(self, data: np.ndarray, seasonal_period: Optional[int] =
None):
    """Create Holt-Winters model based on configuration settings."""
    hw_params = self.model_settings.get('holt_winters_parameters', {})

    # Get model parameters from config
    trend = hw_params.get('trend', 'add')
    seasonal = hw_params.get('seasonal', 'add') if seasonal_period else None
    damped_trend = hw_params.get('damped_trend', False)
    initialization_method = hw_params.get('initialization_method', 'estimated')

    # If auto optimization is enabled, try multiple configurations
    if self.model_settings.get('auto_optimization', {}).get('enabled', True):
        return self._create_auto_optimized_holt_winters(data, seasonal_period)

    # Standard Holt-Winters model
    model = ExponentialSmoothing(
        data,
        trend=trend,
```



```

        seasonal=seasonal,
        seasonal_periods=seasonal_period,
        damped_trend=damped_trend,
        initialization_method=initialization_method
    )

    return model

def _create_auto_optimized_holt_winters(self, data: np.ndarray, seasonal_period:
Optional[int] = None):
    """Create auto-optimized Holt-Winters model by trying multiple configurations."""
    auto_settings = self.model_settings.get('auto_optimization', {})
    trend_options = auto_settings.get('trend_options', ['add', 'mul', None])
    seasonal_options = auto_settings.get('seasonal_options', ['add', 'mul', None]) if
seasonal_period else [None]
    damped_options = auto_settings.get('damped_options', [True, False])

    best_model = None
    best_score = float('inf')
    optimization_method = auto_settings.get('optimization_method', 'mse')

    # Try different model configurations
    for trend in trend_options:
        for seasonal in seasonal_options:
            for damped in damped_options:
                try:
                    model = ExponentialSmoothing(
                        data,
                        trend=trend,
                        seasonal=seasonal,
                        seasonal_periods=seasonal_period,
                        damped_trend=damped,
                        initialization_method='estimated'
                    )

                    fitted_model = model.fit(optimized=True)
                    fitted_values = fitted_model.fittedvalues

                    # Calculate error metric
                    if optimization_method == 'mse':
                        score = np.mean((data[len(data) - len(fitted_values):] -
fitted_values) ** 2)
                    elif optimization_method == 'mae':

```



```

        score = np.mean(np.abs(data[len(data) - len(fitted_values):] -
fitted_values))

    else:
        score = fitted_model.aic # AIC as fallback

    if score < best_score:
        best_score = score
        best_model = model

    except:
        continue

    return best_model if best_model else ExponentialSmoothing(data, trend='add',
seasonal=None)

```

4.5.2 Análisis Estacional Avanzado:

4.5.2.1 Detección Automática de Estacionalidad:

- Identificación de período estacional por autocorrelación.
- Rango del período configurable (mín/máx) con umbral de fuerza.
- Evaluación de la fuerza estacional usando coeficientes de correlación.
- Retroceso al período estacional predeterminado configurado si la detección falla.

```

def _detect_seasonal_period(self, data: np.ndarray) -> Optional[int]:
    """Detect seasonal period using autocorrelation if enabled."""
    if not self.model_settings.get('seasonal_detection', {}).get('enabled', True):
        return self.general_settings.get('seasonal_period', 4)

    min_period = self.model_settings['seasonal_detection'].get('min_seasonal_period', 2)
    max_period = min(self.model_settings['seasonal_detection'].get('max_seasonal_period',
12), len(data) // 2)

    if len(data) < 2 * max_period:
        return self.general_settings.get('seasonal_period', 4)

    # Simple seasonal strength detection using autocorrelation
    best_period = None
    best_strength = 0

    for period in range(min_period, max_period + 1):

```



```

    try:
        # Calculate autocorrelation at this lag
        autocorr = np.corrcoef(data[:-period], data[period:])[0, 1]
        if abs(autocorr) > best_strength:
            best_strength = abs(autocorr)
            best_period = period
    except:
        continue

threshold = self.model_settings['seasonal_detection'].get('seasonal_strength_threshold',
0.3)

if best_strength > threshold:
    return best_period
else:
    return None # No strong seasonality detected

```

4.5.2.2 Descomposición Estacional:

- *seasonal_decompose* de statsmodels para separación de tendencia/estacional/residual.
- Métodos de descomposición configurables (aditivo/multiplicativo).
- Eliminación del componente estacional antes del modelado para mejorar la precisión del pronóstico.

```

# Deseasonalize data
if self.preprocessing_settings.get('deseasonalize_data', False):
    try:
        decomposition_method =
self.preprocessing_settings.get('seasonal_decomposition_method', 'additive')
        seasonal_period = self.general_settings.get('seasonal_period', 4)

        if len(processed_data) >= 2 * seasonal_period:
            decomposition = seasonal_decompose(processed_data,
model=decomposition_method, period=seasonal_period)
            processed_data = processed_data - decomposition.seasonal
            preprocessing_info['deseasonalized'] = True
            preprocessing_info['seasonal_component'] = decomposition.seasonal
    except:
        pass # Skip deseasonalization if it fails

return processed_data, preprocessing_info

```


4.5.3 Pipeline de Preprocesado de Datos:

4.5.3.1 Preparación de Datos Multi-Etapa:

- Outlier Removal: Detección basada en IQR con umbrales configurables.
- Log Transformations: Gestión de estacionalidad multiplicativa y estabilización de varianza.
- Box-Cox Transformation: Técnica avanzada de estabilización de varianza.
- Detrending: Eliminación lineal de tendencia para remover tendencias deterministas.
- Deseasonalization: Eliminación del componente estacional usando descomposición.

4.5.3.2 Estabilización de Varianza:

- Transformación Box-Cox para estabilizar varianza en datos heterocedásticos.
- Estimación automática del parámetro lambda.
- Mejorada confiabilidad del pronóstico para series temporales volátiles.

```
def _apply_data_preprocessing(self, data: np.ndarray) -> Tuple[np.ndarray, Dict]:
    """Apply data preprocessing based on configuration settings."""
    processed_data = data.copy()
    preprocessing_info = {}

    # Handle outliers
    if self.preprocessing_settings.get('handle_outliers', False):
        processed_data, outlier_info = self._handle_outliers(processed_data)
        preprocessing_info['outliers_removed'] = outlier_info

    # Apply transformations
    if self.preprocessing_settings.get('apply_data_transformations', False):
        transform_method = self.preprocessing_settings.get('transform_method', 'log')
        if transform_method == 'log':
            # Add small constant to avoid log(0)
            min_val = np.min(processed_data[processed_data > 0])
            if min_val is not None and min_val > 0:
                processed_data = np.log(processed_data + min_val * 0.1)
                preprocessing_info['transformation'] = 'log'
                preprocessing_info['log_offset'] = min_val * 0.1

    # Stabilize variance
    if self.preprocessing_settings.get('stabilize_variance', False):
        method = self.preprocessing_settings.get('variance_stabilization_method', 'boxcox')
        if method == 'boxcox':
            # Box-Cox transformation for variance stabilization
```



```

        try:
            processed_data, lambda_param = stats.boxcox(processed_data + 1e-10) # Add
small constant
            preprocessing_info['boxcox_lambda'] = lambda_param
        except:
            pass # Skip if Box-Cox fails

# Detrend data
if self.preprocessing_settings.get('detrend_data', False):
    detrend_method = self.preprocessing_settings.get('detrend_method', 'linear')
    if detrend_method == 'linear':
        from scipy import signal
        processed_data = signal.detrend(processed_data)
        preprocessing_info['detrended'] = True

# Deseasonalize data
if self.preprocessing_settings.get('deseasonalize_data', False):
    try:
        decomposition_method =
self.preprocessing_settings.get('seasonal_decomposition_method', 'additive')
        seasonal_period = self.general_settings.get('seasonal_period', 4)

        if len(processed_data) >= 2 * seasonal_period:
            decomposition = seasonal_decompose(processed_data,
model=decomposition_method, period=seasonal_period)
            processed_data = processed_data - decomposition.seasonal
            preprocessing_info['deseasonalized'] = True
            preprocessing_info['seasonal_component'] = decomposition.seasonal
    except:
        pass # Skip deseasonalization if it fails

return processed_data, preprocessing_info

```

4.5.4 Entrenamiento del Modelo y Pronóstico:

4.5.4.1 Pronóstico Secuencial Multi-fase:

- Cada año de pronóstico recupera las predicciones anteriores de Holt-Winters.
- Reentrenamiento del modelo en cada paso con datos históricos actualizados.
- Se requieren mínimo 4 puntos de datos para suavizado exponencial confiable.

4.5.4.2 Gestión de Errores Multi-Nivel:

- Primario: Holt-Winters completo con componentes de tendencia y estacionalidad.
- Secundario: Tendencia lineal de Holt (sin estacionalidad).
- Terciario: Suavizado exponencial simple.
- Cuaternario: Extrapolación lineal.
- Registro integral con informe de errores configurable.

```

# Forecast each year sequentially, using previously forecasted values
current_values = historical_values.copy()

for year_col in sorted_forecast_cols:
    forecast_year = extract_year_for_sort(year_col)

    if len(current_values) < 4:
        # Not enough data points for forecasting this year
        forecast_df.at[idx, year_col] = current_values[-1] if current_values
else None

        current_values.append(current_values[-1] if current_values else 0)
        continue

# Apply Holt-Winters exponential smoothing using config-driven approach
try:
    # Convert to numpy array for preprocessing
    ts_data = np.array(current_values)

    # Apply data preprocessing if enabled
    processed_data, preprocessing_info =
self._apply_data_preprocessing(ts_data)

    # Detect seasonal period if configured
    seasonal_period = None
    if self.general_settings.get('enable_seasonal_adjustment', True):
        seasonal_period = self._detect_seasonal_period(processed_data)

    # Create Holt-Winters model with config settings
    model = self._create_holt_winters_model(processed_data, seasonal_period)

    # Fit the model with timeout if configured
    timeout = self.performance_settings.get('model_timeout_seconds', 30) if
'performance_settings' in self.config else 30

    try:

```



```

        # For auto-optimized models, use optimized=True
        if self.model_settings.get('auto_optimization', {}).get('enabled',
True):

            fitted_model = model.fit(optimized=True)
        else:
            fitted_model = model.fit()

        # Generate forecast
        forecast = fitted_model.forecast(steps=1)
        if hasattr(forecast, 'iloc'):
            forecast_value = forecast.iloc[0]
        else:
            forecast_value = forecast[0] if len(forecast) > 0 else None

        if forecast_value is not None and pd.notna(forecast_value) and
np.isfinite(forecast_value):

            # Validate model if accuracy metrics are enabled
            if self.validation_settings.get('enable_accuracy_metrics', True)
and len(processed_data) > 5:

                fitted_values = fitted_model.fittedvalues
                if len(fitted_values) > 0:

                    # Calculate validation metrics
                    residuals = processed_data[len(processed_data) -
len(fitted_values):] - fitted_values

                    rmse = np.sqrt(np.mean(residuals ** 2))
                    mae = np.mean(np.abs(residuals))
                    mape = np.mean(np.abs(residuals /
processed_data[len(processed_data) - len(fitted_values):])) * 100

                    if self.debug:
                        print(f"Debug: Holt-Winters validation for
{category}: RMSE={rmse:.4f}, MAE={mae:.4f}, MAPE={mape:.2f}%")

                    # Check against thresholds
                    rmse_threshold =
self.validation_settings.get('rmse_threshold_warning', 0.20)
                    mae_threshold =
self.validation_settings.get('mae_threshold_warning', 0.15)
                    mape_threshold =
self.validation_settings.get('mape_threshold_warning', 0.25)

                    if rmse > rmse_threshold or mae > mae_threshold or mape
> mape_threshold:

```



```

        if self.logger:
            self.logger.warning(f"Model performance warning
for {category}: poor fit detected")

        # Apply forecast validation and adjustments
        forecast_value = self._validate_and_adjust_forecast(
            forecast_value, ts_data, forecast_year, category
        )

        forecast_df.at[idx, year_col] = forecast_value
        # Add this forecast to available data for next year's
calculation

        current_values.append(forecast_value)
        forecast_count += 1
    else:
        # Model failed to produce valid forecast
        raise ValueError("Invalid forecast value")

except Exception as fit_error:
    if self.error_settings.get('log_errors', True):
        if self.logger:
            self.logger.error(f"Holt-Winters model fitting failed for
{category}: {fit_error}")

        # Use fallback strategy
        if self.error_settings.get('fallback_to_holt_linear', True):
            # Fallback to Holt's linear trend (no seasonal)
            try:
                simple_model = ExponentialSmoothing(processed_data,
trend='add', seasonal=None)

                simple_fit = simple_model.fit()
                simple_forecast = simple_fit.forecast(steps=1)
                forecast_value = simple_forecast.iloc[0] if
hasattr(simple_forecast, 'iloc') else simple_forecast[0]
            except:
                forecast_value = ts_data[-1] # Last value as final fallback
        elif self.error_settings.get('fallback_to_simple_exponential',
True):

            # Simple exponential smoothing fallback
            try:
                ses_model = ExponentialSmoothing(processed_data, trend=None,
seasonal=None)

                ses_fit = ses_model.fit()
                ses_forecast = ses_fit.forecast(steps=1)

```



```

        forecast_value = ses_forecast.iloc[0] if
hasattr(ses_forecast, 'iloc') else ses_forecast[0]
    except:
        forecast_value = ts_data[-1]
    elif self.error_settings.get('fallback_to_linear_extrapolation',
True):
        # Linear extrapolation fallback
        if len(ts_data) >= 2:
            forecast_value = ts_data[-1] + (ts_data[-1] - ts_data[-2])
        else:
            forecast_value = ts_data[-1]
    else:
        forecast_value = ts_data[-1]

    forecast_value = self._validate_and_adjust_forecast(
        forecast_value, ts_data, forecast_year, category
    )
    forecast_df.at[idx, year_col] = forecast_value
    current_values.append(forecast_value)
    fallback_count += 1

except Exception as e:
    if self.debug or self.error_settings.get('log_errors', True):
        print(f"Error: Holt-Winters failed for {category}: {e}")
    # Apply final fallback strategy
    last_value = current_values[-1] if current_values else 0
    forecast_df.at[idx, year_col] = last_value
    current_values.append(last_value)
    fallback_count += 1

```

4.5.5 Sistema de Validación Avanzado:

4.5.5.1 Validación Cruzada de Series Temporales:

- Validación de ventana expansiva: cada pronóstico utiliza todos los datos históricos previos.
- Se requieren mínimo 6 puntos de datos para validación cruzada significativa.
- Análisis separado por horizontes de pronóstico de 3 años y 5 años.
- Evaluación realista usando la validación de la división de series temporales.

```

# Perform cross-validation: forecast each year using only prior years
for horizon_name, max_years in horizons.items():

```



```

metric_predictions = []
metric_actuals = []
metric_naive_errors = []
metric_directional_total = 0
metric_directional_correct = 0

# Start from sixth data point for Holt-Winters (need enough history for
trend and seasonal components)
for i in range(5, len(historical_data)):
    # Use data up to (but not including) the current year
    train_data = historical_data[:i]
    actual_year, actual_value = historical_data[i]

    # Calculate forecast horizon (years ahead)
    forecast_horizon = actual_year - historical_data[i-1][0]
    if forecast_horizon > max_years:
        continue

    if len(train_data) < 5: # Need at least 5 points for Holt-Winters
        continue

    last_known_value = train_data[-1][1]

    # Prepare training data (just the values)
    ts_data = np.array([data[1] for data in train_data])

    try:
        # Apply data preprocessing if enabled
        processed_data, preprocessing_info =
self._apply_data_preprocessing(ts_data)

        # Detect seasonal period if configured
        seasonal_period = None
        if self.general_settings.get('enable_seasonal_adjustment', True):
            seasonal_period = self._detect_seasonal_period(processed_data)

        # Create Holt-Winters model based on config
        model = self._create_holt_winters_model(processed_data,
seasonal_period)

        # Fit the model
        if self.model_settings.get('auto_optimization', {}).get('enabled',
True):

```



```

        fitted_model = model.fit(optimized=True)
    else:
        fitted_model = model.fit()

    # Generate forecast
    forecast = fitted_model.forecast(steps=1)
    if hasattr(forecast, 'iloc'):
        forecast_value = forecast.iloc[0]
    else:
        forecast_value = forecast[0] if len(forecast) > 0 else None

    if forecast_value is not None and pd.notna(forecast_value) and
np.isfinite(forecast_value):
        # Apply forecast validation and adjustments
        forecast_value = self._validate_and_adjust_forecast(
            forecast_value, ts_data, actual_year, category
        )

        metric_predictions.append(forecast_value)
        metric_actuals.append(actual_value)

        # Calculate naive error (for MASE) - difference from previous
value
        naive_forecast = last_known_value
        naive_error = abs(actual_value - naive_forecast)
        metric_naive_errors.append(naive_error)

        # Calculate directional accuracy
        forecast_direction = 1 if forecast_value > last_known_value else
-1
        actual_direction = 1 if actual_value > last_known_value else -1
        if forecast_direction == actual_direction:
            metric_directional_correct += 1
            metric_directional_total += 1

    except Exception as e:
        if self.debug:
            print(f"Debug: Holt-Winters cross-validation failed for
{category} year {actual_year}: {e}")
            continue

    # Store individual metric results for this horizon

```



```

        if horizon_name not in
performance_results[name]['by_metric'][metric_name]:
            performance_results[name]['by_metric'][metric_name][horizon_name] =
{
                'predictions': [],
                'actuals': [],
                'naive_errors': [],
                'directional_accuracy': None
            }

performance_results[name]['by_metric'][metric_name][horizon_name]['predictions'].extend(metric_
predictions)

performance_results[name]['by_metric'][metric_name][horizon_name]['actuals'].extend(metric_actu
als)

performance_results[name]['by_metric'][metric_name][horizon_name]['naive_errors'].extend(metric
_naive_errors)

        if metric_directional_total > 0:
            metric_accuracy = metric_directional_correct /
metric_directional_total

performance_results[name]['by_metric'][metric_name][horizon_name]['directional_accuracy'] =
metric_accuracy

# Calculate overall and horizon-specific metrics from collected data
def calculate_metrics_from_data(predictions, actuals, naive_errors):
    if not predictions or not actuals:
        return {}

    predictions = np.array(predictions)
    actuals = np.array(actuals)

    mae = np.mean(np.abs(predictions - actuals))
    rmse = np.sqrt(np.mean((predictions - actuals) ** 2))

    # MAPE - avoid division by zero
    mask = actuals != 0
    if np.any(mask):
        mape = np.mean(np.abs((predictions[mask] - actuals[mask]) / actuals[mask]))

* 100

```



```

        else:
            mape = 100.0

        # SMAPE
        smape = np.mean(2 * np.abs(predictions - actuals) / (np.abs(predictions) +
np.abs(actuals))) * 100

        # MASE
        if naive_errors:
            mean_naive_error = np.mean(naive_errors)
            mase = mae / mean_naive_error if mean_naive_error != 0 else float('inf')
        else:
            mase = float('inf')

        # R²
        ss_res = np.sum((actuals - predictions) ** 2)
        ss_tot = np.sum((actuals - np.mean(actuals)) ** 2)
        r2 = 1 - (ss_res / ss_tot) if ss_tot != 0 else float('inf')

        return {
            'mae': mae, 'rmse': rmse, 'mape': mape, 'smape': smape,
            'mase': mase, 'r2': r2, 'sample_size': len(predictions)
        }

    # Calculate overall metrics
    all_predictions = []
    all_actuals = []
    all_naive_errors = []
    all_directional_accuracies = []

    for metric_name, horizons_data in performance_results[name]['by_metric'].items():
        for horizon_name, metric_data in horizons_data.items():
            if metric_data['predictions']:
                all_predictions.extend(metric_data['predictions'])
                all_actuals.extend(metric_data['actuals'])
                all_naive_errors.extend(metric_data['naive_errors'])
                if metric_data['directional_accuracy'] is not None:
                    all_directional_accuracies.append(metric_data['directional_accuracy'])

    if all_predictions:
        performance_results[name]['overall'] = calculate_metrics_from_data(
            all_predictions, all_actuals, all_naive_errors

```



```

    )
    performance_results[name]['overall']['directional_accuracy'] = (
        np.mean(all_directional_accuracies) * 100 if all_directional_accuracies else
None
    )

    # Calculate horizon-specific metrics
    for horizon_name in horizons.keys():
        horizon_predictions = []
        horizon_actuals = []
        horizon_naive_errors = []
        horizon_directional_accuracies = []

        for metric_name, horizons_data in
performance_results[name]['by_metric'].items():
            if horizon_name in horizons_data and
horizons_data[horizon_name]['predictions']:
                horizon_predictions.extend(horizons_data[horizon_name]['predictions'])
                horizon_actuals.extend(horizons_data[horizon_name]['actuals'])
                horizon_naive_errors.extend(horizons_data[horizon_name]['naive_errors'])
                if horizons_data[horizon_name]['directional_accuracy'] is not None:
horizon_directional_accuracies.append(horizons_data[horizon_name]['directional_accuracy'])

            if horizon_predictions:
                performance_results[name]['by_horizon'][horizon_name] =
calculate_metrics_from_data(
                    horizon_predictions, horizon_actuals, horizon_naive_errors
                )

        performance_results[name]['by_horizon'][horizon_name]['directional_accuracy'] = (
            np.mean(horizon_directional_accuracies) * 100 if
horizon_directional_accuracies else None
        )

    print(f"✓ Evaluated {name}: {len(performance_results[name]['by_metric'])} metrics")

    return performance_results

```

4.5.5.2 Validación del Rendimiento del Modelo:

- Validación dentro de la muestra con métricas de precisión durante el entrenamiento.

- Umbrales RMSE/MAE/MAPE para la evaluación de calidad del modelo.
- Sistema automático de advertencia para modelos con rendimiento deficiente.
- Umbrales de rendimiento configurables.

```

        if forecast_value is not None and pd.notna(forecast_value) and
np.isfinite(forecast_value):
            # Validate model if accuracy metrics are enabled
            if self.validation_settings.get('enable_accuracy_metrics', True)
and len(processed_data) > 5:
                fitted_values = fitted_model.fittedvalues
                if len(fitted_values) > 0:
                    # Calculate validation metrics
                    residuals = processed_data[len(processed_data) -
len(fitted_values):] - fitted_values
                    rmse = np.sqrt(np.mean(residuals ** 2))
                    mae = np.mean(np.abs(residuals))
                    mape = np.mean(np.abs(residuals /
processed_data[len(processed_data) - len(fitted_values):])) * 100

                    if self.debug:
                        print(f"Debug: Holt-Winters validation for
{category}: RMSE={rmse:.4f}, MAE={mae:.4f}, MAPE={mape:.2f}%")

                    # Check against thresholds
                    rmse_threshold =
self.validation_settings.get('rmse_threshold_warning', 0.20)
                    mae_threshold =
self.validation_settings.get('mae_threshold_warning', 0.15)
                    mape_threshold =
self.validation_settings.get('mape_threshold_warning', 0.25)

                    if rmse > rmse_threshold or mae > mae_threshold or mape
> mape_threshold:
                        if self.logger:
                            self.logger.warning(f"Model performance warning
for {category}: poor fit detected")

```

4.5.6 Variantes de Pronóstico Ajustadas por Puntuación:

4.5.6.1 Integración de Puntaje Cualitativo:

- Puntaje cualitativo (0-10) modula la agresividad del pronóstico.

- Factores de sensibilidad específicos por métrica según categoría financiera.
- Límite máximo de ajuste del 7% con guía direccional inteligente.
- Ajusta y preserva el signo para métricas negativas (gastos, pérdidas).

```
# If this category requires a score-adjusted variant, create a new row
# with Category = "<original> (Score-Adjusted)" and fill forecast columns.
cat_lower = str(category).lower()
wants_score_adjust = any(k in cat_lower for k in score_adjust_keywords)

if wants_score_adjust and self.qual_score is not None:
    if 'profit' in str(category).lower():
        print(f"[DEBUG] Creating score-adjusted for profit category:
{category}")

        # Diagnostic debug prints to understand why adjusted series may equal base
        try:
            print(f"[DEBUG] Creating Score-Adjusted row for category: {category}")
            print(f"    qual_score={self.qual_score}")
        except Exception:
            pass

        # For Holt-Winters, we create score-adjusted variants by applying a
multiplier to the base forecasts

        # Calculate adjustment factor based on score and metric type
        score_adj = 0.0
        try:
            score_adj = (float(self.qual_score) - 5.0) / 5.0 # Ranges from -1.0 to
+1.0

            # Use adjustment factor based on metric type (30% more sensitive)
            # Profit metrics: conservative 0.1498 (less volatile)
            # Working capital: aggressive 0.3370 (operational efficiency impact)
            # Other metrics: standard 0.2246
            if is_profit_metric:
                base_adjustment_factor = 0.1498
            elif is_working_capital:
                base_adjustment_factor = 0.3370
            else:
                base_adjustment_factor = 0.2246
            adjustment_factor = score_adj * base_adjustment_factor

            # Apply 7% maximum adjustment cap
            max_adjustment = 0.07 # 7% maximum adjustment
```



```

        adjustment_factor = max(-max_adjustment, min(max_adjustment,
adjustment_factor))

        # Smart adjustment: for negative metrics (expenses), reverse the
direction

        # Higher score should make negative metrics less negative (better)
        if is_negative_metric and adjustment_factor != 0:
            adjustment_factor = -adjustment_factor # Reverse direction

        # Apply adjustment to the trend multiplier
        multiplier = 1.0 + adjustment_factor

        # Apply softer cap (25% instead of 15%) to allow more natural variation
        try:
            from forecast_config import ForecastConfig
            cfg = ForecastConfig()
            growth_cap = cfg.score_adjustments.growth_multiplier
        except Exception:
            # Fallback cap if config not available
            growth_cap = 0.07

        # Clamp multiplier to within allowed range
        min_multiplier = 1.0 - abs(growth_cap)
        max_multiplier = 1.0 + abs(growth_cap)
        multiplier = max(min_multiplier, min(max_multiplier, multiplier))

    except Exception:
        # If anything fails, use no adjustment
        multiplier = 1.0

    # Create adjusted values by applying multiplier to base forecasts
    adjusted_values = []
    for year_col in sorted_forecast_cols:
        base_value = forecast_df.at[idx, year_col]
        if pd.isna(base_value):
            adjusted_value = base_value * multiplier
            adjusted_values.append(adjusted_value)
        else:
            adjusted_values.append(base_value)

    # Create new row as a copy of original, then set Category and forecast cells
    new_row = row.copy()
    # If original category already contains '(Base Forecast)', replace it

```



```

name
# Create a canonical Score-Adjusted category label using the base metric

import re
base_name_only = re.sub(r'\s*\s*\s*', '', str(category)).strip()
new_row['Category'] = f'{base_name_only} (Score-Adjusted)'
for col_name, adj_val in zip(sorted_forecast_cols, adjusted_values):
    new_row[col_name] = adj_val

# Check if score-adjusted row already exists, update it or append
score_adjusted_name = f'{base_name_only} (Score-Adjusted)'
existing_mask = forecast_df['Category'] == score_adjusted_name
if existing_mask.any():
    # Update existing row
    existing_idx = forecast_df[existing_mask].index[0]
    for col_name, adj_val in zip(sorted_forecast_cols, adjusted_values):
        forecast_df.at[existing_idx, col_name] = adj_val
else:
    # Append new row
    forecast_df = pd.concat([forecast_df, pd.DataFrame([new_row])],
ignore_index=True)

# also store the adjusted row separately (preserve original column order)
try:
    self.score_adjusted_rows[name].append({k: new_row.get(k, np.nan) for k
in forecast_df.columns})
except Exception:
    # non-fatal
    pass

# Print sample comparison of base vs adjusted forecasts for this category
try:
    # Locate base row exactly and the new adjusted row precisely.
    base_row = forecast_df[forecast_df['Category'] == category]
    # Prefer an exact "(Score-Adjusted)" match for this base metric first
    base_name = str(category).split('(')[0].strip()
    exact_adj_name = f'{base_name} (Score-Adjusted)'
    adj_row = forecast_df[forecast_df['Category'].astype(str) ==
exact_adj_name]

    if adj_row.empty:
        # Fallback: prefer rows that start with the base name and include
score-adjust indicator

        mask_score =
forecast_df['Category'].astype(str).str.contains('score-adjust', case=False, regex=False)
        mask_start =
forecast_df['Category'].astype(str).str.lower().str.startswith(base_name.lower())

```



```

adj_candidates = forecast_df.loc[mask_score & mask_start]
if adj_candidates.empty:
    # final fallback: any row that contains both base name and
score-adjusted literal

adj_candidates =
forecast_df.loc[forecast_df['Category'].astype(str).str.contains(base_name, case=False,
regex=False) & forecast_df['Category'].astype(str).str.contains('score-adjust', case=False,
regex=False)]

if not adj_candidates.empty:
    adj_row = adj_candidates.head(1)

if not base_row.empty and not adj_row.empty:
    try:
        base_last = float(base_row.iloc[0][sorted_forecast_cols[-1]])
        adj_last = float(adj_row.iloc[0][sorted_forecast_cols[-1]])
        diff = adj_last - base_last
        pct = (diff / base_last * 100) if base_last not in (0, None,
np.nan) else np.nan

        print(f"[DEBUG] {category} - last forecast after append:
base={base_last:.2f}, score_adj={adj_last:.2f}, diff={diff:.2f} ({pct:.2f}%)")
    except Exception:
        pass
except Exception:
    # Non-fatal; continue
    pass

forecasted_data[name] = forecast_df
print(f"✓ Updated forecast columns in {name}: {forecast_year_cols}")
print(f" Holt-Winters forecasts: {forecast_count}, Fallbacks: {fallback_count}")

```

4.5.7 Implementación Técnica de Prestaciones:

4.5.7.1 Sistema de Configuración:

- Sistema de configuración por JSON (`holt_forecast_config.json`).
- Configuraciones modulares: parámetros del modelo, preprocesamiento, validación, detección estacional.
- Validación de parámetros en tiempo de ejecución con mensajes de error.
- Diseño extensible para diferentes variantes del suavizado exponencial.

```
class HoltForecastProcessor:
```



```
def __init__(self, base_path: str = "analysis_outputs/Forecasts/Holt_Winters",
             debug: bool = False, config: Optional[Dict] = None):
    self.base_path = base_path
    self.debug = debug
    # Qualitative score to generate score-adjusted forecast variants (None until set)
    self.qual_score: Optional[float] = None

    # Load configuration
    if config is None:
        self.config = load_holt_forecast_config()
    else:
        self.config = config

    # Extract configuration sections
    self.general_settings = self.config.get('general_settings', {})
    self.model_settings = self.config.get('model_settings', {})
    self.quality_settings = self.config.get('quality_settings', {})
    self.output_settings = self.config.get('output_settings', {})
    self.error_settings = self.config.get('error_handling', {})
    self.performance_settings = self.config.get('performance_settings', {})
    self.validation_settings = self.config.get('validation_settings', {})
    self.preprocessing_settings = self.config.get('data_preprocessing', {})
    self.advanced_settings = self.config.get('advanced_settings', {})

    # Update base path from config if specified
    if 'excel_output_directory' in self.output_settings:
        self.base_path = self.output_settings['excel_output_directory']

    # Setup logging based on config
    if self.error_settings.get('log_errors', True):
        logging.basicConfig(level=logging.INFO, format='%(asctime)s - %(levelname)s - %(message)s')
        self.logger = logging.getLogger(__name__)
    else:
        self.logger = None

    # Always use raw Holt-Winters files
    self.files = {
        'income_statement': 'income_statement_holt_winter.xls',
        'balance_sheet': 'balance_sheet_holt_winter.xls',
        'cash_flow': 'cash_flow_holt_winter.xls'
    }
```



```
self.dataframes: Dict[str, pd.DataFrame] = {}
self.year_columns: List[str] = []
# Hold score-adjusted rows created during forecasting, keyed by sheet name
self.score_adjusted_rows: Dict[str, List[Dict[str, Any]]] = {}
```

4.5.7.2 Conjunto de Métricas de Rendimiento:

- Sistema de evaluación integral: MAE, RMSE, MAPE, sMAPE, MASE, R^2 .
- Medición de precisión direccional.
- Intervalos de confianza del 95% con significancia estadística.
- Desgloses de rendimiento por horizonte y por métrica.

```
# Calculate overall and horizon-specific metrics from collected data
def calculate_metrics_from_data(predictions, actuals, naive_errors):
    if not predictions or not actuals:
        return {}

    predictions = np.array(predictions)
    actuals = np.array(actuals)

    mae = np.mean(np.abs(predictions - actuals))
    rmse = np.sqrt(np.mean((predictions - actuals) ** 2))

    # MAPE - avoid division by zero
    mask = actuals != 0
    if np.any(mask):
        mape = np.mean(np.abs((predictions[mask] - actuals[mask]) / actuals[mask]))
* 100
    else:
        mape = 100.0

    # sMAPE
    smape = np.mean(2 * np.abs(predictions - actuals) / (np.abs(predictions) +
np.abs(actuals))) * 100

    # MASE
    if naive_errors:
        mean_naive_error = np.mean(naive_errors)
        mase = mae / mean_naive_error if mean_naive_error != 0 else float('inf')
    else:
        mase = float('inf')
```



```

# R2
ss_res = np.sum((actuals - predictions) ** 2)
ss_tot = np.sum((actuals - np.mean(actuals)) ** 2)
r2 = 1 - (ss_res / ss_tot) if ss_tot != 0 else float('inf')

return {
    'mae': mae, 'rmse': rmse, 'mape': mape, 'smape': smape,
    'mase': mase, 'r2': r2, 'sample_size': len(predictions)
}

# Calculate overall metrics
all_predictions = []
all_actuals = []
all_naive_errors = []
all_directional_accuracies = []

for metric_name, horizons_data in performance_results[name]['by_metric'].items():
    for horizon_name, metric_data in horizons_data.items():
        if metric_data['predictions']:
            all_predictions.extend(metric_data['predictions'])
            all_actuals.extend(metric_data['actuals'])
            all_naive_errors.extend(metric_data['naive_errors'])
            if metric_data['directional_accuracy'] is not None:
all_directional_accuracies.append(metric_data['directional_accuracy'])

if all_predictions:
    performance_results[name]['overall'] = calculate_metrics_from_data(
        all_predictions, all_actuals, all_naive_errors
    )
    performance_results[name]['overall']['directional_accuracy'] = (
        np.mean(all_directional_accuracies) * 100 if all_directional_accuracies else
None
    )

# Calculate horizon-specific metrics
for horizon_name in horizons.keys():
    horizon_predictions = []
    horizon_actuals = []
    horizon_naive_errors = []
    horizon_directional_accuracies = []

```



```

        for metric_name, horizons_data in
performance_results[name]['by_metric'].items():
            if horizon_name in horizons_data and
horizons_data[horizon_name]['predictions']:
                horizon_predictions.extend(horizons_data[horizon_name]['predictions'])
                horizon_actuals.extend(horizons_data[horizon_name]['actuals'])
                horizon_naive_errors.extend(horizons_data[horizon_name]['naive_errors'])
                if horizons_data[horizon_name]['directional_accuracy'] is not None:

horizon_directional_accuracies.append(horizons_data[horizon_name]['directional_accuracy'])

            if horizon_predictions:
                performance_results[name]['by_horizon'][horizon_name] =
calculate_metrics_from_data(
                    horizon_predictions, horizon_actuals, horizon_naive_errors
                )

performance_results[name]['by_horizon'][horizon_name]['directional_accuracy'] = (
    np.mean(horizon_directional_accuracies) * 100 if
horizon_directional_accuracies else None
)

print(f"✓ Evaluated {name}: {len(performance_results[name]['by_metric'])} metrics")

