

**Modelo híbrido ARIMA–XGBoost para la predicción de precios y generar señales de
compraventa de activos en mercados bursátiles internacionales**

Juan Sebastian Herrera Beleño

Asesor

Esneider Dejesus Pineda Martínez

Universidad Nacional Abierta y a Distancia UNAD

Escuela de Ciencias Básicas, Tecnología e Ingeniería ECBTI

Especialización en Ciencia de Datos y Analítica

2026

Resumen

Se desarrolló un pipeline integral que incluyó descarga de datos históricos, limpieza, transformación, ingeniería de características y modelado mediante ARIMA y XGBoost. Ambos modelos fueron evaluados con RMSE, MAE y MAPE, y sus predicciones se tradujeron en un sistema de señales validado mediante *backtesting* frente a Buy & Hold. XGBoost mostró un desempeño predictivo superior (MAPE $\approx 2.7\%$), superando ampliamente a ARIMA. No obstante, el modelo híbrido con ponderación fija no mejoró estos resultados, evidenciando que un ensamble mal balanceado puede degradar al componente más fuerte. A pesar de la alta precisión estadística, la estrategia de *trading* generó pérdidas, principalmente por *sobre-trading*, ausencia de filtros de tendencia y desalineación entre métricas predictivas y objetivos financieros. Se concluye que optimizar solo el error no garantiza rentabilidad y se requiere incorporar gestión de riesgo, control operativo y costos transaccionales.

Palabras claves: Maching learning, series de tiempo financieras, backtesting, ARIMA-XGBosst.

Abstract

A comprehensive data pipeline was developed that included historical data download, cleaning, transformation, feature engineering, and modeling using ARIMA and XGBoost. Both models were evaluated using RMSE, MAE, and MAPE, and their predictions were translated into a signal system validated by backtesting against Buy & Hold. XGBoost showed superior predictive performance (MAPE $\approx 2.7\%$), far outperforming ARIMA. However, the fixed-weight hybrid model did not improve these results, demonstrating that a poorly balanced ensemble can degrade the strongest component. Despite high statistical accuracy, the trading strategy generated losses, mainly due to overtrading, lack of trend filters, and misalignment between predictive metrics and financial objectives. It is concluded that optimizing error alone does not guarantee profitability and that risk management, operational control, and transaction costs must be incorporated.

Key words: Machine learning, financial time series, backtesting, ARIMA- XGBosst.

Tabla de Contenido

Introducción	12
Descripción del Problema	15
Problema Fundamental.....	15
Desafíos Identificados	15
Limitaciones de Modelos Tradicionales.....	15
Riesgo de Overfitting en Machine Learning	16
Desconexión Predicción-Rentabilidad	16
Optimización de Métrica Inadecuada	16
Frecuencia de Trading Excesiva	16
Contexto Específico	16
Vacío en la Literatura	17
Relevancia.....	17
Planteamiento del Problema	18
Sistematización del Problema.....	19
Problema de Recolección y Preprocesamiento de Datos	19
Subproblema de Recolección y Preprocesamiento de Datos	19
Sistematización de Recolección y Preprocesamiento de Datos	19
Problema de Modelado ARIMA.....	19
Subproblema Modelado ARIMA	19
Sistematización Modelado ARIMA	19
Problema de Modelado XGBoost.....	20
Subproblema del Modelado XGBoost	20

Sistematización Modelado XGBoost	20
Problema de Combinación Híbrida.....	21
Subproblema Combinación Híbrida	21
Sistematización Combinación Híbrida	21
Problema de Generación de Señales	21
Subproblema de Generación de Señales	21
Sistematización de Generación de Señales	21
Problema de Backtesting.....	22
Subproblema de Backtesting.....	22
Sistematización de Backtesting.....	22
Problema de Análisis de Causas de Underperformance.....	22
Subproblema de Análisis de Causas de Underperformance	22
Sistematización de Análisis de Causas de Underperformance	22
Justificación	23
Justificación Teórica	23
Contribución Al Conocimiento Científico	23
Demuestra Empíricamente la Paradoja Precisión-Rentabilidad.....	23
Identifica y Cuantifica Factores Explicativos	23
Evalúa Validez de Modelos Híbridos en caso Asimétrico	23
Cuestiona Métricas Tradicionales de Evaluación.....	24
Justificación Práctica	24
Aplicabilidad en Mercados Financieros Reales	24
Objetivos.....	25