return performance_results

```

Estos extractos pretenden demostrar el sistema de suavizado exponencial Winters-Holt en activo, que combina métodos estadísticos clásicos con automatización moderna, preprocesamiento avanzado y validación rigurosa. Resulta adecuado para comparar tanto con métodos más simples como con otros enfoques de aprendizaje automático más complejos. La detección automática de estacionalidad y optimización del modelo lo hacen particularmente poderoso para gestionar diversos patrones de series temporales en un contexto de pronóstico financiero y empresarial y de ahí su selección e implementación.

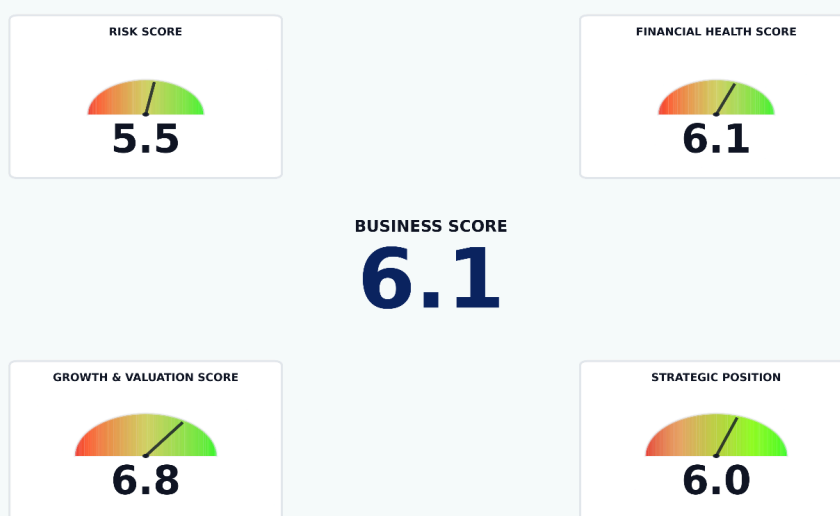


Anexo 5: Cuadros de Mando de Empresa:

CUADROS DE MANDO (ADBE):

EXECUTIVE DASHBOARD

Generated: November 17, 2025



MACRO STRATEGY ANALYSIS

Report Generated: November 17, 2025 at 12:54

STRATEGIC POSITION



6.0

STRATEGIC WORD CLOUD

AI profitability
recurring-revenue
competition margins
cash-flow market-share
innovation pricing
ecosystem
disruption regulatory execution

STRATEGIC STRENGTHS 7.4

- 8 1. Exceptional gross margin of 89% providing strong profitability buff...
- 7 3. Dominant 42% market share in creative software with high switching ...
- 8 1. Exceptional gross margin of 89% providing strong profitability buff...

+ 2 more strategic factors

STRATEGIC WEAKNESSES 3.6

- 3 1. Premium pricing vulnerability against free AI alternatives creating...
- 4 2. Slowing Digital Media ARR growth at 13% YoY indicating market satur...
- 4 3. High dependence on Creative Cloud suite representing 74% of total r...

+ 2 more strategic factors

STRATEGIC OPPORTUNITIES 6.0

- 6 1. Massive generative AI market expansion with 32.5% CAGR and \$5B ARR ...
- 6 2. Enterprise AI penetration through GenStudio and partnerships drivin...
- 6 3. Video software TAM expansion targeting 20% share of \$10B market by ...

+ 2 more strategic factors

STRATEGIC THREATS 3.6

- 3 1. AI commoditization risk eroding premium pricing power for core crea...
- 4 2. Intensifying competition from AI-native players and vertically-inte...
- 4 3. Regulatory headwinds from EU AI Act potentially slowing deployment ...

+ 2 more strategic factors

DSM Macro Strategy Analysis Service

GROWTH & VALUATION ANALYSIS

Report Generated: November 17, 2025 at 12:54

GROWTH & VALUATION SCORE



6.8

REVENUE GROWTH

YoY Revenue Growth:	+7.8%
3-Year CAGR:	+9.6%
5-Year CAGR:	+12.5%
Organic Growth:	+7.8%
Market Share Growth:	+7.8%
Customer Growth:	+7.8%
Geographic Expansion:	+0.0%
Product Line Growth:	+7.8%

Revenue Growth: HIGH

EARNINGS GROWTH

EPS Growth YoY:	+25.1%
Net Income Growth:	+25.1%
EBITDA Growth:	N/A
Operating Income Growth:	N/A
Free Cash Flow Growth:	N/A
Dividend Growth:	N/A
ROE Growth:	+15.7%
ROA Growth:	N/A

Earnings Growth: EXCELLENT

HISTORICAL GROWTH TRENDS

YoY (Most recent):	+7.8%
YoY (Prior 1):	+10.8%
YoY (Prior 2):	+10.2%
YoY (Prior 3):	+11.5%
Historical Average:	17.24%
Historical Median:	18.63%
Historical Volatility:	6.4%
Growth Momentum:	-3.01

Historical Trend: STABLE GROWTH

DUPONT GROWTH ANALYSIS

Profit Margin Change:	-10.9%
Asset Turnover Change:	-166548.8%
Equity Multiplier Change:	+0.0%
ROE Change:	-0.6%
ROE 3-Year Trend:	+40.3%
Profit Margin Trend:	+16.9%
Asset Turnover Trend:	+365428.1%
Leverage Trend:	+0.0%

Dupont Analysis: EXCEPTIONAL ROE GROWTH

ROLLING FORECASTS

Q4 2025 Forecast:	\$7,190,987,032,302.58
Q1 2026 Forecast:	\$7,226,675,408,887.68
Q2 2026 Forecast:	\$7,262,540,904,443.78
Q3 2026 Forecast:	\$7,296,584,398,000.08
Q4 2026 Forecast:	\$7,334,806,772,948.58
Q1 2027 Forecast:	\$7,371,208,917,065.38
18-Month CAGR:	6.3%
Forecast Accuracy:	TBD

Rolling Forecast: MONITOR ACCURACY

GROWTH SUSTAINABILITY

Growth Consistency:	5.7/10
Margin Expansion Trend:	+12.7%
Working Capital Efficiency:	10 days
CapEx Growth Investment:	+5.5%
Market Share Momentum:	+7.8%
Customer Retention Rate:	+80.0%
Innovation Pipeline:	1.0/10
Geographic Diversification:	5.0/10

Sustainability: FAIR SUSTAINABILITY

VALUATION

Book Value:	\$14.1B
Terminal EBIT:	\$-12.3B
Terminal FCF:	\$1.7B
Rule of 40:	24.7x
Confidence Score:	54.0%
Most Conservative:	\$408.0B
Most Aggressive:	\$498.6B
Valuation Range:	\$90.7B

Valuation Assessment: HIGH CONFIDENCE

MARKET TRENDS

Industry Growth:	8.0/10
Market Size:	9.0/10
Market Position:	9.0/10
Competitive Advantage:	8.0/10
Innovation Index:	7.0/10
Brand Strength:	9.0/10
Digital Maturity:	9.0/10
Sustainability Score:	7.0/10

Market Position: LEADER

DSM Growth & Valuation Analysis Service

FINANCIAL HEALTH

Generated: November 17, 2025 at 12:54

FINANCIAL HEALTH SCORE



6.1

PROFITABILITY METRICS

Gross Margin:	89.1%
Operating Margin:	36.2%
EBITDA Margin:	39.9%
Net Margin:	30.0%
Return on Assets:	23.0%
Return on Equity:	49.3%
ROIC:	41.4%
ROCE:	42.6%

Average Margin: 44.0%

LIQUIDITY & WORKING CAPITAL

Current Ratio:	1.07x
Quick Ratio:	1.07x
Cash Ratio:	0.72x
Cash Conversion Cycle:	N/A
Working Capital:	\$+711.0M
DSO:	33 days
DPO:	9 days
DIO:	N/A

Overall Liquidity: POOR

OTHER FINANCIAL RATIOS

Operating CF/Sales:	0.42x
Debt-to-Capital:	0.53x
FCF Yield (Enterprise):	0.60x
Return on Tangible Assets:	28.8%
Cash Return on Assets:	32.4%
Net Debt/EBITDA:	1.02x
EV/FCF:	1.68x
Operating CF/Debt:	0.61x
EBITDA/Interest:	37.47x
FCF Margin:	41.4%
Debt-to-Assets:	0.53x
Cash Conversion Efficiency:	1.14x

LEVERAGE & CAPITAL STRUCTURE

Debt-to-Equity:	1.14x
Debt/EBITDA:	1.74x
Financial Leverage:	2.14x
Interest Coverage:	34.02x
Fixed Charge Coverage:	34.02x
Breakeven Ratio:	0.59x

Overall Leverage: AGGRESSIVE

OPERATIONAL EFFICIENCY

Asset Turnover:	0.77x
Inventory Turnover:	N/A
Receivables Turnover:	11.19x
Payables Turnover:	39.57x
Fixed Asset Turnover:	10.46x
Working Capital Turnover:	32.60x
Operating Leverage:	N/A
Cash Conversion Cycle:	N/A

Overall Efficiency: HIGH

CASH FLOW ANALYSIS

Operating CF:	+9792M
Free CF:	+9599M
Investing CF:	-1022M
Financing CF:	-11006M
CapEx:	+193M
Cash Cycle:	N/A
FCF Yield:	N/A
Cash Flow Margin:	42.2%

Overall Cash Flow: STRONG

FINANCIAL DISTRESS INDICATORS

Altman Z-Score:	2.77	GREEN
Piotroski Score:	5/9	MIDGRADES
Benish H-Score:	-3.22	LOW RISK
Ohlson Q-Score:	-0.61	LOW RISK
Zmijewski Score:	-2.34	ENTRANCE
Springate Score:	23.63	SAFE
Distance to Default:	2.87	MONITOR
Debt Service Coverage:	6.37x	STRONG

ROE BREAKDOWN

ROE = Profit Margin × Asset Turnover × Financial Leverage



Profit Margin

0.30%



Asset Turnover

0.77x



Financial Leverage

2.14x

Calculated ROE: 49.3%

VALUATION MULTIPLES

EV/EBITDA:	1.5x
P/E Ratio:	2.0x
P/B Ratio:	1.0x
Price/Sales:	0.6x
EV/Sales:	0.6x
PEG Ratio:	N/A
Price/Cash Flow:	1.4x
Dividend Yield:	N/A

CAPITAL ALLOCATION & RETURNS

CapEx (annual):	\$193M
Free Cash Flow:	\$9599M
EV/FCF (proxy):	1.68x
ROIC:	41.4%
Net Debt/EBITDA:	1.74x
Cash Runway (mo):	N/A
Near-term Funding Gap:	\$-9406M

DSM Financial Analysis Service

RISK MANAGEMENT

Report Generated: November 17, 2025 at 12:54

RISK SCORE



5.5

STRENGTHS

8.0

- 9 strong operational efficiency and low risk
- 8 superior working capital management and liquidity
- 7 Optimal financial leverage

WEAKNESSES

4.7

- 4 Moderate operating leverage
- 5 Premium pricing model vulnerable to competition and AI
- 5 Market saturation and competitive challenges

KEYWORD CLOUD

Breakeven AI Leverage
Subscriptions Efficiency Margin
Cash CCC Risk
Firefly Innovation Competitive

OPPORTUNITIES

6.3

- 7 AI monetization through Firefly
- 6 Expansion into video and 3D/AR
- 6 Undervaluation

THREATS

4.3

- 4 AI-native competition
- 4 AI commoditization risk potentially
- 5 Macroeconomic sensitivity

DSM Financial Analysis Service

CUADROS DE MANDO (BRKR):

EXECUTIVE DASHBOARD

Generated: December 04, 2025

RISK SCORE



3.6

FINANCIAL HEALTH SCORE



4.5

BUSINESS SCORE

4.3

GROWTH & VALUATION SCORE



4.8

STRATEGIC POSITION

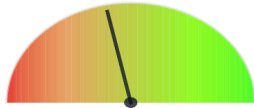


4.3

MACRO STRATEGY ANALYSIS

Report Generated: December 04, 2025 at 12:08

STRATEGIC POSITION



4.3

STRATEGIC WORD CLOUD

transformation
margin-leadership
leverage
liquidity
integration
semiconductor
academic-exposure
risk-reduction
cost-optimization
spatial-biology
M&A
profitability
diversification



STRATEGIC STRENGTHS

7.7

- 8 1. Strong tech leadership, competitive moats
- 8 2. Efficient working capital management
- 7 3. Diversified geographic revenue



STRATEGIC WEAKNESSES

3.0

- 2 1. Liquidity deterioration
- 3 2. Excessive financial leverage
- 4 3. High revenue dependency on government



STRATEGIC OPPORTUNITIES

7.0

- 8 1. Strategic sectors transformation
- 7 2. Cost optimization program
- 6 3. Expansion of high-margin revenue streams



STRATEGIC THREATS

4.7

- 5 1. Complex integration risk
- 4 2. Geopolitical exposure
- 5 3. Intense competitive pressure

DSM Macro Strategy Analysis Service

FINANCIAL HEALTH

Generated: December 04, 2025 at 12:08

FINANCIAL HEALTH SCORE



PROFITABILITY METRICS

Gross Margin:	47.2%
Operating Margin:	8.2%
EBITDA Margin:	14.4%
Net Margin:	-0.7%
Return on Assets:	-0.4%
Return on Equity:	-1.3%
ROIC:	4.8%
ROCE:	6.2%

Average Margin: 9.8%

LIQUIDITY & WORKING CAPITAL

Current Ratio:	1.60x
Quick Ratio:	0.77x
Cash Ratio:	0.14x
Cash Conversion Cycle:	147 days
Working Capital:	\$+771.9M
DSO:	74 days
DPO:	41 days
DIO:	113 days

Overall Liquidity: POOR

OTHER FINANCIAL RATIOS

Operating CF/Sales:	0.03x
Debt-to-Capital:	0.69x
FCF Yield (Enterprise):	-0.00x
Return on Tangible Assets:	-0.5%
Cash Return on Assets:	1.6%
Net Debt/EBITDA:	7.90x
Operating CF/Debt:	0.02x
EBITDA/Interest:	7.67x
FCF Margin:	-0.3%
Debt-to-Assets:	0.69x
Cash Conversion Efficiency:	-0.04x
Capital Intensity:	1.65x

LEVERAGE & CAPITAL STRUCTURE

Debt-to-Equity:	2.20x
Debt/EBITDA:	8.04x
Financial Leverage:	3.20x
Interest Coverage:	4.35x
Fixed Charge Coverage:	4.35x
Breakeven Ratio:	0.82x

Overall Leverage: AGGRESSIVE

OPERATIONAL EFFICIENCY

Asset Turnover:	0.59x
Inventory Turnover:	3.22x
Receivables Turnover:	4.90x
Payables Turnover:	8.90x
Fixed Asset Turnover:	4.22x
Working Capital Turnover:	4.46x
Operating Leverage:	N/A
Cash Conversion Cycle:	147 days

Overall Efficiency: HIGH

CASH FLOW ANALYSIS

Operating CF:	+94M
Free CF:	-11M
Investing CF:	-202M
Financing CF:	+228M
CapEx:	+105M
Cash Cycle:	147 days
FCF Yield:	N/A
Cash Flow Margin:	2.7%

Overall Cash Flow: STRONG

FINANCIAL DISTRESS INDICATORS

Altman Z-Score:	1.82	GOOD
Piotroski Score:	4/9	WEAK
Beneish M-Score:	-3.33	LOW RISK
Ohlson O-Score:	-2.53	LOW RISK
Zmijewski Score:	-0.41	SAFE
Sprinkle Score:	3.39	SAFE
Distance to Default:	1.26	DISTRESS
Debt Service Coverage:	0.35x	WEAK

ROE BREAKDOWN

ROE = Profit Margin × Asset Turnover × Financial Leverage



Calculated ROE: -1.3%

VALUATION MULTIPLES

EV/EBITDA:	3.7x
P/E Ratio:	-76.6x
P/B Ratio:	1.0x
Price/Sales:	0.5x
EV/Sales:	0.5x
PEG Ratio:	N/A
Price/Cash Flow:	19.2x
Dividend Yield:	N/A

CAPITAL ALLOCATION & RETURNS

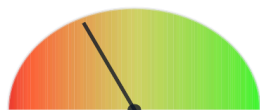
CapEx (annual):	\$105M
Free Cash Flow:	\$-11M
EV/FCF (proxy):	-376.56x
ROIC:	4.8%
Net Debt/EBITDA:	8.04x
Cash Runway (mo):	2324.3773584905657
Near-term Funding Gap:	\$116M

DSM Financial Analysis Service

RISK MANAGEMENT

Report Generated: December 04, 2025 at 12:08

RISK SCORE



3.6

STRENGTHS

6.7

- 7 Conservative financial leverage
- 7 Niche market leadership, technological moat
- 6 Moderate operating leverage

WEAKNESSES

3.3

- 3 Working capital inefficiency
- 3 Margin compression from tariffs
- 4 High exposure to end-markets

KEYWORD CLOUD

efficiency working-capital
liquidity
CCC leverage inventory
margin integration
cyclical headwinds
consolidation
China

OPPORTUNITIES

7.0

- 7 Cost-saving program
- 6 Strategic M&A in high-growth areas
- 8 Reduction of the cash conversion cycle

THREATS

3.7

- 3 Deteriorating end-markets
- 4 Integration failure in acquisitions
- 4 Intense competition from larger rivals

DSM Financial Analysis Service

CUADROS DE MANDO (CROX):

EXECUTIVE DASHBOARD

Generated: December 09, 2025

RISK SCORE



N/A

FINANCIAL HEALTH SCORE



6.2

BUSINESS SCORE

5.6

GROWTH & VALUATION SCORE



6.5

STRATEGIC POSITION

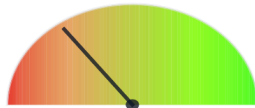


3.0

MACRO STRATEGY ANALYSIS

Report Generated: December 09, 2025 at 16:57

STRATEGIC POSITION



3.0

STRATEGIC WORD CLOUD

leverage liquidity
margins
growth
diversification
brand
sustainability
DTC
HEYDUDE
tariffs
volatility
Gen-Z



STRATEGIC STRENGTHS

8.3

- 9 1. Robust revenue growth, exceeding industry average
- 8 2. High pricing power and operational efficiency
- 8 3. Exceptional asset utilization



STRATEGIC WEAKNESSES

2.0

- 2 1. Excessive financial leverage
- 1 2. Cash balance below \$1,000M threshold
- 3 3. Declining operating margin trend



STRATEGIC OPPORTUNITIES

6.3

- 7 1. Market share expansion potential
- 6 2. Geographic and product diversification
- 6 3. Working capital optimization



STRATEGIC THREATS

2.7

- 1 1. Imminent liquidity crisis
- 3 2. Supply chain concentration in Vietnam
- 4 3. Strategic miscalculation risks

DSM Macro Strategy Analysis Service

GROWTH & VALUATION ANALYSIS

Report Generated: December 09, 2025 at 16:57

GROWTH & VALUATION SCORE



6.5

REVENUE GROWTH

YoY Revenue Growth:	+4.2%
3-Year CAGR:	+3.5%
5-Year CAGR:	+9.5%
Organic Growth:	+4.2%
Market Share Growth:	+4.2%
Customer Growth:	+4.2%
Geographic Expansion:	+0.0%
Product Line Growth:	+4.2%

Revenue Growth: MODERATE

EARNINGS GROWTH

EPS Growth YoY:	+12.0%
Net Income Growth:	+12.6%
EBITDA Growth:	N/A
Operating Income Growth:	N/A
Free Cash Flow Growth:	N/A
Dividend Growth:	N/A
ROE Growth:	+3.9%
ROA Growth:	N/A

Earnings Growth: HIGH

HISTORICAL GROWTH TRENDS

YoY (Most recent):	+4.2%
YoY (Prior 1):	-4.4%
YoY (Prior 2):	+11.4%
YoY (Prior 3):	+16.3%
Historical Average:	5.74%
Historical Median:	7.81%
Historical Volatility:	11.4%
Growth Momentum:	8.55

Historical Trend: MODERATE GROWTH

DUPONT GROWTH ANALYSIS

Profit Margin Change:	+4.8%
Asset Turnover Change:	+6.7%
Equity Multiplier Change:	-0.2%
ROE Change:	+14.4%
ROE 3-Year Trend:	+25.9%
Profit Margin Trend:	+7.5%
Asset Turnover Trend:	-23.5%
Leverage Trend:	+0.7%

Dupont Analysis: EXCEPTIONAL ROE GROWTH

ROLLING FORECASTS

Q4 2025 Forecast:	\$1,008,020,772,483.68
Q1 2026 Forecast:	\$1,013,023,510,601.88
Q2 2026 Forecast:	\$1,018,051,076,967.08
Q3 2026 Forecast:	\$1,023,103,594,800.08
Q4 2026 Forecast:	\$1,028,181,187,933.38
Q1 2027 Forecast:	\$1,033,283,980,813.98
18-Month CAGR:	2.4%
Forecast Accuracy:	TBD

Rolling Forecast: MONITOR ACCURACY

GROWTH SUSTAINABILITY

Growth Consistency:	3.7/10
Margin Expansion Trend:	+6.0%
Working Capital Efficiency:	18 days
CapEx Growth Investment:	N/A
Market Share Momentum:	+4.2%
Customer Retention Rate:	+67.8%
Innovation Pipeline:	5.0/10
Geographic Diversification:	5.0/10

Sustainability: FAIR SUSTAINABILITY

VALUATION

Book Value:	\$1.38
Terminal EBIT:	\$-0.88
Terminal FCF:	\$0.48
Rule of 40:	11.7x
Confidence Score:	54.0%
Most Conservative:	\$5.28
Most Aggressive:	\$6.48
Valuation Range:	\$1.28

Valuation Assessment: MODERATE CONFIDENCE

MARKET TRENDS

Industry Growth:	N/A
Market Size:	N/A
Market Position:	N/A
Competitive Advantage:	N/A
Innovation Index:	N/A
Brand Strength:	N/A
Digital Maturity:	N/A
Sustainability Score:	N/A

Market Position: NO DATA

DSM Growth & Valuation Analysis Service

FINANCIAL HEALTH

Generated: December 09, 2025 at 16:57

FINANCIAL HEALTH SCORE



6.2

PROFITABILITY METRICS

Gross Margin:	36.6%
Operating Margin:	17.4%
EBITDA Margin:	19.7%
Net Margin:	12.6%
Return on Assets:	14.9%
Return on Equity:	39.5%
ROIC:	25.7%
ROCE:	27.5%

Average Margin: 24.2%

LIQUIDITY & WORKING CAPITAL

Current Ratio:	1.87x
Quick Ratio:	1.25x
Cash Ratio:	0.43x
Cash Conversion Cycle:	60 days
Working Capital:	\$+766.5M
D90:	42 days
D120:	30 days
D180:	48 days

Overall Liquidity: ADEQUATE

OTHER FINANCIAL RATIOS

Operating CF/Sales:	0.16x
Debt-to-Capital:	0.62x
FCF Yield (Enterprise):	N/A
Return on Tangible Assets:	18.6%
Cash Return on Assets:	18.8%
Net Debt/EBITDA:	2.23x
Operating CF/Debt:	0.30x
EBITDA/Interest:	16.55x
Debt-to-Assets:	0.62x
Capital Intensity:	0.84x

LEVERAGE & CAPITAL STRUCTURE

Debt-to-Equity:	1.65x
Debt/EBITDA:	2.66x
Financial Leverage:	2.65x
Interest Coverage:	14.62x
Fixed Charge Coverage:	14.62x
Break-even Ratio:	0.52x

Overall Leverage: AGGRESSIVE

OPERATIONAL EFFICIENCY

Asset Turnover:	1.19x
Inventory Turnover:	7.66x
Receivables Turnover:	8.66x
Payables Turnover:	12.33x
Fixed Asset Turnover:	6.74x
Working Capital Turnover:	5.45x
Operating Leverage:	N/A
Cash Conversion Cycle:	60 days

Overall Efficiency: HIGH

CASH FLOW ANALYSIS

Operating CF:	+662M
Free CF:	N/A
Investing CF:	-240M
Financing CF:	-518M
CapEx:	N/A
Cash Cycle:	60 days
FCF Yield:	N/A
Cash Flow Margin:	15.9%

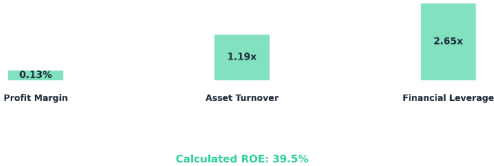
Overall Cash Flow: NEGATIVE

FINANCIAL DISTRESS INDICATORS

Altman Z-Score:	3.10	SAFE
Piotroski Score:	7/9	STRONG
Beneish M-Score:	-3.20	LOW RISK
Ohlson O-Score:	-0.25	LOW RISK
Zmijewski Score:	-1.46	MONITOR
Springate Score:	10.98	SAFE
Distance to Default:	2.12	MONITOR
Debt Service Coverage:	2.58x	STRONG

ROE BREAKDOWN

ROE = Profit Margin × Asset Turnover × Financial Leverage



VALUATION MULTIPLES

EV/EBITDA:	1.6x
P/E Ratio:	2.5x
P/B Ratio:	1.0x
Price/Sales:	0.3x
EV/Sales:	0.3x
PEG Ratio:	N/A
Price/Cash Flow:	2.0x
Dividend Yield:	N/A

CAPITAL ALLOCATION & RETURNS

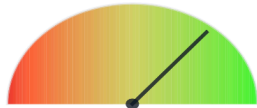
CapEx (annual):	N/A
Free Cash Flow:	N/A
EV/FCF (prevy):	N/A
ROIC:	25.7%
Net Debt/EBITDA:	2.66x
Cash Runway (mo):	N/A
Near-term Funding Gap:	N/A

DSM Financial Analysis Service

RISK MANAGEMENT

Report Generated: December 09, 2025 at 16:57

RISK SCORE



7.2

STRENGTHS

9.0

- 9 Operational resilience and profitability cushion.
- 10 Equity-heavy capital structure
- 8 Efficient cash conversion cycle, good working capital management

WEAKNESSES

4.0

- 4 High product concentration, portfolio risk
- 3 Underperformance of the last acquisition
- 5 Geographic saturation in North American market

KEYWORD CLOUD

Resilience Leverage
Break-even Tariffs
CCC HEYDUDE
Concentration Saturation
Diversification Margin Gen-Z
DTC

OPPORTUNITIES

7.0

- 7 Substantial white space for geographic diversification
- 8 Expanding direct-to-consumer (DTC) channel
- 6 Market expansion into adjacent categories like sandals

THREATS

4.7

- 4 Material tariff exposure from concentrated Vietnam sourcing
- 5 Emerging signs of brand sentiment fatigue
- 5 Macroeconomic uncertainty and volatile trade policy

DSM Financial Analysis Service

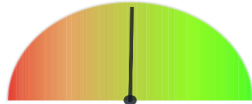
CUADROS DE MANDO (IDXX):



MACRO STRATEGY ANALYSIS

Report Generated: December 06, 2025 at 00:12

STRATEGIC POSITION



5.0

STRATEGIC WORD CLOUD

liquidity leverage
diagnostics
innovation ecosystem
profitability international
competitiveness working-capital
leverage-risk execution veterinary

✓ STRATEGIC STRENGTHS 8.0

- 8 1. Dominant market share, high customer retention
- 8 2. Excellent gross margins, razor-and-blade model
- 8 3. Strong international growth

⚠ STRATEGIC WEAKNESSES 2.7

- 3 1. Liquidity constraints
- 2 2. Highly leveraged capital structure
- 3 3. Operational inefficiencies in working capital management

📄 STRATEGIC OPPORTUNITIES 7.7

- 8 2. Accelerated international expansion
- 7 3. Working capital improvement
- 8 1. Deleveraging

⚡ STRATEGIC THREATS 2.7

- 2 1. High financial risk
- 3 2. Intensifying competitive landscape
- 3 3. Execution risk, amidst declining clients

DSM Macro Strategy Analysis Service

GROWTH & VALUATION ANALYSIS

Report Generated: December 06, 2025 at 00:12

GROWTH & VALUATION SCORE



6.7

REVENUE GROWTH

YoY Revenue Growth:	+6.9%
3-Year CAGR:	+7.4%
5-Year CAGR:	+9.0%
Organic Growth:	+6.9%
Market Share Growth:	+6.9%
Customer Growth:	+6.9%
Geographic Expansion:	+0.0%
Product Line Growth:	+6.9%

Revenue Growth: MODERATE

EARNINGS GROWTH

EPS Growth YoY:	+15.7%
Net Income Growth:	+15.7%
EBITDA Growth:	+4.1%
Operating Income Growth:	N/A
Free Cash Flow Growth:	N/A
Dividend Growth:	N/A
ROE Growth:	+4.6%
ROA Growth:	N/A

Earnings Growth: HIGH

HISTORICAL GROWTH TRENDS

YoY (Most recent):	+6.9%
YoY (Prior 1):	+6.5%
YoY (Prior 2):	+8.7%
YoY (Prior 3):	+4.7%
Historical Average:	10.1%
Historical Median:	9.79%
Historical Volatility:	3.8%
Growth Momentum:	0.46

Historical Trend: STABLE GROWTH

DUPONT GROWTH ANALYSIS

Profit Margin Change:	+3.1%
Asset Turnover Change:	N/A
Equity Multiplier Change:	+0.1%
ROE Change:	-263.0%
ROE 5-Year Trend:	+293.1%
Profit Margin Trend:	+12.7%
Asset Turnover Trend:	N/A
Leverage Trend:	-0.1%

Dupont Analysis: EXCEPTIONAL ROE GROWTH

ROLLING FORECASTS

Q4 2025 Forecast:	\$1,939,253,694,115.38
Q1 2026 Forecast:	\$1,948,878,077,502.38
Q2 2026 Forecast:	\$1,958,550,226,045.58
Q3 2026 Forecast:	\$1,968,270,376,800.08
Q4 2026 Forecast:	\$1,978,038,767,997.68
Q1 2027 Forecast:	\$1,987,855,639,052.38
18-Month CAGR:	4.8%
Forecast Accuracy:	TBD

Rolling Forecast: MONITOR ACCURACY

GROWTH SUSTAINABILITY

Growth Consistency:	5.4/10
Margin Expansion Trend:	+6.0%
Working Capital Efficiency:	28 days
CapEx Growth Investment:	+3.0%
Market Share Momentum:	+6.9%
Customer Retention Rate:	+75.8%
Innovation Pipeline:	1.0/10
Geographic Diversification:	5.0/10

Sustainability: FAIR SUSTAINABILITY

VALUATION

Book Value:	\$1.6B
Terminal EBIT:	\$-1.3B
Terminal FCF:	\$0.3B
Rule of 40:	19.6x
Confidence Score:	54.0%
Most Conservative:	\$11.0B
Most Aggressive:	\$13.4B
Valuation Range:	\$2.4B

Valuation Assessment: MODERATE CONFIDENCE

MARKET TRENDS

Industry Growth:	N/A
Market Size:	N/A
Market Position:	N/A
Competitive Advantage:	N/A
Innovation Index:	N/A
Brand Strength:	N/A
Digital Maturity:	N/A
Sustainability Score:	N/A

Market Position: NO DATA

DSM Growth & Valuation Analysis Service

FINANCIAL HEALTH

Generated: December 06, 2025 at 00:12

FINANCIAL HEALTH SCORE



6.8

PROFITABILITY METRICS

Gross Margin:	61.7%
Operating Margin:	31.3%
EBITDA Margin:	34.7%
Net Margin:	24.7%
Return on Assets:	31.2%
Return on Equity:	64.4%
ROIC:	57.0%
ROCE:	58.7%

Average Margin: 45.5%

LIQUIDITY & WORKING CAPITAL

Current Ratio:	1.31x
Quick Ratio:	0.95x
Cash Ratio:	0.27x
Cash Conversion Cycle:	67 days
Working Capital:	\$+332.0M
DSO:	49 days
DPO:	15 days
DIO:	33 days

Overall Liquidity: POOR

OTHER FINANCIAL RATIOS

Operating CF/Sales:	0.26x
Debt-to-Capital:	0.52x
FCF Yield (Enterprise):	0.57x
Return on Tangible Assets:	39.0%
Cash Return on Assets:	33.0%
Net Debt/EBITDA:	0.90x
EV/FCF:	1.76x
Operating CF/Debt:	0.64x
EBITDA/Interest:	38.94x
FCF Margin:	23.1%
Debt-to-Assets:	0.52x
Cash Conversion Efficiency:	0.74x

LEVERAGE & CAPITAL STRUCTURE

Debt-to-Equity:	1.06x
Debt/EBITDA:	1.17x
Financial Leverage:	2.06x
Interest Coverage:	35.17x
Fixed Charge Coverage:	35.17x
Break-even Ratio:	0.49x

Overall Leverage: AGGRESSIVE

OPERATIONAL EFFICIENCY

Asset Turnover:	1.27x
Inventory Turnover:	10.91x
Receivables Turnover:	7.42x
Payables Turnover:	22.62x
Fixed Asset Turnover:	5.03x
Working Capital Turnover:	12.55x
Operating Leverage:	N/A
Cash Conversion Cycle:	67 days

Overall Efficiency: HIGH

CASH FLOW ANALYSIS

Operating CF:	+1088M
Free CF:	+964M
Investing CF:	-148M
Financing CF:	-1042M
Capex:	+124M
Cash Cycle:	67 days
FCF Yield:	N/A
Cash Flow Margin:	26.1%

Overall Cash Flow: STRONG

FINANCIAL DISTRESS INDICATORS

Altman Z-Score:	3.77	SAFE
Pittman Score:	7/9	STRONG
Beneish M-Score:	-2.84	LOW RISK
Ohlson O-Score:	0.32	LOW RISK
Zmijewski Score:	-2.81	DISTRESS
Springate Score:	25.04	SAFE
Distance to Default:	3.38	SAFE
Debt Service Coverage:	6.29x	STRONG

ROE BREAKDOWN

ROE = Profit Margin × Asset Turnover × Financial Leverage



Calculated ROE: 64.4%

VALUATION MULTIPLES

EV/EBITDA:	1.1x
P/B Ratio:	1.6x
P/B Ratio:	1.0x
P/Book Value:	0.4x
EV/Sales:	0.4x
PEG Ratio:	N/A
Price/Cash Flow:	1.5x
Dividend Yield:	N/A

CAPITAL ALLOCATION & RETURNS

CapEx (annual):	\$124M
Free Cash Flow:	\$964M
EV/FCF (primary):	1.76x
ROIC:	57.0%
Net Debt/EBITDA:	1.17x
Cash Runway (mo):	N/A
Near-term Funding Gap:	\$-839M

DSM Financial Analysis Service

RISK MANAGEMENT

Report Generated: December 06, 2025 at 00:12

RISK SCORE



5.5

STRENGTHS

8.0

- 8 Dominant market leadership, competitive moat
- 8 Robust financial performance
- 8 Strong revenue model, high customer retention

WEAKNESSES

3.7

- 3 Working-capital mismanagement
- 4 Concentration risk
- 4 Sensitivity to growth deceleration

KEYWORD CLOUD

Diagnostics **Ecosystem**
Market-Leadership
Innovation
Dominance **Competitive**
Recurring-Revenue **Margins**
Moat **Consolidation**
Valuation
Working-Capital

OPPORTUNITIES

7.7

- 8 Untapped international markets
- 8 Innovation pipeline expanding
- 7 Secular trends driving structural demand

THREATS

4.0

- 4 Intensifying competition
- 4 Persistent decline clients, volume headwinds
- 4 Macroeconomic pressure

DSM Financial Analysis Service

CUADROS DE MANDO (INTU):

EXECUTIVE DASHBOARD

Generated: December 07, 2025

RISK SCORE



FINANCIAL HEALTH SCORE



BUSINESS SCORE

6.3

GROWTH & VALUATION SCORE



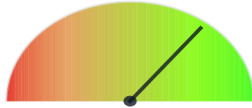
STRATEGIC POSITION



MACRO STRATEGY ANALYSIS

Report Generated: December 07, 2025 at 15:01

STRATEGIC POSITION



7.1

STRATEGIC WORD CLOUD

AI
dominance
ecosystem
market-share
subscription
growth
regulation
execution
automation
integration
profitability
competition
valuation

✓ **STRATEGIC STRENGTHS** 8.7

- 9 1. Dominant market share, pricing power and moat
- 9 2. Exceptional gross margin, efficient operational scaling
- 8 3. Robust free cash flow, capital structure, supporting investments

⚠ **STRATEGIC WEAKNESSES** 6.7

- 6 1. High execution risk, stock is priced for perfect delivery
- 7 2. Significant customer concentration
- 7 3. Execution challenges and complexity

□ **STRATEGIC OPPORTUNITIES** 7.7

- 8 1. Large untapped mid-market opportunity
- 8 2. AI-driven platform transformation
- 7 3. Expansion into international markets and embedded fintech

⚡ **STRATEGIC THREATS** 6.7

- 6 1. Regulatory and political scrutiny
- 7 3. Execution risk, complex strategy
- 7 2. Intensifying competition(fintechs & large tech

DSM Macro Strategy Analysis Service

GROWTH & VALUATION ANALYSIS

Report Generated: December 07, 2025 at 15:01

GROWTH & VALUATION SCORE



REVENUE GROWTH

YoY Revenue Growth:	+0.0%
3-Year CAGR:	+9.4%
5-Year CAGR:	+14.3%
Organic Growth:	+0.0%
Market Share Growth:	+0.0%
Customer Growth:	+0.0%
Geographic Expansion:	+0.0%
Product Line Growth:	+0.0%

Revenue Growth: MODERATE

EARNINGS GROWTH

EPS Growth YoY:	+0.0%
Net Income Growth:	+0.0%
EBITDA Growth:	N/A
Operating Income Growth:	N/A
Free Cash Flow Growth:	N/A
Dividend Growth:	N/A
ROE Growth:	-1.4%
ROA Growth:	N/A

Earnings Growth: MODERATE

HISTORICAL GROWTH TRENDS

YoY (Most recent):	+0.0%
YoY (Prior 1):	+15.6%
YoY (Prior 2):	+13.3%
YoY (Prior 3):	+12.9%
Historical Average:	15.19%
Historical Median:	13.55%
Historical Volatility:	8.1%
Growth Momentum:	-15.63

Historical Trend: STABLE GROWTH

DUPONT GROWTH ANALYSIS

Profit Margin Change:	-0.9%
Asset Turnover Change:	-37.2%
Equity Multiplier Change:	-0.3%
ROE Change:	-20.8%
ROE 5-Year Trend:	-49.8%
Profit Margin Trend:	+3.4%
Asset Turnover Trend:	-59.5%
Leverage Trend:	-1.8%

Dupont Analysis: ROE DECLINE

ROLLING FORECASTS

Q4 2025 Forecast:	\$6,681,315,157,236.88
Q1 2026 Forecast:	\$6,714,474,067,181.18
Q2 2026 Forecast:	\$6,747,797,542,526.68
Q3 2026 Forecast:	\$6,781,286,400,000.08
Q4 2026 Forecast:	\$6,814,941,460,381.58
Q1 2027 Forecast:	\$6,848,763,548,524.88
18-Month CAGR:	6.2%
Forecast Accuracy:	TBD

Rolling Forecast: MONITOR ACCURACY

GROWTH SUSTAINABILITY

Growth Consistency:	5.2/10
Margin Expansion Trend:	+4.2%
Working Capital Efficiency:	23 days
CapEx Growth Investment:	+0.0%
Market Share Momentum:	+0.0%
Customer Retention Rate:	+77.3%
Innovation Pipeline:	1.0/10
Geographic Diversification:	5.0/10

Sustainability: FAIR SUSTAINABILITY

VALUATION

Book Value:	\$19.78
Terminal EBIT:	\$-10.08
Terminal FCF:	\$1.88
Rule of 40:	3.4x
Confidence Score:	54.0%
Most Conservative:	\$319.68
Most Aggressive:	\$390.68
Valuation Range:	\$71.08

Valuation Assessment: MODERATE CONFIDENCE

MARKET TRENDS

Industry Growth:	N/A
Market Size:	N/A
Market Position:	N/A
Competitive Advantage:	N/A
Innovation Index:	N/A
Brand Strength:	N/A
Digital Maturity:	N/A
Sustainability Score:	N/A

Market Position: NO DATA

DSM Growth & Valuation Analysis Service

FINANCIAL HEALTH

Generated: December 07, 2025 at 15:01

FINANCIAL HEALTH SCORE



PROFITABILITY METRICS

Gross Margin:	79.6%
Operating Margin:	26.2%
EBITDA Margin:	30.5%
Net Margin:	20.5%
Return on Assets:	10.5%
Return on Equity:	19.6%
ROIC:	17.6%
ROCE:	18.6%

Average Margin: 27.9%

LIQUIDITY & WORKING CAPITAL

Current Ratio:	1.36x
Quick Ratio:	1.36x
Cash Ratio:	0.28x
Cash Conversion Cycle:	N/A
Working Capital:	\$+3757.0M
DSO:	176 days
DPO:	18 days
DIO:	N/A

Overall Liquidity: POOR

OTHER FINANCIAL RATIOS

Operating CF/Sales:	0.33x
Debt-to-Capital:	0.47x
FCF Yield (Enterprise):	0.35x
Return on Tangible Assets:	13.1%
Cash Return on Assets:	16.8%
Net Debt/EBITDA:	2.72x
EV/FCF:	2.82x
Operating CF/Debt:	0.36x
EBITDA/Interest:	23.27x
FCF Margin:	32.5%
Debt-to-Assets:	0.47x
Cash Conversion Efficiency:	1.24x

LEVERAGE & CAPITAL STRUCTURE

Debt-to-Equity:	0.88x
Debt/EBITDA:	3.00x
Financial Leverage:	1.88x
Interest Coverage:	19.99x
Fixed Charge Coverage:	19.99x
Break-even Ratio:	0.64x

Overall Leverage: AGGRESSIVE

OPERATIONAL EFFICIENCY

Asset Turnover:	0.51x
Inventory Turnover:	N/A
Receivables Turnover:	2.08x
Payables Turnover:	20.20x
Fixed Asset Turnover:	12.54x
Working Capital Turnover:	5.04x
Operating Leverage:	N/A
Cash Conversion Cycle:	N/A

Overall Efficiency: HIGH

CASH FLOW ANALYSIS

Operating CF:	+6207M
Free CF:	+6123M
Investing CF:	-2318M
Financing CF:	-1510M
CapEx:	+84M
Cash Cycle:	N/A
FCF Yield:	N/A
Cash Flow Margin:	33.0%

Overall Cash Flow: STRONG

FINANCIAL DISTRESS INDICATORS

Altman Z-Score:	2.19	WEAK
Piotroski Score:	7/9	STRONG
Beneish M-Score:	-2.99	LOW RISK
Oikari D-Score:	-1.97	LOW RISK
Zmijewski Score:	-2.15	DISASTROUS
Springate Score:	13.91	SAFE
Distance to Default:	2.84	MONITOR
Debt Service Coverage:	3.82x	STRONG

ROE BREAKDOWN

ROE = Profit Margin × Asset Turnover × Financial Leverage



Calculated ROE: 19.6%

VALUATION MULTIPLES

EV/EBITDA:	3.4x
P/E Ratio:	5.1x
P/B Ratio:	1.0x
Price/Sales:	1.0x
EV/Sales:	1.0x
P/S Ratio:	N/A
Price/Cash Flow:	3.2x
Dividend Yield:	N/A

CAPITAL ALLOCATION & RETURNS

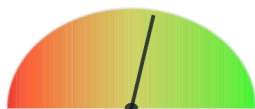
CapEx (annual):	\$84M
Free Cash Flow:	\$6123M
EV/FCF (prev):	2.82x
ROIC:	17.6%
Net Debt/EBITDA:	3.00x
Cash Runway (mo):	N/A
Near-term Funding Gap:	\$-6039M

DSM Financial Analysis Service

RISK MANAGEMENT

Report Generated: December 07, 2025 at 15:01

RISK SCORE



5.6

STRENGTHS

8.0

- 9 Superior operational efficiency
- 8 Strong financial profile robust gross margin
- 7 Balanced operating leverage

WEAKNESSES

5.7

- 6 Breakeven point in moderate risk
- 6 Financial leverage ratio, moderate balance sheet risk
- 5 Dependence on current margins

KEYWORD CLOUD

platform ecosystem AI
QuickBooks market-share TurboTax
integration growth competition
SaaS valuation margin

OPPORTUNITIES

7.3

- 8 Further optimize accounts payable management
- 7 Reduce fixed costs or increase gross margin
- 7 Leverage operational efficiency to strategic investments

THREATS

5.3

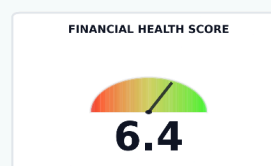
- 6 Revenue slowdown could impact the breakeven point further
- 5 Rising interest rates or credit tightening
- 5 Downturn compressing margins, weakening operating leverage

DSM Financial Analysis Service

CUADROS DE MANDO (LECO):

EXECUTIVE DASHBOARD

Generated: December 08, 2025



BUSINESS SCORE

6.0



MACRO STRATEGY ANALYSIS

Report Generated: December 08, 2025 at 22:06

STRATEGIC POSITION



6.4

STRATEGIC WORD CLOUD

Automation
Cyclicalit
Integration
Liquidity
Innovation
Sustainability
Profitability
Labor-shortage
Efficiency
Growth
Resilience
Competition



STRATEGIC STRENGTHS

8.3

- 9 1. Exceptional profitability with superior gross and operating margins
- 8 2. Dominant market leadership and integrated automation ecosystem
- 8 3. Global operational scale and diversified footprint



STRATEGIC WEAKNESSES

4.0

- 3 1. Declining quick ratio trend
- 4 2. Persistent operational inefficiencies in inventory
- 5 3. High revenue volatility and negative recent growth trends



STRATEGIC OPPORTUNITIES

7.0

- 8 1. Automation platform acceleration and robotic solutions
- 7 2. Strategic acquisitions in MRO and mobile power segments
- 6 3. Working capital optimization and efficiency improvements



STRATEGIC THREATS

4.8

- 5 1. Heavy industry cyclicality and delayed rebound
- 4 2. Geopolitical tensions and critical raw material dependencies
- 5 3. Intensifying competition in energy-efficient equipment and automati...

+ 1 more strategic factors

DSM Macro Strategy Analysis Service

GROWTH & VALUATION ANALYSIS

Report Generated: December 08, 2025 at 22:06

GROWTH & VALUATION SCORE



6.4

REVENUE GROWTH

YoY Revenue Growth:	+4.2%
3-Year CAGR:	+3.5%
5-Year CAGR:	+9.5%
Organic Growth:	+4.2%
Market Share Growth:	+4.2%
Customer Growth:	+4.2%
Geographic Expansion:	+0.0%
Product Line Growth:	+4.2%

Revenue Growth: MODERATE

EARNINGS GROWTH

EPS Growth YoY:	+12.6%
Net Income Growth:	+12.6%
EBITDA Growth:	N/A
Operating Income Growth:	N/A
Free Cash Flow Growth:	N/A
Dividend Growth:	N/A
ROE Growth:	+3.9%
ROA Growth:	N/A

Earnings Growth: HIGH

HISTORICAL GROWTH TRENDS

YoY (Most recent):	+4.2%
YoY (Prior 1):	-4.4%
YoY (Prior 2):	+11.4%
YoY (Prior 3):	+16.3%
Historical Average:	5.74%
Historical Median:	7.81%
Historical Volatility:	11.4%
Growth Momentum:	8.55

Historical Trend: MODERATE GROWTH

DUPONT GROWTH ANALYSIS

Profit Margin Change:	+4.8%
Asset Turnover Change:	+6.7%
Equity Multiplier Change:	-0.2%
ROE Change:	+14.4%
ROE 3-Year Trend:	+25.9%
Profit Margin Trend:	+7.5%
Asset Turnover Trend:	-23.5%
Leverage Trend:	+0.7%

Dupont Analysis: EXCEPTIONAL ROE GROWTH

ROLLING FORECASTS

Q4 2025 Forecast:	\$2,584,706,754,042.28
Q1 2026 Forecast:	\$2,597,534,476,799.38
Q2 2026 Forecast:	\$2,610,425,862,666.78
Q3 2026 Forecast:	\$2,623,381,227,600.08
Q4 2026 Forecast:	\$2,636,400,889,123.08
Q1 2027 Forecast:	\$2,649,485,166,335.38
18-Month CAGR:	2.4%
Forecast Accuracy:	TBD

Rolling Forecast: MONITOR ACCURACY

GROWTH SUSTAINABILITY

Growth Consistency:	3.7/10
Margin Expansion Trend:	+6.0%
Working Capital Efficiency:	18 days
CapEx Growth Investment:	N/A
Market Share Momentum:	+4.2%
Customer Retention Rate:	+67.8%
Innovation Pipeline:	5.0/10
Geographic Diversification:	5.0/10

Sustainability: FAIR SUSTAINABILITY

VALUATION

Book Value:	\$1.38
Terminal EBIT:	\$-0.88
Terminal FCF:	\$0.48
Rule of 40:	11.7x
Confidence Score:	54.0%
Most Conservative:	\$5.28
Most Aggressive:	\$6.48
Valuation Range:	\$1.28

Valuation Assessment: MODERATE CONFIDENCE

MARKET TRENDS

Industry Growth:	N/A
Market Size:	N/A
Market Position:	N/A
Competitive Advantage:	N/A
Innovation Index:	N/A
Brand Strength:	N/A
Digital Maturity:	N/A
Sustainability Score:	N/A

Market Position: NO DATA

DSM Growth & Valuation Analysis Service

FINANCIAL HEALTH

Generated: December 08, 2025 at 22:06

FINANCIAL HEALTH SCORE



PROFITABILITY METRICS

Gross Margin:	36.6%
Operating Margin:	17.4%
EBITDA Margin:	19.7%
Net Margin:	12.6%
Return on Assets:	14.9%
Return on Equity:	39.5%
ROIC:	25.7%
ROCE:	27.5%

Average Margin: 24.2%

LIQUIDITY & WORKING CAPITAL

Current Ratio:	1.87x
Quick Ratio:	1.25x
Cash Ratio:	0.43x
Cash Conversion Cycle:	60 days
Working Capital:	\$+766.5M
DSO:	42 days
DPO:	30 days
DIO:	48 days

Overall Liquidity: ADEQUATE

OTHER FINANCIAL RATIOS

Operating CF/Sales:	0.16x
Debt-to-Capital:	0.62x
FCF Yield (Enterprise):	N/A
Return on Tangible Assets:	18.6%
Cash Return on Assets:	18.8%
Net Debt/EBITDA:	2.23x
Operating CF/Debt:	0.30x
EBITDA/Interest:	16.55x
Debt-to-Assets:	0.62x
Capital Intensity:	0.84x

LEVERAGE & CAPITAL STRUCTURE

Debt-to-Equity:	1.65x
Debt/EBITDA:	2.66x
Financial Leverage:	2.65x
Interest Coverage:	14.62x
Fixed Charge Coverage:	14.62x
Break-even Ratio:	0.52x

Overall Leverage: AGGRESSIVE

OPERATIONAL EFFICIENCY

Asset Turnover:	1.19x
Inventory Turnover:	7.68x
Receivables Turnover:	8.66x
Payables Turnover:	12.33x
Fixed Asset Turnover:	6.74x
Working Capital Turnover:	5.45x
Operating Leverage:	N/A
Cash Conversion Cycle:	60 days

Overall Efficiency: HIGH

CASH FLOW ANALYSIS

Operating CF:	+662M
Free CF:	N/A
Investing CF:	-240M
Financing CF:	-518M
CapEx:	N/A
Cash Cycle:	60 days
FCF Yield:	N/A
Cash Flow Margin:	15.9%

Overall Cash Flow: NEGATIVE

FINANCIAL DISTRESS INDICATORS

Altman Z-Score:	3.10	SAFE
Piotroski Score:	7/9	STRONG
Beneish M-Score:	-3.20	LOW RISK
O'Hara D-Score:	-0.25	LOW RISK
Zmijewski Score:	-1.46	MONITOR
Springate Score:	10.98	SAFE
Distance to Default:	2.12	MONITOR
Debt Service Coverage:	2.94x	STRONG

ROE BREAKDOWN

ROE = Profit Margin × Asset Turnover × Financial Leverage



Calculated ROE: 39.5%

VALUATION MULTIPLES

EV/EBITDA:	1.6x
P/E Ratio:	2.5x
P/B Ratio:	1.0x
Price/Sales:	0.3x
EV/Sales:	0.3x
P/S Ratio:	N/A
Price/Cash Flow:	2.0x
Dividend Yield:	N/A

CAPITAL ALLOCATION & RETURNS

CapEx (annual):	N/A
Free Cash Flow:	N/A
EV/FCF (prevy):	N/A
ROIC:	25.7%
Net Debt/EBITDA:	2.66x
Cash Runway (mo):	N/A
Near-term Funding Gap:	N/A

DSM Financial Analysis Service

RISK MANAGEMENT

Report Generated: December 08, 2025 at 22:06

RISK SCORE



5.4

STRENGTHS

8.7

- 9 Substantial profitability buffer against revenue volatility
- 9 Conservative financial leverage
- 8 Dominant market leadership and high vertical integration

WEAKNESSES

4.3

- 5 MBalanced but limited earnings sensitivity to revenue changes
- 5 Working capital efficiency challenges
- 3 Heavy exposure to cyclical end-markets

KEYWORD CLOUD

automation breakeven
leverage
risk conservatism cycle
integration resilience
scale diversification
profitability mitigation

OPPORTUNITIES

7.0

- 7 Strategic pivot to high-margin automation platform
- 8 Global infrastructure and energy transition investments
- 6 Disciplined acquisitions in MRO and electrification

THREATS

3.0

- 2 Geopolitical tensions and trade policy uncertainty
- 3 Supply chain vulnerabilities for critical raw materials
- 4 Intense competition from global rivals

DSM Financial Analysis Service

CUADROS DE MANDO (NEOG):

EXECUTIVE DASHBOARD

Generated: November 18, 2025

RISK SCORE



2.8

FINANCIAL HEALTH SCORE



3.5

BUSINESS SCORE

3.5

GROWTH & VALUATION SCORE



4.5

STRATEGIC POSITION

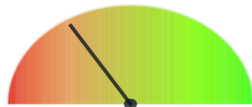


3.2

MACRO STRATEGY ANALYSIS

Report Generated: November 18, 2025 at 11:40

STRATEGIC POSITION



3.2

STRATEGIC WORD CLOUD

Competition
Strategy
Market
Position
Growth
Technology
Economy
Digital
Global
Disruption
Advantage
Innovation
Regulation
Industry
Sustainability

STRATEGIC STRENGTHS 7.7

- 8 1. Strong gross margin, pricing power
- 8 2. Market-leading, high customer retention
- 7 3. Robust liquidity position

STRATEGIC WEAKNESSES 2.3

- 2 1. Catastrophic impairment and negative free cash flow
- 2 2. Profitability deterioration and inefficiencies
- 3 3. Revenue growth deceleration, market share erosion

STRATEGIC OPPORTUNITIES 6.3

- 7 1. Portfolio optimization to diagnostics
- 6 2. Geographic expansion in Asia Pacific and Latin America
- 6 3. Integration and distribution optimization under new leadership

STRATEGIC THREATS 2.3

- 2 1. Liquidity crisis risk
- 2 2. Intensifying competition from better positioned rivals
- 3 3. Additional goodwill impairment exposure

DSM Macro Strategy Analysis Service

GROWTH & VALUATION ANALYSIS

Report Generated: November 18, 2025 at 11:40

GROWTH & VALUATION SCORE



REVENUE GROWTH

YoY Revenue Growth:	-0.9%
3-Year CAGR:	+2.5%
5-Year CAGR:	+13.6%
Organic Growth:	-0.9%
Market Share Growth:	-0.9%
Customer Growth:	-0.9%
Geographic Expansion:	+0.0%
Product Line Growth:	-0.9%

Revenue Growth: MODERATE

EARNINGS GROWTH

EPS Growth YoY:	-4.5%
Net Income Growth:	-4.5%
EBITDA Growth:	-4.5%
Operating Income Growth:	N/A
Free Cash Flow Growth:	N/A
Dividend Growth:	N/A
ROE Growth:	-15.6%
ROA Growth:	N/A

Earnings Growth: LOW

HISTORICAL GROWTH TRENDS

YoY (Most recent):	-0.9%
YoY (Prior 1):	-3.2%
YoY (Prior 2):	+12.4%
YoY (Prior 3):	+56.0%
Historical Average:	11.66%
Historical Median:	11.64%
Historical Volatility:	15.9%
Growth Momentum:	2.33

Historical Trend: MODERATE GROWTH

DUPONT GROWTH ANALYSIS

Profit Margin Change:	-130.6%
Asset Turnover Change:	-33.0%
Equity Multiplier Change:	+0.6%
ROE Change:	-58.8%
ROE 5-Year Trend:	-59.4%
Profit Margin Trend:	-129.0%
Asset Turnover Trend:	-45.4%
Leverage Trend:	+0.6%

Dupont Analysis: ROE DECLINE

ROLLING FORECASTS

Q4 2025 Forecast:	\$407,916,745,241.0B
Q1 2026 Forecast:	\$409,941,208,135.2B
Q2 2026 Forecast:	\$411,975,718,300.3B
Q3 2026 Forecast:	\$414,020,325,600.0B
Q4 2026 Forecast:	\$416,075,080,145.9B
Q1 2027 Forecast:	\$418,140,032,297.9B
18-Month CAGR:	1.7%
Forecast Accuracy:	TBD

Rolling Forecast: MONITOR ACCURACY

GROWTH SUSTAINABILITY

Growth Consistency:	3.5/10
Margin Expansion Trend:	-18.6%
Working Capital Efficiency:	36 days
CapEx Growth Investment:	-13.8%
Market Share Momentum:	-0.9%
Customer Retention Rate:	+69.8%
Innovation Pipeline:	1.0/10
Geographic Diversification:	5.0/10

Sustainability: LOW SUSTAINABILITY

VALUATION

Book Value:	\$2.1B
Terminal EBIT:	\$-0.4B
Terminal FCF:	\$0.0B
Rule of 40:	-129.9x
Confidence Score:	54.0%
Most Conservative:	\$0.6B
Most Aggressive:	\$0.8B
Valuation Range:	\$0.1B

Valuation Assessment: LOW CONFIDENCE

MARKET TRENDS

Industry Growth:	N/A
Market Size:	N/A
Market Position:	N/A
Competitive Advantage:	N/A
Innovation Index:	N/A
Brand Strength:	N/A
Digital Maturity:	N/A
Sustainability Scores:	N/A

Market Position: NO DATA

DSM Growth & Valuation Analysis Service

FINANCIAL HEALTH

Generated: November 18, 2025 at 11:40

FINANCIAL HEALTH SCORE



PROFITABILITY METRICS

Gross Margin:	46.4%
Operating Margin:	-2.3%
EBITDA Margin:	11.8%
Net Margin:	-117.6%
Return on Assets:	-30.3%
Return on Equity:	-50.4%
ROIC:	-0.6%
ROCE:	-0.6%

Average Margin: -17.9%

LIQUIDITY & WORKING CAPITAL

Current Ratio:	3.32x
Quick Ratio:	2.22x
Cash Ratio:	0.74x
Cash Conversion Cycle:	102 days
Working Capital:	\$+402.9M
DSO:	63 days
DPO:	40 days
DIO:	79 days

Overall Liquidity: EXCELLENT

OTHER FINANCIAL RATIOS

Operating CF/Sales:	0.10x
Debt-to-Capital:	0.40x
FCF Yield (Enterprise):	-0.00x
Return on Tangible Assets:	-37.9%
Cash Return on Assets:	2.5%
Net Debt/EBITDA:	12.34x
Operating CF/Debt:	0.06x
EBITDA/Interest:	1.51x
FCF Margin:	-0.4%
Debt-to-Assets:	0.40x
Cash Conversion Efficiency:	0.16x
Capital Intensity:	3.88x

LEVERAGE & CAPITAL STRUCTURE

Debt-to-Equity:	0.66x
Debt/EBITDA:	13.08x
Financial Leverage:	1.66x
Interest Coverage:	-0.29x
Fixed Charge Coverage:	-0.29x
Breakeven Ratio:	1.05x

Overall Leverage: AGGRESSIVE

OPERATIONAL EFFICIENCY

Asset Turnover:	0.26x
Inventory Turnover:	4.65x
Receivables Turnover:	5.78x
Payables Turnover:	9.21x
Fixed Asset Turnover:	2.49x
Working Capital Turnover:	2.20x
Operating Leverage:	N/A
Cash Conversion Cycle:	102 days

Overall Efficiency: HIGH

CASH FLOW ANALYSIS

Operating CF:	+87M
Free CF:	-3M
Investing CF:	+32M
Financing CF:	-102M
CapEx:	+90M
Cash Cycle:	102 days
FCF Yield:	N/A
Cash Flow Margin:	9.8%

Overall Cash Flow: STRONG

FINANCIAL DISTRESS INDICATORS

Altman Z-Score:	1.57	DISTRESS
Piotroski Score:	5/9	Moderate
Beneish M-Score:	-4.15	LOW RISK
O'Brien G-Score:	-4.70	LOW RISK
Zmijewski Score:	-0.69	SAFE
Springate Score:	3.51	SAFE
Distance to Default:	2.90	MONITOR
Debt Service Coverage:	0.48x	WEAK

ROE BREAKDOWN

ROE = Profit Margin × Asset Turnover × Financial Leverage



Calculated ROE: -50.4%

VALUATION MULTIPLES

EV/EBITDA:	19.7x
P/B Ratio:	-2.0x
P/B Ratio:	1.0x
Price/Sales:	2.3x
EV/Sales:	2.3x
PEG Ratio:	N/A
Price/Cash Flow:	25.8x
Dividend Yield:	N/A

CAPITAL ALLOCATION & RETURNS

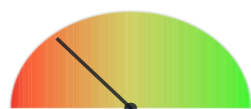
CapEx (annual):	\$90M
Free Cash Flow:	\$-3M
EV/FCF (gross):	-435.33x
ROIC:	-0.6%
Net Debt/EBITDA:	13.08x
Cash Runway (mo):	2195.764034253092
Short-term Funding Gap:	\$93M

DSM Financial Analysis Service

RISK MANAGEMENT

Report Generated: November 18, 2025 at 11:40

RISK SCORE



2.8

STRENGTHS

6.3

- 7 Conservative financial leverage, liquidity buffer
- 6 Favorable product mix
- 6 Attractive markets with regulatory tailwind

WEAKNESSES

2.3

- 2 Critical cash conversion cycle inefficiency
- 2 Critical Negative free cash flo
- 3 No downside protection

KEYWORD CLOUD

cash-conversion integration
impairment
inventory liquidity breakeven
leverage working-capital
goodwill EBITDA
DSO free-cash-flow

OPPORTUNITIES

5.7

- 6 Working capital liberation potential
- 6 New CEO leadership and strategic portfolio optimization
- 5 Market share recapture, runway for stabilization

THREATS

1.7

- 2 Continued revenue deterioration, loss relative market position
- 2 Integration failure, operational bottlenecks, customer attrition, leadership instability
- 1 Goodwill impairment, shareholder litigation destroy management credibility and investor confidence

DSM Financial Analysis Service

CUADROS DE MANDO (NKE):

EXECUTIVE DASHBOARD

Generated: December 08, 2025

RISK SCORE



FINANCIAL HEALTH SCORE



BUSINESS SCORE

6.1

GROWTH & VALUATION SCORE



STRATEGIC POSITION



MACRO STRATEGY ANALYSIS

Report Generated: December 08, 2025 at 19:11

STRATEGIC POSITION



5.2

STRATEGIC WORD CLOUD

Turnaround
Tariffs
Liquidity
Competition
China
Innovation
Wholesale
Cost
Margin
Debt
Restructuring
Nike



STRATEGIC STRENGTHS

8.0

- 7 1. Gross margin and pricing power
- 8 2. Robust free cash flow generation, low liquidity risk
- 9 3. Efficient cash conversion cycle



STRATEGIC WEAKNESSES

3.7

- 3 1. Solvency metrics, leverage risk
- 4 2. Volatile cash flows
- 4 3. Declining operating efficiency and asset turnover



STRATEGIC OPPORTUNITIES

6.3

- 6 1. Cost reduction and restructuring plan to improve margins
- 7 2. Rebalancing distribution channels by rebuilding partnerships
- 6 3. Accelerating product innovation pipeline



STRATEGIC THREATS

3.0

- 3 1. Intensified competitive disruption
- 2 2. Structural market share loss in China
- 4 3. Persistent tariff headwinds

DSM Macro Strategy Analysis Service

GROWTH & VALUATION ANALYSIS

Report Generated: December 08, 2025 at 19:11

GROWTH & VALUATION SCORE



REVENUE GROWTH

YoY Revenue Growth:	+0.3%
3-Year CAGR:	-3.2%
5-Year CAGR:	+0.8%
Organic Growth:	+0.3%
Market Share Growth:	+0.3%
Customer Growth:	+0.3%
Geographic Expansion:	+0.0%
Product Line Growth:	+0.3%

Revenue Growth: MODERATE

EARNINGS GROWTH

EPS Growth YoY:	-10.1%
Net Income Growth:	-10.1%
EBITDA Growth:	N/A
Operating Income Growth:	N/A
Free Cash Flow Growth:	N/A
Dividend Growth:	N/A
RDE Growth:	-0.4%
ROA Growth:	N/A

Earnings Growth: LOW

HISTORICAL GROWTH TRENDS

YoY (Most recent):	+0.3%
YoY (Prior 1):	-9.8%
YoY (Prior 2):	+0.3%
YoY (Prior 3):	+9.7%
Historical Average:	3.95%
Historical Median:	5.42%
Historical Volatility:	7.5%
Growth Momentum:	10.12

Historical Trend: MODERATE GROWTH

DUPONT GROWTH ANALYSIS

Profit Margin Change:	-6.6%
Asset Turnover Change:	-15.2%
Equity Multiplier Change:	-1.1%
ROE Change:	-49.2%
ROE 3-Year Trend:	-8.8%
Profit Margin Trend:	-5.4%
Asset Turnover Trend:	-24.4%
Leverage Trend:	+1.0%

Dupont Analysis: WEAK ROE GROWTH

ROLLING FORECASTS

Q4 2025 Forecast:	\$39,524,912,294,462.08
Q1 2026 Forecast:	\$35,716,108,798,043.18
Q2 2026 Forecast:	\$38,908,254,196,788.58
Q3 2026 Forecast:	\$39,101,353,200,000.08
Q4 2026 Forecast:	\$39,295,410,540,351.38
Q1 2027 Forecast:	\$39,490,430,974,004.08
18-Month CAGR:	-2.1%
Forecast Accuracy:	TBD

Rolling Forecast: MONITOR ACCURACY

GROWTH SUSTAINABILITY

Growth Consistency:	4.1/10
Margin Expansion Trend:	-2.2%
Working Capital Efficiency:	34 days
CapEx Growth Investment:	+20.2%
Market Share Momentum:	+0.3%
Customer Retention Rate:	+68.6%
Innovation Pipeline:	5.0/10
Geographic Diversification:	5.0/10

Sustainability: FAIR SUSTAINABILITY

VALUATION

Book Value:	\$13.28
Terminal EBIT:	\$-16.18
Terminal FCF:	\$4.28
Rule of 40:	-5.1x
Confidence Score:	54.0%
Most Conservative:	\$28.08
Most Aggressive:	\$34.28
Valuation Range:	\$6.28

Valuation Assessment: LOW CONFIDENCE

MARKET TRENDS

Industry Growth:	N/A
Market Size:	N/A
Market Position:	N/A
Competitive Advantage:	N/A
Innovation Index:	N/A
Brand Strength:	N/A
Digital Maturity:	N/A
Sustainability Score:	N/A

Market Position: NO DATA

DSM Growth & Valuation Analysis Service

FINANCIAL HEALTH

Generated: December 08, 2025 at 19:11

FINANCIAL HEALTH SCORE



PROFITABILITY METRICS

Gross Margin:	41.9%
Operating Margin:	7.4%
EBITDA Margin:	9.1%
Net Margin:	6.2%
Return on Assets:	7.9%
Return on Equity:	21.9%
ROIC:	12.8%
ROCE:	13.1%

Average Margin: 15.1%

LIQUIDITY & WORKING CAPITAL

Current Ratio:	2.21x
Quick Ratio:	1.50x
Cash Ratio:	0.71x
Cash Conversion Cycle:	59 days
Working Capital:	\$+12796.0M
DSO:	37 days
DPO:	37 days
DIO:	59 days

Overall Liquidity: ADEQUATE

OTHER FINANCIAL RATIOS

Operating CF/Sales:	0.08x
Debt to Capital:	0.64x
FCF Yield (Enterprise):	0.13x
Return on Tangible Assets:	9.9%
Cash Return on Assets:	9.6%
Net Debt/EBITDA:	4.80x
EV/FCF:	7.77x
Operating CF/Debt:	0.15x
EBITDA/Interest:	\$1.73x
FCF Margin:	6.5%
Debt to Assets:	0.64x
Cash Conversion Efficiency:	0.88x

LEVERAGE & CAPITAL STRUCTURE

Debt-to-Equity:	1.77x
Debt/EBITDA:	5.51x
Financial Leverage:	2.77x
Interest Coverage:	41.71x
Fixed Charge Coverage:	41.71x
Break-even Ratio:	0.82x

Overall Leverage: AGGRESSIVE

OPERATIONAL EFFICIENCY

Asset Turnover:	1.27x
Inventory Turnover:	6.20x
Receivables Turnover:	9.85x
Payables Turnover:	9.79x
Fixed Asset Turnover:	6.16x
Working Capital Turnover:	3.63x
Operating Leverage:	N/A
Cash Conversion Cycle:	59 days

Overall Efficiency: HIGH

CASH FLOW ANALYSIS

Operating CF:	+3526M
Free CF:	+3009M
Investing CF:	-166M
Financing CF:	-4796M
CapEx:	+517M
Cash Cycle:	59 days
FCF Yield:	N/A
Cash Flow Margin:	7.6%

Overall Cash Flow: STRONG

FINANCIAL DISTRESS INDICATORS

Altman Z-Score:	-2.55	CRISK
Piotroski Score:	7/9	STRONG
Beneish M-Score:	-3.17	LOW RISK
O'Keefe D-Score:	-1.75	LOW RISK
Zajackowski Score:	-1.06	MODERATE
Springate Score:	28.68	SAFE
Distance to Default:	1.66	MODERATE
Debt Service Coverage:	1.81x	ADEQUATE

ROE BREAKDOWN

ROE = Profit Margin × Asset Turnover × Financial Leverage



Calculated ROE: 21.9%

VALUATION MULTIPLES

EV/EBITDA:	3.1x
P/B Ratio:	4.6x
P/B Ratio:	1.0x
Price/Sales:	0.3x
EV/Sales:	0.3x
PEG Ratio:	N/A
Price/Cash Flow:	3.7x
Dividend Yield:	N/A

CAPITAL ALLOCATION & RETURNS

CapEx (annual):	\$517M
Free Cash Flow:	\$3009M
EV/FCF (gross):	7.77x
ROIC:	12.8%
Net Debt/EBITDA:	5.51x
Cash Runway (mo):	N/A
Near-term Funding Gap:	\$-2492M

DSM Financial Analysis Service

RISK MANAGEMENT

Report Generated: December 08, 2025 at 19:11

RISK SCORE



5.7

STRENGTHS

8.0

- 9 Conservative capital structur, minimal bankruptcy risk
- 8 Strong brand equity, global market leadership
- 7 Early signs of operational recovery

WEAKNESSES

4.7

- 5 Sensitivity to volatility & fixed cost pressure
- 5 Inventory management inefficiencies
- 4 Deteriorating financial performance

KEYWORD CLOUD



OPPORTUNITIES

7.3

- 7 Strategic partnership restoration improving
- 7 Accelerated innovation pipeline
- 8 Cost reduction program targeting supply chain optimization

THREATS

3.0

- 3 Intensifying competition from specialist brands
- 4 Structural challenges in Greater China
- 2 Tariff exposure

DSM Financial Analysis Service

CUADRO DE MANDO (QLYS):

EXECUTIVE DASHBOARD

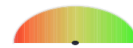
Generated: December 02, 2025

RISK SCORE



N/A

FINANCIAL HEALTH SCORE



N/A

BUSINESS SCORE

4.7

GROWTH & VALUATION SCORE



N/A

STRATEGIC POSITION



N/A

MACRO STRATEGY ANALYSIS

Report Generated: December 02, 2025 at 18:49

STRATEGIC POSITION



6.0

STRATEGIC WORD CLOUD

liquidity platform
profitability
vulnerability consolidations
disruptors retention
remediation efficiency
channel AI-regulated expansion



STRATEGIC STRENGTHS

7.3

- 8 1. Strong profitability, sector-leading margins
- 7 2. Low financial leverage, no debt or solvency risk
- 7 3. Strong growth, high future value potential



STRATEGIC WEAKNESSES

3.0

- 3 1. Low liquidity position
- 3 2. Declining operational efficiency
- 3 3. High cost of capital, market skepticism



STRATEGIC OPPORTUNITIES

6.3

- 7 1. Optimizing payable terms
- 6 2. Re-engineering the operating cost structure
- 6 3. Market expansions and product innovation



STRATEGIC THREATS

2.7

- 2 1. Liquidity crisis due to low cash reserves
- 3 2. Persistent erosion of profitability
- 3 3. Underlying cost integrity

DSM Macro Strategy Analysis Service

GROWTH & VALUATION ANALYSIS

Report Generated: December 02, 2025 at 18:49

GROWTH & VALUATION SCORE



6.6

REVENUE GROWTH

YoY Revenue Growth:	+7.5%
3-Year CAGR:	+10.1%
5-Year CAGR:	+12.5%
Organic Growth:	+7.5%
Market Share Growth:	+7.5%
Customer Growth:	+7.5%
Geographic Expansion:	+0.0%
Product Line Growth:	+7.5%

Revenue Growth: HIGH

EARNINGS GROWTH

EPS Growth YoY:	+8.9%
Net Income Growth:	+8.9%
EBITDA Growth:	+7.8%
Operating Income Growth:	N/A
Free Cash Flow Growth:	N/A
Dividend Growth:	N/A
ROE Growth:	-7.5%
ROA Growth:	N/A

Earnings Growth: MODERATE

HISTORICAL GROWTH TRENDS

YoY (Most recent):	+7.5%
YoY (Prior 1):	+9.6%
YoY (Prior 2):	+13.2%
YoY (Prior 3):	+19.1%
Historical Average:	14.88%
Historical Median:	14.3%
Historical Volatility:	4.2%
Growth Momentum:	-2.1

Historical Trend: STABLE GROWTH

DUPONT GROWTH ANALYSIS

Profit Margin Change:	+3.7%
Asset Turnover Change:	+13.3%
Equity Multiplier Change:	+0.3%
ROE Change:	+16.0%
ROE 5-Year Trend:	+31.5%
Profit Margin Trend:	+19.3%
Asset Turnover Trend:	+16.3%
Leverage Trend:	+0.4%

Dupont Analysis: EXCEPTIONAL ROE GROWTH

ROLLING FORECASTS

Q4 2025 Forecast:	\$227,334,111,897.88
Q1 2026 Forecast:	\$228,462,355,539.48
Q2 2026 Forecast:	\$229,596,198,577.08
Q3 2026 Forecast:	\$230,735,668,800.08
Q4 2026 Forecast:	\$231,880,794,135.88
Q1 2027 Forecast:	\$233,031,602,650.28
18-Month CAGR:	6.6%
Forecast Accuracy:	TBD

Rolling Forecast: MONITOR ACCURACY

GROWTH SUSTAINABILITY

Growth Consistency:	5.8/10
Margin Expansion Trend:	+14.4%
Working Capital Efficiency:	31 days
CapEx Growth Investment:	-18.1%
Market Share Momentum:	+7.5%
Customer Retention Rate:	+79.4%
Innovation Pipeline:	1.0/10
Geographic Diversification:	5.0/10

Sustainability: FAIR SUSTAINABILITY

VALUATION

Book Value:	\$0.58
Terminal EBIT:	\$-0.38
Terminal FCF:	\$0.18
Rule of 40:	54.2x
Confidence Score:	54.0%
Most Conservative:	N/A
Most Aggressive:	N/A
Valuation Range:	N/A

Valuation Assessment: HIGH CONFIDENCE

MARKET TRENDS

Industry Growth:	N/A
Market Size:	N/A
Market Position:	N/A
Competitive Advantage:	N/A
Innovation Index:	N/A
Brand Strength:	N/A
Digital Maturity:	N/A
Sustainability Score:	N/A

Market Position: NO DATA

DSM Growth & Valuation Analysis Service

FINANCIAL HEALTH

Generated: December 02, 2025 at 18:49

FINANCIAL HEALTH SCORE



PROFITABILITY METRICS

Gross Margin:	82.4%
Operating Margin:	32.5%
EBITDA Margin:	34.9%
Net Margin:	29.0%
Return on Assets:	19.4%
Return on Equity:	39.6%
ROIC:	34.4%
ROCE:	39.0%

Average Margin: 38.9%

LIQUIDITY & WORKING CAPITAL

Current Ratio:	1.37x
Quick Ratio:	1.37x
Cash Ratio:	0.54x
Cash Conversion Cycle:	N/A
Working Capital:	\$+157.3M
DSO:	92 days
DPO:	1 days
DIO:	N/A

Overall Liquidity: ADEQUATE

OTHER FINANCIAL RATIOS

Operating CF/Sales:	0.43x
Debt-to-Capital:	0.51x
FCF Yield (Enterprise):	0.55x
Return on Tangible Assets:	24.3%
Cash Return on Assets:	28.9%
Net Debt/EBITDA:	1.64x
EV/FCF:	1.83x
Operating CF/Debt:	0.57x
EBITDA/Interest:	9.08x
FCF Margin:	41.6%
Debt-to-Assets:	0.51x
Cash Conversion Efficiency:	1.28x

LEVERAGE & CAPITAL STRUCTURE

Debt-to-Equity:	1.04x
Debt/EBITDA:	2.18x
Financial Leverage:	2.04x
Interest Coverage:	8.47x
Fixed Charge Coverage:	8.47x
Break-even Ratio:	0.61x

Overall Leverage: AGGRESSIVE

OPERATIONAL EFFICIENCY

Asset Turnover:	0.67x
Inventory Turnover:	N/A
Receivables Turnover:	3.97x
Payables Turnover:	514.26x
Fixed Asset Turnover:	9.16x
Working Capital Turnover:	4.15x
Operating Leverage:	N/A
Cash Conversion Cycle:	N/A

Overall Efficiency: HIGH

CASH FLOW ANALYSIS

Operating CF:	+281M
Free CF:	+271M
Investing CF:	-109M
Financing CF:	-182M
CapEx:	+10M
Cash Cycle:	N/A
FCF Yield:	N/A
Cash Flow Margin:	43.1%

Overall Cash Flow: STRONG

FINANCIAL DISTRESS INDICATORS

Zmian Z-score:	2.67	GREEN
Piotroski Score:	6/9	MODERATE
Beneish M-Score:	-3.19	LOW RISK
Altman Z-Score:	0.28	LOW RISK
Zajackowski Score:	-2.31	UNSTRESS
Springate Score:	6.69	SAFE
Distance to Default:	2.82	MODERATE
Debt Service Coverage:	4.34x	STRONG

ROE BREAKDOWN

ROE = Profit Margin × Asset Turnover × Financial Leverage



Calculated ROE: 39.6%

VALUATION MULTIPLES

EV/EBITDA:	2.1x
P/E Ratio:	2.5x
P/B Ratio:	1.0x
Price/Sales:	0.7x
EV/Sales:	0.7x
P/D Ratio:	N/A
Price/Cash Flow:	1.7x
Dividend Yield:	N/A

CAPITAL ALLOCATION & RETURNS

CapEx (annual):	\$10M
Free Cash Flow:	\$271M
EV/FCF (pre-tax):	1.83x
ROIC:	34.4%
Net Debt/EBITDA:	2.18x
Cash Runway (mo):	N/A
Near-term Funding Gap:	\$-261M

DSM Financial Analysis Service

RISK MANAGEMENT

Report Generated: December 02, 2025 at 18:49

RISK SCORE



5.5

STRENGTHS

7.3

- 8 Highly profitable and scalable SaaS business model.
- 7 Operational resilience, predictable earnings growth.
- 7 Prudent financial leverage

WEAKNESSES

3.3

- 3 Working capital inefficiency and liquidity constraints.
- 3 Very low days payables outstanding
- 4 Operational risks

KEYWORD CLOUD

Working capital
leverage CCC cybersecurity
vulnerability
breakeven SaaS receivables
margin liquidity AI

OPPORTUNITIES

6.3

- 7 High-growth cybersecurity market with AI
- 6 FedRAMP High authorization and strong government positioning moat
- 6 Robust profitability and low financial leverage

THREATS

4.0

- 4 Intense competitive pressure from larger players
- 4 Vulnerability to economic uncertainties

DSM Financial Analysis Service

CUADROS DE MANDO (RRX):

EXECUTIVE DASHBOARD

Generated: December 06, 2025

RISK SCORE



5.7

FINANCIAL HEALTH SCORE



6.0

BUSINESS SCORE

6.2

GROWTH & VALUATION SCORE



6.7

STRATEGIC POSITION

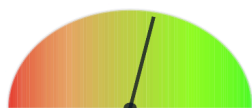


5.7

MACRO STRATEGY ANALYSIS

Report Generated: December 06, 2025 at 11:28

STRATEGIC POSITION



5.7

STRATEGIC WORD CLOUD

Transformation
automation
robotics
margins
portfolio
tariffs
aerospace
volatility
data-center
synergy
liquidity
integration

STRATEGIC STRENGTHS 7.7

- 8 1. Strong market positioning, synergistic portfolio
- 8 2. Effective synergy from integration efforts
- 7 3. Robust free cash flow generation

STRATEGIC WEAKNESSES 3.7

- 3 1. Elevated financial leverage
- 4 2. Profitability margins under pressure
- 4 3. Constrained liquidity position limiting operational resilience and ...

STRATEGIC OPPORTUNITIES 7.3

- 8 1. Focused R&D and partnerships
- 7 2. Further portfolio optimization and integration
- 7 3. Reducing debt to investment-grade metrics

STRATEGIC THREATS 4.7

- 5 1. Exposure to cyclical demand and macroeconomic volatility
- 5 2. Material tariff and geopolitical risks
- 4 3. Execution risk in competitive markets against global competitors

DSM Macro Strategy Analysis Service

GROWTH & VALUATION ANALYSIS

Report Generated: December 06, 2025 at 11:28

GROWTH & VALUATION SCORE



REVENUE GROWTH

YoY Revenue Growth:	-2.7%
3-Year CAGR:	+4.0%
5-Year CAGR:	+15.1%
Organic Growth:	-2.7%
Market Share Growth:	-2.7%
Customer Growth:	-2.7%
Geographic Expansion:	+0.0%
Product Line Growth:	-2.7%

Revenue Growth: MODERATE

EARNINGS GROWTH

EPS Growth YoY:	+30.5%
Net Income Growth:	+30.5%
EBITDA Growth:	-1.1%
Operating Income Growth:	N/A
Free Cash Flow Growth:	N/A
Dividend Growth:	N/A
ROE Growth:	+1.0%
ROA Growth:	N/A

Earnings Growth: HIGH

HISTORICAL GROWTH TRENDS

YoY (Most recent):	-2.7%
YoY (Prior 1):	-3.5%
YoY (Prior 2):	+19.8%
YoY (Prior 3):	+36.9%
Historical Average:	6.48%
Historical Median:	0.77%
Historical Volatility:	16.4%
Growth Momentum:	0.79

Historical Trend: MODERATE GROWTH

DUPONT GROWTH ANALYSIS

Profit Margin Change:	-2.3%
Asset Turnover Change:	-109538.3%
Equity Multiplier Change:	+0.0%
ROE Change:	-4.0%
ROE 3-Year Trend:	-3.4%
Profit Margin Trend:	+0.2%
Asset Turnover Trend:	-101811.4%
Leverage Trend:	+0.0%

Dupont Analysis: MODERATE ROE GROWTH

ROLLING FORECASTS

Q4 2025 Forecast:	\$3,309,437,400,186.18
Q1 2026 Forecast:	\$3,325,861,911,549.18
Q2 2026 Forecast:	\$3,342,367,936,638.08
Q3 2026 Forecast:	\$3,358,955,880,000.08
Q4 2026 Forecast:	\$3,375,626,148,189.98
Q1 2027 Forecast:	\$3,392,379,149,780.08
18-Month CAGR:	2.7%
Forecast Accuracy:	TBD

Rolling Forecast: MONITOR ACCURACY

GROWTH SUSTAINABILITY

Growth Consistency:	2.9/10
Margin Expansion Trend:	-8.5%
Working Capital Efficiency:	16 days
CapEx Growth Investment:	-8.9%
Market Share Momentum:	-2.7%
Customer Retention Rate:	+65.3%
Innovation Pipeline:	5.0/10
Geographic Diversification:	5.0/10

Sustainability: FAIR SUSTAINABILITY

VALUATION

Book Value:	\$6.38
Terminal EBIT:	\$-1.58
Terminal FCF:	\$0.58
Rule of 40:	-2.5x
Confidence Score:	54.0%
Most Conservative:	\$7.48
Most Aggressive:	\$9.18
Valuation Range:	\$1.68

Valuation Assessment: LOW CONFIDENCE

MARKET TRENDS

Industry Growth:	N/A
Market Size:	N/A
Market Position:	N/A
Competitive Advantage:	N/A
Innovation Index:	N/A
Brand Strength:	N/A
Digital Maturity:	N/A
Sustainability Score:	N/A

Market Position: NO DATA

DSM Growth & Valuation Analysis Service

FINANCIAL HEALTH

Generated: December 06, 2025 at 11:28

FINANCIAL HEALTH SCORE



PROFITABILITY METRICS

Gross Margin:	36.7%
Operating Margin:	11.1%
EBITDA Margin:	19.7%
Net Margin:	4.4%
Return on Assets:	1.8%
Return on Equity:	4.1%
ROIC:	2.3%
ROCE:	5.1%

Average Margin: 10.7%

LIQUIDITY & WORKING CAPITAL

Current Ratio:	2.26x
Quick Ratio:	1.25x
Cash Ratio:	0.32x
Cash Conversion Cycle:	88 days
Working Capital:	\$+1535.6M
DSO:	52 days
DPO:	40 days
DIO:	76 days

Overall Liquidity: ADEQUATE

OTHER FINANCIAL RATIOS

Operating CF/Sales:	0.18x
Debt-to-Capital:	0.55x
FCF Yield (Enterprise):	0.12x
Return on Tangible Assets:	2.3%
Cash Return on Assets:	7.4%
Net Debt/EBITDA:	6.39x
EV/FCF:	8.29x
Operating CF/Debt:	0.13x
EBITDA/Interest:	5.24x
FCF Margin:	15.9%
Debt-to-Assets:	0.55x
Cash Conversion Efficiency:	1.44x

LEVERAGE & CAPITAL STRUCTURE

Debt-to-Equity:	1.24x
Debt/EBITDA:	6.71x
Financial Leverage:	2.24x
Interest Coverage:	1.82x
Fixed Charge Coverage:	1.82x
Break-even Ratio:	N/A

Overall Leverage: AGGRESSIVE

OPERATIONAL EFFICIENCY

Asset Turnover:	0.42x
Inventory Turnover:	4.78x
Receivables Turnover:	6.97x
Payables Turnover:	9.03x
Fixed Asset Turnover:	5.53x
Working Capital Turnover:	3.62x
Operating Leverage:	N/A
Cash Conversion Cycle:	88 days

Overall Efficiency: HIGH

CASH FLOW ANALYSIS

Operating CF:	+1036M
Free CF:	+936M
Investing CF:	-73M
Financing CF:	-1019M
CapEx:	+100M
Cash Cycle:	88 days
FCF Yield:	N/A
Cash Flow Margin:	17.6%

Overall Cash Flow: STRONG

FINANCIAL DISTRESS INDICATORS

Altman Z-Score:	1.74	DISTRESS
Piotroski Score:	6/9	MODERATE
Banach H-Score:	-2.85	LOW RISK
Ohlson O-Score:	-1.70	LOW RISK
Zajackowski Score:	-1.27	MONITOR
Springate Score:	1.83	SAFE
Distance to Default:	1.98	MONITOR
Debt Service Coverage:	1.06x	WEAK

ROE BREAKDOWN

ROE = Profit Margin × Asset Turnover × Financial Leverage



Calculated ROE: 4.1%

VALUATION MULTIPLES

EV/EBITDA:	5.4x
P/E Ratio:	24.4x
P/B Ratio:	1.0x
P/Book Value:	1.1x
EV/Sales:	1.1x
P/S Ratio:	N/A
Price/Cash Flow:	6.0x
Dividend Yield:	N/A

CAPITAL ALLOCATION & RETURNS

CapEx (annual):	\$100M
Free Cash Flow:	\$936M
EV/FCF (avg):	8.29x
ROIC:	2.3%
Net Debt/EBITDA:	6.71x
Cash Runway (mo):	N/A
Short-term Funding Gap:	\$-837M

DSM Financial Analysis Service

RISK MANAGEMENT

Report Generated: December 06, 2025 at 11:28

RISK SCORE



5.7

STRENGTHS

7.3

- 8 Conservative financial leverage
- 7 Low operational risk and cost flexibility
- 7 Stable sensitivity to revenue fluctuations

WEAKNESSES

4.0

- 3 Working capital inefficiencies and liquidity drag
- 4 Inventory management challenges, potential obsolescence risk
- 5 Moderate breakeven risk

KEYWORD CLOUD

transformation automation
electrification
integration robotics deleveraging
cybersecurity synergy
end-markets volatility
inflation margin

OPPORTUNITIES

6.3

- 6 Reducing DSO through improved receivables
- 7 Extending DPO to optimize working capital and supplier terms
- 6 Lowering the breakeven ratio, fixed cost reduction or margin improvement

THREATS

4.3

- 4 Risk of liquidity constraints
- 4 Inventory carrying costs, potential write-downs in downturns
- 5 Operating leverage sensitivity to revenue declines

DSM Financial Analysis Service

CUADROS DE MANDO (SBUX):

EXECUTIVE DASHBOARD

Generated: December 07, 2025

RISK SCORE



5.2

FINANCIAL HEALTH SCORE



6.0

BUSINESS SCORE

5.4

GROWTH & VALUATION SCORE



6.0

STRATEGIC POSITION



4.2

MACRO STRATEGY ANALYSIS

Report Generated: December 07, 2025 at 01:03

STRATEGIC POSITION



4.2

STRATEGIC WORD CLOUD

turnaround restructuring
competition
labor margins
commoditization digitalization
international brand premiumization
unionization supply-chain

STRATEGIC STRENGTHS

7.0

- 8 1. Dominant global market share and powerful brand
- 7 2. Robust digital loyalty ecosystem
- 6 3. Clear strategic turnaround

+ 1 more strategic factors



STRATEGIC WEAKNESSES

2.5

- 2 1. Deteriorated profitability
- 3 2. Structural decline in customer traffic and negative comparable stor...
- 2 3. Protracted and escalating labor union crisis

+ 1 more strategic factors



STRATEGIC OPPORTUNITIES

6.5

- 6 1. Portfolio rationalization and strategic expansion
- 7 2. Optimization of digital ordering, inventory management, and labor s...
- 6 3. Coffee market tailwind

+ 1 more strategic factors



STRATEGIC THREATS

2.3

- 3 1. Intense and multifaceted competition
- 2 2. Unprecedented commodity cost inflation
- 2 3. Financial leverage, negative equity, and pressured free cash flow

DSM Macro Strategy Analysis Service

GROWTH & VALUATION ANALYSIS

Report Generated: December 07, 2025 at 01:03

GROWTH & VALUATION SCORE



REVENUE GROWTH

YoY Revenue Growth:	+0.0%
3-Year CAGR:	+1.1%
5-Year CAGR:	+5.0%
Organic Growth:	+0.0%
Market Share Growth:	+0.0%
Customer Growth:	+0.0%
Geographic Expansion:	+0.0%
Product Line Growth:	+0.0%

Revenue Growth: MODERATE

EARNINGS GROWTH

EPS Growth YoY:	+0.0%
Net Income Growth:	+0.0%
EBITDA Growth:	N/A
Operating Income Growth:	N/A
Free Cash Flow Growth:	N/A
Dividend Growth:	N/A
ROE Growth:	+2.0%
ROA Growth:	N/A

Earnings Growth: MODERATE

HISTORICAL GROWTH TRENDS

YoY (Most recent):	+0.0%
YoY (Prior 1):	+2.8%
YoY (Prior 2):	+0.6%
YoY (Prior 3):	+11.6%
Historical Average:	6.08%
Historical Median:	6.13%
Historical Volatility:	8.7%
Growth Momentum:	-2.79

Historical Trend: STABLE GROWTH

DUPONT GROWTH ANALYSIS

Profit Margin Change:	-9.5%
Asset Turnover Change:	+17.2%
Equity Multiplier Change:	-0.2%
ROE Change:	+30.9%
ROE 5-Year Trend:	-70.8%
Profit Margin Trend:	-8.2%
Asset Turnover Trend:	-32.6%
Leverage Trend:	-6.4%

Dupont Analysis: ROE DECLINE

ROLLING FORECASTS

Q4 2025 Forecast:	\$26,107,357,160,543.48
Q1 2026 Forecast:	\$26,236,926,187,688.48
Q2 2026 Forecast:	\$26,367,138,257,049.18
Q3 2026 Forecast:	\$26,497,996,560,000.08
Q4 2026 Forecast:	\$26,629,504,303,754.38
Q1 2027 Forecast:	\$26,761,664,711,442.18
18-Month CAGR:	0.7%
Forecast Accuracy:	TBD

Rolling Forecast: MONITOR ACCURACY

GROWTH SUSTAINABILITY

Growth Consistency:	4.2/10
Margin Expansion Trend:	+0.4%
Working Capital Efficiency:	21 days
CapEx Growth Investment:	+0.0%
Market Share Momentum:	+0.0%
Customer Retention Rate:	+69.6%
Innovation Pipeline:	5.0/10
Geographic Diversification:	5.0/10

Sustainability: FAIR SUSTAINABILITY

VALUATION

Book Value:	\$-8.1B
Terminal EBIT:	\$-4.9B
Terminal FCF:	\$5.8B
Rule of 40:	-8.2x
Confidence Score:	54.0%
Most Conservative:	\$59.3B
Most Aggressive:	\$72.5B
Valuation Range:	\$13.2B

Valuation Assessment: LOW CONFIDENCE

MARKET TRENDS

Industry Growth:	N/A
Market Size:	N/A
Market Position:	N/A
Competitive Advantage:	N/A
Innovation Index:	N/A
Brand Strength:	N/A
Digital Maturity:	N/A
Sustainability Score:	N/A

Market Position: NO DATA

DSM Growth & Valuation Analysis Service

FINANCIAL HEALTH

Generated: December 07, 2025, at 01:03

FINANCIAL HEALTH SCORE



PROFITABILITY METRICS

Gross Margin:	22.8%
Operating Margin:	9.6%
EBITDA Margin:	14.4%
Net Margin:	5.0%
Return on Assets:	N/A
Return on Equity:	0.0%
ROIC:	N/A
ROCE:	N/A

Average Margin: 10.4%

LIQUIDITY & WORKING CAPITAL

Current Ratio:	0.72x
Quick Ratio:	0.51x
Cash Ratio:	0.32x
Cash Conversion Cycle:	5 days
Working Capital:	\$-2828.1M
DSO:	13 days
DPO:	29 days
DIO:	21 days

Overall Liquidity: POOR

OTHER FINANCIAL RATIOS

Operating CF/Sales:	0.13x
Debt-to-Capital:	N/A
FCF Yield (Enterprise):	N/A
EBITDA/Interest:	9.85x
FCF Margin:	6.5%
Cash Conversion Efficiency:	0.68x

LEVERAGE & CAPITAL STRUCTURE

Debt-to-Equity:	N/A
Debt/EBITDA:	N/A
Financial Leverage:	0.17x
Interest Coverage:	6.60x
Fixed Charge Coverage:	6.60x
Breakeven Ratio:	0.31x

Overall Leverage: AGGRESSIVE

OPERATIONAL EFFICIENCY

Asset Turnover:	0.00x
Inventory Turnover:	17.01x
Receivables Turnover:	29.11x
Payables Turnover:	12.63x
Fixed Asset Turnover:	2.09x
Working Capital Turnover:	N/A
Operating Leverage:	N/A
Cash Conversion Cycle:	5 days

Overall Efficiency: HIGH

CASH FLOW ANALYSIS

Operating CF:	+4748M
Free CF:	+2442M
Investing CF:	-2485M
Financing CF:	-2298M
Capex:	-2306M
Cash Cycle:	5 days
FCF Yield:	N/A
Cash Flow Margin:	12.8%

Overall Cash Flow: STRONG

FINANCIAL DISTRESS INDICATORS

Altman Z-Score:	N/A	N/A
Pittetvski Score:	4/9	WEAK
Beneish M-Score:	-2.70	LOW RISK
Ohlson O-Score:	N/A	N/A
Zmijewski Score:	N/A	N/A
Springate Score:	N/A	N/A
Distance to Default:	N/A	N/A
Debt Service Coverage:	0.75x	STRONG

ROE BREAKDOWN

ROE = Profit Margin × Asset Turnover × Financial Leverage

Profit Margin	0.05%
Asset Turnover	0.00x
Financial Leverage	0.17x

Calculated ROE: 0.0%

VALUATION MULTIPLES

EV/EBITDA:	N/A
P/E Ratio:	N/A
P/B Ratio:	N/A
P/Book Value:	N/A
EV/Sales:	N/A
PEG Ratio:	N/A
Price/Cash Flow:	N/A
Dividend Yield:	N/A

CAPITAL ALLOCATION & RETURNS

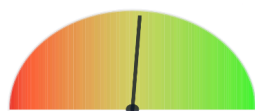
CapEx (annual):	\$2306M
Free Cash Flow:	\$2442M
EV/FCF (proxy):	N/A
ROIC:	N/A
Net Debt/EBITDA:	N/A
Cash Runway (mo):	N/A
Near-term Funding Gap:	\$-136M

DSM Financial Analysis Service

RISK MANAGEMENT

Report Generated: December 07, 2025 at 01:03

RISK SCORE



5.2

STRENGTHS

9.3

- 9 Exceptional margin of safety against revenue volatility
- 9 Conservative financial leverage
- 10 Superior working capital management and operational efficiency

WEAKNESSES

2.3

- 3 Severe profit compression
- 2 Structural traffic declines in core North American market
- 2 Labor relations crisis involving strikes

KEYWORD CLOUD

Breakeven **Leverage**
Risk
Cycle **Cash**
Operating
Financial **Margin**
Assessment
Conversion **Profile**
DSO

OPPORTUNITIES

6.7

- 7 Repositioning on operational excellence, store portfolio, and brand
- 7 Digital and technology initiatives
- 6 Long-term store expansion potential in underserved regions

THREATS

2.7

- 3 Intensifying competition from value-focused chains
- 2 Rising Arabica coffee costs
- 3 Consumer discretionary pullback

DSM Financial Analysis Service

CUADROS DE MANDO (WWD):

EXECUTIVE DASHBOARD

Generated: December 06, 2025

RISK SCORE



5.3

FINANCIAL HEALTH SCORE



5.8

BUSINESS SCORE

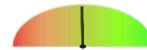
5.6

GROWTH & VALUATION SCORE



5.9

STRATEGIC POSITION

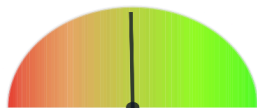


5.0

MACRO STRATEGY ANALYSIS

Report Generated: December 06, 2025 at 18:59

STRATEGIC POSITION



5.0

STRATEGIC WORD CLOUD

Liquidity Aerospace
Actuation
ROE
Aftermarket Defense
Innovation Solvency
Data-WACC China
Competitive-Capex



STRATEGIC STRENGTHS

9.0

- 10 1. Exceptional Return on Equity
- 9 2. Highly efficient accounts receivable & Cash management
- 8 3. Strong and consistent operating cash flow



STRATEGIC WEAKNESSES

2.3

- 3 2. Historical unprofitability, operational or pricing challenges
- 2 1. Low liquidity
- 2 3. Solvency issues



STRATEGIC OPPORTUNITIES

6.7

- 7 1. New markets or product innovations
- 6 2. Optimization of working capital management
- 7 3. Capitalizing on market tailwinds



STRATEGIC THREATS

4.0

- 5 1. Investment hurdle rates
- 3 2. Difficulty accessing external financing
- 4 3. Intensified competitive pressure

DSM Macro Strategy Analysis Service

GROWTH & VALUATION ANALYSIS

Report Generated: December 06, 2025 at 18:59

GROWTH & VALUATION SCORE



REVENUE GROWTH

YoY Revenue Growth:	+3.1%
3-Year CAGR:	+12.9%
5-Year CAGR:	+6.5%
Organic Growth:	+3.1%
Market Share Growth:	+3.1%
Customer Growth:	+3.1%
Geographic Expansion:	+0.0%
Product Line Growth:	+3.1%

Revenue Growth: MODERATE

EARNINGS GROWTH

EPS Growth YoY:	+4.0%
Net Income Growth:	+4.0%
EBITDA Growth:	-1.3%
Operating Income Growth:	N/A
Free Cash Flow Growth:	N/A
Dividend Growth:	N/A
ROE Growth:	-0.2%
ROA Growth:	N/A

Earnings Growth: MODERATE

HISTORICAL GROWTH TRENDS

YoY (Most recent):	+3.1%
YoY (Prior 1):	+14.1%
YoY (Prior 2):	+22.3%
YoY (Prior 3):	+6.1%
Historical Average:	6.01%
Historical Median:	4.92%
Historical Volatility:	11.9%
Growth Momentum:	-10.99

Historical Trend: MODERATE GROWTH

DUPONT GROWTH ANALYSIS

Profit Margin Change:	+1.7%
Asset Turnover Change:	N/A
Equity Multiplier Change:	+0.0%
ROE Change:	+3.9%
ROE 5-Year Trend:	+2.1%
Profit Margin Trend:	+2.4%
Asset Turnover Trend:	N/A
Leverage Trend:	+0.0%

Dupont Analysis: MODERATE ROE GROWTH

ROLLING FORECASTS

Q4 2025 Forecast:	\$2,066,918,617,447.78
Q1 2026 Forecast:	\$2,077,176,593,113.58
Q2 2026 Forecast:	\$2,087,485,478,410.68
Q3 2026 Forecast:	\$2,097,845,526,000.08
Q4 2026 Forecast:	\$2,108,256,989,796.78
Q1 2027 Forecast:	\$2,118,720,124,975.88
18-Month CAGR:	8.4%
Forecast Accuracy:	TBD

Rolling Forecast: MONITOR ACCURACY

GROWTH SUSTAINABILITY

Growth Consistency:	3.6/10
Margin Expansion Trend:	+2.2%
Working Capital Efficiency:	35 days
CapEx Growth Investment:	+6.6%
Market Share Momentum:	+3.1%
Customer Retention Rate:	+67.7%
Innovation Pipeline:	1.0/10
Geographic Diversification:	5.0/10

Sustainability: LOW SUSTAINABILITY

VALUATION

Book Value:	\$2.28
Terminal EBIT:	\$-0.58
Terminal FCF:	\$0.68
Rule of 40:	5.5x
Confidence Score:	54.0%
Most Conservative:	N/A
Most Aggressive:	N/A
Valuation Range:	N/A

Valuation Assessment: MODERATE CONFIDENCE

MARKET TRENDS

Industry Growth:	N/A
Market Size:	N/A
Market Position:	N/A
Competitive Advantage:	N/A
Innovation Index:	N/A
Brand Strength:	N/A
Digital Maturity:	N/A
Sustainability Score:	N/A

Market Position: NO DATA

DSM Growth & Valuation Analysis Service

FINANCIAL HEALTH

Generated: December 06, 2025 at 18:59

FINANCIAL HEALTH SCORE



5.8

PROFITABILITY METRICS

Gross Margin:	25.9%
Operating Margin:	12.4%
EBITDA Margin:	15.7%
Net Margin:	11.3%
Return on Assets:	8.9%
Return on Equity:	17.8%
ROIC:	10.9%
ROCE:	12.3%

Average Margin: 14.4%

LIQUIDITY & WORKING CAPITAL

Current Ratio:	1.89x
Quick Ratio:	1.23x
Cash Ratio:	0.31x
Cash Conversion Cycle:	111 days
Working Capital:	\$4820.1M
DSO:	84 days
DPO:	38 days
DIO:	65 days

Overall Liquidity: ADEQUATE

OTHER FINANCIAL RATIOS

Operating CF/Sales:	0.11x
Debt-to-Capital:	0.50x
FCF Yield (Enterprise):	0.13x
Return on Tangible Assets:	11.1%
Cash Return on Assets:	8.7%
Net Debt/EBITDA:	3.78x
EV/FCF:	7.91x
Operating CF/Debt:	0.17x
EBITDA/Interest:	10.97x
FCF Margin:	8.1%
Debt-to-Assets:	0.50x
Cash Conversion Efficiency:	0.65x

LEVERAGE & CAPITAL STRUCTURE

Debt-to-Equity:	1.01x
Debt/EBITDA:	4.08x
Financial Leverage:	2.01x
Interest Coverage:	8.65x
Fixed Charge Coverage:	8.65x
Breakeven Ratio:	0.52x

Overall Leverage: AGGRESSIVE

OPERATIONAL EFFICIENCY

Asset Turnover:	0.78x
Inventory Turnover:	5.63x
Receivables Turnover:	4.33x
Payables Turnover:	9.62x
Fixed Asset Turnover:	3.64x
Working Capital Turnover:	4.18x
Operating Leverage:	N/A
Cash Conversion Cycle:	111 days

Overall Efficiency: HIGH

CASH FLOW ANALYSIS

Operating CF:	+380M
Free CF:	+277M
Investing CF:	-48M
Financing CF:	-185M
CapEx:	+103M
Cash Cycle:	111 days
FCF Yield:	N/A
Cash Flow Margin:	11.1%

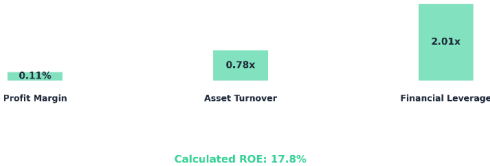
Overall Cash Flow: STRONG

FINANCIAL DISTRESS INDICATORS

Altman Z-Score:	2.42	GREEN
Piotroski Score:	5/9	MODERATE
Beneish M-Score:	-2.80	LOW RISK
Ohlson O-Score:	-1.41	LOW RISK
Zmijewski Score:	-1.88	MODERATE
Springate Score:	6.52	SAFE
Distance to Default:	2.47	MODERATE
Debt Service Coverage:	1.09x	ADEQUATE

ROE BREAKDOWN

ROE = Profit Margin × Asset Turnover × Financial Leverage



VALUATION MULTIPLES

EV/EBITDA:	4.1x
P/E Ratio:	5.6x
P/B Ratio:	1.0x
P/W/Balance:	0.6x
EV/Sales:	0.6x
PEG Ratio:	N/A
Price/Cash Flow:	5.7x
Dividend Yield:	N/A

CAPITAL ALLOCATION & RETURNS

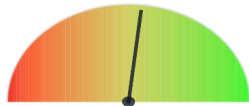
CapEx (annual):	\$103M
Free Cash Flow:	\$277M
EV/FCF (proxy):	7.91x
ROIC:	10.9%
Net Debt/EBITDA:	4.08x
Cash Runway (mo):	N/A
Near-term Funding Gap:	\$-174M

DSM Financial Analysis Service

RISK MANAGEMENT

Report Generated: December 06, 2025 at 18:59

RISK SCORE



5.3

STRENGTHS

8.7

- 9 Low financial leverage
- 9 High operational resilience
- 8 Robust backlog, de-risked

WEAKNESSES

3.3

- 3 Significant working capital inefficiencies
- 4 Excessive capital allocation to buffer stock
- 3 Material geographic and segment concentration risk

KEYWORD CLOUD

Conversion
Defense
Leverage
Aftermarket
Breakeven
Inventory
Industrial
Supply
Backlog
China
Aerospace
Safran

OPPORTUNITIES

7.7

- 8 Global defense rearmament cycle
- 8 Commercial aircraft fleet aging
- 7 Strategic Safran acquisition

THREATS

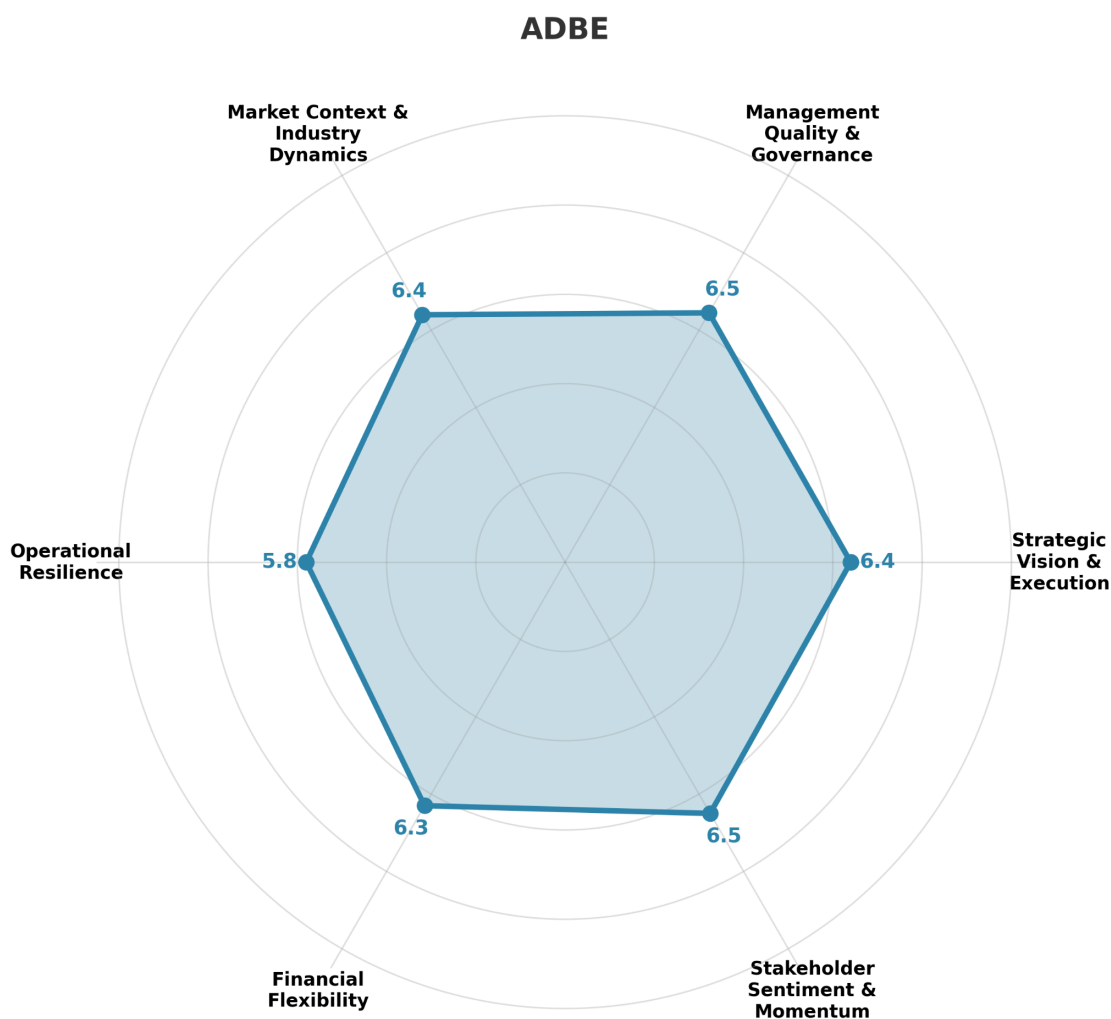
5.0

- 4 Persistent supply chain disruptions
- 5 Boeing 737 MAX production caps
- 6 Intensifying competition from larger conglomerates

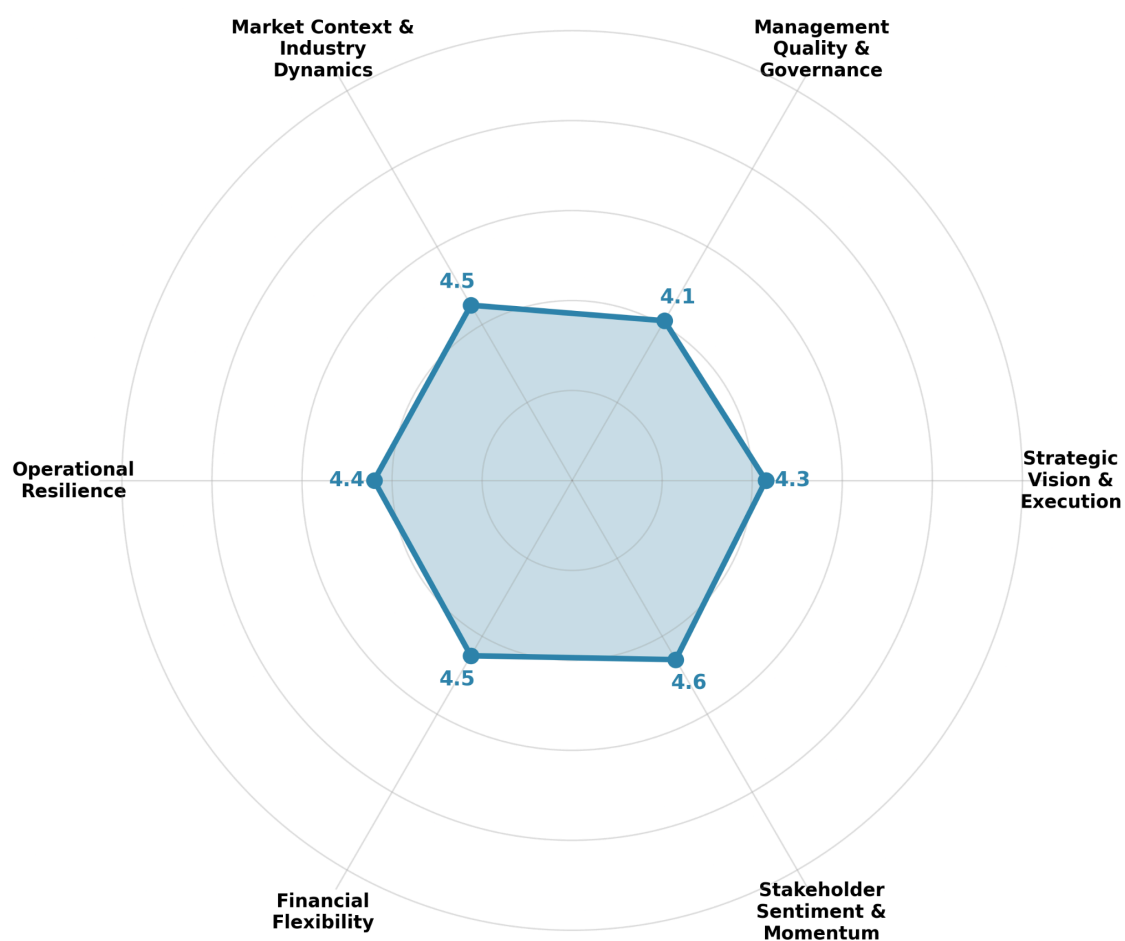
DSM Financial Analysis Service

Anexo 6: Gráficos de Araña de Empresa:

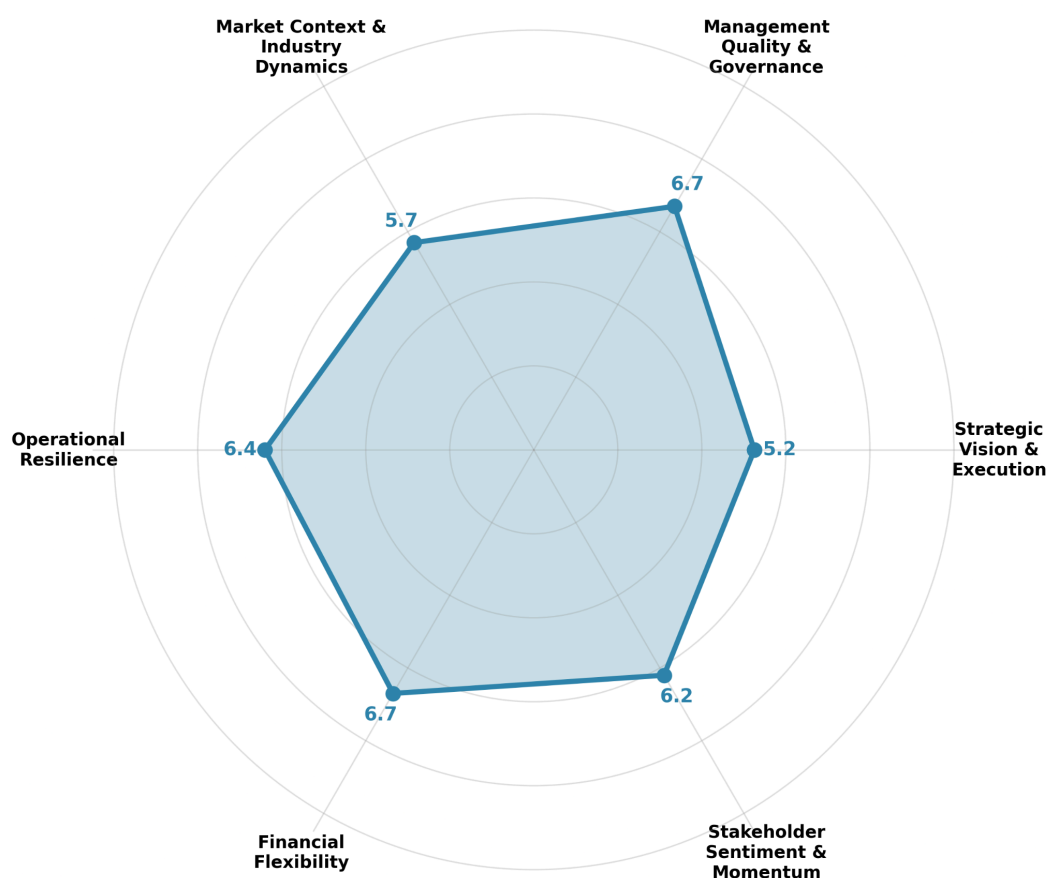
(Puntuaciones aproximadas discrecionalmente; no siguen exactamente el patrón cualitativo porque son aproximaciones incompletas y/o calculadas de las interrelaciones entre las puntuaciones de los ítems en los mandos de control).



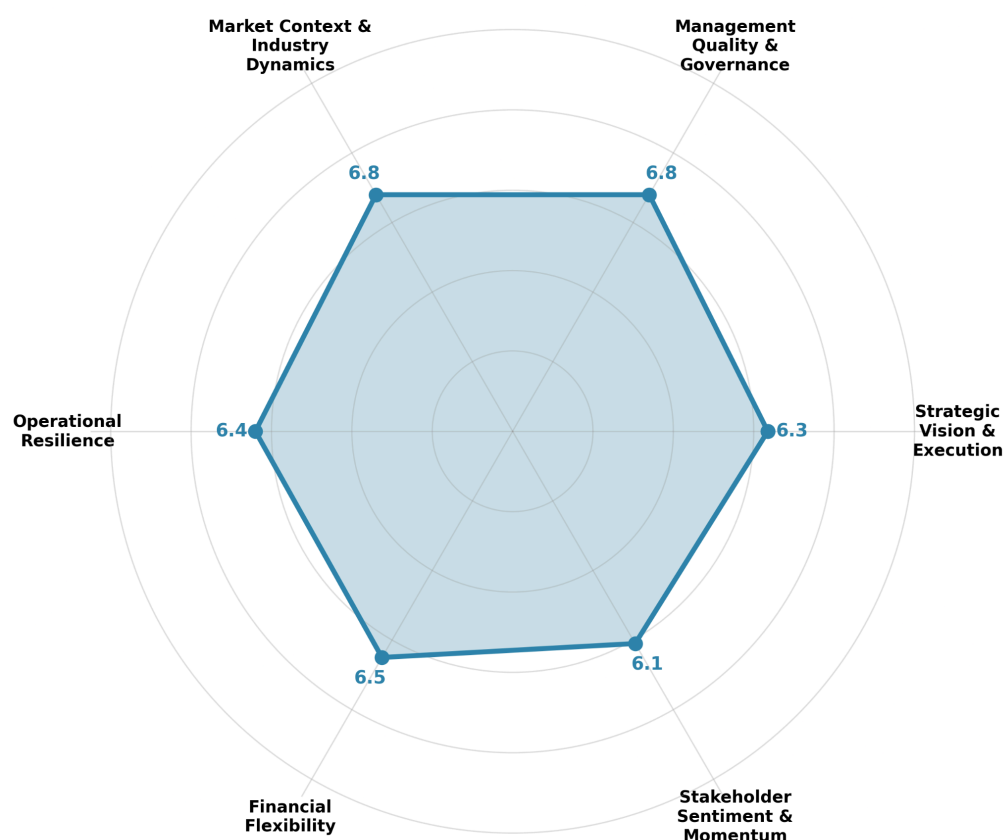
BRKR



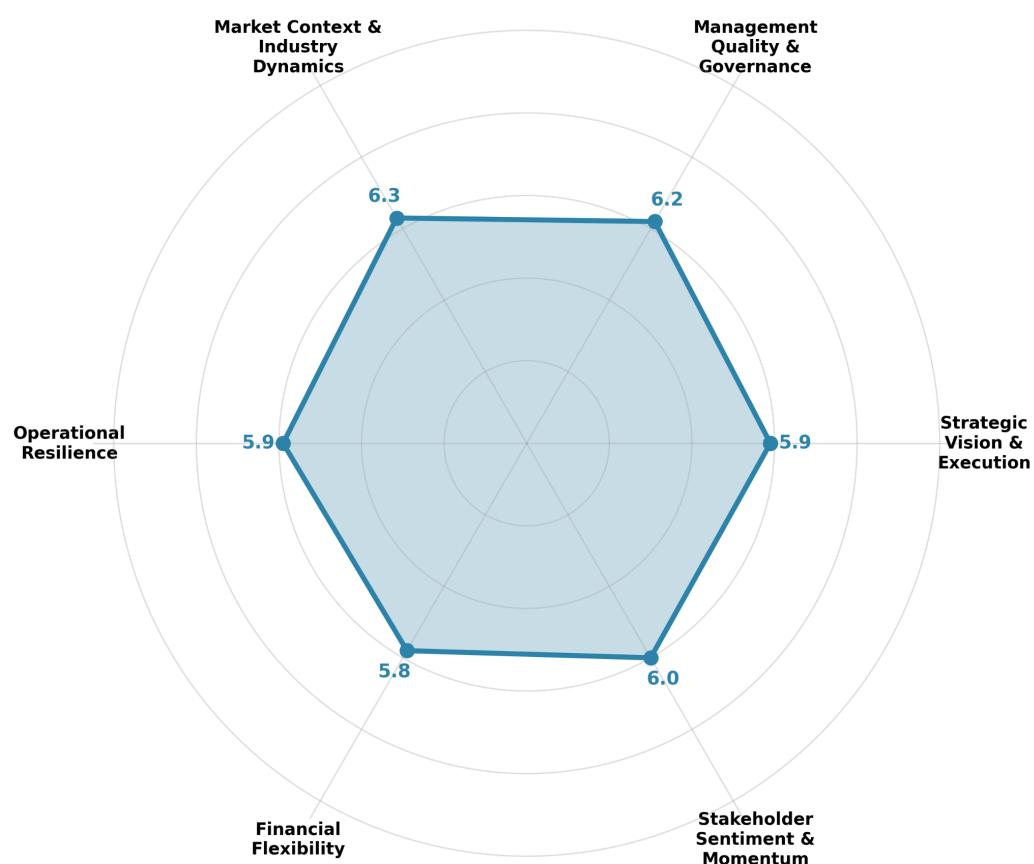
IDXX



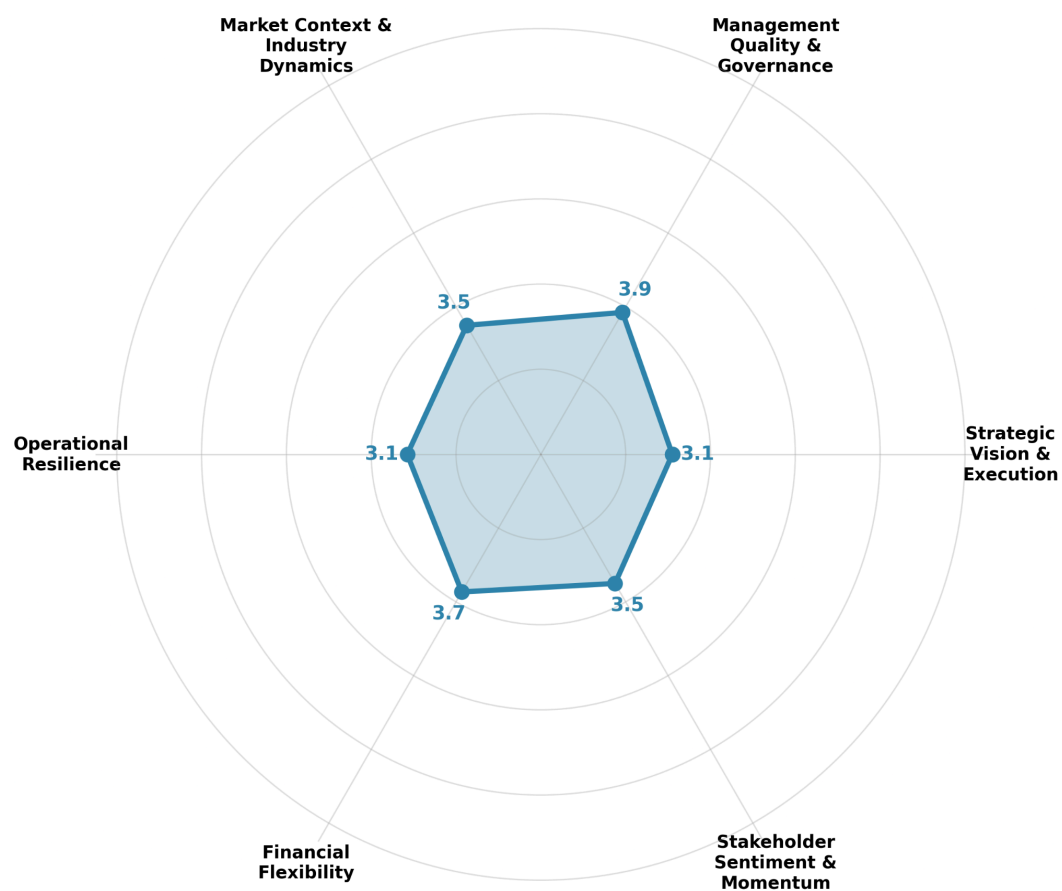
INTU



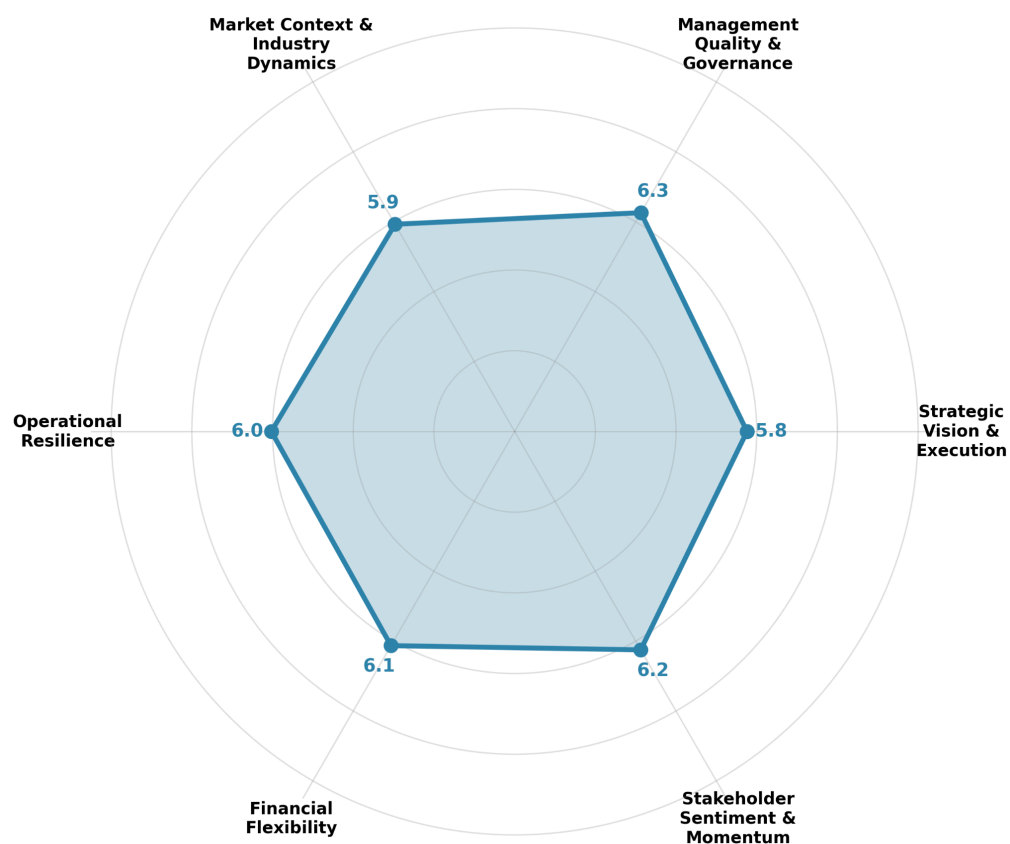
LECo

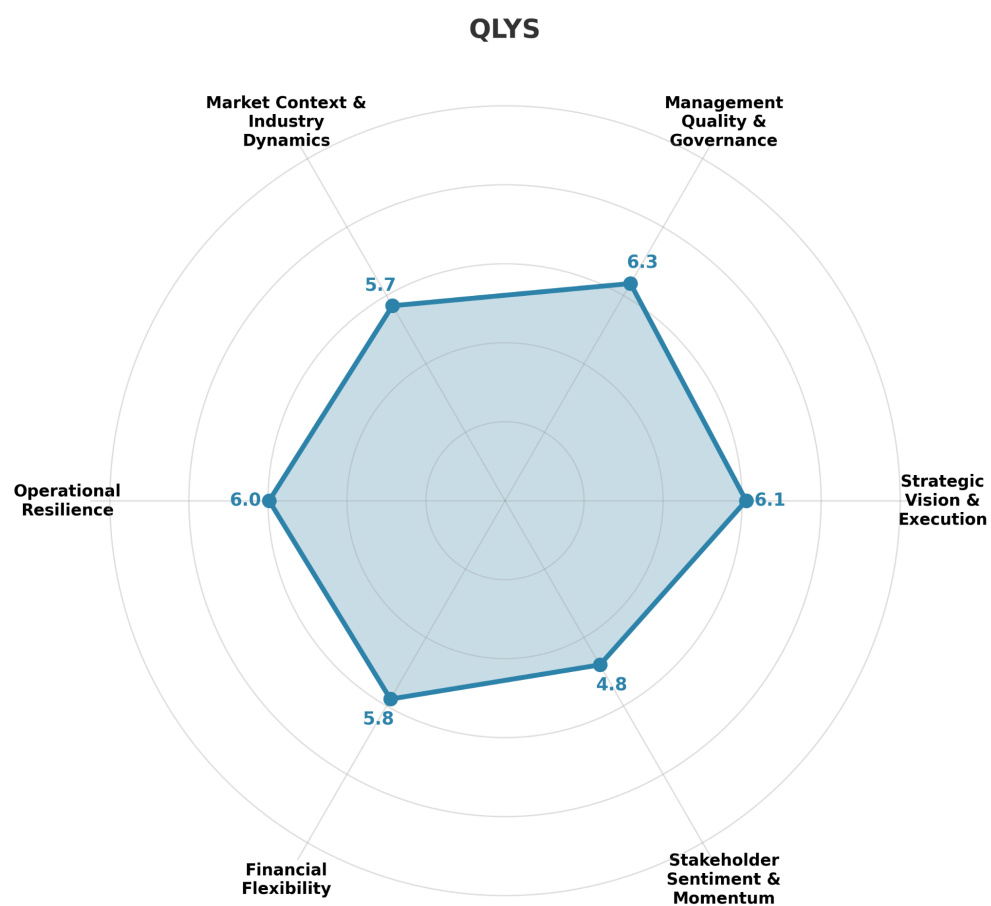


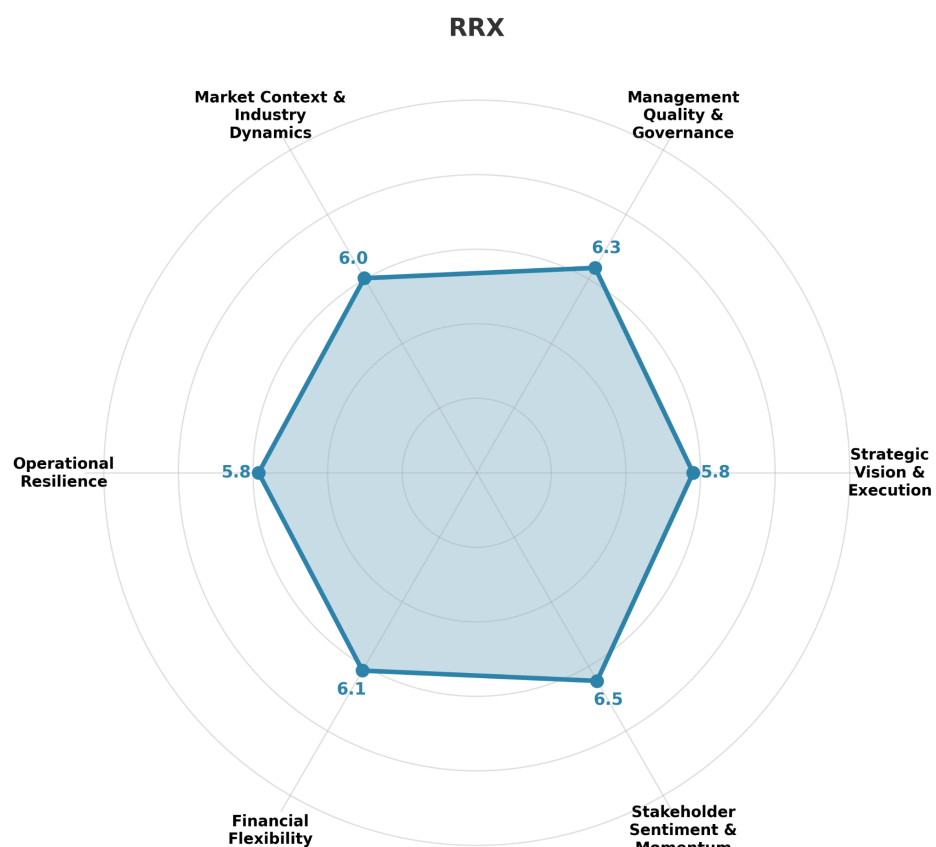
NEOG



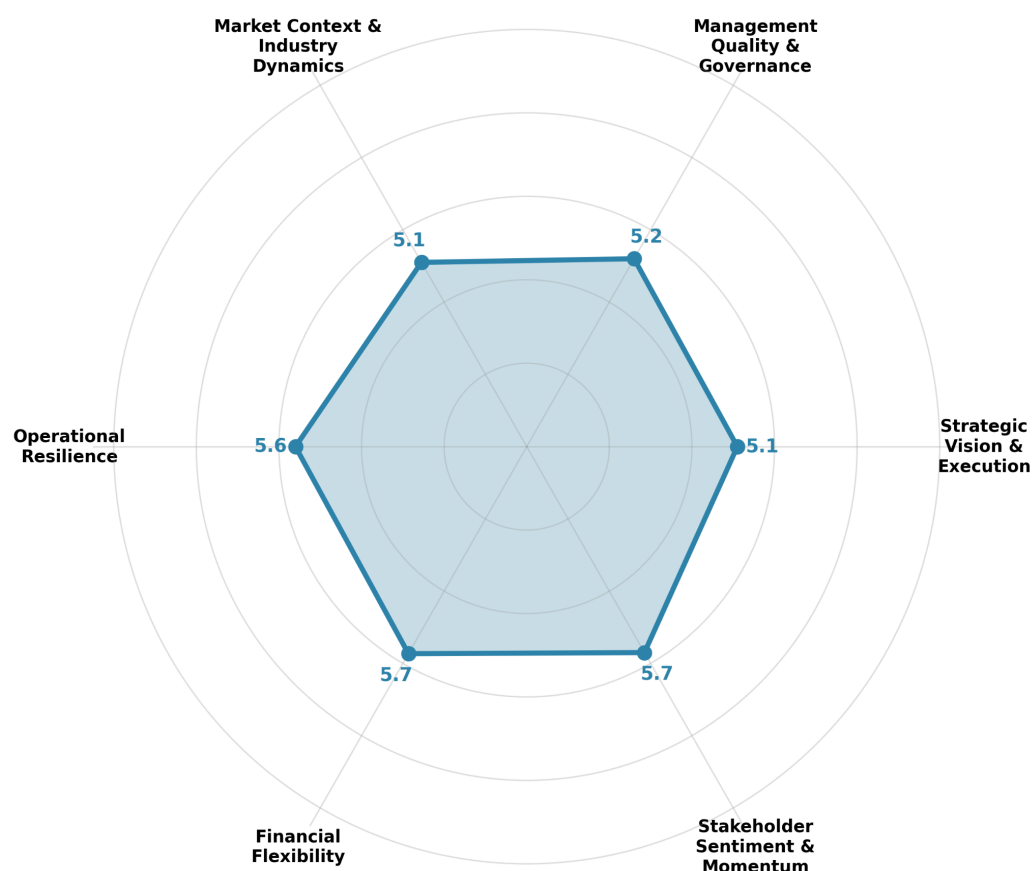
NKE

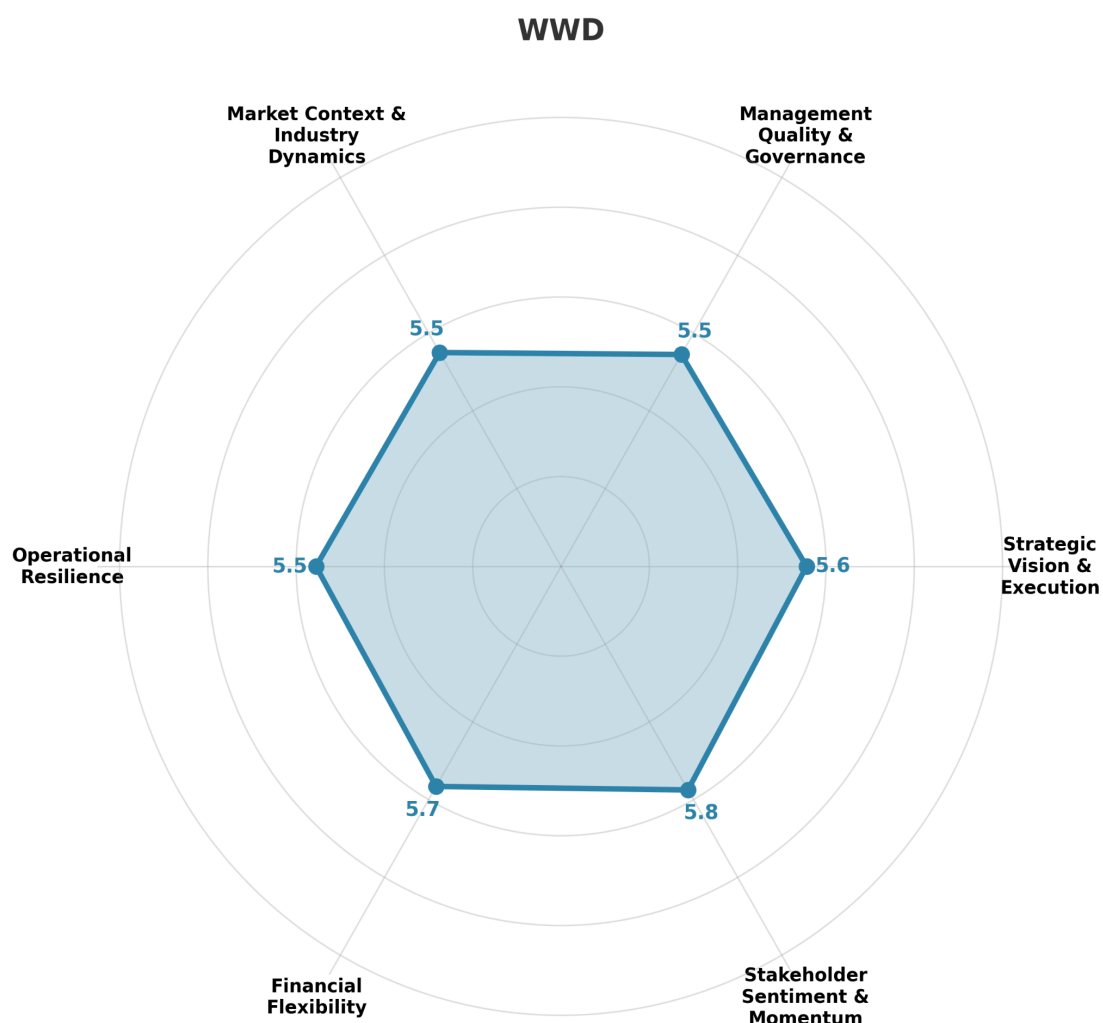






SBUX





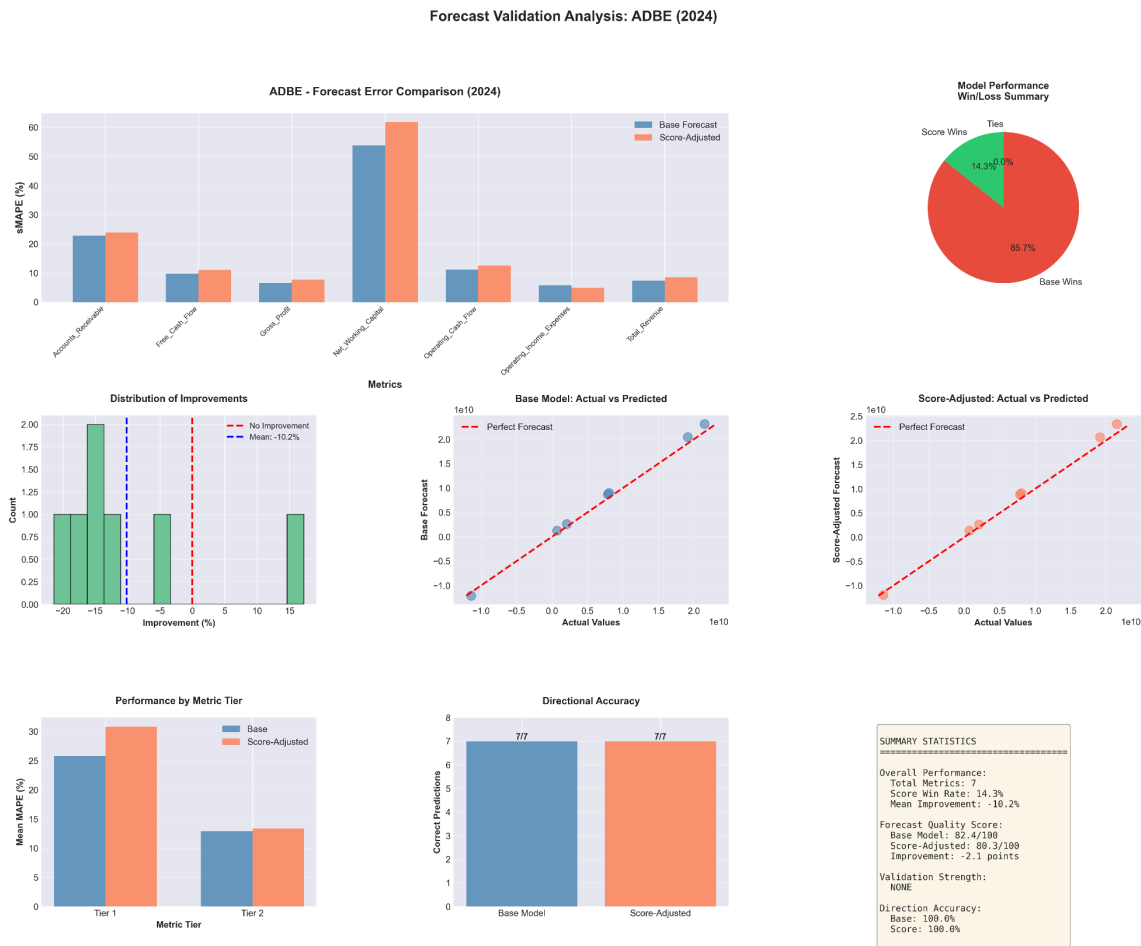
Anexo 7: Gráficos de Pronósticos:

Por voluminosidad, adjuntos en archivo aparte.

Anexo 7: Evaluación de Pronósticos:

Gráficos e informes (generados automáticamente con pre-configuraciones en Python).

ADBE (NAIVE):



FORECAST VALIDATION REPORT

Company: ADBE
Forecast Year: 2024
Analysis Date: 2025-12-11 01:43:14

=====

EXECUTIVE SUMMARY

=====

Validation Strength: NONE

Score-Adjusted method does not validate. Base model performs as well or better.

Overall Win Rate: 14.3% (Score-Adjusted wins)

Mean Improvement: -10.2%

Forecast Quality Scores:

- Base Model: 82.4/100
- Score-Adjusted Model: 80.3/100
- Improvement: -2.1 points

=====

DETAILED METRICS COMPARISON

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 20.28%
- Median MAPE: 10.20%
- Mean sMAPE: 16.76%

Score-Adjusted Model:

- Mean MAPE: 23.34%
- Median MAPE: 11.74%
- Mean sMAPE: 18.61%

Win/Loss Record:

- Score-Adjusted Wins: 1
- Base Model Wins: 6
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%


```
=====
=====
PERFORMANCE BY METRIC TIER
=====
=====
```

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 25.83%
- Score Mean MAPE: 30.83%
- Mean Improvement: -16.24%
- Score Win Rate: 0.0%
- Wins: Score=0, Base=4

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 12.86%
- Score Mean MAPE: 13.36%
- Mean Improvement: -2.05%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

```
=====
=====
STATISTICAL SIGNIFICANCE TESTS
=====
=====
```

Paired T-Test:

- T-Statistic: -1.7408
- P-Value: 0.1324
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 1.0000
- P-Value: 0.0312
- Significant at $\alpha=0.05$: True

Effect Size:

- Cohen's d: -0.6580
- Interpretation: Medium effect

```
=====
=====
```


METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 2,072,000,000.00

Base Forecast: 2,607,438,641.00 (Error: +535,438,641.00)

Score Forecast: 2,634,279,346.00 (Error: +562,279,346.00)

Accuracy:

- Base sMAPE: 22.88%
- Score sMAPE: 23.89%
- Improvement: -5.01%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 7,873,000,000.00

Base Forecast: 8,675,753,929.00 (Error: +802,753,929.00)

Score Forecast: 8,797,116,704.00 (Error: +924,116,704.00)

Accuracy:

- Base sMAPE: 9.70%
- Score sMAPE: 11.09%
- Improvement: -15.12%

Direction Correct: Base=True, Score=True

Winner: Base

Gross_Profit (Tier 2):

Actual Value: 19,147,000,000.00

Base Forecast: 20,436,507,451.00 (Error: +1,289,507,451.00)

Score Forecast: 20,673,212,973.00 (Error: +1,526,212,973.00)

Accuracy:

- Base sMAPE: 6.52%
- Score sMAPE: 7.67%
- Improvement: -18.36%

Direction Correct: Base=True, Score=True

Winner: Base

Net_Working_Capital (Tier 1):

Actual Value: 711,000,000.00

Base Forecast: 1,234,633,279.00 (Error: +523,633,279.00)

Score Forecast: 1,346,518,950.00 (Error: +635,518,950.00)

Accuracy:

- Base sMAPE: 53.83%
- Score sMAPE: 61.78%
- Improvement: -21.37%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 8,056,000,000.00

Base Forecast: 9,013,522,101.00 (Error: +957,522,101.00)

Score Forecast: 9,133,328,648.00 (Error: +1,077,328,648.00)

Accuracy:

- Base sMAPE: 11.22%
- Score sMAPE: 12.53%
- Improvement: -12.51%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Income_Expenses (Tier 2):

Actual Value: -11,406,000,000.00

Base Forecast: -12,091,924,179.00 (Error: -685,924,179.00)

Score Forecast: -11,973,839,486.00 (Error: -567,839,486.00)

Accuracy:

- Base sMAPE: 5.84%
- Score sMAPE: 4.86%
- Improvement: +17.22%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 21,505,000,000.00

Base Forecast: 23,140,844,533.00 (Error: +1,635,844,533.00)

Score Forecast: 23,402,073,650.00 (Error: +1,897,073,650.00)

Accuracy:

- Base sMAPE: 7.33%
- Score sMAPE: 8.45%
- Improvement: -15.97%

Direction Correct: Base=True, Score=True

Winner: Base

```
=====
=====
RECOMMENDATIONS
=====
=====
```

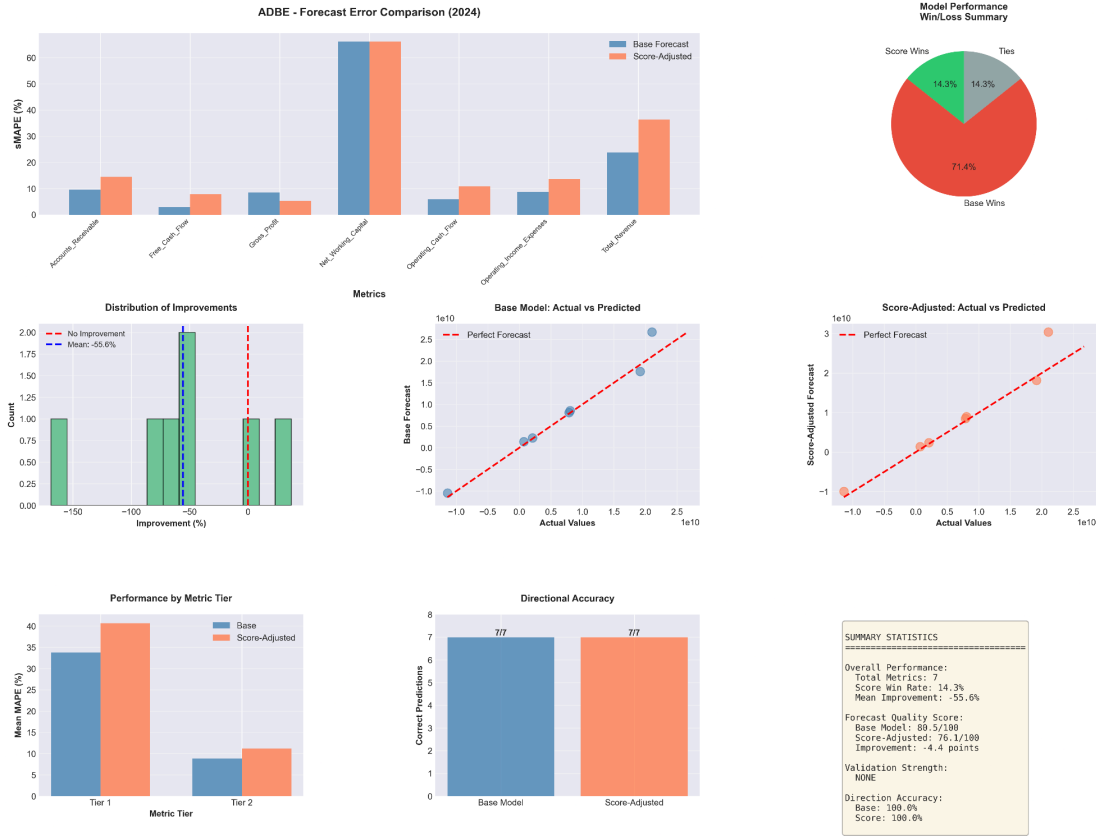
X No validation. Base model performs as well or better. Recommend:

1. Critically review the qualitative scoring framework
2. Check if qualitative factors are already reflected in historical data
3. Consider that market efficiency may limit qualitative alpha
4. Explore if timing or sector-specific factors affected results

```
=====
=====
END OF REPORT
=====
=====
```


ADBE (LINEAR):

Forecast Validation Analysis: ADBE (2024)



FORECAST VALIDATION REPORT

Company: ADBE
Forecast Year: 2024
Analysis Date: 2025-12-11 01:45:46

EXECUTIVE SUMMARY

Validation Strength: NONE

Score-Adjusted method does not validate. Base model performs as well or better.

Overall Win Rate: 14.3% (Score-Adjusted wins)

Mean Improvement: -55.6%

Forecast Quality Scores:

- Base Model: 80.5/100
- Score-Adjusted Model: 76.1/100
- Improvement: -4.4 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 23.08%
- Median MAPE: 8.32%
- Mean sMAPE: 17.96%

Score-Adjusted Model:

- Mean MAPE: 28.05%
- Median MAPE: 12.85%
- Mean sMAPE: 22.08%

Win/Loss Record:

- Score-Adjusted Wins: 1
- Base Model Wins: 5
- Ties: 1

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 33.75%
- Score Mean MAPE: 40.70%
- Mean Improvement: -79.49%
- Score Win Rate: 0.0%
- Wins: Score=0, Base=3

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 8.86%
- Score Mean MAPE: 11.18%
- Mean Improvement: -23.66%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

=====

=====

STATISTICAL SIGNIFICANCE TESTS

=====

=====

Paired T-Test:

- T-Statistic: -2.2154
- P-Value: 0.0686
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 1.0000
- P-Value: 0.0464
- Significant at $\alpha=0.05$: True

Effect Size:

- Cohen's d: -0.8373
- Interpretation: Large effect

=====

=====

METRIC-BY-METRIC BREAKDOWN

=====

=====

Accounts_Receivable (Tier 2):

Actual Value: 2,072,000,000.00

Base Forecast: 2,281,917,548.00 (Error: +209,917,548.00)
Score Forecast: 2,394,671,658.00 (Error: +322,671,658.00)

Accuracy:

- Base sMAPE: 9.64%
- Score sMAPE: 14.45%
- Improvement: -53.71%

Direction Correct: Base=True, Score=True
Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 7,873,000,000.00
Base Forecast: 8,110,425,922.00 (Error: +237,425,922.00)
Score Forecast: 8,511,178,287.00 (Error: +638,178,287.00)

Accuracy:

- Base sMAPE: 2.97%
- Score sMAPE: 7.79%
- Improvement: -168.79%

Direction Correct: Base=True, Score=True
Winner: Base

Gross_Profit (Tier 2):

Actual Value: 19,147,000,000.00
Base Forecast: 17,588,197,795.00 (Error: -1,558,802,205.00)
Score Forecast: 18,167,834,442.00 (Error: -979,165,558.00)

Accuracy:

- Base sMAPE: 8.49%
- Score sMAPE: 5.25%
- Improvement: +37.18%

Direction Correct: Base=True, Score=True
Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 711,000,000.00
Base Forecast: 1,412,958,750.00 (Error: +701,958,750.00)
Score Forecast: 1,412,958,750.00 (Error: +701,958,750.00)

Accuracy:

- Base sMAPE: 66.10%
- Score sMAPE: 66.10%
- Improvement: +0.00%

Direction Correct: Base=True, Score=True

Winner: Tie

Operating_Cash_Flow (Tier 1):

Actual Value: 8,056,000,000.00

Base Forecast: 8,556,628,730.00 (Error: +500,628,730.00)

Score Forecast: 8,979,428,869.00 (Error: +923,428,869.00)

Accuracy:

- Base sMAPE: 6.03%
- Score sMAPE: 10.84%
- Improvement: -84.45%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Income_Expenses (Tier 2):

Actual Value: -11,406,000,000.00

Base Forecast: -10,457,076,602.00 (Error: +948,923,398.00)

Score Forecast: -9,940,371,533.00 (Error: +1,465,628,467.00)

Accuracy:

- Base sMAPE: 8.68%
- Score sMAPE: 13.73%
- Improvement: -54.45%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 21,013,972,474.00

Base Forecast: 26,693,111,782.00 (Error: +5,679,139,308.00)

Score Forecast: 30,369,112,487.00 (Error: +9,355,140,013.00)

Accuracy:

- Base sMAPE: 23.81%
- Score sMAPE: 36.41%

- Improvement: -64.73%

Direction Correct: Base=True, Score=True

Winner: Base

=====

=====

RECOMMENDATIONS

=====

=====

X No validation. Base model performs as well or better. Recommend:

1. Critically review the qualitative scoring framework
2. Check if qualitative factors are already reflected in historical data
3. Consider that market efficiency may limit qualitative alpha
4. Explore if timing or sector-specific factors affected results

=====

=====

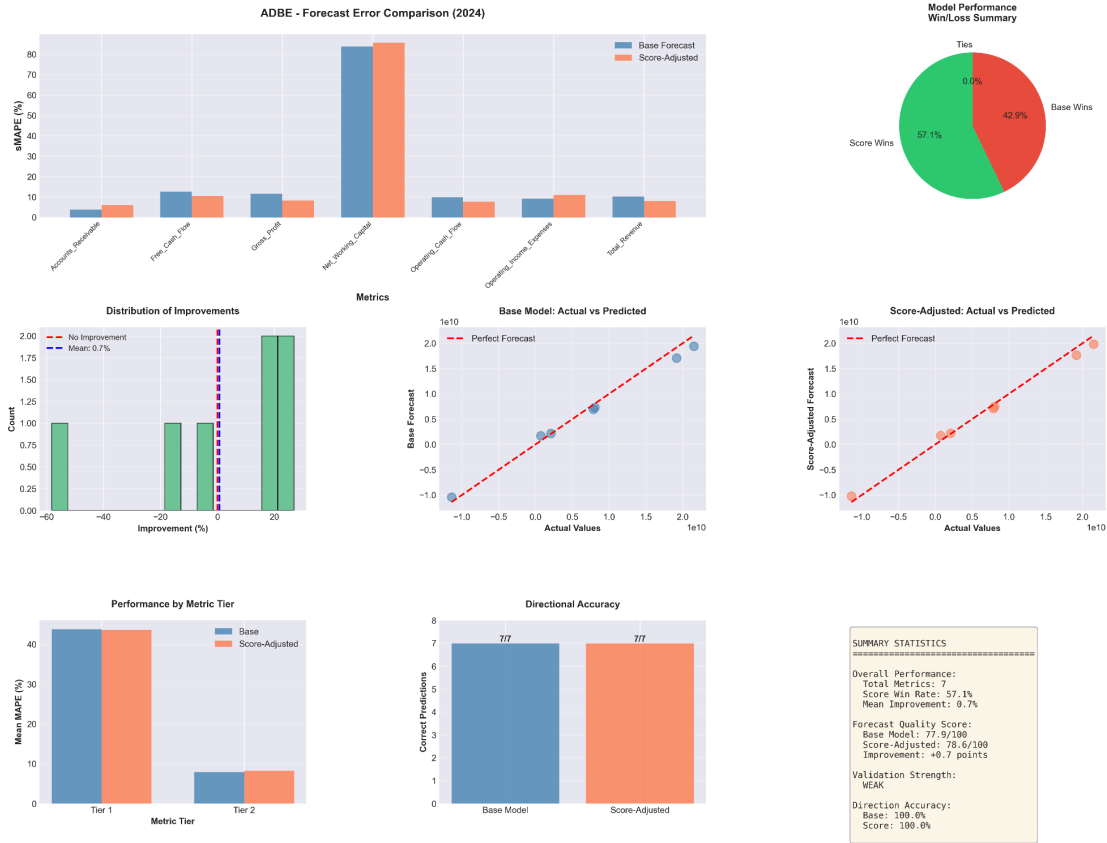
END OF REPORT

=====

=====

ADBE (XGBOOST):

Forecast Validation Analysis: ADBE (2024)



FORECAST VALIDATION REPORT

Company: ADBE
Forecast Year: 2024
Analysis Date: 2025-12-11 01:40:41

EXECUTIVE SUMMARY

Validation Strength: WEAK

Score-Adjusted method shows weak validation. Qualitative factors provide marginal improvements.

Overall Win Rate: 57.1% (Score-Adjusted wins)

Mean Improvement: 0.7%

Forecast Quality Scores:

- Base Model: 77.9/100
- Score-Adjusted Model: 78.6/100
- Improvement: +0.7 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 28.41%
- Median MAPE: 9.75%
- Mean sMAPE: 20.15%

Score-Adjusted Model:

- Mean MAPE: 28.47%
- Median MAPE: 7.99%
- Mean sMAPE: 19.57%

Win/Loss Record:

- Score-Adjusted Wins: 4
- Base Model Wins: 3
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 43.81%
- Score Mean MAPE: 43.67%
- Mean Improvement: 13.59%
- Score Win Rate: 75.0%
- Wins: Score=3, Base=1

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 7.88%
- Score Mean MAPE: 8.19%
- Mean Improvement: -16.54%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

STATISTICAL SIGNIFICANCE TESTS

Paired T-Test:

- T-Statistic: 0.6442
- P-Value: 0.5433
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 9.0000
- P-Value: 0.4688
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.2435
- Interpretation: Small effect

METRIC-BY-METRIC BREAKDOWN

Accounts_Receivable (Tier 2):

Actual Value: 2,072,000,000.00

Base Forecast: 2,153,343,232.00 (Error: +81,343,232.00)

Score Forecast: 2,200,716,783.00 (Error: +128,716,783.00)

Accuracy:

- Base sMAPE: 3.85%
- Score sMAPE: 6.03%
- Improvement: -58.24%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 7,873,000,000.00

Base Forecast: 6,941,999,616.00 (Error: -931,000,384.00)

Score Forecast: 7,094,723,607.00 (Error: -778,276,393.00)

Accuracy:

- Base sMAPE: 12.57%
- Score sMAPE: 10.40%
- Improvement: +16.40%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 19,147,000,000.00

Base Forecast: 17,054,996,480.00 (Error: -2,092,003,520.00)

Score Forecast: 17,617,811,363.00 (Error: -1,529,188,637.00)

Accuracy:

- Base sMAPE: 11.56%
- Score sMAPE: 8.32%
- Improvement: +26.90%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 711,000,000.00

Base Forecast: 1,736,999,808.00 (Error: +1,025,999,808.00)

Score Forecast: 1,775,213,803.00 (Error: +1,064,213,803.00)

Accuracy:

- Base sMAPE: 83.82%
- Score sMAPE: 85.61%
- Improvement: -3.72%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 8,056,000,000.00

Base Forecast: 7,301,999,616.00 (Error: -754,000,384.00)

Score Forecast: 7,462,643,607.00 (Error: -593,356,393.00)

Accuracy:

- Base sMAPE: 9.82%
- Score sMAPE: 7.65%
- Improvement: +21.31%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -11,406,000,000.00

Base Forecast: -10,404,997,120.00 (Error: +1,001,002,880.00)

Score Forecast: -10,221,869,170.00 (Error: +1,184,130,830.00)

Accuracy:

- Base sMAPE: 9.18%
- Score sMAPE: 10.95%
- Improvement: -18.29%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 21,505,000,000.00

Base Forecast: 19,408,994,304.00 (Error: -2,096,005,696.00)

Score Forecast: 19,835,992,178.00 (Error: -1,669,007,822.00)

Accuracy:

- Base sMAPE: 10.25%
- Score sMAPE: 8.07%
- Improvement: +20.37%

Direction Correct: Base=True, Score=True
Winner: Score

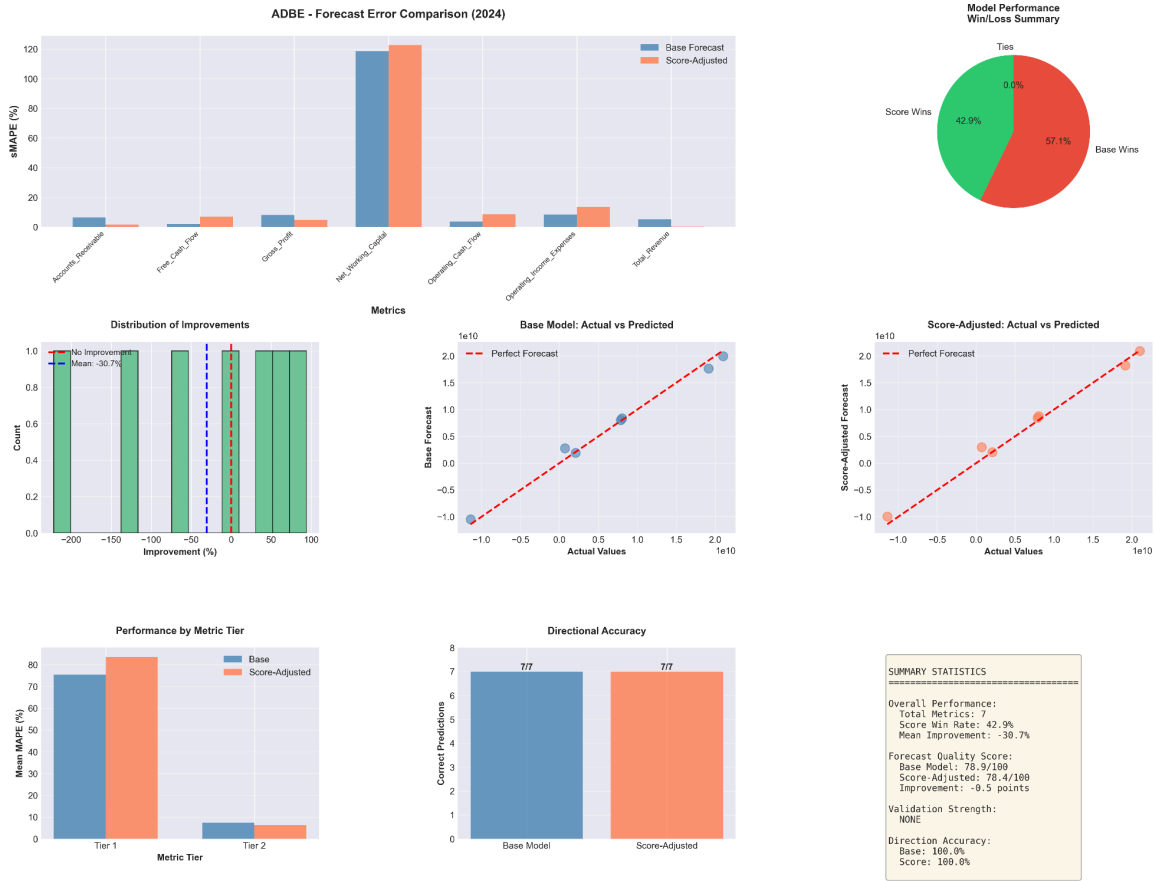
```
=====
=====
RECOMMENDATIONS
=====
=====
```

- △ Weak validation. Qualitative adjustments show marginal impact. Recommend:
1. Review and refine the qualitative scoring methodology
 2. Investigate if certain metric types benefit more from adjustments
 3. Consider whether the base model is already capturing qualitative factors
 4. Test alternative adjustment algorithms

```
=====
=====
END OF REPORT
=====
=====
```


ADBE (ARIMA):

Forecast Validation Analysis: ADBE (2024)



FORECAST VALIDATION REPORT

Company: ADBE
Forecast Year: 2024
Analysis Date: 2025-12-11 01:48:07

EXECUTIVE SUMMARY

Validation Strength: NONE

Score-Adjusted method does not validate. Base model performs as well or better.

Overall Win Rate: 42.9% (Score-Adjusted wins)

Mean Improvement: -30.7%

Forecast Quality Scores:

- Base Model: 78.9/100
- Score-Adjusted Model: 78.4/100
- Improvement: -0.5 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 46.29%
- Median MAPE: 6.44%
- Mean sMAPE: 21.85%

Score-Adjusted Model:

- Mean MAPE: 50.54%
- Median MAPE: 7.33%
- Mean sMAPE: 22.72%

Win/Loss Record:

- Score-Adjusted Wins: 3
- Base Model Wins: 4
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 75.41%
- Score Mean MAPE: 83.62%
- Mean Improvement: -67.27%
- Score Win Rate: 25.0%
- Wins: Score=1, Base=3

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 7.47%
- Score Mean MAPE: 6.43%
- Mean Improvement: 18.10%
- Score Win Rate: 66.7%
- Wins: Score=2, Base=1

STATISTICAL SIGNIFICANCE TESTS

Paired T-Test:

- T-Statistic: -0.4748
- P-Value: 0.6517
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 12.0000
- P-Value: 0.8125
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: -0.1794
- Interpretation: Small effect

METRIC-BY-METRIC BREAKDOWN

Accounts_Receivable (Tier 2):

Actual Value: 2,072,000,000.00

Base Forecast: 1,938,628,341.00 (Error: -133,371,659.00)

Score Forecast: 2,034,419,845.00 (Error: -37,580,155.00)

Accuracy:

- Base sMAPE: 6.65%
- Score sMAPE: 1.83%
- Improvement: +71.82%

Direction Correct: Base=True, Score=True

Winner: Score

Free_Cash_Flow (Tier 1):

Actual Value: 7,873,000,000.00

Base Forecast: 8,052,343,360.00 (Error: +179,343,360.00)

Score Forecast: 8,450,225,750.00 (Error: +577,225,750.00)

Accuracy:

- Base sMAPE: 2.25%
- Score sMAPE: 7.07%
- Improvement: -221.86%

Direction Correct: Base=True, Score=True

Winner: Base

Gross_Profit (Tier 2):

Actual Value: 19,147,000,000.00

Base Forecast: 17,639,383,251.00 (Error: -1,507,616,749.00)

Score Forecast: 18,220,706,765.00 (Error: -926,293,235.00)

Accuracy:

- Base sMAPE: 8.20%
- Score sMAPE: 4.96%
- Improvement: +38.56%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 711,000,000.00

Base Forecast: 2,776,323,728.00 (Error: +2,065,323,728.00)

Score Forecast: 2,970,666,389.00 (Error: +2,259,666,389.00)

Accuracy:

- Base sMAPE: 118.45%
- Score sMAPE: 122.75%
- Improvement: -9.41%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 8,056,000,000.00

Base Forecast: 8,369,586,428.00 (Error: +313,586,428.00)

Score Forecast: 8,783,144,433.00 (Error: +727,144,433.00)

Accuracy:

- Base sMAPE: 3.82%
- Score sMAPE: 8.64%
- Improvement: -131.88%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Income_Expenses (Tier 2):

Actual Value: -11,406,000,000.00

Base Forecast: -10,482,203,871.00 (Error: +923,796,129.00)

Score Forecast: -9,964,257,213.00 (Error: +1,441,742,787.00)

Accuracy:

- Base sMAPE: 8.44%
- Score sMAPE: 13.49%
- Improvement: -56.07%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 21,013,972,474.00

Base Forecast: 19,965,403,064.00 (Error: -1,048,569,410.00)

Score Forecast: 20,951,933,560.00 (Error: -62,038,914.00)

Accuracy:

- Base sMAPE: 5.12%
- Score sMAPE: 0.30%
- Improvement: +94.08%

Direction Correct: Base=True, Score=True
Winner: Score

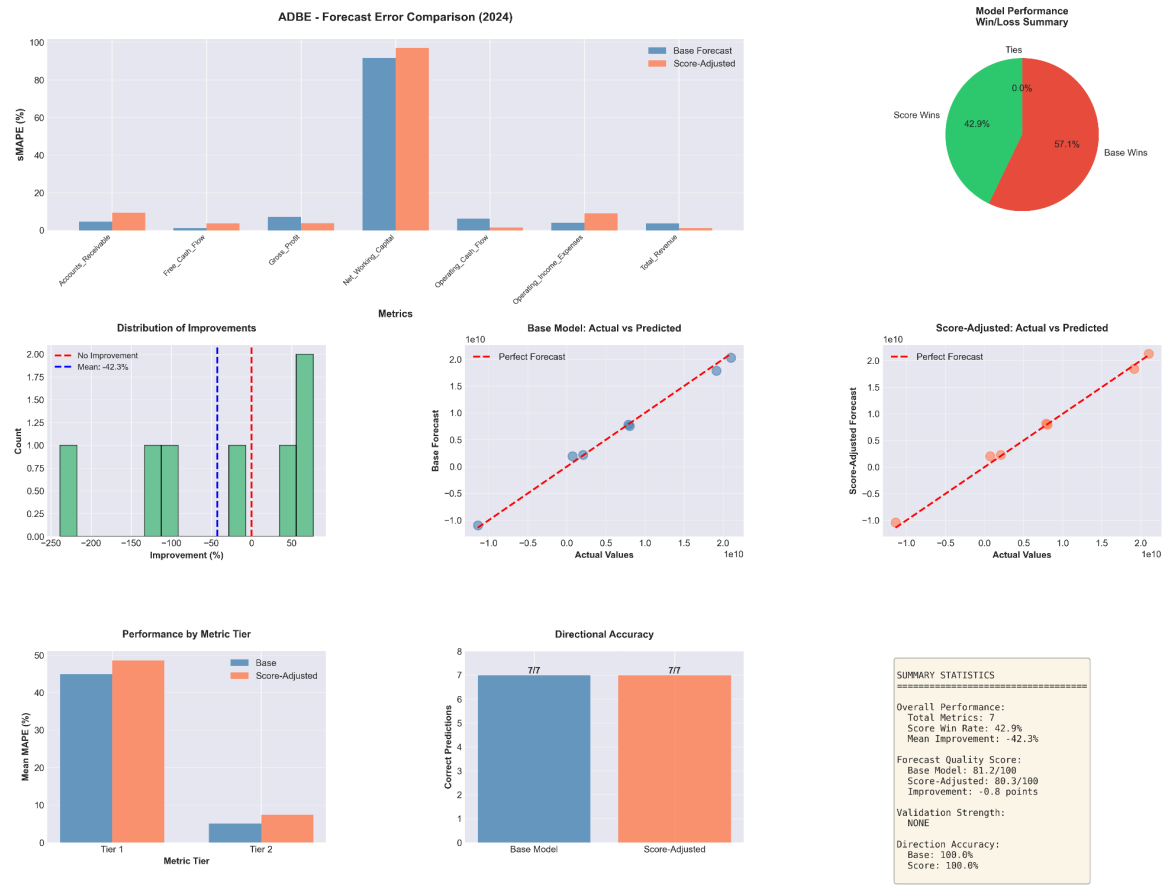
```
=====
=====
RECOMMENDATIONS
=====
=====
```

- X** No validation. Base model performs as well or better. Recommend:
1. Critically review the qualitative scoring framework
 2. Check if qualitative factors are already reflected in historical data
 3. Consider that market efficiency may limit qualitative alpha
 4. Explore if timing or sector-specific factors affected results

```
=====
=====
END OF REPORT
=====
=====
```


ADBE (WINTERS-HOLT):

Forecast Validation Analysis: ADBE (2024)



FORECAST VALIDATION REPORT

Company: ADBE
Forecast Year: 2024
Analysis Date: 2025-12-11 01:50:39

EXECUTIVE SUMMARY

Validation Strength: NONE

Score-Adjusted method does not validate. Base model performs as well or better.

Overall Win Rate: 42.9% (Score-Adjusted wins)

Mean Improvement: -42.3%

Forecast Quality Scores:

- Base Model: 81.2/100
- Score-Adjusted Model: 80.3/100
- Improvement: -0.8 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 27.89%
- Median MAPE: 4.61%
- Mean sMAPE: 16.87%

Score-Adjusted Model:

- Mean MAPE: 30.94%
- Median MAPE: 3.77%
- Mean sMAPE: 17.91%

Win/Loss Record:

- Score-Adjusted Wins: 3
- Base Model Wins: 4
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 44.99%
- Score Mean MAPE: 48.62%
- Mean Improvement: -26.83%
- Score Win Rate: 50.0%
- Wins: Score=2, Base=2

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 5.09%
- Score Mean MAPE: 7.37%
- Mean Improvement: -63.03%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

STATISTICAL SIGNIFICANCE TESTS

Paired T-Test:

- T-Statistic: -0.6272
- P-Value: 0.5537
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 9.0000
- P-Value: 0.4688
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: -0.2370
- Interpretation: Small effect

METRIC-BY-METRIC BREAKDOWN

Accounts_Receivable (Tier 2):

Actual Value: 2,072,000,000.00

Base Forecast: 2,167,473,704.00 (Error: +95,473,704.00)

Score Forecast: 2,274,572,915.00 (Error: +202,572,915.00)

Accuracy:

- Base sMAPE: 4.50%
- Score sMAPE: 9.32%
- Improvement: -112.18%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 7,873,000,000.00

Base Forecast: 7,785,306,672.00 (Error: -87,693,328.00)

Score Forecast: 8,169,994,245.00 (Error: +296,994,245.00)

Accuracy:

- Base sMAPE: 1.12%
- Score sMAPE: 3.70%
- Improvement: -238.67%

Direction Correct: Base=True, Score=True

Winner: Base

Gross_Profit (Tier 2):

Actual Value: 19,147,000,000.00

Base Forecast: 17,849,624,986.00 (Error: -1,297,375,014.00)

Score Forecast: 18,437,877,227.00 (Error: -709,122,773.00)

Accuracy:

- Base sMAPE: 7.01%
- Score sMAPE: 3.77%
- Improvement: +45.34%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 711,000,000.00

Base Forecast: 1,914,366,744.00 (Error: +1,203,366,744.00)

Score Forecast: 2,048,372,416.00 (Error: +1,337,372,416.00)

Accuracy:

- Base sMAPE: 91.67%
- Score sMAPE: 96.93%
- Improvement: -11.14%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 8,056,000,000.00

Base Forecast: 7,569,961,846.00 (Error: -486,038,154.00)

Score Forecast: 7,944,008,800.00 (Error: -111,991,200.00)

Accuracy:

- Base sMAPE: 6.22%
- Score sMAPE: 1.40%
- Improvement: +76.96%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -11,406,000,000.00

Base Forecast: -10,962,931,846.00 (Error: +443,068,154.00)

Score Forecast: -10,421,231,457.00 (Error: +984,768,543.00)

Accuracy:

- Base sMAPE: 3.96%
- Score sMAPE: 9.02%
- Improvement: -122.26%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 21,013,972,474.00

Base Forecast: 20,269,210,793.00 (Error: -744,761,681.00)

Score Forecast: 21,270,753,036.00 (Error: +256,780,562.00)

Accuracy:

- Base sMAPE: 3.61%
- Score sMAPE: 1.21%
- Improvement: +65.52%

Direction Correct: Base=True, Score=True
Winner: Score

=====

=====

RECOMMENDATIONS

=====

=====

- X No validation. Base model performs as well or better. Recommend:
1. Critically review the qualitative scoring framework
 2. Check if qualitative factors are already reflected in historical data
 3. Consider that market efficiency may limit qualitative alpha
 4. Explore if timing or sector-specific factors affected results

=====

=====

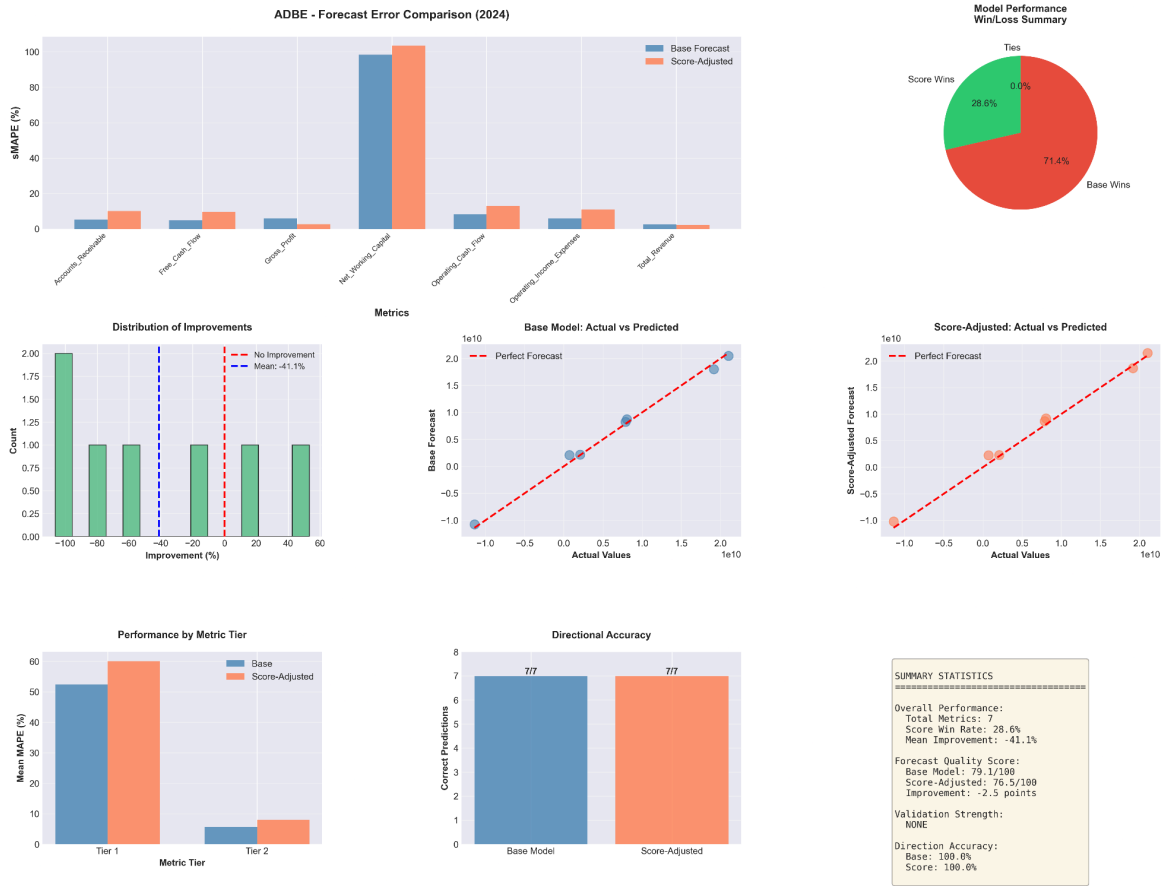
END OF REPORT

=====

=====

ADBE (LINEAR + WINTERS-HOLT):

Forecast Validation Analysis: ADBE (2024)



FORECAST VALIDATION REPORT

Company: ADBE
Forecast Year: 2024
Analysis Date: 2025-12-11 01:54:54

EXECUTIVE SUMMARY

=====

Validation Strength: NONE

Score-Adjusted method does not validate. Base model performs as well or better.

Overall Win Rate: 28.6% (Score-Adjusted wins)

Mean Improvement: -41.1%

Forecast Quality Scores:

- Base Model: 79.1/100
- Score-Adjusted Model: 76.5/100
- Improvement: -2.5 points

=====

DETAILED METRICS COMPARISON

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 32.38%
- Median MAPE: 5.76%
- Mean sMAPE: 18.75%

Score-Adjusted Model:

- Mean MAPE: 37.74%
- Median MAPE: 10.42%
- Mean sMAPE: 21.74%

Win/Loss Record:

- Score-Adjusted Wins: 2
- Base Model Wins: 5
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 52.41%
- Score Mean MAPE: 60.09%
- Mean Improvement: -41.18%
- Score Win Rate: 25.0%
- Wins: Score=1, Base=3

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 5.68%
- Score Mean MAPE: 7.93%
- Mean Improvement: -41.06%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: -2.3283
- P-Value: 0.0588
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 3.0000
- P-Value: 0.0781
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: -0.8800
- Interpretation: Large effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 2,072,000,000.00

Base Forecast: 2,184,843,712.00 (Error: +112,843,712.00)

Score Forecast: 2,292,801,210.00 (Error: +220,801,210.00)

Accuracy:

- Base sMAPE: 5.30%
- Score sMAPE: 10.12%
- Improvement: -95.67%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 7,873,000,000.00

Base Forecast: 8,257,021,282.00 (Error: +384,021,282.00)

Score Forecast: 8,665,017,217.00 (Error: +792,017,217.00)

Accuracy:

- Base sMAPE: 4.76%
- Score sMAPE: 9.58%
- Improvement: -106.24%

Direction Correct: Base=True, Score=True

Winner: Base

Gross_Profit (Tier 2):

Actual Value: 19,147,000,000.00

Base Forecast: 18,031,854,435.00 (Error: -1,115,145,565.00)

Score Forecast: 18,626,112,229.00 (Error: -520,887,771.00)

Accuracy:

- Base sMAPE: 6.00%
- Score sMAPE: 2.76%
- Improvement: +53.29%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 711,000,000.00
Base Forecast: 2,086,652,647.00 (Error: +1,375,652,647.00)
Score Forecast: 2,232,718,333.00 (Error: +1,521,718,333.00)

Accuracy:

- Base sMAPE: 98.34%
- Score sMAPE: 103.39%
- Improvement: -10.62%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 8,056,000,000.00
Base Forecast: 8,754,965,413.00 (Error: +698,965,413.00)
Score Forecast: 9,187,565,764.00 (Error: +1,131,565,764.00)

Accuracy:

- Base sMAPE: 8.32%
- Score sMAPE: 13.12%
- Improvement: -61.89%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Income_Expenses (Tier 2):

Actual Value: -11,406,000,000.00
Base Forecast: -10,748,704,739.00 (Error: +657,295,261.00)
Score Forecast: -10,217,589,741.00 (Error: +1,188,410,259.00)

Accuracy:

- Base sMAPE: 5.93%
- Score sMAPE: 10.99%
- Improvement: -80.80%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 21,013,972,474.00
Base Forecast: 20,470,042,891.00 (Error: -543,929,583.00)
Score Forecast: 21,481,508,650.00 (Error: +467,536,176.00)

Accuracy:

- Base sMAPE: 2.62%
- Score sMAPE: 2.20%
- Improvement: +14.04%

Direction Correct: Base=True, Score=True

Winner: Score

=====

=====

RECOMMENDATIONS

=====

=====

X No validation. Base model performs as well or better. Recommend:

1. Critically review the qualitative scoring framework
2. Check if qualitative factors are already reflected in historical data
3. Consider that market efficiency may limit qualitative alpha
4. Explore if timing or sector-specific factors affected results

=====

=====

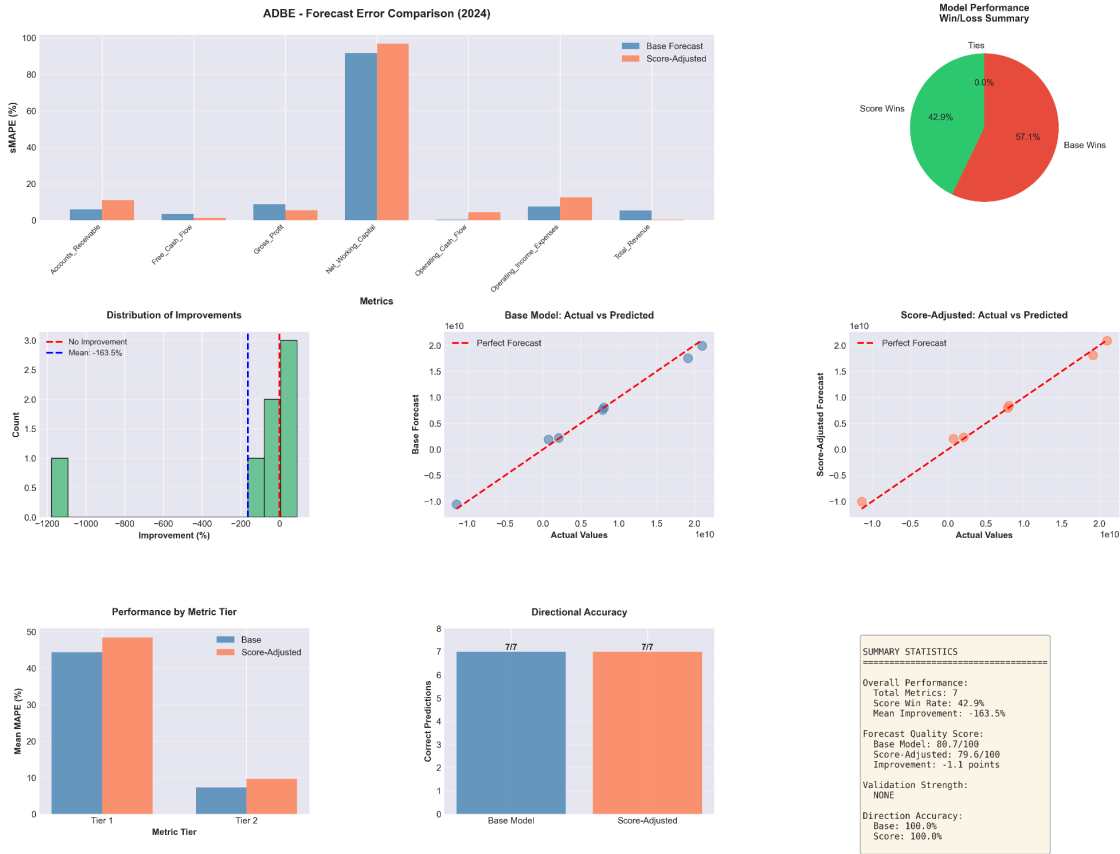
END OF REPORT

=====

=====

ADBE (LINEAR + WINTERS-HOLT + XGBOOST):

Forecast Validation Analysis: ADBE (2024)



FORECAST VALIDATION REPORT

Company: ADBE
Forecast Year: 2024
Analysis Date: 2025-12-11 02:02:04

EXECUTIVE SUMMARY

=====

Validation Strength: NONE

Score-Adjusted method does not validate. Base model performs as well or better.

Overall Win Rate: 42.9% (Score-Adjusted wins)

Mean Improvement: -163.5%

Forecast Quality Scores:

- Base Model: 80.7/100
- Score-Adjusted Model: 79.6/100
- Improvement: -1.1 points

=====

DETAILED METRICS COMPARISON

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 28.56%
- Median MAPE: 6.32%
- Mean sMAPE: 17.62%

Score-Adjusted Model:

- Mean MAPE: 31.84%
- Median MAPE: 5.43%
- Mean sMAPE: 18.89%

Win/Loss Record:

- Score-Adjusted Wins: 3
- Base Model Wins: 4
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 44.45%
- Score Mean MAPE: 48.49%
- Mean Improvement: -258.65%
- Score Win Rate: 50.0%
- Wins: Score=2, Base=2

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 7.37%
- Score Mean MAPE: 9.64%
- Mean Improvement: -36.59%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: -0.7493
- P-Value: 0.4820
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 8.0000
- P-Value: 0.3750
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: -0.2832
- Interpretation: Small effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 2,072,000,000.00

Base Forecast: 2,202,960,480.00 (Error: +130,960,480.00)

Score Forecast: 2,311,813,163.00 (Error: +239,813,163.00)

Accuracy:

- Base sMAPE: 6.13%
- Score sMAPE: 10.94%
- Improvement: -83.12%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 7,873,000,000.00

Base Forecast: 7,598,303,665.00 (Error: -274,696,335.00)

Score Forecast: 7,973,751,045.00 (Error: +100,751,045.00)

Accuracy:

- Base sMAPE: 3.55%
- Score sMAPE: 1.27%
- Improvement: +63.32%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 19,147,000,000.00

Base Forecast: 17,529,513,393.00 (Error: -1,617,486,607.00)

Score Forecast: 18,107,216,036.00 (Error: -1,039,783,964.00)

Accuracy:

- Base sMAPE: 8.82%
- Score sMAPE: 5.58%
- Improvement: +35.72%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 711,000,000.00

Base Forecast: 1,911,029,427.00 (Error: +1,200,029,427.00)

Score Forecast: 2,044,801,487.00 (Error: +1,333,801,487.00)

Accuracy:

- Base sMAPE: 91.53%
- Score sMAPE: 96.80%
- Improvement: -11.15%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 8,056,000,000.00

Base Forecast: 8,027,197,906.00 (Error: -28,802,094.00)

Score Forecast: 8,423,837,809.00 (Error: +367,837,809.00)

Accuracy:

- Base sMAPE: 0.36%
- Score sMAPE: 4.46%
- Improvement: -1177.12%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Income_Expenses (Tier 2):

Actual Value: -11,406,000,000.00

Base Forecast: -10,568,599,025.00 (Error: +837,400,975.00)

Score Forecast: -10,046,383,410.00 (Error: +1,359,616,590.00)

Accuracy:

- Base sMAPE: 7.62%
- Score sMAPE: 12.68%
- Improvement: -62.36%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 21,013,972,474.00

Base Forecast: 19,924,078,725.00 (Error: -1,089,893,749.00)

Score Forecast: 20,908,567,303.00 (Error: -105,405,171.00)

Accuracy:

- Base sMAPE: 5.32%
- Score sMAPE: 0.50%
- Improvement: +90.33%

Direction Correct: Base=True, Score=True

Winner: Score

=====

=====

RECOMMENDATIONS

=====

=====

X No validation. Base model performs as well or better. Recommend:

1. Critically review the qualitative scoring framework
2. Check if qualitative factors are already reflected in historical data
3. Consider that market efficiency may limit qualitative alpha
4. Explore if timing or sector-specific factors affected results

=====

=====

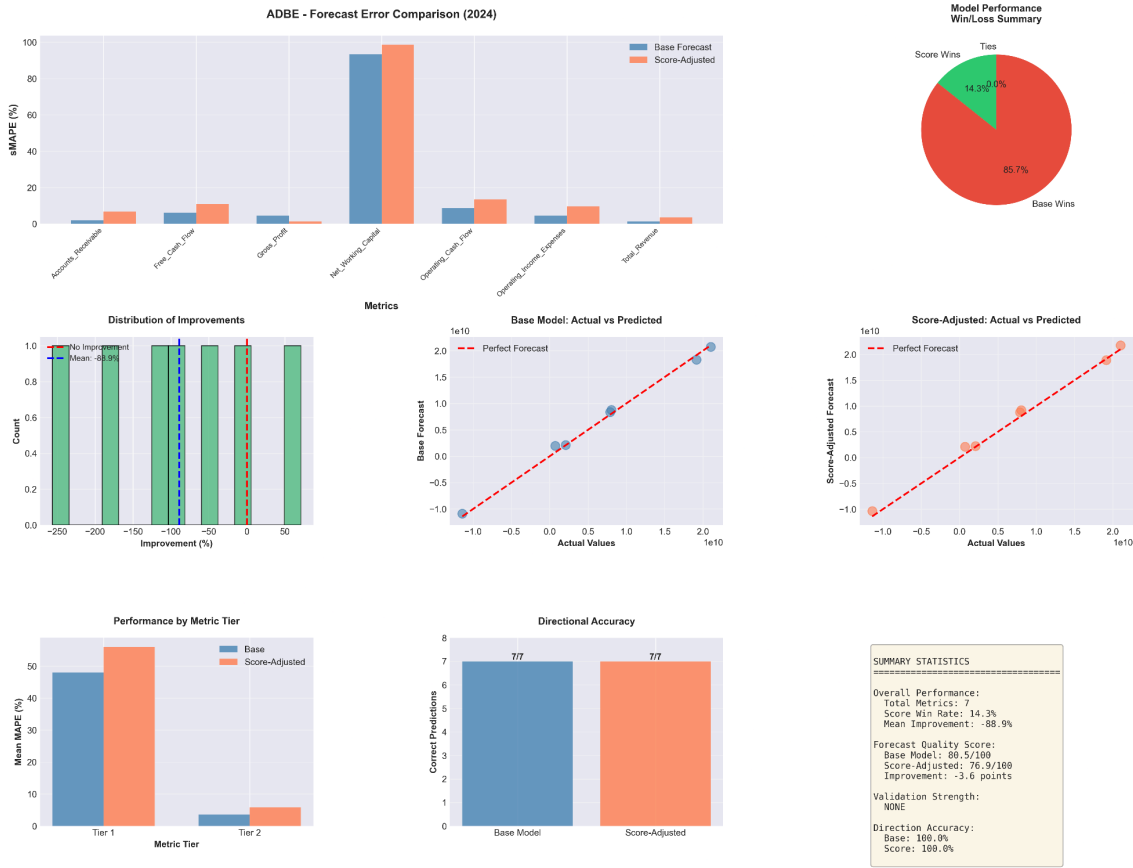
END OF REPORT

=====

=====

ADBE (LINEAR + WINTERS-HOLT + ARIMA):

Forecast Validation Analysis: ADBE (2024)



FORECAST VALIDATION REPORT

Company: ADBE
Forecast Year: 2024
Analysis Date: 2025-12-11 02:07:41

EXECUTIVE SUMMARY

Validation Strength: NONE

Score-Adjusted method does not validate. Base model performs as well or better.

Overall Win Rate: 14.3% (Score-Adjusted wins)

Mean Improvement: -88.9%

Forecast Quality Scores:

- Base Model: 80.5/100
- Score-Adjusted Model: 76.9/100
- Improvement: -3.6 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 28.98%
- Median MAPE: 4.51%
- Mean sMAPE: 17.22%

Score-Adjusted Model:

- Mean MAPE: 34.52%
- Median MAPE: 9.23%
- Mean sMAPE: 20.60%

Win/Loss Record:

- Score-Adjusted Wins: 1
- Base Model Wins: 6
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 47.98%
- Score Mean MAPE: 56.04%
- Mean Improvement: -83.23%
- Score Win Rate: 0.0%
- Wins: Score=0, Base=4

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 3.63%
- Score Mean MAPE: 5.84%
- Mean Improvement: -96.51%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

STATISTICAL SIGNIFICANCE TESTS

Paired T-Test:

- T-Statistic: -2.8930
- P-Value: 0.0276
- Significant at $\alpha=0.05$: True
- Interpretation: Score-Adjusted significantly different

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 2.0000
- P-Value: 0.0469
- Significant at $\alpha=0.05$: True

Effect Size:

- Cohen's d: -1.0934
- Interpretation: Large effect

METRIC-BY-METRIC BREAKDOWN

Accounts_Receivable (Tier 2):

Actual Value: 2,072,000,000.00

Base Forecast: 2,112,752,562.00 (Error: +40,752,562.00)

Score Forecast: 2,217,147,892.00 (Error: +145,147,892.00)

Accuracy:

- Base sMAPE: 1.95%
- Score sMAPE: 6.77%
- Improvement: -256.17%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 7,873,000,000.00

Base Forecast: 8,366,586,094.00 (Error: +493,586,094.00)

Score Forecast: 8,779,995,846.00 (Error: +906,995,846.00)

Accuracy:

- Base sMAPE: 6.08%
- Score sMAPE: 10.89%
- Improvement: -83.76%

Direction Correct: Base=True, Score=True

Winner: Base

Gross_Profit (Tier 2):

Actual Value: 19,147,000,000.00

Base Forecast: 18,299,813,139.00 (Error: -847,186,861.00)

Score Forecast: 18,902,901,781.00 (Error: -244,098,219.00)

Accuracy:

- Base sMAPE: 4.52%
- Score sMAPE: 1.28%
- Improvement: +71.19%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 711,000,000.00

Base Forecast: 1,957,747,959.00 (Error: +1,246,747,959.00)

Score Forecast: 2,094,790,317.00 (Error: +1,383,790,317.00)

Accuracy:

- Base sMAPE: 93.43%
- Score sMAPE: 98.64%
- Improvement: -10.99%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 8,056,000,000.00

Base Forecast: 8,782,528,195.00 (Error: +726,528,195.00)

Score Forecast: 9,216,490,479.00 (Error: +1,160,490,479.00)

Accuracy:

- Base sMAPE: 8.63%
- Score sMAPE: 13.44%
- Improvement: -59.73%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Income_Expenses (Tier 2):

Actual Value: -11,406,000,000.00

Base Forecast: -10,891,301,775.00 (Error: +514,698,225.00)

Score Forecast: -10,353,140,772.00 (Error: +1,052,859,228.00)

Accuracy:

- Base sMAPE: 4.62%
- Score sMAPE: 9.68%
- Improvement: -104.56%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 21,013,972,474.00

Base Forecast: 20,743,128,321.00 (Error: -270,844,153.00)

Score Forecast: 21,768,087,778.00 (Error: +754,115,304.00)

Accuracy:

- Base sMAPE: 1.30%
- Score sMAPE: 3.53%
- Improvement: -178.43%

Direction Correct: Base=True, Score=True
Winner: Base

=====

=====

RECOMMENDATIONS

=====

=====

- X No validation. Base model performs as well or better. Recommend:
1. Critically review the qualitative scoring framework
 2. Check if qualitative factors are already reflected in historical data
 3. Consider that market efficiency may limit qualitative alpha
 4. Explore if timing or sector-specific factors affected results

=====

=====

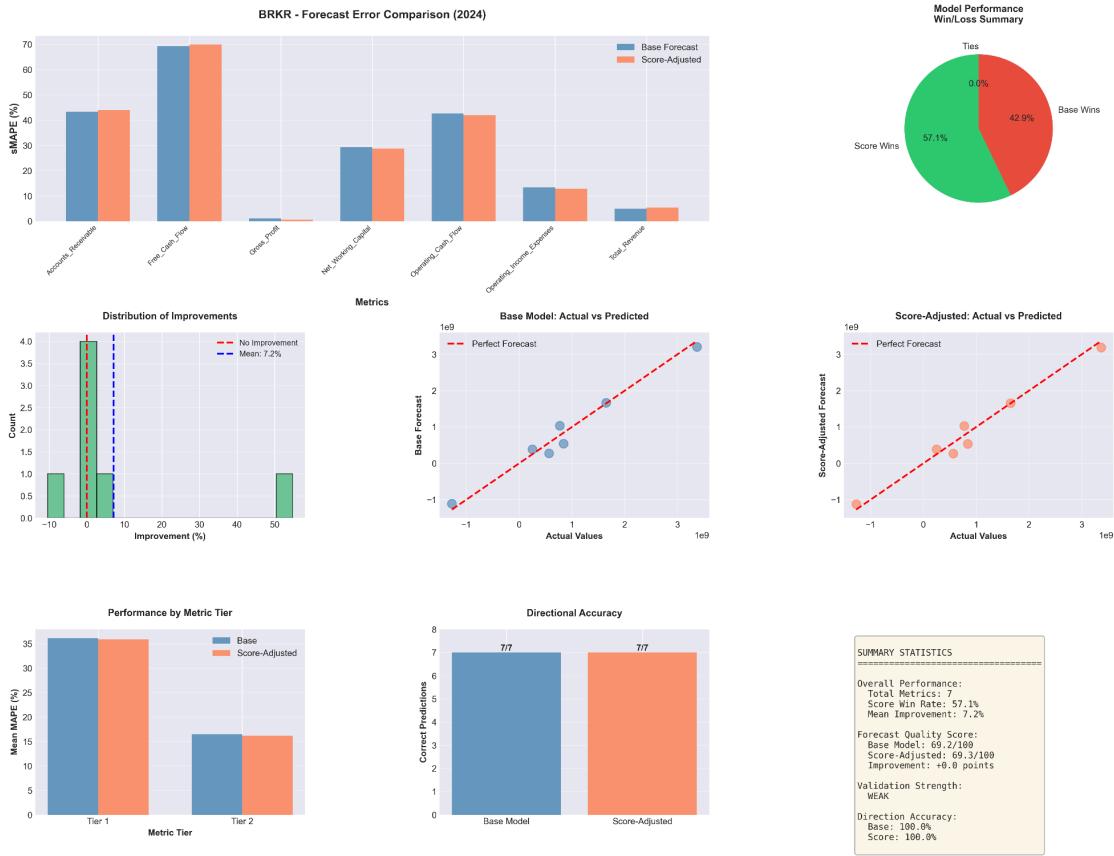
END OF REPORT

=====

=====

BRKR (NAIVE):

Forecast Validation Analysis: BRKR (2024)



FORECAST VALIDATION REPORT

Company: BRKR
Forecast Year: 2024
Analysis Date: 2025-12-12 20:27:38

EXECUTIVE SUMMARY

Validation Strength: WEAK

Score-Adjusted method shows weak validation. Qualitative factors provide marginal improvements.

Overall Win Rate: 57.1% (Score-Adjusted wins)

Mean Improvement: 7.2%

Forecast Quality Scores:

- Base Model: 69.2/100
- Score-Adjusted Model: 69.3/100
- Improvement: +0.0 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 27.71%
- Median MAPE: 34.30%
- Mean sMAPE: 29.12%

Score-Adjusted Model:

- Mean MAPE: 27.49%
- Median MAPE: 33.62%
- Mean sMAPE: 29.05%

Win/Loss Record:

- Score-Adjusted Wins: 4
- Base Model Wins: 3
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 36.14%
- Score Mean MAPE: 35.94%
- Mean Improvement: -1.82%
- Score Win Rate: 50.0%
- Wins: Score=2, Base=2

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 16.46%
- Score Mean MAPE: 16.22%
- Mean Improvement: 19.13%
- Score Win Rate: 66.7%
- Wins: Score=2, Base=1

STATISTICAL SIGNIFICANCE TESTS

Paired T-Test:

- T-Statistic: 0.2736
- P-Value: 0.7936
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 13.0000
- P-Value: 0.9375
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.1034
- Interpretation: Small effect

METRIC-BY-METRIC BREAKDOWN

Accounts_Receivable (Tier 2):

Actual Value: 842,800,000.00

Base Forecast: 542,725,636.00 (Error: -300,074,364.00)

Score Forecast: 539,174,841.00 (Error: -303,625,159.00)

Accuracy:

- Base sMAPE: 43.32%
- Score sMAPE: 43.94%
- Improvement: -1.18%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 565,500,000.00

Base Forecast: 274,857,225.00 (Error: -290,642,775.00)

Score Forecast: 272,641,219.00 (Error: -292,858,781.00)

Accuracy:

- Base sMAPE: 69.17%
- Score sMAPE: 69.88%
- Improvement: -0.76%

Direction Correct: Base=True, Score=True

Winner: Base

Gross_Profit (Tier 2):

Actual Value: 1,649,500,000.00

Base Forecast: 1,669,444,841.00 (Error: +19,944,841.00)

Score Forecast: 1,658,514,702.00 (Error: +9,014,702.00)

Accuracy:

- Base sMAPE: 1.20%
- Score sMAPE: 0.55%
- Improvement: +54.80%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 771,900,000.00

Base Forecast: 1,036,630,493.00 (Error: +264,730,493.00)

Score Forecast: 1,031,413,358.00 (Error: +259,513,358.00)

Accuracy:

- Base sMAPE: 29.28%
- Score sMAPE: 28.78%
- Improvement: +1.97%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 251,300,000.00

Base Forecast: 387,218,773.00 (Error: +135,918,773.00)

Score Forecast: 384,620,459.00 (Error: +133,320,459.00)

Accuracy:

- Base sMAPE: 42.57%
- Score sMAPE: 41.93%
- Improvement: +1.91%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -1,270,300,000.00

Base Forecast: -1,110,778,277.00 (Error: +159,521,723.00)

Score Forecast: -1,116,796,756.00 (Error: +153,503,244.00)

Accuracy:

- Base sMAPE: 13.40%
- Score sMAPE: 12.86%
- Improvement: +3.77%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 3,366,400,000.00

Base Forecast: 3,204,858,500.00 (Error: -161,541,500.00)

Score Forecast: 3,188,033,405.00 (Error: -178,366,595.00)

Accuracy:

- Base sMAPE: 4.92%
- Score sMAPE: 5.44%
- Improvement: -10.42%

Direction Correct: Base=True, Score=True
Winner: Base

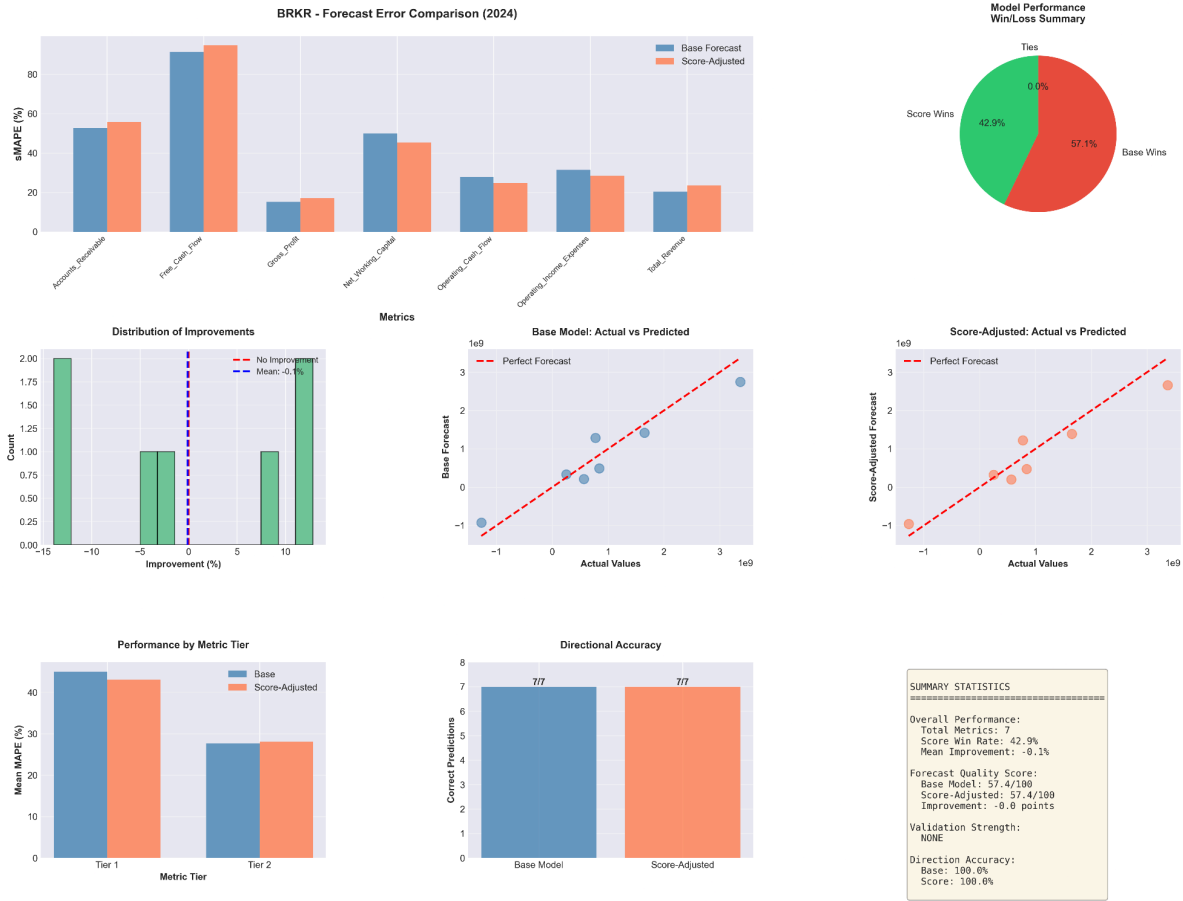
```
=====
=====
RECOMMENDATIONS
=====
=====
```

- △ Weak validation. Qualitative adjustments show marginal impact. Recommend:
1. Review and refine the qualitative scoring methodology
 2. Investigate if certain metric types benefit more from adjustments
 3. Consider whether the base model is already capturing qualitative factors
 4. Test alternative adjustment algorithms

```
=====
=====
END OF REPORT
=====
=====
```


BRKR (LINEAL):

Forecast Validation Analysis: BRKR (2024)



FORECAST VALIDATION REPORT

Company: BRKR
Forecast Year: 2024
Analysis Date: 2025-12-12 20:28:24

EXECUTIVE SUMMARY

Validation Strength: NONE

Score-Adjusted method does not validate. Base model performs as well or better.

Overall Win Rate: 42.9% (Score-Adjusted wins)

Mean Improvement: -0.1%

Forecast Quality Scores:

- Base Model: 57.4/100
- Score-Adjusted Model: 57.4/100
- Improvement: -0.0 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 37.58%
- Median MAPE: 32.47%
- Mean sMAPE: 41.24%

Score-Adjusted Model:

- Mean MAPE: 36.65%
- Median MAPE: 28.31%
- Mean sMAPE: 41.37%

Win/Loss Record:

- Score-Adjusted Wins: 3
- Base Model Wins: 4
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 45.03%
- Score Mean MAPE: 43.06%
- Mean Improvement: 2.05%
- Score Win Rate: 50.0%
- Wins: Score=2, Base=2

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 27.64%
- Score Mean MAPE: 28.09%
- Mean Improvement: -2.93%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

STATISTICAL SIGNIFICANCE TESTS

Paired T-Test:

- T-Statistic: -0.0991
- P-Value: 0.9243
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 14.0000
- P-Value: 1.0000
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: -0.0374
- Interpretation: Small effect

METRIC-BY-METRIC BREAKDOWN

Accounts_Receivable (Tier 2):

Actual Value: 842,800,000.00

Base Forecast: 490,925,694.00 (Error: -351,874,306.00)

Score Forecast: 475,489,026.00 (Error: -367,310,974.00)

Accuracy:

- Base sMAPE: 52.77%

- Score sMAPE: 55.73%

- Improvement: -4.39%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 565,500,000.00

Base Forecast: 211,219,833.00 (Error: -354,280,167.00)

Score Forecast: 202,186,964.00 (Error: -363,313,036.00)

Accuracy:

- Base sMAPE: 91.22%

- Score sMAPE: 94.65%

- Improvement: -2.55%

Direction Correct: Base=True, Score=True

Winner: Base

Gross_Profit (Tier 2):

Actual Value: 1,649,500,000.00

Base Forecast: 1,418,017,500.00 (Error: -231,482,500.00)

Score Forecast: 1,388,278,836.00 (Error: -261,221,164.00)

Accuracy:

- Base sMAPE: 15.09%

- Score sMAPE: 17.20%

- Improvement: -12.85%

Direction Correct: Base=True, Score=True

Winner: Base

Net_Working_Capital (Tier 1):

Actual Value: 771,900,000.00

Base Forecast: 1,285,434,444.00 (Error: +513,534,444.00)

Score Forecast: 1,224,787,647.00 (Error: +452,887,647.00)

Accuracy:

- Base sMAPE: 49.92%
- Score sMAPE: 45.36%
- Improvement: +11.81%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 251,300,000.00

Base Forecast: 332,898,472.00 (Error: +81,598,472.00)

Score Forecast: 322,430,812.00 (Error: +71,130,812.00)

Accuracy:

- Base sMAPE: 27.94%
- Score sMAPE: 24.80%
- Improvement: +12.83%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -1,270,300,000.00

Base Forecast: -925,429,305.00 (Error: +344,870,695.00)

Score Forecast: -954,528,504.00 (Error: +315,771,496.00)

Accuracy:

- Base sMAPE: 31.41%
- Score sMAPE: 28.39%
- Improvement: +8.44%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 3,366,400,000.00

Base Forecast: 2,745,233,472.00 (Error: -621,166,528.00)

Score Forecast: 2,658,912,350.00 (Error: -707,487,650.00)

Accuracy:

- Base sMAPE: 20.33%
- Score sMAPE: 23.48%
- Improvement: -13.90%

Direction Correct: Base=True, Score=True
Winner: Base

=====

=====

RECOMMENDATIONS

=====

=====

- X No validation. Base model performs as well or better. Recommend:
1. Critically review the qualitative scoring framework
 2. Check if qualitative factors are already reflected in historical data
 3. Consider that market efficiency may limit qualitative alpha
 4. Explore if timing or sector-specific factors affected results

=====

=====

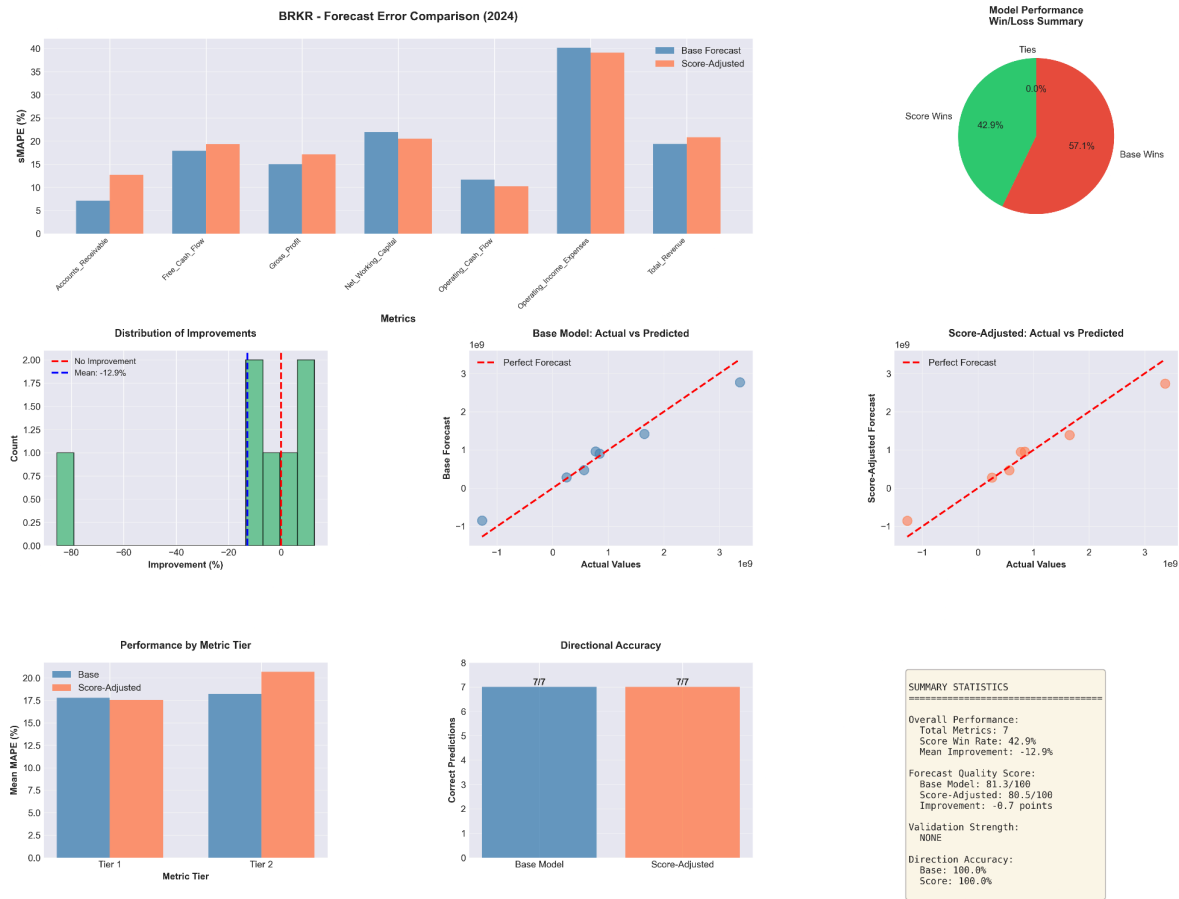
END OF REPORT

=====

=====

BRKR (XGBOOST):

Forecast Validation Analysis: BRKR (2024)



FORECAST VALIDATION REPORT

Company: BRKR
Forecast Year: 2024
Analysis Date: 2025-12-12 20:28:56

EXECUTIVE SUMMARY

Validation Strength: NONE

Score-Adjusted method does not validate. Base model performs as well or better.

Overall Win Rate: 42.9% (Score-Adjusted wins)

Mean Improvement: -12.9%

Forecast Quality Scores:

- Base Model: 81.3/100
- Score-Adjusted Model: 80.5/100
- Improvement: -0.7 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 17.97%
- Median MAPE: 16.41%
- Mean sMAPE: 19.01%

Score-Adjusted Model:

- Mean MAPE: 18.88%
- Median MAPE: 17.58%
- Mean sMAPE: 19.96%

Win/Loss Record:

- Score-Adjusted Wins: 3
- Base Model Wins: 4
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 17.78%
- Score Mean MAPE: 17.53%
- Mean Improvement: 1.54%
- Score Win Rate: 50.0%
- Wins: Score=2, Base=2

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 18.23%
- Score Mean MAPE: 20.67%
- Mean Improvement: -32.10%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

STATISTICAL SIGNIFICANCE TESTS

Paired T-Test:

- T-Statistic: -0.9915
- P-Value: 0.3597
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 8.0000
- P-Value: 0.3750
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: -0.3748
- Interpretation: Small effect

METRIC-BY-METRIC BREAKDOWN

Accounts_Receivable (Tier 2):

Actual Value: 842,800,000.00

Base Forecast: 904,554,839.00 (Error: +61,754,839.00)

Score Forecast: 957,377,223.00 (Error: +114,577,223.00)

Accuracy:

- Base sMAPE: 7.07%
- Score sMAPE: 12.73%
- Improvement: -85.54%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 565,500,000.00

Base Forecast: 472,699,968.00 (Error: -92,800,032.00)

Score Forecast: 466,082,168.00 (Error: -99,417,832.00)

Accuracy:

- Base sMAPE: 17.88%
- Score sMAPE: 19.27%
- Improvement: -7.13%

Direction Correct: Base=True, Score=True

Winner: Base

Gross_Profit (Tier 2):

Actual Value: 1,649,500,000.00

Base Forecast: 1,419,805,568.00 (Error: -229,694,432.00)

Score Forecast: 1,389,989,651.00 (Error: -259,510,349.00)

Accuracy:

- Base sMAPE: 14.97%
- Score sMAPE: 17.08%
- Improvement: -12.98%

Direction Correct: Base=True, Score=True

Winner: Base

Net_Working_Capital (Tier 1):

Actual Value: 771,900,000.00

Base Forecast: 962,101,632.00 (Error: +190,201,632.00)

Score Forecast: 948,632,209.00 (Error: +176,732,209.00)

Accuracy:

- Base sMAPE: 21.94%
- Score sMAPE: 20.54%
- Improvement: +7.08%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 251,300,000.00

Base Forecast: 282,399,968.00 (Error: +31,099,968.00)

Score Forecast: 278,446,368.00 (Error: +27,146,368.00)

Accuracy:

- Base sMAPE: 11.65%
- Score sMAPE: 10.25%
- Improvement: +12.71%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -1,270,300,000.00

Base Forecast: -845,701,056.00 (Error: +424,598,944.00)

Score Forecast: -855,172,907.00 (Error: +415,127,093.00)

Accuracy:

- Base sMAPE: 40.13%
- Score sMAPE: 39.06%
- Improvement: +2.23%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 3,366,400,000.00

Base Forecast: 2,770,671,104.00 (Error: -595,728,896.00)

Score Forecast: 2,731,881,708.00 (Error: -634,518,292.00)

Accuracy:

- Base sMAPE: 19.41%
- Score sMAPE: 20.81%
- Improvement: -6.51%

Direction Correct: Base=True, Score=True
Winner: Base

=====

=====

RECOMMENDATIONS

=====

=====

- X No validation. Base model performs as well or better. Recommend:
1. Critically review the qualitative scoring framework
 2. Check if qualitative factors are already reflected in historical data
 3. Consider that market efficiency may limit qualitative alpha
 4. Explore if timing or sector-specific factors affected results

=====

=====

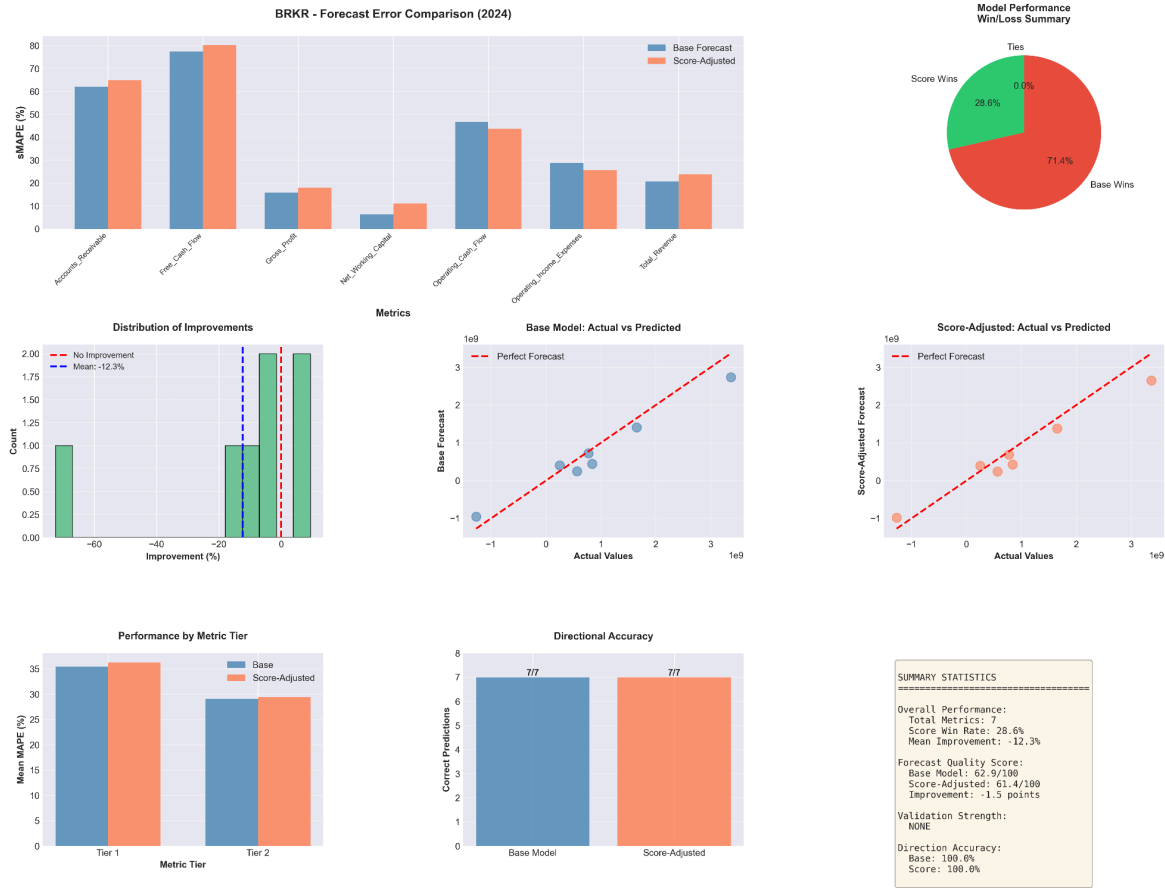
END OF REPORT

=====

=====

BRKR (ARIMA):

Forecast Validation Analysis: BRKR (2024)



FORECAST VALIDATION REPORT

Company: BRKR
Forecast Year: 2024
Analysis Date: 2025-12-12 20:29:29

EXECUTIVE SUMMARY

Validation Strength: NONE

Score-Adjusted method does not validate. Base model performs as well or better.

Overall Win Rate: 28.6% (Score-Adjusted wins)

Mean Improvement: -12.3%

Forecast Quality Scores:

- Base Model: 62.9/100
- Score-Adjusted Model: 61.4/100
- Improvement: -1.5 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 32.73%
- Median MAPE: 25.09%
- Mean sMAPE: 36.87%

Score-Adjusted Model:

- Mean MAPE: 33.35%
- Median MAPE: 22.73%
- Mean sMAPE: 38.24%

Win/Loss Record:

- Score-Adjusted Wins: 2
- Base Model Wins: 5
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 35.47%
- Score Mean MAPE: 36.29%
- Mean Improvement: -20.05%
- Score Win Rate: 25.0%
- Wins: Score=1, Base=3

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 29.07%
- Score Mean MAPE: 29.43%
- Mean Improvement: -2.08%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

STATISTICAL SIGNIFICANCE TESTS

Paired T-Test:

- T-Statistic: -1.1602
- P-Value: 0.2900
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 9.0000
- P-Value: 0.4688
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: -0.4385
- Interpretation: Small effect

METRIC-BY-METRIC BREAKDOWN

Accounts_Receivable (Tier 2):

Actual Value: 842,800,000.00
Base Forecast: 443,376,434.00 (Error: -399,423,566.00)
Score Forecast: 429,434,905.00 (Error: -413,365,095.00)

Accuracy:

- Base sMAPE: 62.11%
- Score sMAPE: 64.98%
- Improvement: -3.49%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 565,500,000.00
Base Forecast: 249,672,151.00 (Error: -315,827,849.00)
Score Forecast: 241,821,459.00 (Error: -323,678,541.00)

Accuracy:

- Base sMAPE: 77.49%
- Score sMAPE: 80.19%
- Improvement: -2.49%

Direction Correct: Base=True, Score=True

Winner: Base

Gross_Profit (Tier 2):

Actual Value: 1,649,500,000.00
Base Forecast: 1,406,521,014.00 (Error: -242,978,986.00)
Score Forecast: 1,377,023,456.00 (Error: -272,476,544.00)

Accuracy:

- Base sMAPE: 15.90%
- Score sMAPE: 18.01%
- Improvement: -12.14%

Direction Correct: Base=True, Score=True

Winner: Base

Net_Working_Capital (Tier 1):

Actual Value: 771,900,000.00
Base Forecast: 724,696,779.00 (Error: -47,203,221.00)
Score Forecast: 690,505,585.00 (Error: -81,394,415.00)

Accuracy:

- Base sMAPE: 6.31%
- Score sMAPE: 11.13%
- Improvement: -72.43%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 251,300,000.00

Base Forecast: 404,811,458.00 (Error: +153,511,458.00)

Score Forecast: 392,082,566.00 (Error: +140,782,566.00)

Accuracy:

- Base sMAPE: 46.79%
- Score sMAPE: 43.76%
- Improvement: +8.29%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -1,270,300,000.00

Base Forecast: -951,599,771.00 (Error: +318,700,229.00)

Score Forecast: -981,521,874.00 (Error: +288,778,126.00)

Accuracy:

- Base sMAPE: 28.69%
- Score sMAPE: 25.65%
- Improvement: +9.39%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 3,366,400,000.00

Base Forecast: 2,732,861,780.00 (Error: -633,538,220.00)

Score Forecast: 2,646,929,674.00 (Error: -719,470,326.00)

Accuracy:

- Base sMAPE: 20.77%

- Score sMAPE: 23.93%
- Improvement: -13.56%

Direction Correct: Base=True, Score=True

Winner: Base

=====

=====

RECOMMENDATIONS

=====

=====

✗ No validation. Base model performs as well or better. Recommend:

1. Critically review the qualitative scoring framework
2. Check if qualitative factors are already reflected in historical data
3. Consider that market efficiency may limit qualitative alpha
4. Explore if timing or sector-specific factors affected results

=====

=====

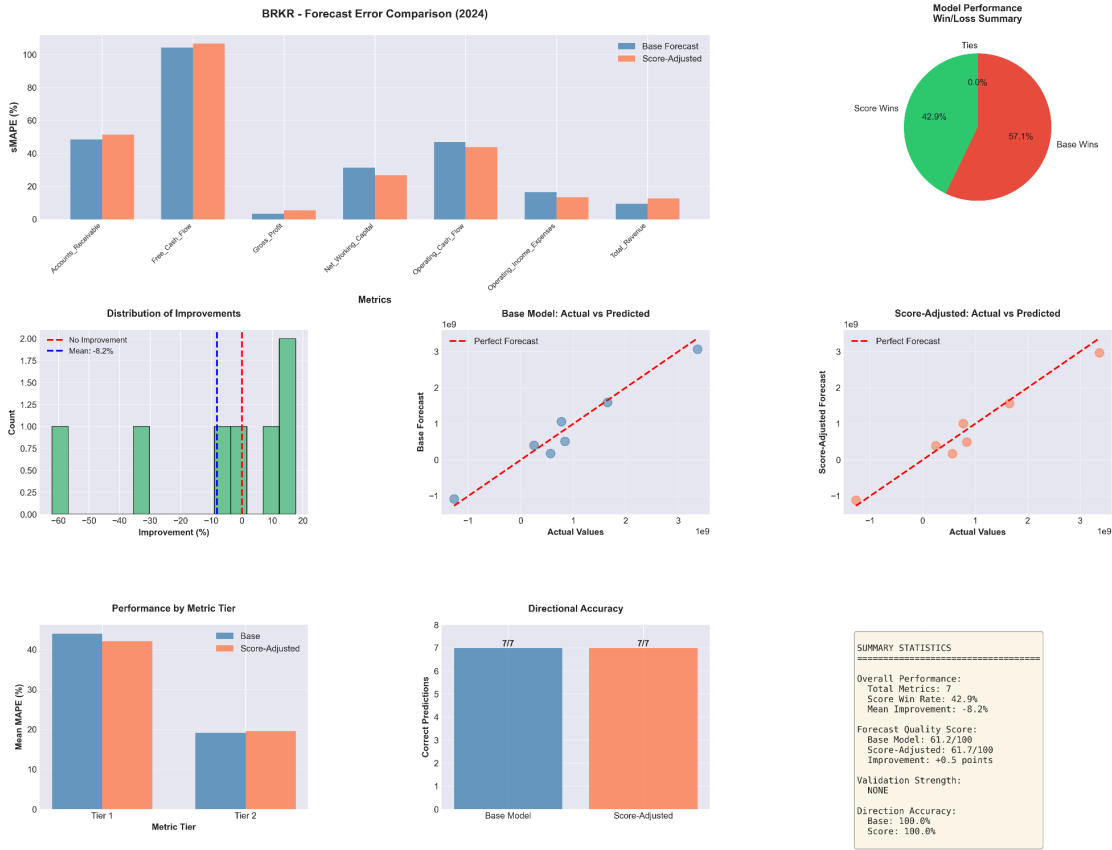
END OF REPORT

=====

=====

BRKR (WINTERS-HOLT):

Forecast Validation Analysis: BRKR (2024)



FORECAST VALIDATION REPORT

Company: BRKR
Forecast Year: 2024
Analysis Date: 2025-12-12 20:30:18

EXECUTIVE SUMMARY

Validation Strength: NONE

Score-Adjusted method does not validate. Base model performs as well or better.

Overall Win Rate: 42.9% (Score-Adjusted wins)

Mean Improvement: -8.2%

Forecast Quality Scores:

- Base Model: 61.2/100
- Score-Adjusted Model: 61.7/100
- Improvement: +0.5 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 33.31%
- Median MAPE: 37.28%
- Mean sMAPE: 37.13%

Score-Adjusted Model:

- Mean MAPE: 32.40%
- Median MAPE: 30.81%
- Mean sMAPE: 37.09%

Win/Loss Record:

- Score-Adjusted Wins: 3
- Base Model Wins: 4
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 43.96%
- Score Mean MAPE: 42.04%
- Mean Improvement: -1.93%
- Score Win Rate: 50.0%
- Wins: Score=2, Base=2

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 19.12%
- Score Mean MAPE: 19.54%
- Mean Improvement: -16.46%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

STATISTICAL SIGNIFICANCE TESTS

Paired T-Test:

- T-Statistic: 0.0258
- P-Value: 0.9803
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 12.0000
- P-Value: 0.8125
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.0097
- Interpretation: Small effect

METRIC-BY-METRIC BREAKDOWN

Accounts_Receivable (Tier 2):

Actual Value: 842,800,000.00
Base Forecast: 514,543,501.00 (Error: -328,256,499.00)
Score Forecast: 498,364,196.00 (Error: -344,435,804.00)

Accuracy:

- Base sMAPE: 48.37%
- Score sMAPE: 51.36%
- Improvement: -4.93%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 565,500,000.00
Base Forecast: 178,047,866.00 (Error: -387,452,134.00)
Score Forecast: 172,449,329.00 (Error: -393,050,671.00)

Accuracy:

- Base sMAPE: 104.22%
- Score sMAPE: 106.53%
- Improvement: -1.44%

Direction Correct: Base=True, Score=True

Winner: Base

Gross_Profit (Tier 2):

Actual Value: 1,649,500,000.00
Base Forecast: 1,595,590,665.00 (Error: -53,909,335.00)
Score Forecast: 1,562,127,938.00 (Error: -87,372,062.00)

Accuracy:

- Base sMAPE: 3.32%
- Score sMAPE: 5.44%
- Improvement: -62.07%

Direction Correct: Base=True, Score=True

Winner: Base

Net_Working_Capital (Tier 1):

Actual Value: 771,900,000.00
Base Forecast: 1,059,689,009.00 (Error: +287,789,009.00)
Score Forecast: 1,009,692,881.00 (Error: +237,792,881.00)

Accuracy:

- Base sMAPE: 31.43%
- Score sMAPE: 26.69%
- Improvement: +17.37%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 251,300,000.00

Base Forecast: 404,811,458.00 (Error: +153,511,458.00)

Score Forecast: 392,082,566.00 (Error: +140,782,566.00)

Accuracy:

- Base sMAPE: 46.79%
- Score sMAPE: 43.76%
- Improvement: +8.29%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -1,270,300,000.00

Base Forecast: -1,078,020,556.00 (Error: +192,279,444.00)

Score Forecast: -1,111,917,834.00 (Error: +158,382,166.00)

Accuracy:

- Base sMAPE: 16.38%
- Score sMAPE: 13.30%
- Improvement: +17.63%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 3,366,400,000.00

Base Forecast: 3,064,710,072.00 (Error: -301,689,928.00)

Score Forecast: 2,968,343,328.00 (Error: -398,056,672.00)

Accuracy:

- Base sMAPE: 9.38%

- Score sMAPE: 12.57%
- Improvement: -31.94%

Direction Correct: Base=True, Score=True

Winner: Base

=====

=====

RECOMMENDATIONS

=====

=====

✗ No validation. Base model performs as well or better. Recommend:

1. Critically review the qualitative scoring framework
2. Check if qualitative factors are already reflected in historical data
3. Consider that market efficiency may limit qualitative alpha
4. Explore if timing or sector-specific factors affected results

=====

=====

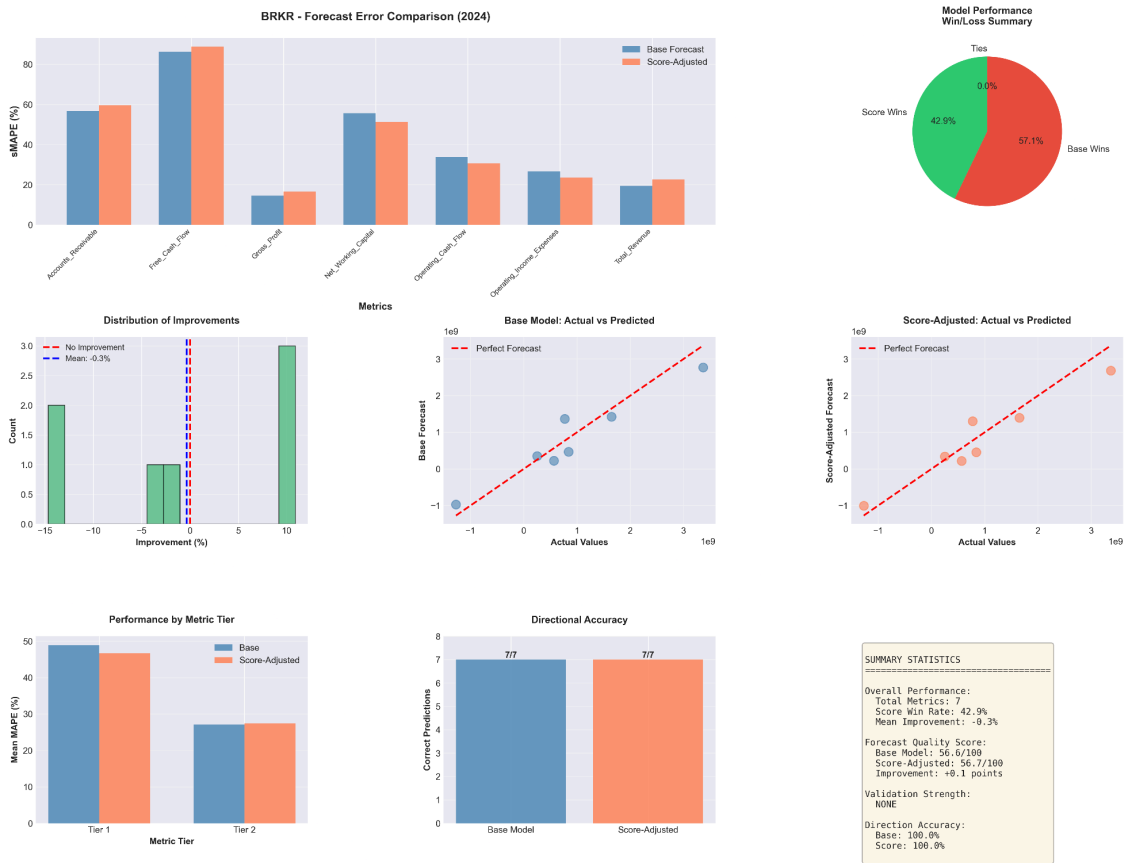
END OF REPORT

=====

=====

BRKR (LINEAR + WINTERS-HOLT):

Forecast Validation Analysis: BRKR (2024)



FORECAST VALIDATION REPORT

Company: BRKR
Forecast Year: 2024
Analysis Date: 2025-12-12 20:30:46

EXECUTIVE SUMMARY

Validation Strength: NONE

Score-Adjusted method does not validate. Base model performs as well or better.

Overall Win Rate: 42.9% (Score-Adjusted wins)

Mean Improvement: -0.3%

Forecast Quality Scores:

- Base Model: 56.6/100
- Score-Adjusted Model: 56.7/100
- Improvement: +0.1 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 39.55%
- Median MAPE: 40.57%
- Mean sMAPE: 41.83%

Score-Adjusted Model:

- Mean MAPE: 38.44%
- Median MAPE: 36.15%
- Mean sMAPE: 41.85%

Win/Loss Record:

- Score-Adjusted Wins: 3
- Base Model Wins: 4
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 48.91%
- Score Mean MAPE: 46.68%
- Mean Improvement: 1.25%
- Score Win Rate: 50.0%
- Wins: Score=2, Base=2

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 27.06%
- Score Mean MAPE: 27.45%
- Mean Improvement: -2.41%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

STATISTICAL SIGNIFICANCE TESTS

Paired T-Test:

- T-Statistic: -0.0144
- P-Value: 0.9890
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 12.0000
- P-Value: 0.8125
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: -0.0054
- Interpretation: Small effect

METRIC-BY-METRIC BREAKDOWN

Accounts_Receivable (Tier 2):

Actual Value: 842,800,000.00

Base Forecast: 470,678,042.00 (Error: -372,121,958.00)

Score Forecast: 455,878,041.00 (Error: -386,921,959.00)

Accuracy:

- Base sMAPE: 56.66%

- Score sMAPE: 59.59%

- Improvement: -3.98%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 565,500,000.00

Base Forecast: 224,980,599.00 (Error: -340,519,401.00)

Score Forecast: 217,906,309.00 (Error: -347,593,691.00)

Accuracy:

- Base sMAPE: 86.16%

- Score sMAPE: 88.74%

- Improvement: -2.08%

Direction Correct: Base=True, Score=True

Winner: Base

Gross_Profit (Tier 2):

Actual Value: 1,649,500,000.00

Base Forecast: 1,427,125,210.00 (Error: -222,374,790.00)

Score Forecast: 1,397,195,540.00 (Error: -252,304,460.00)

Accuracy:

- Base sMAPE: 14.46%

- Score sMAPE: 16.56%

- Improvement: -13.46%

Direction Correct: Base=True, Score=True

Winner: Base

Net_Working_Capital (Tier 1):

Actual Value: 771,900,000.00

Base Forecast: 1,367,767,025.00 (Error: +595,867,025.00)

Score Forecast: 1,303,235,777.00 (Error: +531,335,777.00)

Accuracy:

- Base sMAPE: 55.70%
- Score sMAPE: 51.21%
- Improvement: +10.83%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 251,300,000.00

Base Forecast: 353,261,878.00 (Error: +101,961,878.00)

Score Forecast: 342,153,911.00 (Error: +90,853,911.00)

Accuracy:

- Base sMAPE: 33.73%
- Score sMAPE: 30.62%
- Improvement: +10.89%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -1,270,300,000.00

Base Forecast: -971,016,518.00 (Error: +299,283,482.00)

Score Forecast: -1,001,549,162.00 (Error: +268,750,838.00)

Accuracy:

- Base sMAPE: 26.71%
- Score sMAPE: 23.66%
- Improvement: +10.20%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 3,366,400,000.00

Base Forecast: 2,771,611,938.00 (Error: -594,788,062.00)

Score Forecast: 2,684,461,372.00 (Error: -681,938,628.00)

Accuracy:

- Base sMAPE: 19.38%
- Score sMAPE: 22.54%
- Improvement: -14.65%

Direction Correct: Base=True, Score=True
Winner: Base

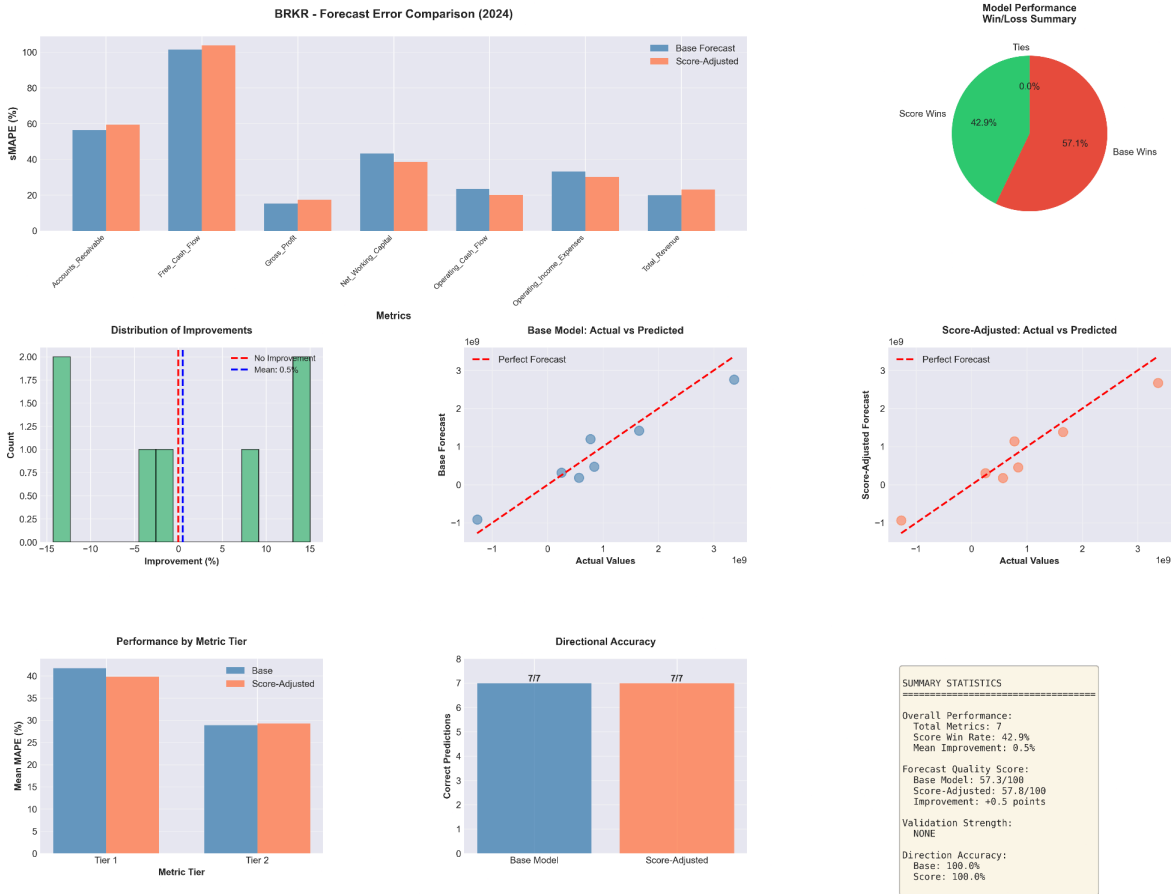
```
=====
=====
RECOMMENDATIONS
=====
=====
```

- X** No validation. Base model performs as well or better. Recommend:
1. Critically review the qualitative scoring framework
 2. Check if qualitative factors are already reflected in historical data
 3. Consider that market efficiency may limit qualitative alpha
 4. Explore if timing or sector-specific factors affected results

```
=====
=====
END OF REPORT
=====
=====
```


BRKR (LINEAR + WINTERS-HOLT + XGBOOST):

Forecast Validation Analysis: BRKR (2024)



FORECAST VALIDATION REPORT

Company: BRKR
Forecast Year: 2024
Analysis Date: 2025-12-12 20:31:17

EXECUTIVE SUMMARY

=====

Validation Strength: NONE

Score-Adjusted method does not validate. Base model performs as well or better.

Overall Win Rate: 42.9% (Score-Adjusted wins)

Mean Improvement: 0.5%

Forecast Quality Scores:

- Base Model: 57.3/100
- Score-Adjusted Model: 57.8/100
- Improvement: +0.5 points

=====

DETAILED METRICS COMPARISON

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 36.24%
- Median MAPE: 28.49%
- Mean sMAPE: 41.82%

Score-Adjusted Model:

- Mean MAPE: 35.33%
- Median MAPE: 26.24%
- Mean sMAPE: 41.78%

Win/Loss Record:

- Score-Adjusted Wins: 3
- Base Model Wins: 4
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 41.76%
- Score Mean MAPE: 39.83%
- Mean Improvement: 3.12%
- Score Win Rate: 50.0%
- Wins: Score=2, Base=2

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 28.88%
- Score Mean MAPE: 29.32%
- Mean Improvement: -2.96%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 0.0289
- P-Value: 0.9779
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 12.0000
- P-Value: 0.8125
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.0109
- Interpretation: Small effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 842,800,000.00

Base Forecast: 471,629,133.00 (Error: -371,170,867.00)

Score Forecast: 456,799,226.00 (Error: -386,000,774.00)

Accuracy:

- Base sMAPE: 56.48%

- Score sMAPE: 59.40%

- Improvement: -4.00%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 565,500,000.00

Base Forecast: 185,120,387.00 (Error: -380,379,613.00)

Score Forecast: 179,299,462.00 (Error: -386,200,538.00)

Accuracy:

- Base sMAPE: 101.35%

- Score sMAPE: 103.71%

- Improvement: -1.53%

Direction Correct: Base=True, Score=True

Winner: Base

Gross_Profit (Tier 2):

Actual Value: 1,649,500,000.00

Base Forecast: 1,416,709,805.00 (Error: -232,790,195.00)

Score Forecast: 1,386,998,567.00 (Error: -262,501,433.00)

Accuracy:

- Base sMAPE: 15.18%

- Score sMAPE: 17.29%

- Improvement: -12.76%

Direction Correct: Base=True, Score=True

Winner: Base

Net_Working_Capital (Tier 1):

Actual Value: 771,900,000.00

Base Forecast: 1,198,173,016.00 (Error: +426,273,016.00)

Score Forecast: 1,141,643,213.00 (Error: +369,743,213.00)

Accuracy:

- Base sMAPE: 43.27%
- Score sMAPE: 38.64%
- Improvement: +13.26%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 251,300,000.00

Base Forecast: 317,829,867.00 (Error: +66,529,867.00)

Score Forecast: 307,836,024.00 (Error: +56,536,024.00)

Accuracy:

- Base sMAPE: 23.38%
- Score sMAPE: 20.22%
- Improvement: +15.02%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -1,270,300,000.00

Base Forecast: -908,402,563.00 (Error: +361,897,437.00)

Score Forecast: -936,966,373.00 (Error: +333,333,627.00)

Accuracy:

- Base sMAPE: 33.22%
- Score sMAPE: 30.20%
- Improvement: +7.89%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 3,366,400,000.00

Base Forecast: 2,758,092,177.00 (Error: -608,307,823.00)

Score Forecast: 2,671,366,726.00 (Error: -695,033,274.00)

Accuracy:

- Base sMAPE: 19.86%
- Score sMAPE: 23.02%
- Improvement: -14.26%

Direction Correct: Base=True, Score=True

Winner: Base

=====

=====

RECOMMENDATIONS

=====

=====

X No validation. Base model performs as well or better. Recommend:

1. Critically review the qualitative scoring framework
2. Check if qualitative factors are already reflected in historical data
3. Consider that market efficiency may limit qualitative alpha
4. Explore if timing or sector-specific factors affected results

=====

=====

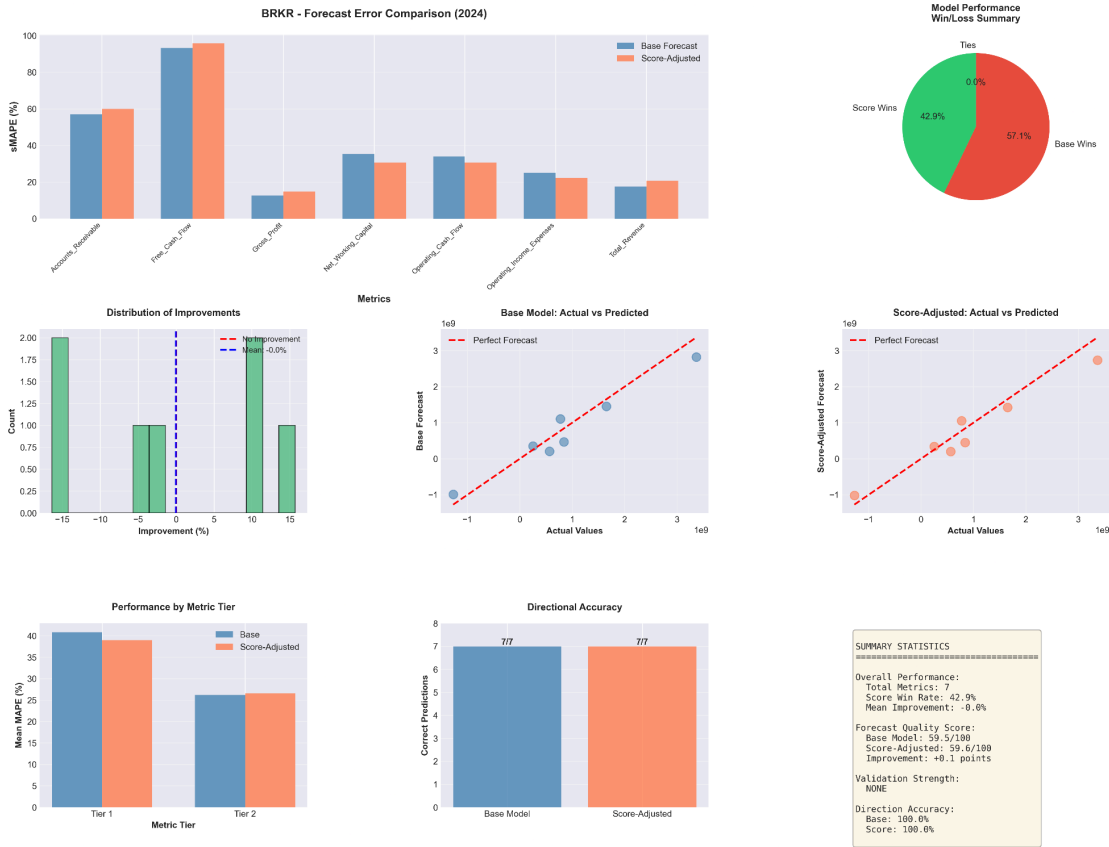
END OF REPORT

=====

=====

BRKR (LINEAR + WINTERS-HOLT + ARIMA):

Forecast Validation Analysis: BRKR (2024)



FORECAST VALIDATION REPORT

Company: BRKR
Forecast Year: 2024
Analysis Date: 2025-12-12 20:32:06

EXECUTIVE SUMMARY

Validation Strength: NONE

Score-Adjusted method does not validate. Base model performs as well or better.

Overall Win Rate: 42.9% (Score-Adjusted wins)

Mean Improvement: -0.0%

Forecast Quality Scores:

- Base Model: 59.5/100
- Score-Adjusted Model: 59.6/100
- Improvement: +0.1 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 34.57%
- Median MAPE: 40.77%
- Mean sMAPE: 39.26%

Score-Adjusted Model:

- Mean MAPE: 33.68%
- Median MAPE: 36.22%
- Mean sMAPE: 39.23%

Win/Loss Record:

- Score-Adjusted Wins: 3
- Base Model Wins: 4
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 40.85%
- Score Mean MAPE: 39.01%
- Mean Improvement: 2.10%
- Score Win Rate: 50.0%
- Wins: Score=2, Base=2

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 26.20%
- Score Mean MAPE: 26.59%
- Mean Improvement: -2.87%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

STATISTICAL SIGNIFICANCE TESTS

Paired T-Test:

- T-Statistic: 0.0206
- P-Value: 0.9842
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 12.0000
- P-Value: 0.8125
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.0078
- Interpretation: Small effect

METRIC-BY-METRIC BREAKDOWN

Accounts_Receivable (Tier 2):

Actual Value: 842,800,000.00

Base Forecast: 468,695,039.00 (Error: -374,104,961.00)

Score Forecast: 453,957,392.00 (Error: -388,842,608.00)

Accuracy:

- Base sMAPE: 57.05%

- Score sMAPE: 59.97%

- Improvement: -3.94%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 565,500,000.00

Base Forecast: 205,929,735.00 (Error: -359,570,265.00)

Score Forecast: 199,454,480.00 (Error: -366,045,520.00)

Accuracy:

- Base sMAPE: 93.22%

- Score sMAPE: 95.70%

- Improvement: -1.80%

Direction Correct: Base=True, Score=True

Winner: Base

Gross_Profit (Tier 2):

Actual Value: 1,649,500,000.00

Base Forecast: 1,453,836,823.00 (Error: -195,663,177.00)

Score Forecast: 1,423,346,957.00 (Error: -226,153,043.00)

Accuracy:

- Base sMAPE: 12.61%

- Score sMAPE: 14.72%

- Improvement: -15.58%

Direction Correct: Base=True, Score=True

Winner: Base

Net_Working_Capital (Tier 1):

Actual Value: 771,900,000.00

Base Forecast: 1,103,517,439.00 (Error: +331,617,439.00)

Score Forecast: 1,051,453,486.00 (Error: +279,553,486.00)

Accuracy:

- Base sMAPE: 35.36%
- Score sMAPE: 30.66%
- Improvement: +15.70%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 251,300,000.00

Base Forecast: 353,742,694.00 (Error: +102,442,694.00)

Score Forecast: 342,619,609.00 (Error: +91,319,609.00)

Accuracy:

- Base sMAPE: 33.86%
- Score sMAPE: 30.75%
- Improvement: +10.86%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -1,270,300,000.00

Base Forecast: -986,350,244.00 (Error: +283,949,756.00)

Score Forecast: -1,017,365,041.00 (Error: +252,934,959.00)

Accuracy:

- Base sMAPE: 25.17%
- Score sMAPE: 22.11%
- Improvement: +10.92%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 3,366,400,000.00

Base Forecast: 2,824,154,274.00 (Error: -542,245,726.00)

Score Forecast: 2,735,351,567.00 (Error: -631,048,433.00)

Accuracy:

- Base sMAPE: 17.52%
- Score sMAPE: 20.68%
- Improvement: -16.38%

Direction Correct: Base=True, Score=True
Winner: Base

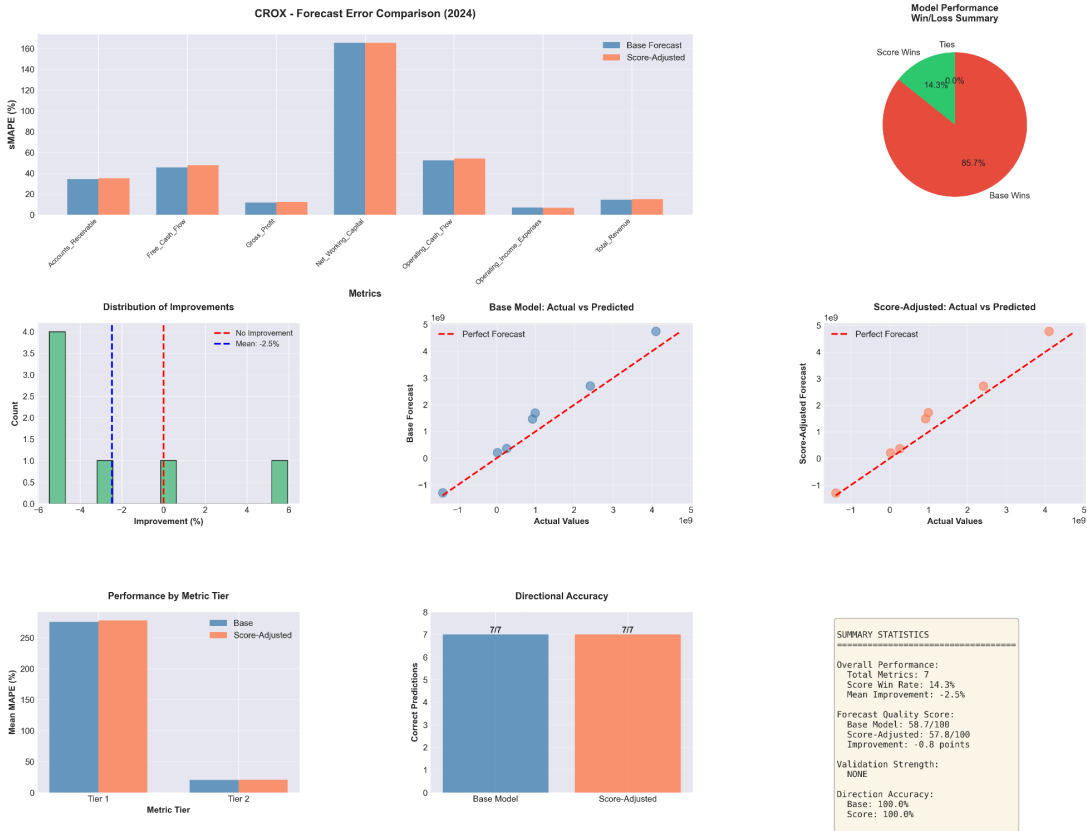
```
=====
=====
RECOMMENDATIONS
=====
=====
```

- X** No validation. Base model performs as well or better. Recommend:
1. Critically review the qualitative scoring framework
 2. Check if qualitative factors are already reflected in historical data
 3. Consider that market efficiency may limit qualitative alpha
 4. Explore if timing or sector-specific factors affected results

```
=====
=====
END OF REPORT
=====
=====
```


CROX (NAIVE):

Forecast Validation Analysis: CROX (2024)



FORECAST VALIDATION REPORT

Company: CROX
Forecast Year: 2024
Analysis Date: 2025-11-20 17:35:09

EXECUTIVE SUMMARY

Validation Strength: NONE

Score-Adjusted method does not validate. Base model performs as well or better.

Overall Win Rate: 14.3% (Score-Adjusted wins)

Mean Improvement: -2.5%

Forecast Quality Scores:

- Base Model: 58.7/100
- Score-Adjusted Model: 57.8/100
- Improvement: -0.8 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 166.19%
- Median MAPE: 41.46%
- Mean sMAPE: 47.30%

Score-Adjusted Model:

- Mean MAPE: 167.51%
- Median MAPE: 42.48%
- Mean sMAPE: 48.05%

Win/Loss Record:

- Score-Adjusted Wins: 1
- Base Model Wins: 6
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 275.63%
- Score Mean MAPE: 277.65%
- Mean Improvement: -3.95%
- Score Win Rate: 0.0%
- Wins: Score=0, Base=4

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 20.26%
- Score Mean MAPE: 20.67%
- Mean Improvement: -0.51%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

STATISTICAL SIGNIFICANCE TESTS

Paired T-Test:

- T-Statistic: -2.3079
- P-Value: 0.0604
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 2.0000
- P-Value: 0.0469
- Significant at $\alpha=0.05$: True

Effect Size:

- Cohen's d: -0.8723
- Interpretation: Large effect

METRIC-BY-METRIC BREAKDOWN

Accounts_Receivable (Tier 2):

Actual Value: 257,657,000.00
Base Forecast: 364,484,063.00 (Error: +106,827,063.00)
Score Forecast: 367,122,531.00 (Error: +109,465,531.00)

Accuracy:

- Base sMAPE: 34.34%
- Score sMAPE: 35.04%
- Improvement: -2.47%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 923,139,000.00
Base Forecast: 1,469,285,598.00 (Error: +546,146,598.00)
Score Forecast: 1,498,684,238.00 (Error: +575,545,238.00)

Accuracy:

- Base sMAPE: 45.66%
- Score sMAPE: 47.53%
- Improvement: -5.38%

Direction Correct: Base=True, Score=True

Winner: Base

Gross_Profit (Tier 2):

Actual Value: 2,410,258,000.00
Base Forecast: 2,708,598,060.00 (Error: +298,340,060.00)
Score Forecast: 2,723,535,758.00 (Error: +313,277,758.00)

Accuracy:

- Base sMAPE: 11.66%
- Score sMAPE: 12.20%
- Improvement: -5.01%

Direction Correct: Base=True, Score=True

Winner: Base

Net_Working_Capital (Tier 1):

Actual Value: 20,259,000.00
Base Forecast: 214,116,796.00 (Error: +193,857,796.00)
Score Forecast: 214,231,969.00 (Error: +193,972,969.00)

Accuracy:

- Base sMAPE: 165.42%
- Score sMAPE: 165.44%
- Improvement: -0.06%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 992,486,000.00

Base Forecast: 1,695,159,349.00 (Error: +702,673,349.00)

Score Forecast: 1,729,510,363.00 (Error: +737,024,363.00)

Accuracy:

- Base sMAPE: 52.29%
- Score sMAPE: 54.15%
- Improvement: -4.89%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Income_Expenses (Tier 2):

Actual Value: -1,388,347,000.00

Base Forecast: -1,291,809,700.00 (Error: +96,537,300.00)

Score Forecast: -1,297,553,607.00 (Error: +90,793,393.00)

Accuracy:

- Base sMAPE: 7.20%
- Score sMAPE: 6.76%
- Improvement: +5.95%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 4,102,108,000.00

Base Forecast: 4,744,220,587.00 (Error: +642,112,587.00)

Score Forecast: 4,779,342,349.00 (Error: +677,234,349.00)

Accuracy:

- Base sMAPE: 14.52%

- Score sMAPE: 15.25%
- Improvement: -5.47%

Direction Correct: Base=True, Score=True

Winner: Base

=====

=====

RECOMMENDATIONS

=====

=====

✗ No validation. Base model performs as well or better. Recommend:

1. Critically review the qualitative scoring framework
2. Check if qualitative factors are already reflected in historical data
3. Consider that market efficiency may limit qualitative alpha
4. Explore if timing or sector-specific factors affected results

=====

=====

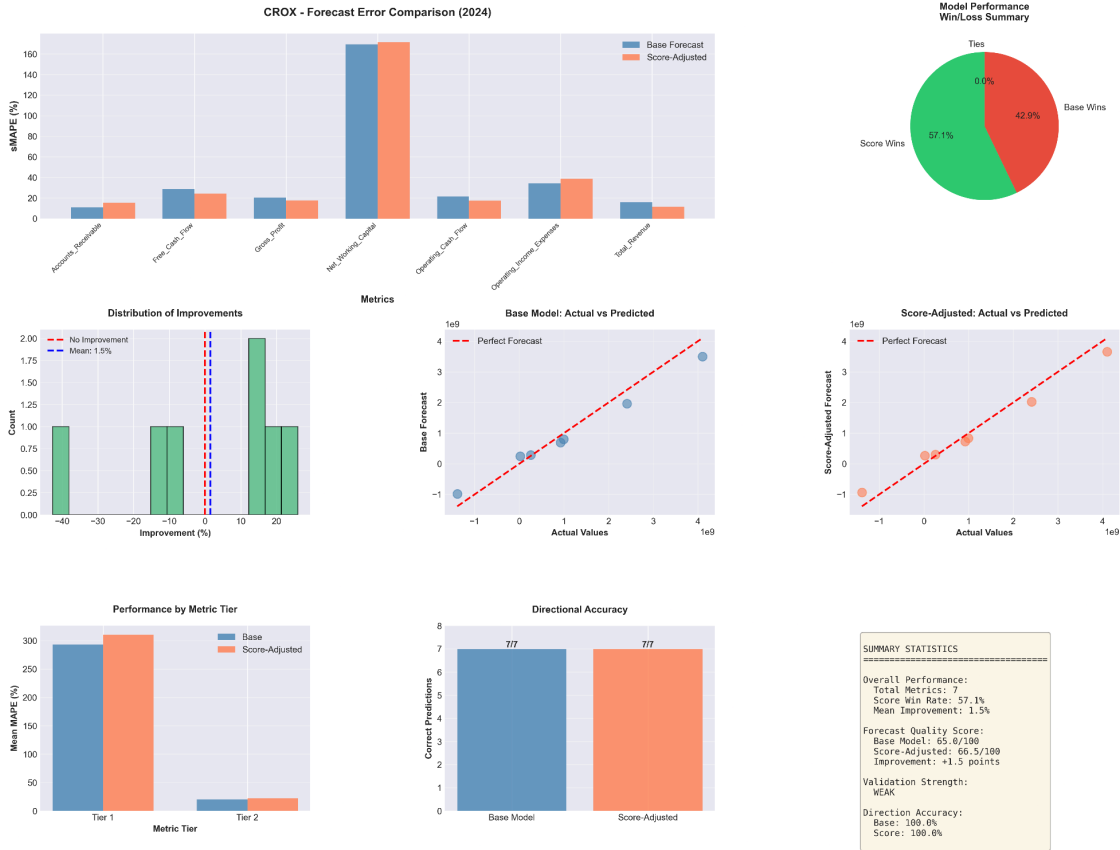
END OF REPORT

=====

=====

CROX (LINEAR):

Forecast Validation Analysis: CROX (2024)



FORECAST VALIDATION REPORT

Company: CROX
Forecast Year: 2024
Analysis Date: 2025-11-20 17:57:34

EXECUTIVE SUMMARY

Validation Strength: WEAK

Score-Adjusted method shows weak validation. Qualitative factors provide marginal improvements.

Overall Win Rate: 57.1% (Score-Adjusted wins)

Mean Improvement: 1.5%

Forecast Quality Scores:

- Base Model: 65.0/100
- Score-Adjusted Model: 66.5/100
- Improvement: +1.5 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 175.97%
- Median MAPE: 19.52%
- Mean sMAPE: 43.07%

Score-Adjusted Model:

- Mean MAPE: 186.92%
- Median MAPE: 16.79%
- Mean sMAPE: 42.31%

Win/Loss Record:

- Score-Adjusted Wins: 4
- Base Model Wins: 3
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 293.04%
- Score Mean MAPE: 310.77%
- Mean Improvement: 12.65%
- Score Win Rate: 75.0%
- Wins: Score=3, Base=1

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 19.87%
- Score Mean MAPE: 21.79%
- Mean Improvement: -13.44%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

STATISTICAL SIGNIFICANCE TESTS

Paired T-Test:

- T-Statistic: 0.4890
- P-Value: 0.6422
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 14.0000
- P-Value: 1.0000
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.1848
- Interpretation: Small effect

METRIC-BY-METRIC BREAKDOWN

Accounts_Receivable (Tier 2):

Actual Value: 257,657,000.00

Base Forecast: 287,978,434.00 (Error: +30,321,434.00)

Score Forecast: 300,914,426.00 (Error: +43,257,426.00)

Accuracy:

- Base sMAPE: 11.11%
- Score sMAPE: 15.49%
- Improvement: -42.66%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 923,139,000.00

Base Forecast: 692,653,048.00 (Error: -230,485,952.00)

Score Forecast: 723,767,023.00 (Error: -199,371,977.00)

Accuracy:

- Base sMAPE: 28.53%
- Score sMAPE: 24.21%
- Improvement: +13.50%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 2,410,258,000.00

Base Forecast: 1,963,395,030.00 (Error: -446,862,970.00)

Score Forecast: 2,022,218,345.00 (Error: -388,039,655.00)

Accuracy:

- Base sMAPE: 20.43%
- Score sMAPE: 17.51%
- Improvement: +13.16%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 20,259,000.00

Base Forecast: 245,722,513.00 (Error: +225,463,513.00)

Score Forecast: 262,284,211.00 (Error: +242,025,211.00)

Accuracy:

- Base sMAPE: 169.53%
- Score sMAPE: 171.32%
- Improvement: -7.35%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 992,486,000.00

Base Forecast: 798,765,963.00 (Error: -193,720,037.00)

Score Forecast: 834,646,530.00 (Error: -157,839,470.00)

Accuracy:

- Base sMAPE: 21.63%
- Score sMAPE: 17.28%
- Improvement: +18.52%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -1,388,347,000.00

Base Forecast: -981,469,620.00 (Error: +406,877,380.00)

Score Forecast: -937,382,005.00 (Error: +450,964,995.00)

Accuracy:

- Base sMAPE: 34.34%
- Score sMAPE: 38.78%
- Improvement: -10.84%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 4,102,108,000.00

Base Forecast: 3,496,668,166.00 (Error: -605,439,834.00)

Score Forecast: 3,653,738,500.00 (Error: -448,369,500.00)

Accuracy:

- Base sMAPE: 15.94%
- Score sMAPE: 11.56%
- Improvement: +25.94%

Direction Correct: Base=True, Score=True
Winner: Score

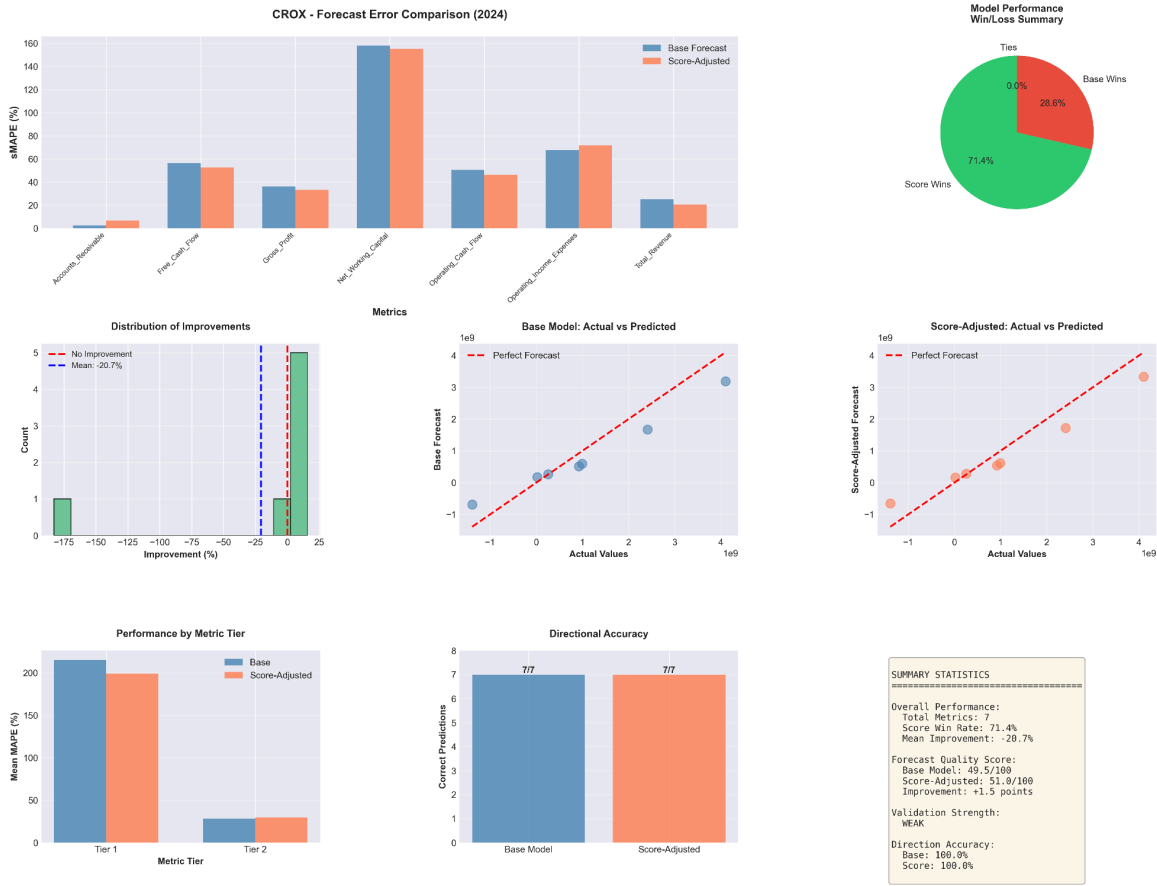
```
=====
=====
RECOMMENDATIONS
=====
=====
```

- △ Weak validation. Qualitative adjustments show marginal impact. Recommend:
1. Review and refine the qualitative scoring methodology
 2. Investigate if certain metric types benefit more from adjustments
 3. Consider whether the base model is already capturing qualitative factors
 4. Test alternative adjustment algorithms

```
=====
=====
END OF REPORT
=====
=====
```


CROX (XGBOOST):

Forecast Validation Analysis: CROX (2024)



FORECAST VALIDATION REPORT

Company: CROX
Forecast Year: 2024
Analysis Date: 2025-11-20 18:17:10

EXECUTIVE SUMMARY

=====

Validation Strength: WEAK

Score-Adjusted method shows weak validation. Qualitative factors provide marginal improvements.

Overall Win Rate: 71.4% (Score-Adjusted wins)

Mean Improvement: -20.7%

Forecast Quality Scores:

- Base Model: 49.5/100
- Score-Adjusted Model: 51.0/100
- Improvement: +1.5 points

=====

DETAILED METRICS COMPARISON

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 135.01%
- Median MAPE: 40.37%
- Mean sMAPE: 56.66%

Score-Adjusted Model:

- Mean MAPE: 126.22%
- Median MAPE: 37.70%
- Mean sMAPE: 55.28%

Win/Loss Record:

- Score-Adjusted Wins: 5
- Base Model Wins: 2
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 215.34%
- Score Mean MAPE: 198.77%
- Mean Improvement: 8.91%
- Score Win Rate: 100.0%
- Wins: Score=4, Base=0

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 27.90%
- Score Mean MAPE: 29.48%
- Mean Improvement: -60.11%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 0.9424
- P-Value: 0.3824
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 10.0000
- P-Value: 0.5781
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.3562
- Interpretation: Small effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 257,657,000.00

Base Forecast: 264,150,312.00 (Error: +6,493,312.00)

Score Forecast: 276,015,944.00 (Error: +18,358,944.00)

Accuracy:

- Base sMAPE: 2.49%
- Score sMAPE: 6.88%
- Improvement: -182.74%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 923,139,000.00

Base Forecast: 515,817,836.00 (Error: -407,321,164.00)

Score Forecast: 538,988,374.00 (Error: -384,150,626.00)

Accuracy:

- Base sMAPE: 56.61%
- Score sMAPE: 52.55%
- Improvement: +5.69%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 2,410,258,000.00

Base Forecast: 1,672,517,408.00 (Error: -737,740,592.00)

Score Forecast: 1,722,626,029.00 (Error: -687,631,971.00)

Accuracy:

- Base sMAPE: 36.14%
- Score sMAPE: 33.28%
- Improvement: +6.79%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 20,259,000.00

Base Forecast: 173,138,214.00 (Error: +152,879,214.00)

Score Forecast: 161,468,698.00 (Error: +141,209,698.00)

Accuracy:

- Base sMAPE: 158.10%
- Score sMAPE: 155.41%
- Improvement: +7.63%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 992,486,000.00

Base Forecast: 591,775,094.00 (Error: -400,710,906.00)

Score Forecast: 618,357,631.00 (Error: -374,128,369.00)

Accuracy:

- Base sMAPE: 50.59%
- Score sMAPE: 46.45%
- Improvement: +6.63%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -1,388,347,000.00

Base Forecast: -686,203,513.00 (Error: +702,143,487.00)

Score Forecast: -655,379,251.00 (Error: +732,967,749.00)

Accuracy:

- Base sMAPE: 67.69%
- Score sMAPE: 71.73%
- Improvement: -4.39%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 4,102,108,000.00

Base Forecast: 3,189,573,350.00 (Error: -912,534,650.00)

Score Forecast: 3,332,848,985.00 (Error: -769,259,015.00)

Accuracy:

- Base sMAPE: 25.03%
- Score sMAPE: 20.69%
- Improvement: +15.70%

Direction Correct: Base=True, Score=True

Winner: Score

```
=====
=====
RECOMMENDATIONS
=====
=====
```

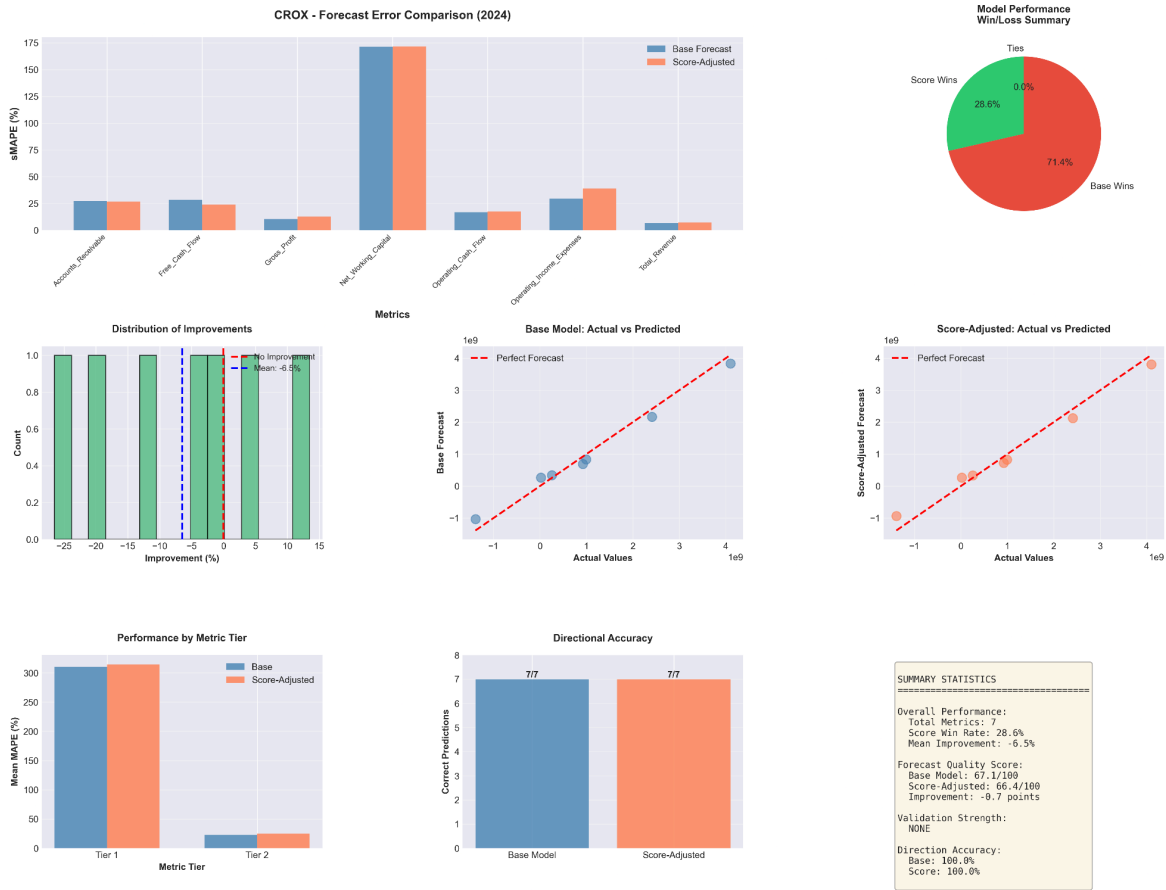
△ Weak validation. Qualitative adjustments show marginal impact. Recommend:

1. Review and refine the qualitative scoring methodology
2. Investigate if certain metric types benefit more from adjustments
3. Consider whether the base model is already capturing qualitative factors
4. Test alternative adjustment algorithms

```
=====
=====
END OF REPORT
=====
=====
```


CROX (ARIMA):

Forecast Validation Analysis: CROX (2024)



FORECAST VALIDATION REPORT

Company: CROX
Forecast Year: 2024
Analysis Date: 2025-11-20 18:03:55

EXECUTIVE SUMMARY

Validation Strength: NONE

Score-Adjusted method does not validate. Base model performs as well or better.

Overall Win Rate: 28.6% (Score-Adjusted wins)

Mean Improvement: -6.5%

Forecast Quality Scores:

- Base Model: 67.1/100
- Score-Adjusted Model: 66.4/100
- Improvement: -0.7 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 187.21%
- Median MAPE: 24.97%
- Mean sMAPE: 41.59%

Score-Adjusted Model:

- Mean MAPE: 190.63%
- Median MAPE: 21.60%
- Mean sMAPE: 42.79%

Win/Loss Record:

- Score-Adjusted Wins: 2
- Base Model Wins: 5
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 310.71%
- Score Mean MAPE: 314.74%
- Mean Improvement: -0.64%
- Score Win Rate: 25.0%
- Wins: Score=1, Base=3

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 22.54%
- Score Mean MAPE: 25.15%
- Mean Improvement: -14.29%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

STATISTICAL SIGNIFICANCE TESTS

Paired T-Test:

- T-Statistic: -0.7623
- P-Value: 0.4748
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 8.0000
- P-Value: 0.3750
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: -0.2881
- Interpretation: Small effect

METRIC-BY-METRIC BREAKDOWN

Accounts_Receivable (Tier 2):

Actual Value: 257,657,000.00

Base Forecast: 339,383,096.00 (Error: +81,726,096.00)

Score Forecast: 336,896,775.00 (Error: +79,239,775.00)

Accuracy:

- Base sMAPE: 27.38%

- Score sMAPE: 26.66%

- Improvement: +3.04%

Direction Correct: Base=True, Score=True

Winner: Score

Free_Cash_Flow (Tier 1):

Actual Value: 923,139,000.00

Base Forecast: 692,653,048.00 (Error: -230,485,952.00)

Score Forecast: 723,767,023.00 (Error: -199,371,977.00)

Accuracy:

- Base sMAPE: 28.53%

- Score sMAPE: 24.21%

- Improvement: +13.50%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 2,410,258,000.00

Base Forecast: 2,169,776,645.00 (Error: -240,481,355.00)

Score Forecast: 2,123,043,996.00 (Error: -287,214,004.00)

Accuracy:

- Base sMAPE: 10.50%

- Score sMAPE: 12.67%

- Improvement: -19.43%

Direction Correct: Base=True, Score=True

Winner: Base

Net_Working_Capital (Tier 1):

Actual Value: 20,259,000.00

Base Forecast: 262,519,031.00 (Error: +242,260,031.00)

Score Forecast: 266,202,173.00 (Error: +245,943,173.00)

Accuracy:

- Base sMAPE: 171.34%
- Score sMAPE: 171.71%
- Improvement: -1.52%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 992,486,000.00

Base Forecast: 837,771,767.00 (Error: -154,714,233.00)

Score Forecast: 831,634,251.00 (Error: -160,851,749.00)

Accuracy:

- Base sMAPE: 16.91%
- Score sMAPE: 17.64%
- Improvement: -3.97%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Income_Expenses (Tier 2):

Actual Value: -1,388,347,000.00

Base Forecast: -1,028,489,602.00 (Error: +359,857,398.00)

Score Forecast: -933,175,356.00 (Error: +455,171,644.00)

Accuracy:

- Base sMAPE: 29.78%
- Score sMAPE: 39.21%
- Improvement: -26.49%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 4,102,108,000.00

Base Forecast: 3,836,267,929.00 (Error: -265,840,071.00)

Score Forecast: 3,808,163,430.00 (Error: -293,944,570.00)

Accuracy:

- Base sMAPE: 6.70%
- Score sMAPE: 7.43%
- Improvement: -10.57%

Direction Correct: Base=True, Score=True
Winner: Base

=====

=====

RECOMMENDATIONS

=====

=====

- X No validation. Base model performs as well or better. Recommend:
1. Critically review the qualitative scoring framework
 2. Check if qualitative factors are already reflected in historical data
 3. Consider that market efficiency may limit qualitative alpha
 4. Explore if timing or sector-specific factors affected results

=====

=====

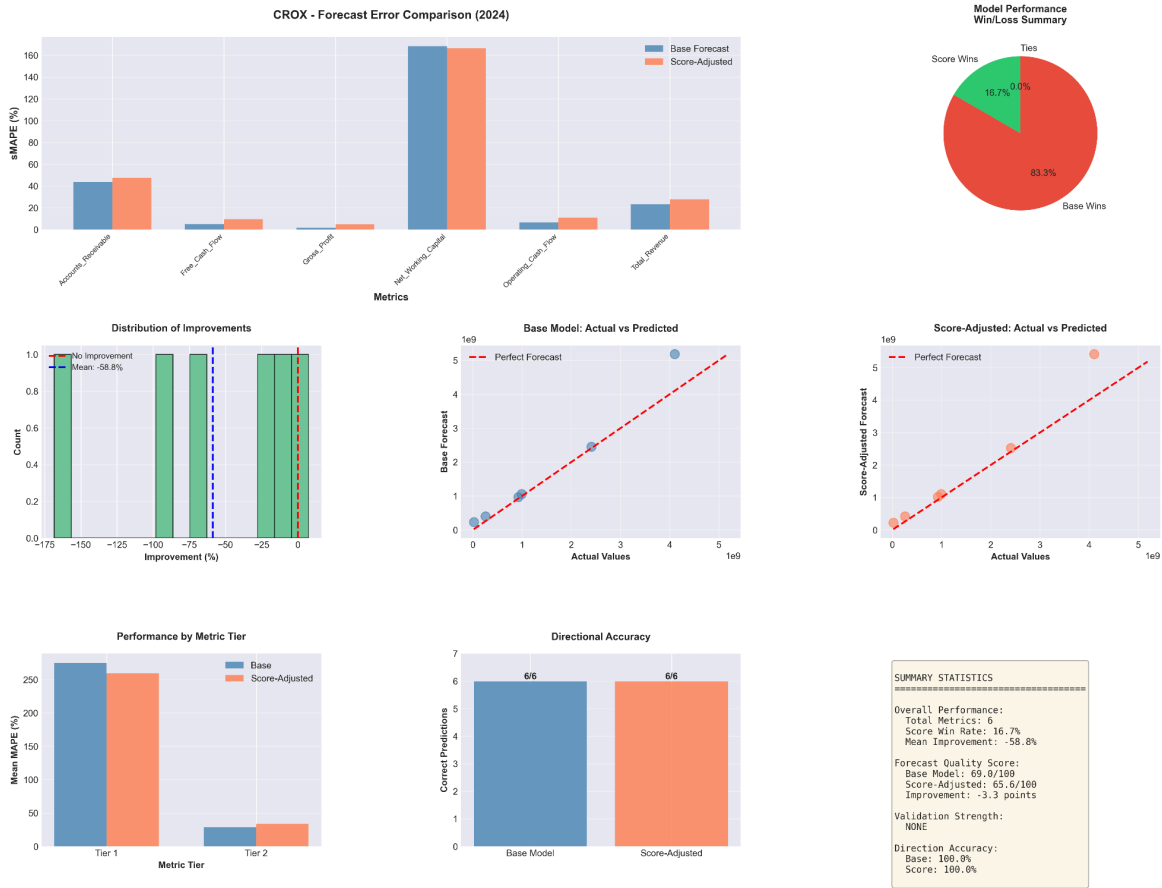
END OF REPORT

=====

=====

CROX (WINTERS-HOLT):

Forecast Validation Analysis: CROX (2024)



FORECAST VALIDATION REPORT

Company: CROX
Forecast Year: 2024
Analysis Date: 2025-11-20 18:24:27

EXECUTIVE SUMMARY

=====

Validation Strength: NONE

Score-Adjusted method does not validate. Base model performs as well or better.

Overall Win Rate: 16.7% (Score-Adjusted wins)

Mean Improvement: -58.8%

Forecast Quality Scores:

- Base Model: 69.0/100
- Score-Adjusted Model: 65.6/100
- Improvement: -3.3 points

=====

DETAILED METRICS COMPARISON

=====

Total Metrics Analyzed: 6

Accuracy Metrics:

Base Model:

- Mean MAPE: 192.98%
- Median MAPE: 16.71%
- Mean sMAPE: 41.48%

Score-Adjusted Model:

- Mean MAPE: 184.14%
- Median MAPE: 21.95%
- Mean sMAPE: 44.50%

Win/Loss Record:

- Score-Adjusted Wins: 1
- Base Model Wins: 5
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 275.13%
- Score Mean MAPE: 259.36%
- Mean Improvement: -42.90%
- Score Win Rate: 25.0%
- Wins: Score=1, Base=3

Tier 2 (Important) Metrics:

- Count: 2
- Base Mean MAPE: 28.69%
- Score Mean MAPE: 33.71%
- Mean Improvement: -90.46%
- Score Win Rate: 0.0%
- Wins: Score=0, Base=2

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: -2.8813
- P-Value: 0.0345
- Significant at $\alpha=0.05$: True
- Interpretation: Score-Adjusted significantly different

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 1.0000
- P-Value: 0.0625
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: -1.1763
- Interpretation: Large effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 257,657,000.00

Base Forecast: 400,842,040.00 (Error: +143,185,040.00)

Score Forecast: 418,847,865.00 (Error: +161,190,865.00)

Accuracy:

- Base sMAPE: 43.49%
- Score sMAPE: 47.65%
- Improvement: -12.58%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 923,139,000.00

Base Forecast: 972,400,456.00 (Error: +49,261,456.00)

Score Forecast: 1,016,080,685.00 (Error: +92,941,685.00)

Accuracy:

- Base sMAPE: 5.20%
- Score sMAPE: 9.59%
- Improvement: -88.67%

Direction Correct: Base=True, Score=True

Winner: Base

Gross_Profit (Tier 2):

Actual Value: 2,410,258,000.00

Base Forecast: 2,453,931,563.00 (Error: +43,673,563.00)

Score Forecast: 2,527,451,353.00 (Error: +117,193,353.00)

Accuracy:

- Base sMAPE: 1.80%
- Score sMAPE: 4.75%
- Improvement: -168.34%

Direction Correct: Base=True, Score=True

Winner: Base

Net_Working_Capital (Tier 1):

Actual Value: 20,259,000.00
Base Forecast: 235,359,906.00 (Error: +215,100,906.00)
Score Forecast: 219,496,648.00 (Error: +199,237,648.00)

Accuracy:

- Base sMAPE: 168.30%
- Score sMAPE: 166.20%
- Improvement: +7.37%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 992,486,000.00
Base Forecast: 1,061,783,431.00 (Error: +69,297,431.00)
Score Forecast: 1,109,478,743.00 (Error: +116,992,743.00)

Accuracy:

- Base sMAPE: 6.75%
- Score sMAPE: 11.13%
- Improvement: -68.83%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 4,102,108,000.00
Base Forecast: 5,186,584,267.00 (Error: +1,084,476,267.00)
Score Forecast: 5,419,565,632.00 (Error: +1,317,457,632.00)

Accuracy:

- Base sMAPE: 23.35%
- Score sMAPE: 27.67%
- Improvement: -21.48%

Direction Correct: Base=True, Score=True

Winner: Base

=====

RECOMMENDATIONS

=====

x No validation. Base model performs as well or better. Recommend:

1. Critically review the qualitative scoring framework
2. Check if qualitative factors are already reflected in historical data
3. Consider that market efficiency may limit qualitative alpha
4. Explore if timing or sector-specific factors affected results

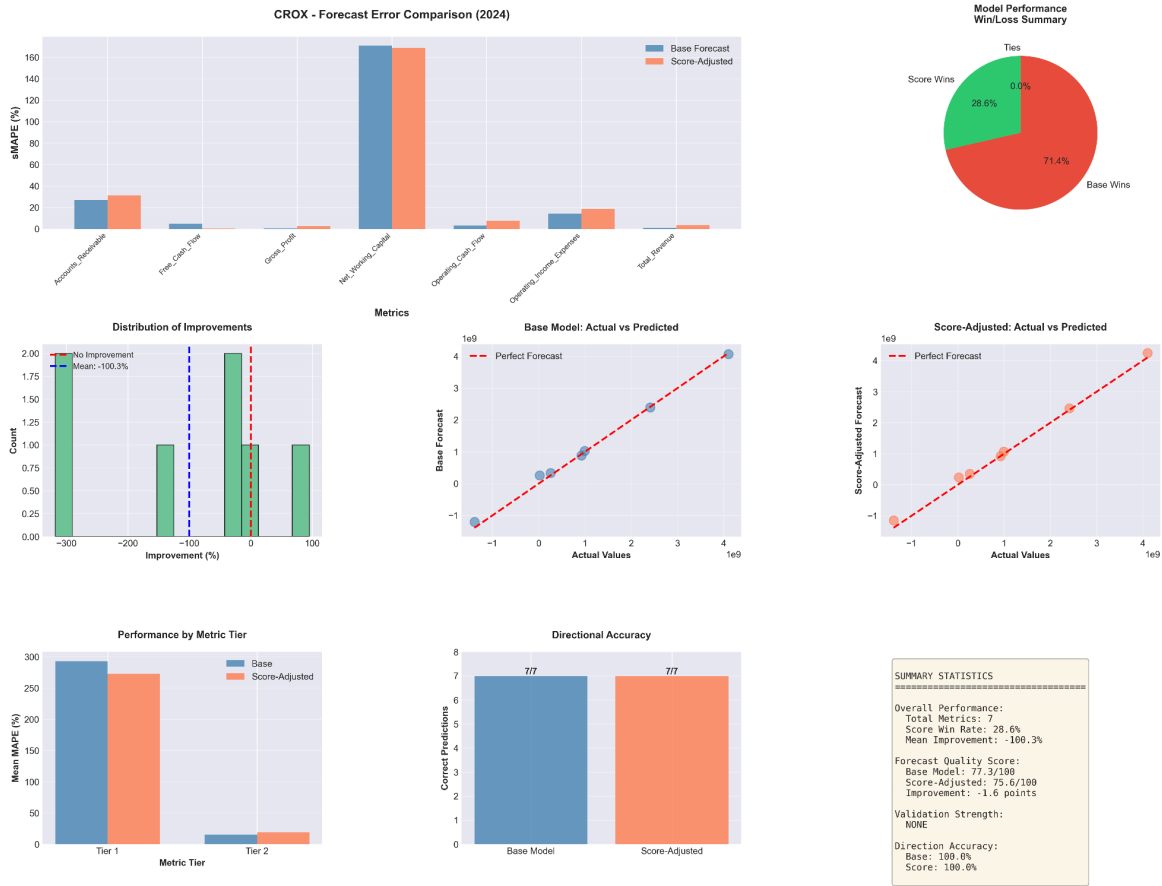
=====

END OF REPORT

=====

CROX (LINEAR + WINTERS-HOLT):

Forecast Validation Analysis: CROX (2024)



FORECAST VALIDATION REPORT

Company: CROX
Forecast Year: 2024
Analysis Date: 2025-11-20 18:50:57

EXECUTIVE SUMMARY

Validation Strength: NONE

Score-Adjusted method does not validate. Base model performs as well or better.

Overall Win Rate: 28.6% (Score-Adjusted wins)

Mean Improvement: -100.3%

Forecast Quality Scores:

- Base Model: 77.3/100
- Score-Adjusted Model: 75.6/100
- Improvement: -1.6 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 173.93%
- Median MAPE: 4.51%
- Mean sMAPE: 31.57%

Score-Adjusted Model:

- Mean MAPE: 163.84%
- Median MAPE: 7.85%
- Mean sMAPE: 33.18%

Win/Loss Record:

- Score-Adjusted Wins: 2
- Base Model Wins: 5
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 293.16%
- Score Mean MAPE: 272.61%
- Mean Improvement: -84.00%
- Score Win Rate: 50.0%
- Wins: Score=2, Base=2

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 14.94%
- Score Mean MAPE: 18.81%
- Mean Improvement: -122.13%
- Score Win Rate: 0.0%
- Wins: Score=0, Base=3

STATISTICAL SIGNIFICANCE TESTS

Paired T-Test:

- T-Statistic: -1.2201
- P-Value: 0.2682
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 8.0000
- P-Value: 0.3750
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: -0.4612
- Interpretation: Small effect

METRIC-BY-METRIC BREAKDOWN

Accounts_Receivable (Tier 2):

Actual Value: 257,657,000.00

Base Forecast: 337,561,807.00 (Error: +79,904,807.00)

Score Forecast: 352,725,083.00 (Error: +95,068,083.00)

Accuracy:

- Base sMAPE: 26.85%

- Score sMAPE: 31.15%

- Improvement: -18.98%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 923,139,000.00

Base Forecast: 881,524,650.00 (Error: -41,614,350.00)

Score Forecast: 921,122,737.00 (Error: -2,016,263.00)

Accuracy:

- Base sMAPE: 4.61%

- Score sMAPE: 0.22%

- Improvement: +95.15%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 2,410,258,000.00

Base Forecast: 2,396,397,503.00 (Error: -13,860,497.00)

Score Forecast: 2,468,193,573.00 (Error: +57,935,573.00)

Accuracy:

- Base sMAPE: 0.58%

- Score sMAPE: 2.38%

- Improvement: -317.99%

Direction Correct: Base=True, Score=True

Winner: Base

Net_Working_Capital (Tier 1):

Actual Value: 20,259,000.00

Base Forecast: 256,078,955.00 (Error: +235,819,955.00)

Score Forecast: 238,819,234.00 (Error: +218,560,234.00)

Accuracy:

- Base sMAPE: 170.68%
- Score sMAPE: 168.72%
- Improvement: +7.32%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 992,486,000.00

Base Forecast: 1,024,420,147.00 (Error: +31,934,147.00)

Score Forecast: 1,070,437,100.00 (Error: +77,951,100.00)

Accuracy:

- Base sMAPE: 3.17%
- Score sMAPE: 7.56%
- Improvement: -144.10%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Income_Expenses (Tier 2):

Actual Value: -1,388,347,000.00

Base Forecast: -1,204,529,370.00 (Error: +183,817,630.00)

Score Forecast: -1,150,421,911.00 (Error: +237,925,089.00)

Accuracy:

- Base sMAPE: 14.18%
- Score sMAPE: 18.74%
- Improvement: -29.44%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 4,102,108,000.00

Base Forecast: 4,065,171,670.00 (Error: -36,936,330.00)

Score Forecast: 4,247,779,182.00 (Error: +145,671,182.00)

Accuracy:

- Base sMAPE: 0.90%
- Score sMAPE: 3.49%
- Improvement: -294.38%

Direction Correct: Base=True, Score=True
Winner: Base

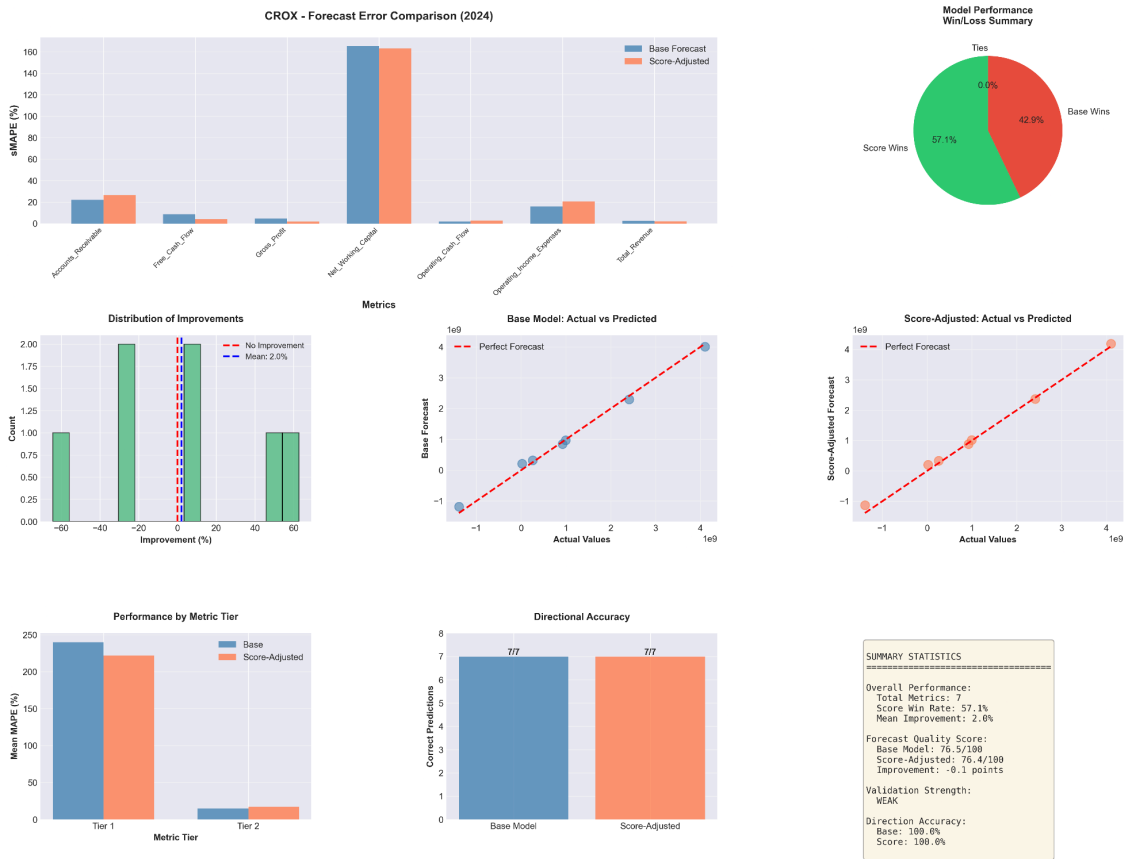
```
=====
=====
RECOMMENDATIONS
=====
=====
```

- X** No validation. Base model performs as well or better. Recommend:
1. Critically review the qualitative scoring framework
 2. Check if qualitative factors are already reflected in historical data
 3. Consider that market efficiency may limit qualitative alpha
 4. Explore if timing or sector-specific factors affected results

```
=====
=====
END OF REPORT
=====
=====
```


CROX (LINEAR + WINTERS-HOLT + XGBOOST):

Forecast Validation Analysis: CROX (2024)



FORECAST VALIDATION REPORT

Company: CROX
Forecast Year: 2024
Analysis Date: 2025-11-20 18:51:36

EXECUTIVE SUMMARY

Validation Strength: WEAK

Score-Adjusted method shows weak validation. Qualitative factors provide marginal improvements.

Overall Win Rate: 57.1% (Score-Adjusted wins)

Mean Improvement: 2.0%

Forecast Quality Scores:

- Base Model: 76.5/100
- Score-Adjusted Model: 76.4/100
- Improvement: -0.1 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 143.50%
- Median MAPE: 8.27%
- Mean sMAPE: 31.49%

Score-Adjusted Model:

- Mean MAPE: 133.88%
- Median MAPE: 4.15%
- Mean sMAPE: 31.50%

Win/Loss Record:

- Score-Adjusted Wins: 4
- Base Model Wins: 3
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 240.09%
- Score Mean MAPE: 221.63%
- Mean Improvement: -0.06%
- Score Win Rate: 75.0%
- Wins: Score=3, Base=1

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 14.70%
- Score Mean MAPE: 16.89%
- Mean Improvement: 4.72%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: -0.0089
- P-Value: 0.9932
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 14.0000
- P-Value: 1.0000
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: -0.0034
- Interpretation: Small effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 257,657,000.00

Base Forecast: 321,383,863.00 (Error: +63,726,863.00)

Score Forecast: 335,820,426.00 (Error: +78,163,426.00)

Accuracy:

- Base sMAPE: 22.01%
- Score sMAPE: 26.34%
- Improvement: -22.65%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 923,139,000.00

Base Forecast: 846,818,805.00 (Error: -76,320,195.00)

Score Forecast: 884,857,905.00 (Error: -38,281,095.00)

Accuracy:

- Base sMAPE: 8.62%
- Score sMAPE: 4.23%
- Improvement: +49.84%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 2,410,258,000.00

Base Forecast: 2,300,296,543.00 (Error: -109,961,457.00)

Score Forecast: 2,369,213,428.00 (Error: -41,044,572.00)

Accuracy:

- Base sMAPE: 4.67%
- Score sMAPE: 1.72%
- Improvement: +62.67%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 20,259,000.00
Base Forecast: 212,348,085.00 (Error: +192,089,085.00)
Score Forecast: 198,035,824.00 (Error: +177,776,824.00)

Accuracy:

- Base sMAPE: 165.16%
- Score sMAPE: 162.88%
- Improvement: +7.45%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 992,486,000.00
Base Forecast: 975,902,905.00 (Error: -16,583,095.00)
Score Forecast: 1,019,740,464.00 (Error: +27,254,464.00)

Accuracy:

- Base sMAPE: 1.68%
- Score sMAPE: 2.71%
- Improvement: -64.35%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Income_Expenses (Tier 2):

Actual Value: -1,388,347,000.00
Base Forecast: -1,182,942,509.00 (Error: +205,404,491.00)
Score Forecast: -1,129,804,731.00 (Error: +258,542,269.00)

Accuracy:

- Base sMAPE: 15.98%
- Score sMAPE: 20.53%
- Improvement: -25.87%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 4,102,108,000.00
Base Forecast: 4,008,882,027.00 (Error: -93,225,973.00)
Score Forecast: 4,188,961,008.00 (Error: +86,853,008.00)

Accuracy:

- Base sMAPE: 2.30%
- Score sMAPE: 2.10%
- Improvement: +6.84%

Direction Correct: Base=True, Score=True

Winner: Score

=====

=====

RECOMMENDATIONS

=====

=====

△ Weak validation. Qualitative adjustments show marginal impact. Recommend:

1. Review and refine the qualitative scoring methodology
2. Investigate if certain metric types benefit more from adjustments
3. Consider whether the base model is already capturing qualitative factors
4. Test alternative adjustment algorithms

=====

=====

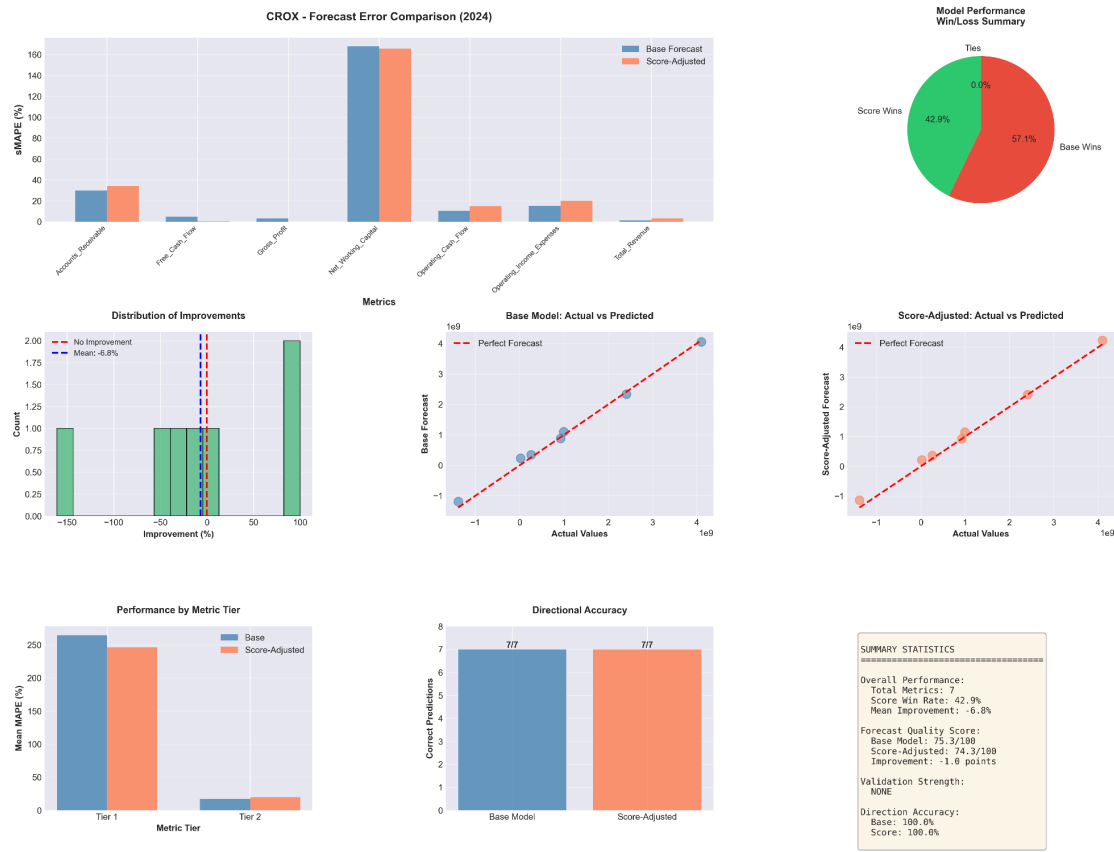
END OF REPORT

=====

=====

CROX (LINEAR + WINTERS-HOLT + ARIMA):

Forecast Validation Analysis: CROX (2024)



FORECAST VALIDATION REPORT

Company: CROX
Forecast Year: 2024
Analysis Date: 2025-11-20 18:52:19

EXECUTIVE SUMMARY

Validation Strength: NONE

Score-Adjusted method does not validate. Base model performs as well or better.

Overall Win Rate: 42.9% (Score-Adjusted wins)

Mean Improvement: -6.8%

Forecast Quality Scores:

- Base Model: 75.3/100
- Score-Adjusted Model: 74.3/100
- Improvement: -1.0 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 158.72%
- Median MAPE: 11.01%
- Mean sMAPE: 33.13%

Score-Adjusted Model:

- Mean MAPE: 149.10%
- Median MAPE: 16.00%
- Mean sMAPE: 33.94%

Win/Loss Record:

- Score-Adjusted Wins: 3
- Base Model Wins: 4
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 264.79%
- Score Mean MAPE: 246.20%
- Mean Improvement: -25.80%
- Score Win Rate: 50.0%
- Wins: Score=2, Base=2

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 17.30%
- Score Mean MAPE: 19.63%
- Mean Improvement: 18.44%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

STATISTICAL SIGNIFICANCE TESTS

Paired T-Test:

- T-Statistic: -0.5537
- P-Value: 0.5998
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 11.0000
- P-Value: 0.6875
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: -0.2093
- Interpretation: Small effect

METRIC-BY-METRIC BREAKDOWN

Accounts_Receivable (Tier 2):

Actual Value: 257,657,000.00

Base Forecast: 347,403,586.00 (Error: +89,746,586.00)

Score Forecast: 363,008,955.00 (Error: +105,351,955.00)

Accuracy:

- Base sMAPE: 29.67%

- Score sMAPE: 33.95%

- Improvement: -17.39%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 923,139,000.00

Base Forecast: 881,694,833.00 (Error: -41,444,167.00)

Score Forecast: 921,300,565.00 (Error: -1,838,435.00)

Accuracy:

- Base sMAPE: 4.59%

- Score sMAPE: 0.20%

- Improvement: +95.56%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 2,410,258,000.00

Base Forecast: 2,340,186,460.00 (Error: -70,071,540.00)

Score Forecast: 2,410,298,446.00 (Error: +40,446.00)

Accuracy:

- Base sMAPE: 2.95%

- Score sMAPE: 0.00%

- Improvement: +99.94%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 20,259,000.00

Base Forecast: 231,441,785.00 (Error: +211,182,785.00)

Score Forecast: 215,842,608.00 (Error: +195,583,608.00)

Accuracy:

- Base sMAPE: 167.80%
- Score sMAPE: 165.68%
- Improvement: +7.39%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 992,486,000.00

Base Forecast: 1,101,762,166.00 (Error: +109,276,166.00)

Score Forecast: 1,151,253,322.00 (Error: +158,767,322.00)

Accuracy:

- Base sMAPE: 10.44%
- Score sMAPE: 14.81%
- Improvement: -45.29%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Income_Expenses (Tier 2):

Actual Value: -1,388,347,000.00

Base Forecast: -1,191,851,175.00 (Error: +196,495,825.00)

Score Forecast: -1,138,313,221.00 (Error: +250,033,779.00)

Accuracy:

- Base sMAPE: 15.23%
- Score sMAPE: 19.79%
- Improvement: -27.25%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 4,102,108,000.00

Base Forecast: 4,051,674,219.00 (Error: -50,433,781.00)

Score Forecast: 4,233,675,425.00 (Error: +131,567,425.00)

Accuracy:

- Base sMAPE: 1.24%
- Score sMAPE: 3.16%
- Improvement: -160.87%

Direction Correct: Base=True, Score=True
Winner: Base

=====

=====

RECOMMENDATIONS

=====

=====

- X No validation. Base model performs as well or better. Recommend:
1. Critically review the qualitative scoring framework
 2. Check if qualitative factors are already reflected in historical data
 3. Consider that market efficiency may limit qualitative alpha
 4. Explore if timing or sector-specific factors affected results

=====

=====

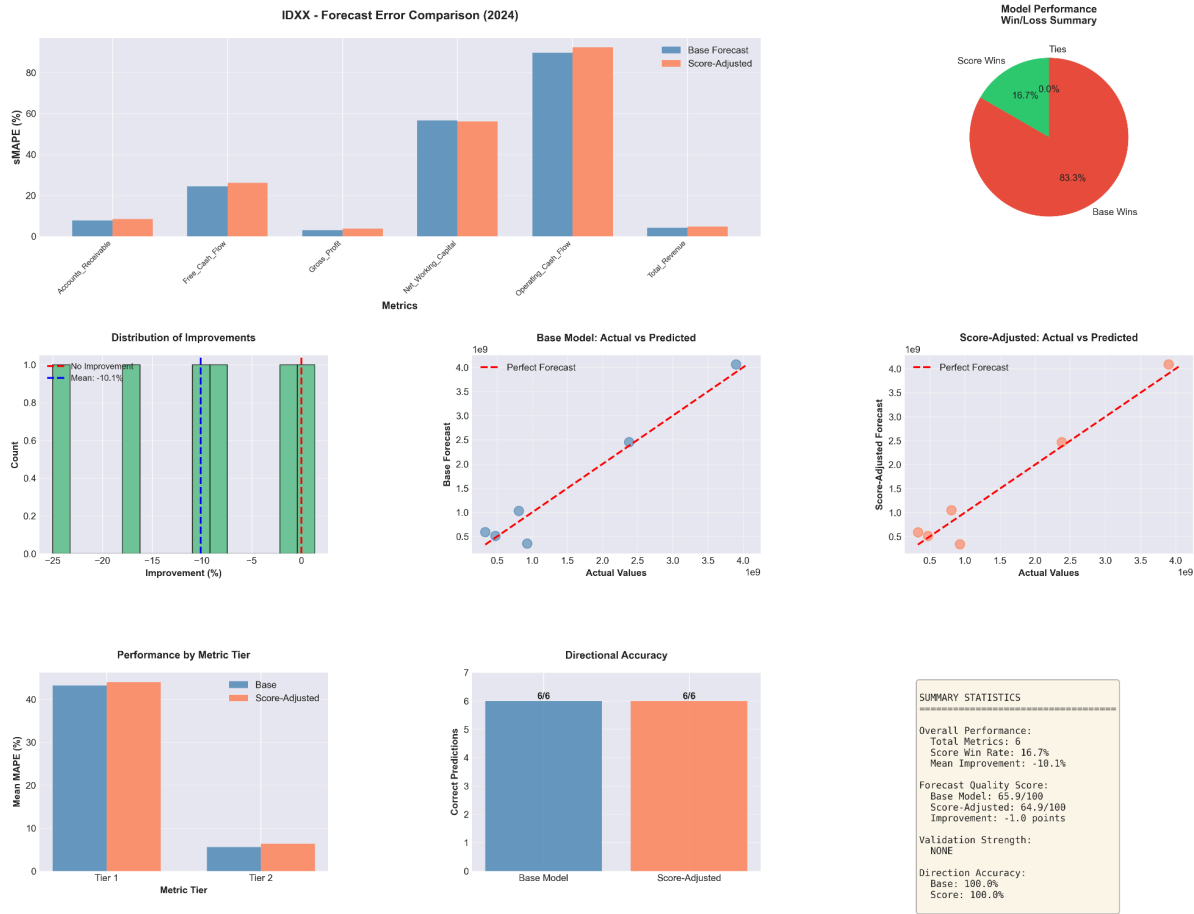
END OF REPORT

=====

=====

IDXX (NAIVE):

Forecast Validation Analysis: IDXX (2024)



FORECAST VALIDATION REPORT

Company: IDXX
Forecast Year: 2024
Analysis Date: 2025-12-12 23:56:05

EXECUTIVE SUMMARY

=====

Validation Strength: NONE

Score-Adjusted method does not validate. Base model performs as well or better.

Overall Win Rate: 16.7% (Score-Adjusted wins)

Mean Improvement: -10.1%

Forecast Quality Scores:

- Base Model: 65.9/100
- Score-Adjusted Model: 64.9/100
- Improvement: -1.0 points

=====

DETAILED METRICS COMPARISON

=====

Total Metrics Analyzed: 6

Accuracy Metrics:

Base Model:

- Mean MAPE: 30.69%
- Median MAPE: 17.89%
- Mean sMAPE: 30.94%

Score-Adjusted Model:

- Mean MAPE: 31.46%
- Median MAPE: 19.42%
- Mean sMAPE: 31.92%

Win/Loss Record:

- Score-Adjusted Wins: 1
- Base Model Wins: 5
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 43.25%
- Score Mean MAPE: 44.02%
- Mean Improvement: -6.44%
- Score Win Rate: 25.0%
- Wins: Score=1, Base=3

Tier 2 (Important) Metrics:

- Count: 2
- Base Mean MAPE: 5.56%
- Score Mean MAPE: 6.35%
- Mean Improvement: -17.49%
- Score Win Rate: 0.0%
- Wins: Score=0, Base=2

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: -2.2828
- P-Value: 0.0713
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 1.0000
- P-Value: 0.0625
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: -0.9319
- Interpretation: Large effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 473,575,000.00

Base Forecast: 511,628,516.00 (Error: +38,053,516.00)

Score Forecast: 515,421,362.00 (Error: +41,846,362.00)

Accuracy:

- Base sMAPE: 7.73%
- Score sMAPE: 8.46%
- Improvement: -9.97%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 808,079,000.00

Base Forecast: 1,032,331,670.00 (Error: +224,252,670.00)

Score Forecast: 1,050,493,357.00 (Error: +242,414,357.00)

Accuracy:

- Base sMAPE: 24.37%
- Score sMAPE: 26.09%
- Improvement: -8.10%

Direction Correct: Base=True, Score=True

Winner: Base

Gross_Profit (Tier 2):

Actual Value: 2,378,927,000.00

Base Forecast: 2,452,356,991.00 (Error: +73,429,991.00)

Score Forecast: 2,470,724,080.00 (Error: +91,797,080.00)

Accuracy:

- Base sMAPE: 3.04%
- Score sMAPE: 3.79%
- Improvement: -25.01%

Direction Correct: Base=True, Score=True

Winner: Base

Net_Working_Capital (Tier 1):

Actual Value: 331,981,000.00
Base Forecast: 594,618,489.00 (Error: +262,637,489.00)
Score Forecast: 591,054,685.00 (Error: +259,073,685.00)

Accuracy:

- Base sMAPE: 56.69%
- Score sMAPE: 56.14%
- Improvement: +1.36%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 929,001,000.00
Base Forecast: 353,742,694.00 (Error: -575,258,306.00)
Score Forecast: 342,619,609.00 (Error: -586,381,391.00)

Accuracy:

- Base sMAPE: 89.69%
- Score sMAPE: 92.23%
- Improvement: -1.93%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 3,897,504,000.00
Base Forecast: 4,061,758,865.00 (Error: +164,254,865.00)
Score Forecast: 4,089,815,276.00 (Error: +192,311,276.00)

Accuracy:

- Base sMAPE: 4.13%
- Score sMAPE: 4.82%
- Improvement: -17.08%

Direction Correct: Base=True, Score=True

Winner: Base

=====

RECOMMENDATIONS

=====

x No validation. Base model performs as well or better. Recommend:

1. Critically review the qualitative scoring framework
2. Check if qualitative factors are already reflected in historical data
3. Consider that market efficiency may limit qualitative alpha
4. Explore if timing or sector-specific factors affected results

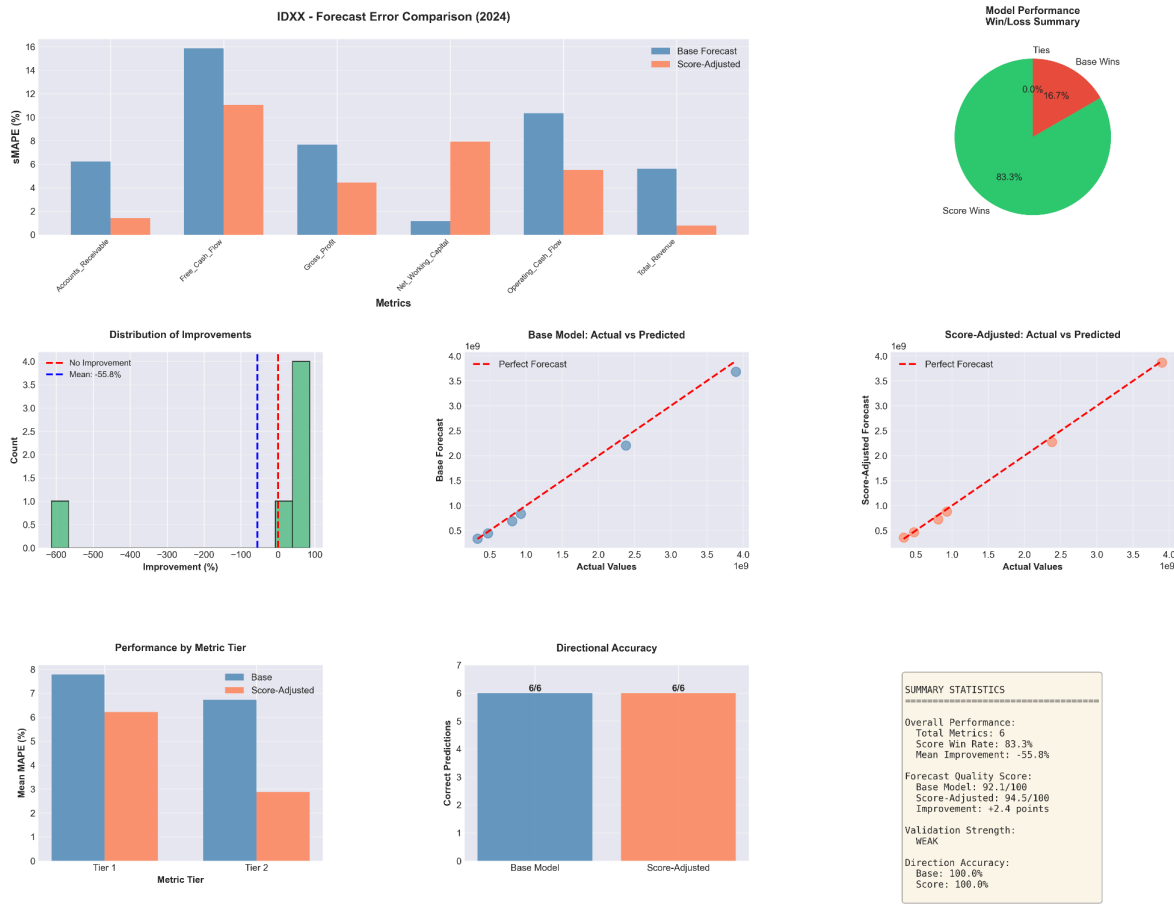
=====

END OF REPORT

=====

IDXX (LINEAR):

Forecast Validation Analysis: IDXX (2024)



FORECAST VALIDATION REPORT

Company: IDXX
Forecast Year: 2024
Analysis Date: 2025-12-12 23:56:22

EXECUTIVE SUMMARY

=====

Validation Strength: WEAK

Score-Adjusted method shows weak validation. Qualitative factors provide marginal improvements.

Overall Win Rate: 83.3% (Score-Adjusted wins)

Mean Improvement: -55.8%

Forecast Quality Scores:

- Base Model: 92.1/100
- Score-Adjusted Model: 94.5/100
- Improvement: +2.4 points

=====

DETAILED METRICS COMPARISON

=====

Total Metrics Analyzed: 6

Accuracy Metrics:

Base Model:

- Mean MAPE: 7.43%
- Median MAPE: 6.72%
- Mean sMAPE: 7.81%

Score-Adjusted Model:

- Mean MAPE: 5.10%
- Median MAPE: 4.85%
- Mean sMAPE: 5.19%

Win/Loss Record:

- Score-Adjusted Wins: 5
- Base Model Wins: 1
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 7.78%
- Score Mean MAPE: 6.21%
- Mean Improvement: -113.22%
- Score Win Rate: 75.0%
- Wins: Score=3, Base=1

Tier 2 (Important) Metrics:

- Count: 2
- Base Mean MAPE: 6.72%
- Score Mean MAPE: 2.87%
- Mean Improvement: 58.99%
- Score Win Rate: 100.0%
- Wins: Score=2, Base=0

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 1.3844
- P-Value: 0.2248
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 6.0000
- P-Value: 0.4375
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.5652
- Interpretation: Medium effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 473,575,000.00

Base Forecast: 444,878,270.00 (Error: -28,696,730.00)

Score Forecast: 466,860,595.00 (Error: -6,714,405.00)

Accuracy:

- Base sMAPE: 6.25%
- Score sMAPE: 1.43%
- Improvement: +76.60%

Direction Correct: Base=True, Score=True

Winner: Score

Free_Cash_Flow (Tier 1):

Actual Value: 808,079,000.00

Base Forecast: 689,396,738.00 (Error: -118,682,262.00)

Score Forecast: 723,461,210.00 (Error: -84,617,790.00)

Accuracy:

- Base sMAPE: 15.85%
- Score sMAPE: 11.05%
- Improvement: +28.70%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 2,378,927,000.00

Base Forecast: 2,203,396,961.00 (Error: -175,530,039.00)

Score Forecast: 2,276,012,111.00 (Error: -102,914,889.00)

Accuracy:

- Base sMAPE: 7.66%
- Score sMAPE: 4.42%
- Improvement: +41.37%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 331,981,000.00

Base Forecast: 335,817,320.00 (Error: +3,836,320.00)

Score Forecast: 359,324,533.00 (Error: +27,343,533.00)

Accuracy:

- Base sMAPE: 1.15%
- Score sMAPE: 7.91%
- Improvement: -612.75%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 929,001,000.00

Base Forecast: 837,616,319.00 (Error: -91,384,681.00)

Score Forecast: 879,004,617.00 (Error: -49,996,383.00)

Accuracy:

- Base sMAPE: 10.35%
- Score sMAPE: 5.53%
- Improvement: +45.29%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 3,897,504,000.00

Base Forecast: 3,685,502,622.00 (Error: -212,001,378.00)

Score Forecast: 3,867,610,677.00 (Error: -29,893,323.00)

Accuracy:

- Base sMAPE: 5.59%
- Score sMAPE: 0.77%
- Improvement: +85.90%

Direction Correct: Base=True, Score=True

Winner: Score

=====

RECOMMENDATIONS

=====

=====

△ Weak validation. Qualitative adjustments show marginal impact. Recommend:

1. Review and refine the qualitative scoring methodology
2. Investigate if certain metric types benefit more from adjustments
3. Consider whether the base model is already capturing qualitative factors
4. Test alternative adjustment algorithms

=====

=====

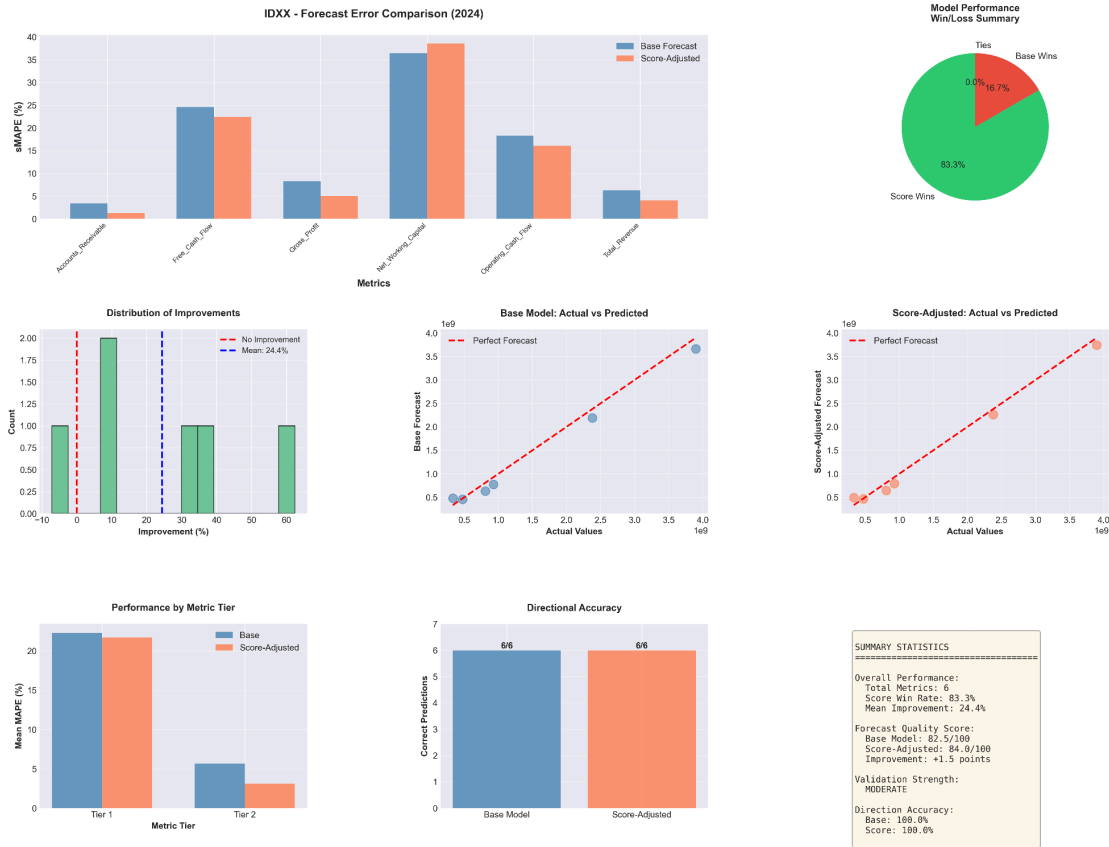
END OF REPORT

=====

=====

IDXX (XGBOOST):

Forecast Validation Analysis: IDXX (2024)



FORECAST VALIDATION REPORT

Company: IDXX
Forecast Year: 2024
Analysis Date: 2025-12-12 23:56:44

EXECUTIVE SUMMARY

Validation Strength: MODERATE

Score-Adjusted method shows moderate validation. Qualitative factors provide meaningful improvements.

Overall Win Rate: 83.3% (Score-Adjusted wins)

Mean Improvement: 24.4%

Forecast Quality Scores:

- Base Model: 82.5/100
- Score-Adjusted Model: 84.0/100
- Improvement: +1.5 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 6

Accuracy Metrics:

Base Model:

- Mean MAPE: 16.77%
- Median MAPE: 12.33%
- Mean sMAPE: 16.21%

Score-Adjusted Model:

- Mean MAPE: 15.50%
- Median MAPE: 9.90%
- Mean sMAPE: 14.58%

Win/Loss Record:

- Score-Adjusted Wins: 5
- Base Model Wins: 1
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 22.32%
- Score Mean MAPE: 21.71%
- Mean Improvement: 11.43%
- Score Win Rate: 75.0%
- Wins: Score=3, Base=1

Tier 2 (Important) Metrics:

- Count: 2
- Base Mean MAPE: 5.67%
- Score Mean MAPE: 3.09%
- Mean Improvement: 50.32%
- Score Win Rate: 100.0%
- Wins: Score=2, Base=0

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 2.1292
- P-Value: 0.0865
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 1.0000
- P-Value: 0.0625
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.8693
- Interpretation: Large effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 473,575,000.00

Base Forecast: 457,444,896.00 (Error: -16,130,104.00)

Score Forecast: 467,508,683.00 (Error: -6,066,317.00)

Accuracy:

- Base sMAPE: 3.47%
- Score sMAPE: 1.29%
- Improvement: +62.39%

Direction Correct: Base=True, Score=True

Winner: Score

Free_Cash_Flow (Tier 1):

Actual Value: 808,079,000.00

Base Forecast: 631,120,064.00 (Error: -176,958,936.00)

Score Forecast: 645,004,705.00 (Error: -163,074,295.00)

Accuracy:

- Base sMAPE: 24.59%
- Score sMAPE: 22.45%
- Improvement: +7.85%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 2,378,927,000.00

Base Forecast: 2,189,969,152.00 (Error: -188,957,848.00)

Score Forecast: 2,262,238,134.00 (Error: -116,688,866.00)

Accuracy:

- Base sMAPE: 8.27%
- Score sMAPE: 5.03%
- Improvement: +38.25%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 331,981,000.00
Base Forecast: 479,976,960.00 (Error: +147,995,960.00)
Score Forecast: 490,536,453.00 (Error: +158,555,453.00)

Accuracy:

- Base sMAPE: 36.45%
- Score sMAPE: 38.55%
- Improvement: -7.13%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 929,001,000.00
Base Forecast: 773,679,168.00 (Error: -155,321,832.00)
Score Forecast: 790,700,109.00 (Error: -138,300,891.00)

Accuracy:

- Base sMAPE: 18.24%
- Score sMAPE: 16.08%
- Improvement: +10.96%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 3,897,504,000.00
Base Forecast: 3,660,952,320.00 (Error: -236,551,680.00)
Score Forecast: 3,741,493,271.00 (Error: -156,010,729.00)

Accuracy:

- Base sMAPE: 6.26%
- Score sMAPE: 4.08%
- Improvement: +34.05%

Direction Correct: Base=True, Score=True

Winner: Score

RECOMMENDATIONS

=====

- Moderate validation achieved. Qualitative adjustments provide meaningful but not dramatic improvements. Recommend:
 1. Continue using score-adjusted approach with refinements
 2. Investigate which specific qualitative factors are most predictive
 3. Consider sector-specific adjustments to scoring methodology
 4. Test on additional companies to confirm consistency

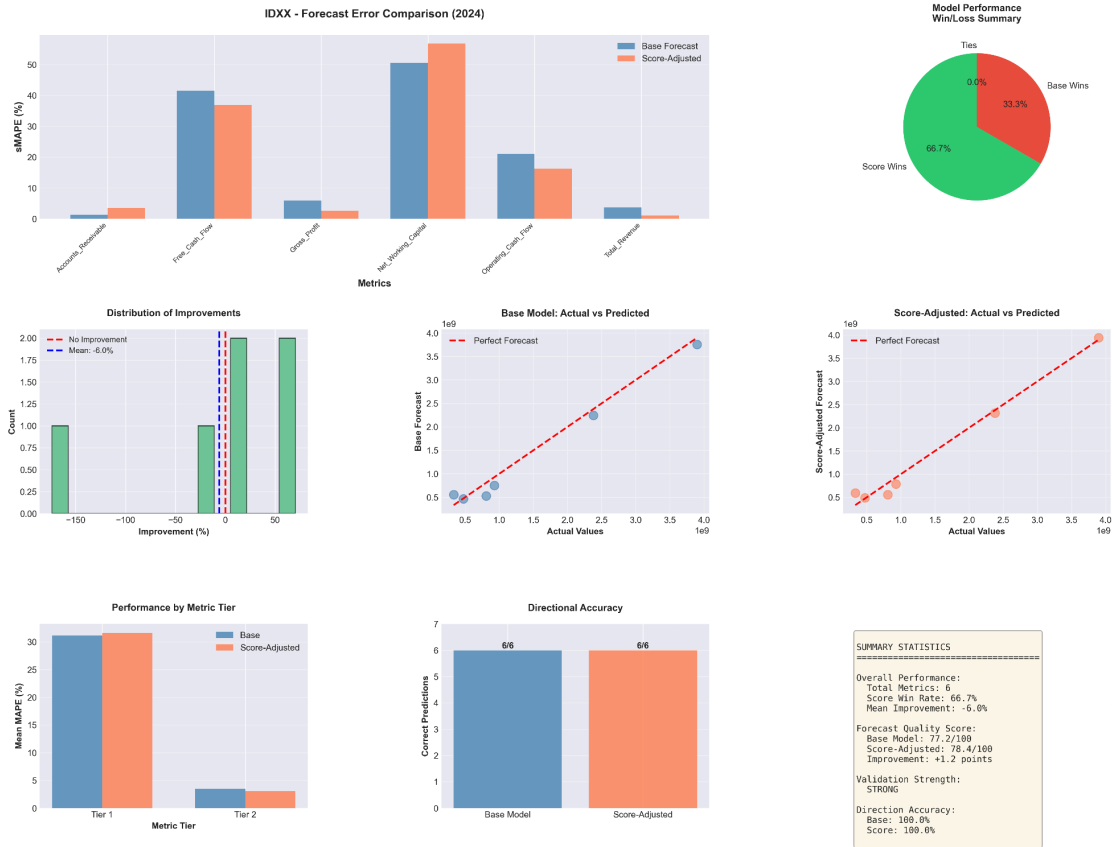
=====

END OF REPORT

=====

IDXX (ARIMA):

Forecast Validation Analysis: IDXX (2024)



FORECAST VALIDATION REPORT

Company: IDXX
Forecast Year: 2024
Analysis Date: 2025-12-12 23:57:14

EXECUTIVE SUMMARY

Validation Strength: STRONG

Score-Adjusted method shows strong validation. Qualitative factors significantly improve forecast accuracy.

Overall Win Rate: 66.7% (Score-Adjusted wins)

Mean Improvement: -6.0%

Forecast Quality Scores:

- Base Model: 77.2/100
- Score-Adjusted Model: 78.4/100
- Improvement: +1.2 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 6

Accuracy Metrics:

Base Model:

- Mean MAPE: 21.96%
- Median MAPE: 12.41%
- Mean sMAPE: 20.69%

Score-Adjusted Model:

- Mean MAPE: 22.14%
- Median MAPE: 9.33%
- Mean sMAPE: 19.55%

Win/Loss Record:

- Score-Adjusted Wins: 4
- Base Model Wins: 2
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 31.19%
- Score Mean MAPE: 31.67%
- Mean Improvement: 20.85%
- Score Win Rate: 75.0%
- Wins: Score=3, Base=1

Tier 2 (Important) Metrics:

- Count: 2
- Base Mean MAPE: 3.52%
- Score Mean MAPE: 3.10%
- Mean Improvement: -59.75%
- Score Win Rate: 50.0%
- Wins: Score=1, Base=1

STATISTICAL SIGNIFICANCE TESTS

Paired T-Test:

- T-Statistic: 0.6300
- P-Value: 0.5564
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 7.0000
- P-Value: 0.5625
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.2572
- Interpretation: Small effect

METRIC-BY-METRIC BREAKDOWN

Accounts_Receivable (Tier 2):

Actual Value: 473,575,000.00

Base Forecast: 467,395,728.00 (Error: -6,179,272.00)

Score Forecast: 490,490,686.00 (Error: +16,915,686.00)

Accuracy:

- Base sMAPE: 1.31%
- Score sMAPE: 3.51%
- Improvement: -173.75%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 808,079,000.00

Base Forecast: 530,110,400.00 (Error: -277,968,600.00)

Score Forecast: 556,304,215.00 (Error: -251,774,785.00)

Accuracy:

- Base sMAPE: 41.54%
- Score sMAPE: 36.91%
- Improvement: +9.42%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 2,378,927,000.00

Base Forecast: 2,242,686,025.00 (Error: -136,240,975.00)

Score Forecast: 2,316,595,986.00 (Error: -62,331,014.00)

Accuracy:

- Base sMAPE: 5.90%
- Score sMAPE: 2.65%
- Improvement: +54.25%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 331,981,000.00

Base Forecast: 556,377,884.00 (Error: +224,396,884.00)

Score Forecast: 595,324,336.00 (Error: +263,343,336.00)

Accuracy:

- Base sMAPE: 50.52%
- Score sMAPE: 56.80%
- Improvement: -17.36%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 929,001,000.00

Base Forecast: 751,636,017.00 (Error: -177,364,983.00)

Score Forecast: 788,775,856.00 (Error: -140,225,144.00)

Accuracy:

- Base sMAPE: 21.11%
- Score sMAPE: 16.33%
- Improvement: +20.94%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 3,897,504,000.00

Base Forecast: 3,754,370,237.00 (Error: -143,133,763.00)

Score Forecast: 3,939,881,180.00 (Error: +42,377,180.00)

Accuracy:

- Base sMAPE: 3.74%
- Score sMAPE: 1.08%
- Improvement: +70.39%

Direction Correct: Base=True, Score=True

Winner: Score

=====

=====

RECOMMENDATIONS

=====

=====

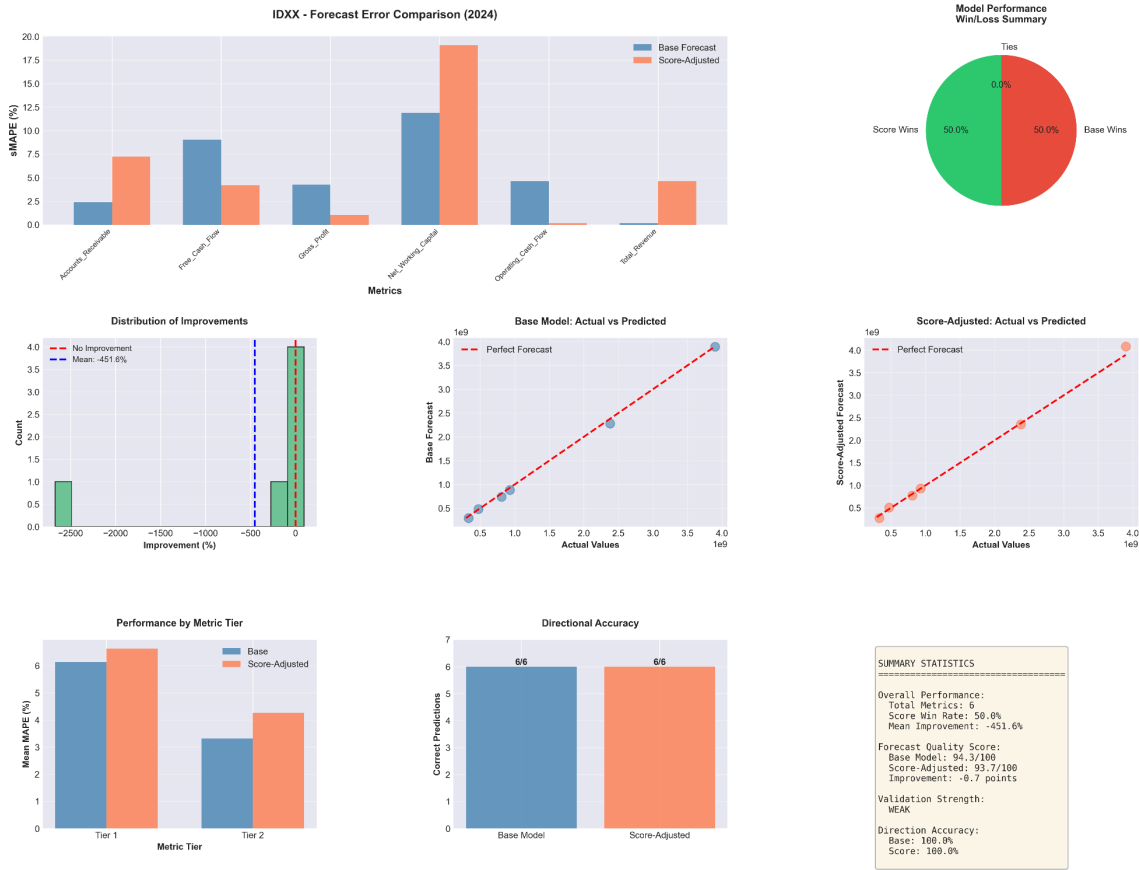
✓ Strong validation achieved. Qualitative score adjustments significantly improve forecast accuracy. Recommend:

1. Continue using score-adjusted methodology for future forecasts
2. Analyze which qualitative factors drove the best improvements
3. Consider increasing weight given to qualitative adjustments
4. Document the successful factors for replication

=====
=====
END OF REPORT
=====
=====

IDXX (WINTERS-HOLT):

Forecast Validation Analysis: IDXX (2024)



FORECAST VALIDATION REPORT

Company: IDXX
Forecast Year: 2024
Analysis Date: 2025-12-12 23:57:33

EXECUTIVE SUMMARY

=====

Validation Strength: WEAK

Score-Adjusted method shows weak validation. Qualitative factors provide marginal improvements.

Overall Win Rate: 50.0% (Score-Adjusted wins)

Mean Improvement: -451.6%

Forecast Quality Scores:

- Base Model: 94.3/100
- Score-Adjusted Model: 93.7/100
- Improvement: -0.7 points

=====

DETAILED METRICS COMPARISON

=====

Total Metrics Analyzed: 6

Accuracy Metrics:

Base Model:

- Mean MAPE: 5.20%
- Median MAPE: 4.35%
- Mean sMAPE: 5.40%

Score-Adjusted Model:

- Mean MAPE: 5.84%
- Median MAPE: 4.44%
- Mean sMAPE: 6.07%

Win/Loss Record:

- Score-Adjusted Wins: 3
- Base Model Wins: 3
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 6.14%
- Score Mean MAPE: 6.62%
- Mean Improvement: -644.72%
- Score Win Rate: 50.0%
- Wins: Score=2, Base=2

Tier 2 (Important) Metrics:

- Count: 2
- Base Mean MAPE: 3.32%
- Score Mean MAPE: 4.27%
- Mean Improvement: -65.34%
- Score Win Rate: 50.0%
- Wins: Score=1, Base=1

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: -0.3022
- P-Value: 0.7746
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 7.0000
- P-Value: 0.5625
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: -0.1234
- Interpretation: Small effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 473,575,000.00

Base Forecast: 485,197,656.00 (Error: +11,622,656.00)

Score Forecast: 509,172,243.00 (Error: +35,597,243.00)

Accuracy:

- Base sMAPE: 2.42%
- Score sMAPE: 7.24%
- Improvement: -206.27%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 808,079,000.00

Base Forecast: 738,343,418.00 (Error: -69,735,582.00)

Score Forecast: 774,826,443.00 (Error: -33,252,557.00)

Accuracy:

- Base sMAPE: 9.02%
- Score sMAPE: 4.20%
- Improvement: +52.32%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 2,378,927,000.00

Base Forecast: 2,279,549,907.00 (Error: -99,377,093.00)

Score Forecast: 2,354,674,754.00 (Error: -24,252,246.00)

Accuracy:

- Base sMAPE: 4.27%
- Score sMAPE: 1.02%
- Improvement: +75.60%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 331,981,000.00

Base Forecast: 294,753,037.00 (Error: -37,227,963.00)

Score Forecast: 274,120,324.00 (Error: -57,860,676.00)

Accuracy:

- Base sMAPE: 11.88%
- Score sMAPE: 19.09%
- Improvement: -55.42%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 929,001,000.00

Base Forecast: 886,911,422.00 (Error: -42,089,578.00)

Score Forecast: 930,735,489.00 (Error: +1,734,489.00)

Accuracy:

- Base sMAPE: 4.64%
- Score sMAPE: 0.19%
- Improvement: +95.88%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 3,897,504,000.00

Base Forecast: 3,890,809,199.00 (Error: -6,694,801.00)

Score Forecast: 4,083,061,863.00 (Error: +185,557,863.00)

Accuracy:

- Base sMAPE: 0.17%
- Score sMAPE: 4.65%
- Improvement: -2671.67%

Direction Correct: Base=True, Score=True

Winner: Base

=====

RECOMMENDATIONS

=====

=====

△ Weak validation. Qualitative adjustments show marginal impact. Recommend:

1. Review and refine the qualitative scoring methodology
2. Investigate if certain metric types benefit more from adjustments
3. Consider whether the base model is already capturing qualitative factors
4. Test alternative adjustment algorithms

=====

=====

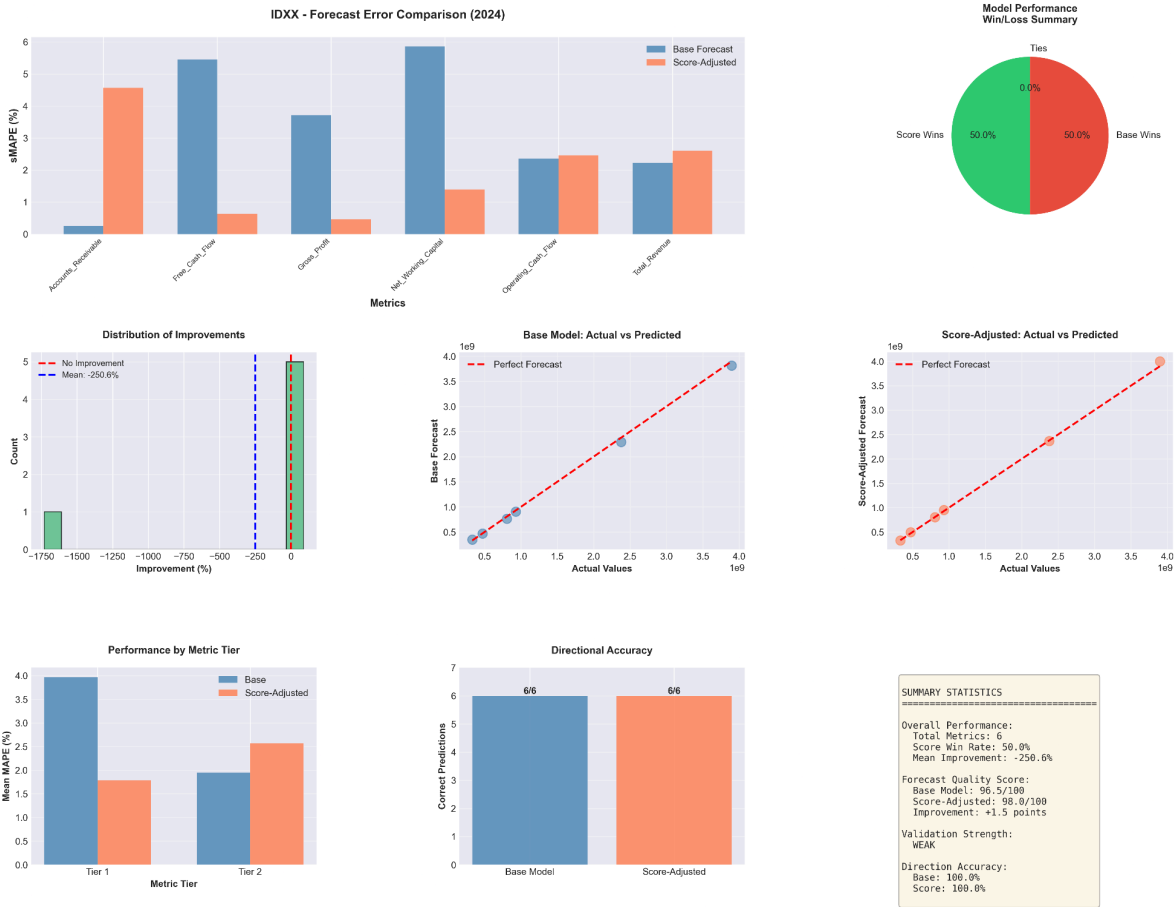
END OF REPORT

=====

=====

IDXX (LINEAR + WINTERS-HOLT):

Forecast Validation Analysis: IDXX (2024)



FORECAST VALIDATION REPORT

Company: IDXX
Forecast Year: 2024
Analysis Date: 2025-12-12 23:57:55

EXECUTIVE SUMMARY

=====

Validation Strength: WEAK

Score-Adjusted method shows weak validation. Qualitative factors provide marginal improvements.

Overall Win Rate: 50.0% (Score-Adjusted wins)

Mean Improvement: -250.6%

Forecast Quality Scores:

- Base Model: 96.5/100
- Score-Adjusted Model: 98.0/100
- Improvement: +1.5 points

=====

DETAILED METRICS COMPARISON

=====

Total Metrics Analyzed: 6

Accuracy Metrics:

Base Model:

- Mean MAPE: 3.30%
- Median MAPE: 2.99%
- Mean sMAPE: 3.31%

Score-Adjusted Model:

- Mean MAPE: 2.05%
- Median MAPE: 1.94%
- Mean sMAPE: 2.02%

Win/Loss Record:

- Score-Adjusted Wins: 3
- Base Model Wins: 3
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 3.97%
- Score Mean MAPE: 1.78%
- Mean Improvement: 34.74%
- Score Win Rate: 50.0%
- Wins: Score=2, Base=2

Tier 2 (Important) Metrics:

- Count: 2
- Base Mean MAPE: 1.95%
- Score Mean MAPE: 2.57%
- Mean Improvement: -821.30%
- Score Win Rate: 50.0%
- Wins: Score=1, Base=1

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 0.8951
- P-Value: 0.4117
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 7.0000
- P-Value: 0.5625
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.3654
- Interpretation: Small effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 473,575,000.00

Base Forecast: 472,365,530.00 (Error: -1,209,470.00)

Score Forecast: 495,706,055.00 (Error: +22,131,055.00)

Accuracy:

- Base sMAPE: 0.26%

- Score sMAPE: 4.57%

- Improvement: -1729.81%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 808,079,000.00

Base Forecast: 765,190,064.00 (Error: -42,888,936.00)

Score Forecast: 802,999,635.00 (Error: -5,079,365.00)

Accuracy:

- Base sMAPE: 5.45%

- Score sMAPE: 0.63%

- Improvement: +88.16%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 2,378,927,000.00

Base Forecast: 2,292,312,774.00 (Error: -86,614,226.00)

Score Forecast: 2,367,858,234.00 (Error: -11,068,766.00)

Accuracy:

- Base sMAPE: 3.71%

- Score sMAPE: 0.47%

- Improvement: +87.22%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 331,981,000.00

Base Forecast: 352,017,221.00 (Error: +20,036,221.00)

Score Forecast: 327,376,015.00 (Error: -4,604,985.00)

Accuracy:

- Base sMAPE: 5.86%
- Score sMAPE: 1.40%
- Improvement: +77.02%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 929,001,000.00

Base Forecast: 907,284,290.00 (Error: -21,716,710.00)

Score Forecast: 952,115,021.00 (Error: +23,114,021.00)

Accuracy:

- Base sMAPE: 2.37%
- Score sMAPE: 2.46%
- Improvement: -6.43%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 3,897,504,000.00

Base Forecast: 3,811,811,696.00 (Error: -85,692,304.00)

Score Forecast: 4,000,160,936.00 (Error: +102,656,936.00)

Accuracy:

- Base sMAPE: 2.22%
- Score sMAPE: 2.60%
- Improvement: -19.80%

Direction Correct: Base=True, Score=True

Winner: Base

=====

RECOMMENDATIONS

=====

△ Weak validation. Qualitative adjustments show marginal impact. Recommend:

1. Review and refine the qualitative scoring methodology
2. Investigate if certain metric types benefit more from adjustments
3. Consider whether the base model is already capturing qualitative factors
4. Test alternative adjustment algorithms

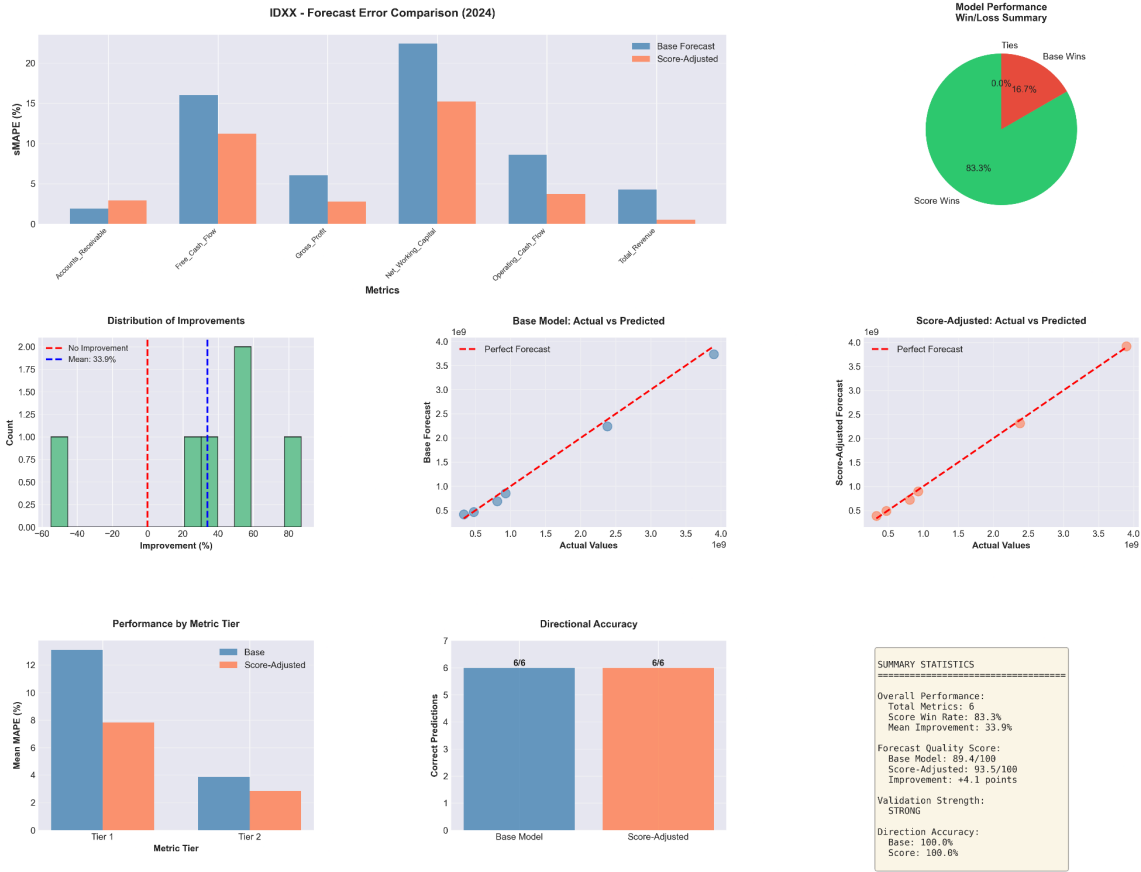
=====

END OF REPORT

=====

IDXX (LINEAR + WINTERS-HOLT + XGBOOST):

Forecast Validation Analysis: IDXX (2024)



FORECAST VALIDATION REPORT

Company: IDXX
Forecast Year: 2024
Analysis Date: 2025-12-13 00:32:26

EXECUTIVE SUMMARY

Validation Strength: STRONG

Score-Adjusted method shows strong validation. Qualitative factors significantly improve forecast accuracy.

Overall Win Rate: 83.3% (Score-Adjusted wins)

Mean Improvement: 33.9%

Forecast Quality Scores:

- Base Model: 89.4/100
- Score-Adjusted Model: 93.5/100
- Improvement: +4.1 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 6

Accuracy Metrics:

Base Model:

- Mean MAPE: 10.04%
- Median MAPE: 7.04%
- Mean sMAPE: 9.87%

Score-Adjusted Model:

- Mean MAPE: 6.17%
- Median MAPE: 3.32%
- Mean sMAPE: 6.07%

Win/Loss Record:

- Score-Adjusted Wins: 5
- Base Model Wins: 1
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 13.12%
- Score Mean MAPE: 7.83%
- Mean Improvement: 51.34%
- Score Win Rate: 100.0%
- Wins: Score=4, Base=0

Tier 2 (Important) Metrics:

- Count: 2
- Base Mean MAPE: 3.88%
- Score Mean MAPE: 2.85%
- Mean Improvement: -0.89%
- Score Win Rate: 50.0%
- Wins: Score=1, Base=1

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 3.4366
- P-Value: 0.0185
- Significant at $\alpha=0.05$: True
- Interpretation: Score-Adjusted significantly different

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 1.0000
- P-Value: 0.0625
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 1.4030
- Interpretation: Large effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 473,575,000.00

Base Forecast: 464,565,037.00 (Error: -9,009,963.00)

Score Forecast: 487,520,124.00 (Error: +13,945,124.00)

Accuracy:

- Base sMAPE: 1.92%
- Score sMAPE: 2.90%
- Improvement: -54.77%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 808,079,000.00

Base Forecast: 688,297,400.00 (Error: -119,781,600.00)

Score Forecast: 722,307,551.00 (Error: -85,771,449.00)

Accuracy:

- Base sMAPE: 16.01%
- Score sMAPE: 11.21%
- Improvement: +28.39%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 2,378,927,000.00

Base Forecast: 2,239,670,371.00 (Error: -139,256,629.00)

Score Forecast: 2,313,480,948.00 (Error: -65,446,052.00)

Accuracy:

- Base sMAPE: 6.03%
- Score sMAPE: 2.79%
- Improvement: +53.00%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 331,981,000.00
Base Forecast: 415,747,026.00 (Error: +83,766,026.00)
Score Forecast: 386,644,734.00 (Error: +54,663,734.00)

Accuracy:

- Base sMAPE: 22.41%
- Score sMAPE: 15.21%
- Improvement: +34.74%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 929,001,000.00
Base Forecast: 852,627,361.00 (Error: -76,373,639.00)
Score Forecast: 894,757,384.00 (Error: -34,243,616.00)

Accuracy:

- Base sMAPE: 8.57%
- Score sMAPE: 3.76%
- Improvement: +55.16%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 3,897,504,000.00
Base Forecast: 3,734,111,672.00 (Error: -163,392,328.00)
Score Forecast: 3,918,621,598.00 (Error: +21,117,598.00)

Accuracy:

- Base sMAPE: 4.28%
- Score sMAPE: 0.54%
- Improvement: +87.08%

Direction Correct: Base=True, Score=True

Winner: Score

=====

=====

RECOMMENDATIONS

=====

✓ Strong validation achieved. Qualitative score adjustments significantly improve forecast accuracy. Recommend:

1. Continue using score-adjusted methodology for future forecasts
2. Analyze which qualitative factors drove the best improvements
3. Consider increasing weight given to qualitative adjustments
4. Document the successful factors for replication

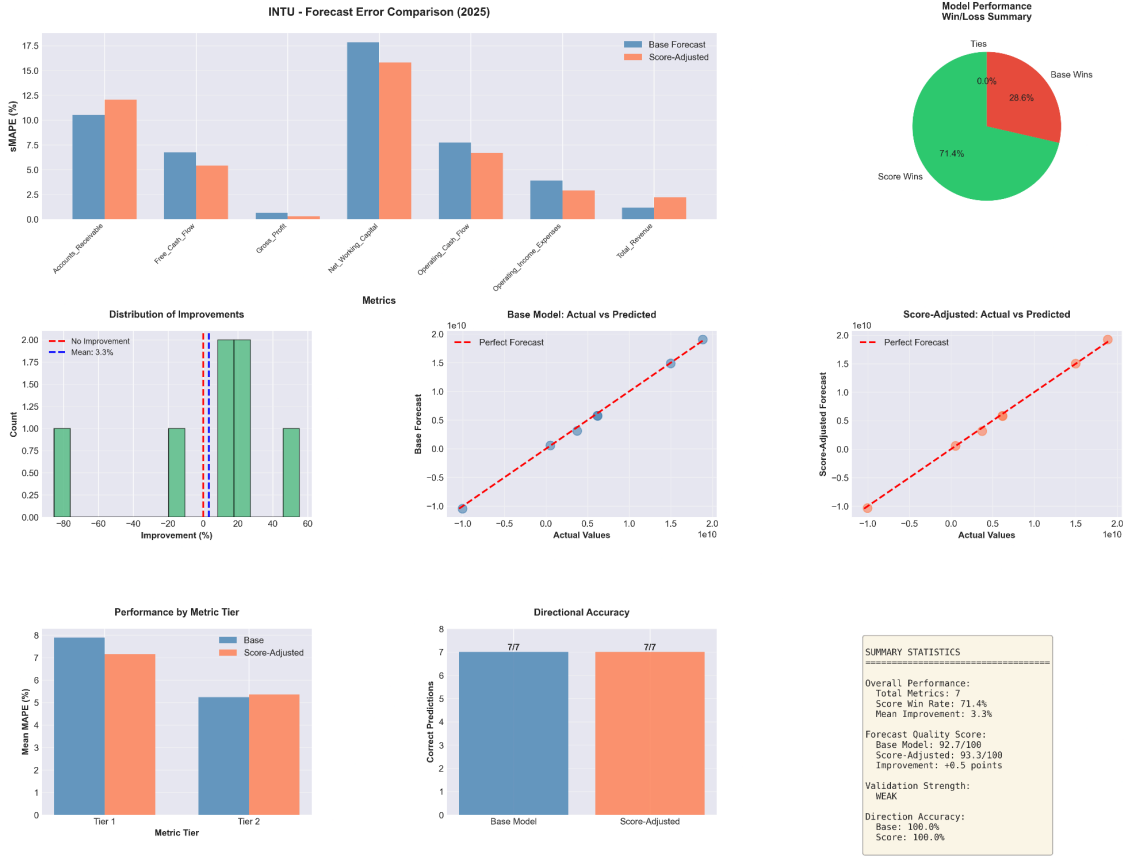
=====

END OF REPORT

=====

INTU (NAIVE):

Forecast Validation Analysis: INTU (2025)



FORECAST VALIDATION REPORT

Company: INTU
Forecast Year: 2025
Analysis Date: 2025-12-11 15:58:00

EXECUTIVE SUMMARY

Validation Strength: WEAK

Score-Adjusted method shows weak validation. Qualitative factors provide marginal improvements.

Overall Win Rate: 71.4% (Score-Adjusted wins)

Mean Improvement: 3.3%

Forecast Quality Scores:

- Base Model: 92.7/100
- Score-Adjusted Model: 93.3/100
- Improvement: +0.5 points

```
=====
=====
DETAILED METRICS COMPARISON
=====
=====
```

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 6.76%
- Median MAPE: 6.51%
- Mean sMAPE: 6.94%

Score-Adjusted Model:

- Mean MAPE: 6.39%
- Median MAPE: 5.30%
- Mean sMAPE: 6.48%

Win/Loss Record:

- Score-Adjusted Wins: 5
- Base Model Wins: 2
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

```
=====
=====
PERFORMANCE BY METRIC TIER
=====
=====
```


Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 7.89%
- Score Mean MAPE: 7.16%
- Mean Improvement: -10.76%
- Score Win Rate: 75.0%
- Wins: Score=3, Base=1

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 5.25%
- Score Mean MAPE: 5.35%
- Mean Improvement: 21.96%
- Score Win Rate: 66.7%
- Wins: Score=2, Base=1

STATISTICAL SIGNIFICANCE TESTS

Paired T-Test:

- T-Statistic: 0.9361
- P-Value: 0.3854
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 9.0000
- P-Value: 0.4688
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.3538
- Interpretation: Small effect

METRIC-BY-METRIC BREAKDOWN

Accounts_Receivable (Tier 2):

Actual Value: 530,000,000.00
Base Forecast: 588,778,899.00 (Error: +58,778,899.00)
Score Forecast: 598,003,422.00 (Error: +68,003,422.00)

Accuracy:

- Base sMAPE: 10.51%
- Score sMAPE: 12.06%
- Improvement: -15.69%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 6,167,000,000.00
Base Forecast: 5,765,268,253.00 (Error: -401,731,747.00)
Score Forecast: 5,840,327,031.00 (Error: -326,672,969.00)

Accuracy:

- Base sMAPE: 6.73%
- Score sMAPE: 5.44%
- Improvement: +18.68%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 14,983,000,000.00
Base Forecast: 14,883,196,485.00 (Error: -99,803,515.00)
Score Forecast: 15,027,620,239.00 (Error: +44,620,239.00)

Accuracy:

- Base sMAPE: 0.67%
- Score sMAPE: 0.30%
- Improvement: +55.29%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 3,737,000,000.00
Base Forecast: 3,124,347,889.00 (Error: -612,652,111.00)
Score Forecast: 3,189,962,242.00 (Error: -547,037,758.00)

Accuracy:

- Base sMAPE: 17.86%
- Score sMAPE: 15.79%
- Improvement: +10.71%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 6,207,000,000.00

Base Forecast: 5,744,553,453.00 (Error: -462,446,547.00)

Score Forecast: 5,804,792,195.00 (Error: -402,207,805.00)

Accuracy:

- Base sMAPE: 7.74%
- Score sMAPE: 6.70%
- Improvement: +13.03%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -10,045,000,000.00

Base Forecast: -10,444,850,802.00 (Error: -399,850,802.00)

Score Forecast: -10,339,721,246.00 (Error: -294,721,246.00)

Accuracy:

- Base sMAPE: 3.90%
- Score sMAPE: 2.89%
- Improvement: +26.29%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 18,831,000,000.00

Base Forecast: 19,058,189,368.00 (Error: +227,189,368.00)

Score Forecast: 19,252,312,624.00 (Error: +421,312,624.00)

Accuracy:

- Base sMAPE: 1.20%

- Score sMAPE: 2.21%
- Improvement: -85.45%

Direction Correct: Base=True, Score=True

Winner: Base

=====

=====

RECOMMENDATIONS

=====

=====

△ Weak validation. Qualitative adjustments show marginal impact. Recommend:

1. Review and refine the qualitative scoring methodology
2. Investigate if certain metric types benefit more from adjustments
3. Consider whether the base model is already capturing qualitative factors
4. Test alternative adjustment algorithms

=====

=====

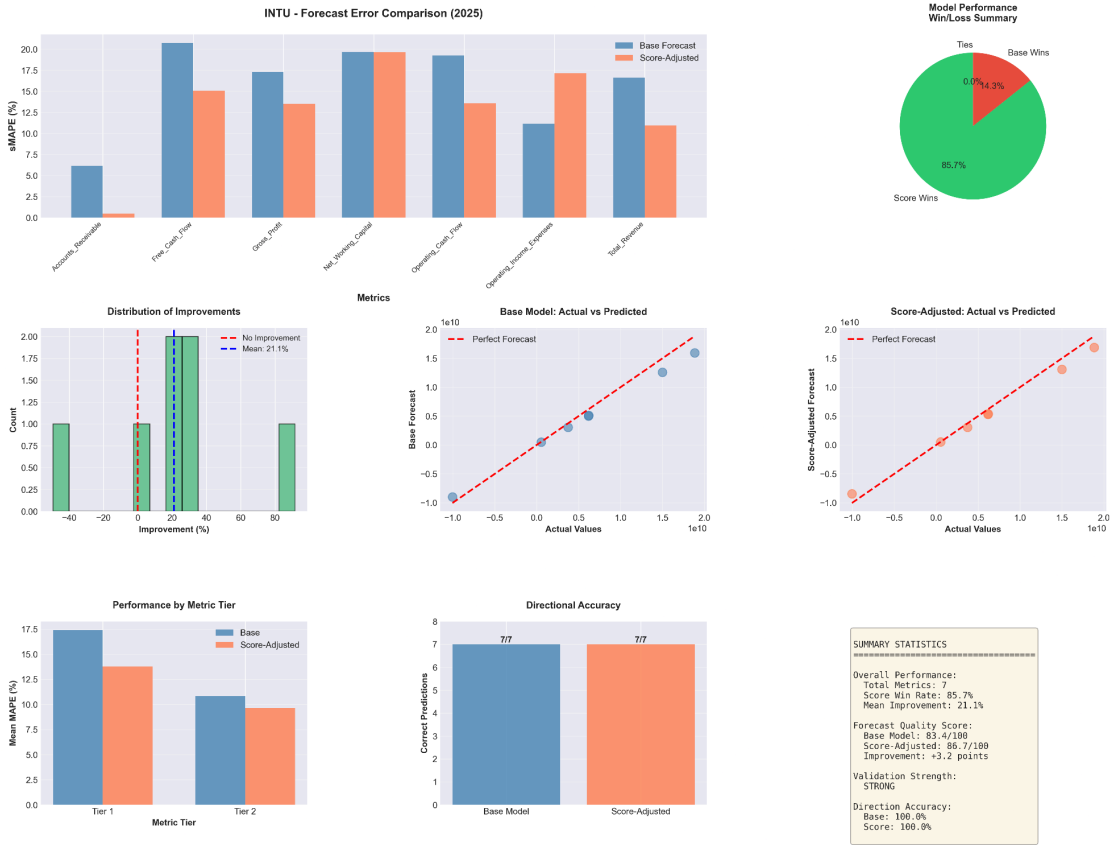
END OF REPORT

=====

=====

INTU (LINEAR):

Forecast Validation Analysis: INTU (2025)



FORECAST VALIDATION REPORT

Company: INTU
Forecast Year: 2025
Analysis Date: 2025-12-11 16:26:44

EXECUTIVE SUMMARY

Validation Strength: **STRONG**

Score-Adjusted method shows strong validation. Qualitative factors significantly improve forecast accuracy.

Overall Win Rate: 85.7% (Score-Adjusted wins)

Mean Improvement: 21.1%

Forecast Quality Scores:

- Base Model: 83.4/100
- Score-Adjusted Model: 86.7/100
- Improvement: +3.2 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 14.58%
- Median MAPE: 15.92%
- Mean sMAPE: 15.84%

Score-Adjusted Model:

- Mean MAPE: 12.00%
- Median MAPE: 12.73%
- Mean sMAPE: 12.92%

Win/Loss Record:

- Score-Adjusted Wins: 6
- Base Model Wins: 1
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 17.39%
- Score Mean MAPE: 13.76%
- Mean Improvement: 21.27%
- Score Win Rate: 100.0%
- Wins: Score=4, Base=0

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 10.83%
- Score Mean MAPE: 9.64%
- Mean Improvement: 20.95%
- Score Win Rate: 66.7%
- Wins: Score=2, Base=1

STATISTICAL SIGNIFICANCE TESTS

Paired T-Test:

- T-Statistic: 1.7400
- P-Value: 0.1325
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 7.0000
- P-Value: 0.2969
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.6576
- Interpretation: Medium effect

METRIC-BY-METRIC BREAKDOWN

Accounts_Receivable (Tier 2):

Actual Value: 530,000,000.00

Base Forecast: 498,275,000.00 (Error: -31,725,000.00)

Score Forecast: 527,372,266.00 (Error: -2,627,734.00)

Accuracy:

- Base sMAPE: 6.17%
- Score sMAPE: 0.50%
- Improvement: +91.72%

Direction Correct: Base=True, Score=True

Winner: Score

Free_Cash_Flow (Tier 1):

Actual Value: 6,167,000,000.00

Base Forecast: 5,008,875,000.00 (Error: -1,158,125,000.00)

Score Forecast: 5,301,373,264.00 (Error: -865,626,736.00)

Accuracy:

- Base sMAPE: 20.73%
- Score sMAPE: 15.10%
- Improvement: +25.26%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 14,983,000,000.00

Base Forecast: 12,597,158,333.00 (Error: -2,385,841,667.00)

Score Forecast: 13,087,792,456.00 (Error: -1,895,207,544.00)

Accuracy:

- Base sMAPE: 17.30%
- Score sMAPE: 13.50%
- Improvement: +20.56%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 3,737,000,000.00

Base Forecast: 3,067,550,000.00 (Error: -669,450,000.00)

Score Forecast: 3,068,533,750.00 (Error: -668,466,250.00)

Accuracy:

- Base sMAPE: 19.68%
- Score sMAPE: 19.64%
- Improvement: +0.15%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 6,207,000,000.00

Base Forecast: 5,117,729,166.00 (Error: -1,089,270,834.00)

Score Forecast: 5,416,584,079.00 (Error: -790,415,921.00)

Accuracy:

- Base sMAPE: 19.24%
- Score sMAPE: 13.60%
- Improvement: +27.44%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -10,045,000,000.00

Base Forecast: -8,983,622,222.00 (Error: +1,061,377,778.00)

Score Forecast: -8,459,014,618.00 (Error: +1,585,985,382.00)

Accuracy:

- Base sMAPE: 11.16%
- Score sMAPE: 17.14%
- Improvement: -49.43%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 18,831,000,000.00

Base Forecast: 15,943,375,000.00 (Error: -2,887,625,000.00)

Score Forecast: 16,874,404,326.00 (Error: -1,956,595,674.00)

Accuracy:

- Base sMAPE: 16.61%
- Score sMAPE: 10.96%
- Improvement: +32.24%

Direction Correct: Base=True, Score=True
Winner: Score

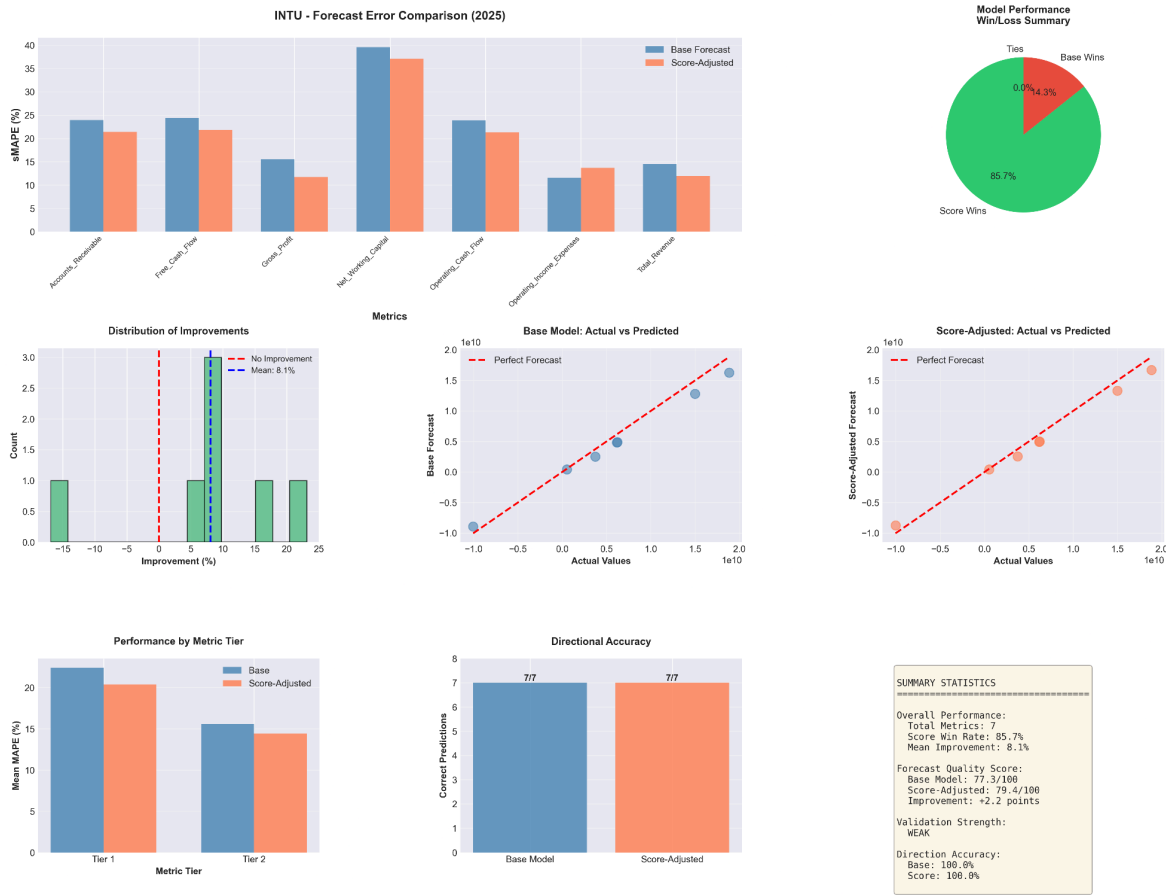
```
=====
=====
RECOMMENDATIONS
=====
=====
```

- ✓ Strong validation achieved. Qualitative score adjustments significantly improve forecast accuracy. Recommend:
1. Continue using score-adjusted methodology for future forecasts
 2. Analyze which qualitative factors drove the best improvements
 3. Consider increasing weight given to qualitative adjustments
 4. Document the successful factors for replication

```
=====
=====
END OF REPORT
=====
=====
```


INTU (XGBOOST):

Forecast Validation Analysis: INTU (2025)



FORECAST VALIDATION REPORT

Company: INTU
Forecast Year: 2025
Analysis Date: 2025-12-11 15:20:32

EXECUTIVE SUMMARY

=====

Validation Strength: WEAK

Score-Adjusted method shows weak validation. Qualitative factors provide marginal improvements.

Overall Win Rate: 85.7% (Score-Adjusted wins)

Mean Improvement: 8.1%

Forecast Quality Scores:

- Base Model: 77.3/100
- Score-Adjusted Model: 79.4/100
- Improvement: +2.2 points

=====

DETAILED METRICS COMPARISON

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 19.49%
- Median MAPE: 21.31%
- Mean sMAPE: 21.92%

Score-Adjusted Model:

- Mean MAPE: 17.83%
- Median MAPE: 19.27%
- Mean sMAPE: 19.88%

Win/Loss Record:

- Score-Adjusted Wins: 6
- Base Model Wins: 1
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 22.41%
- Score Mean MAPE: 20.39%
- Mean Improvement: 10.21%
- Score Win Rate: 100.0%
- Wins: Score=4, Base=0

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 15.60%
- Score Mean MAPE: 14.42%
- Mean Improvement: 5.27%
- Score Win Rate: 66.7%
- Wins: Score=2, Base=1

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 2.8708
- P-Value: 0.0284
- Significant at $\alpha=0.05$: True
- Interpretation: Score-Adjusted significantly different

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 1.0000
- P-Value: 0.0312
- Significant at $\alpha=0.05$: True

Effect Size:

- Cohen's d: 1.0851
- Interpretation: Large effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 530,000,000.00

Base Forecast: 416,683,936.00 (Error: -113,316,064.00)

Score Forecast: 427,517,718.00 (Error: -102,482,282.00)

Accuracy:

- Base sMAPE: 23.94%
- Score sMAPE: 21.41%
- Improvement: +9.56%

Direction Correct: Base=True, Score=True

Winner: Score

Free_Cash_Flow (Tier 1):

Actual Value: 6,167,000,000.00

Base Forecast: 4,824,999,424.00 (Error: -1,342,000,576.00)

Score Forecast: 4,950,449,409.00 (Error: -1,216,550,591.00)

Accuracy:

- Base sMAPE: 24.42%
- Score sMAPE: 21.89%
- Improvement: +9.35%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 14,983,000,000.00

Base Forecast: 12,819,996,672.00 (Error: -2,163,003,328.00)

Score Forecast: 13,319,976,542.00 (Error: -1,663,023,458.00)

Accuracy:

- Base sMAPE: 15.56%
- Score sMAPE: 11.75%
- Improvement: +23.12%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 3,737,000,000.00

Base Forecast: 2,501,999,872.00 (Error: -1,235,000,128.00)

Score Forecast: 2,567,051,868.00 (Error: -1,169,948,132.00)

Accuracy:

- Base sMAPE: 39.59%
- Score sMAPE: 37.12%
- Improvement: +5.27%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 6,207,000,000.00

Base Forecast: 4,883,999,232.00 (Error: -1,323,000,768.00)

Score Forecast: 5,010,983,212.00 (Error: -1,196,016,788.00)

Accuracy:

- Base sMAPE: 23.86%
- Score sMAPE: 21.32%
- Improvement: +9.60%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -10,045,000,000.00

Base Forecast: -8,942,996,480.00 (Error: +1,102,003,520.00)

Score Forecast: -8,756,982,153.00 (Error: +1,288,017,847.00)

Accuracy:

- Base sMAPE: 11.61%
- Score sMAPE: 13.70%
- Improvement: -16.88%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 18,831,000,000.00

Base Forecast: 16,284,996,608.00 (Error: -2,546,003,392.00)
Score Forecast: 16,708,406,519.00 (Error: -2,122,593,481.00)

Accuracy:

- Base sMAPE: 14.50%
- Score sMAPE: 11.95%
- Improvement: +16.63%

Direction Correct: Base=True, Score=True

Winner: Score

```
=====
=====
RECOMMENDATIONS
=====
=====
```

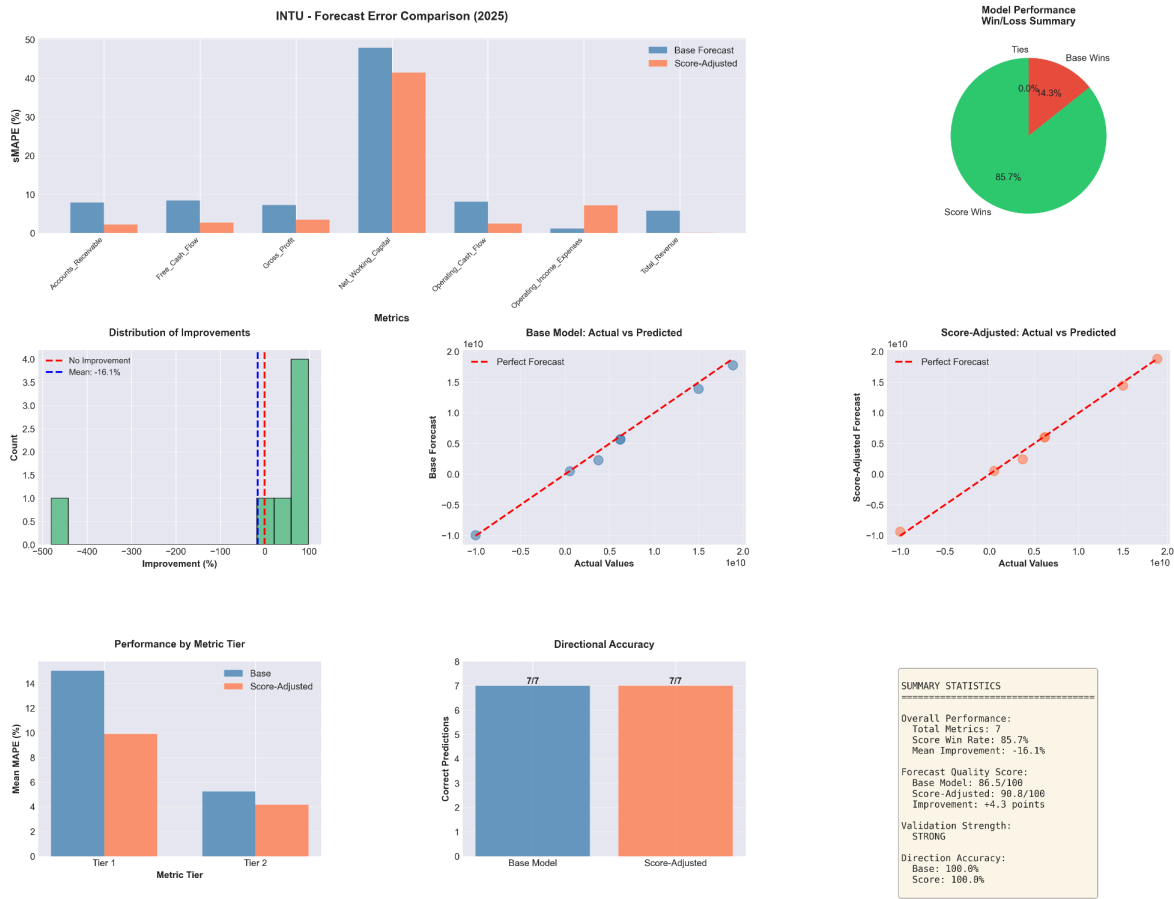
△ Weak validation. Qualitative adjustments show marginal impact. Recommend:

1. Review and refine the qualitative scoring methodology
2. Investigate if certain metric types benefit more from adjustments
3. Consider whether the base model is already capturing qualitative factors
4. Test alternative adjustment algorithms

```
=====
=====
END OF REPORT
=====
=====
```


INTU (ARIMA):

Forecast Validation Analysis: INTU (2025)



FORECAST VALIDATION REPORT

Company: INTU
Forecast Year: 2025
Analysis Date: 2025-12-11 16:48:19

EXECUTIVE SUMMARY

=====

Validation Strength: **STRONG**

Score-Adjusted method shows strong validation. Qualitative factors significantly improve forecast accuracy.

Overall Win Rate: 85.7% (Score-Adjusted wins)

Mean Improvement: -16.1%

Forecast Quality Scores:

- Base Model: 86.5/100
- Score-Adjusted Model: 90.8/100
- Improvement: +4.3 points

=====

DETAILED METRICS COMPARISON

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 10.84%
- Median MAPE: 7.58%
- Mean sMAPE: 12.36%

Score-Adjusted Model:

- Mean MAPE: 7.44%
- Median MAPE: 2.71%
- Mean sMAPE: 8.52%

Win/Loss Record:

- Score-Adjusted Wins: 6
- Base Model Wins: 1
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 15.03%
- Score Mean MAPE: 9.89%
- Mean Improvement: 61.24%
- Score Win Rate: 100.0%
- Wins: Score=4, Base=0

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 5.26%
- Score Mean MAPE: 4.17%
- Mean Improvement: -119.26%
- Score Win Rate: 66.7%
- Wins: Score=2, Base=1

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 2.3019
- P-Value: 0.0609
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 6.0000
- P-Value: 0.2188
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.8700
- Interpretation: Large effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 530,000,000.00

Base Forecast: 489,846,070.00 (Error: -40,153,930.00)

Score Forecast: 518,451,121.00 (Error: -11,548,879.00)

Accuracy:

- Base sMAPE: 7.87%
- Score sMAPE: 2.20%
- Improvement: +71.24%

Direction Correct: Base=True, Score=True

Winner: Score

Free_Cash_Flow (Tier 1):

Actual Value: 6,167,000,000.00

Base Forecast: 5,668,675,342.00 (Error: -498,324,658.00)

Score Forecast: 5,999,703,307.00 (Error: -167,296,693.00)

Accuracy:

- Base sMAPE: 8.42%
- Score sMAPE: 2.75%
- Improvement: +66.43%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 14,983,000,000.00

Base Forecast: 13,935,667,679.00 (Error: -1,047,332,321.00)

Score Forecast: 14,478,434,064.00 (Error: -504,565,936.00)

Accuracy:

- Base sMAPE: 7.24%
- Score sMAPE: 3.43%
- Improvement: +51.82%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 3,737,000,000.00

Base Forecast: 2,292,629,512.00 (Error: -1,444,370,488.00)

Score Forecast: 2,453,113,578.00 (Error: -1,283,886,422.00)

Accuracy:

- Base sMAPE: 47.91%
- Score sMAPE: 41.48%
- Improvement: +11.11%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 6,207,000,000.00

Base Forecast: 5,722,981,653.00 (Error: -484,018,347.00)

Score Forecast: 6,057,180,890.00 (Error: -149,819,110.00)

Accuracy:

- Base sMAPE: 8.11%
- Score sMAPE: 2.44%
- Improvement: +69.05%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -10,045,000,000.00

Base Forecast: -9,924,468,903.00 (Error: +120,531,097.00)

Score Forecast: -9,344,919,617.00 (Error: +700,080,383.00)

Accuracy:

- Base sMAPE: 1.21%
- Score sMAPE: 7.22%
- Improvement: -480.83%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 18,831,000,000.00

Base Forecast: 17,775,698,573.00 (Error: -1,055,301,427.00)

Score Forecast: 18,813,728,267.00 (Error: -17,271,733.00)

Accuracy:

- Base sMAPE: 5.77%
- Score sMAPE: 0.09%
- Improvement: +98.36%

Direction Correct: Base=True, Score=True

Winner: Score

```
=====
=====
RECOMMENDATIONS
=====
=====
```

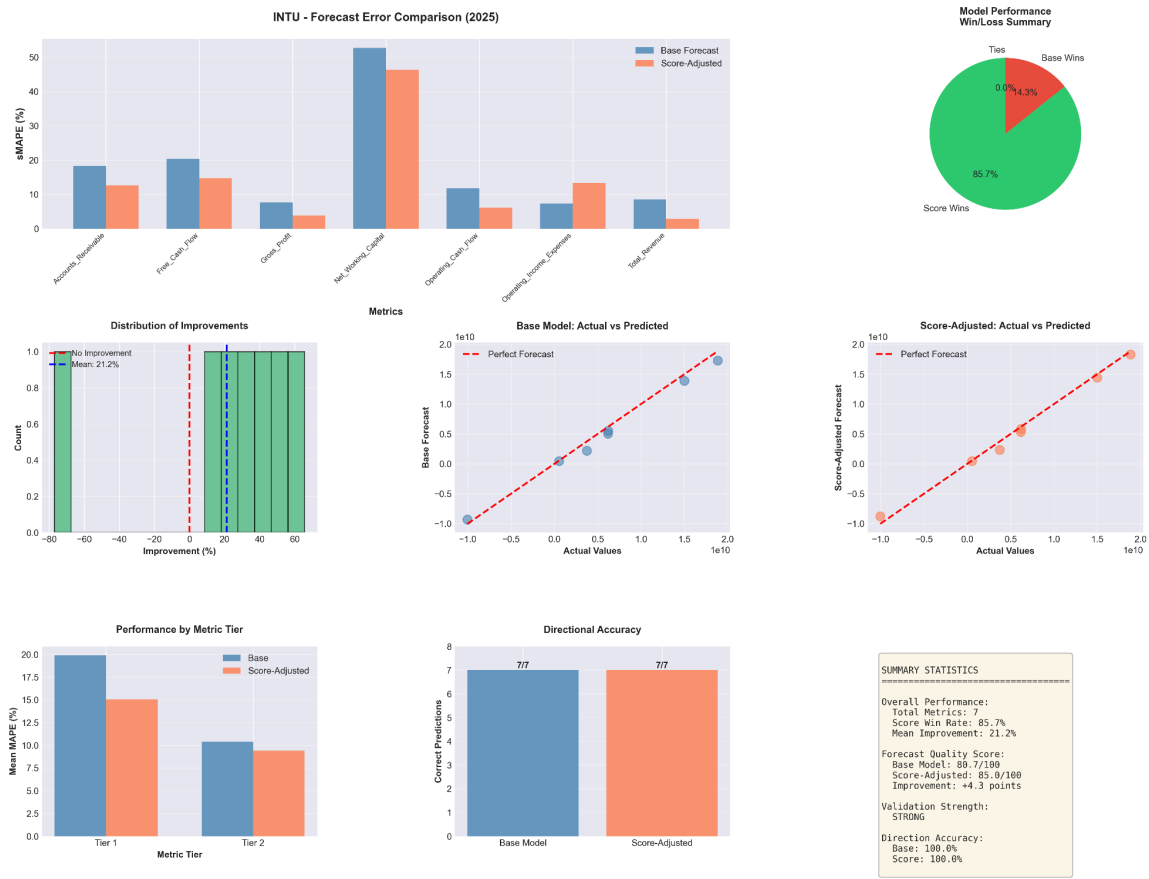
✓ Strong validation achieved. Qualitative score adjustments significantly improve forecast accuracy. Recommend:

1. Continue using score-adjusted methodology for future forecasts
2. Analyze which qualitative factors drove the best improvements
3. Consider increasing weight given to qualitative adjustments
4. Document the successful factors for replication

```
=====
=====
END OF REPORT
=====
=====
```


INTU (WINTERS-HOLT):

Forecast Validation Analysis: INTU (2025)



FORECAST VALIDATION REPORT

Company: INTU
Forecast Year: 2025
Analysis Date: 2025-12-11 16:42:41

EXECUTIVE SUMMARY

Validation Strength: **STRONG**

Score-Adjusted method shows strong validation. Qualitative factors significantly improve forecast accuracy.

Overall Win Rate: 85.7% (Score-Adjusted wins)

Mean Improvement: 21.2%

Forecast Quality Scores:

- Base Model: 80.7/100
- Score-Adjusted Model: 85.0/100
- Improvement: +4.3 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 15.82%
- Median MAPE: 11.18%
- Mean sMAPE: 18.10%

Score-Adjusted Model:

- Mean MAPE: 12.62%
- Median MAPE: 11.89%
- Mean sMAPE: 14.27%

Win/Loss Record:

- Score-Adjusted Wins: 6
- Base Model Wins: 1
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 19.89%
- Score Mean MAPE: 15.04%
- Mean Improvement: 36.89%
- Score Win Rate: 100.0%
- Wins: Score=4, Base=0

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 10.40%
- Score Mean MAPE: 9.39%
- Mean Improvement: 0.26%
- Score Win Rate: 66.7%
- Wins: Score=2, Base=1

STATISTICAL SIGNIFICANCE TESTS

Paired T-Test:

- T-Statistic: 2.2988
- P-Value: 0.0612
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 6.0000
- P-Value: 0.2188
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.8689
- Interpretation: Large effect

METRIC-BY-METRIC BREAKDOWN

Accounts_Receivable (Tier 2):

Actual Value: 530,000,000.00

Base Forecast: 441,200,422.00 (Error: -88,799,578.00)

Score Forecast: 466,964,762.00 (Error: -63,035,238.00)

Accuracy:

- Base sMAPE: 18.29%

- Score sMAPE: 12.65%

- Improvement: +29.01%

Direction Correct: Base=True, Score=True

Winner: Score

Free_Cash_Flow (Tier 1):

Actual Value: 6,167,000,000.00

Base Forecast: 5,025,779,358.00 (Error: -1,141,220,642.00)

Score Forecast: 5,319,264,770.00 (Error: -847,735,230.00)

Accuracy:

- Base sMAPE: 20.39%

- Score sMAPE: 14.76%

- Improvement: +25.72%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 14,983,000,000.00

Base Forecast: 13,874,915,596.00 (Error: -1,108,084,404.00)

Score Forecast: 14,415,315,809.00 (Error: -567,684,191.00)

Accuracy:

- Base sMAPE: 7.68%

- Score sMAPE: 3.86%

- Improvement: +48.77%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 3,737,000,000.00

Base Forecast: 2,178,938,805.00 (Error: -1,558,061,195.00)

Score Forecast: 2,331,464,521.00 (Error: -1,405,535,479.00)

Accuracy:

- Base sMAPE: 52.67%
- Score sMAPE: 46.32%
- Improvement: +9.79%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 6,207,000,000.00

Base Forecast: 5,512,976,434.00 (Error: -694,023,566.00)

Score Forecast: 5,834,912,205.00 (Error: -372,087,795.00)

Accuracy:

- Base sMAPE: 11.84%
- Score sMAPE: 6.18%
- Improvement: +46.39%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -10,045,000,000.00

Base Forecast: -9,336,937,997.00 (Error: +708,062,003.00)

Score Forecast: -8,791,698,165.00 (Error: +1,253,301,835.00)

Accuracy:

- Base sMAPE: 7.31%
- Score sMAPE: 13.31%
- Improvement: -77.00%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 18,831,000,000.00

Base Forecast: 17,292,986,960.00 (Error: -1,538,013,040.00)

Score Forecast: 18,302,828,226.00 (Error: -528,171,774.00)

Accuracy:

- Base sMAPE: 8.52%
- Score sMAPE: 2.84%
- Improvement: +65.66%

Direction Correct: Base=True, Score=True
Winner: Score

=====

=====

RECOMMENDATIONS

=====

=====

- ✓ Strong validation achieved. Qualitative score adjustments significantly improve forecast accuracy. Recommend:
1. Continue using score-adjusted methodology for future forecasts
 2. Analyze which qualitative factors drove the best improvements
 3. Consider increasing weight given to qualitative adjustments
 4. Document the successful factors for replication

=====

=====

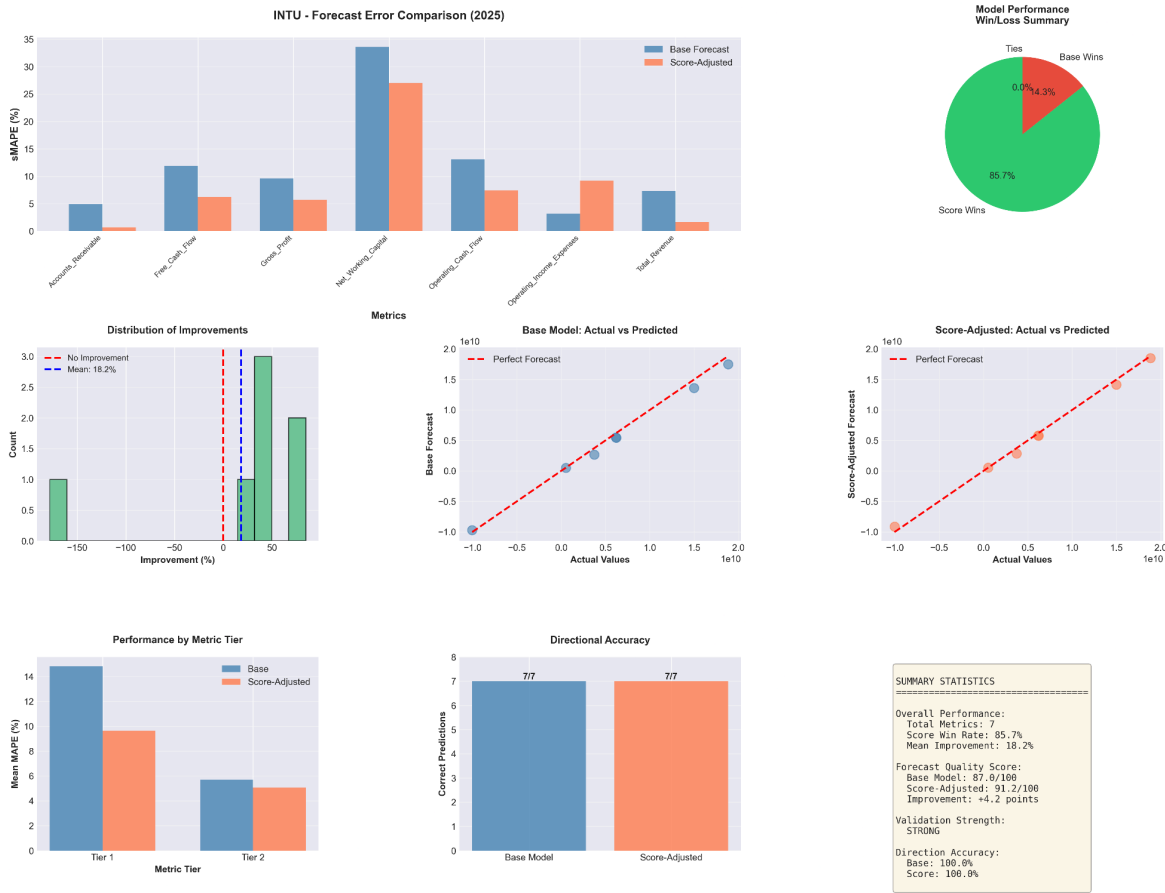
END OF REPORT

=====

=====

INTU (LINEAR + WINTERS-HOLT):

Forecast Validation Analysis: INTU (2025)



FORECAST VALIDATION REPORT

Company: INTU
Forecast Year: 2025
Analysis Date: 2025-12-11 16:44:29

EXECUTIVE SUMMARY

=====

Validation Strength: **STRONG**

Score-Adjusted method shows strong validation. Qualitative factors significantly improve forecast accuracy.

Overall Win Rate: 85.7% (Score-Adjusted wins)

Mean Improvement: 18.2%

Forecast Quality Scores:

- Base Model: 87.0/100
- Score-Adjusted Model: 91.2/100
- Improvement: +4.2 points

=====

DETAILED METRICS COMPARISON

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 10.92%
- Median MAPE: 9.13%
- Mean sMAPE: 11.95%

Score-Adjusted Model:

- Mean MAPE: 7.68%
- Median MAPE: 6.04%
- Mean sMAPE: 8.29%

Win/Loss Record:

- Score-Adjusted Wins: 6
- Base Model Wins: 1
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 14.83%
- Score Mean MAPE: 9.65%
- Mean Improvement: 45.52%
- Score Win Rate: 100.0%
- Wins: Score=4, Base=0

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 5.71%
- Score Mean MAPE: 5.05%
- Mean Improvement: -18.20%
- Score Win Rate: 66.7%
- Wins: Score=2, Base=1

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 2.2151
- P-Value: 0.0687
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 6.0000
- P-Value: 0.2188
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.8372
- Interpretation: Large effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 530,000,000.00

Base Forecast: 504,455,179.00 (Error: -25,544,821.00)

Score Forecast: 533,913,344.00 (Error: +3,913,344.00)

Accuracy:

- Base sMAPE: 4.94%
- Score sMAPE: 0.74%
- Improvement: +84.68%

Direction Correct: Base=True, Score=True

Winner: Score

Free_Cash_Flow (Tier 1):

Actual Value: 6,167,000,000.00

Base Forecast: 5,474,922,224.00 (Error: -692,077,776.00)

Score Forecast: 5,794,635,782.00 (Error: -372,364,218.00)

Accuracy:

- Base sMAPE: 11.89%
- Score sMAPE: 6.23%
- Improvement: +46.20%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 14,983,000,000.00

Base Forecast: 13,615,055,417.00 (Error: -1,367,944,583.00)

Score Forecast: 14,145,334,595.00 (Error: -837,665,405.00)

Accuracy:

- Base sMAPE: 9.57%
- Score sMAPE: 5.75%
- Improvement: +38.76%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 3,737,000,000.00

Base Forecast: 2,661,542,282.00 (Error: -1,075,457,718.00)

Score Forecast: 2,847,850,242.00 (Error: -889,149,758.00)

Accuracy:

- Base sMAPE: 33.62%
- Score sMAPE: 27.01%
- Improvement: +17.32%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 6,207,000,000.00

Base Forecast: 5,445,859,777.00 (Error: -761,140,223.00)

Score Forecast: 5,763,876,204.00 (Error: -443,123,796.00)

Accuracy:

- Base sMAPE: 13.06%
- Score sMAPE: 7.40%
- Improvement: +41.78%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -10,045,000,000.00

Base Forecast: -9,725,990,068.00 (Error: +319,009,932.00)

Score Forecast: -9,158,031,152.00 (Error: +886,968,848.00)

Accuracy:

- Base sMAPE: 3.23%
- Score sMAPE: 9.24%
- Improvement: -178.04%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 18,831,000,000.00

Base Forecast: 17,499,984,363.00 (Error: -1,331,015,637.00)

Score Forecast: 18,521,913,450.00 (Error: -309,086,550.00)

Accuracy:

- Base sMAPE: 7.33%
- Score sMAPE: 1.65%
- Improvement: +76.78%

Direction Correct: Base=True, Score=True

Winner: Score

=====

=====

RECOMMENDATIONS

=====

=====

✓ Strong validation achieved. Qualitative score adjustments significantly improve forecast accuracy. Recommend:

1. Continue using score-adjusted methodology for future forecasts
2. Analyze which qualitative factors drove the best improvements
3. Consider increasing weight given to qualitative adjustments
4. Document the successful factors for replication

=====

=====

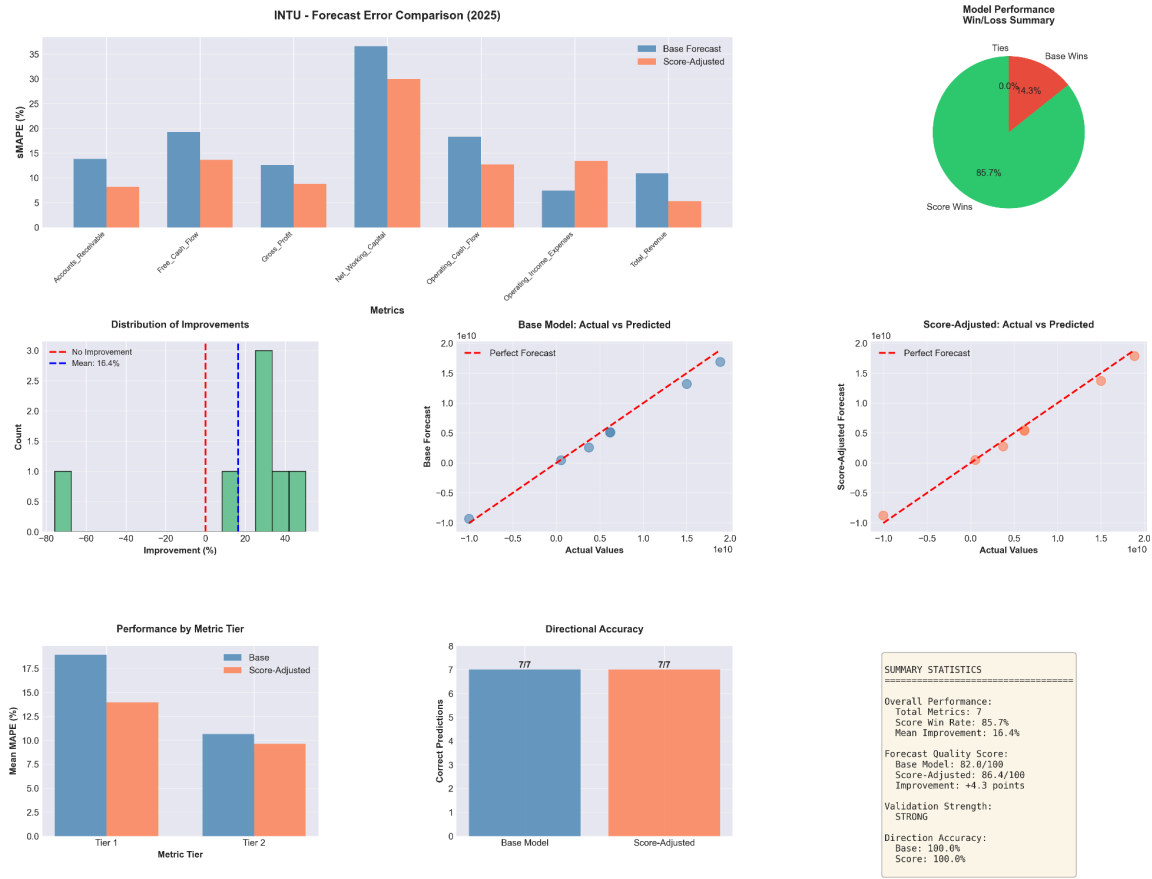
END OF REPORT

=====

=====

INTU (LINEAR + WINTERS-HOLT + XGBOOST):

Forecast Validation Analysis: INTU (2025)



FORECAST VALIDATION REPORT

Company: INTU
Forecast Year: 2025
Analysis Date: 2025-12-11 16:46:14

EXECUTIVE SUMMARY

Validation Strength: STRONG

Score-Adjusted method shows strong validation. Qualitative factors significantly improve forecast accuracy.

Overall Win Rate: 85.7% (Score-Adjusted wins)

Mean Improvement: 16.4%

Forecast Quality Scores:

- Base Model: 82.0/100
- Score-Adjusted Model: 86.4/100
- Improvement: +4.3 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 15.38%
- Median MAPE: 12.94%
- Mean sMAPE: 17.00%

Score-Adjusted Model:

- Mean MAPE: 12.11%
- Median MAPE: 11.94%
- Mean sMAPE: 13.14%

Win/Loss Record:

- Score-Adjusted Wins: 6
- Base Model Wins: 1
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 18.92%
- Score Mean MAPE: 13.99%
- Mean Improvement: 30.59%
- Score Win Rate: 100.0%
- Wins: Score=4, Base=0

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 10.65%
- Score Mean MAPE: 9.62%
- Mean Improvement: -2.53%
- Score Win Rate: 66.7%
- Wins: Score=2, Base=1

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 2.3057
- P-Value: 0.0606
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 6.0000
- P-Value: 0.2188
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.8715
- Interpretation: Large effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 530,000,000.00

Base Forecast: 461,411,077.00 (Error: -68,588,923.00)

Score Forecast: 488,355,639.00 (Error: -41,644,361.00)

Accuracy:

- Base sMAPE: 13.84%
- Score sMAPE: 8.18%
- Improvement: +39.28%

Direction Correct: Base=True, Score=True

Winner: Score

Free_Cash_Flow (Tier 1):

Actual Value: 6,167,000,000.00

Base Forecast: 5,083,137,656.00 (Error: -1,083,862,344.00)

Score Forecast: 5,379,972,562.00 (Error: -787,027,438.00)

Accuracy:

- Base sMAPE: 19.27%
- Score sMAPE: 13.63%
- Improvement: +27.39%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 14,983,000,000.00

Base Forecast: 13,205,949,212.00 (Error: -1,777,050,788.00)

Score Forecast: 13,720,294,522.00 (Error: -1,262,705,478.00)

Accuracy:

- Base sMAPE: 12.61%
- Score sMAPE: 8.80%
- Improvement: +28.94%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 3,737,000,000.00

Base Forecast: 2,581,248,197.00 (Error: -1,155,751,803.00)

Score Forecast: 2,761,935,571.00 (Error: -975,064,429.00)

Accuracy:

- Base sMAPE: 36.58%
- Score sMAPE: 30.01%
- Improvement: +15.63%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 6,207,000,000.00

Base Forecast: 5,164,036,576.00 (Error: -1,042,963,424.00)

Score Forecast: 5,465,595,656.00 (Error: -741,404,344.00)

Accuracy:

- Base sMAPE: 18.34%
- Score sMAPE: 12.70%
- Improvement: +28.91%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -10,045,000,000.00

Base Forecast: -9,326,553,690.00 (Error: +718,446,310.00)

Score Forecast: -8,781,920,260.00 (Error: +1,263,079,740.00)

Accuracy:

- Base sMAPE: 7.42%
- Score sMAPE: 13.42%
- Improvement: -75.81%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 18,831,000,000.00

Base Forecast: 16,877,215,477.00 (Error: -1,953,784,523.00)

Score Forecast: 17,862,777,352.00 (Error: -968,222,648.00)

Accuracy:

- Base sMAPE: 10.94%
- Score sMAPE: 5.28%
- Improvement: +50.44%

Direction Correct: Base=True, Score=True

Winner: Score

=====

=====

RECOMMENDATIONS

=====

=====

✓ Strong validation achieved. Qualitative score adjustments significantly improve forecast accuracy. Recommend:

1. Continue using score-adjusted methodology for future forecasts
2. Analyze which qualitative factors drove the best improvements
3. Consider increasing weight given to qualitative adjustments
4. Document the successful factors for replication

=====

=====

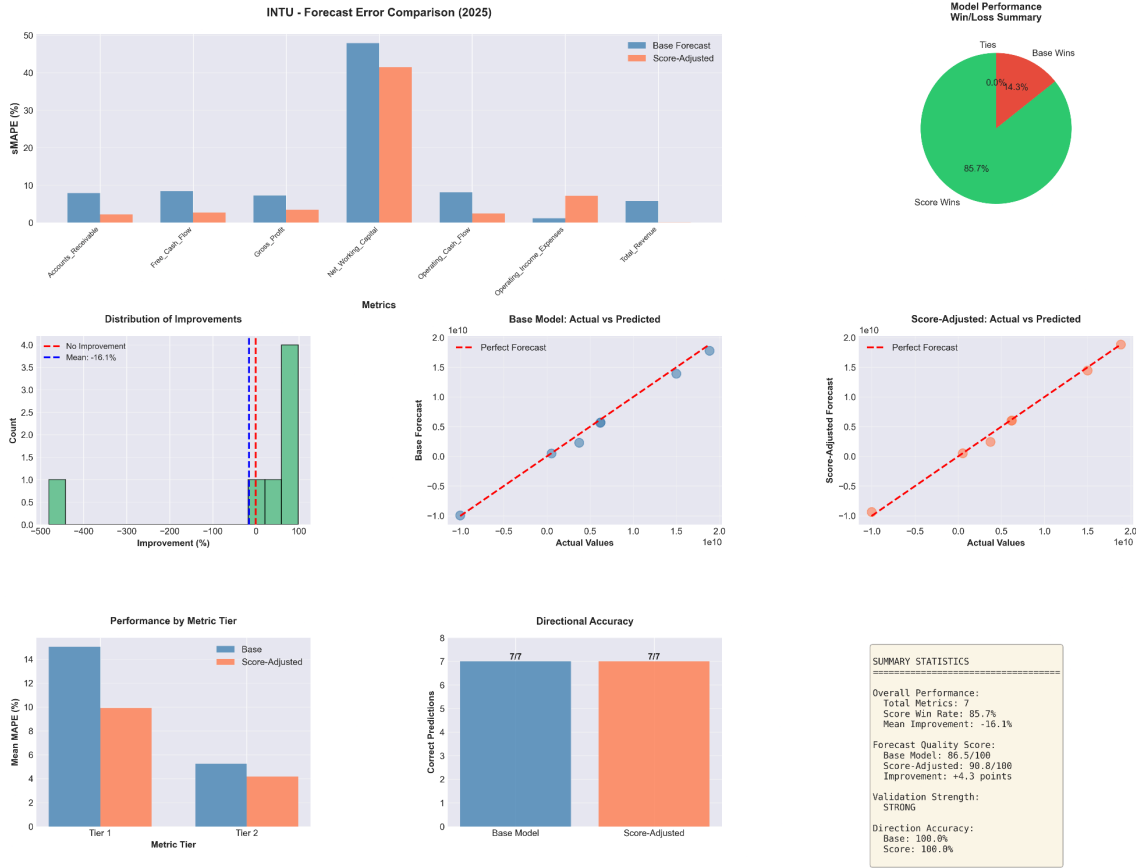
END OF REPORT

=====

=====

INTU (LINEAR + WINTERS-HOLT + ARIMA):

Forecast Validation Analysis: INTU (2025)



FORECAST VALIDATION REPORT

Company: INTU
Forecast Year: 2025
Analysis Date: 2025-12-11 16:48:19

EXECUTIVE SUMMARY

Validation Strength: STRONG

Score-Adjusted method shows strong validation. Qualitative factors significantly improve forecast accuracy.

Overall Win Rate: 85.7% (Score-Adjusted wins)

Mean Improvement: -16.1%

Forecast Quality Scores:

- Base Model: 86.5/100
- Score-Adjusted Model: 90.8/100
- Improvement: +4.3 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 10.84%
- Median MAPE: 7.58%
- Mean sMAPE: 12.36%

Score-Adjusted Model:

- Mean MAPE: 7.44%
- Median MAPE: 2.71%
- Mean sMAPE: 8.52%

Win/Loss Record:

- Score-Adjusted Wins: 6
- Base Model Wins: 1
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 15.03%
- Score Mean MAPE: 9.89%
- Mean Improvement: 61.24%
- Score Win Rate: 100.0%
- Wins: Score=4, Base=0

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 5.26%
- Score Mean MAPE: 4.17%
- Mean Improvement: -119.26%
- Score Win Rate: 66.7%
- Wins: Score=2, Base=1

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 2.3019
- P-Value: 0.0609
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 6.0000
- P-Value: 0.2188
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.8700
- Interpretation: Large effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 530,000,000.00

Base Forecast: 489,846,070.00 (Error: -40,153,930.00)

Score Forecast: 518,451,121.00 (Error: -11,548,879.00)

Accuracy:

- Base sMAPE: 7.87%
- Score sMAPE: 2.20%
- Improvement: +71.24%

Direction Correct: Base=True, Score=True

Winner: Score

Free_Cash_Flow (Tier 1):

Actual Value: 6,167,000,000.00

Base Forecast: 5,668,675,342.00 (Error: -498,324,658.00)

Score Forecast: 5,999,703,307.00 (Error: -167,296,693.00)

Accuracy:

- Base sMAPE: 8.42%
- Score sMAPE: 2.75%
- Improvement: +66.43%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 14,983,000,000.00

Base Forecast: 13,935,667,679.00 (Error: -1,047,332,321.00)

Score Forecast: 14,478,434,064.00 (Error: -504,565,936.00)

Accuracy:

- Base sMAPE: 7.24%
- Score sMAPE: 3.43%
- Improvement: +51.82%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 3,737,000,000.00

Base Forecast: 2,292,629,512.00 (Error: -1,444,370,488.00)

Score Forecast: 2,453,113,578.00 (Error: -1,283,886,422.00)

Accuracy:

- Base sMAPE: 47.91%
- Score sMAPE: 41.48%
- Improvement: +11.11%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 6,207,000,000.00

Base Forecast: 5,722,981,653.00 (Error: -484,018,347.00)

Score Forecast: 6,057,180,890.00 (Error: -149,819,110.00)

Accuracy:

- Base sMAPE: 8.11%
- Score sMAPE: 2.44%
- Improvement: +69.05%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -10,045,000,000.00

Base Forecast: -9,924,468,903.00 (Error: +120,531,097.00)

Score Forecast: -9,344,919,617.00 (Error: +700,080,383.00)

Accuracy:

- Base sMAPE: 1.21%
- Score sMAPE: 7.22%
- Improvement: -480.83%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 18,831,000,000.00

Base Forecast: 17,775,698,573.00 (Error: -1,055,301,427.00)

Score Forecast: 18,813,728,267.00 (Error: -17,271,733.00)

Accuracy:

- Base sMAPE: 5.77%
- Score sMAPE: 0.09%
- Improvement: +98.36%

Direction Correct: Base=True, Score=True

Winner: Score

```
=====
=====
RECOMMENDATIONS
=====
=====
```

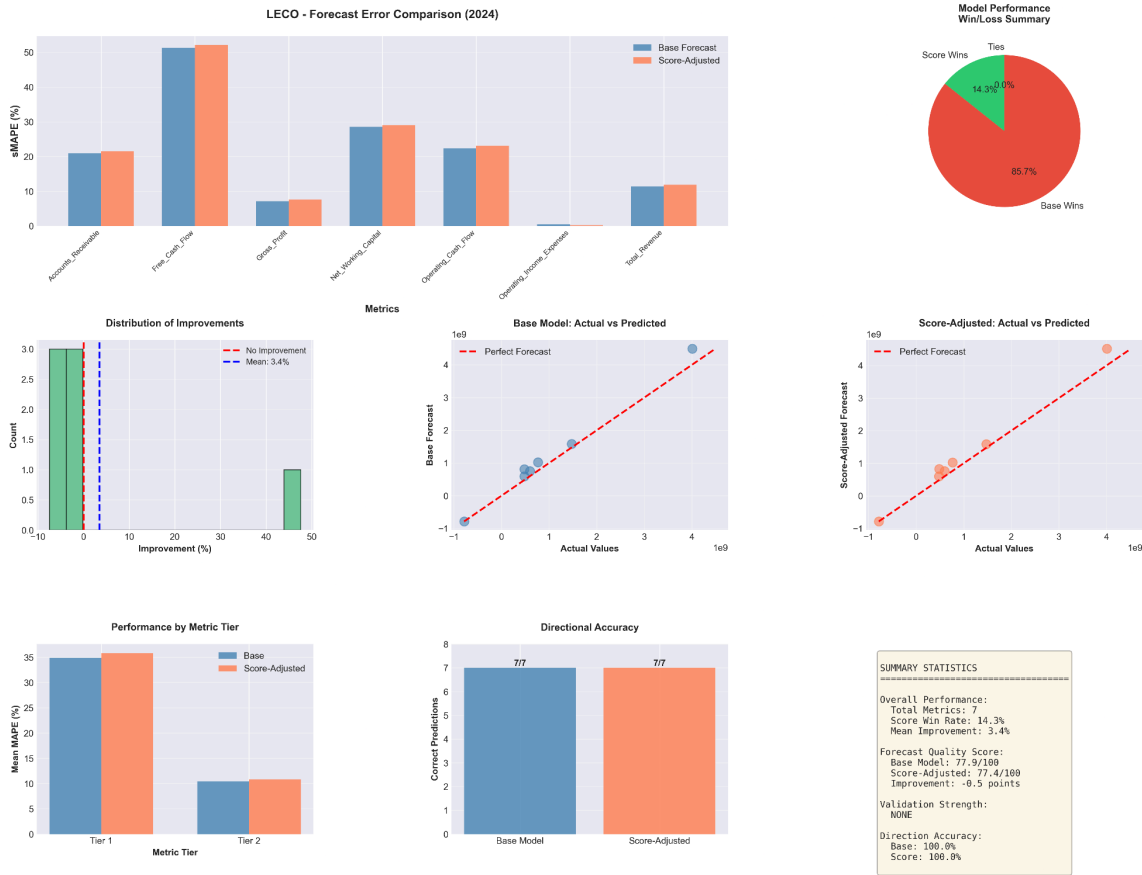
✓ Strong validation achieved. Qualitative score adjustments significantly improve forecast accuracy. Recommend:

1. Continue using score-adjusted methodology for future forecasts
2. Analyze which qualitative factors drove the best improvements
3. Consider increasing weight given to qualitative adjustments
4. Document the successful factors for replication

```
=====
=====
END OF REPORT
=====
=====
```


LECO (NAIVE):

Forecast Validation Analysis: LECO (2024)



FORECAST VALIDATION REPORT

Company: LECO
Forecast Year: 2024
Analysis Date: 2025-12-11 18:29:32

EXECUTIVE SUMMARY

=====

Validation Strength: NONE

Score-Adjusted method does not validate. Base model performs as well or better.

Overall Win Rate: 14.3% (Score-Adjusted wins)

Mean Improvement: 3.4%

Forecast Quality Scores:

- Base Model: 77.9/100
- Score-Adjusted Model: 77.4/100
- Improvement: -0.5 points

=====

DETAILED METRICS COMPARISON

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 24.41%
- Median MAPE: 23.37%
- Mean sMAPE: 20.31%

Score-Adjusted Model:

- Mean MAPE: 25.10%
- Median MAPE: 24.18%
- Mean sMAPE: 20.81%

Win/Loss Record:

- Score-Adjusted Wins: 1
- Base Model Wins: 6
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 34.91%
- Score Mean MAPE: 35.83%
- Mean Improvement: -3.13%
- Score Win Rate: 0.0%
- Wins: Score=0, Base=4

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 10.42%
- Score Mean MAPE: 10.80%
- Mean Improvement: 12.22%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: -3.8648
- P-Value: 0.0083
- Significant at $\alpha=0.05$: True
- Interpretation: Score-Adjusted significantly different

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 1.0000
- P-Value: 0.0312
- Significant at $\alpha=0.05$: True

Effect Size:

- Cohen's d: -1.4608
- Interpretation: Large effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 481,979,000.00

Base Forecast: 594,638,047.00 (Error: +112,659,047.00)

Score Forecast: 598,544,610.00 (Error: +116,565,610.00)

Accuracy:

- Base sMAPE: 20.93%
- Score sMAPE: 21.58%
- Improvement: -3.47%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 482,374,000.00

Base Forecast: 815,485,178.00 (Error: +333,111,178.00)

Score Forecast: 822,376,621.00 (Error: +340,002,621.00)

Accuracy:

- Base sMAPE: 51.33%
- Score sMAPE: 52.12%
- Improvement: -2.07%

Direction Correct: Base=True, Score=True

Winner: Base

Gross_Profit (Tier 2):

Actual Value: 1,472,912,000.00

Base Forecast: 1,581,871,514.00 (Error: +108,959,514.00)

Score Forecast: 1,590,021,370.00 (Error: +117,109,370.00)

Accuracy:

- Base sMAPE: 7.13%
- Score sMAPE: 7.65%
- Improvement: -7.48%

Direction Correct: Base=True, Score=True

Winner: Base

Net_Working_Capital (Tier 1):

Actual Value: 766,479,000.00
Base Forecast: 1,021,390,124.00 (Error: +254,911,124.00)
Score Forecast: 1,027,192,363.00 (Error: +260,713,363.00)

Accuracy:

- Base sMAPE: 28.52%
- Score sMAPE: 29.07%
- Improvement: -2.28%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 598,977,000.00
Base Forecast: 750,099,721.00 (Error: +151,122,721.00)
Score Forecast: 755,878,762.00 (Error: +156,901,762.00)

Accuracy:

- Base sMAPE: 22.40%
- Score sMAPE: 23.16%
- Improvement: -3.82%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Income_Expenses (Tier 2):

Actual Value: -780,590,000.00
Base Forecast: -784,327,973.00 (Error: -3,737,973.00)
Score Forecast: -782,548,715.00 (Error: -1,958,715.00)

Accuracy:

- Base sMAPE: 0.48%
- Score sMAPE: 0.25%
- Improvement: +47.60%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 4,008,670,000.00
Base Forecast: 4,493,415,964.00 (Error: +484,745,964.00)
Score Forecast: 4,514,540,562.00 (Error: +505,870,562.00)

Accuracy:

- Base sMAPE: 11.40%
- Score sMAPE: 11.87%
- Improvement: -4.36%

Direction Correct: Base=True, Score=True

Winner: Base

=====

=====

RECOMMENDATIONS

=====

=====

X No validation. Base model performs as well or better. Recommend:

1. Critically review the qualitative scoring framework
2. Check if qualitative factors are already reflected in historical data
3. Consider that market efficiency may limit qualitative alpha
4. Explore if timing or sector-specific factors affected results

=====

=====

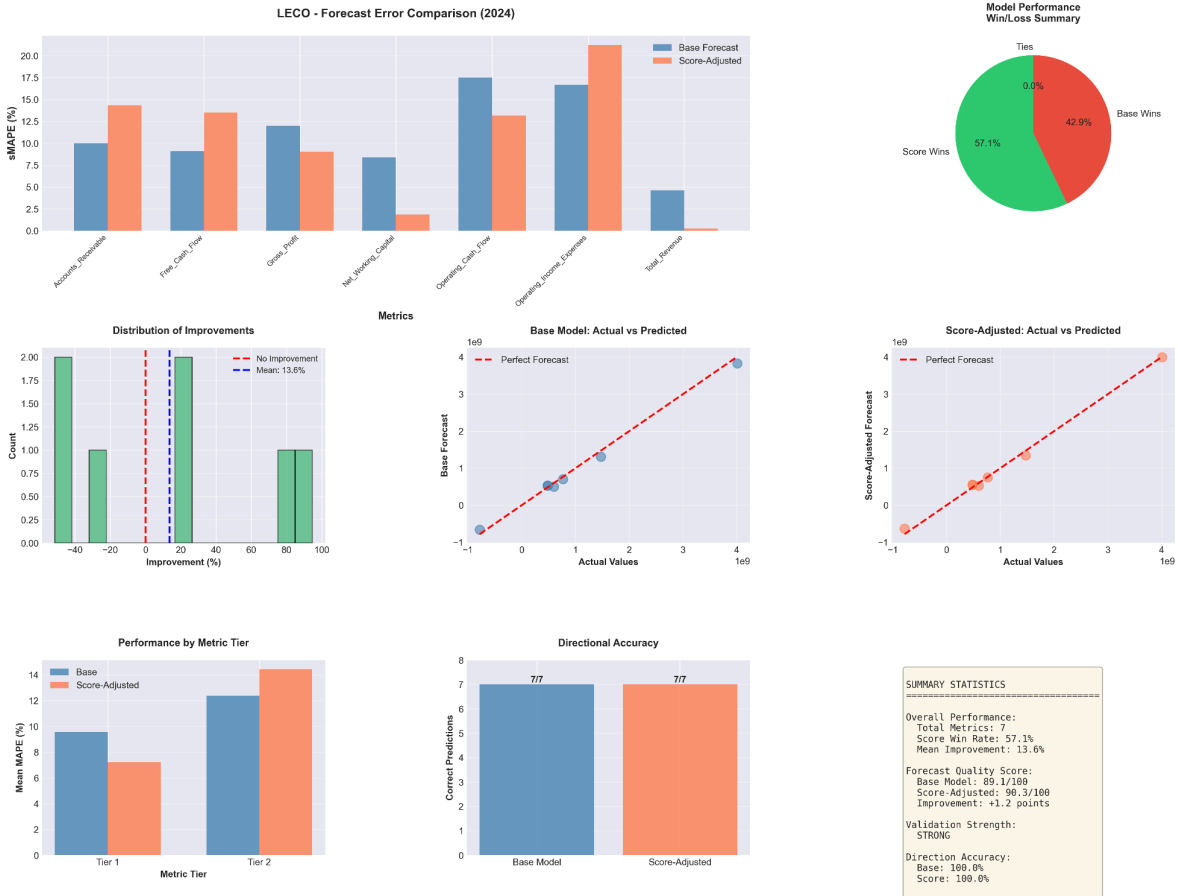
END OF REPORT

=====

=====

LECO (LINEAR):

Forecast Validation Analysis: LECO (2024)



FORECAST VALIDATION REPORT

Company: LECO
Forecast Year: 2024
Analysis Date: 2025-12-11 18:31:34

EXECUTIVE SUMMARY

=====

Validation Strength: **STRONG**

Score-Adjusted method shows strong validation. Qualitative factors significantly improve forecast accuracy.

Overall Win Rate: 57.1% (Score-Adjusted wins)

Mean Improvement: 13.6%

Forecast Quality Scores:

- Base Model: 89.1/100
- Score-Adjusted Model: 90.3/100
- Improvement: +1.2 points

=====

DETAILED METRICS COMPARISON

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 10.77%
- Median MAPE: 10.51%
- Mean sMAPE: 11.18%

Score-Adjusted Model:

- Mean MAPE: 10.31%
- Median MAPE: 12.33%
- Mean sMAPE: 10.48%

Win/Loss Record:

- Score-Adjusted Wins: 4
- Base Model Wins: 3
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 9.56%
- Score Mean MAPE: 7.22%
- Mean Improvement: 35.92%
- Score Win Rate: 75.0%
- Wins: Score=3, Base=1

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 12.40%
- Score Mean MAPE: 14.44%
- Mean Improvement: -16.15%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 0.3773
- P-Value: 0.7189
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 13.0000
- P-Value: 0.9375
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.1426
- Interpretation: Small effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 481,979,000.00

Base Forecast: 532,621,300.00 (Error: +50,642,300.00)

Score Forecast: 556,546,648.00 (Error: +74,567,648.00)

Accuracy:

- Base sMAPE: 9.98%
- Score sMAPE: 14.36%
- Improvement: -47.24%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 482,374,000.00

Base Forecast: 528,462,226.00 (Error: +46,088,226.00)

Score Forecast: 552,200,749.00 (Error: +69,826,749.00)

Accuracy:

- Base sMAPE: 9.12%
- Score sMAPE: 13.50%
- Improvement: -51.51%

Direction Correct: Base=True, Score=True

Winner: Base

Gross_Profit (Tier 2):

Actual Value: 1,472,912,000.00

Base Forecast: 1,306,382,261.00 (Error: -166,529,739.00)

Score Forecast: 1,345,521,473.00 (Error: -127,390,527.00)

Accuracy:

- Base sMAPE: 11.98%
- Score sMAPE: 9.04%
- Improvement: +23.50%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 766,479,000.00

Base Forecast: 704,884,641.00 (Error: -61,594,359.00)

Score Forecast: 752,393,866.00 (Error: -14,085,134.00)

Accuracy:

- Base sMAPE: 8.37%
- Score sMAPE: 1.85%
- Improvement: +77.13%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 598,977,000.00

Base Forecast: 502,528,281.00 (Error: -96,448,719.00)

Score Forecast: 525,101,852.00 (Error: -73,875,148.00)

Accuracy:

- Base sMAPE: 17.51%
- Score sMAPE: 13.14%
- Improvement: +23.40%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -780,590,000.00

Base Forecast: -660,502,462.00 (Error: +120,087,538.00)

Score Forecast: -630,832,691.00 (Error: +149,757,309.00)

Accuracy:

- Base sMAPE: 16.67%
- Score sMAPE: 21.22%
- Improvement: -24.71%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 4,008,670,000.00

Base Forecast: 3,827,049,165.00 (Error: -181,620,835.00)

Score Forecast: 3,998,960,213.00 (Error: -9,709,787.00)

Accuracy:

- Base sMAPE: 4.64%
- Score sMAPE: 0.24%
- Improvement: +94.65%

Direction Correct: Base=True, Score=True

Winner: Score

=====

=====

RECOMMENDATIONS

=====

=====

✓ Strong validation achieved. Qualitative score adjustments significantly improve forecast accuracy. Recommend:

1. Continue using score-adjusted methodology for future forecasts
2. Analyze which qualitative factors drove the best improvements
3. Consider increasing weight given to qualitative adjustments
4. Document the successful factors for replication

=====

=====

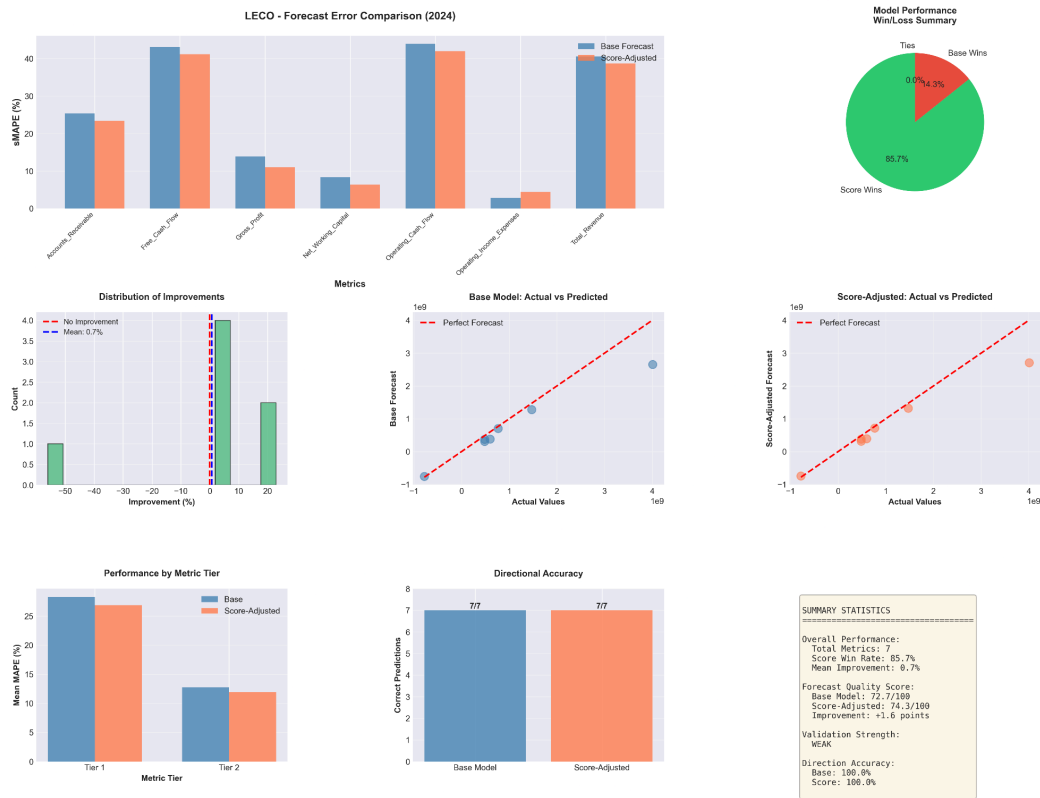
END OF REPORT

=====

=====

LECO (XGBOOST):

Forecast Validation Analysis: LECO (2024)



FORECAST VALIDATION REPORT

Company: LECO
Forecast Year: 2024
Analysis Date: 2025-12-11 18:27:34

EXECUTIVE SUMMARY

Validation Strength: WEAK

Score-Adjusted method shows weak validation. Qualitative factors provide marginal improvements.

Overall Win Rate: 85.7% (Score-Adjusted wins)

Mean Improvement: 0.7%

Forecast Quality Scores:

- Base Model: 72.7/100
- Score-Adjusted Model: 74.3/100
- Improvement: +1.6 points

```
=====
=====
DETAILED METRICS COMPARISON
=====
=====
```

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 21.65%
- Median MAPE: 22.51%
- Mean sMAPE: 25.44%

Score-Adjusted Model:

- Mean MAPE: 20.46%
- Median MAPE: 20.96%
- Mean sMAPE: 23.87%

Win/Loss Record:

- Score-Adjusted Wins: 6
- Base Model Wins: 1
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

```
=====
=====
PERFORMANCE BY METRIC TIER
=====
=====
```


Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 28.30%
- Score Mean MAPE: 26.87%
- Mean Improvement: 8.50%
- Score Win Rate: 100.0%
- Wins: Score=4, Base=0

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 12.78%
- Score Mean MAPE: 11.91%
- Mean Improvement: -9.71%
- Score Win Rate: 66.7%
- Wins: Score=2, Base=1

STATISTICAL SIGNIFICANCE TESTS

Paired T-Test:

- T-Statistic: 2.8495
- P-Value: 0.0292
- Significant at $\alpha=0.05$: True
- Interpretation: Score-Adjusted significantly different

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 1.0000
- P-Value: 0.0312
- Significant at $\alpha=0.05$: True

Effect Size:

- Cohen's d: 1.0770
- Interpretation: Large effect

METRIC-BY-METRIC BREAKDOWN

Accounts_Receivable (Tier 2):

Actual Value: 481,979,000.00
Base Forecast: 373,487,104.00 (Error: -108,491,896.00)
Score Forecast: 380,956,846.00 (Error: -101,022,154.00)

Accuracy:

- Base sMAPE: 25.36%
- Score sMAPE: 23.41%
- Improvement: +6.89%

Direction Correct: Base=True, Score=True

Winner: Score

Free_Cash_Flow (Tier 1):

Actual Value: 482,374,000.00
Base Forecast: 311,502,976.00 (Error: -170,871,024.00)
Score Forecast: 317,733,035.00 (Error: -164,640,965.00)

Accuracy:

- Base sMAPE: 43.05%
- Score sMAPE: 41.15%
- Improvement: +3.65%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 1,472,912,000.00
Base Forecast: 1,280,758,144.00 (Error: -192,153,856.00)
Score Forecast: 1,319,180,888.00 (Error: -153,731,112.00)

Accuracy:

- Base sMAPE: 13.96%
- Score sMAPE: 11.01%
- Improvement: +20.00%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 766,479,000.00
Base Forecast: 704,892,800.00 (Error: -61,586,200.00)
Score Forecast: 718,990,656.00 (Error: -47,488,344.00)

Accuracy:

- Base sMAPE: 8.37%
- Score sMAPE: 6.39%
- Improvement: +22.89%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 598,977,000.00

Base Forecast: 383,385,856.00 (Error: -215,591,144.00)

Score Forecast: 391,053,573.00 (Error: -207,923,427.00)

Accuracy:

- Base sMAPE: 43.89%
- Score sMAPE: 42.00%
- Improvement: +3.56%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -780,590,000.00

Base Forecast: -758,909,760.00 (Error: +21,680,240.00)

Score Forecast: -746,767,203.00 (Error: +33,822,797.00)

Accuracy:

- Base sMAPE: 2.82%
- Score sMAPE: 4.43%
- Improvement: -56.01%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 4,008,670,000.00

Base Forecast: 2,655,400,704.00 (Error: -1,353,269,296.00)

Score Forecast: 2,708,508,718.00 (Error: -1,300,161,282.00)

Accuracy:

- Base sMAPE: 40.61%

- Score sMAPE: 38.71%
- Improvement: +3.92%

Direction Correct: Base=True, Score=True

Winner: Score

=====

=====

RECOMMENDATIONS

=====

=====

△ Weak validation. Qualitative adjustments show marginal impact. Recommend:

1. Review and refine the qualitative scoring methodology
2. Investigate if certain metric types benefit more from adjustments
3. Consider whether the base model is already capturing qualitative factors
4. Test alternative adjustment algorithms

=====

=====

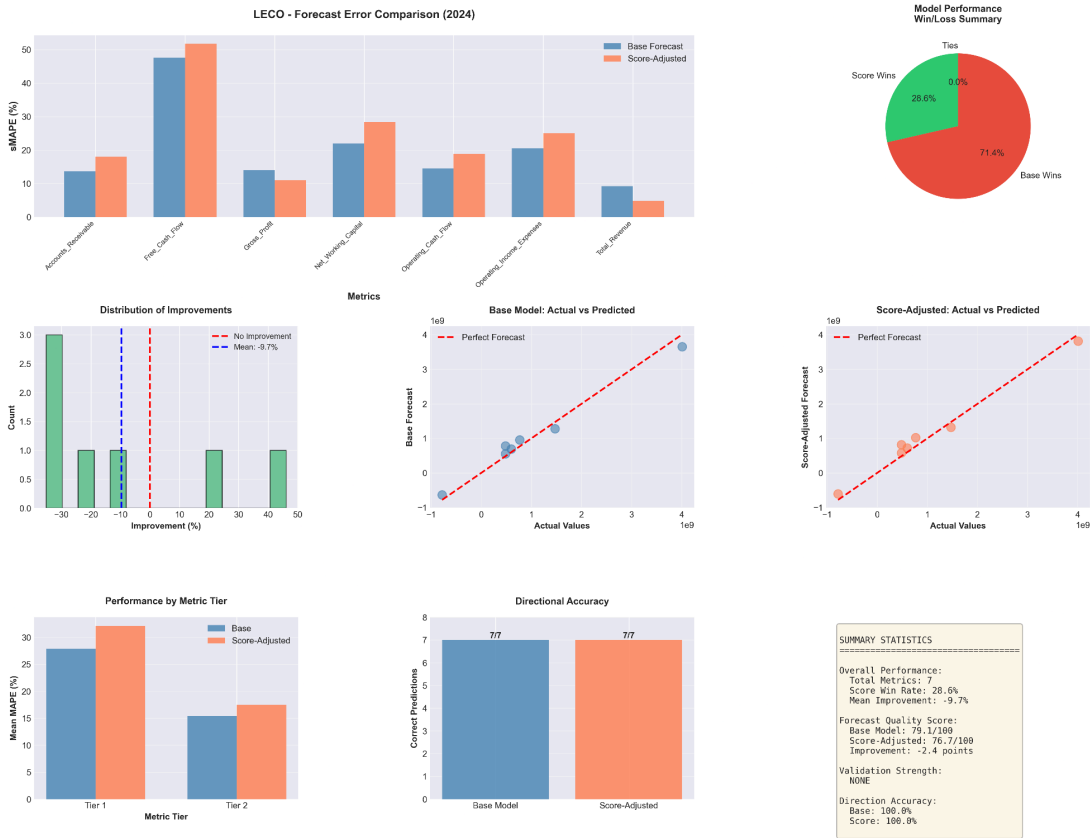
END OF REPORT

=====

=====

LECO (ARIMA):

Forecast Validation Analysis: LECO (2024)



FORECAST VALIDATION REPORT

Company: LECO
Forecast Year: 2024
Analysis Date: 2025-12-11 18:35:44

EXECUTIVE SUMMARY

Validation Strength: NONE

Score-Adjusted method does not validate. Base model performs as well or better.

Overall Win Rate: 28.6% (Score-Adjusted wins)

Mean Improvement: -9.7%

Forecast Quality Scores:

- Base Model: 79.1/100
- Score-Adjusted Model: 76.7/100
- Improvement: -2.4 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 22.58%
- Median MAPE: 15.63%
- Mean sMAPE: 20.22%

Score-Adjusted Model:

- Mean MAPE: 25.86%
- Median MAPE: 20.83%
- Mean sMAPE: 22.58%

Win/Loss Record:

- Score-Adjusted Wins: 2
- Base Model Wins: 5
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 27.92%
- Score Mean MAPE: 32.12%
- Mean Improvement: -8.21%
- Score Win Rate: 25.0%
- Wins: Score=1, Base=3

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 15.46%
- Score Mean MAPE: 17.52%
- Mean Improvement: -11.62%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

STATISTICAL SIGNIFICANCE TESTS

Paired T-Test:

- T-Statistic: -1.4819
- P-Value: 0.1889
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 6.0000
- P-Value: 0.2188
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: -0.5601
- Interpretation: Medium effect

METRIC-BY-METRIC BREAKDOWN

Accounts_Receivable (Tier 2):

Actual Value: 481,979,000.00
Base Forecast: 552,596,052.00 (Error: +70,617,052.00)
Score Forecast: 577,418,666.00 (Error: +95,439,666.00)

Accuracy:

- Base sMAPE: 13.65%
- Score sMAPE: 18.02%
- Improvement: -35.15%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 482,374,000.00
Base Forecast: 783,786,959.00 (Error: +301,412,959.00)
Score Forecast: 818,994,670.00 (Error: +336,620,670.00)

Accuracy:

- Base sMAPE: 47.61%
- Score sMAPE: 51.73%
- Improvement: -11.68%

Direction Correct: Base=True, Score=True

Winner: Base

Gross_Profit (Tier 2):

Actual Value: 1,472,912,000.00
Base Forecast: 1,280,199,667.00 (Error: -192,712,333.00)
Score Forecast: 1,318,554,449.00 (Error: -154,357,551.00)

Accuracy:

- Base sMAPE: 14.00%
- Score sMAPE: 11.06%
- Improvement: +19.90%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 766,479,000.00
Base Forecast: 955,652,987.00 (Error: +189,173,987.00)
Score Forecast: 1,020,063,999.00 (Error: +253,584,999.00)

Accuracy:

- Base sMAPE: 21.97%
- Score sMAPE: 28.39%
- Improvement: -34.05%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 598,977,000.00

Base Forecast: 692,624,817.00 (Error: +93,647,817.00)

Score Forecast: 723,737,523.00 (Error: +124,760,523.00)

Accuracy:

- Base sMAPE: 14.50%
- Score sMAPE: 18.86%
- Improvement: -33.22%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Income_Expenses (Tier 2):

Actual Value: -780,590,000.00

Base Forecast: -635,140,185.00 (Error: +145,449,815.00)

Score Forecast: -606,609,688.00 (Error: +173,980,312.00)

Accuracy:

- Base sMAPE: 20.55%
- Score sMAPE: 25.08%
- Improvement: -19.62%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 4,008,670,000.00

Base Forecast: 3,652,826,374.00 (Error: -355,843,626.00)

Score Forecast: 3,816,911,335.00 (Error: -191,758,665.00)

Accuracy:

- Base sMAPE: 9.29%

- Score sMAPE: 4.90%
- Improvement: +46.11%

Direction Correct: Base=True, Score=True

Winner: Score

```
=====
=====
RECOMMENDATIONS
=====
=====
```

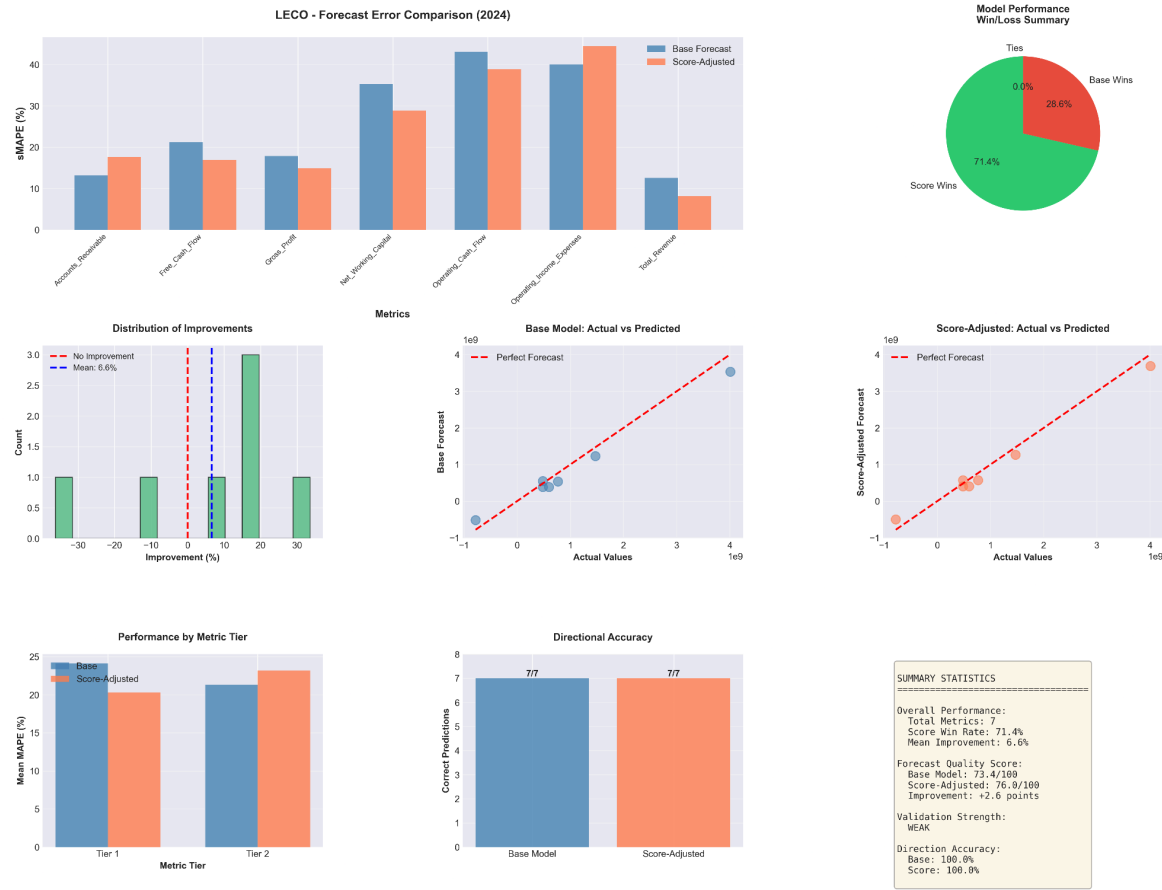
✗ No validation. Base model performs as well or better. Recommend:

1. Critically review the qualitative scoring framework
2. Check if qualitative factors are already reflected in historical data
3. Consider that market efficiency may limit qualitative alpha
4. Explore if timing or sector-specific factors affected results

```
=====
=====
END OF REPORT
=====
=====
```


LECO (WINTERS-HOLT):

Forecast Validation Analysis: LECO (2024)



FORECAST VALIDATION REPORT

Company: LECO
Forecast Year: 2024
Analysis Date: 2025-12-11 18:37:27

EXECUTIVE SUMMARY

=====

Validation Strength: WEAK

Score-Adjusted method shows weak validation. Qualitative factors provide marginal improvements.

Overall Win Rate: 71.4% (Score-Adjusted wins)

Mean Improvement: 6.6%

Forecast Quality Scores:

- Base Model: 73.4/100
- Score-Adjusted Model: 76.0/100
- Improvement: +2.6 points

=====

DETAILED METRICS COMPARISON

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 22.92%
- Median MAPE: 19.24%
- Mean sMAPE: 26.20%

Score-Adjusted Model:

- Mean MAPE: 21.55%
- Median MAPE: 19.29%
- Mean sMAPE: 24.27%

Win/Loss Record:

- Score-Adjusted Wins: 5
- Base Model Wins: 2
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 24.12%
- Score Mean MAPE: 20.32%
- Mean Improvement: 19.08%
- Score Win Rate: 100.0%
- Wins: Score=4, Base=0

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 21.32%
- Score Mean MAPE: 23.19%
- Mean Improvement: -9.97%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 1.1518
- P-Value: 0.2932
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 10.0000
- P-Value: 0.5781
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.4354
- Interpretation: Small effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 481,979,000.00

Base Forecast: 550,245,517.00 (Error: +68,266,517.00)

Score Forecast: 574,962,546.00 (Error: +92,983,546.00)

Accuracy:

- Base sMAPE: 13.23%

- Score sMAPE: 17.59%

- Improvement: -36.21%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 482,374,000.00

Base Forecast: 389,560,140.00 (Error: -92,813,860.00)

Score Forecast: 407,059,182.00 (Error: -75,314,818.00)

Accuracy:

- Base sMAPE: 21.29%

- Score sMAPE: 16.94%

- Improvement: +18.85%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 1,472,912,000.00

Base Forecast: 1,231,043,387.00 (Error: -241,868,613.00)

Score Forecast: 1,267,925,447.00 (Error: -204,986,553.00)

Accuracy:

- Base sMAPE: 17.89%

- Score sMAPE: 14.96%

- Improvement: +15.25%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 766,479,000.00

Base Forecast: 536,749,301.00 (Error: -229,729,699.00)

Score Forecast: 572,926,204.00 (Error: -193,552,796.00)

Accuracy:

- Base sMAPE: 35.26%
- Score sMAPE: 28.90%
- Improvement: +15.75%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 598,977,000.00

Base Forecast: 386,602,244.00 (Error: -212,374,756.00)

Score Forecast: 403,968,417.00 (Error: -195,008,583.00)

Accuracy:

- Base sMAPE: 43.10%
- Score sMAPE: 38.89%
- Improvement: +8.18%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -780,590,000.00

Base Forecast: -520,066,538.00 (Error: +260,523,462.00)

Score Forecast: -496,705,150.00 (Error: +283,884,850.00)

Accuracy:

- Base sMAPE: 40.06%
- Score sMAPE: 44.45%
- Improvement: -8.97%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 4,008,670,000.00

Base Forecast: 3,535,167,444.00 (Error: -473,502,556.00)

Score Forecast: 3,693,967,165.00 (Error: -314,702,835.00)

Accuracy:

- Base sMAPE: 12.55%
- Score sMAPE: 8.17%
- Improvement: +33.54%

Direction Correct: Base=True, Score=True

Winner: Score

```
=====
=====
RECOMMENDATIONS
=====
=====
```

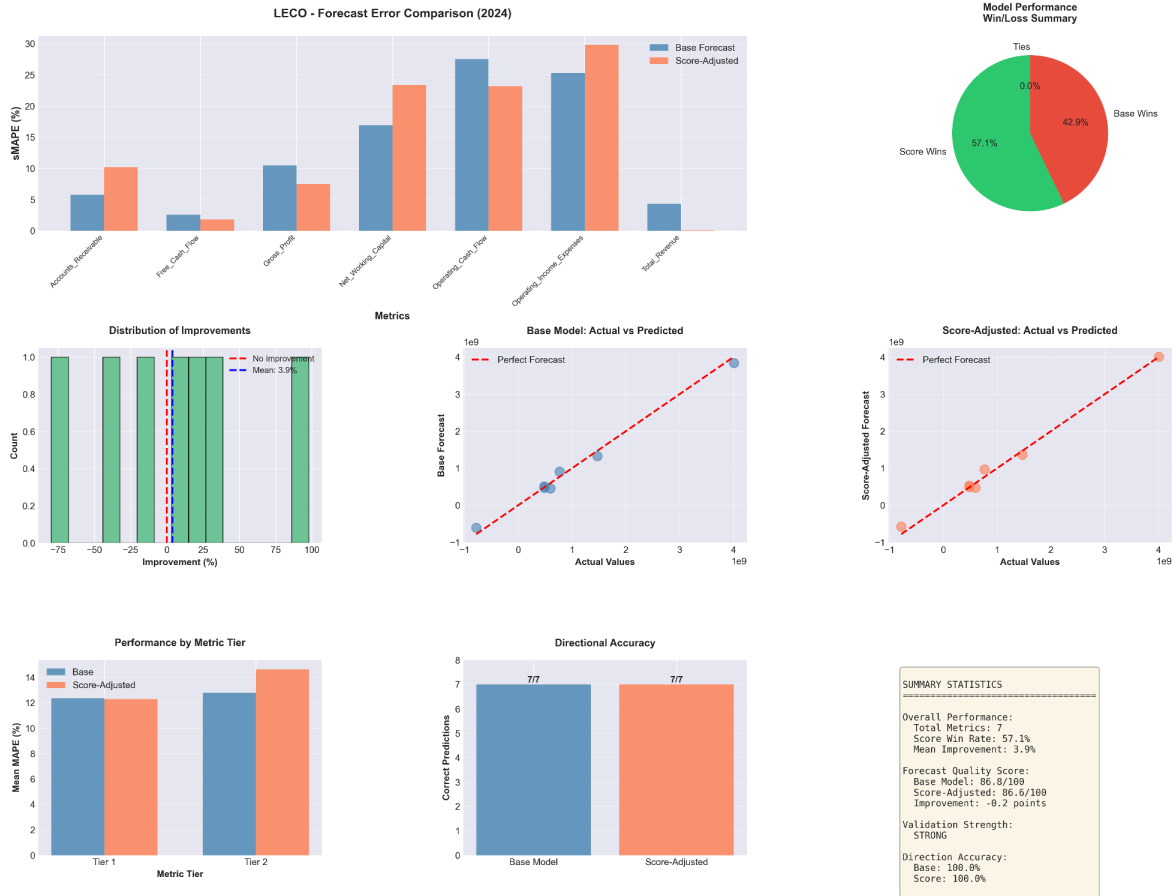
△ Weak validation. Qualitative adjustments show marginal impact. Recommend:

1. Review and refine the qualitative scoring methodology
2. Investigate if certain metric types benefit more from adjustments
3. Consider whether the base model is already capturing qualitative factors
4. Test alternative adjustment algorithms

```
=====
=====
END OF REPORT
=====
=====
```


LECO (LINEAR + WINTERS-HOLT):

Forecast Validation Analysis: LECO (2024)



FORECAST VALIDATION REPORT

Company: LECO
Forecast Year: 2024
Analysis Date: 2025-12-11 18:40:00

EXECUTIVE SUMMARY

=====

Validation Strength: **STRONG**

Score-Adjusted method shows strong validation. Qualitative factors significantly improve forecast accuracy.

Overall Win Rate: 57.1% (Score-Adjusted wins)

Mean Improvement: 3.9%

Forecast Quality Scores:

- Base Model: 86.8/100
- Score-Adjusted Model: 86.6/100
- Improvement: -0.2 points

=====

DETAILED METRICS COMPARISON

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 12.55%
- Median MAPE: 9.97%
- Mean sMAPE: 13.27%

Score-Adjusted Model:

- Mean MAPE: 13.31%
- Median MAPE: 10.73%
- Mean sMAPE: 13.72%

Win/Loss Record:

- Score-Adjusted Wins: 4
- Base Model Wins: 3
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 12.36%
- Score Mean MAPE: 12.30%
- Mean Improvement: 23.83%
- Score Win Rate: 75.0%
- Wins: Score=3, Base=1

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 12.80%
- Score Mean MAPE: 14.64%
- Mean Improvement: -22.74%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: -0.2598
- P-Value: 0.8037
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 10.0000
- P-Value: 0.5781
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: -0.0982
- Interpretation: Small effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 481,979,000.00

Base Forecast: 510,754,382.00 (Error: +28,775,382.00)

Score Forecast: 533,697,468.00 (Error: +51,718,468.00)

Accuracy:

- Base sMAPE: 5.80%
- Score sMAPE: 10.18%
- Improvement: -79.73%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 482,374,000.00

Base Forecast: 470,196,546.00 (Error: -12,177,454.00)

Score Forecast: 491,317,775.00 (Error: +8,943,775.00)

Accuracy:

- Base sMAPE: 2.56%
- Score sMAPE: 1.84%
- Improvement: +26.55%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 1,472,912,000.00

Base Forecast: 1,326,025,103.00 (Error: -146,886,897.00)

Score Forecast: 1,365,752,815.00 (Error: -107,159,185.00)

Accuracy:

- Base sMAPE: 10.50%
- Score sMAPE: 7.55%
- Improvement: +27.05%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 766,479,000.00

Base Forecast: 908,107,466.00 (Error: +141,628,466.00)

Score Forecast: 969,313,909.00 (Error: +202,834,909.00)

Accuracy:

- Base sMAPE: 16.92%
- Score sMAPE: 23.37%
- Improvement: -43.22%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 598,977,000.00

Base Forecast: 453,933,645.00 (Error: -145,043,355.00)

Score Forecast: 474,324,345.00 (Error: -124,652,655.00)

Accuracy:

- Base sMAPE: 27.55%
- Score sMAPE: 23.23%
- Improvement: +14.06%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -780,590,000.00

Base Forecast: -605,392,671.00 (Error: +175,197,329.00)

Score Forecast: -578,198,432.00 (Error: +202,391,568.00)

Accuracy:

- Base sMAPE: 25.28%
- Score sMAPE: 29.79%
- Improvement: -15.52%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 4,008,670,000.00

Base Forecast: 3,839,683,719.00 (Error: -168,986,281.00)

Score Forecast: 4,012,162,312.00 (Error: +3,492,312.00)

Accuracy:

- Base sMAPE: 4.31%
- Score sMAPE: 0.09%
- Improvement: +97.93%

Direction Correct: Base=True, Score=True

Winner: Score

=====

=====

RECOMMENDATIONS

=====

=====

✓ Strong validation achieved. Qualitative score adjustments significantly improve forecast accuracy. Recommend:

1. Continue using score-adjusted methodology for future forecasts
2. Analyze which qualitative factors drove the best improvements
3. Consider increasing weight given to qualitative adjustments
4. Document the successful factors for replication

=====

=====

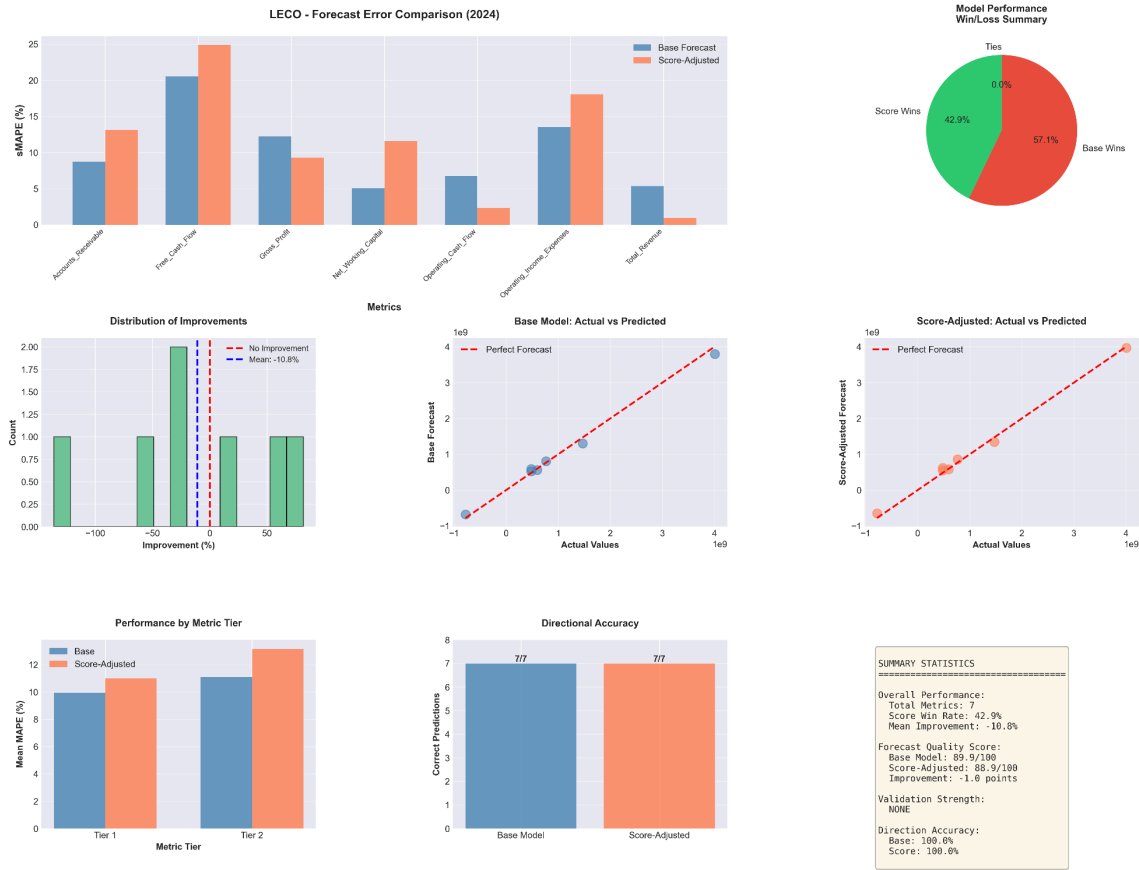
END OF REPORT

=====

=====

LECO (LINEAR + WINTERS-HOLT + XGBOOST):

Forecast Validation Analysis: LECO (2024)



FORECAST VALIDATION REPORT

Company: LECO
Forecast Year: 2024
Analysis Date: 2025-12-11 18:43:12

EXECUTIVE SUMMARY

Validation Strength: NONE

Score-Adjusted method does not validate. Base model performs as well or better.

Overall Win Rate: 42.9% (Score-Adjusted wins)

Mean Improvement: -10.8%

Forecast Quality Scores:

- Base Model: 89.9/100
- Score-Adjusted Model: 88.9/100
- Improvement: -1.0 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 10.45%
- Median MAPE: 9.13%
- Mean sMAPE: 10.31%

Score-Adjusted Model:

- Mean MAPE: 11.92%
- Median MAPE: 12.30%
- Mean sMAPE: 11.46%

Win/Loss Record:

- Score-Adjusted Wins: 3
- Base Model Wins: 4
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 9.95%
- Score Mean MAPE: 10.99%
- Mean Improvement: -3.51%
- Score Win Rate: 50.0%
- Wins: Score=2, Base=2

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 11.10%
- Score Mean MAPE: 13.16%
- Mean Improvement: -20.57%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

STATISTICAL SIGNIFICANCE TESTS

Paired T-Test:

- T-Statistic: -0.6331
- P-Value: 0.5500
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 10.0000
- P-Value: 0.5781
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: -0.2393
- Interpretation: Small effect

METRIC-BY-METRIC BREAKDOWN

Accounts_Receivable (Tier 2):

Actual Value: 481,979,000.00

Base Forecast: 525,965,255.00 (Error: +43,986,255.00)

Score Forecast: 549,591,614.00 (Error: +67,612,614.00)

Accuracy:

- Base sMAPE: 8.73%
- Score sMAPE: 13.11%
- Improvement: -53.71%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 482,374,000.00

Base Forecast: 592,773,121.00 (Error: +110,399,121.00)

Score Forecast: 619,400,489.00 (Error: +137,026,489.00)

Accuracy:

- Base sMAPE: 20.54%
- Score sMAPE: 24.87%
- Improvement: -24.12%

Direction Correct: Base=True, Score=True

Winner: Base

Gross_Profit (Tier 2):

Actual Value: 1,472,912,000.00

Base Forecast: 1,303,123,079.00 (Error: -169,788,921.00)

Score Forecast: 1,342,164,647.00 (Error: -130,747,353.00)

Accuracy:

- Base sMAPE: 12.23%
- Score sMAPE: 9.29%
- Improvement: +22.99%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 766,479,000.00

Base Forecast: 806,393,061.00 (Error: +39,914,061.00)

Score Forecast: 860,743,953.00 (Error: +94,264,953.00)

Accuracy:

- Base sMAPE: 5.08%
- Score sMAPE: 11.59%
- Improvement: -136.17%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 598,977,000.00

Base Forecast: 559,995,046.00 (Error: -38,981,954.00)

Score Forecast: 585,150,024.00 (Error: -13,826,976.00)

Accuracy:

- Base sMAPE: 6.73%
- Score sMAPE: 2.34%
- Improvement: +64.53%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -780,590,000.00

Base Forecast: -681,782,415.00 (Error: +98,807,585.00)

Score Forecast: -651,156,749.00 (Error: +129,433,251.00)

Accuracy:

- Base sMAPE: 13.51%
- Score sMAPE: 18.08%
- Improvement: -31.00%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 4,008,670,000.00

Base Forecast: 3,799,781,827.00 (Error: -208,888,173.00)

Score Forecast: 3,970,468,027.00 (Error: -38,201,973.00)

Accuracy:

- Base sMAPE: 5.35%
- Score sMAPE: 0.96%
- Improvement: +81.71%

Direction Correct: Base=True, Score=True
Winner: Score

=====

=====

RECOMMENDATIONS

=====

=====

- X No validation. Base model performs as well or better. Recommend:
1. Critically review the qualitative scoring framework
 2. Check if qualitative factors are already reflected in historical data
 3. Consider that market efficiency may limit qualitative alpha
 4. Explore if timing or sector-specific factors affected results

=====

=====

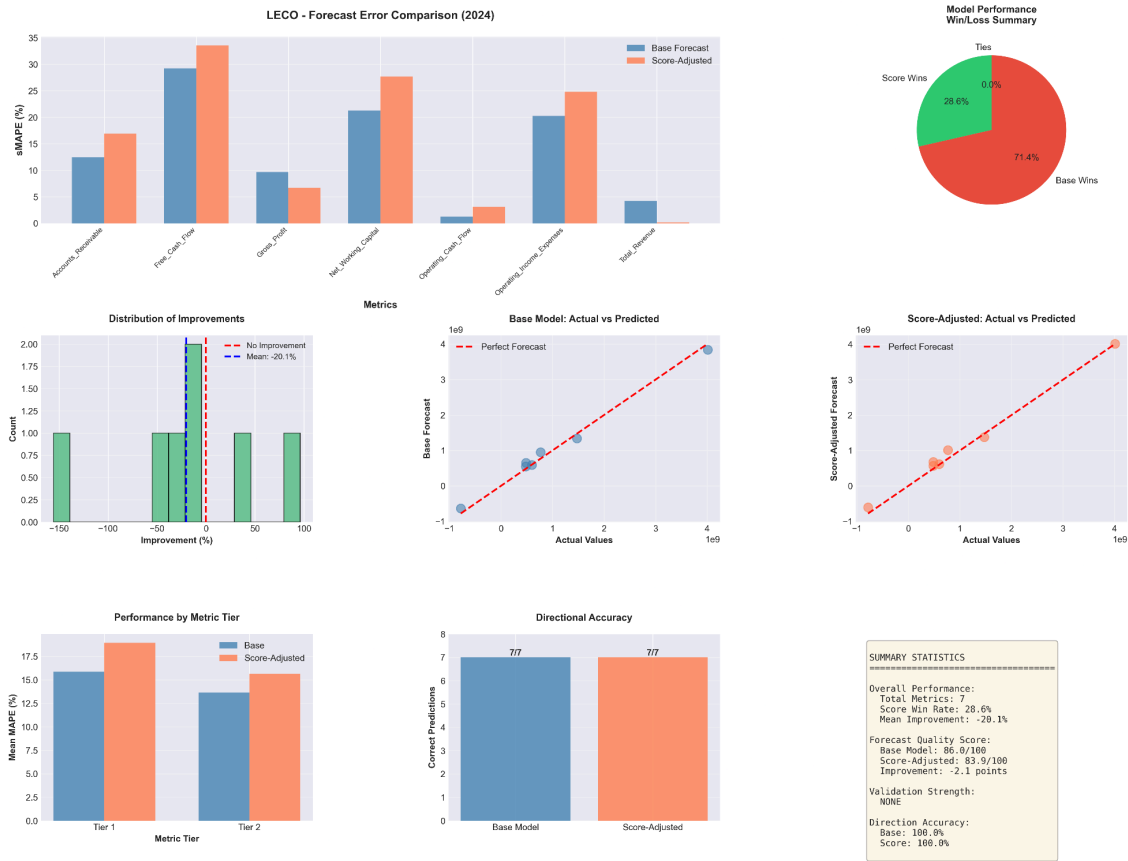
END OF REPORT

=====

=====

LECO (LINEAR + WINTERS-HOLT + ARIMA):

Forecast Validation Analysis: LECO (2024)



FORECAST VALIDATION REPORT

Company: LECO
Forecast Year: 2024
Analysis Date: 2025-12-11 18:45:08

EXECUTIVE SUMMARY

=====

Validation Strength: NONE

Score-Adjusted method does not validate. Base model performs as well or better.

Overall Win Rate: 28.6% (Score-Adjusted wins)

Mean Improvement: -20.1%

Forecast Quality Scores:

- Base Model: 86.0/100
- Score-Adjusted Model: 83.9/100
- Improvement: -2.1 points

=====

DETAILED METRICS COMPARISON

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 14.91%
- Median MAPE: 13.33%
- Mean sMAPE: 14.06%

Score-Adjusted Model:

- Mean MAPE: 17.54%
- Median MAPE: 18.42%
- Mean sMAPE: 16.13%

Win/Loss Record:

- Score-Adjusted Wins: 2
- Base Model Wins: 5
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 15.86%
- Score Mean MAPE: 18.94%
- Mean Improvement: -28.04%
- Score Win Rate: 25.0%
- Wins: Score=1, Base=3

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 13.66%
- Score Mean MAPE: 15.67%
- Mean Improvement: -9.56%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: -1.3520
- P-Value: 0.2251
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 5.0000
- P-Value: 0.1562
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: -0.5110
- Interpretation: Medium effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 481,979,000.00

Base Forecast: 546,217,218.00 (Error: +64,238,218.00)

Score Forecast: 570,753,295.00 (Error: +88,774,295.00)

Accuracy:

- Base sMAPE: 12.50%
- Score sMAPE: 16.87%
- Improvement: -38.20%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 482,374,000.00

Base Forecast: 647,617,725.00 (Error: +165,243,725.00)

Score Forecast: 676,708,713.00 (Error: +194,334,713.00)

Accuracy:

- Base sMAPE: 29.25%
- Score sMAPE: 33.53%
- Improvement: -17.60%

Direction Correct: Base=True, Score=True

Winner: Base

Gross_Profit (Tier 2):

Actual Value: 1,472,912,000.00

Base Forecast: 1,336,801,850.00 (Error: -136,110,150.00)

Score Forecast: 1,376,852,433.00 (Error: -96,059,567.00)

Accuracy:

- Base sMAPE: 9.69%
- Score sMAPE: 6.74%
- Improvement: +29.43%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 766,479,000.00
Base Forecast: 948,728,176.00 (Error: +182,249,176.00)
Score Forecast: 1,012,672,455.00 (Error: +246,193,455.00)

Accuracy:

- Base sMAPE: 21.25%
- Score sMAPE: 27.68%
- Improvement: -35.09%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 598,977,000.00
Base Forecast: 591,506,200.00 (Error: -7,470,800.00)
Score Forecast: 618,076,658.00 (Error: +19,099,658.00)

Accuracy:

- Base sMAPE: 1.26%
- Score sMAPE: 3.14%
- Improvement: -155.66%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Income_Expenses (Tier 2):

Actual Value: -780,590,000.00
Base Forecast: -636,980,643.00 (Error: +143,609,357.00)
Score Forecast: -608,367,473.00 (Error: +172,222,527.00)

Accuracy:

- Base sMAPE: 20.26%
- Score sMAPE: 24.80%
- Improvement: -19.92%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 4,008,670,000.00
Base Forecast: 3,842,382,056.00 (Error: -166,287,944.00)
Score Forecast: 4,014,981,858.00 (Error: +6,311,858.00)

Accuracy:

- Base sMAPE: 4.24%
- Score sMAPE: 0.16%
- Improvement: +96.20%

Direction Correct: Base=True, Score=True

Winner: Score

=====

=====

RECOMMENDATIONS

=====

=====

X No validation. Base model performs as well or better. Recommend:

1. Critically review the qualitative scoring framework
2. Check if qualitative factors are already reflected in historical data
3. Consider that market efficiency may limit qualitative alpha
4. Explore if timing or sector-specific factors affected results

=====

=====

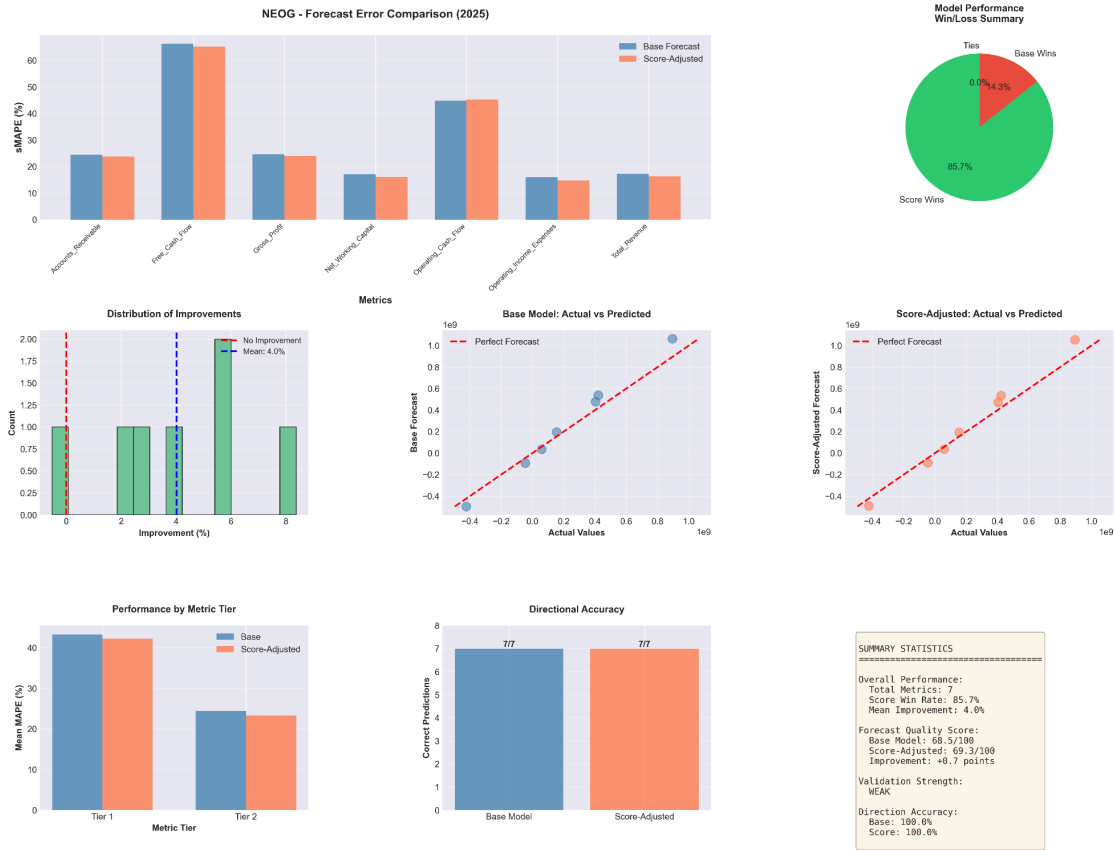
END OF REPORT

=====

=====

NEOG (NAIVE):

Forecast Validation Analysis: NEOG (2025)



FORECAST VALIDATION REPORT

Company: NEOG
Forecast Year: 2025
Analysis Date: 2025-11-21 13:46:13

EXECUTIVE SUMMARY

Validation Strength: WEAK

Score-Adjusted method shows weak validation. Qualitative factors provide marginal improvements.

Overall Win Rate: 85.7% (Score-Adjusted wins)

Mean Improvement: 4.0%

Forecast Quality Scores:

- Base Model: 68.5/100
- Score-Adjusted Model: 69.3/100
- Improvement: +0.7 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 35.21%
- Median MAPE: 27.92%
- Mean sMAPE: 30.06%

Score-Adjusted Model:

- Mean MAPE: 34.13%
- Median MAPE: 26.90%
- Mean sMAPE: 29.32%

Win/Loss Record:

- Score-Adjusted Wins: 6
- Base Model Wins: 1
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 43.30%
- Score Mean MAPE: 42.23%
- Mean Improvement: 3.30%
- Score Win Rate: 75.0%
- Wins: Score=3, Base=1

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 24.42%
- Score Mean MAPE: 23.33%
- Mean Improvement: 4.97%
- Score Win Rate: 100.0%
- Wins: Score=3, Base=0

STATISTICAL SIGNIFICANCE TESTS

Paired T-Test:

- T-Statistic: 3.9990
- P-Value: 0.0071
- Significant at $\alpha=0.05$: True
- Interpretation: Score-Adjusted significantly different

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 1.0000
- P-Value: 0.0312
- Significant at $\alpha=0.05$: True

Effect Size:

- Cohen's d: 1.5115
- Interpretation: Large effect

METRIC-BY-METRIC BREAKDOWN

Accounts_Receivable (Tier 2):

Actual Value: 153,384,000.00

Base Forecast: 196,205,181.00 (Error: +42,821,181.00)

Score Forecast: 194,641,953.00 (Error: +41,257,953.00)

Accuracy:

- Base sMAPE: 24.50%

- Score sMAPE: 23.71%

- Improvement: +3.65%

Direction Correct: Base=True, Score=True

Winner: Score

Free_Cash_Flow (Tier 1):

Actual Value: -46,351,000.00

Base Forecast: -92,255,712.00 (Error: -45,904,712.00)

Score Forecast: -91,170,981.00 (Error: -44,819,981.00)

Accuracy:

- Base sMAPE: 66.24%

- Score sMAPE: 65.18%

- Improvement: +2.36%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 421,376,000.00

Base Forecast: 539,296,175.00 (Error: +117,920,175.00)

Score Forecast: 535,907,871.00 (Error: +114,531,871.00)

Accuracy:

- Base sMAPE: 24.55%

- Score sMAPE: 23.93%

- Improvement: +2.87%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 402,926,000.00

Base Forecast: 477,954,749.00 (Error: +75,028,749.00)

Score Forecast: 473,602,925.00 (Error: +70,676,925.00)

Accuracy:

- Base sMAPE: 17.03%
- Score sMAPE: 16.13%
- Improvement: +5.80%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 58,244,000.00

Base Forecast: 36,883,493.00 (Error: -21,360,507.00)

Score Forecast: 36,774,371.00 (Error: -21,469,629.00)

Accuracy:

- Base sMAPE: 44.91%
- Score sMAPE: 45.19%
- Improvement: -0.51%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Income_Expenses (Tier 2):

Actual Value: -423,052,000.00

Base Forecast: -496,466,277.00 (Error: -73,414,277.00)

Score Forecast: -490,319,248.00 (Error: -67,267,248.00)

Accuracy:

- Base sMAPE: 15.97%
- Score sMAPE: 14.73%
- Improvement: +8.37%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 894,661,000.00

Base Forecast: 1,063,627,629.00 (Error: +168,966,629.00)

Score Forecast: 1,054,234,478.00 (Error: +159,573,478.00)

Accuracy:

- Base sMAPE: 17.26%
- Score sMAPE: 16.38%
- Improvement: +5.56%

Direction Correct: Base=True, Score=True
Winner: Score

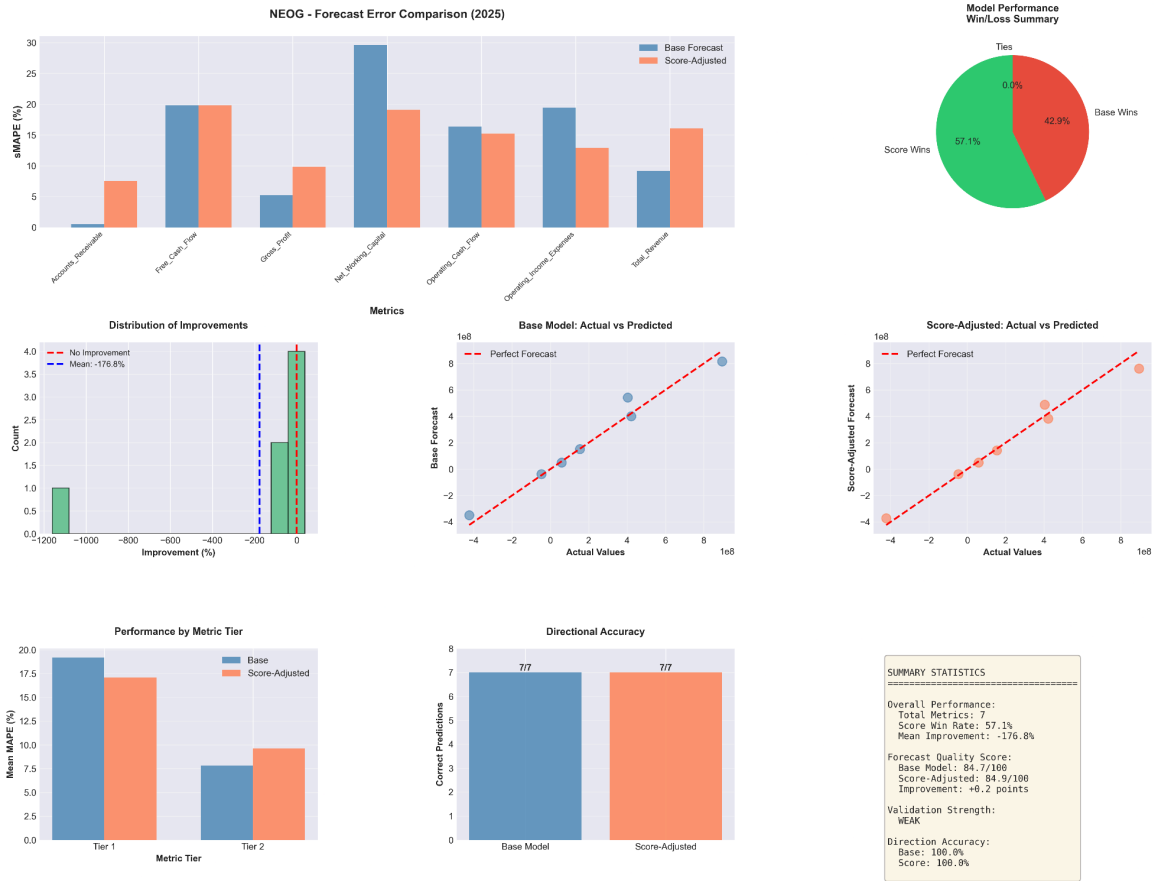
```
=====
=====
RECOMMENDATIONS
=====
=====
```

- △ Weak validation. Qualitative adjustments show marginal impact. Recommend:
1. Review and refine the qualitative scoring methodology
 2. Investigate if certain metric types benefit more from adjustments
 3. Consider whether the base model is already capturing qualitative factors
 4. Test alternative adjustment algorithms

```
=====
=====
END OF REPORT
=====
=====
```


NEOG (LINEAR):

Forecast Validation Analysis: NEOG (2025)



FORECAST VALIDATION REPORT

Company: NEOG
Forecast Year: 2025
Analysis Date: 2025-11-21 13:56:07

EXECUTIVE SUMMARY

Validation Strength: WEAK

Score-Adjusted method shows weak validation. Qualitative factors provide marginal improvements.

Overall Win Rate: 57.1% (Score-Adjusted wins)

Mean Improvement: -176.8%

Forecast Quality Scores:

- Base Model: 84.7/100
- Score-Adjusted Model: 84.9/100
- Improvement: +0.2 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 14.31%
- Median MAPE: 15.16%
- Mean sMAPE: 14.33%

Score-Adjusted Model:

- Mean MAPE: 13.87%
- Median MAPE: 14.19%
- Mean sMAPE: 14.39%

Win/Loss Record:

- Score-Adjusted Wins: 4
- Base Model Wins: 3
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 19.18%
- Score Mean MAPE: 17.07%
- Mean Improvement: -6.16%
- Score Win Rate: 75.0%
- Wins: Score=3, Base=1

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 7.80%
- Score Mean MAPE: 9.61%
- Mean Improvement: -404.22%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: -0.0231
- P-Value: 0.9823
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 14.0000
- P-Value: 1.0000
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: -0.0087
- Interpretation: Small effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 153,384,000.00

Base Forecast: 152,498,961.00 (Error: -885,039.00)

Score Forecast: 142,223,581.00 (Error: -11,160,419.00)

Accuracy:

- Base sMAPE: 0.58%
- Score sMAPE: 7.55%
- Improvement: -1161.01%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: -46,351,000.00

Base Forecast: -37,983,303.00 (Error: +8,367,697.00)

Score Forecast: -37,983,304.00 (Error: +8,367,696.00)

Accuracy:

- Base sMAPE: 19.84%
- Score sMAPE: 19.84%
- Improvement: +0.00%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 421,376,000.00

Base Forecast: 399,726,143.00 (Error: -21,649,857.00)

Score Forecast: 381,762,450.00 (Error: -39,613,550.00)

Accuracy:

- Base sMAPE: 5.27%
- Score sMAPE: 9.86%
- Improvement: -82.97%

Direction Correct: Base=True, Score=True

Winner: Base

Net_Working_Capital (Tier 1):

Actual Value: 402,926,000.00
Base Forecast: 542,971,338.00 (Error: +140,045,338.00)
Score Forecast: 488,076,936.00 (Error: +85,150,936.00)

Accuracy:

- Base sMAPE: 29.61%
- Score sMAPE: 19.11%
- Improvement: +39.20%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 58,244,000.00
Base Forecast: 49,413,680.00 (Error: -8,830,320.00)
Score Forecast: 49,976,301.00 (Error: -8,267,699.00)

Accuracy:

- Base sMAPE: 16.40%
- Score sMAPE: 15.28%
- Improvement: +6.37%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -423,052,000.00
Base Forecast: -348,171,463.00 (Error: +74,880,537.00)
Score Forecast: -371,631,257.00 (Error: +51,420,743.00)

Accuracy:

- Base sMAPE: 19.42%
- Score sMAPE: 12.94%
- Improvement: +31.33%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 894,661,000.00
Base Forecast: 816,323,336.00 (Error: -78,337,664.00)
Score Forecast: 761,319,469.00 (Error: -133,341,531.00)

Accuracy:

- Base sMAPE: 9.16%
- Score sMAPE: 16.10%
- Improvement: -70.21%

Direction Correct: Base=True, Score=True

Winner: Base

=====

=====

RECOMMENDATIONS

=====

=====

△ Weak validation. Qualitative adjustments show marginal impact. Recommend:

1. Review and refine the qualitative scoring methodology
2. Investigate if certain metric types benefit more from adjustments
3. Consider whether the base model is already capturing qualitative factors
4. Test alternative adjustment algorithms

=====

=====

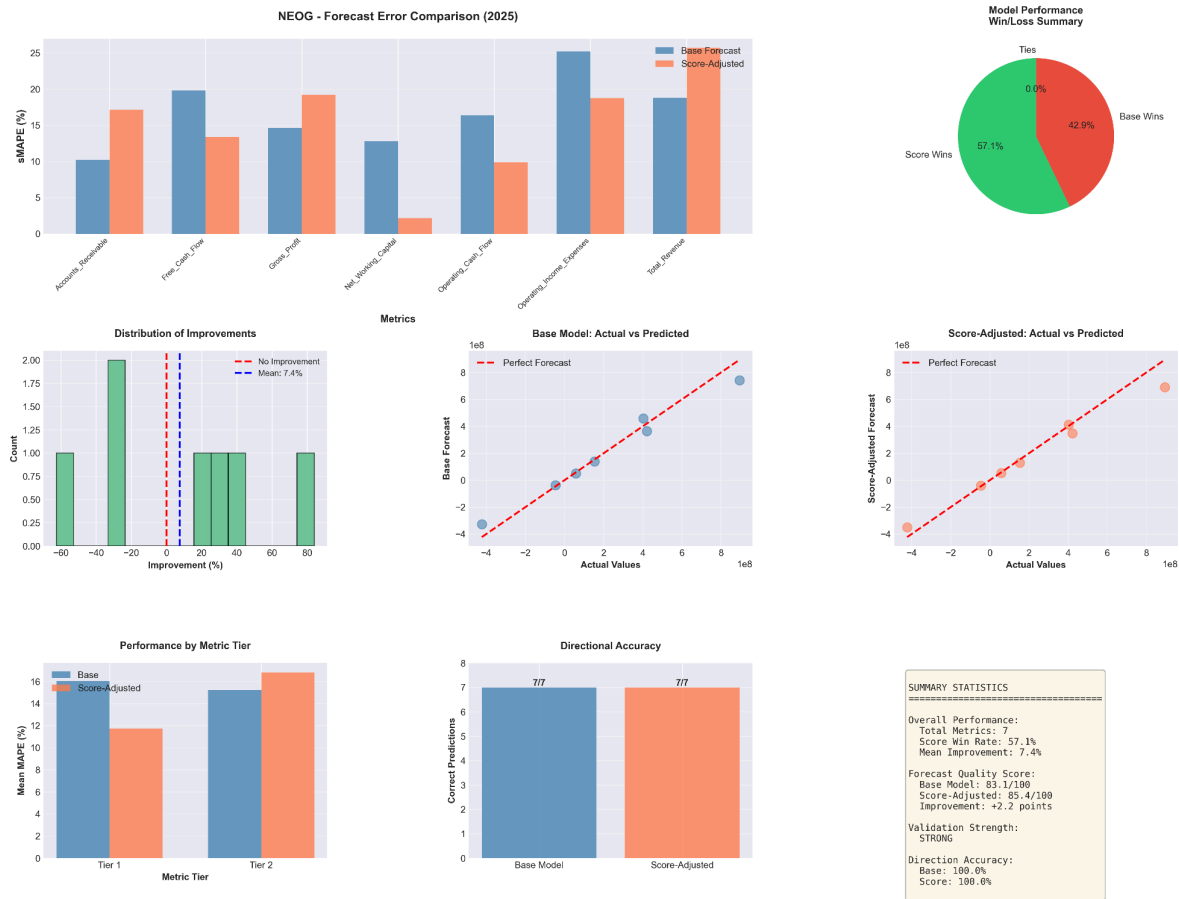
END OF REPORT

=====

=====

NEOG (XGBOOST):

Forecast Validation Analysis: NEOG (2025)



FORECAST VALIDATION REPORT

Company: NEOG
Forecast Year: 2025
Analysis Date: 2025-11-21 13:59:20

EXECUTIVE SUMMARY

=====

Validation Strength: **STRONG**

Score-Adjusted method shows strong validation. Qualitative factors significantly improve forecast accuracy.

Overall Win Rate: 57.1% (Score-Adjusted wins)

Mean Improvement: 7.4%

Forecast Quality Scores:

- Base Model: 83.1/100
- Score-Adjusted Model: 85.4/100
- Improvement: +2.2 points

=====

DETAILED METRICS COMPARISON

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 15.69%
- Median MAPE: 15.16%
- Mean sMAPE: 16.85%

Score-Adjusted Model:

- Mean MAPE: 13.92%
- Median MAPE: 15.80%
- Mean sMAPE: 15.18%

Win/Loss Record:

- Score-Adjusted Wins: 4
- Base Model Wins: 3
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 16.02%
- Score Mean MAPE: 11.73%
- Mean Improvement: 29.95%
- Score Win Rate: 75.0%
- Wins: Score=3, Base=1

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 15.25%
- Score Mean MAPE: 16.83%
- Mean Improvement: -22.56%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 0.5887
- P-Value: 0.5776
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 12.0000
- P-Value: 0.8125
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.2225
- Interpretation: Small effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 153,384,000.00

Base Forecast: 138,473,717.00 (Error: -14,910,283.00)

Score Forecast: 129,143,358.00 (Error: -24,240,642.00)

Accuracy:

- Base sMAPE: 10.22%
- Score sMAPE: 17.16%
- Improvement: -62.58%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: -46,351,000.00

Base Forecast: -37,983,303.00 (Error: +8,367,697.00)

Score Forecast: -40,542,618.00 (Error: +5,808,382.00)

Accuracy:

- Base sMAPE: 19.84%
- Score sMAPE: 13.37%
- Improvement: +30.59%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 421,376,000.00

Base Forecast: 363,907,334.00 (Error: -57,468,666.00)

Score Forecast: 347,553,338.00 (Error: -73,822,662.00)

Accuracy:

- Base sMAPE: 14.64%
- Score sMAPE: 19.20%
- Improvement: -28.46%

Direction Correct: Base=True, Score=True

Winner: Base

Net_Working_Capital (Tier 1):

Actual Value: 402,926,000.00

Base Forecast: 458,076,107.00 (Error: +55,150,107.00)

Score Forecast: 411,764,612.00 (Error: +8,838,612.00)

Accuracy:

- Base sMAPE: 12.81%
- Score sMAPE: 2.17%
- Improvement: +83.97%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 58,244,000.00

Base Forecast: 49,413,680.00 (Error: -8,830,320.00)

Score Forecast: 52,743,173.00 (Error: -5,500,827.00)

Accuracy:

- Base sMAPE: 16.40%
- Score sMAPE: 9.91%
- Improvement: +37.71%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -423,052,000.00

Base Forecast: -328,340,368.00 (Error: +94,711,632.00)

Score Forecast: -350,463,941.00 (Error: +72,588,059.00)

Accuracy:

- Base sMAPE: 25.21%
- Score sMAPE: 18.77%
- Improvement: +23.36%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 894,661,000.00

Base Forecast: 740,901,321.00 (Error: -153,759,679.00)
Score Forecast: 690,979,390.00 (Error: -203,681,610.00)

Accuracy:

- Base sMAPE: 18.80%
- Score sMAPE: 25.69%
- Improvement: -32.47%

Direction Correct: Base=True, Score=True

Winner: Base

=====

=====

RECOMMENDATIONS

=====

=====

✓ Strong validation achieved. Qualitative score adjustments significantly improve forecast accuracy. Recommend:

1. Continue using score-adjusted methodology for future forecasts
2. Analyze which qualitative factors drove the best improvements
3. Consider increasing weight given to qualitative adjustments
4. Document the successful factors for replication

=====

=====

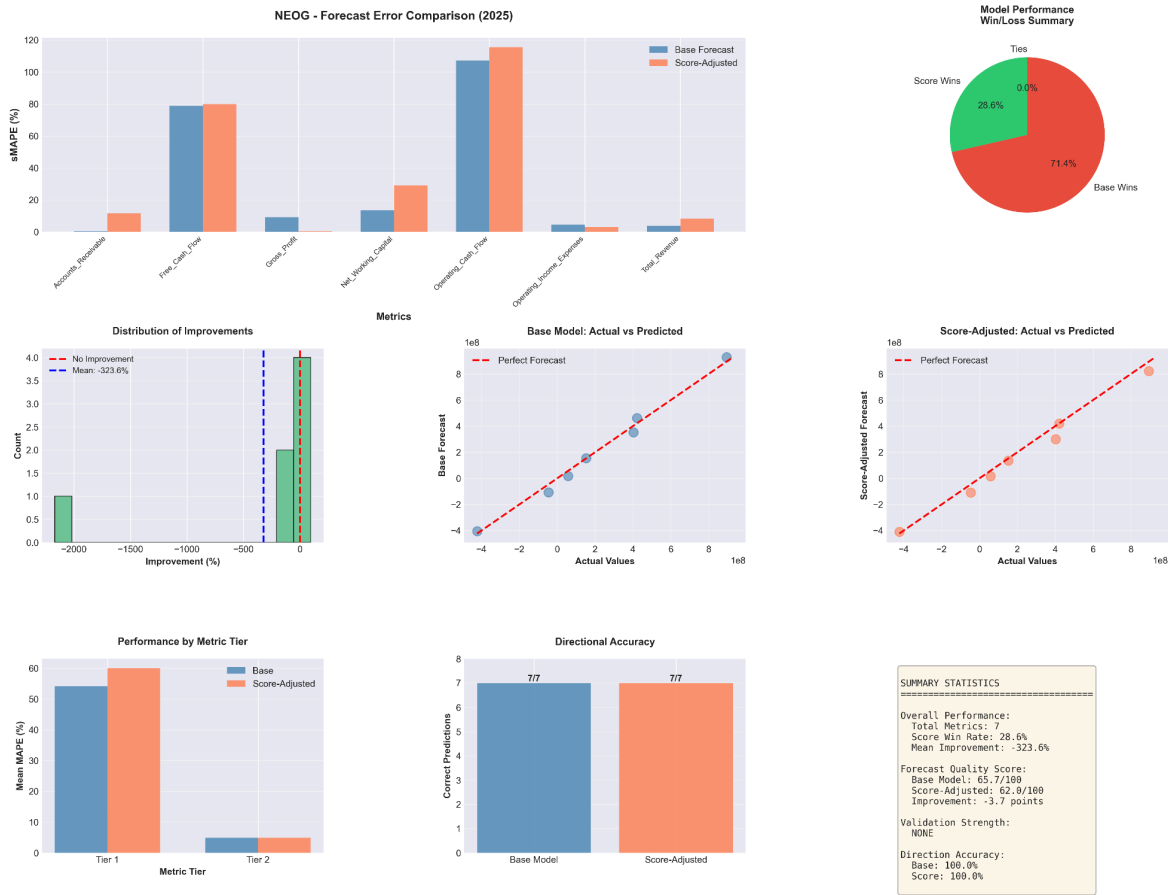
END OF REPORT

=====

=====

NEOG (ARIMA):

Forecast Validation Analysis: NEOG (2025)



FORECAST VALIDATION REPORT

Company: NEOG
Forecast Year: 2025
Analysis Date: 2025-11-21 13:58:24

EXECUTIVE SUMMARY

=====

Validation Strength: NONE

Score-Adjusted method does not validate. Base model performs as well or better.

Overall Win Rate: 28.6% (Score-Adjusted wins)

Mean Improvement: -323.6%

Forecast Quality Scores:

- Base Model: 65.7/100
- Score-Adjusted Model: 62.0/100
- Improvement: -3.7 points

=====

DETAILED METRICS COMPARISON

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 33.02%
- Median MAPE: 9.68%
- Mean sMAPE: 31.09%

Score-Adjusted Model:

- Mean MAPE: 36.40%
- Median MAPE: 10.97%
- Mean sMAPE: 35.49%

Win/Loss Record:

- Score-Adjusted Wins: 2
- Base Model Wins: 5
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 54.12%
- Score Mean MAPE: 60.04%
- Mean Improvement: -55.38%
- Score Win Rate: 0.0%
- Wins: Score=0, Base=4

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 4.88%
- Score Mean MAPE: 4.87%
- Mean Improvement: -681.30%
- Score Win Rate: 66.7%
- Wins: Score=2, Base=1

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: -1.4200
- P-Value: 0.2054
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 7.0000
- P-Value: 0.2969
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: -0.5367
- Interpretation: Medium effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 153,384,000.00

Base Forecast: 154,125,855.00 (Error: +741,855.00)

Score Forecast: 136,553,812.00 (Error: -16,830,188.00)

Accuracy:

- Base sMAPE: 0.48%
- Score sMAPE: 11.61%
- Improvement: -2168.66%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: -46,351,000.00

Base Forecast: -106,714,996.00 (Error: -60,363,996.00)

Score Forecast: -108,210,180.00 (Error: -61,859,180.00)

Accuracy:

- Base sMAPE: 78.87%
- Score sMAPE: 80.04%
- Improvement: -2.48%

Direction Correct: Base=True, Score=True

Winner: Base

Gross_Profit (Tier 2):

Actual Value: 421,376,000.00

Base Forecast: 462,166,710.00 (Error: +40,790,710.00)

Score Forecast: 419,327,091.00 (Error: -2,048,909.00)

Accuracy:

- Base sMAPE: 9.23%
- Score sMAPE: 0.49%
- Improvement: +94.98%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 402,926,000.00
Base Forecast: 351,930,886.00 (Error: -50,995,114.00)
Score Forecast: 300,533,139.00 (Error: -102,392,861.00)

Accuracy:

- Base sMAPE: 13.51%
- Score sMAPE: 29.11%
- Improvement: -100.79%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 58,244,000.00
Base Forecast: 17,587,920.00 (Error: -40,656,080.00)
Score Forecast: 15,582,703.00 (Error: -42,661,297.00)

Accuracy:

- Base sMAPE: 107.23%
- Score sMAPE: 115.57%
- Improvement: -4.93%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Income_Expenses (Tier 2):

Actual Value: -423,052,000.00
Base Forecast: -404,054,251.00 (Error: +18,997,749.00)
Score Forecast: -409,715,455.00 (Error: +13,336,545.00)

Accuracy:

- Base sMAPE: 4.59%
- Score sMAPE: 3.20%
- Improvement: +29.80%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 894,661,000.00
Base Forecast: 928,446,693.00 (Error: +33,785,693.00)
Score Forecast: 822,593,557.00 (Error: -72,067,443.00)

Accuracy:

- Base sMAPE: 3.71%
- Score sMAPE: 8.39%
- Improvement: -113.31%

Direction Correct: Base=True, Score=True

Winner: Base

=====

=====

RECOMMENDATIONS

=====

=====

X No validation. Base model performs as well or better. Recommend:

1. Critically review the qualitative scoring framework
2. Check if qualitative factors are already reflected in historical data
3. Consider that market efficiency may limit qualitative alpha
4. Explore if timing or sector-specific factors affected results

=====

=====

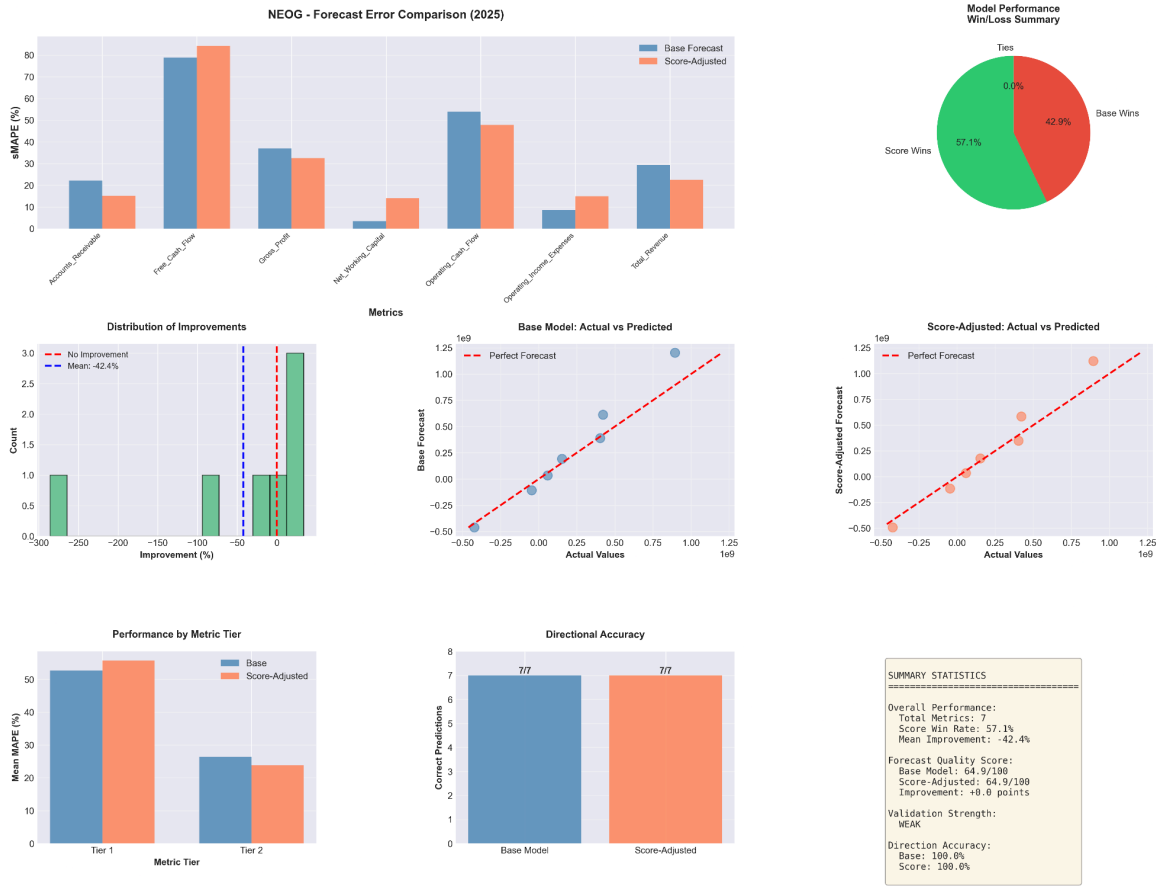
END OF REPORT

=====

=====

NEOG (WINTERS-HOLT):

Forecast Validation Analysis: NEOG (2025)



FORECAST VALIDATION REPORT

Company: NEOG
Forecast Year: 2025
Analysis Date: 2025-11-21 14:02:01

EXECUTIVE SUMMARY

Validation Strength: WEAK

Score-Adjusted method shows weak validation. Qualitative factors provide marginal improvements.

Overall Win Rate: 57.1% (Score-Adjusted wins)

Mean Improvement: -42.4%

Forecast Quality Scores:

- Base Model: 64.9/100
- Score-Adjusted Model: 64.9/100
- Improvement: +0.0 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 41.38%
- Median MAPE: 34.50%
- Mean sMAPE: 33.32%

Score-Adjusted Model:

- Mean MAPE: 42.06%
- Median MAPE: 25.44%
- Mean sMAPE: 33.07%

Win/Loss Record:

- Score-Adjusted Wins: 4
- Base Model Wins: 3
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 52.65%
- Score Mean MAPE: 55.73%
- Mean Improvement: -65.55%
- Score Win Rate: 50.0%
- Wins: Score=2, Base=2

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 26.37%
- Score Mean MAPE: 23.83%
- Mean Improvement: -11.43%
- Score Win Rate: 66.7%
- Wins: Score=2, Base=1

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 0.0890
- P-Value: 0.9320
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 13.0000
- P-Value: 0.9375
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.0336
- Interpretation: Small effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 153,384,000.00

Base Forecast: 191,503,777.00 (Error: +38,119,777.00)

Score Forecast: 178,600,253.00 (Error: +25,216,253.00)

Accuracy:

- Base sMAPE: 22.11%
- Score sMAPE: 15.19%
- Improvement: +33.85%

Direction Correct: Base=True, Score=True

Winner: Score

Free_Cash_Flow (Tier 1):

Actual Value: -46,351,000.00

Base Forecast: -106,714,996.00 (Error: -60,363,996.00)

Score Forecast: -113,905,452.00 (Error: -67,554,452.00)

Accuracy:

- Base sMAPE: 78.87%
- Score sMAPE: 84.31%
- Improvement: -11.91%

Direction Correct: Base=True, Score=True

Winner: Base

Gross_Profit (Tier 2):

Actual Value: 421,376,000.00

Base Forecast: 612,509,789.00 (Error: +191,133,789.00)

Score Forecast: 584,983,599.00 (Error: +163,607,599.00)

Accuracy:

- Base sMAPE: 36.97%
- Score sMAPE: 32.51%
- Improvement: +14.40%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 402,926,000.00
Base Forecast: 389,154,721.00 (Error: -13,771,279.00)
Score Forecast: 349,811,179.00 (Error: -53,114,821.00)

Accuracy:

- Base sMAPE: 3.48%
- Score sMAPE: 14.11%
- Improvement: -285.69%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 58,244,000.00
Base Forecast: 33,528,822.00 (Error: -24,715,178.00)
Score Forecast: 35,787,994.00 (Error: -22,456,006.00)

Accuracy:

- Base sMAPE: 53.86%
- Score sMAPE: 47.76%
- Improvement: +9.14%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -423,052,000.00
Base Forecast: -460,657,558.00 (Error: -37,605,558.00)
Score Forecast: -491,696,664.00 (Error: -68,644,664.00)

Accuracy:

- Base sMAPE: 8.51%
- Score sMAPE: 15.01%
- Improvement: -82.54%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 894,661,000.00
Base Forecast: 1,203,313,149.00 (Error: +308,652,149.00)
Score Forecast: 1,122,233,909.00 (Error: +227,572,909.00)

Accuracy:

- Base sMAPE: 29.42%
- Score sMAPE: 22.57%
- Improvement: +26.27%

Direction Correct: Base=True, Score=True

Winner: Score

```
=====
=====
RECOMMENDATIONS
=====
=====
```

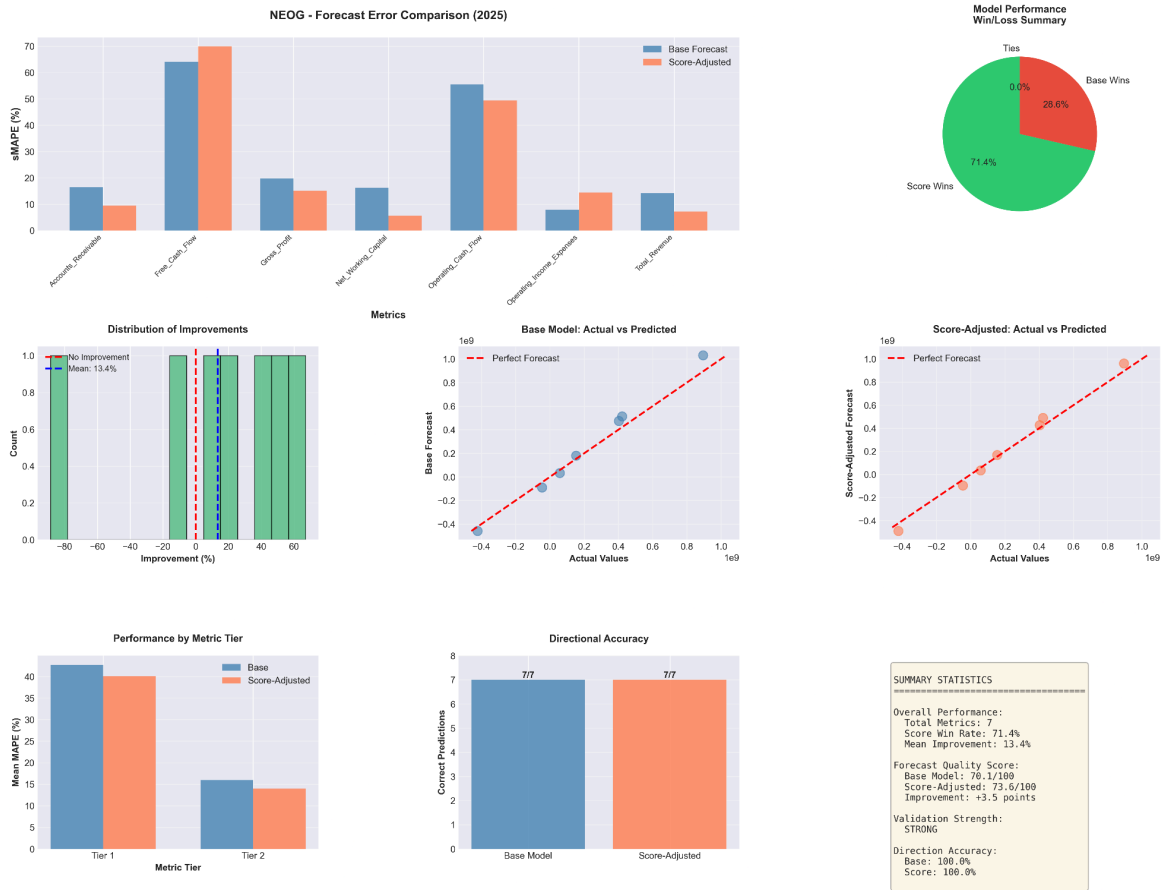
△ Weak validation. Qualitative adjustments show marginal impact. Recommend:

1. Review and refine the qualitative scoring methodology
2. Investigate if certain metric types benefit more from adjustments
3. Consider whether the base model is already capturing qualitative factors
4. Test alternative adjustment algorithms

```
=====
=====
END OF REPORT
=====
=====
```


NEOG (LINEAR + WINTERS-HOLT):

Forecast Validation Analysis: NEOG (2025)



FORECAST VALIDATION REPORT

Company: NEOG
Forecast Year: 2025
Analysis Date: 2025-11-21 14:24:14

EXECUTIVE SUMMARY

=====

Validation Strength: **STRONG**

Score-Adjusted method shows strong validation. Qualitative factors significantly improve forecast accuracy.

Overall Win Rate: 71.4% (Score-Adjusted wins)

Mean Improvement: 13.4%

Forecast Quality Scores:

- Base Model: 70.1/100
- Score-Adjusted Model: 73.6/100
- Improvement: +3.5 points

=====

DETAILED METRICS COMPARISON

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 31.26%
- Median MAPE: 17.87%
- Mean sMAPE: 27.74%

Score-Adjusted Model:

- Mean MAPE: 28.90%
- Median MAPE: 15.52%
- Mean sMAPE: 24.47%

Win/Loss Record:

- Score-Adjusted Wins: 5
- Base Model Wins: 2
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 42.71%
- Score Mean MAPE: 40.11%
- Mean Improvement: 28.19%
- Score Win Rate: 75.0%
- Wins: Score=3, Base=1

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 16.00%
- Score Mean MAPE: 13.96%
- Mean Improvement: -6.38%
- Score Win Rate: 66.7%
- Wins: Score=2, Base=1

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 1.2930
- P-Value: 0.2436
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 6.0000
- P-Value: 0.2188
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.4887
- Interpretation: Small effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 153,384,000.00

Base Forecast: 180,800,333.00 (Error: +27,416,333.00)

Score Forecast: 168,618,007.00 (Error: +15,234,007.00)

Accuracy:

- Base sMAPE: 16.41%

- Score sMAPE: 9.46%

- Improvement: +44.43%

Direction Correct: Base=True, Score=True

Winner: Score

Free_Cash_Flow (Tier 1):

Actual Value: -46,351,000.00

Base Forecast: -90,086,276.00 (Error: -43,735,276.00)

Score Forecast: -96,156,290.00 (Error: -49,805,290.00)

Accuracy:

- Base sMAPE: 64.11%

- Score sMAPE: 69.90%

- Improvement: -13.88%

Direction Correct: Base=True, Score=True

Winner: Base

Gross_Profit (Tier 2):

Actual Value: 421,376,000.00

Base Forecast: 513,617,575.00 (Error: +92,241,575.00)

Score Forecast: 490,535,601.00 (Error: +69,159,601.00)

Accuracy:

- Base sMAPE: 19.73%

- Score sMAPE: 15.17%

- Improvement: +25.02%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 402,926,000.00

Base Forecast: 474,401,440.00 (Error: +71,475,440.00)

Score Forecast: 426,439,454.00 (Error: +23,513,454.00)

Accuracy:

- Base sMAPE: 16.29%
- Score sMAPE: 5.67%
- Improvement: +67.10%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 58,244,000.00

Base Forecast: 32,944,545.00 (Error: -25,299,455.00)

Score Forecast: 35,164,349.00 (Error: -23,079,651.00)

Accuracy:

- Base sMAPE: 55.49%
- Score sMAPE: 49.42%
- Improvement: +8.77%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -423,052,000.00

Base Forecast: -457,877,275.00 (Error: -34,825,275.00)

Score Forecast: -488,729,046.00 (Error: -65,677,046.00)

Accuracy:

- Base sMAPE: 7.91%
- Score sMAPE: 14.41%
- Improvement: -88.59%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 894,661,000.00

Base Forecast: 1,031,649,106.00 (Error: +136,988,106.00)

Score Forecast: 962,136,589.00 (Error: +67,475,589.00)

Accuracy:

- Base sMAPE: 14.22%
- Score sMAPE: 7.27%
- Improvement: +50.74%

Direction Correct: Base=True, Score=True

Winner: Score

=====

=====

RECOMMENDATIONS

=====

=====

✓ Strong validation achieved. Qualitative score adjustments significantly improve forecast accuracy. Recommend:

1. Continue using score-adjusted methodology for future forecasts
2. Analyze which qualitative factors drove the best improvements
3. Consider increasing weight given to qualitative adjustments
4. Document the successful factors for replication

=====

=====

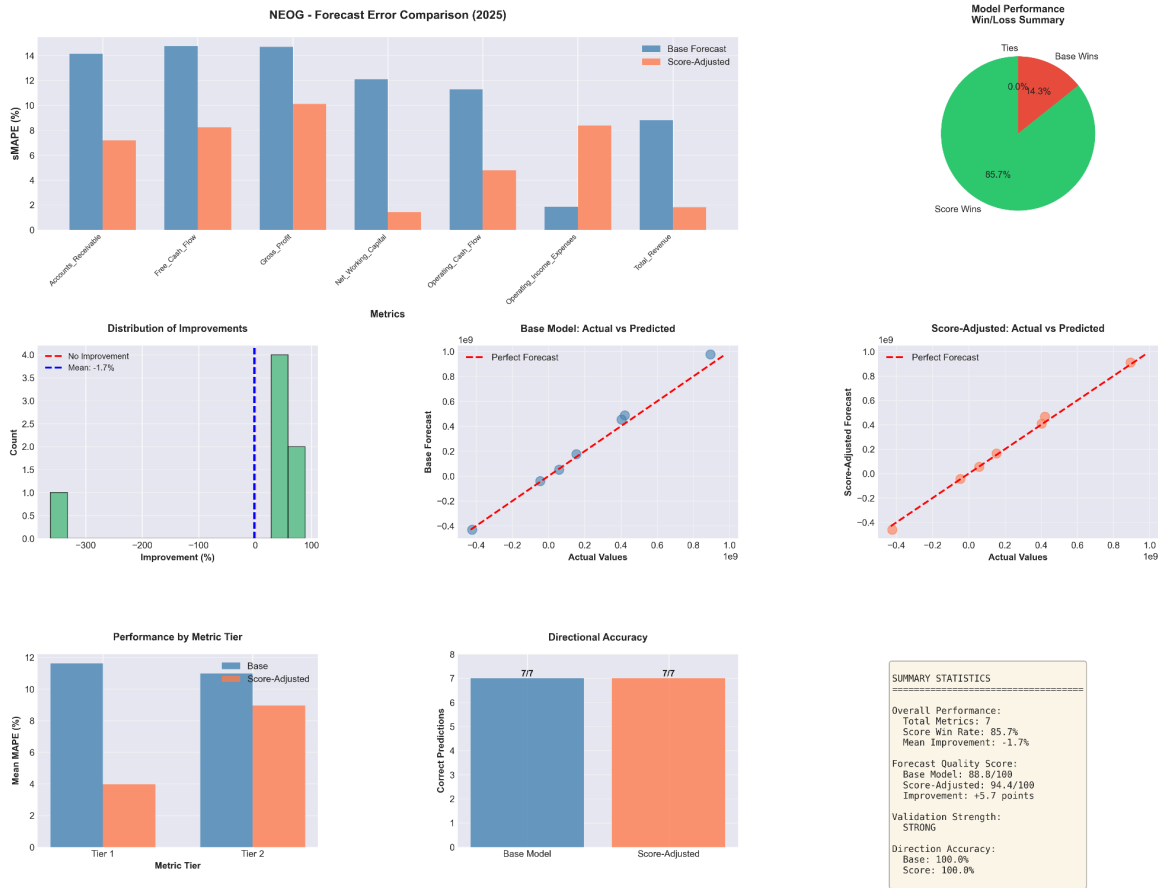
END OF REPORT

=====

=====

NEOG (LINEAR + WINTERS-HOLT + XGBOOST):

Forecast Validation Analysis: NEOG (2025)



FORECAST VALIDATION REPORT

Company: NEOG
Forecast Year: 2025
Analysis Date: 2025-11-21 14:24:53

EXECUTIVE SUMMARY

Validation Strength: STRONG

Score-Adjusted method shows strong validation. Qualitative factors significantly improve forecast accuracy.

Overall Win Rate: 85.7% (Score-Adjusted wins)

Mean Improvement: -1.7%

Forecast Quality Scores:

- Base Model: 88.8/100
- Score-Adjusted Model: 94.4/100
- Improvement: +5.7 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 11.35%
- Median MAPE: 12.87%
- Mean sMAPE: 11.09%

Score-Adjusted Model:

- Mean MAPE: 6.11%
- Median MAPE: 7.46%
- Mean sMAPE: 6.00%

Win/Loss Record:

- Score-Adjusted Wins: 6
- Base Model Wins: 1
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 11.63%
- Score Mean MAPE: 3.98%
- Mean Improvement: 66.80%
- Score Win Rate: 100.0%
- Wins: Score=4, Base=0

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 10.99%
- Score Mean MAPE: 8.96%
- Mean Improvement: -92.99%
- Score Win Rate: 66.7%
- Wins: Score=2, Base=1

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 2.4816
- P-Value: 0.0477
- Significant at $\alpha=0.05$: True
- Interpretation: Score-Adjusted significantly different

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 4.0000
- P-Value: 0.1094
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.9379
- Interpretation: Large effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 153,384,000.00

Base Forecast: 176,727,046.00 (Error: +23,343,046.00)

Score Forecast: 164,819,178.00 (Error: +11,435,178.00)

Accuracy:

- Base sMAPE: 14.14%
- Score sMAPE: 7.19%
- Improvement: +51.01%

Direction Correct: Base=True, Score=True

Winner: Score

Free_Cash_Flow (Tier 1):

Actual Value: -46,351,000.00

Base Forecast: -39,982,425.00 (Error: +6,368,575.00)

Score Forecast: -42,676,440.00 (Error: +3,674,560.00)

Accuracy:

- Base sMAPE: 14.75%
- Score sMAPE: 8.25%
- Improvement: +42.30%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 421,376,000.00

Base Forecast: 488,231,811.00 (Error: +66,855,811.00)

Score Forecast: 466,290,674.00 (Error: +44,914,674.00)

Accuracy:

- Base sMAPE: 14.70%
- Score sMAPE: 10.12%
- Improvement: +32.82%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 402,926,000.00
Base Forecast: 454,780,480.00 (Error: +51,854,480.00)
Score Forecast: 408,802,173.00 (Error: +5,876,173.00)

Accuracy:

- Base sMAPE: 12.09%
- Score sMAPE: 1.45%
- Improvement: +88.67%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 58,244,000.00
Base Forecast: 52,014,400.00 (Error: -6,229,600.00)
Score Forecast: 55,519,130.00 (Error: -2,724,870.00)

Accuracy:

- Base sMAPE: 11.30%
- Score sMAPE: 4.79%
- Improvement: +56.26%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -423,052,000.00
Base Forecast: -431,057,853.00 (Error: -8,005,853.00)
Score Forecast: -460,102,532.00 (Error: -37,050,532.00)

Accuracy:

- Base sMAPE: 1.87%
- Score sMAPE: 8.39%
- Improvement: -362.79%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 894,661,000.00
Base Forecast: 976,967,145.00 (Error: +82,306,145.00)
Score Forecast: 911,139,098.00 (Error: +16,478,098.00)

Accuracy:

- Base sMAPE: 8.80%
- Score sMAPE: 1.83%
- Improvement: +79.98%

Direction Correct: Base=True, Score=True

Winner: Score

```
=====
=====
RECOMMENDATIONS
=====
=====
```

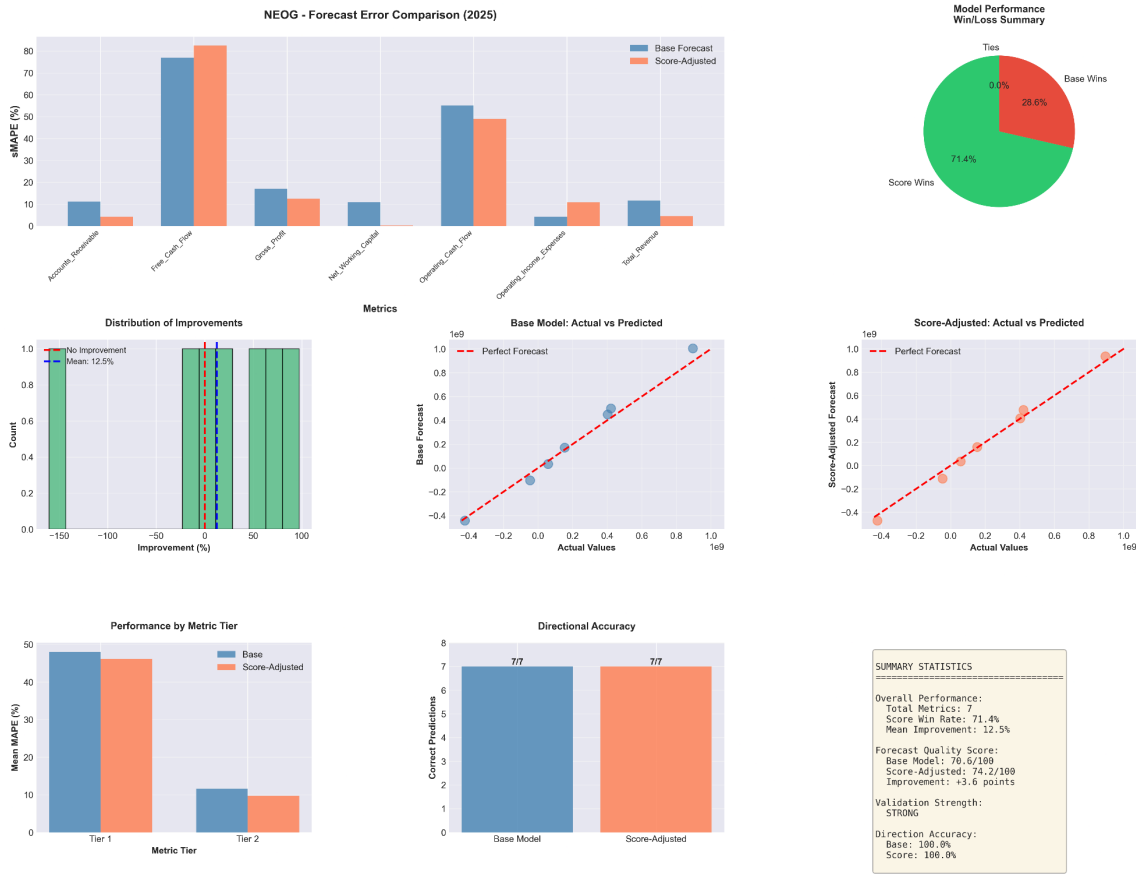
✓ Strong validation achieved. Qualitative score adjustments significantly improve forecast accuracy. Recommend:

1. Continue using score-adjusted methodology for future forecasts
2. Analyze which qualitative factors drove the best improvements
3. Consider increasing weight given to qualitative adjustments
4. Document the successful factors for replication

```
=====
=====
END OF REPORT
=====
=====
```


NEOG (LINEAR + WINTERS-HOLT + ARIMA):

Forecast Validation Analysis: NEOG (2025)



FORECAST VALIDATION REPORT

Company: NEOG
Forecast Year: 2025
Analysis Date: 2025-11-21 14:25:35

EXECUTIVE SUMMARY

Validation Strength: STRONG

Score-Adjusted method shows strong validation. Qualitative factors significantly improve forecast accuracy.

Overall Win Rate: 71.4% (Score-Adjusted wins)

Mean Improvement: 12.5%

Forecast Quality Scores:

- Base Model: 70.6/100
- Score-Adjusted Model: 74.2/100
- Improvement: +3.6 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 32.44%
- Median MAPE: 12.28%
- Mean sMAPE: 26.73%

Score-Adjusted Model:

- Mean MAPE: 30.54%
- Median MAPE: 11.42%
- Mean sMAPE: 23.41%

Win/Loss Record:

- Score-Adjusted Wins: 5
- Base Model Wins: 2
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 48.05%
- Score Mean MAPE: 46.18%
- Mean Improvement: 38.93%
- Score Win Rate: 75.0%
- Wins: Score=3, Base=1

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 11.63%
- Score Mean MAPE: 9.68%
- Mean Improvement: -22.71%
- Score Win Rate: 66.7%
- Wins: Score=2, Base=1

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 1.3248
- P-Value: 0.2334
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 6.0000
- P-Value: 0.2188
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.5007
- Interpretation: Medium effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 153,384,000.00

Base Forecast: 171,519,037.00 (Error: +18,135,037.00)

Score Forecast: 159,962,085.00 (Error: +6,578,085.00)

Accuracy:

- Base sMAPE: 11.16%
- Score sMAPE: 4.20%
- Improvement: +63.73%

Direction Correct: Base=True, Score=True

Winner: Score

Free_Cash_Flow (Tier 1):

Actual Value: -46,351,000.00

Base Forecast: -104,363,064.00 (Error: -58,012,064.00)

Score Forecast: -111,395,047.00 (Error: -65,044,047.00)

Accuracy:

- Base sMAPE: 76.98%
- Score sMAPE: 82.47%
- Improvement: -12.12%

Direction Correct: Base=True, Score=True

Winner: Base

Gross_Profit (Tier 2):

Actual Value: 421,376,000.00

Base Forecast: 500,054,424.00 (Error: +78,678,424.00)

Score Forecast: 477,581,979.00 (Error: +56,205,979.00)

Accuracy:

- Base sMAPE: 17.08%
- Score sMAPE: 12.50%
- Improvement: +28.56%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 402,926,000.00
Base Forecast: 449,610,466.00 (Error: +46,684,466.00)
Score Forecast: 404,154,848.00 (Error: +1,228,848.00)

Accuracy:

- Base sMAPE: 10.95%
- Score sMAPE: 0.30%
- Improvement: +97.37%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 58,244,000.00
Base Forecast: 33,087,730.00 (Error: -25,156,270.00)
Score Forecast: 35,317,182.00 (Error: -22,926,818.00)

Accuracy:

- Base sMAPE: 55.09%
- Score sMAPE: 49.01%
- Improvement: +8.86%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -423,052,000.00
Base Forecast: -441,598,770.00 (Error: -18,546,770.00)
Score Forecast: -471,353,695.00 (Error: -48,301,695.00)

Accuracy:

- Base sMAPE: 4.29%
- Score sMAPE: 10.80%
- Improvement: -160.43%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 894,661,000.00
Base Forecast: 1,004,480,707.00 (Error: +109,819,707.00)
Score Forecast: 936,798,797.00 (Error: +42,137,797.00)

Accuracy:

- Base sMAPE: 11.57%
- Score sMAPE: 4.60%
- Improvement: +61.63%

Direction Correct: Base=True, Score=True

Winner: Score

=====

=====

RECOMMENDATIONS

=====

=====

✓ Strong validation achieved. Qualitative score adjustments significantly improve forecast accuracy. Recommend:

1. Continue using score-adjusted methodology for future forecasts
2. Analyze which qualitative factors drove the best improvements
3. Consider increasing weight given to qualitative adjustments
4. Document the successful factors for replication

=====

=====

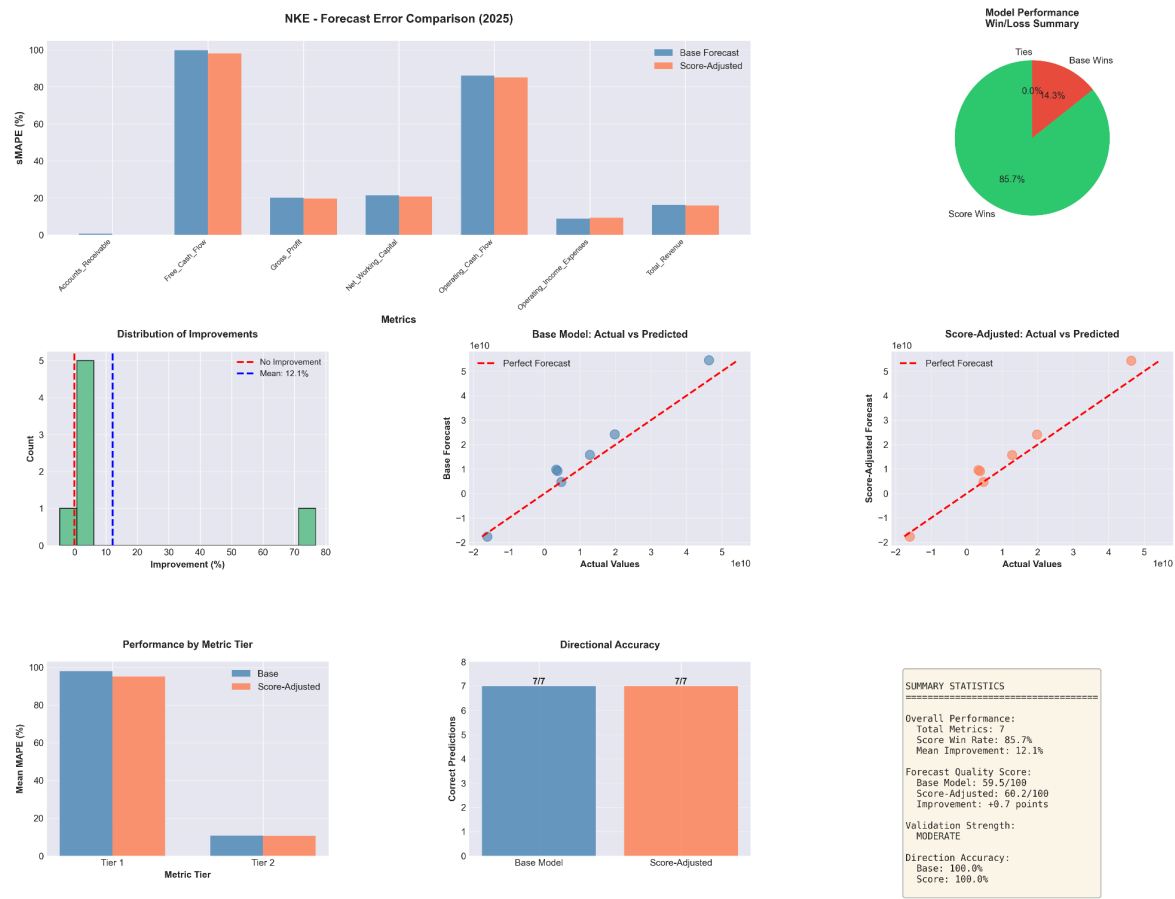
END OF REPORT

=====

=====

NKE (NAIVE):

Forecast Validation Analysis: NKE (2025)



FORECAST VALIDATION REPORT

Company: NKE
Forecast Year: 2025
Analysis Date: 2025-12-11 23:10:48

EXECUTIVE SUMMARY

Validation Strength: MODERATE

Score-Adjusted method shows moderate validation. Qualitative factors provide meaningful improvements.

Overall Win Rate: 85.7% (Score-Adjusted wins)

Mean Improvement: 12.1%

Forecast Quality Scores:

- Base Model: 59.5/100
- Score-Adjusted Model: 60.2/100
- Improvement: +0.7 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 60.58%
- Median MAPE: 22.29%
- Mean sMAPE: 36.14%

Score-Adjusted Model:

- Mean MAPE: 58.89%
- Median MAPE: 21.82%
- Mean sMAPE: 35.54%

Win/Loss Record:

- Score-Adjusted Wins: 6
- Base Model Wins: 1
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 97.98%
- Score Mean MAPE: 95.15%
- Mean Improvement: 2.71%
- Score Win Rate: 100.0%
- Wins: Score=4, Base=0

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 10.72%
- Score Mean MAPE: 10.55%
- Mean Improvement: 24.72%
- Score Win Rate: 66.7%
- Wins: Score=2, Base=1

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 2.3657
- P-Value: 0.0558
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 2.0000
- P-Value: 0.0469
- Significant at $\alpha=0.05$: True

Effect Size:

- Cohen's d: 0.8941
- Interpretation: Large effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 4,717,000,000.00

Base Forecast: 4,746,087,219.00 (Error: +29,087,219.00)

Score Forecast: 4,723,751,113.00 (Error: +6,751,113.00)

Accuracy:

- Base sMAPE: 0.61%
- Score sMAPE: 0.14%
- Improvement: +76.79%

Direction Correct: Base=True, Score=True

Winner: Score

Free_Cash_Flow (Tier 1):

Actual Value: 3,268,000,000.00

Base Forecast: 9,776,931,344.00 (Error: +6,508,931,344.00)

Score Forecast: 9,555,736,150.00 (Error: +6,287,736,150.00)

Accuracy:

- Base sMAPE: 99.79%
- Score sMAPE: 98.06%
- Improvement: +3.40%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 19,790,000,000.00

Base Forecast: 24,200,309,114.00 (Error: +4,410,309,114.00)

Score Forecast: 24,108,377,476.00 (Error: +4,318,377,476.00)

Accuracy:

- Base sMAPE: 20.05%
- Score sMAPE: 19.67%
- Improvement: +2.08%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 12,796,000,000.00

Base Forecast: 15,841,365,051.00 (Error: +3,045,365,051.00)

Score Forecast: 15,767,699,497.00 (Error: +2,971,699,497.00)

Accuracy:

- Base sMAPE: 21.27%
- Score sMAPE: 20.81%
- Improvement: +2.42%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 3,698,000,000.00

Base Forecast: 9,291,088,622.00 (Error: +5,593,088,622.00)

Score Forecast: 9,160,742,419.00 (Error: +5,462,742,419.00)

Accuracy:

- Base sMAPE: 86.12%
- Score sMAPE: 84.97%
- Improvement: +2.33%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -16,088,000,000.00

Base Forecast: -17,576,502,570.00 (Error: -1,488,502,570.00)

Score Forecast: -17,646,537,750.00 (Error: -1,558,537,750.00)

Accuracy:

- Base sMAPE: 8.84%
- Score sMAPE: 9.24%
- Improvement: -4.71%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 46,309,000,000.00

Base Forecast: 54,510,010,354.00 (Error: +8,201,010,354.00)

Score Forecast: 54,289,649,629.00 (Error: +7,980,649,629.00)

Accuracy:

- Base sMAPE: 16.27%
- Score sMAPE: 15.87%
- Improvement: +2.69%

Direction Correct: Base=True, Score=True

Winner: Score

=====

=====

RECOMMENDATIONS

=====

=====

- Moderate validation achieved. Qualitative adjustments provide meaningful but not dramatic improvements. Recommend:
 1. Continue using score-adjusted approach with refinements
 2. Investigate which specific qualitative factors are most predictive
 3. Consider sector-specific adjustments to scoring methodology
 4. Test on additional companies to confirm consistency

=====

=====

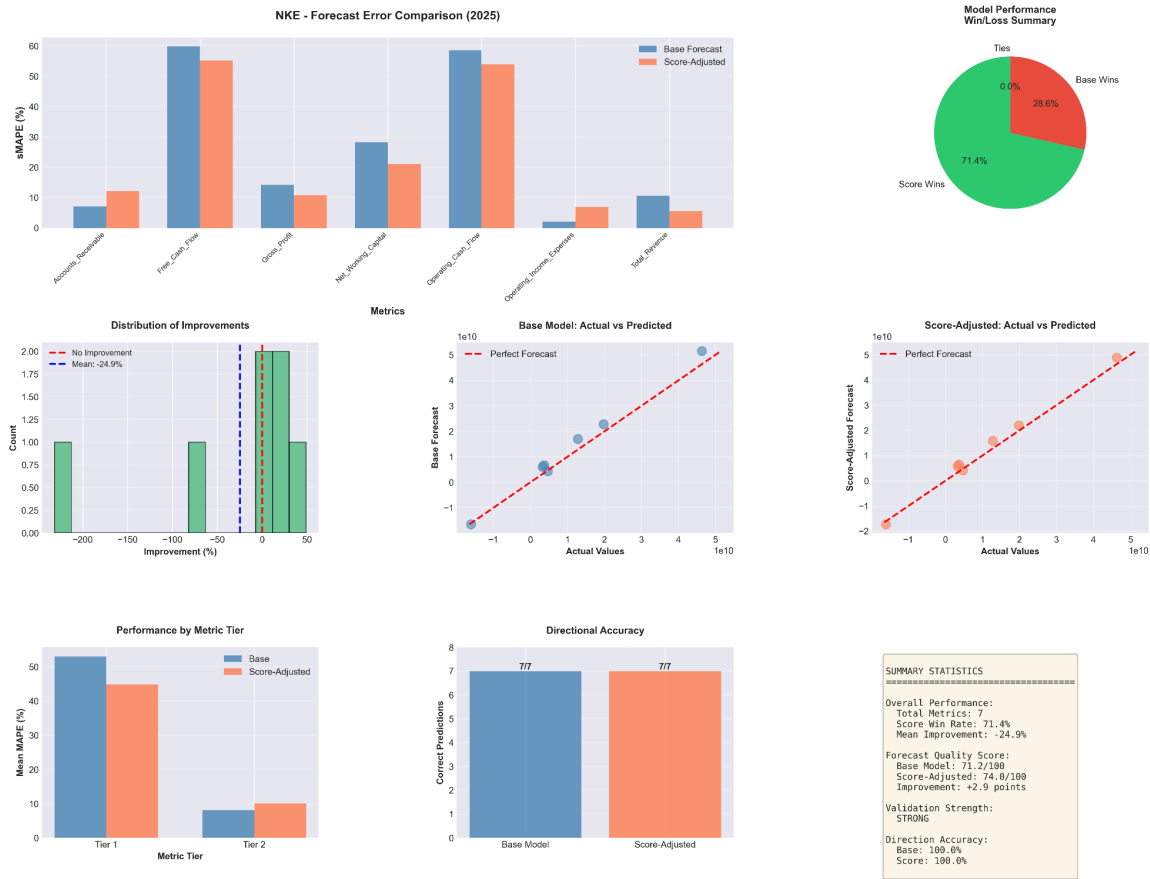
END OF REPORT

=====

=====

NKE (LINEAR):

Forecast Validation Analysis: NKE (2025)



FORECAST VALIDATION REPORT

Company: NKE
Forecast Year: 2025
Analysis Date: 2025-12-11 23:13:33

EXECUTIVE SUMMARY

Validation Strength: STRONG

Score-Adjusted method shows strong validation. Qualitative factors significantly improve forecast accuracy.

Overall Win Rate: 71.4% (Score-Adjusted wins)

Mean Improvement: -24.9%

Forecast Quality Scores:

- Base Model: 71.2/100
- Score-Adjusted Model: 74.0/100
- Improvement: +2.9 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 33.81%
- Median MAPE: 15.30%
- Mean sMAPE: 25.82%

Score-Adjusted Model:

- Mean MAPE: 29.93%
- Median MAPE: 11.51%
- Mean sMAPE: 23.68%

Win/Loss Record:

- Score-Adjusted Wins: 5
- Base Model Wins: 2
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 53.06%
- Score Mean MAPE: 44.81%
- Mean Improvement: 24.73%
- Score Win Rate: 100.0%
- Wins: Score=4, Base=0

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 8.13%
- Score Mean MAPE: 10.08%
- Mean Improvement: -91.08%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

STATISTICAL SIGNIFICANCE TESTS

Paired T-Test:

- T-Statistic: 1.1405
- P-Value: 0.2975
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 9.0000
- P-Value: 0.4688
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.4311
- Interpretation: Small effect

METRIC-BY-METRIC BREAKDOWN

Accounts_Receivable (Tier 2):

Actual Value: 4,717,000,000.00

Base Forecast: 4,391,243,055.00 (Error: -325,756,945.00)

Score Forecast: 4,174,262,953.00 (Error: -542,737,047.00)

Accuracy:

- Base sMAPE: 7.15%
- Score sMAPE: 12.21%
- Improvement: -66.61%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 3,268,000,000.00

Base Forecast: 6,055,431,944.00 (Error: +2,787,431,944.00)

Score Forecast: 5,756,220,941.00 (Error: +2,488,220,941.00)

Accuracy:

- Base sMAPE: 59.79%
- Score sMAPE: 55.15%
- Improvement: +10.73%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 19,790,000,000.00

Base Forecast: 22,818,736,111.00 (Error: +3,028,736,111.00)

Score Forecast: 22,066,721,843.00 (Error: +2,276,721,843.00)

Accuracy:

- Base sMAPE: 14.22%
- Score sMAPE: 10.88%
- Improvement: +24.83%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 12,796,000,000.00

Base Forecast: 17,003,918,055.00 (Error: +4,207,918,055.00)

Score Forecast: 15,813,643,791.00 (Error: +3,017,643,791.00)

Accuracy:

- Base sMAPE: 28.24%
- Score sMAPE: 21.10%
- Improvement: +28.29%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 3,698,000,000.00

Base Forecast: 6,761,915,277.00 (Error: +3,063,915,277.00)

Score Forecast: 6,427,795,520.00 (Error: +2,729,795,520.00)

Accuracy:

- Base sMAPE: 58.58%
- Score sMAPE: 53.92%
- Improvement: +10.90%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -16,088,000,000.00

Base Forecast: -16,438,931,944.00 (Error: -350,931,944.00)

Score Forecast: -17,251,212,449.00 (Error: -1,163,212,449.00)

Accuracy:

- Base sMAPE: 2.16%
- Score sMAPE: 6.98%
- Improvement: -231.46%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 46,309,000,000.00

Base Forecast: 51,502,033,333.00 (Error: +5,193,033,333.00)

Score Forecast: 48,957,214,862.00 (Error: +2,648,214,862.00)

Accuracy:

- Base sMAPE: 10.62%
- Score sMAPE: 5.56%
- Improvement: +49.00%

Direction Correct: Base=True, Score=True
Winner: Score

=====

=====

RECOMMENDATIONS

=====

=====

- ✓ Strong validation achieved. Qualitative score adjustments significantly improve forecast accuracy. Recommend:
1. Continue using score-adjusted methodology for future forecasts
 2. Analyze which qualitative factors drove the best improvements
 3. Consider increasing weight given to qualitative adjustments
 4. Document the successful factors for replication

=====

=====

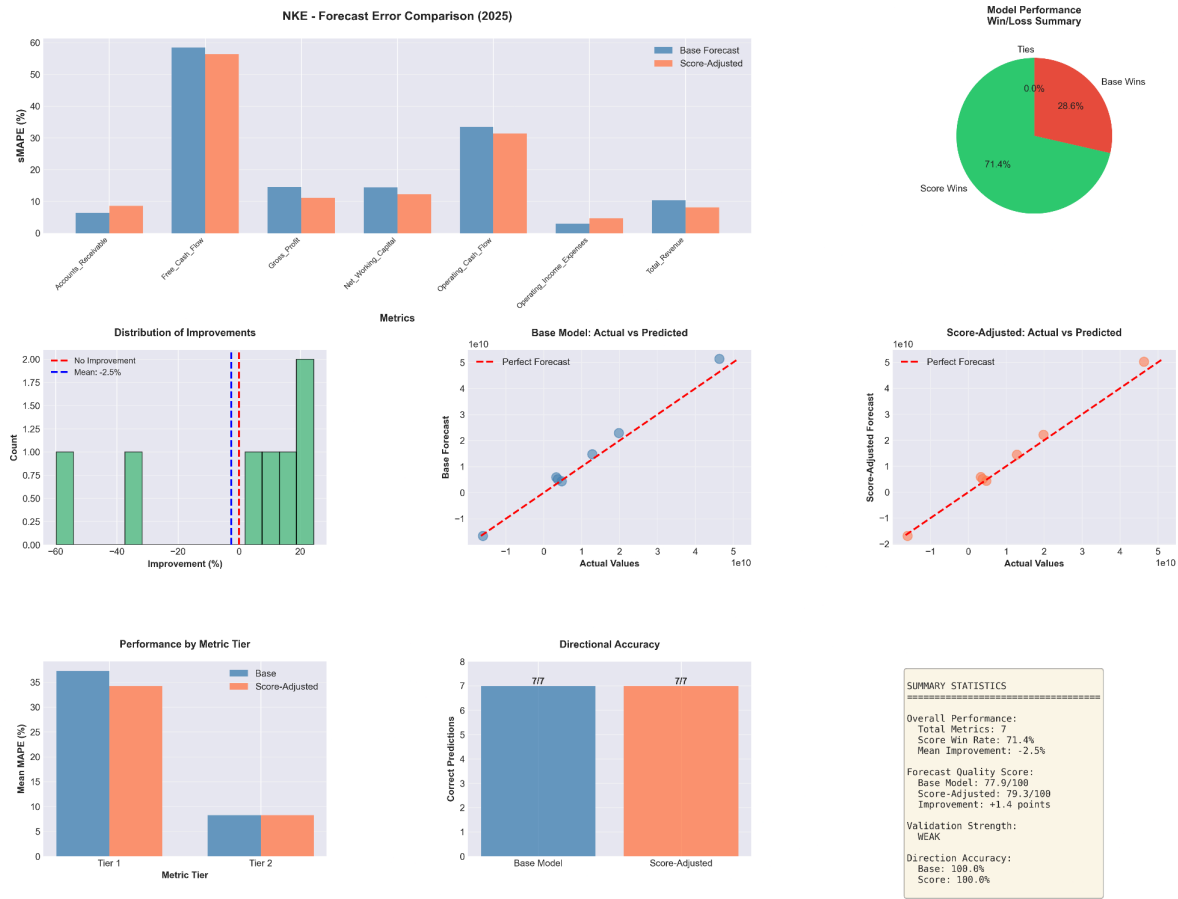
END OF REPORT

=====

=====

NKE (XGBOOST):

Forecast Validation Analysis: NKE (2025)



FORECAST VALIDATION REPORT

Company: NKE
Forecast Year: 2025
Analysis Date: 2025-12-11 23:07:47

EXECUTIVE SUMMARY

=====

Validation Strength: WEAK

Score-Adjusted method shows weak validation. Qualitative factors provide marginal improvements.

Overall Win Rate: 71.4% (Score-Adjusted wins)

Mean Improvement: -2.5%

Forecast Quality Scores:

- Base Model: 77.9/100
- Score-Adjusted Model: 79.3/100
- Improvement: +1.4 points

=====

DETAILED METRICS COMPARISON

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 24.86%
- Median MAPE: 15.58%
- Mean sMAPE: 20.08%

Score-Adjusted Model:

- Mean MAPE: 23.15%
- Median MAPE: 11.83%
- Mean sMAPE: 18.93%

Win/Loss Record:

- Score-Adjusted Wins: 5
- Base Model Wins: 2
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 37.30%
- Score Mean MAPE: 34.28%
- Mean Improvement: 12.80%
- Score Win Rate: 100.0%
- Wins: Score=4, Base=0

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 8.28%
- Score Mean MAPE: 8.30%
- Mean Improvement: -22.99%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 1.3869
- P-Value: 0.2148
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 7.0000
- P-Value: 0.2969
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.5242
- Interpretation: Medium effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 4,717,000,000.00

Base Forecast: 4,426,999,296.00 (Error: -290,000,704.00)

Score Forecast: 4,329,605,311.00 (Error: -387,394,689.00)

Accuracy:

- Base sMAPE: 6.34%
- Score sMAPE: 8.56%
- Improvement: -33.58%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 3,268,000,000.00

Base Forecast: 5,961,999,360.00 (Error: +2,693,999,360.00)

Score Forecast: 5,830,835,374.00 (Error: +2,562,835,374.00)

Accuracy:

- Base sMAPE: 58.37%
- Score sMAPE: 56.33%
- Improvement: +4.87%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 19,790,000,000.00

Base Forecast: 22,886,993,920.00 (Error: +3,096,993,920.00)

Score Forecast: 22,131,723,120.00 (Error: +2,341,723,120.00)

Accuracy:

- Base sMAPE: 14.51%
- Score sMAPE: 11.17%
- Improvement: +24.39%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 12,796,000,000.00

Base Forecast: 14,788,999,168.00 (Error: +1,992,999,168.00)

Score Forecast: 14,463,641,186.00 (Error: +1,667,641,186.00)

Accuracy:

- Base sMAPE: 14.45%
- Score sMAPE: 12.24%
- Improvement: +16.33%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 3,698,000,000.00

Base Forecast: 5,188,000,256.00 (Error: +1,490,000,256.00)

Score Forecast: 5,073,864,250.00 (Error: +1,375,864,250.00)

Accuracy:

- Base sMAPE: 33.54%
- Score sMAPE: 31.37%
- Improvement: +7.66%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -16,088,000,000.00

Base Forecast: -16,575,996,928.00 (Error: -487,996,928.00)

Score Forecast: -16,867,734,473.00 (Error: -779,734,473.00)

Accuracy:

- Base sMAPE: 2.99%
- Score sMAPE: 4.73%
- Improvement: -59.78%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 46,309,000,000.00

Base Forecast: 51,361,988,608.00 (Error: +5,052,988,608.00)

Score Forecast: 50,232,024,858.00 (Error: +3,923,024,858.00)

Accuracy:

- Base sMAPE: 10.35%
- Score sMAPE: 8.13%
- Improvement: +22.36%

Direction Correct: Base=True, Score=True

Winner: Score

```
=====
=====
RECOMMENDATIONS
=====
=====
```

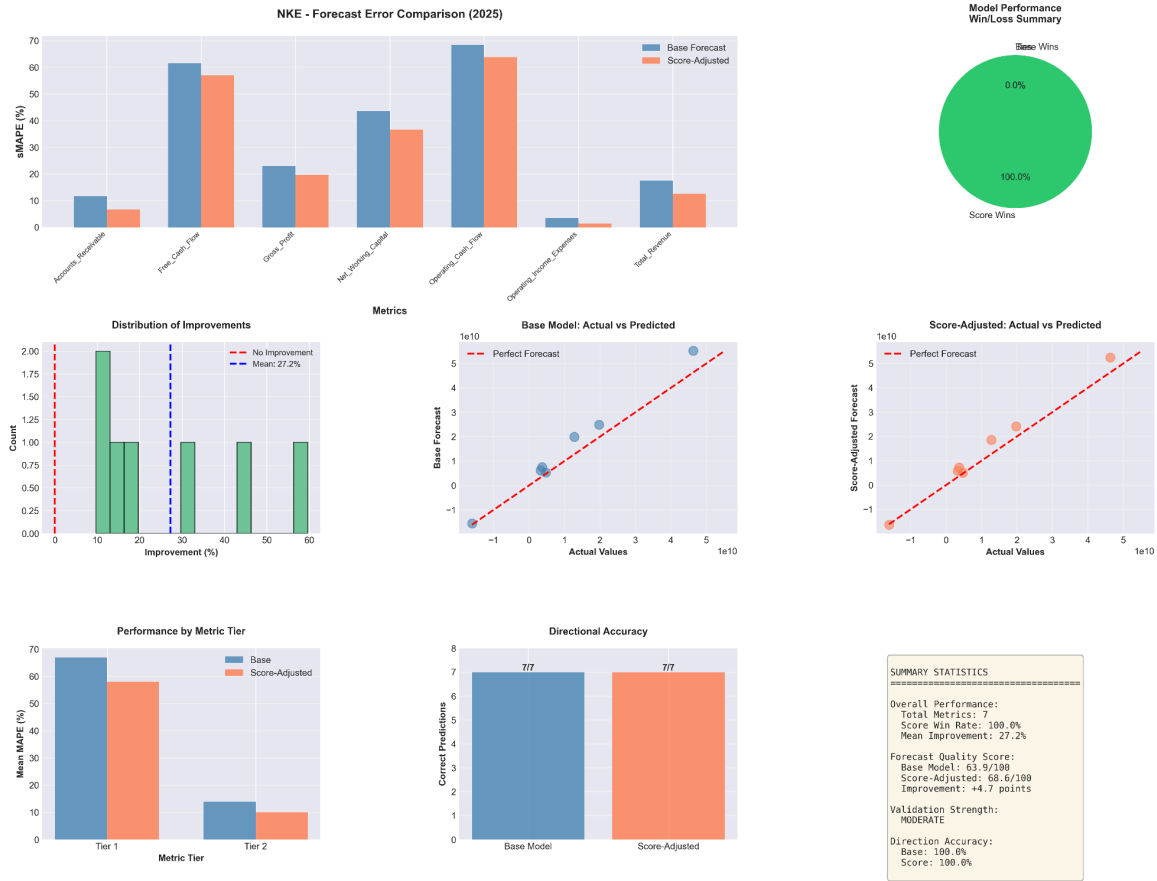
△ Weak validation. Qualitative adjustments show marginal impact. Recommend:

1. Review and refine the qualitative scoring methodology
2. Investigate if certain metric types benefit more from adjustments
3. Consider whether the base model is already capturing qualitative factors
4. Test alternative adjustment algorithms

```
=====
=====
END OF REPORT
=====
=====
```


NKE (ARIMA):

Forecast Validation Analysis: NKE (2025)



FORECAST VALIDATION REPORT

Company: NKE
Forecast Year: 2025
Analysis Date: 2025-12-11 23:38:46

EXECUTIVE SUMMARY

Validation Strength: MODERATE

Score-Adjusted method shows moderate validation. Qualitative factors provide meaningful improvements.

Overall Win Rate: 100.0% (Score-Adjusted wins)

Mean Improvement: 27.2%

Forecast Quality Scores:

- Base Model: 63.9/100
- Score-Adjusted Model: 68.6/100
- Improvement: +4.7 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 44.25%
- Median MAPE: 25.95%
- Mean sMAPE: 32.76%

Score-Adjusted Model:

- Mean MAPE: 37.40%
- Median MAPE: 21.80%
- Mean sMAPE: 28.24%

Win/Loss Record:

- Score-Adjusted Wins: 7
- Base Model Wins: 0
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 66.98%
- Score Mean MAPE: 57.92%
- Mean Improvement: 17.59%
- Score Win Rate: 100.0%
- Wins: Score=4, Base=0

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 13.94%
- Score Mean MAPE: 10.03%
- Mean Improvement: 40.05%
- Score Win Rate: 100.0%
- Wins: Score=3, Base=0

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 7.8542
- P-Value: 0.0002
- Significant at $\alpha=0.05$: True
- Interpretation: Score-Adjusted significantly different

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 0.0000
- P-Value: 0.0156
- Significant at $\alpha=0.05$: True

Effect Size:

- Cohen's d: 2.9686
- Interpretation: Large effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 4,717,000,000.00

Base Forecast: 5,305,688,143.00 (Error: +588,688,143.00)

Score Forecast: 5,043,523,481.00 (Error: +326,523,481.00)

Accuracy:

- Base sMAPE: 11.75%
- Score sMAPE: 6.69%
- Improvement: +44.53%

Direction Correct: Base=True, Score=True

Winner: Score

Free_Cash_Flow (Tier 1):

Actual Value: 3,268,000,000.00

Base Forecast: 6,178,272,325.00 (Error: +2,910,272,325.00)

Score Forecast: 5,872,991,533.00 (Error: +2,604,991,533.00)

Accuracy:

- Base sMAPE: 61.62%
- Score sMAPE: 57.00%
- Improvement: +10.49%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 19,790,000,000.00

Base Forecast: 24,926,052,682.00 (Error: +5,136,052,682.00)

Score Forecast: 24,104,589,690.00 (Error: +4,314,589,690.00)

Accuracy:

- Base sMAPE: 22.97%
- Score sMAPE: 19.66%
- Improvement: +15.99%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 12,796,000,000.00

Base Forecast: 19,927,758,263.00 (Error: +7,131,758,263.00)

Score Forecast: 18,532,815,185.00 (Error: +5,736,815,185.00)

Accuracy:

- Base sMAPE: 43.59%
- Score sMAPE: 36.62%
- Improvement: +19.56%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 3,698,000,000.00

Base Forecast: 7,538,517,616.00 (Error: +3,840,517,616.00)

Score Forecast: 7,166,024,383.00 (Error: +3,468,024,383.00)

Accuracy:

- Base sMAPE: 68.36%
- Score sMAPE: 63.84%
- Improvement: +9.70%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -16,088,000,000.00

Base Forecast: -15,540,970,996.00 (Error: +547,029,004.00)

Score Forecast: -16,308,881,455.00 (Error: -220,881,455.00)

Accuracy:

- Base sMAPE: 3.46%
- Score sMAPE: 1.36%
- Improvement: +59.62%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 46,309,000,000.00

Base Forecast: 55,228,254,814.00 (Error: +8,919,254,814.00)

Score Forecast: 52,499,316,287.00 (Error: +6,190,316,287.00)

Accuracy:

- Base sMAPE: 17.57%
- Score sMAPE: 12.53%
- Improvement: +30.60%

Direction Correct: Base=True, Score=True

Winner: Score

=====

=====

RECOMMENDATIONS

=====

=====

- Moderate validation achieved. Qualitative adjustments provide meaningful but not dramatic improvements. Recommend:
 1. Continue using score-adjusted approach with refinements
 2. Investigate which specific qualitative factors are most predictive
 3. Consider sector-specific adjustments to scoring methodology
 4. Test on additional companies to confirm consistency

=====

=====

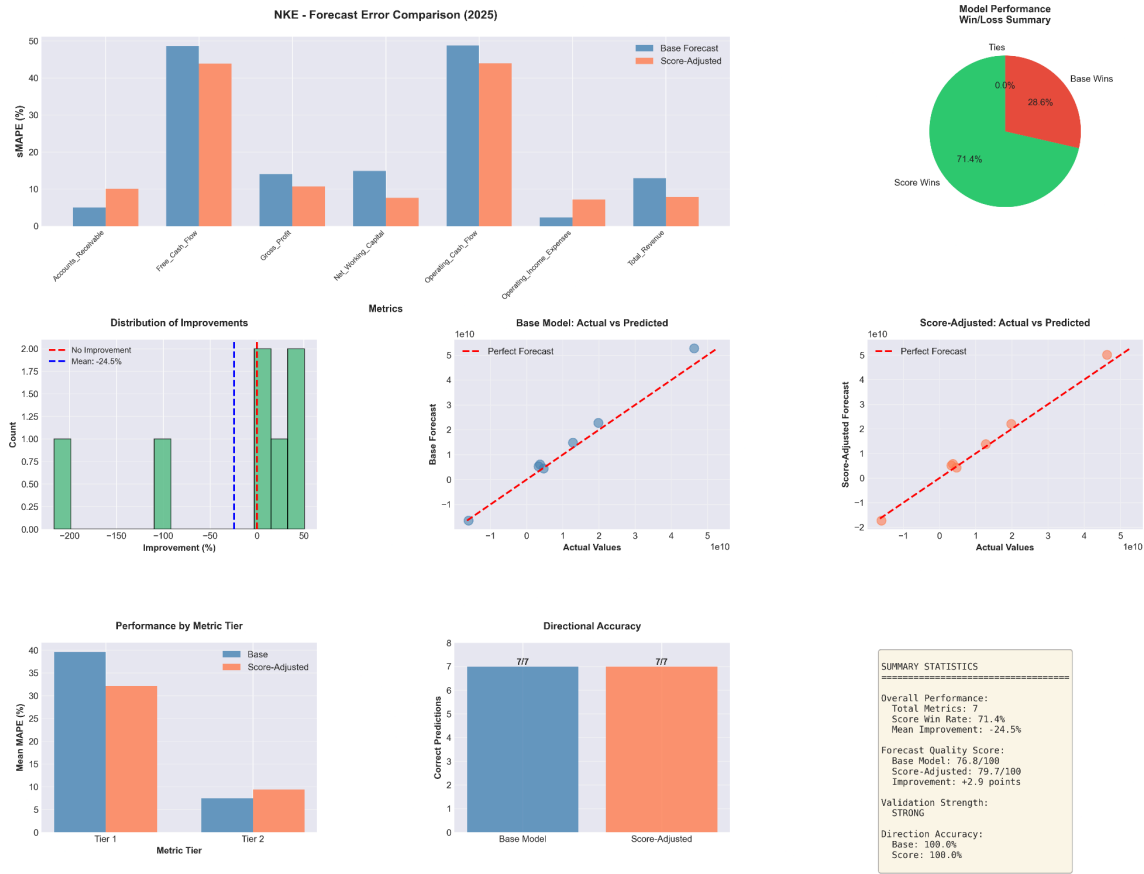
END OF REPORT

=====

=====

NKE (WINTERS-HOLT):

Forecast Validation Analysis: NKE (2025)



FORECAST VALIDATION REPORT

Company: NKE
Forecast Year: 2025
Analysis Date: 2025-12-11 23:41:46

EXECUTIVE SUMMARY

Validation Strength: STRONG

Score-Adjusted method shows strong validation. Qualitative factors significantly improve forecast accuracy.

Overall Win Rate: 71.4% (Score-Adjusted wins)

Mean Improvement: -24.5%

Forecast Quality Scores:

- Base Model: 76.8/100
- Score-Adjusted Model: 79.7/100
- Improvement: +2.9 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 25.84%
- Median MAPE: 15.12%
- Mean sMAPE: 20.93%

Score-Adjusted Model:

- Mean MAPE: 22.41%
- Median MAPE: 9.58%
- Mean sMAPE: 18.74%

Win/Loss Record:

- Score-Adjusted Wins: 5
- Base Model Wins: 2
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 39.63%
- Score Mean MAPE: 32.14%
- Mean Improvement: 29.12%
- Score Win Rate: 100.0%
- Wins: Score=4, Base=0

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 7.45%
- Score Mean MAPE: 9.43%
- Mean Improvement: -95.90%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 1.1584
- P-Value: 0.2907
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 10.0000
- P-Value: 0.5781
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.4378
- Interpretation: Small effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 4,717,000,000.00

Base Forecast: 4,486,786,129.00 (Error: -230,213,871.00)

Score Forecast: 4,265,085,053.00 (Error: -451,914,947.00)

Accuracy:

- Base sMAPE: 5.00%
- Score sMAPE: 10.06%
- Improvement: -96.30%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 3,268,000,000.00

Base Forecast: 5,367,049,312.00 (Error: +2,099,049,312.00)

Score Forecast: 5,101,852,672.00 (Error: +1,833,852,672.00)

Accuracy:

- Base sMAPE: 48.62%
- Score sMAPE: 43.82%
- Improvement: +12.63%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 19,790,000,000.00

Base Forecast: 22,782,455,151.00 (Error: +2,992,455,151.00)

Score Forecast: 22,031,636,559.00 (Error: +2,241,636,559.00)

Accuracy:

- Base sMAPE: 14.06%
- Score sMAPE: 10.72%
- Improvement: +25.09%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 12,796,000,000.00

Base Forecast: 14,849,274,104.00 (Error: +2,053,274,104.00)

Score Forecast: 13,809,824,916.00 (Error: +1,013,824,916.00)

Accuracy:

- Base sMAPE: 14.85%
- Score sMAPE: 7.62%
- Improvement: +50.62%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 3,698,000,000.00

Base Forecast: 6,079,735,461.00 (Error: +2,381,735,461.00)

Score Forecast: 5,779,323,573.00 (Error: +2,081,323,573.00)

Accuracy:

- Base sMAPE: 48.72%
- Score sMAPE: 43.92%
- Improvement: +12.61%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -16,088,000,000.00

Base Forecast: -16,463,753,225.00 (Error: -375,753,225.00)

Score Forecast: -17,277,260,199.00 (Error: -1,189,260,199.00)

Accuracy:

- Base sMAPE: 2.31%
- Score sMAPE: 7.13%
- Improvement: -216.50%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 46,309,000,000.00

Base Forecast: 52,725,289,770.00 (Error: +6,416,289,770.00)

Score Forecast: 50,120,027,752.00 (Error: +3,811,027,752.00)

Accuracy:

- Base sMAPE: 12.96%
- Score sMAPE: 7.90%
- Improvement: +40.60%

Direction Correct: Base=True, Score=True

Winner: Score

```
=====
=====
RECOMMENDATIONS
=====
=====
```

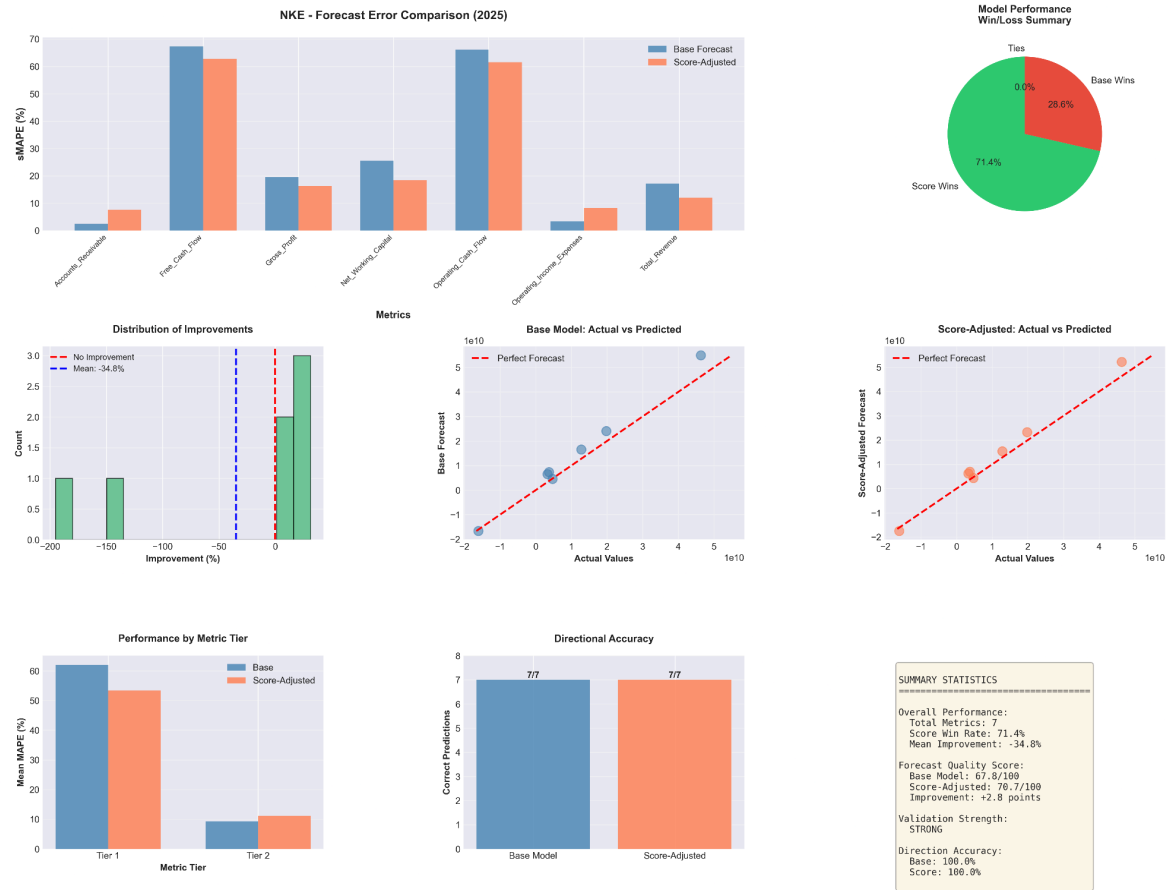
✓ Strong validation achieved. Qualitative score adjustments significantly improve forecast accuracy. Recommend:

1. Continue using score-adjusted methodology for future forecasts
2. Analyze which qualitative factors drove the best improvements
3. Consider increasing weight given to qualitative adjustments
4. Document the successful factors for replication

```
=====
=====
END OF REPORT
=====
=====
```


NKE (LINEAR + WINTERS-HOLT):

Forecast Validation Analysis: NKE (2025)



FORECAST VALIDATION REPORT

Company: NKE
Forecast Year: 2025
Analysis Date: 2025-12-11 23:44:06

EXECUTIVE SUMMARY

=====

Validation Strength: **STRONG**

Score-Adjusted method shows strong validation. Qualitative factors significantly improve forecast accuracy.

Overall Win Rate: 71.4% (Score-Adjusted wins)

Mean Improvement: -34.8%

Forecast Quality Scores:

- Base Model: 67.8/100
- Score-Adjusted Model: 70.7/100
- Improvement: +2.8 points

=====

DETAILED METRICS COMPARISON

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 39.39%
- Median MAPE: 21.76%
- Mean sMAPE: 28.78%

Score-Adjusted Model:

- Mean MAPE: 35.28%
- Median MAPE: 17.75%
- Mean sMAPE: 26.68%

Win/Loss Record:

- Score-Adjusted Wins: 5
- Base Model Wins: 2
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 62.01%
- Score Mean MAPE: 53.34%
- Mean Improvement: 20.53%
- Score Win Rate: 100.0%
- Wins: Score=4, Base=0

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 9.22%
- Score Mean MAPE: 11.19%
- Mean Improvement: -108.48%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 1.1252
- P-Value: 0.3035
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 10.0000
- P-Value: 0.5781
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.4253
- Interpretation: Small effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 4,717,000,000.00

Base Forecast: 4,600,498,997.00 (Error: -116,501,003.00)

Score Forecast: 4,373,179,140.00 (Error: -343,820,860.00)

Accuracy:

- Base sMAPE: 2.50%
- Score sMAPE: 7.56%
- Improvement: -195.12%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 3,268,000,000.00

Base Forecast: 6,580,811,621.00 (Error: +3,312,811,621.00)

Score Forecast: 6,255,640,557.00 (Error: +2,987,640,557.00)

Accuracy:

- Base sMAPE: 67.27%
- Score sMAPE: 62.74%
- Improvement: +9.82%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 19,790,000,000.00

Base Forecast: 24,096,427,114.00 (Error: +4,306,427,114.00)

Score Forecast: 23,302,305,262.00 (Error: +3,512,305,262.00)

Accuracy:

- Base sMAPE: 19.63%
- Score sMAPE: 16.30%
- Improvement: +18.44%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 12,796,000,000.00

Base Forecast: 16,538,033,352.00 (Error: +3,742,033,352.00)

Score Forecast: 15,380,371,017.00 (Error: +2,584,371,017.00)

Accuracy:

- Base sMAPE: 25.51%
- Score sMAPE: 18.34%
- Improvement: +30.94%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 3,698,000,000.00

Base Forecast: 7,350,417,767.00 (Error: +3,652,417,767.00)

Score Forecast: 6,987,218,924.00 (Error: +3,289,218,924.00)

Accuracy:

- Base sMAPE: 66.12%
- Score sMAPE: 61.57%
- Improvement: +9.94%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -16,088,000,000.00

Base Forecast: -16,640,721,966.00 (Error: -552,721,966.00)

Score Forecast: -17,462,973,320.00 (Error: -1,374,973,320.00)

Accuracy:

- Base sMAPE: 3.38%
- Score sMAPE: 8.20%
- Improvement: -148.76%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 46,309,000,000.00

Base Forecast: 54,949,127,855.00 (Error: +8,640,127,855.00)

Score Forecast: 52,233,981,549.00 (Error: +5,924,981,549.00)

Accuracy:

- Base sMAPE: 17.07%
- Score sMAPE: 12.03%
- Improvement: +31.42%

Direction Correct: Base=True, Score=True

Winner: Score

```
=====
=====
RECOMMENDATIONS
=====
=====
```

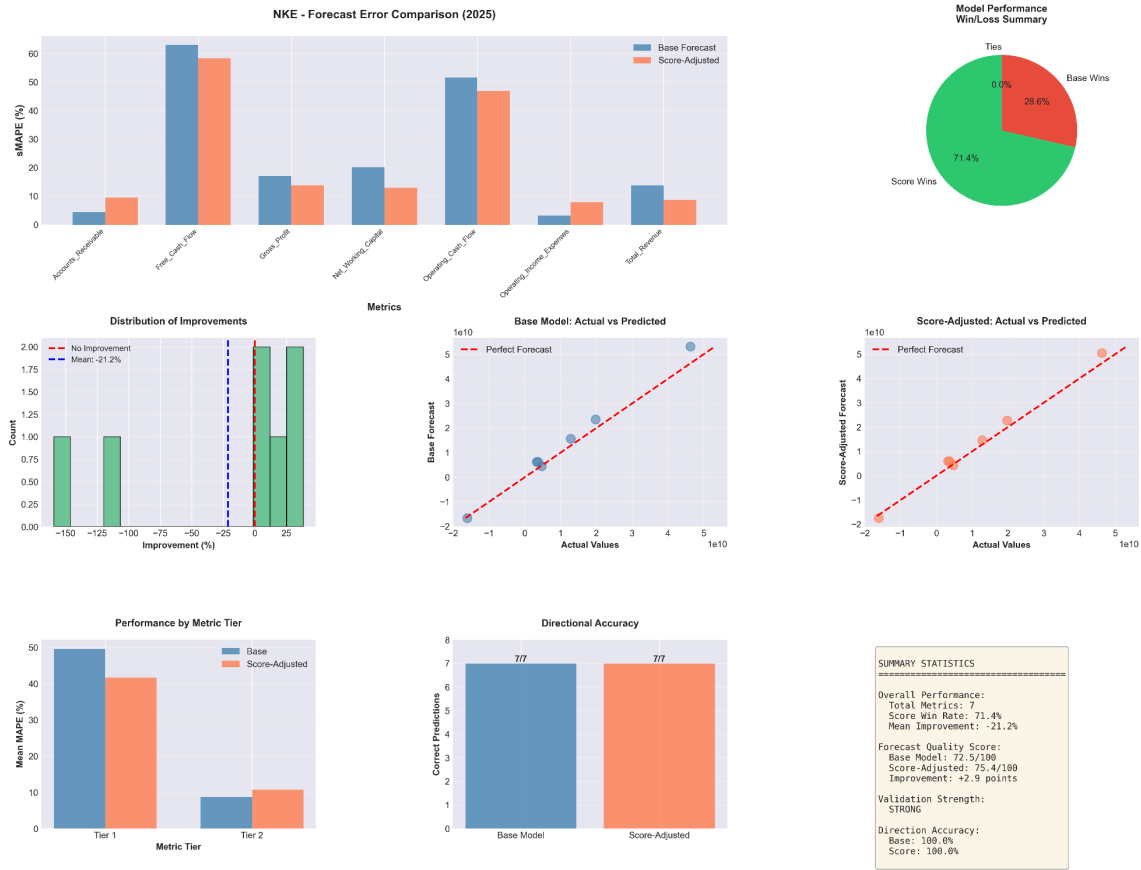
✓ Strong validation achieved. Qualitative score adjustments significantly improve forecast accuracy. Recommend:

1. Continue using score-adjusted methodology for future forecasts
2. Analyze which qualitative factors drove the best improvements
3. Consider increasing weight given to qualitative adjustments
4. Document the successful factors for replication

```
=====
=====
END OF REPORT
=====
=====
```


NKE (LINEAR + WINTERS-HOLT + XGBOOST):

Forecast Validation Analysis: NKE (2025)



FORECAST VALIDATION REPORT

Company: NKE
Forecast Year: 2025
Analysis Date: 2025-12-11 23:46:56

EXECUTIVE SUMMARY

Validation Strength: STRONG

Score-Adjusted method shows strong validation. Qualitative factors significantly improve forecast accuracy.

Overall Win Rate: 71.4% (Score-Adjusted wins)

Mean Improvement: -21.2%

Forecast Quality Scores:

- Base Model: 72.5/100
- Score-Adjusted Model: 75.4/100
- Improvement: +2.9 points

DETAILED METRICS COMPARISON

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 32.11%
- Median MAPE: 18.66%
- Mean sMAPE: 24.73%

Score-Adjusted Model:

- Mean MAPE: 28.37%
- Median MAPE: 13.84%
- Mean sMAPE: 22.57%

Win/Loss Record:

- Score-Adjusted Wins: 5
- Base Model Wins: 2
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 49.65%
- Score Mean MAPE: 41.62%
- Mean Improvement: 24.75%
- Score Win Rate: 100.0%
- Wins: Score=4, Base=0

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 8.73%
- Score Mean MAPE: 10.70%
- Mean Improvement: -82.58%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 1.1439
- P-Value: 0.2962
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 10.0000
- P-Value: 0.5781
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.4324
- Interpretation: Small effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 4,717,000,000.00

Base Forecast: 4,513,329,050.00 (Error: -203,670,950.00)

Score Forecast: 4,290,316,435.00 (Error: -426,683,565.00)

Accuracy:

- Base sMAPE: 4.41%
- Score sMAPE: 9.47%
- Improvement: -109.50%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 3,268,000,000.00

Base Forecast: 6,269,973,170.00 (Error: +3,001,973,170.00)

Score Forecast: 5,960,161,256.00 (Error: +2,692,161,256.00)

Accuracy:

- Base sMAPE: 62.95%
- Score sMAPE: 58.35%
- Improvement: +10.32%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 19,790,000,000.00

Base Forecast: 23,483,560,501.00 (Error: +3,693,560,501.00)

Score Forecast: 22,709,636,281.00 (Error: +2,919,636,281.00)

Accuracy:

- Base sMAPE: 17.07%
- Score sMAPE: 13.74%
- Improvement: +20.95%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 12,796,000,000.00

Base Forecast: 15,663,339,108.00 (Error: +2,867,339,108.00)

Score Forecast: 14,566,905,370.00 (Error: +1,770,905,370.00)

Accuracy:

- Base sMAPE: 20.15%
- Score sMAPE: 12.94%
- Improvement: +38.24%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 3,698,000,000.00

Base Forecast: 6,270,380,723.00 (Error: +2,572,380,723.00)

Score Forecast: 5,960,548,671.00 (Error: +2,262,548,671.00)

Accuracy:

- Base sMAPE: 51.61%
- Score sMAPE: 46.85%
- Improvement: +12.04%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -16,088,000,000.00

Base Forecast: -16,603,295,767.00 (Error: -515,295,767.00)

Score Forecast: -17,423,697,817.00 (Error: -1,335,697,817.00)

Accuracy:

- Base sMAPE: 3.15%
- Score sMAPE: 7.97%
- Improvement: -159.21%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 46,309,000,000.00

Base Forecast: 53,144,562,519.00 (Error: +6,835,562,519.00)

Score Forecast: 50,518,583,396.00 (Error: +4,209,583,396.00)

Accuracy:

- Base sMAPE: 13.75%
- Score sMAPE: 8.70%
- Improvement: +38.42%

Direction Correct: Base=True, Score=True

Winner: Score

=====

=====

=====

RECOMMENDATIONS

=====

=====

✓ Strong validation achieved. Qualitative score adjustments significantly improve forecast accuracy. Recommend:

1. Continue using score-adjusted methodology for future forecasts
2. Analyze which qualitative factors drove the best improvements
3. Consider increasing weight given to qualitative adjustments
4. Document the successful factors for replication

=====

=====

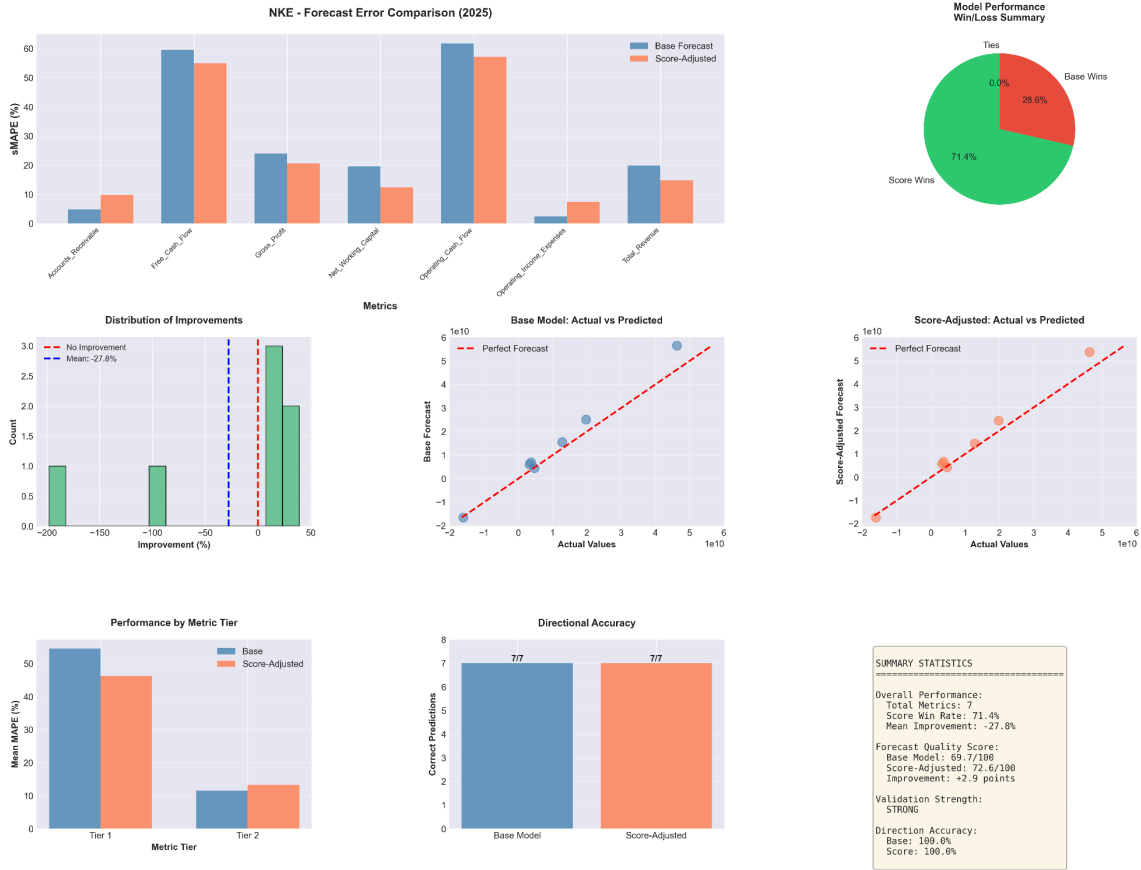
END OF REPORT

=====

=====

NKE (LINEAR + WINTERS-HOLT + ARIMA):

Forecast Validation Analysis: NKE (2025)



FORECAST VALIDATION REPORT

Company: NKE
Forecast Year: 2025
Analysis Date: 2025-12-12 00:02:11

EXECUTIVE SUMMARY

Validation Strength: STRONG

Score-Adjusted method shows strong validation. Qualitative factors significantly improve forecast accuracy.

Overall Win Rate: 71.4% (Score-Adjusted wins)

Mean Improvement: -27.8%

Forecast Quality Scores:

- Base Model: 69.7/100
- Score-Adjusted Model: 72.6/100
- Improvement: +2.9 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 36.08%
- Median MAPE: 22.10%
- Mean sMAPE: 27.46%

Score-Adjusted Model:

- Mean MAPE: 32.16%
- Median MAPE: 16.06%
- Mean sMAPE: 25.32%

Win/Loss Record:

- Score-Adjusted Wins: 5
- Base Model Wins: 2
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 54.53%
- Score Mean MAPE: 46.27%
- Mean Improvement: 21.92%
- Score Win Rate: 100.0%
- Wins: Score=4, Base=0

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 11.49%
- Score Mean MAPE: 13.35%
- Mean Improvement: -94.07%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 1.1363
- P-Value: 0.2992
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 10.0000
- P-Value: 0.5781
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.4295
- Interpretation: Small effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 4,717,000,000.00

Base Forecast: 4,494,152,648.00 (Error: -222,847,352.00)

Score Forecast: 4,272,087,577.00 (Error: -444,912,423.00)

Accuracy:

- Base sMAPE: 4.84%
- Score sMAPE: 9.90%
- Improvement: -99.65%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 3,268,000,000.00

Base Forecast: 6,044,253,933.00 (Error: +2,776,253,933.00)

Score Forecast: 5,745,595,258.00 (Error: +2,477,595,258.00)

Accuracy:

- Base sMAPE: 59.63%
- Score sMAPE: 54.97%
- Improvement: +10.76%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 19,790,000,000.00

Base Forecast: 25,167,188,653.00 (Error: +5,377,188,653.00)

Score Forecast: 24,337,778,784.00 (Error: +4,547,778,784.00)

Accuracy:

- Base sMAPE: 23.92%
- Score sMAPE: 20.61%
- Improvement: +15.42%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 12,796,000,000.00

Base Forecast: 15,582,459,929.00 (Error: +2,786,459,929.00)

Score Forecast: 14,491,687,734.00 (Error: +1,695,687,734.00)

Accuracy:

- Base sMAPE: 19.64%
- Score sMAPE: 12.43%
- Improvement: +39.15%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 3,698,000,000.00

Base Forecast: 7,000,618,221.00 (Error: +3,302,618,221.00)

Score Forecast: 6,654,703,673.00 (Error: +2,956,703,673.00)

Accuracy:

- Base sMAPE: 61.74%
- Score sMAPE: 57.12%
- Improvement: +10.47%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -16,088,000,000.00

Base Forecast: -16,499,819,402.00 (Error: -411,819,402.00)

Score Forecast: -17,315,108,479.00 (Error: -1,227,108,479.00)

Accuracy:

- Base sMAPE: 2.53%
- Score sMAPE: 7.35%
- Improvement: -197.97%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 46,309,000,000.00

Base Forecast: 56,542,066,461.00 (Error: +10,233,066,461.00)

Score Forecast: 53,748,209,873.00 (Error: +7,439,209,873.00)

Accuracy:

- Base sMAPE: 19.90%
- Score sMAPE: 14.87%
- Improvement: +27.30%

Direction Correct: Base=True, Score=True

Winner: Score

```
=====
=====
RECOMMENDATIONS
=====
=====
```

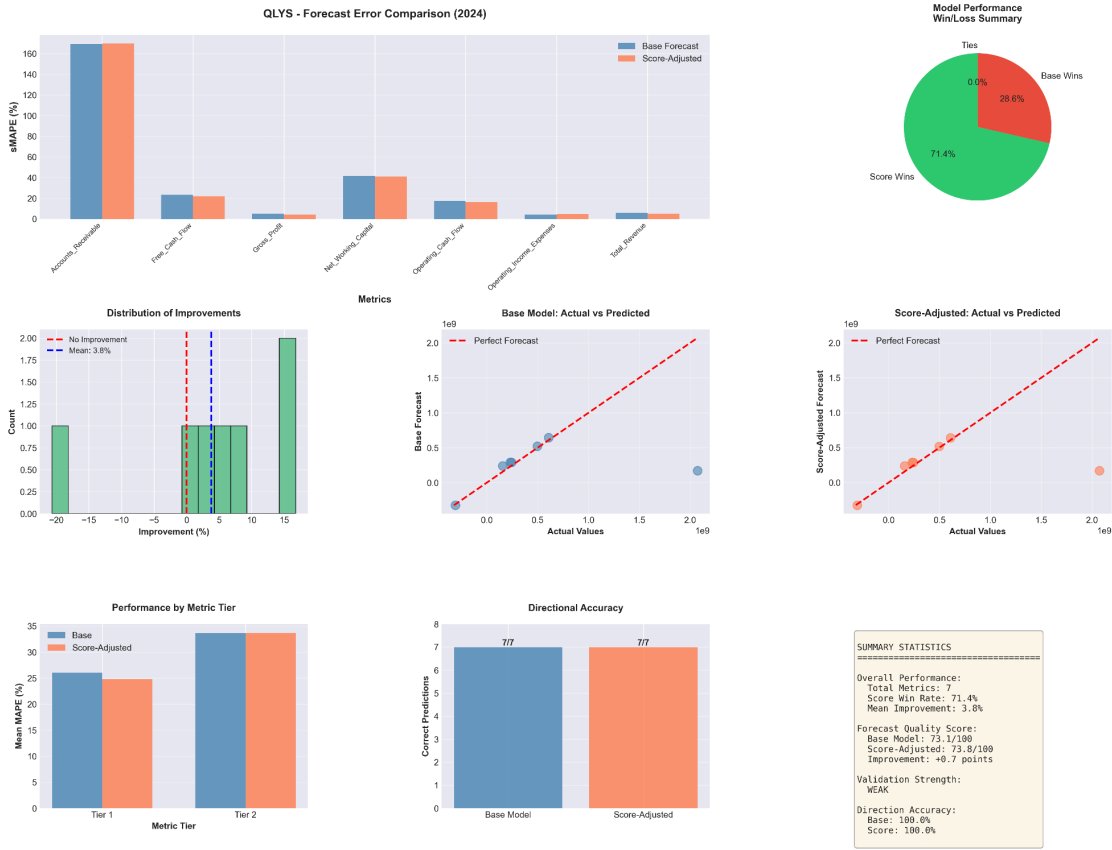
✓ Strong validation achieved. Qualitative score adjustments significantly improve forecast accuracy. Recommend:

1. Continue using score-adjusted methodology for future forecasts
2. Analyze which qualitative factors drove the best improvements
3. Consider increasing weight given to qualitative adjustments
4. Document the successful factors for replication

```
=====
=====
END OF REPORT
=====
=====
```


QLYS (NAIVE):

Forecast Validation Analysis: QLYS (2024)



FORECAST VALIDATION REPORT

Company: QLYS
Forecast Year: 2024
Analysis Date: 2025-12-10 20:34:10

EXECUTIVE SUMMARY

Validation Strength: WEAK

Score-Adjusted method shows weak validation. Qualitative factors provide marginal improvements.

Overall Win Rate: 71.4% (Score-Adjusted wins)

Mean Improvement: 3.8%

Forecast Quality Scores:

- Base Model: 73.1/100
- Score-Adjusted Model: 73.8/100
- Improvement: +0.7 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 29.33%
- Median MAPE: 19.19%
- Mean sMAPE: 38.14%

Score-Adjusted Model:

- Mean MAPE: 28.63%
- Median MAPE: 17.86%
- Mean sMAPE: 37.61%

Win/Loss Record:

- Score-Adjusted Wins: 5
- Base Model Wins: 2
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 26.08%
- Score Mean MAPE: 24.82%
- Mean Improvement: 7.84%
- Score Win Rate: 100.0%
- Wins: Score=4, Base=0

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 33.67%
- Score Mean MAPE: 33.71%
- Mean Improvement: -1.65%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

STATISTICAL SIGNIFICANCE TESTS

Paired T-Test:

- T-Statistic: 1.8178
- P-Value: 0.1190
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 4.0000
- P-Value: 0.1094
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.6871
- Interpretation: Medium effect

METRIC-BY-METRIC BREAKDOWN

Accounts_Receivable (Tier 2):

Actual Value: 2,072,000,000.00

Base Forecast: 171,393,812.00 (Error: -1,900,606,188.00)

Score Forecast: 169,632,065.00 (Error: -1,902,367,935.00)

Accuracy:

- Base sMAPE: 169.44%

- Score sMAPE: 169.73%

- Improvement: -0.09%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 231,760,000.00

Base Forecast: 292,666,068.00 (Error: +60,906,068.00)

Score Forecast: 289,488,171.00 (Error: +57,728,171.00)

Accuracy:

- Base sMAPE: 23.23%

- Score sMAPE: 22.15%

- Improvement: +5.22%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 496,089,000.00

Base Forecast: 521,521,918.00 (Error: +25,432,918.00)

Score Forecast: 517,504,525.00 (Error: +21,415,525.00)

Accuracy:

- Base sMAPE: 5.00%

- Score sMAPE: 4.23%

- Improvement: +15.80%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 157,301,000.00

Base Forecast: 240,002,273.00 (Error: +82,701,273.00)

Score Forecast: 237,957,414.00 (Error: +80,656,414.00)

Accuracy:

- Base sMAPE: 41.63%
- Score sMAPE: 40.81%
- Improvement: +2.47%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 244,094,000.00

Base Forecast: 290,935,377.00 (Error: +46,841,377.00)

Score Forecast: 287,692,251.00 (Error: +43,598,251.00)

Accuracy:

- Base sMAPE: 17.51%
- Score sMAPE: 16.40%
- Improvement: +6.92%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -308,893,000.00

Base Forecast: -321,700,661.00 (Error: -12,807,661.00)

Score Forecast: -324,346,428.00 (Error: -15,453,428.00)

Accuracy:

- Base sMAPE: 4.06%
- Score sMAPE: 4.88%
- Improvement: -20.66%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 607,571,000.00

Base Forecast: 645,738,257.00 (Error: +38,167,257.00)

Score Forecast: 639,348,639.00 (Error: +31,777,639.00)

Accuracy:

- Base sMAPE: 6.09%
- Score sMAPE: 5.10%
- Improvement: +16.74%

Direction Correct: Base=True, Score=True
Winner: Score

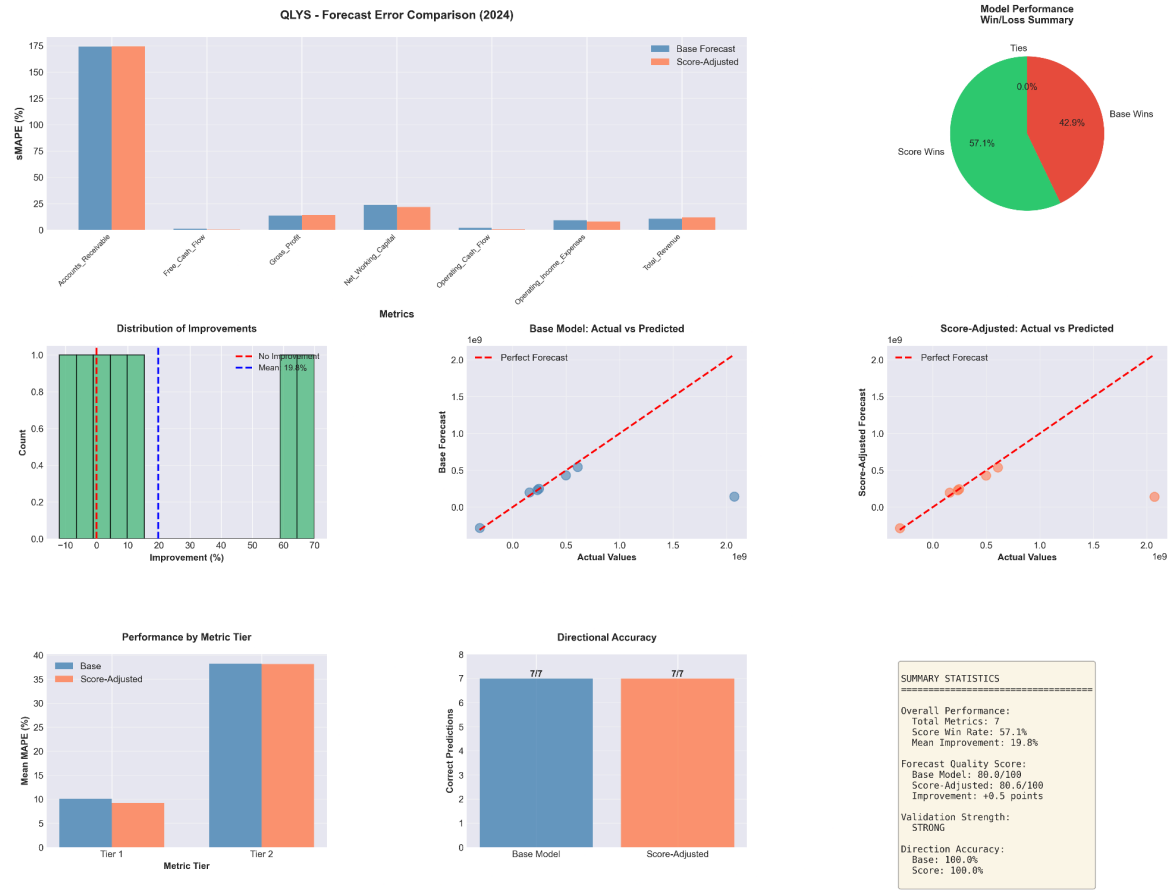
```
=====
=====
RECOMMENDATIONS
=====
=====
```

- △ Weak validation. Qualitative adjustments show marginal impact. Recommend:
1. Review and refine the qualitative scoring methodology
 2. Investigate if certain metric types benefit more from adjustments
 3. Consider whether the base model is already capturing qualitative factors
 4. Test alternative adjustment algorithms

```
=====
=====
END OF REPORT
=====
=====
```


QLYS (LINEAR):

Forecast Validation Analysis: QLYS (2024)



FORECAST VALIDATION REPORT

Company: QLYS
Forecast Year: 2024
Analysis Date: 2025-12-10 20:38:02

EXECUTIVE SUMMARY

Validation Strength: STRONG

Score-Adjusted method shows strong validation. Qualitative factors significantly improve forecast accuracy.

Overall Win Rate: 57.1% (Score-Adjusted wins)

Mean Improvement: 19.8%

Forecast Quality Scores:

- Base Model: 80.0/100
- Score-Adjusted Model: 80.6/100
- Improvement: +0.5 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 22.18%
- Median MAPE: 10.14%
- Mean sMAPE: 33.56%

Score-Adjusted Model:

- Mean MAPE: 21.64%
- Median MAPE: 11.35%
- Mean sMAPE: 33.16%

Win/Loss Record:

- Score-Adjusted Wins: 4
- Base Model Wins: 3
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 10.13%
- Score Mean MAPE: 9.27%
- Mean Improvement: 32.75%
- Score Win Rate: 75.0%
- Wins: Score=3, Base=1

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 38.25%
- Score Mean MAPE: 38.13%
- Mean Improvement: 2.43%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 0.8455
- P-Value: 0.4303
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 9.0000
- P-Value: 0.4688
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.3196
- Interpretation: Small effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 2,072,000,000.00

Base Forecast: 142,489,470.00 (Error: -1,929,510,530.00)

Score Forecast: 140,569,282.00 (Error: -1,931,430,718.00)

Accuracy:

- Base sMAPE: 174.26%
- Score sMAPE: 174.59%
- Improvement: -0.10%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 231,760,000.00

Base Forecast: 234,184,183.00 (Error: +2,424,183.00)

Score Forecast: 231,028,317.00 (Error: -731,683.00)

Accuracy:

- Base sMAPE: 1.04%
- Score sMAPE: 0.32%
- Improvement: +69.82%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 496,089,000.00

Base Forecast: 433,498,933.00 (Error: -62,590,067.00)

Score Forecast: 429,602,644.00 (Error: -66,486,356.00)

Accuracy:

- Base sMAPE: 13.47%
- Score sMAPE: 14.36%
- Improvement: -6.23%

Direction Correct: Base=True, Score=True

Winner: Base

Net_Working_Capital (Tier 1):

Actual Value: 157,301,000.00
Base Forecast: 200,063,165.00 (Error: +42,762,165.00)
Score Forecast: 196,017,888.00 (Error: +38,716,888.00)

Accuracy:

- Base sMAPE: 23.93%
- Score sMAPE: 21.92%
- Improvement: +9.46%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 244,094,000.00
Base Forecast: 249,372,730.00 (Error: +5,278,730.00)
Score Forecast: 246,012,183.00 (Error: +1,918,183.00)

Accuracy:

- Base sMAPE: 2.14%
- Score sMAPE: 0.78%
- Improvement: +63.66%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -308,893,000.00
Base Forecast: -281,089,193.00 (Error: +27,803,807.00)
Score Forecast: -284,877,151.00 (Error: +24,015,849.00)

Accuracy:

- Base sMAPE: 9.43%
- Score sMAPE: 8.09%
- Improvement: +13.62%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 607,571,000.00
Base Forecast: 545,984,844.00 (Error: -61,586,156.00)
Score Forecast: 538,627,152.00 (Error: -68,943,848.00)

Accuracy:

- Base sMAPE: 10.68%
- Score sMAPE: 12.03%
- Improvement: -11.95%

Direction Correct: Base=True, Score=True

Winner: Base

```
=====
=====
RECOMMENDATIONS
=====
=====
```

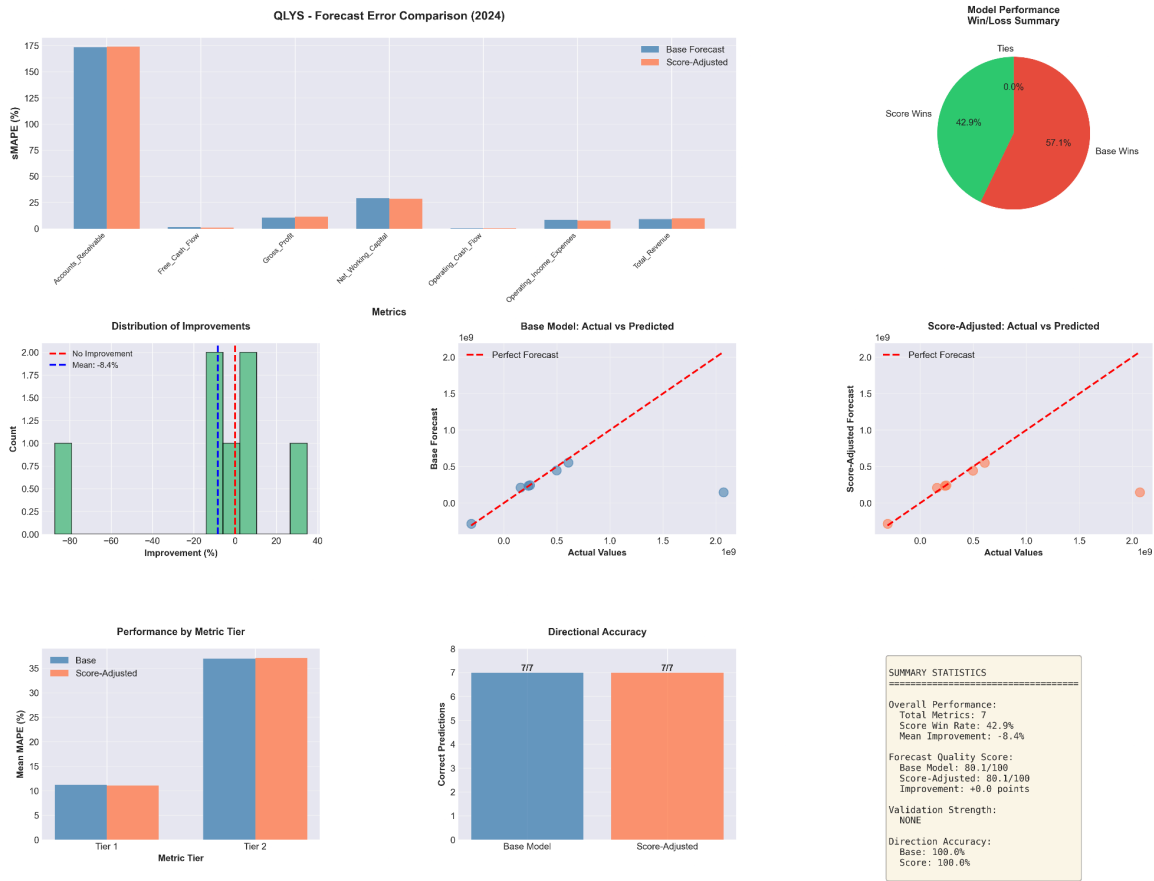
✓ Strong validation achieved. Qualitative score adjustments significantly improve forecast accuracy. Recommend:

1. Continue using score-adjusted methodology for future forecasts
2. Analyze which qualitative factors drove the best improvements
3. Consider increasing weight given to qualitative adjustments
4. Document the successful factors for replication

```
=====
=====
END OF REPORT
=====
=====
```


QLYS (XGBOOST):

Forecast Validation Analysis: QLYS (2024)



FORECAST VALIDATION REPORT

Company: QLYS
Forecast Year: 2024
Analysis Date: 2025-12-10 20:31:21

EXECUTIVE SUMMARY

=====

Validation Strength: NONE

Score-Adjusted method does not validate. Base model performs as well or better.

Overall Win Rate: 42.9% (Score-Adjusted wins)

Mean Improvement: -8.4%

Forecast Quality Scores:

- Base Model: 80.1/100
- Score-Adjusted Model: 80.1/100
- Improvement: +0.0 points

=====

DETAILED METRICS COMPARISON

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 22.23%
- Median MAPE: 8.74%
- Mean sMAPE: 33.23%

Score-Adjusted Model:

- Mean MAPE: 22.20%
- Median MAPE: 9.29%
- Mean sMAPE: 33.26%

Win/Loss Record:

- Score-Adjusted Wins: 3
- Base Model Wins: 4
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 11.18%
- Score Mean MAPE: 11.01%
- Mean Improvement: -14.07%
- Score Win Rate: 50.0%
- Wins: Score=2, Base=2

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 36.98%
- Score Mean MAPE: 37.12%
- Mean Improvement: -0.93%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: -0.1033
- P-Value: 0.9211
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 13.0000
- P-Value: 0.9375
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: -0.0391
- Interpretation: Small effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 2,072,000,000.00

Base Forecast: 146,225,952.00 (Error: -1,925,774,048.00)

Score Forecast: 145,348,596.00 (Error: -1,926,651,404.00)

Accuracy:

- Base sMAPE: 173.63%
- Score sMAPE: 173.78%
- Improvement: -0.05%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 231,760,000.00

Base Forecast: 235,818,960.00 (Error: +4,058,960.00)

Score Forecast: 234,404,046.00 (Error: +2,644,046.00)

Accuracy:

- Base sMAPE: 1.74%
- Score sMAPE: 1.13%
- Improvement: +34.86%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 496,089,000.00

Base Forecast: 446,972,896.00 (Error: -49,116,104.00)

Score Forecast: 442,950,139.00 (Error: -53,138,861.00)

Accuracy:

- Base sMAPE: 10.42%
- Score sMAPE: 11.32%
- Improvement: -8.19%

Direction Correct: Base=True, Score=True

Winner: Base

Net_Working_Capital (Tier 1):

Actual Value: 157,301,000.00
Base Forecast: 210,789,984.00 (Error: +53,488,984.00)
Score Forecast: 209,525,244.00 (Error: +52,224,244.00)

Accuracy:

- Base sMAPE: 29.06%
- Score sMAPE: 28.47%
- Improvement: +2.36%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 244,094,000.00
Base Forecast: 244,604,912.00 (Error: +510,912.00)
Score Forecast: 243,137,282.00 (Error: -956,718.00)

Accuracy:

- Base sMAPE: 0.21%
- Score sMAPE: 0.39%
- Improvement: -87.26%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Income_Expenses (Tier 2):

Actual Value: -308,893,000.00
Base Forecast: -283,903,904.00 (Error: +24,989,096.00)
Score Forecast: -285,266,642.00 (Error: +23,626,358.00)

Accuracy:

- Base sMAPE: 8.43%
- Score sMAPE: 7.95%
- Improvement: +5.45%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 607,571,000.00
Base Forecast: 554,457,792.00 (Error: -53,113,208.00)
Score Forecast: 551,131,045.00 (Error: -56,439,955.00)

Accuracy:

- Base sMAPE: 9.14%
- Score sMAPE: 9.74%
- Improvement: -6.26%

Direction Correct: Base=True, Score=True

Winner: Base

=====

=====

RECOMMENDATIONS

=====

=====

X No validation. Base model performs as well or better. Recommend:

1. Critically review the qualitative scoring framework
2. Check if qualitative factors are already reflected in historical data
3. Consider that market efficiency may limit qualitative alpha
4. Explore if timing or sector-specific factors affected results

=====

=====

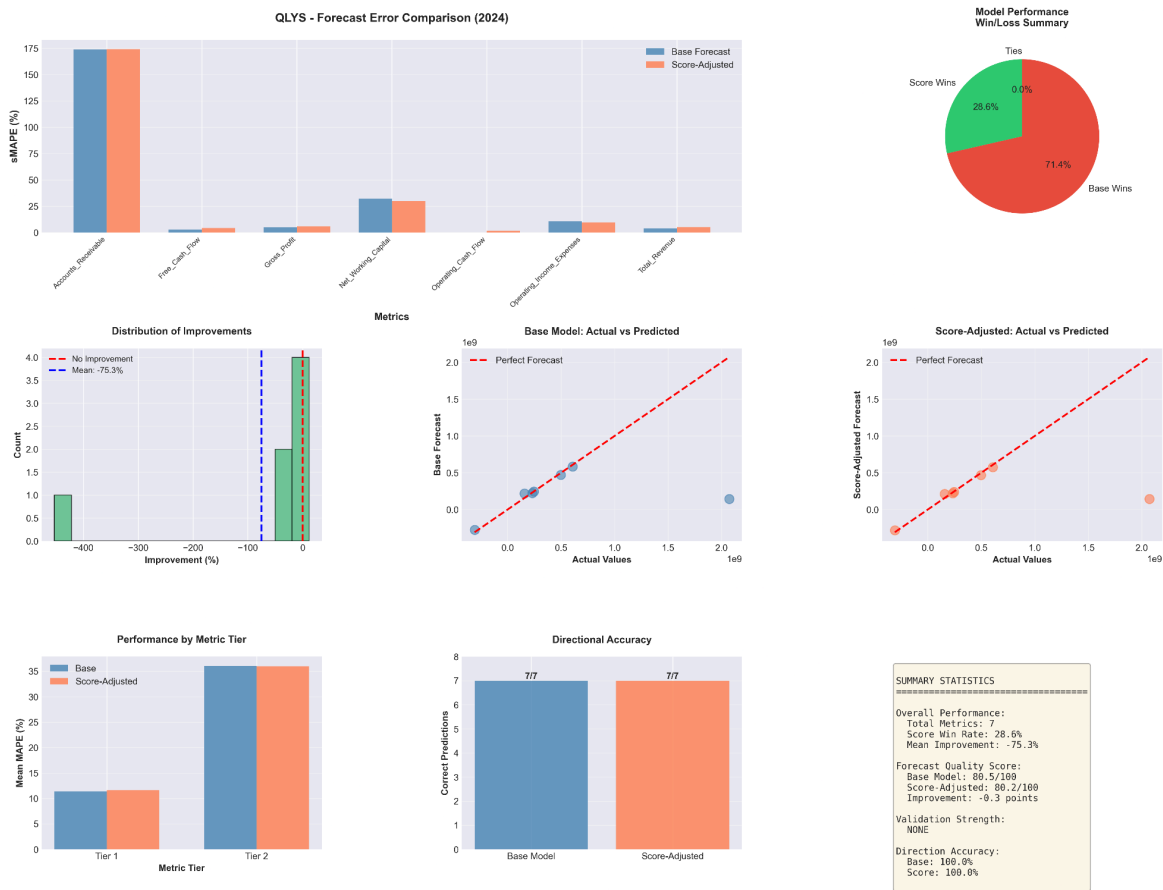
END OF REPORT

=====

=====

QLYS (ARIMA):

Forecast Validation Analysis: QLYS (2024)



FORECAST VALIDATION REPORT

Company: QLYS
Forecast Year: 2024
Analysis Date: 2025-12-10 20:48:11

EXECUTIVE SUMMARY

=====

Validation Strength: NONE

Score-Adjusted method does not validate. Base model performs as well or better.

Overall Win Rate: 28.6% (Score-Adjusted wins)

Mean Improvement: -75.3%

Forecast Quality Scores:

- Base Model: 80.5/100
- Score-Adjusted Model: 80.2/100
- Improvement: -0.3 points

=====

DETAILED METRICS COMPARISON

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 21.97%
- Median MAPE: 4.91%
- Mean sMAPE: 32.76%

Score-Adjusted Model:

- Mean MAPE: 22.09%
- Median MAPE: 5.76%
- Mean sMAPE: 33.04%

Win/Loss Record:

- Score-Adjusted Wins: 2
- Base Model Wins: 5
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 11.39%
- Score Mean MAPE: 11.67%
- Mean Improvement: -130.36%
- Score Win Rate: 25.0%
- Wins: Score=1, Base=3

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 36.07%
- Score Mean MAPE: 35.99%
- Mean Improvement: -1.93%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: -0.5352
- P-Value: 0.6118
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 10.0000
- P-Value: 0.5781
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: -0.2023
- Interpretation: Small effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 2,072,000,000.00

Base Forecast: 144,827,661.00 (Error: -1,927,172,339.00)

Score Forecast: 142,875,963.00 (Error: -1,929,124,037.00)

Accuracy:

- Base sMAPE: 173.87%
- Score sMAPE: 174.20%
- Improvement: -0.10%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 231,760,000.00

Base Forecast: 224,690,886.00 (Error: -7,069,114.00)

Score Forecast: 221,662,952.00 (Error: -10,097,048.00)

Accuracy:

- Base sMAPE: 3.10%
- Score sMAPE: 4.45%
- Improvement: -42.83%

Direction Correct: Base=True, Score=True

Winner: Base

Gross_Profit (Tier 2):

Actual Value: 496,089,000.00

Base Forecast: 471,739,038.00 (Error: -24,349,962.00)

Score Forecast: 467,499,048.00 (Error: -28,589,952.00)

Accuracy:

- Base sMAPE: 5.03%
- Score sMAPE: 5.93%
- Improvement: -17.41%

Direction Correct: Base=True, Score=True

Winner: Base

Net_Working_Capital (Tier 1):

Actual Value: 157,301,000.00
Base Forecast: 217,488,137.00 (Error: +60,187,137.00)
Score Forecast: 213,090,527.00 (Error: +55,789,527.00)

Accuracy:

- Base sMAPE: 32.12%
- Score sMAPE: 30.12%
- Improvement: +7.31%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 244,094,000.00
Base Forecast: 243,370,030.00 (Error: -723,970.00)
Score Forecast: 240,090,375.00 (Error: -4,003,625.00)

Accuracy:

- Base sMAPE: 0.30%
- Score sMAPE: 1.65%
- Improvement: -453.01%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Income_Expenses (Tier 2):

Actual Value: -308,893,000.00
Base Forecast: -277,063,578.00 (Error: +31,829,422.00)
Score Forecast: -280,797,287.00 (Error: +28,095,713.00)

Accuracy:

- Base sMAPE: 10.86%
- Score sMAPE: 9.53%
- Improvement: +11.73%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 607,571,000.00
Base Forecast: 583,660,132.00 (Error: -23,910,868.00)
Score Forecast: 575,794,728.00 (Error: -31,776,272.00)

Accuracy:

- Base sMAPE: 4.01%
- Score sMAPE: 5.37%
- Improvement: -32.89%

Direction Correct: Base=True, Score=True

Winner: Base

=====

=====

RECOMMENDATIONS

=====

=====

X No validation. Base model performs as well or better. Recommend:

1. Critically review the qualitative scoring framework
2. Check if qualitative factors are already reflected in historical data
3. Consider that market efficiency may limit qualitative alpha
4. Explore if timing or sector-specific factors affected results

=====

=====

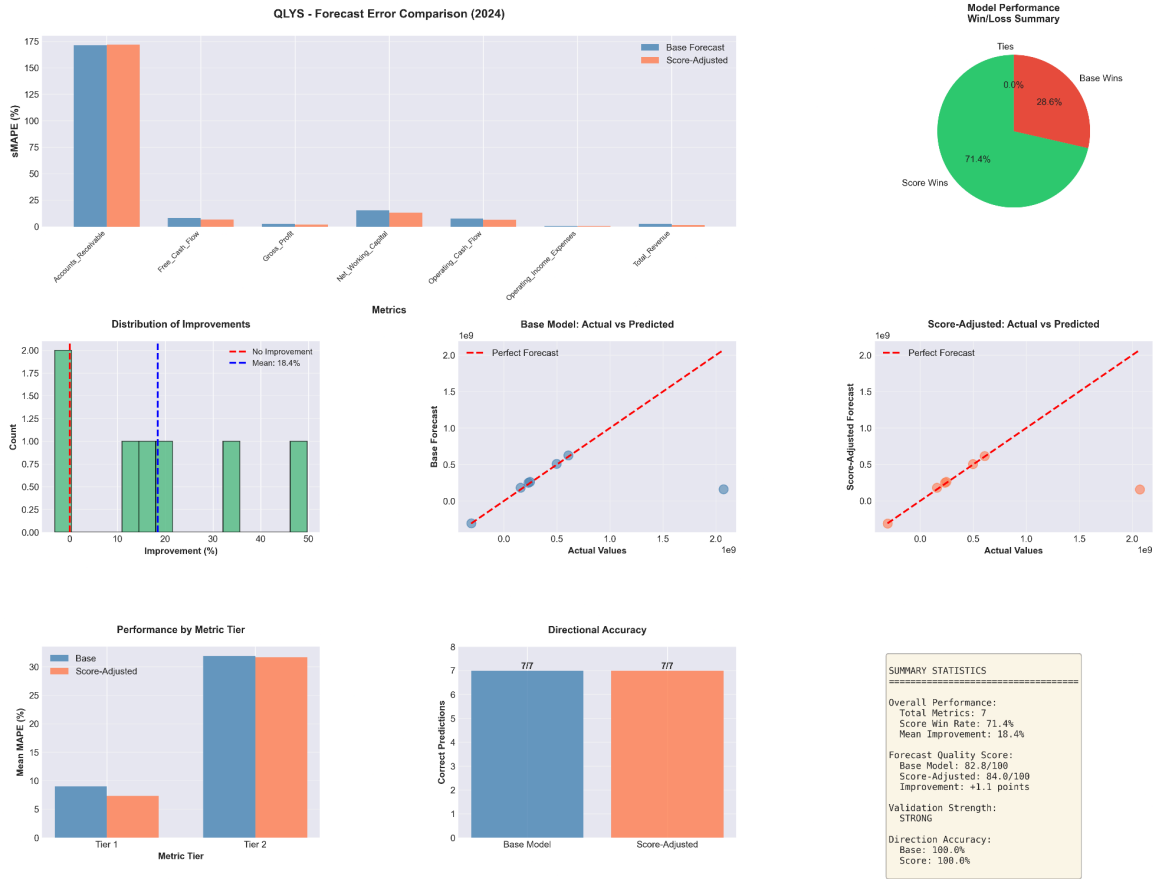
END OF REPORT

=====

=====

QLYS (WINTERS-HOLT):

Forecast Validation Analysis: QLYS (2024)



FORECAST VALIDATION REPORT

Company: QLYS
Forecast Year: 2024
Analysis Date: 2025-12-10 21:37:21

EXECUTIVE SUMMARY

=====

Validation Strength: **STRONG**

Score-Adjusted method shows strong validation. Qualitative factors significantly improve forecast accuracy.

Overall Win Rate: 71.4% (Score-Adjusted wins)

Mean Improvement: 18.4%

Forecast Quality Scores:

- Base Model: 82.8/100
- Score-Adjusted Model: 84.0/100
- Improvement: +1.1 points

=====

DETAILED METRICS COMPARISON

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 18.81%
- Median MAPE: 8.04%
- Mean sMAPE: 29.82%

Score-Adjusted Model:

- Mean MAPE: 17.75%
- Median MAPE: 6.58%
- Mean sMAPE: 28.87%

Win/Loss Record:

- Score-Adjusted Wins: 5
- Base Model Wins: 2
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 8.99%
- Score Mean MAPE: 7.32%
- Mean Improvement: 24.83%
- Score Win Rate: 100.0%
- Wins: Score=4, Base=0

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 31.91%
- Score Mean MAPE: 31.65%
- Mean Improvement: 9.92%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 2.9541
- P-Value: 0.0255
- Significant at $\alpha=0.05$: True
- Interpretation: Score-Adjusted significantly different

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 3.0000
- P-Value: 0.0781
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 1.1166
- Interpretation: Large effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 2,072,000,000.00

Base Forecast: 159,952,596.00 (Error: -1,912,047,404.00)

Score Forecast: 157,797,075.00 (Error: -1,914,202,925.00)

Accuracy:

- Base sMAPE: 171.33%
- Score sMAPE: 171.69%
- Improvement: -0.11%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 231,760,000.00

Base Forecast: 251,309,833.00 (Error: +19,549,833.00)

Score Forecast: 247,923,182.00 (Error: +16,163,182.00)

Accuracy:

- Base sMAPE: 8.09%
- Score sMAPE: 6.74%
- Improvement: +17.32%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 496,089,000.00

Base Forecast: 509,989,857.00 (Error: +13,900,857.00)

Score Forecast: 505,406,069.00 (Error: +9,317,069.00)

Accuracy:

- Base sMAPE: 2.76%
- Score sMAPE: 1.86%
- Improvement: +32.97%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 157,301,000.00

Base Forecast: 183,560,927.00 (Error: +26,259,927.00)

Score Forecast: 179,849,325.00 (Error: +22,548,325.00)

Accuracy:

- Base sMAPE: 15.41%
- Score sMAPE: 13.38%
- Improvement: +14.13%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 244,094,000.00

Base Forecast: 263,712,761.00 (Error: +19,618,761.00)

Score Forecast: 260,158,968.00 (Error: +16,064,968.00)

Accuracy:

- Base sMAPE: 7.73%
- Score sMAPE: 6.37%
- Improvement: +18.11%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -308,893,000.00

Base Forecast: -306,856,931.00 (Error: +2,036,069.00)

Score Forecast: -310,992,135.00 (Error: -2,099,135.00)

Accuracy:

- Base sMAPE: 0.66%
- Score sMAPE: 0.68%
- Improvement: -3.10%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 607,571,000.00

Base Forecast: 624,494,156.00 (Error: +16,923,156.00)

Score Forecast: 616,078,473.00 (Error: +8,507,473.00)

Accuracy:

- Base sMAPE: 2.75%
- Score sMAPE: 1.39%
- Improvement: +49.73%

Direction Correct: Base=True, Score=True

Winner: Score

```
=====
=====
RECOMMENDATIONS
=====
=====
```

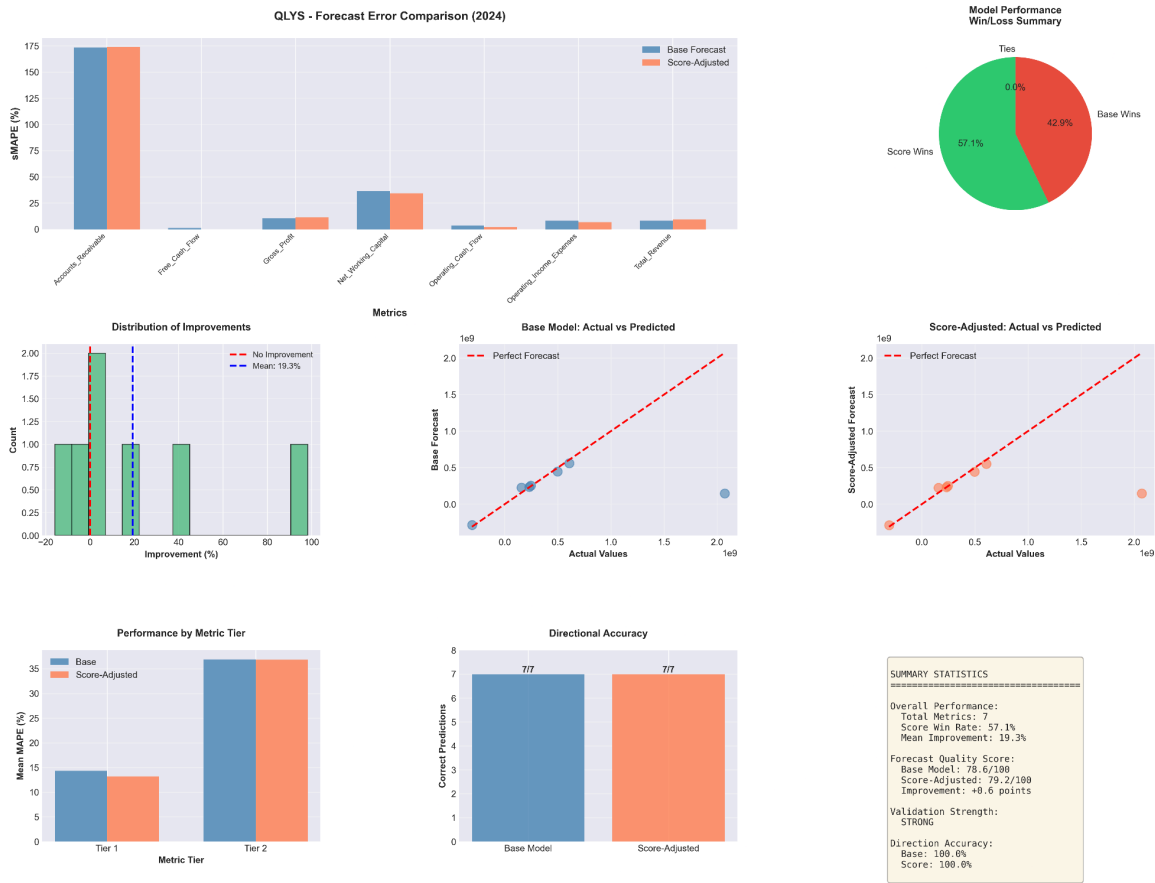
✓ Strong validation achieved. Qualitative score adjustments significantly improve forecast accuracy. Recommend:

1. Continue using score-adjusted methodology for future forecasts
2. Analyze which qualitative factors drove the best improvements
3. Consider increasing weight given to qualitative adjustments
4. Document the successful factors for replication

```
=====
=====
END OF REPORT
=====
=====
```


QLYS (LINEAR + WINTERS-HOLT):

Forecast Validation Analysis: QLYS (2024)



FORECAST VALIDATION REPORT

Company: QLYS
Forecast Year: 2024
Analysis Date: 2025-12-10 21:47:48

EXECUTIVE SUMMARY

=====

Validation Strength: **STRONG**

Score-Adjusted method shows strong validation. Qualitative factors significantly improve forecast accuracy.

Overall Win Rate: 57.1% (Score-Adjusted wins)

Mean Improvement: 19.3%

Forecast Quality Scores:

- Base Model: 78.6/100
- Score-Adjusted Model: 79.2/100
- Improvement: +0.6 points

=====

DETAILED METRICS COMPARISON

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 23.99%
- Median MAPE: 7.84%
- Mean sMAPE: 34.51%

Score-Adjusted Model:

- Mean MAPE: 23.31%
- Median MAPE: 9.09%
- Mean sMAPE: 34.02%

Win/Loss Record:

- Score-Adjusted Wins: 4
- Base Model Wins: 3
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 14.31%
- Score Mean MAPE: 13.21%
- Mean Improvement: 31.83%
- Score Win Rate: 75.0%
- Wins: Score=3, Base=1

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 36.89%
- Score Mean MAPE: 36.78%
- Mean Improvement: 2.54%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 0.9756
- P-Value: 0.3669
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 8.0000
- P-Value: 0.3750
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.3687
- Interpretation: Small effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 2,072,000,000.00

Base Forecast: 147,006,561.00 (Error: -1,924,993,439.00)

Score Forecast: 145,025,501.00 (Error: -1,926,974,499.00)

Accuracy:

- Base sMAPE: 173.50%
- Score sMAPE: 173.83%
- Improvement: -0.10%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 231,760,000.00

Base Forecast: 234,869,658.00 (Error: +3,109,658.00)

Score Forecast: 231,704,555.00 (Error: -55,445.00)

Accuracy:

- Base sMAPE: 1.33%
- Score sMAPE: 0.02%
- Improvement: +98.22%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 496,089,000.00

Base Forecast: 446,751,958.00 (Error: -49,337,042.00)

Score Forecast: 442,736,552.00 (Error: -53,352,448.00)

Accuracy:

- Base sMAPE: 10.47%
- Score sMAPE: 11.37%
- Improvement: -8.14%

Direction Correct: Base=True, Score=True

Winner: Base

Net_Working_Capital (Tier 1):

Actual Value: 157,301,000.00

Base Forecast: 227,172,176.00 (Error: +69,871,176.00)

Score Forecast: 222,578,755.00 (Error: +65,277,755.00)

Accuracy:

- Base sMAPE: 36.35%
- Score sMAPE: 34.37%
- Improvement: +6.57%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 244,094,000.00

Base Forecast: 252,983,428.00 (Error: +8,889,428.00)

Score Forecast: 249,574,224.00 (Error: +5,480,224.00)

Accuracy:

- Base sMAPE: 3.58%
- Score sMAPE: 2.22%
- Improvement: +38.35%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -308,893,000.00

Base Forecast: -284,694,214.00 (Error: +24,198,786.00)

Score Forecast: -288,530,754.00 (Error: +20,362,246.00)

Accuracy:

- Base sMAPE: 8.15%
- Score sMAPE: 6.82%
- Improvement: +15.85%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 607,571,000.00

Base Forecast: 559,907,572.00 (Error: -47,663,428.00)

Score Forecast: 552,362,258.00 (Error: -55,208,742.00)

Accuracy:

- Base sMAPE: 8.17%
- Score sMAPE: 9.52%
- Improvement: -15.83%

Direction Correct: Base=True, Score=True

Winner: Base

=====

=====

RECOMMENDATIONS

=====

=====

✓ Strong validation achieved. Qualitative score adjustments significantly improve forecast accuracy. Recommend:

1. Continue using score-adjusted methodology for future forecasts
2. Analyze which qualitative factors drove the best improvements
3. Consider increasing weight given to qualitative adjustments
4. Document the successful factors for replication

=====

=====

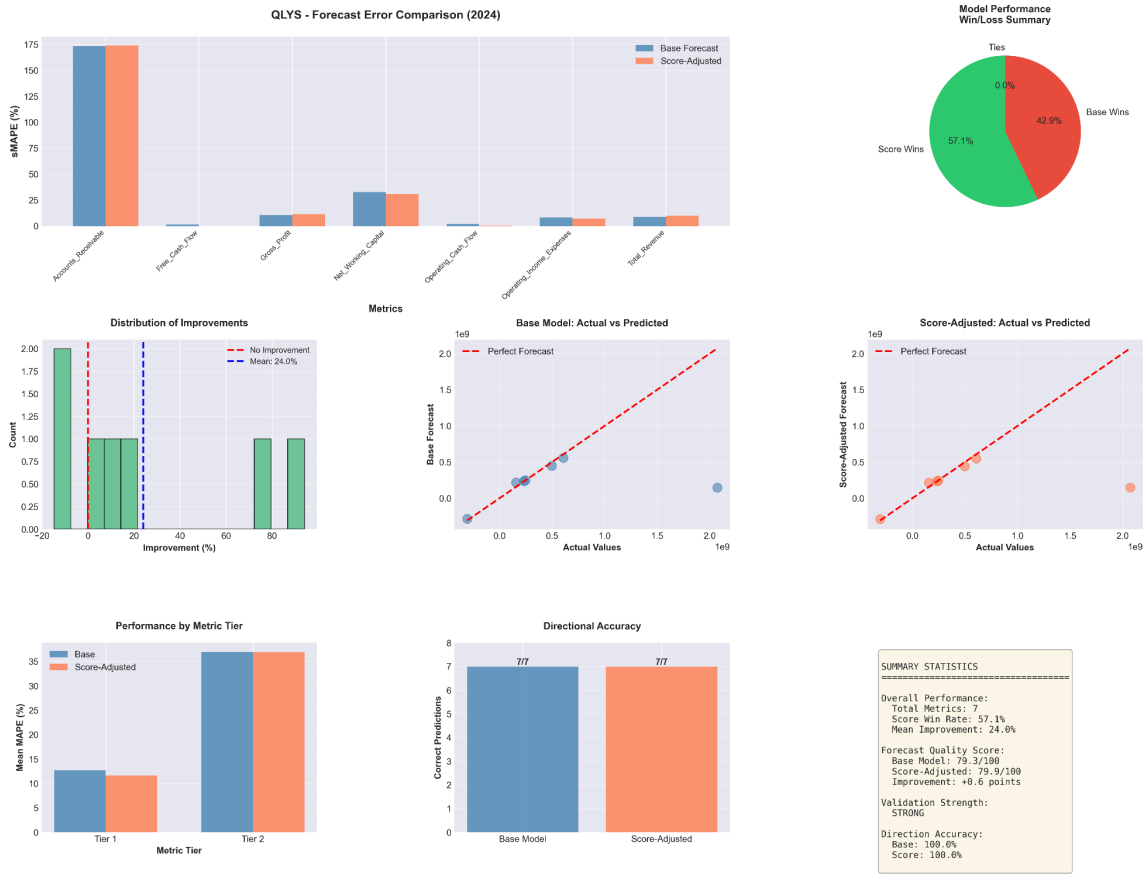
END OF REPORT

=====

=====

QLYS (LINEAR + WINTERS-HOLT + XGBOOST):

Forecast Validation Analysis: QLYS (2024)



FORECAST VALIDATION REPORT

Company: QLYS
Forecast Year: 2024
Analysis Date: 2025-12-10 22:39:00

EXECUTIVE SUMMARY

Validation Strength: STRONG

Score-Adjusted method shows strong validation. Qualitative factors significantly improve forecast accuracy.

Overall Win Rate: 57.1% (Score-Adjusted wins)

Mean Improvement: 24.0%

Forecast Quality Scores:

- Base Model: 79.3/100
- Score-Adjusted Model: 79.9/100
- Improvement: +0.6 points

DETAILED METRICS COMPARISON

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 23.12%
- Median MAPE: 8.37%
- Mean sMAPE: 33.88%

Score-Adjusted Model:

- Mean MAPE: 22.45%
- Median MAPE: 9.60%
- Mean sMAPE: 33.39%

Win/Loss Record:

- Score-Adjusted Wins: 4
- Base Model Wins: 3
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 12.72%
- Score Mean MAPE: 11.64%
- Mean Improvement: 40.14%
- Score Win Rate: 75.0%
- Wins: Score=3, Base=1

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 36.98%
- Score Mean MAPE: 36.87%
- Mean Improvement: 2.42%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 0.9875
- P-Value: 0.3615
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 7.0000
- P-Value: 0.2969
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.3733
- Interpretation: Small effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 2,072,000,000.00

Base Forecast: 147,006,561.00 (Error: -1,924,993,439.00)

Score Forecast: 145,025,501.00 (Error: -1,926,974,499.00)

Accuracy:

- Base sMAPE: 173.50%
- Score sMAPE: 173.83%
- Improvement: -0.10%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 231,760,000.00

Base Forecast: 235,128,701.00 (Error: +3,368,701.00)

Score Forecast: 231,960,106.00 (Error: +200,106.00)

Accuracy:

- Base sMAPE: 1.44%
- Score sMAPE: 0.09%
- Improvement: +94.06%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 496,089,000.00

Base Forecast: 446,467,067.00 (Error: -49,621,933.00)

Score Forecast: 442,454,221.00 (Error: -53,634,779.00)

Accuracy:

- Base sMAPE: 10.53%
- Score sMAPE: 11.43%
- Improvement: -8.09%

Direction Correct: Base=True, Score=True

Winner: Base

Net_Working_Capital (Tier 1):

Actual Value: 157,301,000.00
Base Forecast: 218,970,792.00 (Error: +61,669,792.00)
Score Forecast: 214,543,202.00 (Error: +57,242,202.00)

Accuracy:

- Base sMAPE: 32.78%
- Score sMAPE: 30.79%
- Improvement: +7.18%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 244,094,000.00
Base Forecast: 248,617,882.00 (Error: +4,523,882.00)
Score Forecast: 245,267,507.00 (Error: +1,173,507.00)

Accuracy:

- Base sMAPE: 1.84%
- Score sMAPE: 0.48%
- Improvement: +74.06%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -308,893,000.00
Base Forecast: -284,096,067.00 (Error: +24,796,933.00)
Score Forecast: -287,924,546.00 (Error: +20,968,454.00)

Accuracy:

- Base sMAPE: 8.36%
- Score sMAPE: 7.03%
- Improvement: +15.44%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 607,571,000.00
Base Forecast: 556,724,474.00 (Error: -50,846,526.00)
Score Forecast: 549,222,055.00 (Error: -58,348,945.00)

Accuracy:

- Base sMAPE: 8.73%
- Score sMAPE: 10.09%
- Improvement: -14.76%

Direction Correct: Base=True, Score=True

Winner: Base

=====

=====

=====

RECOMMENDATIONS

=====

=====

✓ Strong validation achieved. Qualitative score adjustments significantly improve forecast accuracy. Recommend:

1. Continue using score-adjusted methodology for future forecasts
2. Analyze which qualitative factors drove the best improvements
3. Consider increasing weight given to qualitative adjustments
4. Document the successful factors for replication

=====

=====

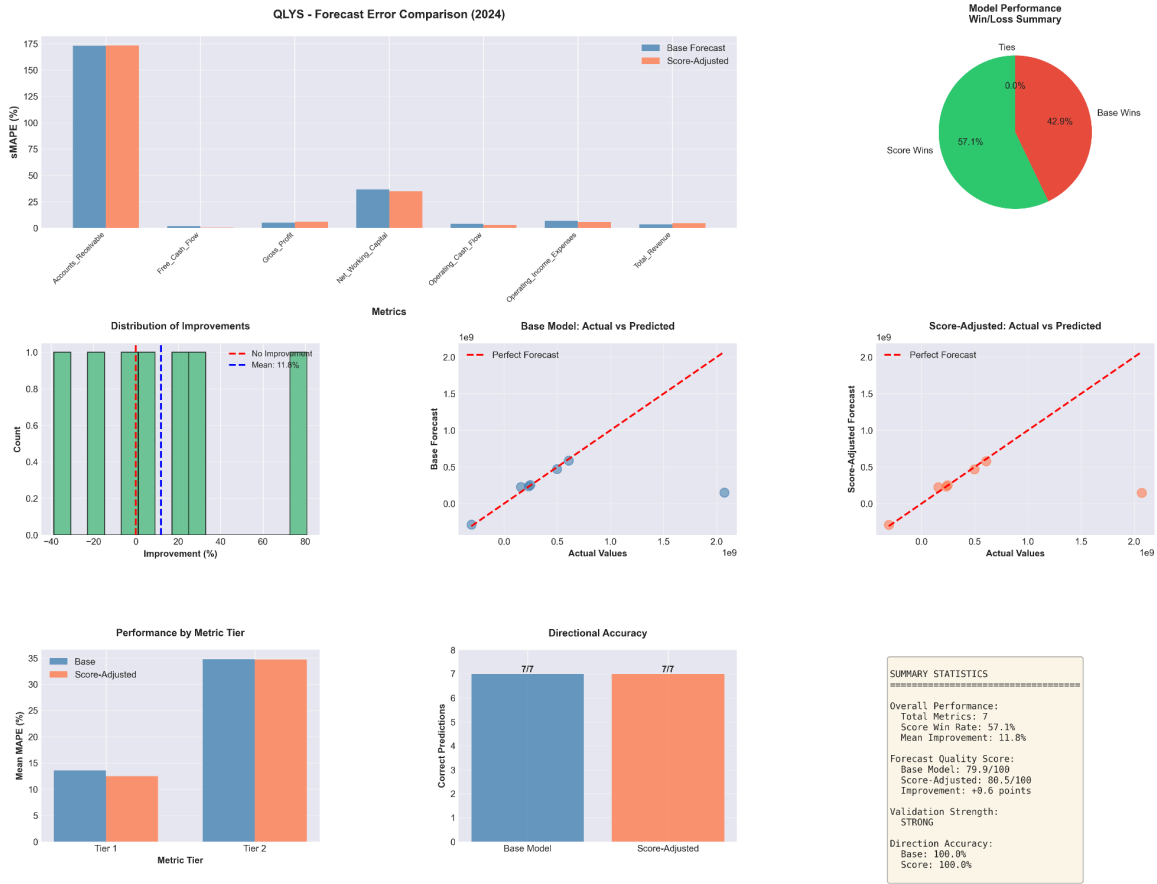
END OF REPORT

=====

=====

QLYS (LINEAR + WINTERS-HOLT + ARIMA):

Forecast Validation Analysis: QLYS (2024)



FORECAST VALIDATION REPORT

Company: QLYS
Forecast Year: 2024
Analysis Date: 2025-12-10 23:12:58

EXECUTIVE SUMMARY

=====

Validation Strength: **STRONG**

Score-Adjusted method shows strong validation. Qualitative factors significantly improve forecast accuracy.

Overall Win Rate: 57.1% (Score-Adjusted wins)

Mean Improvement: 11.8%

Forecast Quality Scores:

- Base Model: 79.9/100
- Score-Adjusted Model: 80.5/100
- Improvement: +0.6 points

=====

DETAILED METRICS COMPARISON

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 22.68%
- Median MAPE: 4.92%
- Mean sMAPE: 33.01%

Score-Adjusted Model:

- Mean MAPE: 22.00%
- Median MAPE: 5.45%
- Mean sMAPE: 32.52%

Win/Loss Record:

- Score-Adjusted Wins: 4
- Base Model Wins: 3
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 13.58%
- Score Mean MAPE: 12.48%
- Mean Improvement: 20.31%
- Score Win Rate: 75.0%
- Wins: Score=3, Base=1

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 34.80%
- Score Mean MAPE: 34.70%
- Mean Improvement: 0.43%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 0.9817
- P-Value: 0.3642
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 7.0000
- P-Value: 0.2969
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.3710
- Interpretation: Small effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 2,072,000,000.00

Base Forecast: 149,728,365.00 (Error: -1,922,271,635.00)

Score Forecast: 147,710,626.00 (Error: -1,924,289,374.00)

Accuracy:

- Base sMAPE: 173.04%
- Score sMAPE: 173.38%
- Improvement: -0.10%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 231,760,000.00

Base Forecast: 235,693,190.00 (Error: +3,933,190.00)

Score Forecast: 232,516,989.00 (Error: +756,989.00)

Accuracy:

- Base sMAPE: 1.68%
- Score sMAPE: 0.33%
- Improvement: +80.75%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 496,089,000.00

Base Forecast: 471,659,684.00 (Error: -24,429,316.00)

Score Forecast: 467,420,406.00 (Error: -28,668,594.00)

Accuracy:

- Base sMAPE: 5.05%
- Score sMAPE: 5.95%
- Improvement: -17.35%

Direction Correct: Base=True, Score=True

Winner: Base

Net_Working_Capital (Tier 1):

Actual Value: 157,301,000.00

Base Forecast: 228,053,529.00 (Error: +70,752,529.00)

Score Forecast: 223,442,286.00 (Error: +66,141,286.00)

Accuracy:

- Base sMAPE: 36.72%
- Score sMAPE: 34.74%
- Improvement: +6.52%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 244,094,000.00

Base Forecast: 254,581,203.00 (Error: +10,487,203.00)

Score Forecast: 251,150,467.00 (Error: +7,056,467.00)

Accuracy:

- Base sMAPE: 4.21%
- Score sMAPE: 2.85%
- Improvement: +32.71%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -308,893,000.00

Base Forecast: -288,170,043.00 (Error: +20,722,957.00)

Score Forecast: -292,053,423.00 (Error: +16,839,577.00)

Accuracy:

- Base sMAPE: 6.94%
- Score sMAPE: 5.60%
- Improvement: +18.74%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 607,571,000.00

Base Forecast: 587,143,329.00 (Error: -20,427,671.00)

Score Forecast: 579,230,986.00 (Error: -28,340,014.00)

Accuracy:

- Base sMAPE: 3.42%
- Score sMAPE: 4.78%
- Improvement: -38.73%

Direction Correct: Base=True, Score=True

Winner: Base

=====

=====

RECOMMENDATIONS

=====

=====

✓ Strong validation achieved. Qualitative score adjustments significantly improve forecast accuracy. Recommend:

1. Continue using score-adjusted methodology for future forecasts
2. Analyze which qualitative factors drove the best improvements
3. Consider increasing weight given to qualitative adjustments
4. Document the successful factors for replication

=====

=====

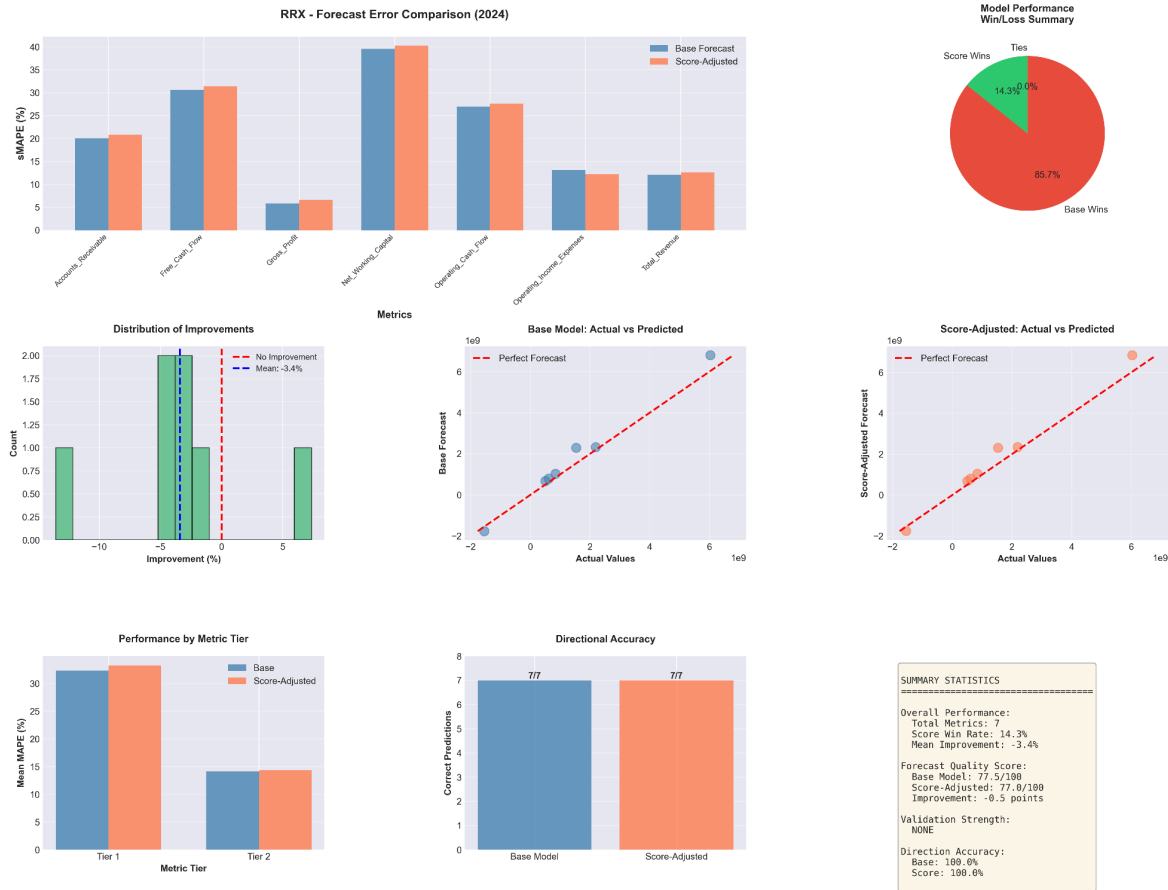
END OF REPORT

=====

=====

RRX (NAIVE):

Forecast Validation Analysis: RRX (2024)



FORECAST VALIDATION REPORT

Company: RRX
Forecast Year: 2024
Analysis Date: 2025-12-12 17:48:41

EXECUTIVE SUMMARY

Validation Strength: NONE

Score-Adjusted method does not validate. Base model performs as well or better.

Overall Win Rate: 14.3% (Score-Adjusted wins)

Mean Improvement: -3.4%

Forecast Quality Scores:

- Base Model: 77.5/100
- Score-Adjusted Model: 77.0/100
- Improvement: -0.5 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 24.52%
- Median MAPE: 22.30%
- Mean sMAPE: 21.15%

Score-Adjusted Model:

- Mean MAPE: 25.17%
- Median MAPE: 23.21%
- Mean sMAPE: 21.64%

Win/Loss Record:

- Score-Adjusted Wins: 1
- Base Model Wins: 6
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 32.31%
- Score Mean MAPE: 33.27%
- Mean Improvement: -3.39%
- Score Win Rate: 0.0%
- Wins: Score=0, Base=4

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 14.13%
- Score Mean MAPE: 14.36%
- Mean Improvement: -3.42%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

STATISTICAL SIGNIFICANCE TESTS

Paired T-Test:

- T-Statistic: -2.0732
- P-Value: 0.0835
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 7.0000
- P-Value: 0.2969
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: -0.7836
- Interpretation: Medium effect

METRIC-BY-METRIC BREAKDOWN

Accounts_Receivable (Tier 2):

Actual Value: 842,800,000.00

Base Forecast: 1,030,776,530.00 (Error: +187,976,530.00)

Score Forecast: 1,038,418,887.00 (Error: +195,618,887.00)

Accuracy:

- Base sMAPE: 20.07%
- Score sMAPE: 20.80%
- Improvement: -4.07%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 499,900,000.00

Base Forecast: 680,463,920.00 (Error: +180,563,920.00)

Score Forecast: 686,362,394.00 (Error: +186,462,394.00)

Accuracy:

- Base sMAPE: 30.59%
- Score sMAPE: 31.44%
- Improvement: -3.27%

Direction Correct: Base=True, Score=True

Winner: Base

Gross_Profit (Tier 2):

Actual Value: 2,191,000,000.00

Base Forecast: 2,323,245,859.00 (Error: +132,245,859.00)

Score Forecast: 2,341,162,069.00 (Error: +150,162,069.00)

Accuracy:

- Base sMAPE: 5.86%
- Score sMAPE: 6.63%
- Improvement: -13.55%

Direction Correct: Base=True, Score=True

Winner: Base

Net_Working_Capital (Tier 1):

Actual Value: 1,535,600,000.00

Base Forecast: 2,291,855,429.00 (Error: +756,255,429.00)

Score Forecast: 2,308,253,309.00 (Error: +772,653,309.00)

Accuracy:

- Base sMAPE: 39.52%
- Score sMAPE: 40.20%
- Improvement: -2.17%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 609,400,000.00

Base Forecast: 798,725,042.00 (Error: +189,325,042.00)

Score Forecast: 804,564,795.00 (Error: +195,164,795.00)

Accuracy:

- Base sMAPE: 26.89%
- Score sMAPE: 27.61%
- Improvement: -3.08%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Income_Expenses (Tier 2):

Actual Value: -1,548,500,000.00

Base Forecast: -1,766,138,731.00 (Error: -217,638,731.00)

Score Forecast: -1,750,127,020.00 (Error: -201,627,020.00)

Accuracy:

- Base sMAPE: 13.13%
- Score sMAPE: 12.22%
- Improvement: +7.36%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 6,033,800,000.00

Base Forecast: 6,805,304,730.00 (Error: +771,504,730.00)

Score Forecast: 6,844,127,061.00 (Error: +810,327,061.00)

Accuracy:

- Base sMAPE: 12.02%

- Score sMAPE: 12.58%
- Improvement: -5.03%

Direction Correct: Base=True, Score=True

Winner: Base

=====

=====

RECOMMENDATIONS

=====

=====

✗ No validation. Base model performs as well or better. Recommend:

1. Critically review the qualitative scoring framework
2. Check if qualitative factors are already reflected in historical data
3. Consider that market efficiency may limit qualitative alpha
4. Explore if timing or sector-specific factors affected results

=====

=====

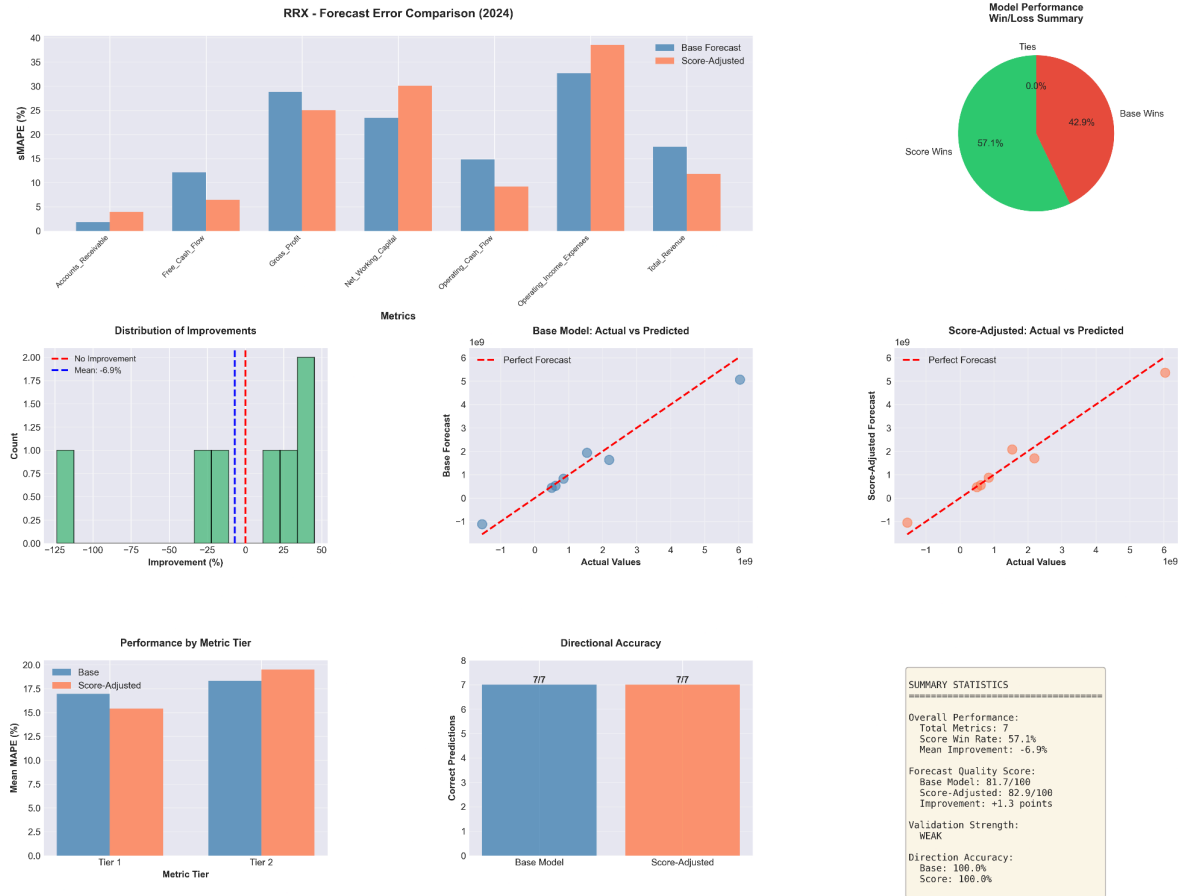
END OF REPORT

=====

=====

RRX (LINEAR):

Forecast Validation Analysis: RRX (2024)



FORECAST VALIDATION REPORT

Company: RRX
Forecast Year: 2024
Analysis Date: 2025-12-12 17:48:59

EXECUTIVE SUMMARY

Validation Strength: WEAK

Score-Adjusted method shows weak validation. Qualitative factors provide marginal improvements.

Overall Win Rate: 57.1% (Score-Adjusted wins)

Mean Improvement: -6.9%

Forecast Quality Scores:

- Base Model: 81.7/100
- Score-Adjusted Model: 82.9/100
- Improvement: +1.3 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 17.55%
- Median MAPE: 16.04%
- Mean sMAPE: 18.72%

Score-Adjusted Model:

- Mean MAPE: 17.15%
- Median MAPE: 11.14%
- Mean sMAPE: 17.84%

Win/Loss Record:

- Score-Adjusted Wins: 4
- Base Model Wins: 3
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 16.95%
- Score Mean MAPE: 15.39%
- Mean Improvement: 19.75%
- Score Win Rate: 75.0%
- Wins: Score=3, Base=1

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 18.34%
- Score Mean MAPE: 19.50%
- Mean Improvement: -42.35%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 0.4159
- P-Value: 0.6919
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 14.0000
- P-Value: 1.0000
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.1572
- Interpretation: Small effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 842,800,000.00

Base Forecast: 827,864,305.00 (Error: -14,935,695.00)

Score Forecast: 876,208,269.00 (Error: +33,408,269.00)

Accuracy:

- Base sMAPE: 1.79%
- Score sMAPE: 3.89%
- Improvement: -123.68%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 499,900,000.00

Base Forecast: 442,821,388.00 (Error: -57,078,612.00)

Score Forecast: 468,680,386.00 (Error: -31,219,614.00)

Accuracy:

- Base sMAPE: 12.11%
- Score sMAPE: 6.45%
- Improvement: +45.30%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 2,191,000,000.00

Base Forecast: 1,640,027,222.00 (Error: -550,972,778.00)

Score Forecast: 1,703,903,002.00 (Error: -487,096,998.00)

Accuracy:

- Base sMAPE: 28.76%
- Score sMAPE: 25.01%
- Improvement: +11.59%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 1,535,600,000.00
Base Forecast: 1,943,451,944.00 (Error: +407,851,944.00)
Score Forecast: 2,079,493,580.00 (Error: +543,893,580.00)

Accuracy:

- Base sMAPE: 23.45%
- Score sMAPE: 30.09%
- Improvement: -33.36%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 609,400,000.00
Base Forecast: 525,305,138.00 (Error: -84,094,862.00)
Score Forecast: 555,980,857.00 (Error: -53,419,143.00)

Accuracy:

- Base sMAPE: 14.82%
- Score sMAPE: 9.17%
- Improvement: +36.48%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -1,548,500,000.00
Base Forecast: -1,113,518,750.00 (Error: +434,981,250.00)
Score Forecast: -1,048,493,709.00 (Error: +500,006,291.00)

Accuracy:

- Base sMAPE: 32.68%
- Score sMAPE: 38.51%
- Improvement: -14.95%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 6,033,800,000.00
Base Forecast: 5,065,948,888.00 (Error: -967,851,112.00)
Score Forecast: 5,361,780,040.00 (Error: -672,019,960.00)

Accuracy:

- Base sMAPE: 17.44%
- Score sMAPE: 11.79%
- Improvement: +30.57%

Direction Correct: Base=True, Score=True

Winner: Score

```
=====
=====
RECOMMENDATIONS
=====
=====
```

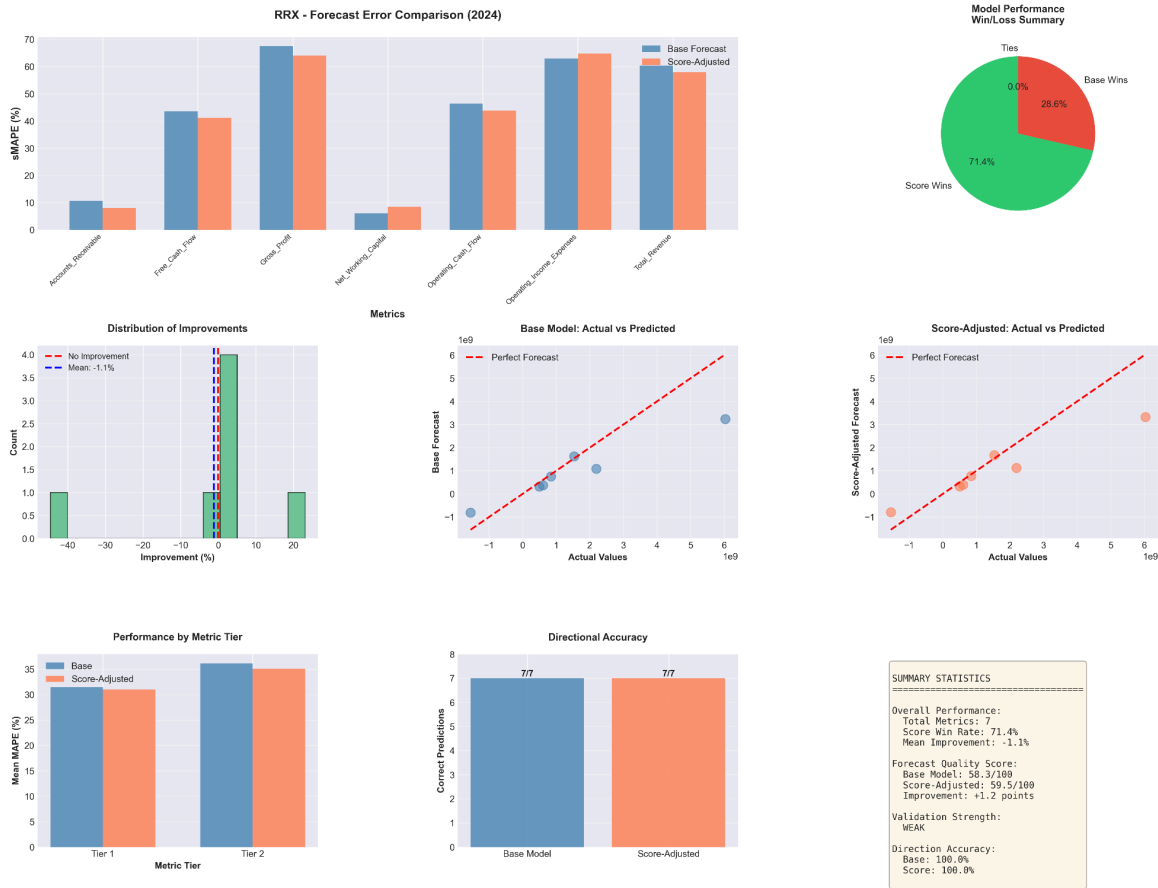
△ Weak validation. Qualitative adjustments show marginal impact. Recommend:

1. Review and refine the qualitative scoring methodology
2. Investigate if certain metric types benefit more from adjustments
3. Consider whether the base model is already capturing qualitative factors
4. Test alternative adjustment algorithms

```
=====
=====
END OF REPORT
=====
=====
```


RRX (XGBOOST):

Forecast Validation Analysis: RRX (2024)



FORECAST VALIDATION REPORT

Company: RRX
Forecast Year: 2024
Analysis Date: 2025-12-12 17:49:20

EXECUTIVE SUMMARY

=====

Validation Strength: WEAK

Score-Adjusted method shows weak validation. Qualitative factors provide marginal improvements.

Overall Win Rate: 71.4% (Score-Adjusted wins)

Mean Improvement: -1.1%

Forecast Quality Scores:

- Base Model: 58.3/100
- Score-Adjusted Model: 59.5/100
- Improvement: +1.2 points

=====

DETAILED METRICS COMPARISON

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 33.49%
- Median MAPE: 37.64%
- Mean sMAPE: 42.48%

Score-Adjusted Model:

- Mean MAPE: 32.76%
- Median MAPE: 36.02%
- Mean sMAPE: 41.23%

Win/Loss Record:

- Score-Adjusted Wins: 5
- Base Model Wins: 2
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 31.49%
- Score Mean MAPE: 31.01%
- Mean Improvement: -8.15%
- Score Win Rate: 75.0%
- Wins: Score=3, Base=1

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 36.16%
- Score Mean MAPE: 35.10%
- Mean Improvement: 8.20%
- Score Win Rate: 66.7%
- Wins: Score=2, Base=1

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 1.3731
- P-Value: 0.2188
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 7.0000
- P-Value: 0.2969
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.5190
- Interpretation: Medium effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 842,800,000.00

Base Forecast: 757,367,232.00 (Error: -85,432,768.00)

Score Forecast: 777,058,780.00 (Error: -65,741,220.00)

Accuracy:

- Base sMAPE: 10.68%

- Score sMAPE: 8.12%

- Improvement: +23.05%

Direction Correct: Base=True, Score=True

Winner: Score

Free_Cash_Flow (Tier 1):

Actual Value: 499,900,000.00

Base Forecast: 321,015,808.00 (Error: -178,884,192.00)

Score Forecast: 329,362,219.00 (Error: -170,537,781.00)

Accuracy:

- Base sMAPE: 43.58%

- Score sMAPE: 41.13%

- Improvement: +4.67%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 2,191,000,000.00

Base Forecast: 1,085,332,500.00 (Error: -1,105,667,500.00)

Score Forecast: 1,127,660,467.00 (Error: -1,063,339,533.00)

Accuracy:

- Base sMAPE: 67.49%

- Score sMAPE: 64.08%

- Improvement: +3.83%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 1,535,600,000.00

Base Forecast: 1,630,729,856.00 (Error: +95,129,856.00)

Score Forecast: 1,673,128,832.00 (Error: +137,528,832.00)

Accuracy:

- Base sMAPE: 6.01%
- Score sMAPE: 8.57%
- Improvement: -44.57%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 609,400,000.00

Base Forecast: 380,033,056.00 (Error: -229,366,944.00)

Score Forecast: 389,913,915.00 (Error: -219,486,085.00)

Accuracy:

- Base sMAPE: 46.36%
- Score sMAPE: 43.93%
- Improvement: +4.31%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -1,548,500,000.00

Base Forecast: -807,135,000.00 (Error: +741,365,000.00)

Score Forecast: -790,346,592.00 (Error: +758,153,408.00)

Accuracy:

- Base sMAPE: 62.94%
- Score sMAPE: 64.83%
- Improvement: -2.26%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 6,033,800,000.00

Base Forecast: 3,238,000,384.00 (Error: -2,795,799,616.00)

Score Forecast: 3,322,188,393.00 (Error: -2,711,611,607.00)

Accuracy:

- Base sMAPE: 60.31%
- Score sMAPE: 57.97%
- Improvement: +3.01%

Direction Correct: Base=True, Score=True

Winner: Score

```
=====
=====
RECOMMENDATIONS
=====
=====
```

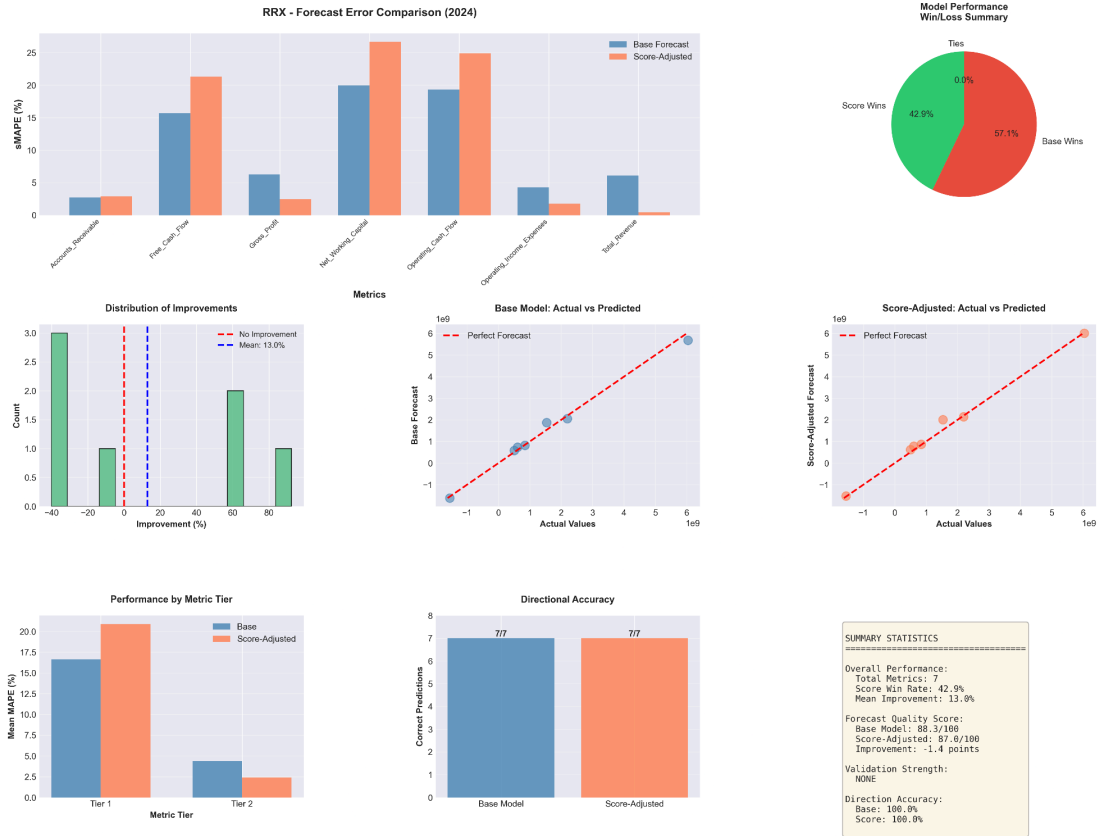
△ Weak validation. Qualitative adjustments show marginal impact. Recommend:

1. Review and refine the qualitative scoring methodology
2. Investigate if certain metric types benefit more from adjustments
3. Consider whether the base model is already capturing qualitative factors
4. Test alternative adjustment algorithms

```
=====
=====
END OF REPORT
=====
=====
```


RRX (ARIMA):

Forecast Validation Analysis: RRX (2024)



FORECAST VALIDATION REPORT

Company: RRX
Forecast Year: 2024
Analysis Date: 2025-12-12 17:49:39

EXECUTIVE SUMMARY

Validation Strength: NONE

Score-Adjusted method does not validate. Base model performs as well or better.

Overall Win Rate: 42.9% (Score-Adjusted wins)

Mean Improvement: 13.0%

Forecast Quality Scores:

- Base Model: 88.3/100
- Score-Adjusted Model: 87.0/100
- Improvement: -1.4 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 11.38%
- Median MAPE: 6.10%
- Mean sMAPE: 10.63%

Score-Adjusted Model:

- Mean MAPE: 12.95%
- Median MAPE: 2.97%
- Mean sMAPE: 11.50%

Win/Loss Record:

- Score-Adjusted Wins: 3
- Base Model Wins: 4
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 16.63%
- Score Mean MAPE: 20.88%
- Mean Improvement: -4.83%
- Score Win Rate: 25.0%
- Wins: Score=1, Base=3

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 4.38%
- Score Mean MAPE: 2.39%
- Mean Improvement: 36.66%
- Score Win Rate: 66.7%
- Wins: Score=2, Base=1

=====

=====

STATISTICAL SIGNIFICANCE TESTS

=====

=====

Paired T-Test:

- T-Statistic: -0.4537
- P-Value: 0.6660
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 11.0000
- P-Value: 0.6875
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: -0.1715
- Interpretation: Small effect

=====

=====

METRIC-BY-METRIC BREAKDOWN

=====

=====

Accounts_Receivable (Tier 2):

Actual Value: 842,800,000.00

Base Forecast: 819,972,246.00 (Error: -22,827,754.00)
Score Forecast: 867,855,346.00 (Error: +25,055,346.00)

Accuracy:

- Base sMAPE: 2.75%
- Score sMAPE: 2.93%
- Improvement: -9.76%

Direction Correct: Base=True, Score=True
Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 499,900,000.00
Base Forecast: 584,946,031.00 (Error: +85,046,031.00)
Score Forecast: 619,104,540.00 (Error: +119,204,540.00)

Accuracy:

- Base sMAPE: 15.68%
- Score sMAPE: 21.31%
- Improvement: -40.16%

Direction Correct: Base=True, Score=True
Winner: Base

Gross_Profit (Tier 2):

Actual Value: 2,191,000,000.00
Base Forecast: 2,057,381,746.00 (Error: -133,618,254.00)
Score Forecast: 2,137,512,650.00 (Error: -53,487,350.00)

Accuracy:

- Base sMAPE: 6.29%
- Score sMAPE: 2.47%
- Improvement: +59.97%

Direction Correct: Base=True, Score=True
Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 1,535,600,000.00
Base Forecast: 1,876,588,666.00 (Error: +340,988,666.00)
Score Forecast: 2,007,949,873.00 (Error: +472,349,873.00)

Accuracy:

- Base sMAPE: 19.99%
- Score sMAPE: 26.66%
- Improvement: -38.52%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 609,400,000.00

Base Forecast: 739,641,130.00 (Error: +130,241,130.00)

Score Forecast: 782,833,213.00 (Error: +173,433,213.00)

Accuracy:

- Base sMAPE: 19.31%
- Score sMAPE: 24.91%
- Improvement: -33.16%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Income_Expenses (Tier 2):

Actual Value: -1,548,500,000.00

Base Forecast: -1,615,787,979.00 (Error: -67,287,979.00)

Score Forecast: -1,521,432,424.00 (Error: +27,067,576.00)

Accuracy:

- Base sMAPE: 4.25%
- Score sMAPE: 1.76%
- Improvement: +59.77%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 6,033,800,000.00

Base Forecast: 5,675,568,274.00 (Error: -358,231,726.00)

Score Forecast: 6,006,998,759.00 (Error: -26,801,241.00)

Accuracy:

- Base sMAPE: 6.12%
- Score sMAPE: 0.45%

- Improvement: +92.52%

Direction Correct: Base=True, Score=True

Winner: Score

=====

=====

RECOMMENDATIONS

=====

=====

X No validation. Base model performs as well or better. Recommend:

1. Critically review the qualitative scoring framework
2. Check if qualitative factors are already reflected in historical data
3. Consider that market efficiency may limit qualitative alpha
4. Explore if timing or sector-specific factors affected results

=====

=====

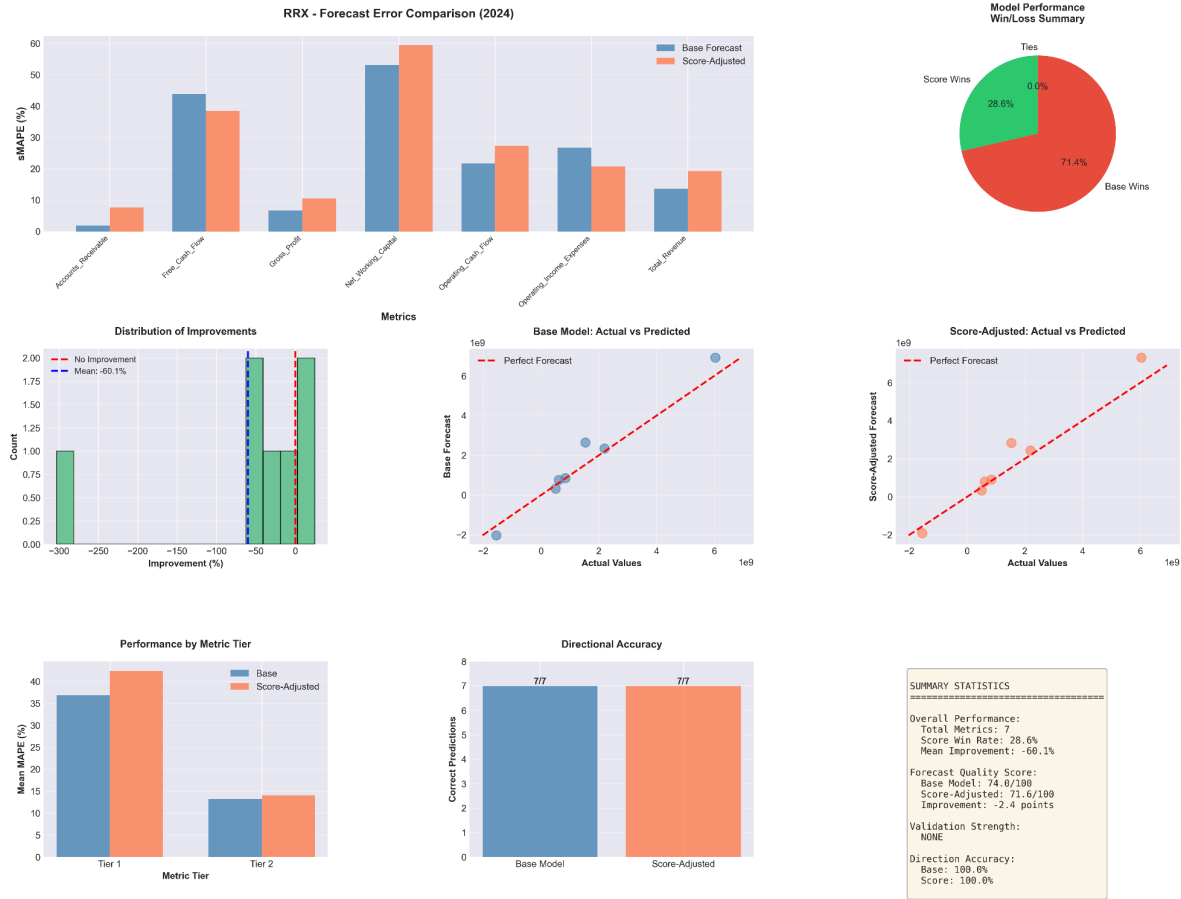
END OF REPORT

=====

=====

RRX (WINTERS-HOLT):

Forecast Validation Analysis: RRX (2024)



FORECAST VALIDATION REPORT

Company: RRX
Forecast Year: 2024
Analysis Date: 2025-12-12 17:49:55

EXECUTIVE SUMMARY

Validation Strength: NONE

Score-Adjusted method does not validate. Base model performs as well or better.

Overall Win Rate: 28.6% (Score-Adjusted wins)

Mean Improvement: -60.1%

Forecast Quality Scores:

- Base Model: 74.0/100
- Score-Adjusted Model: 71.6/100
- Improvement: -2.4 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 26.70%
- Median MAPE: 24.38%
- Mean sMAPE: 23.94%

Score-Adjusted Model:

- Mean MAPE: 30.24%
- Median MAPE: 23.09%
- Mean sMAPE: 26.16%

Win/Loss Record:

- Score-Adjusted Wins: 2
- Base Model Wins: 5
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 36.83%
- Score Mean MAPE: 42.40%
- Mean Improvement: -20.45%
- Score Win Rate: 25.0%
- Wins: Score=1, Base=3

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 13.19%
- Score Mean MAPE: 14.02%
- Mean Improvement: -112.88%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

STATISTICAL SIGNIFICANCE TESTS

Paired T-Test:

- T-Statistic: -1.0788
- P-Value: 0.3221
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 8.0000
- P-Value: 0.3750
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: -0.4078
- Interpretation: Small effect

METRIC-BY-METRIC BREAKDOWN

Accounts_Receivable (Tier 2):

Actual Value: 842,800,000.00

Base Forecast: 859,361,026.00 (Error: +16,561,026.00)

Score Forecast: 909,544,272.00 (Error: +66,744,272.00)

Accuracy:

- Base sMAPE: 1.95%
- Score sMAPE: 7.62%
- Improvement: -303.02%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 499,900,000.00

Base Forecast: 320,267,268.00 (Error: -179,632,732.00)

Score Forecast: 338,969,596.00 (Error: -160,930,404.00)

Accuracy:

- Base sMAPE: 43.80%
- Score sMAPE: 38.37%
- Improvement: +10.41%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 2,191,000,000.00

Base Forecast: 2,341,865,043.00 (Error: +150,865,043.00)

Score Forecast: 2,433,076,003.00 (Error: +242,076,003.00)

Accuracy:

- Base sMAPE: 6.66%
- Score sMAPE: 10.47%
- Improvement: -60.46%

Direction Correct: Base=True, Score=True

Winner: Base

Net_Working_Capital (Tier 1):

Actual Value: 1,535,600,000.00

Base Forecast: 2,647,002,928.00 (Error: +1,111,402,928.00)

Score Forecast: 2,832,293,133.00 (Error: +1,296,693,133.00)

Accuracy:

- Base sMAPE: 53.14%
- Score sMAPE: 59.37%
- Improvement: -16.67%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 609,400,000.00

Base Forecast: 757,958,898.00 (Error: +148,558,898.00)

Score Forecast: 802,220,665.00 (Error: +192,820,665.00)

Accuracy:

- Base sMAPE: 21.73%
- Score sMAPE: 27.32%
- Improvement: -29.79%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Income_Expenses (Tier 2):

Actual Value: -1,548,500,000.00

Base Forecast: -2,024,214,392.00 (Error: -475,714,392.00)

Score Forecast: -1,906,008,368.00 (Error: -357,508,368.00)

Accuracy:

- Base sMAPE: 26.63%
- Score sMAPE: 20.70%
- Improvement: +24.85%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 6,033,800,000.00

Base Forecast: 6,917,084,601.00 (Error: +883,284,601.00)

Score Forecast: 7,321,014,673.00 (Error: +1,287,214,673.00)

Accuracy:

- Base sMAPE: 13.64%
- Score sMAPE: 19.28%
- Improvement: -45.73%

Direction Correct: Base=True, Score=True
Winner: Base

=====

=====

RECOMMENDATIONS

=====

=====

- X No validation. Base model performs as well or better. Recommend:
1. Critically review the qualitative scoring framework
 2. Check if qualitative factors are already reflected in historical data
 3. Consider that market efficiency may limit qualitative alpha
 4. Explore if timing or sector-specific factors affected results

=====

=====

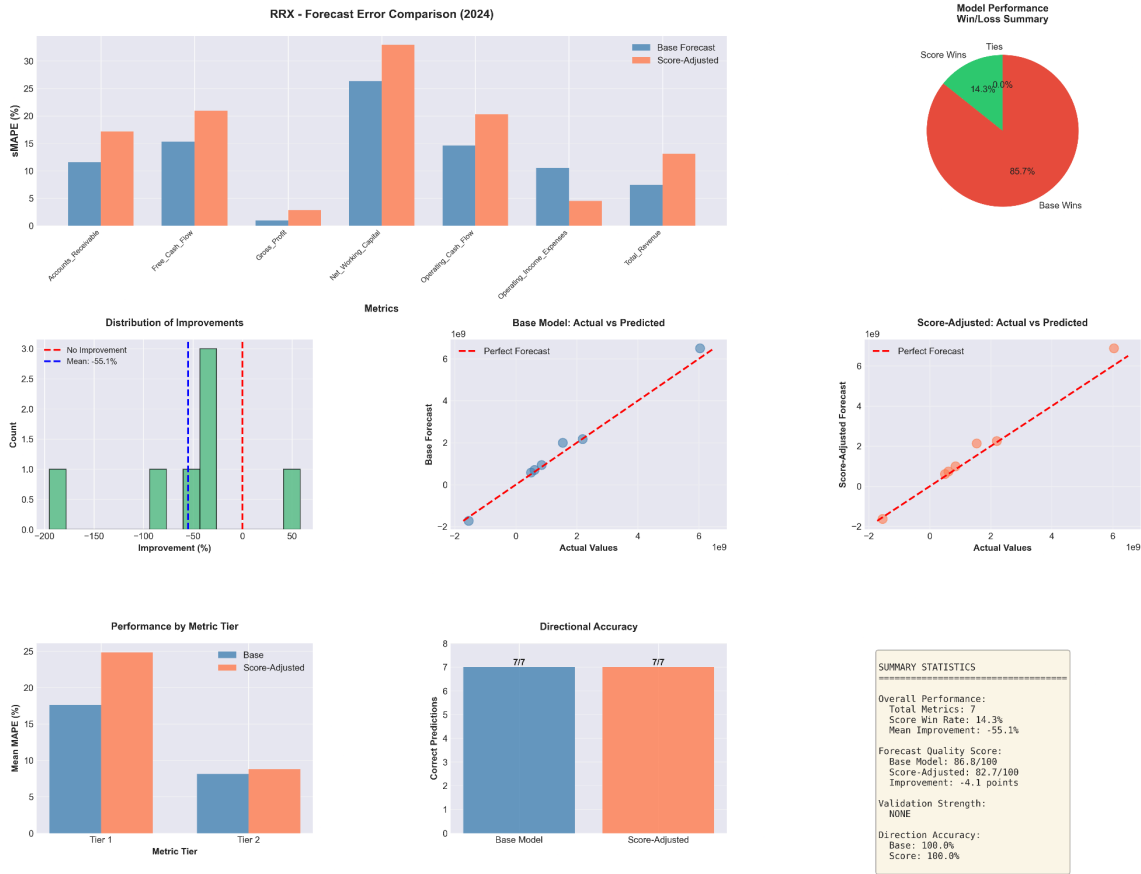
END OF REPORT

=====

=====

RRX (LINEAR + WINTERS-HOLT):

Forecast Validation Analysis: RRX (2024)



FORECAST VALIDATION REPORT

Company: RRX
Forecast Year: 2024
Analysis Date: 2025-12-12 17:50:18

EXECUTIVE SUMMARY

Validation Strength: NONE

Score-Adjusted method does not validate. Base model performs as well or better.

Overall Win Rate: 14.3% (Score-Adjusted wins)

Mean Improvement: -55.1%

Forecast Quality Scores:

- Base Model: 86.8/100
- Score-Adjusted Model: 82.7/100
- Improvement: -4.1 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 13.53%
- Median MAPE: 12.24%
- Mean sMAPE: 12.39%

Score-Adjusted Model:

- Mean MAPE: 17.96%
- Median MAPE: 18.80%
- Mean sMAPE: 15.97%

Win/Loss Record:

- Score-Adjusted Wins: 1
- Base Model Wins: 6
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 17.60%
- Score Mean MAPE: 24.84%
- Mean Improvement: -48.80%
- Score Win Rate: 0.0%
- Wins: Score=0, Base=4

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 8.12%
- Score Mean MAPE: 8.77%
- Mean Improvement: -63.43%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

STATISTICAL SIGNIFICANCE TESTS

Paired T-Test:

- T-Statistic: -2.1060
- P-Value: 0.0798
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 6.0000
- P-Value: 0.2188
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: -0.7960
- Interpretation: Medium effect

METRIC-BY-METRIC BREAKDOWN

Accounts_Receivable (Tier 2):

Actual Value: 842,800,000.00

Base Forecast: 945,980,998.00 (Error: +103,180,998.00)

Score Forecast: 1,001,222,504.00 (Error: +158,422,504.00)

Accuracy:

- Base sMAPE: 11.54%
- Score sMAPE: 17.18%
- Improvement: -53.54%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 499,900,000.00

Base Forecast: 582,731,514.00 (Error: +82,831,514.00)

Score Forecast: 616,760,703.00 (Error: +116,860,703.00)

Accuracy:

- Base sMAPE: 15.30%
- Score sMAPE: 20.93%
- Improvement: -41.08%

Direction Correct: Base=True, Score=True

Winner: Base

Gross_Profit (Tier 2):

Actual Value: 2,191,000,000.00

Base Forecast: 2,169,610,437.00 (Error: -21,389,563.00)

Score Forecast: 2,254,112,424.00 (Error: +63,112,424.00)

Accuracy:

- Base sMAPE: 0.98%
- Score sMAPE: 2.84%
- Improvement: -195.06%

Direction Correct: Base=True, Score=True

Winner: Base

Net_Working_Capital (Tier 1):

Actual Value: 1,535,600,000.00

Base Forecast: 2,000,753,790.00 (Error: +465,153,790.00)

Score Forecast: 2,140,806,555.00 (Error: +605,206,555.00)

Accuracy:

- Base sMAPE: 26.31%
- Score sMAPE: 32.92%
- Improvement: -30.11%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 609,400,000.00

Base Forecast: 705,532,844.00 (Error: +96,132,844.00)

Score Forecast: 746,733,140.00 (Error: +137,333,140.00)

Accuracy:

- Base sMAPE: 14.62%
- Score sMAPE: 20.25%
- Improvement: -42.86%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Income_Expenses (Tier 2):

Actual Value: -1,548,500,000.00

Base Forecast: -1,720,878,104.00 (Error: -172,378,104.00)

Score Forecast: -1,620,385,706.00 (Error: -71,885,706.00)

Accuracy:

- Base sMAPE: 10.55%
- Score sMAPE: 4.54%
- Improvement: +58.30%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 6,033,800,000.00

Base Forecast: 6,501,716,122.00 (Error: +467,916,122.00)

Score Forecast: 6,881,390,336.00 (Error: +847,590,336.00)

Accuracy:

- Base sMAPE: 7.47%

- Score sMAPE: 13.13%
- Improvement: -81.14%

Direction Correct: Base=True, Score=True

Winner: Base

=====

=====

RECOMMENDATIONS

=====

=====

✗ No validation. Base model performs as well or better. Recommend:

1. Critically review the qualitative scoring framework
2. Check if qualitative factors are already reflected in historical data
3. Consider that market efficiency may limit qualitative alpha
4. Explore if timing or sector-specific factors affected results

=====

=====

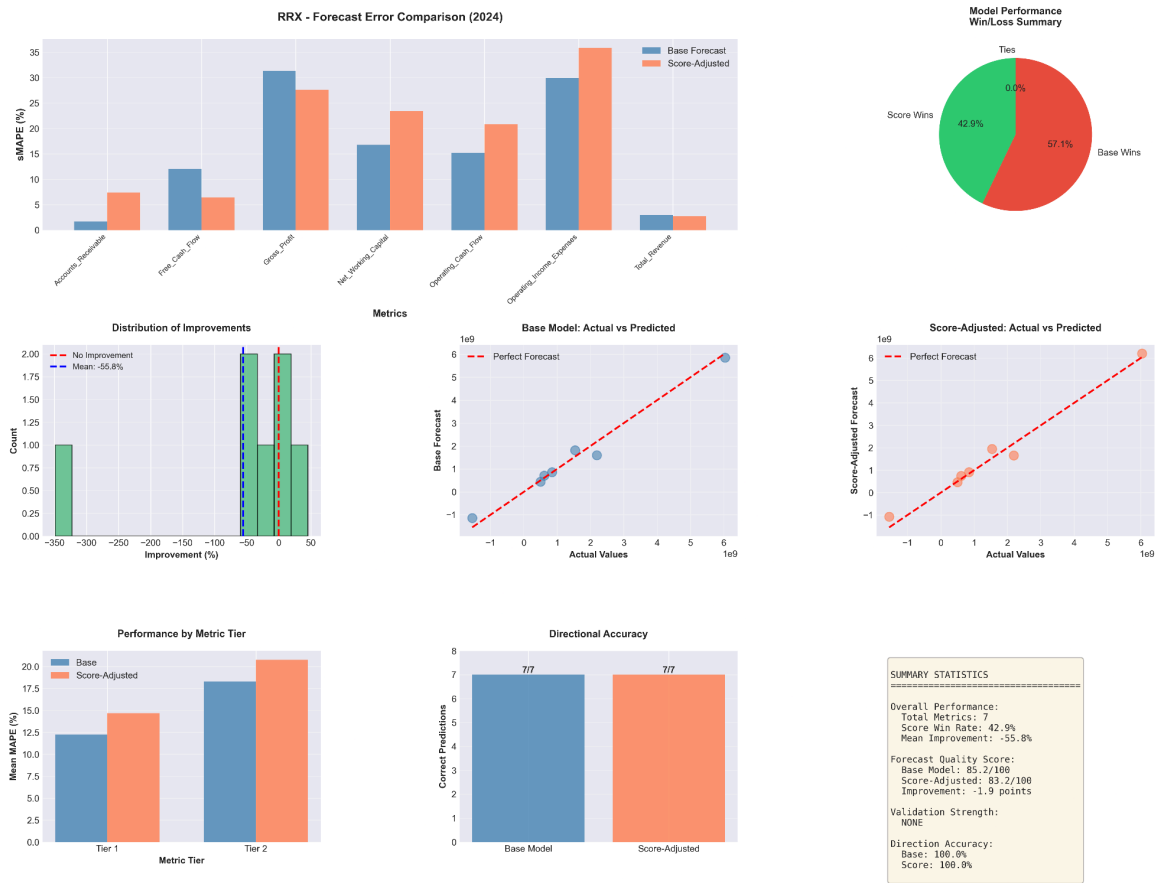
END OF REPORT

=====

=====

RRX (LINEAR + WINTERS-HOLT + XGBOOST):

Forecast Validation Analysis: RRX (2024)



FORECAST VALIDATION REPORT

Company: RRX
Forecast Year: 2024
Analysis Date: 2025-12-12 17:50:41

EXECUTIVE SUMMARY

=====

Validation Strength: NONE

Score-Adjusted method does not validate. Base model performs as well or better.

Overall Win Rate: 42.9% (Score-Adjusted wins)

Mean Improvement: -55.8%

Forecast Quality Scores:

- Base Model: 85.2/100
- Score-Adjusted Model: 83.2/100
- Improvement: -1.9 points

=====

DETAILED METRICS COMPARISON

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 14.84%
- Median MAPE: 16.45%
- Mean sMAPE: 15.70%

Score-Adjusted Model:

- Mean MAPE: 17.29%
- Median MAPE: 23.25%
- Mean sMAPE: 17.74%

Win/Loss Record:

- Score-Adjusted Wins: 3
- Base Model Wins: 4
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 12.25%
- Score Mean MAPE: 14.69%
- Mean Improvement: -9.01%
- Score Win Rate: 50.0%
- Wins: Score=2, Base=2

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 18.28%
- Score Mean MAPE: 20.75%
- Mean Improvement: -118.29%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: -1.0405
- P-Value: 0.3382
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 7.0000
- P-Value: 0.2969
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: -0.3933
- Interpretation: Small effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 842,800,000.00

Base Forecast: 857,149,827.00 (Error: +14,349,827.00)

Score Forecast: 907,203,948.00 (Error: +64,403,948.00)

Accuracy:

- Base sMAPE: 1.69%
- Score sMAPE: 7.36%
- Improvement: -348.81%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 499,900,000.00

Base Forecast: 443,060,493.00 (Error: -56,839,507.00)

Score Forecast: 468,933,453.00 (Error: -30,966,547.00)

Accuracy:

- Base sMAPE: 12.06%
- Score sMAPE: 6.39%
- Improvement: +45.52%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 2,191,000,000.00

Base Forecast: 1,597,827,554.00 (Error: -593,172,446.00)

Score Forecast: 1,660,059,742.00 (Error: -530,940,258.00)

Accuracy:

- Base sMAPE: 31.31%
- Score sMAPE: 27.57%
- Improvement: +10.49%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 1,535,600,000.00

Base Forecast: 1,816,284,351.00 (Error: +280,684,351.00)

Score Forecast: 1,943,424,255.00 (Error: +407,824,255.00)

Accuracy:

- Base sMAPE: 16.75%
- Score sMAPE: 23.44%
- Improvement: -45.30%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 609,400,000.00

Base Forecast: 709,638,838.00 (Error: +100,238,838.00)

Score Forecast: 751,078,907.00 (Error: +141,678,907.00)

Accuracy:

- Base sMAPE: 15.20%
- Score sMAPE: 20.83%
- Improvement: -41.34%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Income_Expenses (Tier 2):

Actual Value: -1,548,500,000.00

Base Forecast: -1,144,832,556.00 (Error: +403,667,444.00)

Score Forecast: -1,077,978,914.00 (Error: +470,521,086.00)

Accuracy:

- Base sMAPE: 29.98%
- Score sMAPE: 35.83%
- Improvement: -16.56%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 6,033,800,000.00

Base Forecast: 5,858,305,613.00 (Error: -175,494,387.00)

Score Forecast: 6,200,407,227.00 (Error: +166,607,227.00)

Accuracy:

- Base sMAPE: 2.95%
- Score sMAPE: 2.72%
- Improvement: +5.06%

Direction Correct: Base=True, Score=True

Winner: Score

=====

=====

RECOMMENDATIONS

=====

=====

X No validation. Base model performs as well or better. Recommend:

1. Critically review the qualitative scoring framework
2. Check if qualitative factors are already reflected in historical data
3. Consider that market efficiency may limit qualitative alpha
4. Explore if timing or sector-specific factors affected results

=====

=====

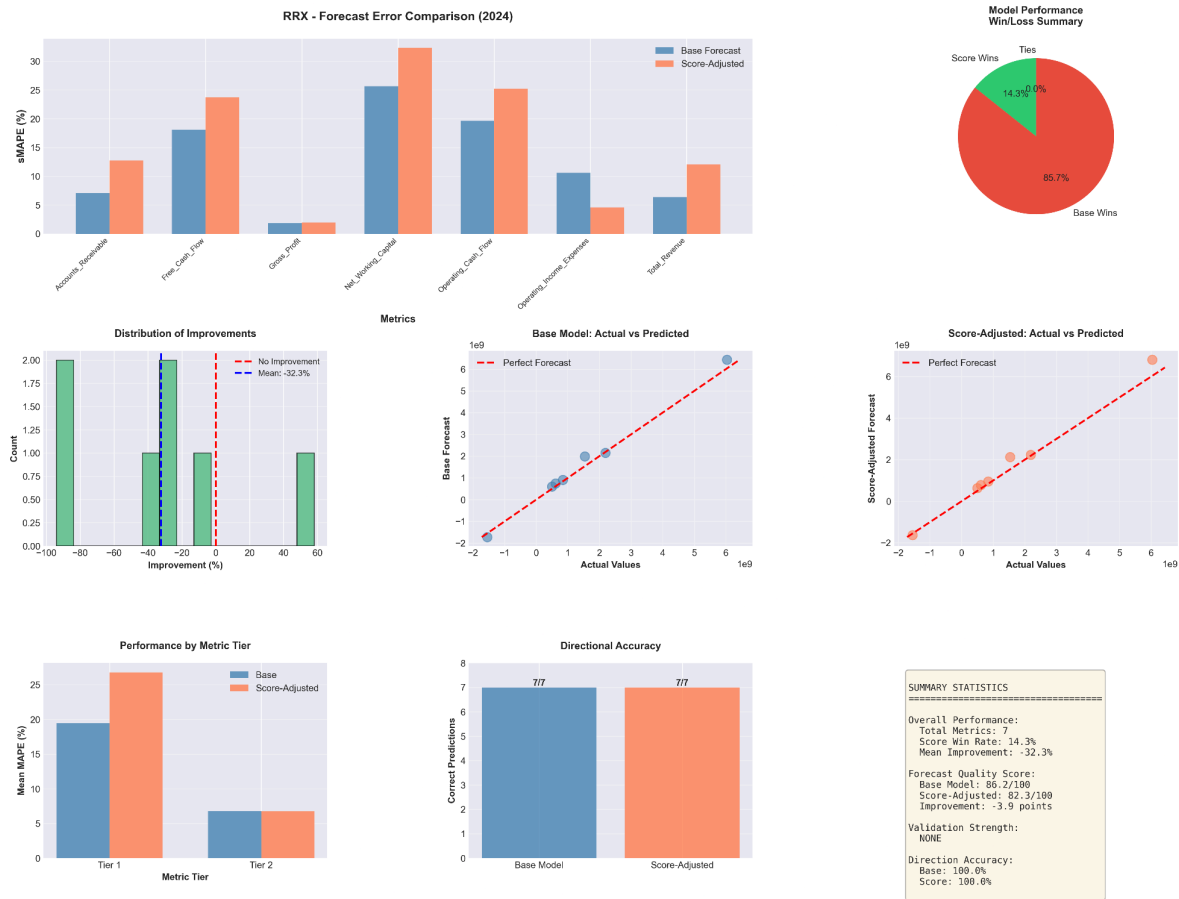
END OF REPORT

=====

=====

RRX (LINEAR + WINTERS-HOLT + ARIMA):

Forecast Validation Analysis: RRX (2024)



FORECAST VALIDATION REPORT

Company: RRX
Forecast Year: 2024
Analysis Date: 2025-12-12 17:51:04

EXECUTIVE SUMMARY

=====

Validation Strength: NONE

Score-Adjusted method does not validate. Base model performs as well or better.

Overall Win Rate: 14.3% (Score-Adjusted wins)

Mean Improvement: -32.3%

Forecast Quality Scores:

- Base Model: 86.2/100
- Score-Adjusted Model: 82.3/100
- Improvement: -3.9 points

=====

DETAILED METRICS COMPARISON

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 14.01%
- Median MAPE: 11.17%
- Mean sMAPE: 12.76%

Score-Adjusted Model:

- Mean MAPE: 18.20%
- Median MAPE: 13.59%
- Mean sMAPE: 16.08%

Win/Loss Record:

- Score-Adjusted Wins: 1
- Base Model Wins: 6
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 19.43%
- Score Mean MAPE: 26.78%
- Mean Improvement: -48.16%
- Score Win Rate: 0.0%
- Wins: Score=0, Base=4

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 6.79%
- Score Mean MAPE: 6.75%
- Mean Improvement: -11.06%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: -1.8865
- P-Value: 0.1082
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 6.0000
- P-Value: 0.2188
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: -0.7130
- Interpretation: Medium effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 842,800,000.00

Base Forecast: 904,554,839.00 (Error: +61,754,839.00)

Score Forecast: 957,377,223.00 (Error: +114,577,223.00)

Accuracy:

- Base sMAPE: 7.07%
- Score sMAPE: 12.73%
- Improvement: -85.54%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 499,900,000.00

Base Forecast: 599,232,089.00 (Error: +99,332,089.00)

Score Forecast: 634,224,846.00 (Error: +134,324,846.00)

Accuracy:

- Base sMAPE: 18.07%
- Score sMAPE: 23.69%
- Improvement: -35.23%

Direction Correct: Base=True, Score=True

Winner: Base

Gross_Profit (Tier 2):

Actual Value: 2,191,000,000.00

Base Forecast: 2,150,297,799.00 (Error: -40,702,201.00)

Score Forecast: 2,234,047,598.00 (Error: +43,047,598.00)

Accuracy:

- Base sMAPE: 1.88%
- Score sMAPE: 1.95%
- Improvement: -5.76%

Direction Correct: Base=True, Score=True

Winner: Base

Net_Working_Capital (Tier 1):

Actual Value: 1,535,600,000.00

Base Forecast: 1,988,055,140.00 (Error: +452,455,140.00)

Score Forecast: 2,127,219,000.00 (Error: +591,619,000.00)

Accuracy:

- Base sMAPE: 25.68%
- Score sMAPE: 32.30%
- Improvement: -30.76%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 609,400,000.00

Base Forecast: 742,051,227.00 (Error: +132,651,227.00)

Score Forecast: 785,384,051.00 (Error: +175,984,051.00)

Accuracy:

- Base sMAPE: 19.63%
- Score sMAPE: 25.23%
- Improvement: -32.67%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Income_Expenses (Tier 2):

Actual Value: -1,548,500,000.00

Base Forecast: -1,721,509,055.00 (Error: -173,009,055.00)

Score Forecast: -1,620,979,812.00 (Error: -72,479,812.00)

Accuracy:

- Base sMAPE: 10.58%
- Score sMAPE: 4.57%
- Improvement: +58.11%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 6,033,800,000.00

Base Forecast: 6,433,463,843.00 (Error: +399,663,843.00)

Score Forecast: 6,809,152,398.00 (Error: +775,352,398.00)

Accuracy:

- Base sMAPE: 6.41%
- Score sMAPE: 12.07%
- Improvement: -94.00%

Direction Correct: Base=True, Score=True

Winner: Base

=====

=====

RECOMMENDATIONS

=====

=====

X No validation. Base model performs as well or better. Recommend:

1. Critically review the qualitative scoring framework
2. Check if qualitative factors are already reflected in historical data
3. Consider that market efficiency may limit qualitative alpha
4. Explore if timing or sector-specific factors affected results

=====

=====

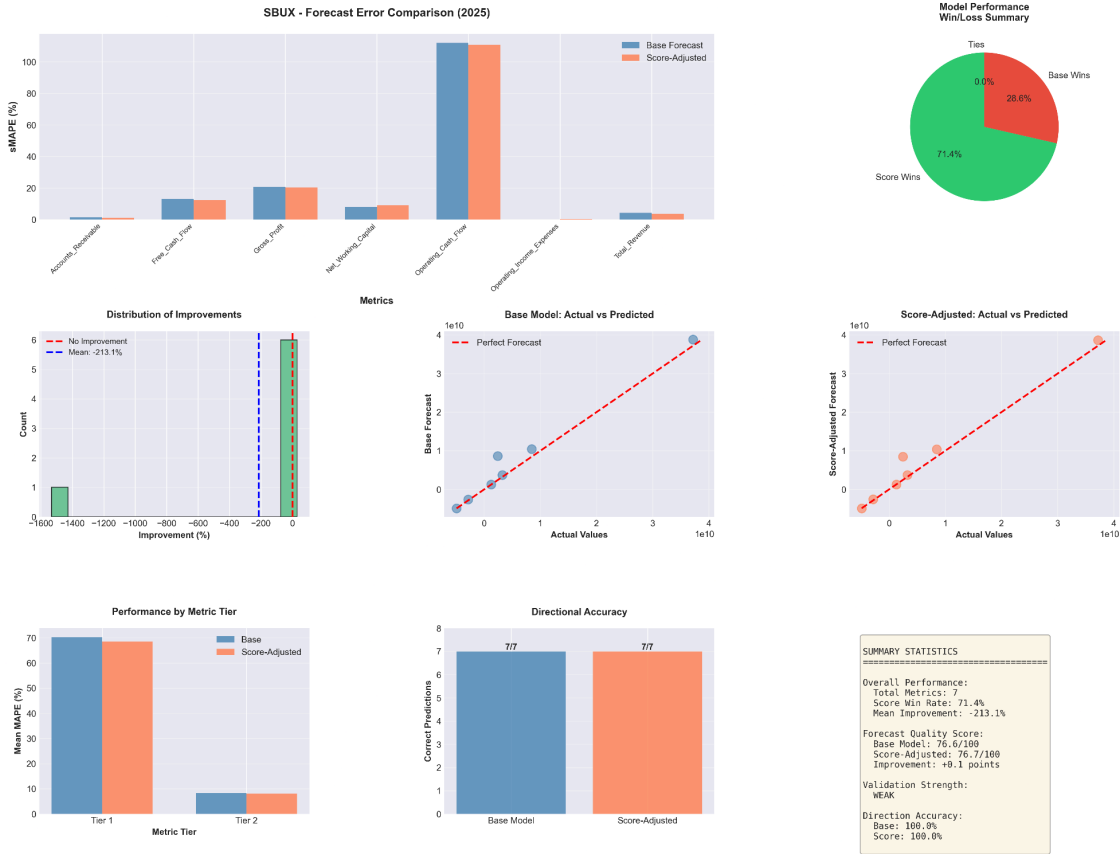
END OF REPORT

=====

=====

SBUX (NAIVE):

Forecast Validation Analysis: SBUX (2025)



FORECAST VALIDATION REPORT

Company: SBUX
Forecast Year: 2025
Analysis Date: 2025-12-12 14:34:04

EXECUTIVE SUMMARY

Validation Strength: WEAK

Score-Adjusted method shows weak validation. Qualitative factors provide marginal improvements.

Overall Win Rate: 71.4% (Score-Adjusted wins)

Mean Improvement: -213.1%

Forecast Quality Scores:

- Base Model: 76.6/100
- Score-Adjusted Model: 76.7/100
- Improvement: +0.1 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 43.71%
- Median MAPE: 7.76%
- Mean sMAPE: 22.85%

Score-Adjusted Model:

- Mean MAPE: 42.59%
- Median MAPE: 8.72%
- Mean sMAPE: 22.55%

Win/Loss Record:

- Score-Adjusted Wins: 5
- Base Model Wins: 2
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 70.28%
- Score Mean MAPE: 68.49%
- Mean Improvement: 1.98%
- Score Win Rate: 75.0%
- Wins: Score=3, Base=1

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 8.28%
- Score Mean MAPE: 8.05%
- Mean Improvement: -499.99%
- Score Win Rate: 66.7%
- Wins: Score=2, Base=1

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 1.0294
- P-Value: 0.3430
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 7.0000
- P-Value: 0.2969
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.3891
- Interpretation: Small effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 1,277,500,000.00

Base Forecast: 1,297,057,027.00 (Error: +19,557,027.00)

Score Forecast: 1,291,229,035.00 (Error: +13,729,035.00)

Accuracy:

- Base sMAPE: 1.52%
- Score sMAPE: 1.07%
- Improvement: +29.80%

Direction Correct: Base=True, Score=True

Winner: Score

Free_Cash_Flow (Tier 1):

Actual Value: 3,268,000,000.00

Base Forecast: 3,728,512,483.00 (Error: +460,512,483.00)

Score Forecast: 3,699,783,609.00 (Error: +431,783,609.00)

Accuracy:

- Base sMAPE: 13.16%
- Score sMAPE: 12.39%
- Improvement: +6.24%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 8,467,300,000.00

Base Forecast: 10,439,865,701.00 (Error: +1,972,565,701.00)

Score Forecast: 10,388,712,102.00 (Error: +1,921,412,102.00)

Accuracy:

- Base sMAPE: 20.87%
- Score sMAPE: 20.38%
- Improvement: +2.59%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: -2,828,100,000.00

Base Forecast: -2,608,599,839.00 (Error: +219,500,161.00)

Score Forecast: -2,581,579,850.00 (Error: +246,520,150.00)

Accuracy:

- Base sMAPE: 8.07%
- Score sMAPE: 9.11%
- Improvement: -12.31%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 2,442,000,000.00

Base Forecast: 8,667,227,045.00 (Error: +6,225,227,045.00)

Score Forecast: 8,502,938,434.00 (Error: +6,060,938,434.00)

Accuracy:

- Base sMAPE: 112.07%
- Score sMAPE: 110.75%
- Improvement: +2.64%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -4,886,500,000.00

Base Forecast: -4,885,353,022.00 (Error: +1,146,978.00)

Score Forecast: -4,905,222,734.00 (Error: -18,722,734.00)

Accuracy:

- Base sMAPE: 0.02%
- Score sMAPE: 0.38%
- Improvement: -1532.35%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 37,184,400,000.00

Base Forecast: 38,801,336,497.00 (Error: +1,616,936,497.00)

Score Forecast: 38,617,576,942.00 (Error: +1,433,176,942.00)

Accuracy:

- Base sMAPE: 4.26%
- Score sMAPE: 3.78%
- Improvement: +11.36%

Direction Correct: Base=True, Score=True

Winner: Score

=====

=====

RECOMMENDATIONS

=====

=====

△ Weak validation. Qualitative adjustments show marginal impact. Recommend:

1. Review and refine the qualitative scoring methodology
2. Investigate if certain metric types benefit more from adjustments
3. Consider whether the base model is already capturing qualitative factors
4. Test alternative adjustment algorithms

=====

=====

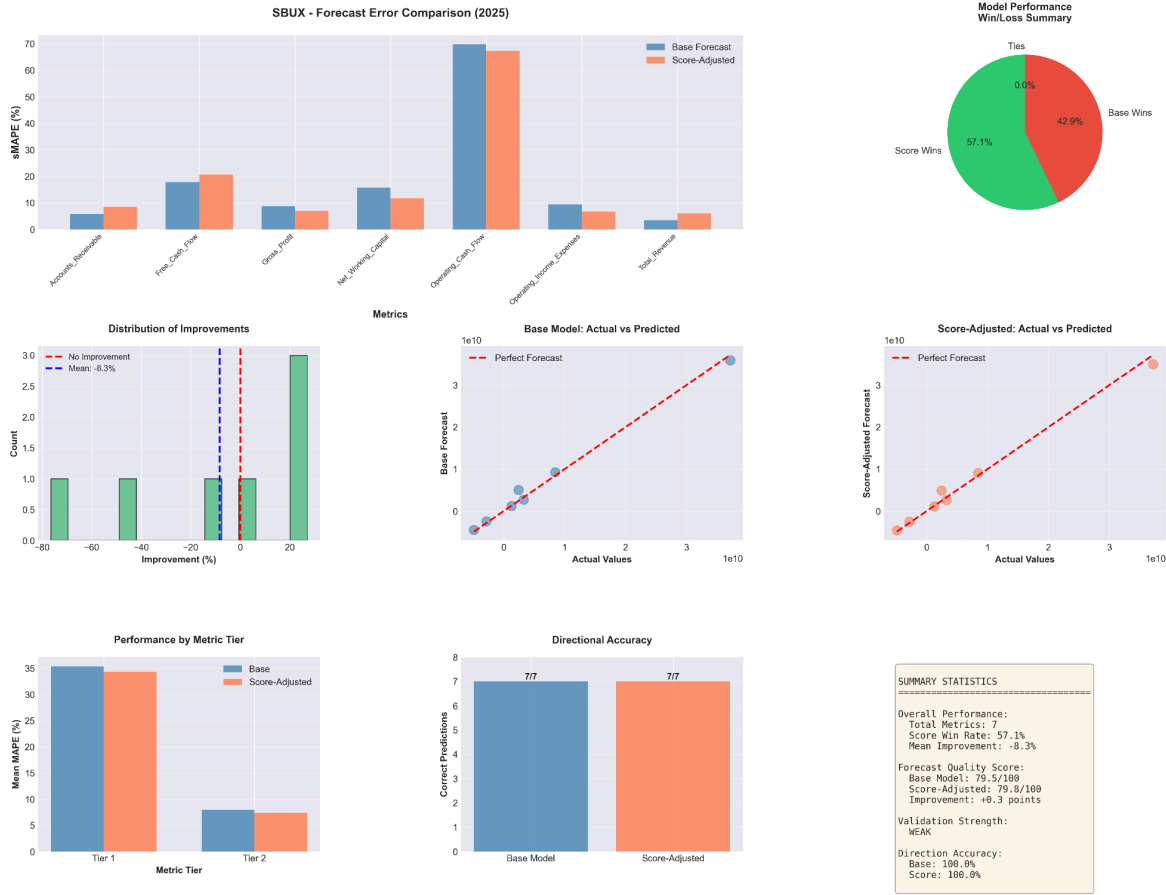
END OF REPORT

=====

=====

SBUX (LINEAR):

Forecast Validation Analysis: SBUX (2025)



FORECAST VALIDATION REPORT

Company: SBUX
Forecast Year: 2025
Analysis Date: 2025-12-12 14:37:53

EXECUTIVE SUMMARY

Validation Strength: WEAK

Score-Adjusted method shows weak validation. Qualitative factors provide marginal improvements.

Overall Win Rate: 57.1% (Score-Adjusted wins)

Mean Improvement: -8.3%

Forecast Quality Scores:

- Base Model: 79.5/100
- Score-Adjusted Model: 79.8/100
- Improvement: +0.3 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 23.64%
- Median MAPE: 9.27%
- Mean sMAPE: 18.72%

Score-Adjusted Model:

- Mean MAPE: 22.78%
- Median MAPE: 8.20%
- Mean sMAPE: 18.34%

Win/Loss Record:

- Score-Adjusted Wins: 4
- Base Model Wins: 3
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 35.38%
- Score Mean MAPE: 34.34%
- Mean Improvement: -15.32%
- Score Win Rate: 50.0%
- Wins: Score=2, Base=2

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 7.99%
- Score Mean MAPE: 7.37%
- Mean Improvement: 1.09%
- Score Win Rate: 66.7%
- Wins: Score=2, Base=1

STATISTICAL SIGNIFICANCE TESTS

Paired T-Test:

- T-Statistic: 0.3381
- P-Value: 0.7468
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 13.0000
- P-Value: 0.9375
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.1278
- Interpretation: Small effect

METRIC-BY-METRIC BREAKDOWN

Accounts_Receivable (Tier 2):

Actual Value: 1,277,500,000.00

Base Forecast: 1,205,267,638.00 (Error: -72,232,362.00)

Score Forecast: 1,172,783,265.00 (Error: -104,716,735.00)

Accuracy:

- Base sMAPE: 5.82%
- Score sMAPE: 8.55%
- Improvement: -44.97%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 3,268,000,000.00

Base Forecast: 2,730,751,250.00 (Error: -537,248,750.00)

Score Forecast: 2,657,152,042.00 (Error: -610,847,958.00)

Accuracy:

- Base sMAPE: 17.91%
- Score sMAPE: 20.62%
- Improvement: -13.70%

Direction Correct: Base=True, Score=True

Winner: Base

Gross_Profit (Tier 2):

Actual Value: 8,467,300,000.00

Base Forecast: 9,252,329,722.00 (Error: +785,029,722.00)

Score Forecast: 9,086,009,843.00 (Error: +618,709,843.00)

Accuracy:

- Base sMAPE: 8.86%
- Score sMAPE: 7.05%
- Improvement: +21.19%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: -2,828,100,000.00

Base Forecast: -2,415,137,500.00 (Error: +412,962,500.00)

Score Forecast: -2,512,805,660.00 (Error: +315,294,340.00)

Accuracy:

- Base sMAPE: 15.75%
- Score sMAPE: 11.81%
- Improvement: +23.65%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 2,442,000,000.00

Base Forecast: 5,056,894,861.00 (Error: +2,614,894,861.00)

Score Forecast: 4,920,601,430.00 (Error: +2,478,601,430.00)

Accuracy:

- Base sMAPE: 69.74%
- Score sMAPE: 67.33%
- Improvement: +5.21%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -4,886,500,000.00

Base Forecast: -4,443,815,000.00 (Error: +442,685,000.00)

Score Forecast: -4,563,584,701.00 (Error: +322,915,299.00)

Accuracy:

- Base sMAPE: 9.49%
- Score sMAPE: 6.83%
- Improvement: +27.06%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 37,184,400,000.00

Base Forecast: 35,918,170,000.00 (Error: -1,266,230,000.00)

Score Forecast: 34,950,103,482.00 (Error: -2,234,296,518.00)

Accuracy:

- Base sMAPE: 3.46%
- Score sMAPE: 6.19%
- Improvement: -76.45%

Direction Correct: Base=True, Score=True
Winner: Base

=====

=====

RECOMMENDATIONS

=====

=====

- △ Weak validation. Qualitative adjustments show marginal impact. Recommend:
1. Review and refine the qualitative scoring methodology
 2. Investigate if certain metric types benefit more from adjustments
 3. Consider whether the base model is already capturing qualitative factors
 4. Test alternative adjustment algorithms

=====

=====

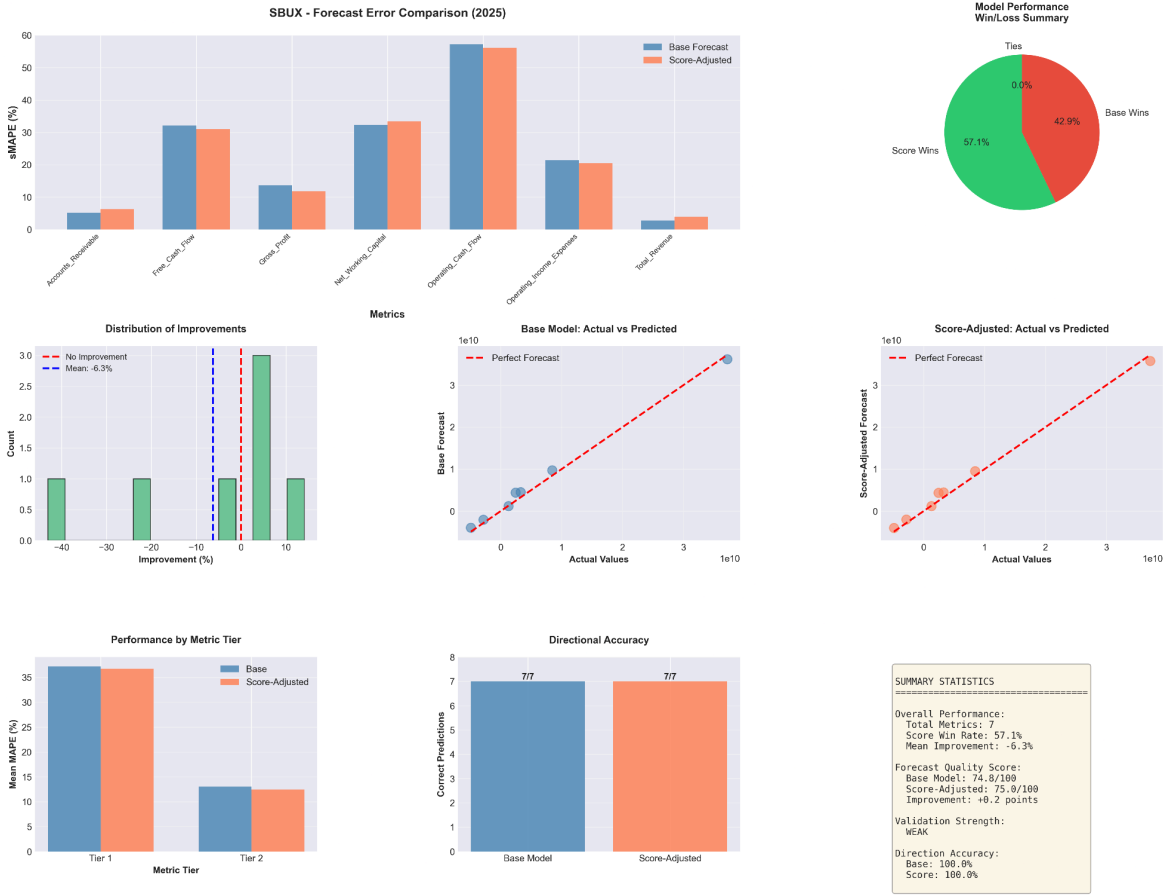
END OF REPORT

=====

=====

SBUX (XGBOOST):

Forecast Validation Analysis: SBUX (2025)



FORECAST VALIDATION REPORT

Company: SBUX
Forecast Year: 2025
Analysis Date: 2025-12-12 14:45:56

EXECUTIVE SUMMARY

Validation Strength: WEAK

Score-Adjusted method shows weak validation. Qualitative factors provide marginal improvements.

Overall Win Rate: 57.1% (Score-Adjusted wins)

Mean Improvement: -6.3%

Forecast Quality Scores:

- Base Model: 74.8/100
- Score-Adjusted Model: 75.0/100
- Improvement: +0.2 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 26.83%
- Median MAPE: 19.34%
- Mean sMAPE: 23.50%

Score-Adjusted Model:

- Mean MAPE: 26.34%
- Median MAPE: 18.57%
- Mean sMAPE: 23.30%

Win/Loss Record:

- Score-Adjusted Wins: 4
- Base Model Wins: 3
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 37.22%
- Score Mean MAPE: 36.77%
- Mean Improvement: -9.79%
- Score Win Rate: 50.0%
- Wins: Score=2, Base=2

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 12.99%
- Score Mean MAPE: 12.43%
- Mean Improvement: -1.59%
- Score Win Rate: 66.7%
- Wins: Score=2, Base=1

STATISTICAL SIGNIFICANCE TESTS

Paired T-Test:

- T-Statistic: 0.4101
- P-Value: 0.6960
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 14.0000
- P-Value: 1.0000
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.1550
- Interpretation: Small effect

METRIC-BY-METRIC BREAKDOWN

Accounts_Receivable (Tier 2):

Actual Value: 1,277,500,000.00

Base Forecast: 1,213,799,552.00 (Error: -63,700,448.00)

Score Forecast: 1,199,233,957.00 (Error: -78,266,043.00)

Accuracy:

- Base sMAPE: 5.11%
- Score sMAPE: 6.32%
- Improvement: -22.87%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 3,268,000,000.00

Base Forecast: 4,519,098,368.00 (Error: +1,251,098,368.00)

Score Forecast: 4,464,869,187.00 (Error: +1,196,869,187.00)

Accuracy:

- Base sMAPE: 32.13%
- Score sMAPE: 30.96%
- Improvement: +4.33%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 8,467,300,000.00

Base Forecast: 9,707,749,376.00 (Error: +1,240,449,376.00)

Score Forecast: 9,533,009,887.00 (Error: +1,065,709,887.00)

Accuracy:

- Base sMAPE: 13.65%
- Score sMAPE: 11.84%
- Improvement: +14.09%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: -2,828,100,000.00

Base Forecast: -2,041,901,696.00 (Error: +786,198,304.00)

Score Forecast: -2,017,398,875.00 (Error: +810,701,125.00)

Accuracy:

- Base sMAPE: 32.29%
- Score sMAPE: 33.46%
- Improvement: -3.12%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 2,442,000,000.00

Base Forecast: 4,397,301,760.00 (Error: +1,955,301,760.00)

Score Forecast: 4,344,534,138.00 (Error: +1,902,534,138.00)

Accuracy:

- Base sMAPE: 57.18%
- Score sMAPE: 56.07%
- Improvement: +2.70%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -4,886,500,000.00

Base Forecast: -3,941,399,808.00 (Error: +945,100,192.00)

Score Forecast: -3,979,237,246.00 (Error: +907,262,754.00)

Accuracy:

- Base sMAPE: 21.41%
- Score sMAPE: 20.47%
- Improvement: +4.00%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 37,184,400,000.00

Base Forecast: 36,176,187,392.00 (Error: -1,008,212,608.00)

Score Forecast: 35,742,073,143.00 (Error: -1,442,326,857.00)

Accuracy:

- Base sMAPE: 2.75%
- Score sMAPE: 3.96%
- Improvement: -43.06%

Direction Correct: Base=True, Score=True
Winner: Base

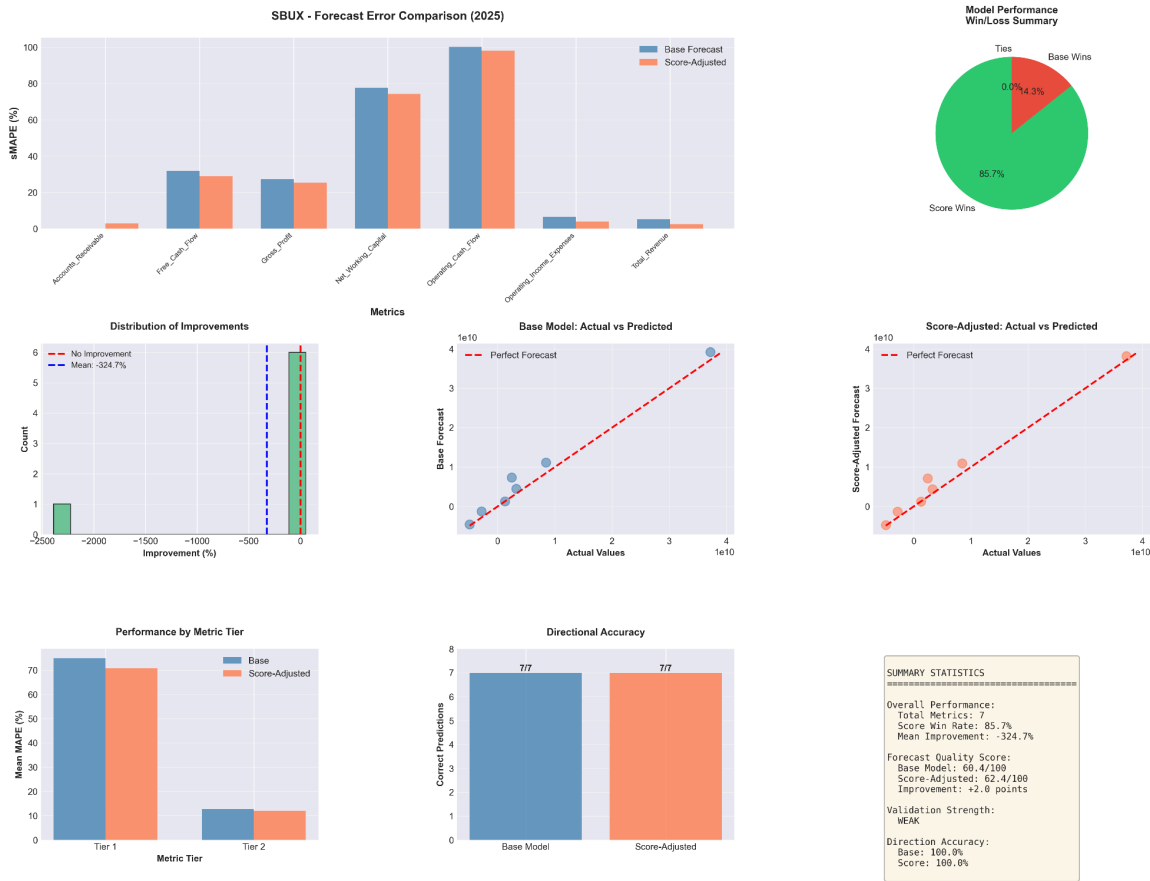
```
=====
=====
RECOMMENDATIONS
=====
=====
```

- △ Weak validation. Qualitative adjustments show marginal impact. Recommend:
1. Review and refine the qualitative scoring methodology
 2. Investigate if certain metric types benefit more from adjustments
 3. Consider whether the base model is already capturing qualitative factors
 4. Test alternative adjustment algorithms

```
=====
=====
END OF REPORT
=====
=====
```


SBUX (ARIMA):

Forecast Validation Analysis: SBUX (2025)



FORECAST VALIDATION REPORT

Company: SBUX
Forecast Year: 2025
Analysis Date: 2025-12-12 14:56:03

EXECUTIVE SUMMARY

Validation Strength: WEAK

Score-Adjusted method shows weak validation. Qualitative factors provide marginal improvements.

Overall Win Rate: 85.7% (Score-Adjusted wins)

Mean Improvement: -324.7%

Forecast Quality Scores:

- Base Model: 60.4/100
- Score-Adjusted Model: 62.4/100
- Improvement: +2.0 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 48.28%
- Median MAPE: 31.53%
- Mean sMAPE: 35.55%

Score-Adjusted Model:

- Mean MAPE: 45.62%
- Median MAPE: 29.17%
- Mean sMAPE: 33.76%

Win/Loss Record:

- Score-Adjusted Wins: 6
- Base Model Wins: 1
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 74.98%
- Score Mean MAPE: 70.87%
- Mean Improvement: 17.36%
- Score Win Rate: 100.0%
- Wins: Score=4, Base=0

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 12.69%
- Score Mean MAPE: 11.95%
- Mean Improvement: -780.89%
- Score Win Rate: 66.7%
- Wins: Score=2, Base=1

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 2.3029
- P-Value: 0.0609
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 6.0000
- P-Value: 0.2188
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.8704
- Interpretation: Large effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 1,277,500,000.00

Base Forecast: 1,276,060,699.00 (Error: -1,439,301.00)

Score Forecast: 1,241,668,311.00 (Error: -35,831,689.00)

Accuracy:

- Base sMAPE: 0.11%
- Score sMAPE: 2.84%
- Improvement: -2389.52%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 3,268,000,000.00

Base Forecast: 4,501,430,611.00 (Error: +1,233,430,611.00)

Score Forecast: 4,380,108,053.00 (Error: +1,112,108,053.00)

Accuracy:

- Base sMAPE: 31.75%
- Score sMAPE: 29.08%
- Improvement: +9.84%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 8,467,300,000.00

Base Forecast: 11,137,270,015.00 (Error: +2,669,970,015.00)

Score Forecast: 10,937,066,449.00 (Error: +2,469,766,449.00)

Accuracy:

- Base sMAPE: 27.24%
- Score sMAPE: 25.46%
- Improvement: +7.50%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: -2,828,100,000.00
Base Forecast: -1,246,072,777.00 (Error: +1,582,027,223.00)
Score Forecast: -1,296,463,960.00 (Error: +1,531,636,040.00)

Accuracy:

- Base sMAPE: 77.66%
- Score sMAPE: 74.27%
- Improvement: +3.19%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 2,442,000,000.00
Base Forecast: 7,345,980,117.00 (Error: +4,903,980,117.00)
Score Forecast: 7,147,991,260.00 (Error: +4,705,991,260.00)

Accuracy:

- Base sMAPE: 100.20%
- Score sMAPE: 98.14%
- Improvement: +4.04%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -4,886,500,000.00
Base Forecast: -4,573,278,648.00 (Error: +313,221,352.00)
Score Forecast: -4,696,537,654.00 (Error: +189,962,346.00)

Accuracy:

- Base sMAPE: 6.62%
- Score sMAPE: 3.96%
- Improvement: +39.35%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 37,184,400,000.00
Base Forecast: 39,201,921,515.00 (Error: +2,017,521,515.00)
Score Forecast: 38,145,351,327.00 (Error: +960,951,327.00)

Accuracy:

- Base sMAPE: 5.28%
- Score sMAPE: 2.55%
- Improvement: +52.37%

Direction Correct: Base=True, Score=True

Winner: Score

=====

=====

RECOMMENDATIONS

=====

=====

△ Weak validation. Qualitative adjustments show marginal impact. Recommend:

1. Review and refine the qualitative scoring methodology
2. Investigate if certain metric types benefit more from adjustments
3. Consider whether the base model is already capturing qualitative factors
4. Test alternative adjustment algorithms

=====

=====

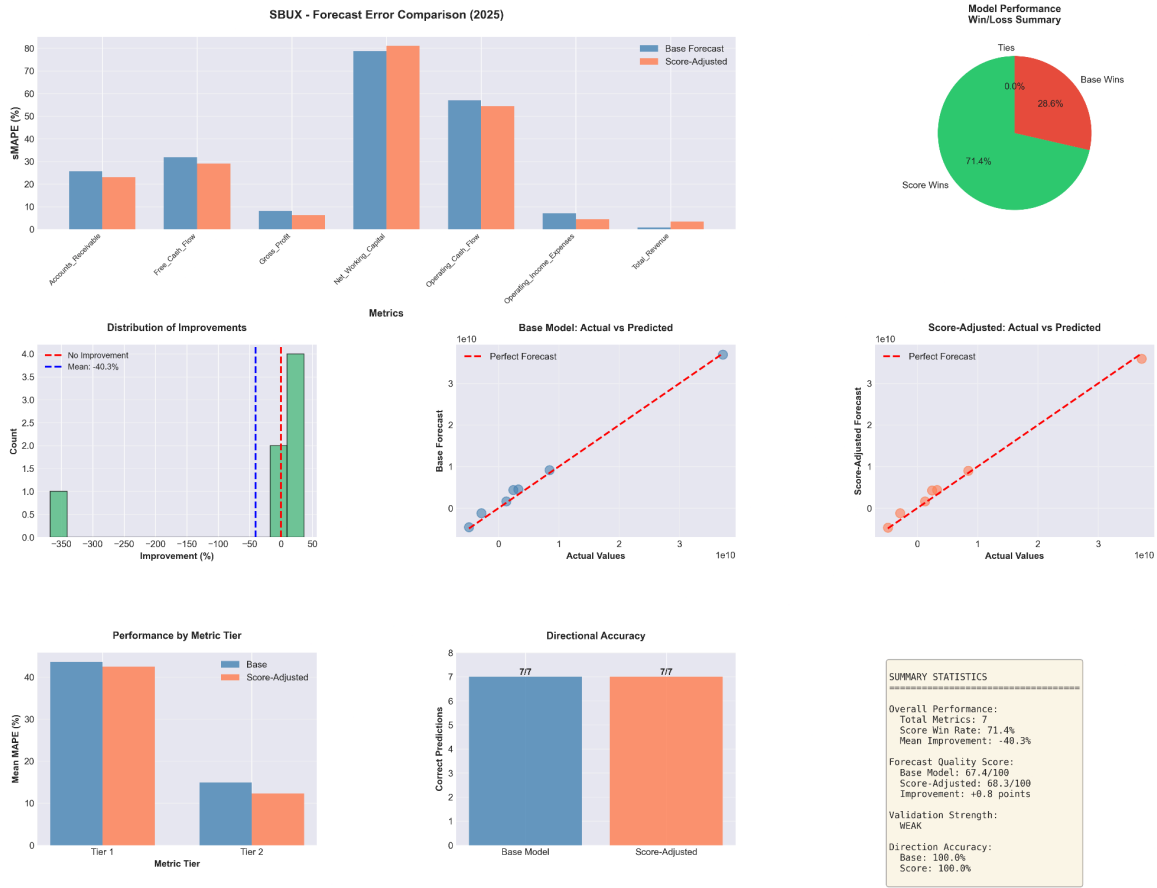
END OF REPORT

=====

=====

SBUX (WINTERS-HOLT):

Forecast Validation Analysis: SBUX (2025)



FORECAST VALIDATION REPORT

Company: SBUX
Forecast Year: 2025
Analysis Date: 2025-12-12 15:20:46

EXECUTIVE SUMMARY

Validation Strength: WEAK

Score-Adjusted method shows weak validation. Qualitative factors provide marginal improvements.

Overall Win Rate: 71.4% (Score-Adjusted wins)

Mean Improvement: -40.3%

Forecast Quality Scores:

- Base Model: 67.4/100
- Score-Adjusted Model: 68.3/100
- Improvement: +0.8 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 31.34%
- Median MAPE: 29.54%
- Mean sMAPE: 29.86%

Score-Adjusted Model:

- Mean MAPE: 29.53%
- Median MAPE: 26.05%
- Mean sMAPE: 28.81%

Win/Loss Record:

- Score-Adjusted Wins: 5
- Base Model Wins: 2
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 43.64%
- Score Mean MAPE: 42.47%
- Mean Improvement: -88.47%
- Score Win Rate: 50.0%
- Wins: Score=2, Base=2

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 14.92%
- Score Mean MAPE: 12.27%
- Mean Improvement: 23.91%
- Score Win Rate: 100.0%
- Wins: Score=3, Base=0

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 1.1275
- P-Value: 0.3026
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 9.0000
- P-Value: 0.4688
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.4261
- Interpretation: Small effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 1,277,500,000.00

Base Forecast: 1,654,902,375.00 (Error: +377,402,375.00)

Score Forecast: 1,610,299,446.00 (Error: +332,799,446.00)

Accuracy:

- Base sMAPE: 25.74%
- Score sMAPE: 23.05%
- Improvement: +11.82%

Direction Correct: Base=True, Score=True

Winner: Score

Free_Cash_Flow (Tier 1):

Actual Value: 3,268,000,000.00

Base Forecast: 4,501,430,611.00 (Error: +1,233,430,611.00)

Score Forecast: 4,380,108,053.00 (Error: +1,112,108,053.00)

Accuracy:

- Base sMAPE: 31.75%
- Score sMAPE: 29.08%
- Improvement: +9.84%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 8,467,300,000.00

Base Forecast: 9,176,607,922.00 (Error: +709,307,922.00)

Score Forecast: 9,011,649,218.00 (Error: +544,349,218.00)

Accuracy:

- Base sMAPE: 8.04%
- Score sMAPE: 6.23%
- Improvement: +23.26%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: -2,828,100,000.00
Base Forecast: -1,230,809,902.00 (Error: +1,597,290,098.00)
Score Forecast: -1,197,637,114.00 (Error: +1,630,462,886.00)

Accuracy:

- Base sMAPE: 78.71%
- Score sMAPE: 81.00%
- Improvement: -2.08%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 2,442,000,000.00
Base Forecast: 4,386,546,165.00 (Error: +1,944,546,165.00)
Score Forecast: 4,268,319,972.00 (Error: +1,826,319,972.00)

Accuracy:

- Base sMAPE: 56.95%
- Score sMAPE: 54.43%
- Improvement: +6.08%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -4,886,500,000.00
Base Forecast: -4,551,873,947.00 (Error: +334,626,053.00)
Score Forecast: -4,674,556,054.00 (Error: +211,943,946.00)

Accuracy:

- Base sMAPE: 7.09%
- Score sMAPE: 4.43%
- Improvement: +36.66%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 37,184,400,000.00
Base Forecast: 36,913,849,029.00 (Error: -270,550,971.00)
Score Forecast: 35,918,946,970.00 (Error: -1,265,453,030.00)

Accuracy:

- Base sMAPE: 0.73%
- Score sMAPE: 3.46%
- Improvement: -367.73%

Direction Correct: Base=True, Score=True

Winner: Base

=====

=====

RECOMMENDATIONS

=====

=====

△ Weak validation. Qualitative adjustments show marginal impact. Recommend:

1. Review and refine the qualitative scoring methodology
2. Investigate if certain metric types benefit more from adjustments
3. Consider whether the base model is already capturing qualitative factors
4. Test alternative adjustment algorithms

=====

=====

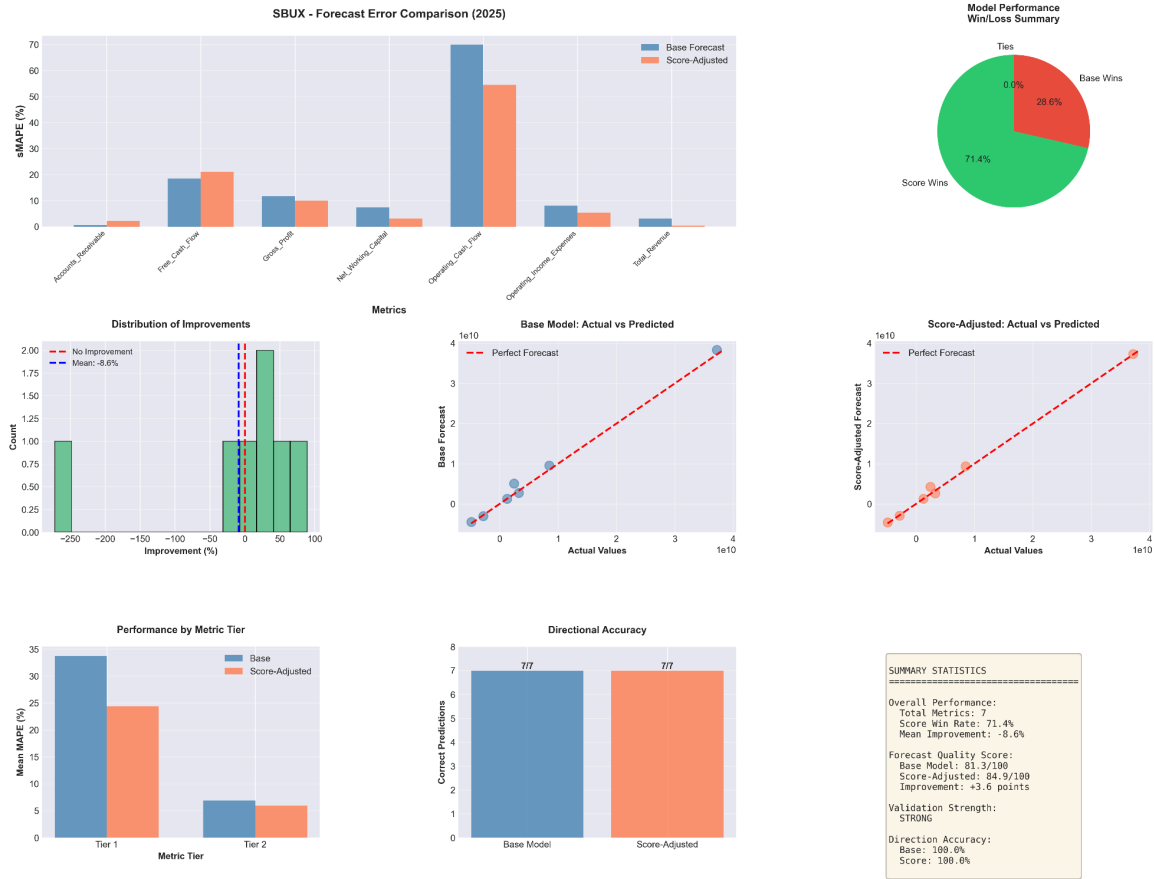
END OF REPORT

=====

=====

SBUX (LINEAR + WINTERS-HOLT):

Forecast Validation Analysis: SBUX (2025)



FORECAST VALIDATION REPORT

Company: SBUX
Forecast Year: 2025
Analysis Date: 2025-12-12 15:21:14

EXECUTIVE SUMMARY

Validation Strength: STRONG

Score-Adjusted method shows strong validation. Qualitative factors significantly improve forecast accuracy.

Overall Win Rate: 71.4% (Score-Adjusted wins)

Mean Improvement: -8.6%

Forecast Quality Scores:

- Base Model: 81.3/100
- Score-Adjusted Model: 84.9/100
- Improvement: +3.6 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 22.27%
- Median MAPE: 7.73%
- Mean sMAPE: 17.02%

Score-Adjusted Model:

- Mean MAPE: 16.48%
- Median MAPE: 5.24%
- Mean sMAPE: 13.80%

Win/Loss Record:

- Score-Adjusted Wins: 5
- Base Model Wins: 2
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 33.76%
- Score Mean MAPE: 24.37%
- Mean Improvement: 40.85%
- Score Win Rate: 75.0%
- Wins: Score=3, Base=1

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 6.94%
- Score Mean MAPE: 5.96%
- Mean Improvement: -74.60%
- Score Win Rate: 66.7%
- Wins: Score=2, Base=1

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 1.4318
- P-Value: 0.2022
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 5.0000
- P-Value: 0.1562
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.5412
- Interpretation: Medium effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 1,277,500,000.00

Base Forecast: 1,284,834,733.00 (Error: +7,334,733.00)

Score Forecast: 1,250,205,868.00 (Error: -27,294,132.00)

Accuracy:

- Base sMAPE: 0.57%
- Score sMAPE: 2.16%
- Improvement: -272.12%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 3,268,000,000.00

Base Forecast: 2,717,145,495.00 (Error: -550,854,505.00)

Score Forecast: 2,643,912,990.00 (Error: -624,087,010.00)

Accuracy:

- Base sMAPE: 18.41%
- Score sMAPE: 21.11%
- Improvement: -13.29%

Direction Correct: Base=True, Score=True

Winner: Base

Gross_Profit (Tier 2):

Actual Value: 8,467,300,000.00

Base Forecast: 9,527,385,913.00 (Error: +1,060,085,913.00)

Score Forecast: 9,356,121,624.00 (Error: +888,821,624.00)

Accuracy:

- Base sMAPE: 11.78%
- Score sMAPE: 9.97%
- Improvement: +16.16%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: -2,828,100,000.00

Base Forecast: -3,043,246,279.00 (Error: -215,146,279.00)

Score Forecast: -2,920,177,399.00 (Error: -92,077,399.00)

Accuracy:

- Base sMAPE: 7.33%
- Score sMAPE: 3.20%
- Improvement: +57.20%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 2,442,000,000.00

Base Forecast: 5,066,346,342.00 (Error: +2,624,346,342.00)

Score Forecast: 4,268,319,972.00 (Error: +1,826,319,972.00)

Accuracy:

- Base sMAPE: 69.90%
- Score sMAPE: 54.43%
- Improvement: +30.41%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -4,886,500,000.00

Base Forecast: -4,508,718,633.00 (Error: +377,781,367.00)

Score Forecast: -4,630,237,618.00 (Error: +256,262,382.00)

Accuracy:

- Base sMAPE: 8.04%
- Score sMAPE: 5.39%
- Improvement: +32.17%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 37,184,400,000.00

Base Forecast: 38,344,408,454.00 (Error: +1,160,008,454.00)

Score Forecast: 37,310,949,958.00 (Error: +126,549,958.00)

Accuracy:

- Base sMAPE: 3.07%
- Score sMAPE: 0.34%
- Improvement: +89.09%

Direction Correct: Base=True, Score=True

Winner: Score

```
=====
=====
RECOMMENDATIONS
=====
=====
```

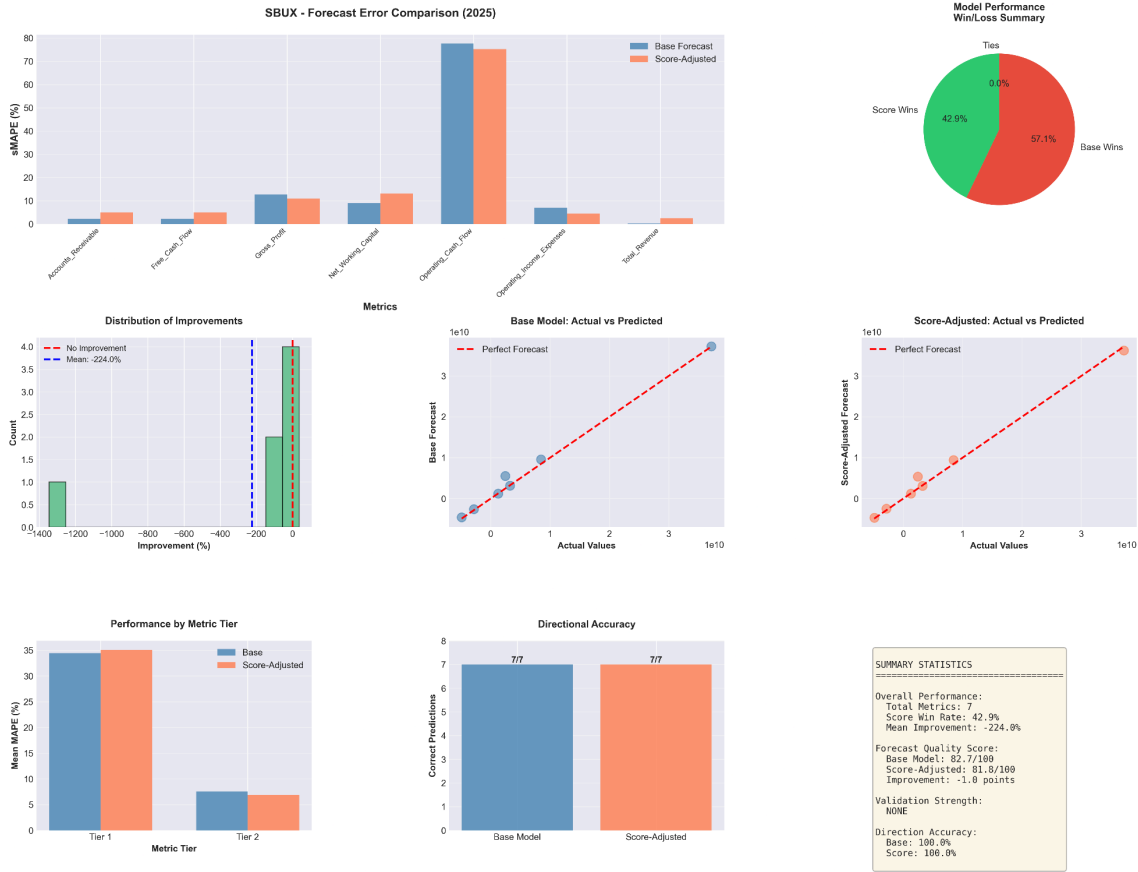
✓ Strong validation achieved. Qualitative score adjustments significantly improve forecast accuracy. Recommend:

1. Continue using score-adjusted methodology for future forecasts
2. Analyze which qualitative factors drove the best improvements
3. Consider increasing weight given to qualitative adjustments
4. Document the successful factors for replication

```
=====
=====
END OF REPORT
=====
=====
```


SBUX (LINEAR + WINTERS-HOLT + XGBOOST):

Forecast Validation Analysis: SBUX (2025)



FORECAST VALIDATION REPORT

Company: SBUX
Forecast Year: 2025
Analysis Date: 2025-12-12 15:24:26

EXECUTIVE SUMMARY

Validation Strength: NONE

Score-Adjusted method does not validate. Base model performs as well or better.

Overall Win Rate: 42.9% (Score-Adjusted wins)

Mean Improvement: -224.0%

Forecast Quality Scores:

- Base Model: 82.7/100
- Score-Adjusted Model: 81.8/100
- Improvement: -1.0 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 22.91%
- Median MAPE: 6.82%
- Mean sMAPE: 15.86%

Score-Adjusted Model:

- Mean MAPE: 23.00%
- Median MAPE: 4.88%
- Mean sMAPE: 16.60%

Win/Loss Record:

- Score-Adjusted Wins: 3
- Base Model Wins: 4
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 34.43%
- Score Mean MAPE: 35.07%
- Mean Improvement: -375.64%
- Score Win Rate: 25.0%
- Wins: Score=1, Base=3

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 7.55%
- Score Mean MAPE: 6.91%
- Mean Improvement: -21.89%
- Score Win Rate: 66.7%
- Wins: Score=2, Base=1

STATISTICAL SIGNIFICANCE TESTS

Paired T-Test:

- T-Statistic: -0.6793
- P-Value: 0.5222
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 7.0000
- P-Value: 0.2969
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: -0.2568
- Interpretation: Small effect

METRIC-BY-METRIC BREAKDOWN

Accounts_Receivable (Tier 2):

Actual Value: 1,277,500,000.00

Base Forecast: 1,248,861,206.00 (Error: -28,638,794.00)

Score Forecast: 1,215,201,899.00 (Error: -62,298,101.00)

Accuracy:

- Base sMAPE: 2.27%
- Score sMAPE: 5.00%
- Improvement: -117.53%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 3,268,000,000.00

Base Forecast: 3,195,036,875.00 (Error: -72,963,125.00)

Score Forecast: 3,108,924,241.00 (Error: -159,075,759.00)

Accuracy:

- Base sMAPE: 2.26%
- Score sMAPE: 4.99%
- Improvement: -118.02%

Direction Correct: Base=True, Score=True

Winner: Base

Gross_Profit (Tier 2):

Actual Value: 8,467,300,000.00

Base Forecast: 9,617,204,636.00 (Error: +1,149,904,636.00)

Score Forecast: 9,444,325,766.00 (Error: +977,025,766.00)

Accuracy:

- Base sMAPE: 12.72%
- Score sMAPE: 10.91%
- Improvement: +15.03%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: -2,828,100,000.00

Base Forecast: -2,584,647,203.00 (Error: +243,452,797.00)

Score Forecast: -2,480,124,070.00 (Error: +347,975,930.00)

Accuracy:

- Base sMAPE: 9.00%
- Score sMAPE: 13.11%
- Improvement: -42.93%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 2,442,000,000.00

Base Forecast: 5,536,274,099.00 (Error: +3,094,274,099.00)

Score Forecast: 5,387,060,439.00 (Error: +2,945,060,439.00)

Accuracy:

- Base sMAPE: 77.57%
- Score sMAPE: 75.23%
- Improvement: +4.82%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -4,886,500,000.00

Base Forecast: -4,553,149,460.00 (Error: +333,350,540.00)

Score Forecast: -4,675,865,945.00 (Error: +210,634,055.00)

Accuracy:

- Base sMAPE: 7.06%
- Score sMAPE: 4.41%
- Improvement: +36.81%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 37,184,400,000.00

Base Forecast: 37,249,320,643.00 (Error: +64,920,643.00)

Score Forecast: 36,245,376,953.00 (Error: -939,023,047.00)

Accuracy:

- Base sMAPE: 0.17%

- Score sMAPE: 2.56%
- Improvement: -1346.42%

Direction Correct: Base=True, Score=True

Winner: Base

=====

=====

RECOMMENDATIONS

=====

=====

✗ No validation. Base model performs as well or better. Recommend:

1. Critically review the qualitative scoring framework
2. Check if qualitative factors are already reflected in historical data
3. Consider that market efficiency may limit qualitative alpha
4. Explore if timing or sector-specific factors affected results

=====

=====

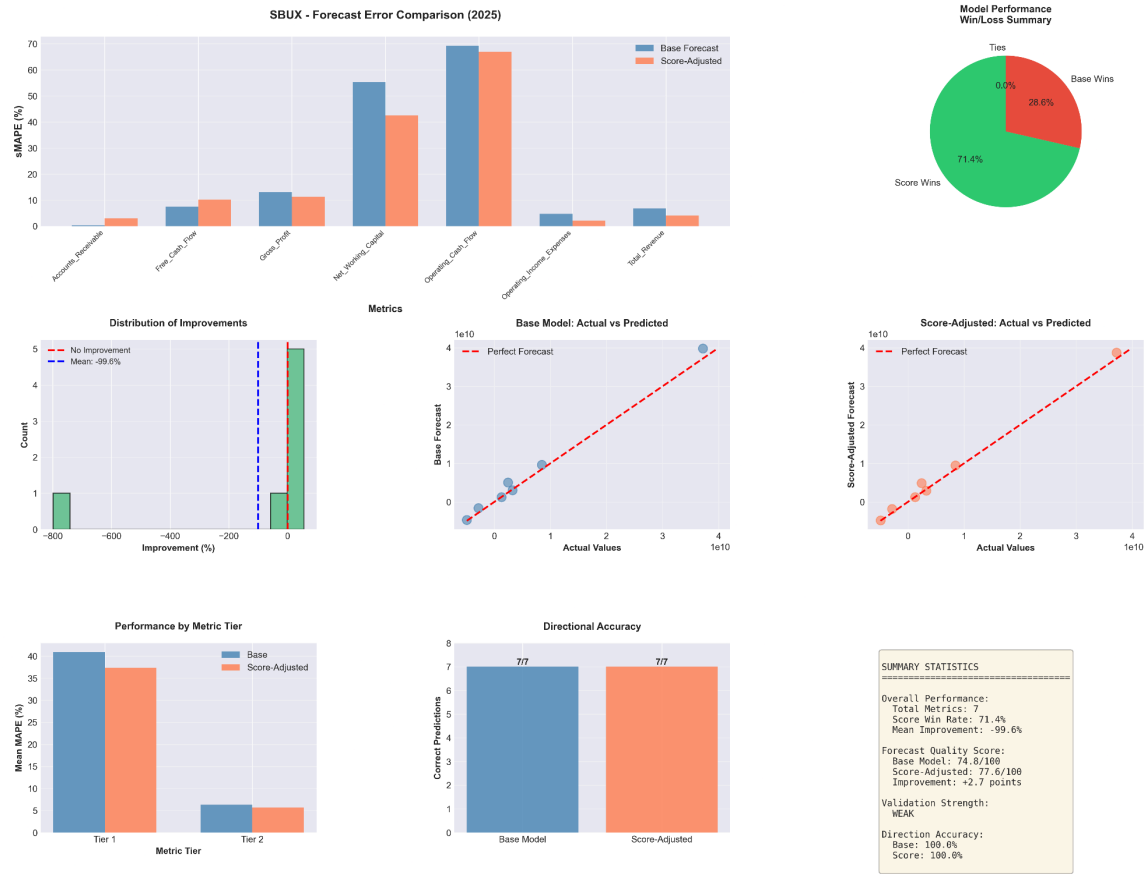
END OF REPORT

=====

=====

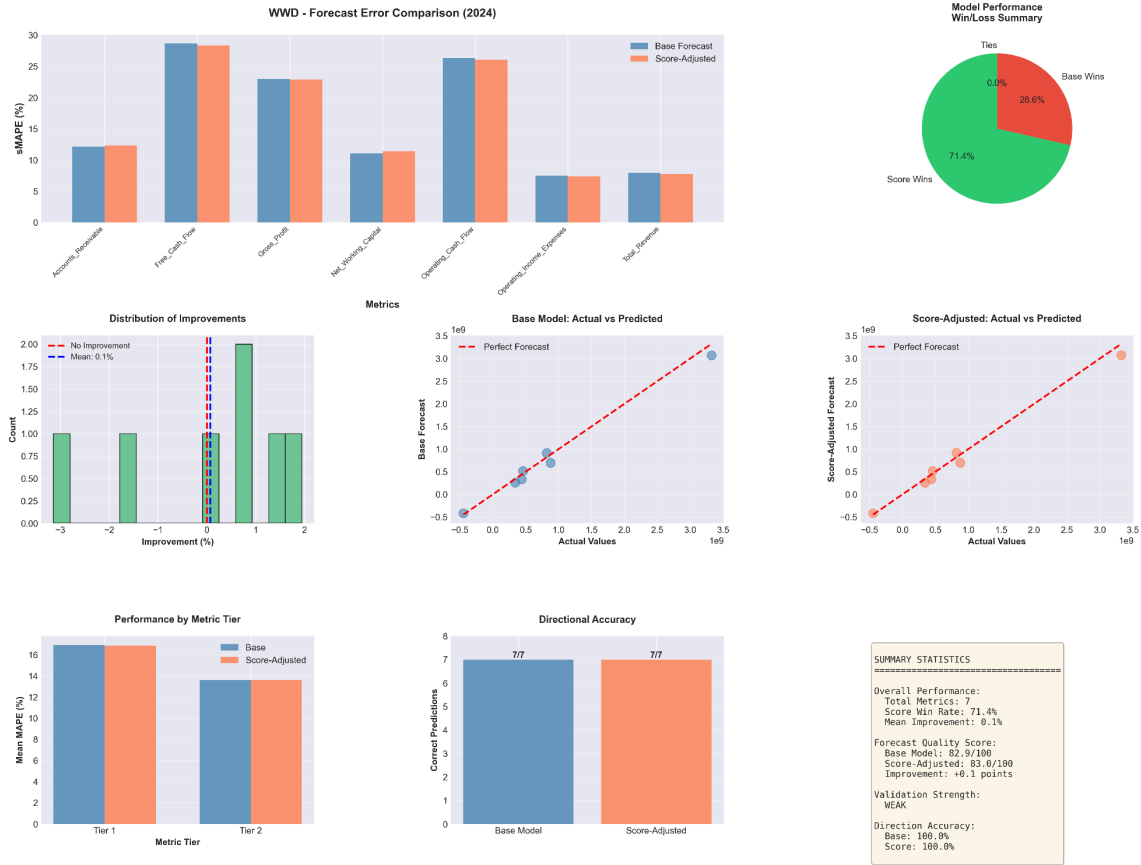
SBUX (LINEAR + WINTERS-HOLT + ARIMA):

Forecast Validation Analysis: SBUX (2025)



WWD (NAIVE):

Forecast Validation Analysis: WWD (2024)



FORECAST VALIDATION REPORT

Company: WWD
Forecast Year: 2024
Analysis Date: 2025-12-06 22:06:47

EXECUTIVE SUMMARY

Validation Strength: WEAK

Score-Adjusted method shows weak validation. Qualitative factors provide marginal improvements.

Overall Win Rate: 71.4% (Score-Adjusted wins)

Mean Improvement: 0.1%

Forecast Quality Scores:

- Base Model: 82.9/100
- Score-Adjusted Model: 83.0/100
- Improvement: +0.1 points

```
=====
=====
DETAILED METRICS COMPARISON
=====
=====
```

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 15.49%
- Median MAPE: 12.92%
- Mean sMAPE: 16.66%

Score-Adjusted Model:

- Mean MAPE: 15.47%
- Median MAPE: 13.13%
- Mean sMAPE: 16.61%

Win/Loss Record:

- Score-Adjusted Wins: 5
- Base Model Wins: 2
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

```
=====
=====
PERFORMANCE BY METRIC TIER
=====
=====
```


Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 16.92%
- Score Mean MAPE: 16.87%
- Mean Improvement: 0.15%
- Score Win Rate: 75.0%
- Wins: Score=3, Base=1

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 13.59%
- Score Mean MAPE: 13.61%
- Mean Improvement: -0.03%
- Score Win Rate: 66.7%
- Wins: Score=2, Base=1

STATISTICAL SIGNIFICANCE TESTS

Paired T-Test:

- T-Statistic: 0.5893
- P-Value: 0.5771
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 11.0000
- P-Value: 0.6875
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.2227
- Interpretation: Small effect

METRIC-BY-METRIC BREAKDOWN

Accounts_Receivable (Tier 2):

Actual Value: 457,234,000.00
Base Forecast: 516,291,062.00 (Error: +59,057,062.00)
Score Forecast: 517,267,841.00 (Error: +60,033,841.00)

Accuracy:

- Base sMAPE: 12.13%
- Score sMAPE: 12.32%
- Improvement: -1.65%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 342,809,000.00
Base Forecast: 256,886,473.00 (Error: -85,922,527.00)
Score Forecast: 257,667,652.00 (Error: -85,141,348.00)

Accuracy:

- Base sMAPE: 28.66%
- Score sMAPE: 28.36%
- Improvement: +0.91%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 876,479,000.00
Base Forecast: 695,961,771.00 (Error: -180,517,229.00)
Score Forecast: 696,347,210.00 (Error: -180,131,790.00)

Accuracy:

- Base sMAPE: 22.96%
- Score sMAPE: 22.91%
- Improvement: +0.21%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 820,101,000.00
Base Forecast: 916,209,077.00 (Error: +96,108,077.00)
Score Forecast: 919,226,384.00 (Error: +99,125,384.00)

Accuracy:

- Base sMAPE: 11.07%
- Score sMAPE: 11.40%
- Improvement: -3.14%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 439,089,000.00

Base Forecast: 336,844,786.00 (Error: -102,244,214.00)

Score Forecast: 337,734,707.00 (Error: -101,354,293.00)

Accuracy:

- Base sMAPE: 26.35%
- Score sMAPE: 26.09%
- Improvement: +0.87%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -448,175,000.00

Base Forecast: -415,640,841.00 (Error: +32,534,159.00)

Score Forecast: -416,076,461.00 (Error: +32,098,539.00)

Accuracy:

- Base sMAPE: 7.53%
- Score sMAPE: 7.43%
- Improvement: +1.34%

Direction Correct: Base=True, Score=True

Winner: Score

Total_Revenue (Tier 1):

Actual Value: 3,324,249,000.00

Base Forecast: 3,071,096,221.00 (Error: -253,152,779.00)

Score Forecast: 3,076,018,157.00 (Error: -248,230,843.00)

Accuracy:

- Base sMAPE: 7.92%

- Score sMAPE: 7.76%
- Improvement: +1.94%

Direction Correct: Base=True, Score=True

Winner: Score

```
=====
=====
RECOMMENDATIONS
=====
=====
```

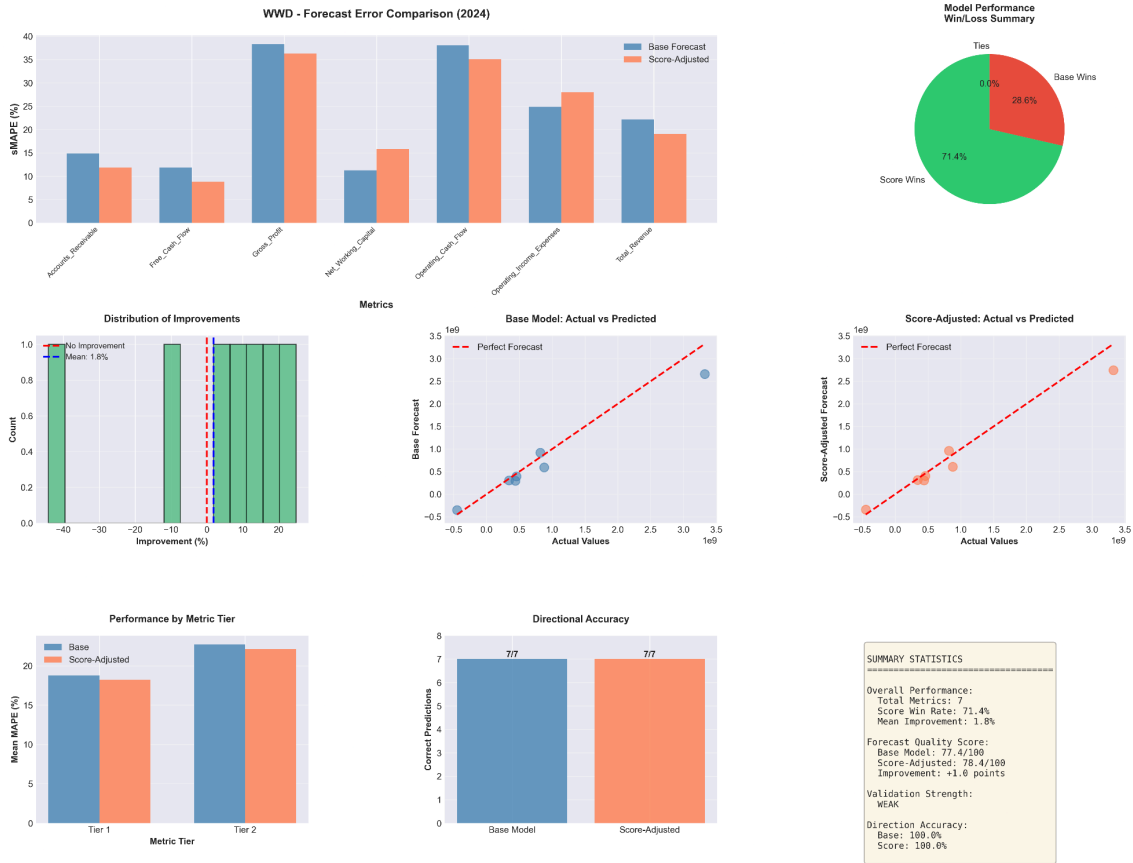
△ Weak validation. Qualitative adjustments show marginal impact. Recommend:

1. Review and refine the qualitative scoring methodology
2. Investigate if certain metric types benefit more from adjustments
3. Consider whether the base model is already capturing qualitative factors
4. Test alternative adjustment algorithms

```
=====
=====
END OF REPORT
=====
=====
```


WWD (LINEAR):

Forecast Validation Analysis: WWD (2024)



FORECAST VALIDATION REPORT

Company: WWD
Forecast Year: 2024
Analysis Date: 2025-12-06 22:15:04

EXECUTIVE SUMMARY

Validation Strength: WEAK

Score-Adjusted method shows weak validation. Qualitative factors provide marginal improvements.

Overall Win Rate: 71.4% (Score-Adjusted wins)

Mean Improvement: 1.8%

Forecast Quality Scores:

- Base Model: 77.4/100
- Score-Adjusted Model: 78.4/100
- Improvement: +1.0 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 20.46%
- Median MAPE: 19.95%
- Mean sMAPE: 23.07%

Score-Adjusted Model:

- Mean MAPE: 19.91%
- Median MAPE: 17.43%
- Mean sMAPE: 22.14%

Win/Loss Record:

- Score-Adjusted Wins: 5
- Base Model Wins: 2
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 18.78%
- Score Mean MAPE: 18.24%
- Mean Improvement: 0.00%
- Score Win Rate: 75.0%
- Wins: Score=3, Base=1

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 22.71%
- Score Mean MAPE: 22.15%
- Mean Improvement: 4.29%
- Score Win Rate: 66.7%
- Wins: Score=2, Base=1

STATISTICAL SIGNIFICANCE TESTS

Paired T-Test:

- T-Statistic: 0.7394
- P-Value: 0.4876
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 13.0000
- P-Value: 0.9375
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.2795
- Interpretation: Small effect

METRIC-BY-METRIC BREAKDOWN

Accounts_Receivable (Tier 2):

Actual Value: 457,234,000.00

Base Forecast: 393,798,697.00 (Error: -63,435,303.00)

Score Forecast: 406,181,303.00 (Error: -51,052,697.00)

Accuracy:

- Base sMAPE: 14.91%

- Score sMAPE: 11.83%

- Improvement: +19.52%

Direction Correct: Base=True, Score=True

Winner: Score

Free_Cash_Flow (Tier 1):

Actual Value: 342,809,000.00

Base Forecast: 304,323,870.00 (Error: -38,485,130.00)

Score Forecast: 313,893,030.00 (Error: -28,915,970.00)

Accuracy:

- Base sMAPE: 11.89%

- Score sMAPE: 8.81%

- Improvement: +24.86%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 876,479,000.00

Base Forecast: 594,791,200.00 (Error: -281,687,800.00)

Score Forecast: 607,265,161.00 (Error: -269,213,839.00)

Accuracy:

- Base sMAPE: 38.29%

- Score sMAPE: 36.29%

- Improvement: +4.43%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 820,101,000.00

Base Forecast: 918,212,076.00 (Error: +98,111,076.00)

Score Forecast: 961,533,322.00 (Error: +141,432,322.00)

Accuracy:

- Base sMAPE: 11.29%
- Score sMAPE: 15.88%
- Improvement: -44.16%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 439,089,000.00

Base Forecast: 298,654,455.00 (Error: -140,434,545.00)

Score Forecast: 308,045,346.00 (Error: -131,043,654.00)

Accuracy:

- Base sMAPE: 38.07%
- Score sMAPE: 35.08%
- Improvement: +6.69%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -448,175,000.00

Base Forecast: -349,058,447.00 (Error: +99,116,553.00)

Score Forecast: -338,082,653.00 (Error: +110,092,347.00)

Accuracy:

- Base sMAPE: 24.87%
- Score sMAPE: 28.00%
- Improvement: -11.07%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 3,324,249,000.00

Base Forecast: 2,661,181,641.00 (Error: -663,067,359.00)

Score Forecast: 2,744,859,837.00 (Error: -579,389,163.00)

Accuracy:

- Base sMAPE: 22.16%
- Score sMAPE: 19.09%
- Improvement: +12.62%

Direction Correct: Base=True, Score=True
Winner: Score

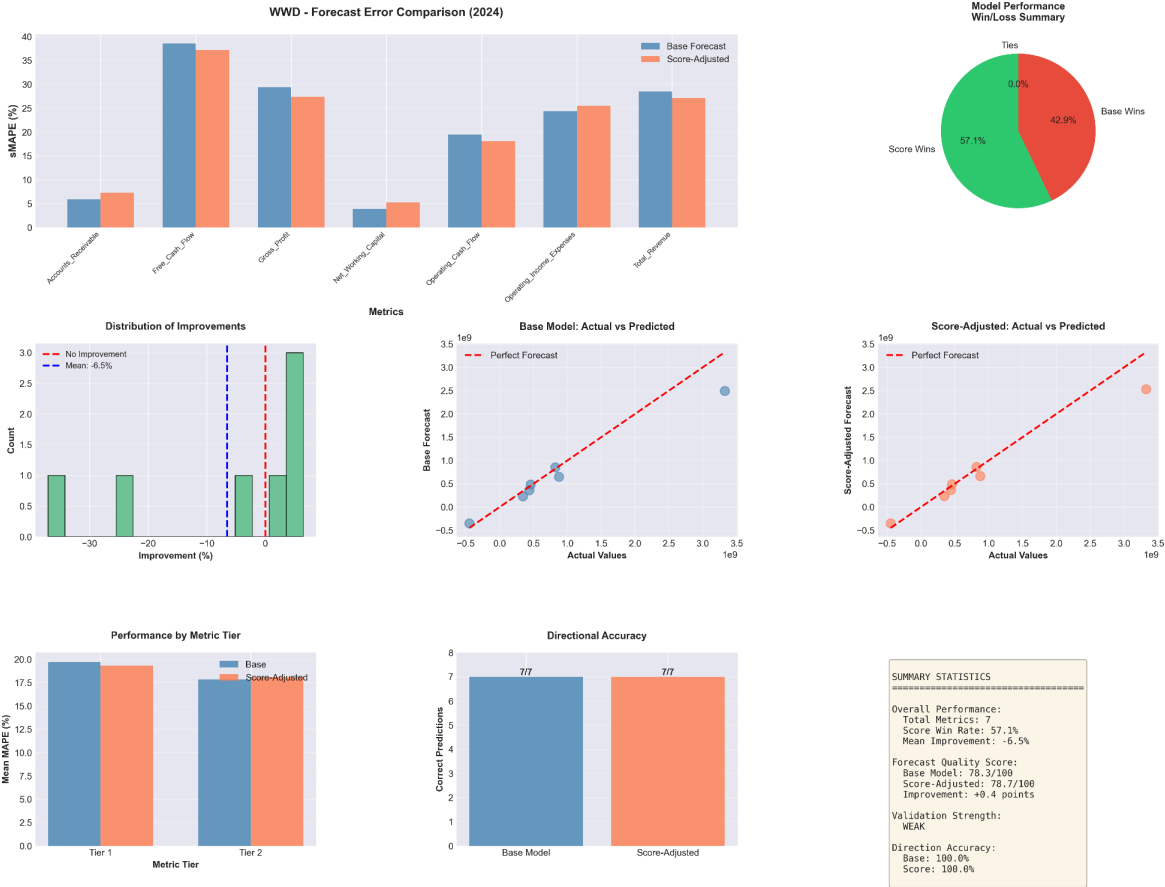
```
=====
=====
RECOMMENDATIONS
=====
=====
```

- △ Weak validation. Qualitative adjustments show marginal impact. Recommend:
1. Review and refine the qualitative scoring methodology
 2. Investigate if certain metric types benefit more from adjustments
 3. Consider whether the base model is already capturing qualitative factors
 4. Test alternative adjustment algorithms

```
=====
=====
END OF REPORT
=====
=====
```


WWD (XGBOOST):

Forecast Validation Analysis: WWD (2024)



FORECAST VALIDATION REPORT

Company: WWD
Forecast Year: 2024
Analysis Date: 2025-12-06 22:03:53

EXECUTIVE SUMMARY

Validation Strength: WEAK

Score-Adjusted method shows weak validation. Qualitative factors provide marginal improvements.

Overall Win Rate: 57.1% (Score-Adjusted wins)

Mean Improvement: -6.5%

Forecast Quality Scores:

- Base Model: 78.3/100
- Score-Adjusted Model: 78.7/100
- Improvement: +0.4 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 18.91%
- Median MAPE: 21.72%
- Mean sMAPE: 21.44%

Score-Adjusted Model:

- Mean MAPE: 18.79%
- Median MAPE: 22.59%
- Mean sMAPE: 21.12%

Win/Loss Record:

- Score-Adjusted Wins: 4
- Base Model Wins: 3
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 19.72%
- Score Mean MAPE: 19.30%
- Mean Improvement: -5.87%
- Score Win Rate: 75.0%
- Wins: Score=3, Base=1

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 17.84%
- Score Mean MAPE: 18.10%
- Mean Improvement: -7.41%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 0.5505
- P-Value: 0.6019
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 12.0000
- P-Value: 0.8125
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.2081
- Interpretation: Small effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 457,234,000.00

Base Forecast: 485,226,912.00 (Error: +27,992,912.00)

Score Forecast: 492,020,088.00 (Error: +34,786,088.00)

Accuracy:

- Base sMAPE: 5.94%
- Score sMAPE: 7.33%
- Improvement: -24.27%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 342,809,000.00

Base Forecast: 232,043,024.00 (Error: -110,765,976.00)

Score Forecast: 235,291,626.00 (Error: -107,517,374.00)

Accuracy:

- Base sMAPE: 38.54%
- Score sMAPE: 37.20%
- Improvement: +2.93%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 876,479,000.00

Base Forecast: 651,501,696.00 (Error: -224,977,304.00)

Score Forecast: 665,183,231.00 (Error: -211,295,769.00)

Accuracy:

- Base sMAPE: 29.45%
- Score sMAPE: 27.41%
- Improvement: +6.08%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 820,101,000.00
Base Forecast: 852,255,936.00 (Error: +32,154,936.00)
Score Forecast: 864,187,519.00 (Error: +44,086,519.00)

Accuracy:

- Base sMAPE: 3.85%
- Score sMAPE: 5.24%
- Improvement: -37.11%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 439,089,000.00
Base Forecast: 361,227,520.00 (Error: -77,861,480.00)
Score Forecast: 366,284,705.00 (Error: -72,804,295.00)

Accuracy:

- Base sMAPE: 19.46%
- Score sMAPE: 18.08%
- Improvement: +6.50%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -448,175,000.00
Base Forecast: -350,843,968.00 (Error: +97,331,032.00)
Score Forecast: -346,914,515.00 (Error: +101,260,485.00)

Accuracy:

- Base sMAPE: 24.36%
- Score sMAPE: 25.47%
- Improvement: -4.04%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 3,324,249,000.00
Base Forecast: 2,495,665,152.00 (Error: -828,583,848.00)
Score Forecast: 2,530,604,464.00 (Error: -793,644,536.00)

Accuracy:

- Base sMAPE: 28.47%
- Score sMAPE: 27.11%
- Improvement: +4.22%

Direction Correct: Base=True, Score=True

Winner: Score

=====

=====

RECOMMENDATIONS

=====

=====

△ Weak validation. Qualitative adjustments show marginal impact. Recommend:

1. Review and refine the qualitative scoring methodology
2. Investigate if certain metric types benefit more from adjustments
3. Consider whether the base model is already capturing qualitative factors
4. Test alternative adjustment algorithms

=====

=====

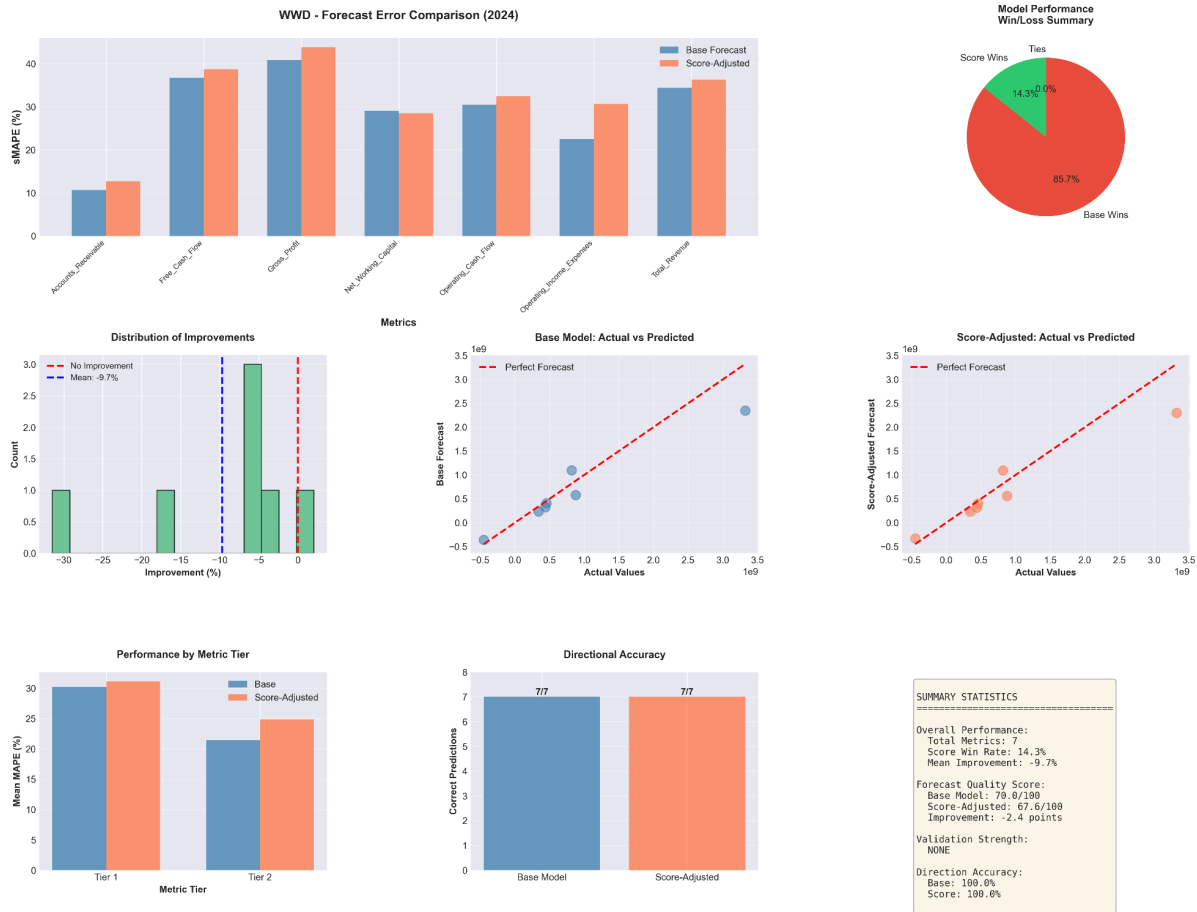
END OF REPORT

=====

=====

WWD (ARIMA):

Forecast Validation Analysis: WWD (2024)



FORECAST VALIDATION REPORT

Company: WWD
Forecast Year: 2024
Analysis Date: 2025-12-06 22:17:14

EXECUTIVE SUMMARY

=====

Validation Strength: NONE

Score-Adjusted method does not validate. Base model performs as well or better.

Overall Win Rate: 14.3% (Score-Adjusted wins)

Mean Improvement: -9.7%

Forecast Quality Scores:

- Base Model: 70.0/100
- Score-Adjusted Model: 67.6/100
- Improvement: -2.4 points

=====

DETAILED METRICS COMPARISON

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 26.45%
- Median MAPE: 29.33%
- Mean sMAPE: 29.25%

Score-Adjusted Model:

- Mean MAPE: 28.42%
- Median MAPE: 30.75%
- Mean sMAPE: 31.90%

Win/Loss Record:

- Score-Adjusted Wins: 1
- Base Model Wins: 6
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 30.21%
- Score Mean MAPE: 31.11%
- Mean Improvement: -3.22%
- Score Win Rate: 25.0%
- Wins: Score=1, Base=3

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 21.44%
- Score Mean MAPE: 24.83%
- Mean Improvement: -18.39%
- Score Win Rate: 0.0%
- Wins: Score=0, Base=3

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: -2.6353
- P-Value: 0.0388
- Significant at $\alpha=0.05$: True
- Interpretation: Score-Adjusted significantly different

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 1.0000
- P-Value: 0.0312
- Significant at $\alpha=0.05$: True

Effect Size:

- Cohen's d: -0.9961
- Interpretation: Large effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 457,234,000.00

Base Forecast: 410,952,851.00 (Error: -46,281,149.00)

Score Forecast: 402,681,109.00 (Error: -54,552,891.00)

Accuracy:

- Base sMAPE: 10.66%
- Score sMAPE: 12.69%
- Improvement: -17.87%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 342,809,000.00

Base Forecast: 236,401,729.00 (Error: -106,407,271.00)

Score Forecast: 231,643,388.00 (Error: -111,165,612.00)

Accuracy:

- Base sMAPE: 36.74%
- Score sMAPE: 38.70%
- Improvement: -4.47%

Direction Correct: Base=True, Score=True

Winner: Base

Gross_Profit (Tier 2):

Actual Value: 876,479,000.00

Base Forecast: 578,926,775.00 (Error: -297,552,225.00)

Score Forecast: 561,514,626.00 (Error: -314,964,374.00)

Accuracy:

- Base sMAPE: 40.89%
- Score sMAPE: 43.81%
- Improvement: -5.85%

Direction Correct: Base=True, Score=True

Winner: Base

Net_Working_Capital (Tier 1):

Actual Value: 820,101,000.00

Base Forecast: 1,099,018,046.00 (Error: +278,917,046.00)

Score Forecast: 1,093,326,232.00 (Error: +273,225,232.00)

Accuracy:

- Base sMAPE: 29.07%
- Score sMAPE: 28.56%
- Improvement: +2.04%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Cash_Flow (Tier 1):

Actual Value: 439,089,000.00

Base Forecast: 322,894,159.00 (Error: -116,194,841.00)

Score Forecast: 316,394,880.00 (Error: -122,694,120.00)

Accuracy:

- Base sMAPE: 30.50%
- Score sMAPE: 32.48%
- Improvement: -5.59%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Income_Expenses (Tier 2):

Actual Value: -448,175,000.00

Base Forecast: -357,417,394.00 (Error: +90,757,606.00)

Score Forecast: -328,869,823.00 (Error: +119,305,177.00)

Accuracy:

- Base sMAPE: 22.53%
- Score sMAPE: 30.71%
- Improvement: -31.45%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 3,324,249,000.00

Base Forecast: 2,349,301,678.00 (Error: -974,947,322.00)

Score Forecast: 2,302,014,464.00 (Error: -1,022,234,536.00)

Accuracy:

- Base sMAPE: 34.37%
- Score sMAPE: 36.34%
- Improvement: -4.85%

Direction Correct: Base=True, Score=True

Winner: Base

=====

=====

RECOMMENDATIONS

=====

=====

X No validation. Base model performs as well or better. Recommend:

1. Critically review the qualitative scoring framework
2. Check if qualitative factors are already reflected in historical data
3. Consider that market efficiency may limit qualitative alpha
4. Explore if timing or sector-specific factors affected results

=====

=====

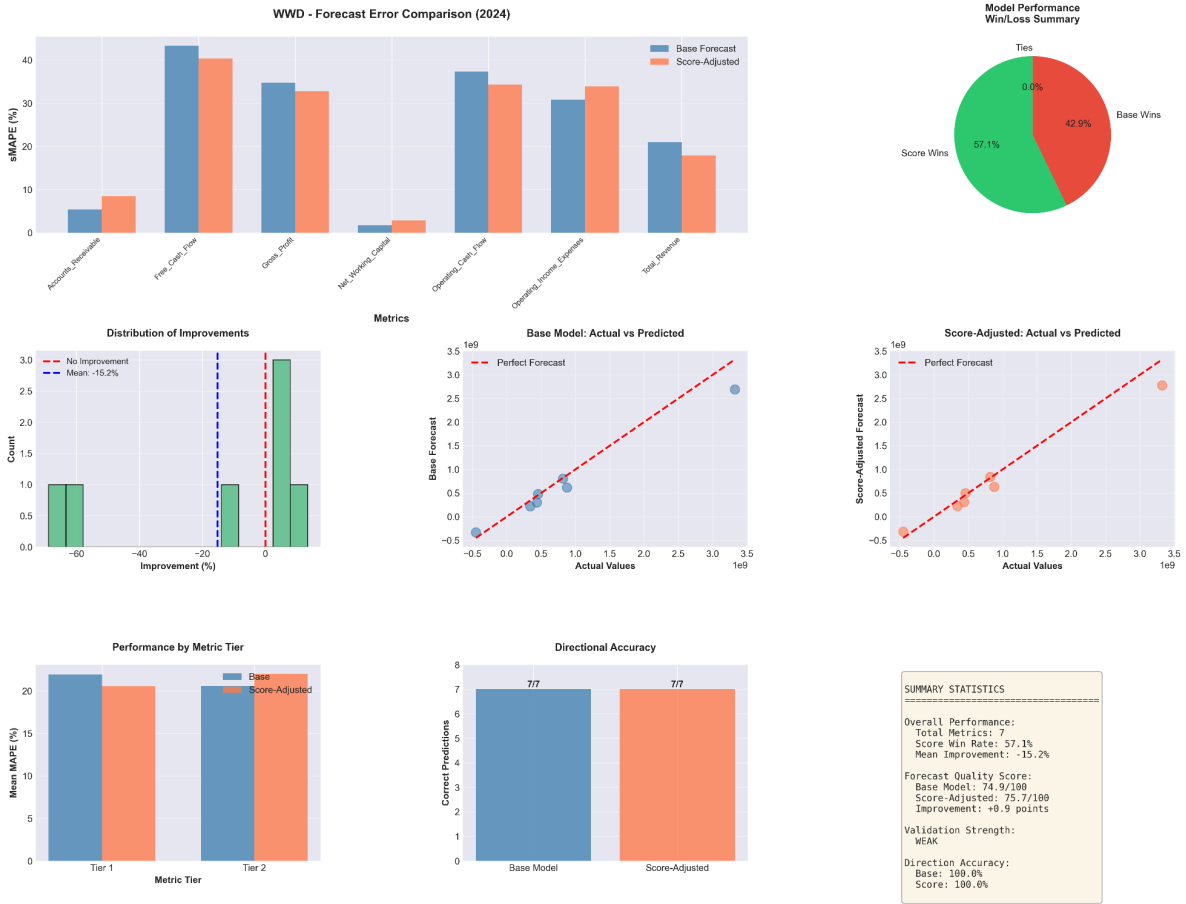
END OF REPORT

=====

=====

WWD (WINTERS-HOLT):

Forecast Validation Analysis: WWD (2024)



FORECAST VALIDATION REPORT

Company: WWD
Forecast Year: 2024
Analysis Date: 2025-12-06 22:26:48

EXECUTIVE SUMMARY

Validation Strength: WEAK

Score-Adjusted method shows weak validation. Qualitative factors provide marginal improvements.

Overall Win Rate: 57.1% (Score-Adjusted wins)

Mean Improvement: -15.2%

Forecast Quality Scores:

- Base Model: 74.9/100
- Score-Adjusted Model: 75.7/100
- Improvement: +0.9 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 21.37%
- Median MAPE: 26.69%
- Mean sMAPE: 24.90%

Score-Adjusted Model:

- Mean MAPE: 21.17%
- Median MAPE: 28.14%
- Mean sMAPE: 24.37%

Win/Loss Record:

- Score-Adjusted Wins: 4
- Base Model Wins: 3
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 21.94%
- Score Mean MAPE: 20.55%
- Mean Improvement: -10.71%
- Score Win Rate: 75.0%
- Wins: Score=3, Base=1

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 20.61%
- Score Mean MAPE: 21.99%
- Mean Improvement: -21.21%
- Score Win Rate: 33.3%
- Wins: Score=1, Base=2

STATISTICAL SIGNIFICANCE TESTS

Paired T-Test:

- T-Statistic: 0.4864
- P-Value: 0.6439
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 14.0000
- P-Value: 1.0000
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.1838
- Interpretation: Small effect

METRIC-BY-METRIC BREAKDOWN

Accounts_Receivable (Tier 2):

Actual Value: 457,234,000.00

Base Forecast: 482,529,825.00 (Error: +25,295,825.00)

Score Forecast: 497,702,492.00 (Error: +40,468,492.00)

Accuracy:

- Base sMAPE: 5.38%
- Score sMAPE: 8.48%
- Improvement: -59.98%

Direction Correct: Base=True, Score=True

Winner: Base

Free_Cash_Flow (Tier 1):

Actual Value: 342,809,000.00

Base Forecast: 220,792,795.00 (Error: -122,016,205.00)

Score Forecast: 227,735,404.00 (Error: -115,073,596.00)

Accuracy:

- Base sMAPE: 43.30%
- Score sMAPE: 40.34%
- Improvement: +5.69%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 876,479,000.00

Base Forecast: 616,910,581.00 (Error: -259,568,419.00)

Score Forecast: 629,848,429.00 (Error: -246,630,571.00)

Accuracy:

- Base sMAPE: 34.76%
- Score sMAPE: 32.75%
- Improvement: +4.98%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 820,101,000.00

Base Forecast: 805,955,696.00 (Error: -14,145,304.00)

Score Forecast: 843,980,686.00 (Error: +23,879,686.00)

Accuracy:

- Base sMAPE: 1.74%
- Score sMAPE: 2.87%
- Improvement: -68.82%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 439,089,000.00

Base Forecast: 300,952,726.00 (Error: -138,136,274.00)

Score Forecast: 310,415,883.00 (Error: -128,673,117.00)

Accuracy:

- Base sMAPE: 37.33%
- Score sMAPE: 34.34%
- Improvement: +6.85%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -448,175,000.00

Base Forecast: -328,557,966.00 (Error: +119,617,034.00)

Score Forecast: -318,226,790.00 (Error: +129,948,210.00)

Accuracy:

- Base sMAPE: 30.80%
- Score sMAPE: 33.91%
- Improvement: -8.64%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 3,324,249,000.00

Base Forecast: 2,693,454,829.00 (Error: -630,794,171.00)

Score Forecast: 2,778,147,823.00 (Error: -546,101,177.00)

Accuracy:

- Base sMAPE: 20.96%
- Score sMAPE: 17.90%
- Improvement: +13.43%

Direction Correct: Base=True, Score=True
Winner: Score

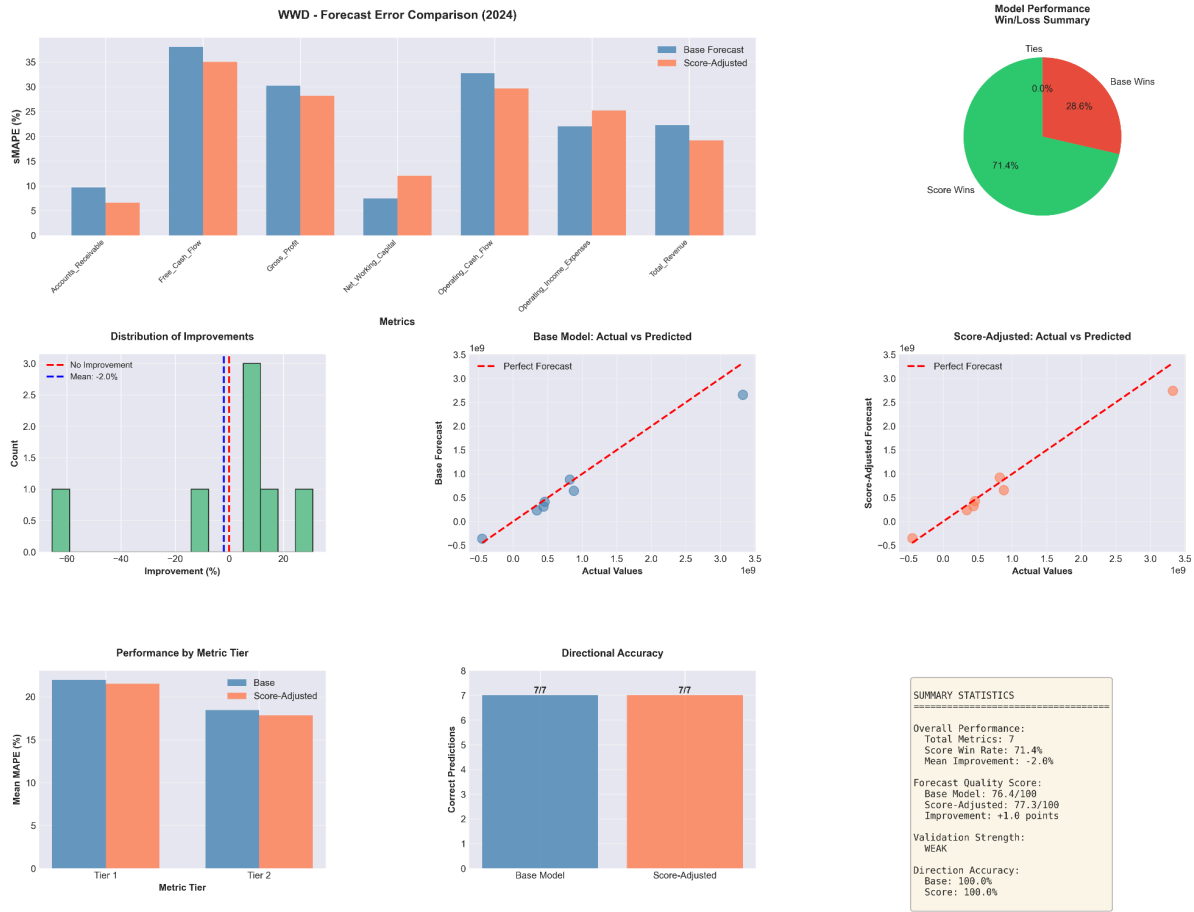
```
=====
=====
RECOMMENDATIONS
=====
=====
```

- △ Weak validation. Qualitative adjustments show marginal impact. Recommend:
1. Review and refine the qualitative scoring methodology
 2. Investigate if certain metric types benefit more from adjustments
 3. Consider whether the base model is already capturing qualitative factors
 4. Test alternative adjustment algorithms

```
=====
=====
END OF REPORT
=====
=====
```


WWD (LINEAR + WINTERS-HOLT + XGBOOST):

Forecast Validation Analysis: WWD (2024)



FORECAST VALIDATION REPORT

Company: WWD
Forecast Year: 2024
Analysis Date: 2025-12-06 22:32:30

EXECUTIVE SUMMARY

Validation Strength: WEAK

Score-Adjusted method shows weak validation. Qualitative factors provide marginal improvements.

Overall Win Rate: 71.4% (Score-Adjusted wins)

Mean Improvement: -2.0%

Forecast Quality Scores:

- Base Model: 76.4/100
- Score-Adjusted Model: 77.3/100
- Improvement: +1.0 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 20.45%
- Median MAPE: 20.02%
- Mean sMAPE: 23.20%

Score-Adjusted Model:

- Mean MAPE: 19.92%
- Median MAPE: 22.37%
- Mean sMAPE: 22.27%

Win/Loss Record:

- Score-Adjusted Wins: 5
- Base Model Wins: 2
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 21.96%
- Score Mean MAPE: 21.50%
- Mean Improvement: -9.52%
- Score Win Rate: 75.0%
- Wins: Score=3, Base=1

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 18.44%
- Score Mean MAPE: 17.82%
- Mean Improvement: 8.02%
- Score Win Rate: 66.7%
- Wins: Score=2, Base=1

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 0.7334
- P-Value: 0.4910
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 13.0000
- P-Value: 0.9375
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.2772
- Interpretation: Small effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 457,234,000.00

Base Forecast: 414,955,340.00 (Error: -42,278,660.00)

Score Forecast: 428,003,195.00 (Error: -29,230,805.00)

Accuracy:

- Base sMAPE: 9.69%
- Score sMAPE: 6.60%
- Improvement: +30.86%

Direction Correct: Base=True, Score=True

Winner: Score

Free_Cash_Flow (Tier 1):

Actual Value: 342,809,000.00

Base Forecast: 233,292,369.00 (Error: -109,516,631.00)

Score Forecast: 240,628,014.00 (Error: -102,180,986.00)

Accuracy:

- Base sMAPE: 38.02%
- Score sMAPE: 35.03%
- Improvement: +6.70%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 876,479,000.00

Base Forecast: 646,548,135.00 (Error: -229,930,865.00)

Score Forecast: 660,107,542.00 (Error: -216,371,458.00)

Accuracy:

- Base sMAPE: 30.19%
- Score sMAPE: 28.16%
- Improvement: +5.90%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 820,101,000.00
Base Forecast: 883,859,089.00 (Error: +63,758,089.00)
Score Forecast: 925,559,561.00 (Error: +105,458,561.00)

Accuracy:

- Base sMAPE: 7.48%
- Score sMAPE: 12.08%
- Improvement: -65.40%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 439,089,000.00
Base Forecast: 315,710,096.00 (Error: -123,378,904.00)
Score Forecast: 325,637,284.00 (Error: -113,451,716.00)

Accuracy:

- Base sMAPE: 32.69%
- Score sMAPE: 29.67%
- Improvement: +8.05%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -448,175,000.00
Base Forecast: -359,218,715.00 (Error: +88,956,285.00)
Score Forecast: -347,923,442.00 (Error: +100,251,558.00)

Accuracy:

- Base sMAPE: 22.04%
- Score sMAPE: 25.19%
- Improvement: -12.70%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 3,324,249,000.00
Base Forecast: 2,658,724,105.00 (Error: -665,524,895.00)
Score Forecast: 2,742,325,026.00 (Error: -581,923,974.00)

Accuracy:

- Base sMAPE: 22.25%
- Score sMAPE: 19.18%
- Improvement: +12.56%

Direction Correct: Base=True, Score=True

Winner: Score

```
=====
=====
RECOMMENDATIONS
=====
=====
```

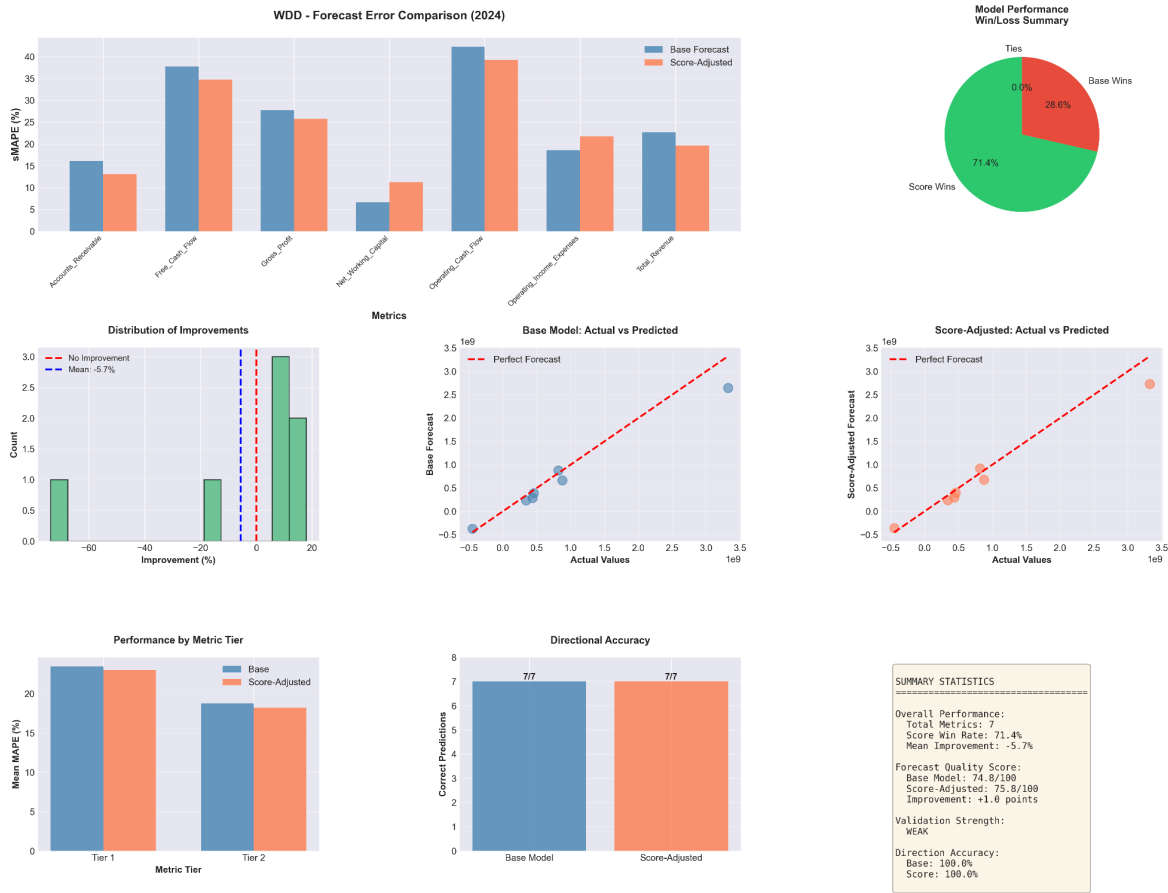
△ Weak validation. Qualitative adjustments show marginal impact. Recommend:

1. Review and refine the qualitative scoring methodology
2. Investigate if certain metric types benefit more from adjustments
3. Consider whether the base model is already capturing qualitative factors
4. Test alternative adjustment algorithms

```
=====
=====
END OF REPORT
=====
=====
```


WDD (LINEAR + WINTERS-HOLT + ARIMA):

Forecast Validation Analysis: WDD (2024)



FORECAST VALIDATION REPORT

Company: WDD
Forecast Year: 2024
Analysis Date: 2025-12-06 22:35:44

EXECUTIVE SUMMARY

Validation Strength: WEAK

Score-Adjusted method shows weak validation. Qualitative factors provide marginal improvements.

Overall Win Rate: 71.4% (Score-Adjusted wins)

Mean Improvement: -5.7%

Forecast Quality Scores:

- Base Model: 74.8/100
- Score-Adjusted Model: 75.8/100
- Improvement: +1.0 points

=====

=====

DETAILED METRICS COMPARISON

=====

=====

Total Metrics Analyzed: 7

Accuracy Metrics:

Base Model:

- Mean MAPE: 21.46%
- Median MAPE: 20.36%
- Mean sMAPE: 24.54%

Score-Adjusted Model:

- Mean MAPE: 20.98%
- Median MAPE: 19.62%
- Mean sMAPE: 23.63%

Win/Loss Record:

- Score-Adjusted Wins: 5
- Base Model Wins: 2
- Ties: 0

Directional Accuracy:

- Base Model: 100.0%
- Score-Adjusted: 100.0%

=====

=====

PERFORMANCE BY METRIC TIER

=====

Tier 1 (Critical) Metrics:

- Count: 4
- Base Mean MAPE: 23.46%
- Score Mean MAPE: 23.05%
- Mean Improvement: -12.19%
- Score Win Rate: 75.0%
- Wins: Score=3, Base=1

Tier 2 (Important) Metrics:

- Count: 3
- Base Mean MAPE: 18.78%
- Score Mean MAPE: 18.23%
- Mean Improvement: 3.02%
- Score Win Rate: 66.7%
- Wins: Score=2, Base=1

=====

STATISTICAL SIGNIFICANCE TESTS

=====

Paired T-Test:

- T-Statistic: 0.7260
- P-Value: 0.4951
- Significant at $\alpha=0.05$: False
- Interpretation: No significant difference

Wilcoxon Signed-Rank Test (Non-Parametric):

- W-Statistic: 13.0000
- P-Value: 0.9375
- Significant at $\alpha=0.05$: False

Effect Size:

- Cohen's d: 0.2744
- Interpretation: Small effect

=====

METRIC-BY-METRIC BREAKDOWN

=====

Accounts_Receivable (Tier 2):

Actual Value: 457,234,000.00

Base Forecast: 388,933,146.00 (Error: -68,300,854.00)

Score Forecast: 401,162,760.00 (Error: -56,071,240.00)

Accuracy:

- Base sMAPE: 16.14%

- Score sMAPE: 13.06%

- Improvement: +17.91%

Direction Correct: Base=True, Score=True

Winner: Score

Free_Cash_Flow (Tier 1):

Actual Value: 342,809,000.00

Base Forecast: 233,909,949.00 (Error: -108,899,051.00)

Score Forecast: 241,265,014.00 (Error: -101,543,986.00)

Accuracy:

- Base sMAPE: 37.77%

- Score sMAPE: 34.77%

- Improvement: +6.75%

Direction Correct: Base=True, Score=True

Winner: Score

Gross_Profit (Tier 2):

Actual Value: 876,479,000.00

Base Forecast: 662,650,406.00 (Error: -213,828,594.00)

Score Forecast: 676,547,511.00 (Error: -199,931,489.00)

Accuracy:

- Base sMAPE: 27.79%

- Score sMAPE: 25.75%

- Improvement: +6.50%

Direction Correct: Base=True, Score=True

Winner: Score

Net_Working_Capital (Tier 1):

Actual Value: 820,101,000.00
Base Forecast: 876,204,187.00 (Error: +56,103,187.00)
Score Forecast: 917,543,501.00 (Error: +97,442,501.00)

Accuracy:

- Base sMAPE: 6.61%
- Score sMAPE: 11.22%
- Improvement: -73.68%

Direction Correct: Base=True, Score=True

Winner: Base

Operating_Cash_Flow (Tier 1):

Actual Value: 439,089,000.00
Base Forecast: 285,989,839.00 (Error: -153,099,161.00)
Score Forecast: 294,982,503.00 (Error: -144,106,497.00)

Accuracy:

- Base sMAPE: 42.23%
- Score sMAPE: 39.26%
- Improvement: +5.87%

Direction Correct: Base=True, Score=True

Winner: Score

Operating_Income_Expenses (Tier 2):

Actual Value: -448,175,000.00
Base Forecast: -371,918,205.00 (Error: +76,256,795.00)
Score Forecast: -360,223,609.00 (Error: +87,951,391.00)

Accuracy:

- Base sMAPE: 18.60%
- Score sMAPE: 21.76%
- Improvement: -15.34%

Direction Correct: Base=True, Score=True

Winner: Base

Total_Revenue (Tier 1):

Actual Value: 3,324,249,000.00
Base Forecast: 2,647,288,017.00 (Error: -676,960,983.00)
Score Forecast: 2,730,529,341.00 (Error: -593,719,659.00)

Accuracy:

- Base sMAPE: 22.67%
- Score sMAPE: 19.61%
- Improvement: +12.30%

Direction Correct: Base=True, Score=True

Winner: Score

=====

=====

RECOMMENDATIONS

=====

=====

△ Weak validation. Qualitative adjustments show marginal impact. Recommend:

1. Review and refine the qualitative scoring methodology
2. Investigate if certain metric types benefit more from adjustments
3. Consider whether the base model is already capturing qualitative factors
4. Test alternative adjustment algorithms

=====

=====

END OF REPORT

=====

=====