Objetivo General	25
Objetivos Específicos.....	25
Marco de Referencia	26
Estado del Arte	26
Modelos ARIMA en Predicción Bursátil.....	26
XGBoost y Machine Learning en Finanzas	26
Modelos Híbridos Estadísticos-ML.....	27
Trading Algorítmico y Backtesting	28
Posicionamiento del Presente Estudio.....	28
Marco Contextual.....	29
Contexto del Mercado Financiero	29
Marco Teórico	30
Teoría de Mercados Eficientes (EMH).....	30
Teoría de Random Walk	31
Teoría de Series Temporales.....	31
Teoría de Ensemble Learning	32
Teoría de Optimización y Trading.....	33
Marco Conceptual	33
Marco Normativo	35
Regulación de Mercados Financieros.....	35
Normativa Académica	36
Ética en Trading Algorítmico	36
Metodología	38

Método	38
Tipo de Estudio	38
Según Objetivo: Estudio Descriptivo-Explicativo (Proyecto Aplicado).....	38
Según Naturaleza De Datos: Investigación Cuantitativa Ex-Post-Facto.....	39
Según Temporalidad: Estudio Longitudinal Retrospectivo.....	39
Según Profundidad: Estudio Aplicado	39
Recolección De Datos	39
Fuente de Datos	39
Período Temporal	40
Variables Recolectadas	40
Preprocesamiento de Datos	40
Ingeniería de Características	41
Validación de Datos.....	42
Resultados.....	44
Resultados del Modelo ARIMA.....	44
Identificación y Especificación del Modelo.....	44
Estimación de Parámetro)	45
Diagnóstico de Residuales.....	45
Predicciones y Métricas de Error	46
Resultados del Modelo XGBoost.....	48
Configuración y Entrenamiento	48
Importancia de Features	49
Predicciones y Métricas de Error	50

Análisis de Predicciones.....	51
Resultados del Modelo Híbrido (ARIMA + Algoritmo de Extrema Gradiente).....	52
Esquema de Combinación.....	52
Métricas de Error.....	53
Análisis del Ensemble	54
Resultados del Backtesting	54
Generación de Señales.....	54
Simulación.....	56
Conclusiones.....	58
Recomendaciones	62
Referencias bibliográficas.....	63
Apéndices.....	67

Lista de Tablas

Tabla 1 <i>Coefficientes Estimados (sobre log-Precios Diferenciados)</i>	45
Tabla 2 <i>Métricas de Desempeño - ARIMA</i>	46
Tabla 3 <i>Hiperparámetros Utilizados para XGBoost</i>	48
Tabla 4 <i>Métricas de Desempeño – XGBoost</i>	51
Tabla 5 <i>Métricas de Desempeño - Modelo Híbrido</i>	53
Tabla 6 <i>Distribución de las Señales</i>	55
Tabla 7 <i>Resultados de la Simulación</i>	57

Lista de Figuras

Figura 1 <i>Predicción del Modelo ARIMA</i>	47
Figura 2 <i>Hiperparámetros Utilizados para XGBoost</i>	50
Figura 3 <i>Predicción Todos los Modelos</i>	54
Figura 4 Sistema de Señales	55
Figura 5 <i>Simulación</i>	56

Lista de Apéndices

Apéndice A <i>Código Python Para el Desarrollo del Proyecto</i>	67
Apéndice B <i>Análisis de Residuales y Prueba de Residuales Para el Modelo ARIMA</i>	67
Apéndice C <i>Análisis de Residuales y Prueba de Residuales Para el Modelo ARIMA</i>	67
Apéndice D <i>Análisis de Residuales y Prueba de Residuales Para el Modelo ARIMA</i>	68
Apéndice E <i>Entrenamiento y Predicción de Modelo ARIMA</i>	69
Apéndice F <i>Entrenamiento y Predicción de Modelo ARIMA</i>	69
Apéndice G <i>Comparación de los Modelos</i>	70
Apéndice H <i>Enlace de Sustentación</i>	70
Apéndice I <i>Reporte Turnitin</i>	70

Introducción

La predicción de precios en mercados financieros constituye uno de los desafíos más complejos en el ámbito de las finanzas cuantitativas y el aprendizaje automático. La naturaleza estocástica, no lineal y heteroscedástica de las series financieras ha motivado décadas de investigación orientada al desarrollo de modelos predictivos cada vez más sofisticados. Los mercados bursátiles presentan comportamientos caracterizados por alta volatilidad, clustering de varianza, colas pesadas en las distribuciones de retornos, y efectos simultáneos de momentum y reversión a la media, fenómenos que dificultan considerablemente la modelización precisa de sus movimientos.

Los modelos econométricos tradicionales, particularmente la familia ARIMA (AutoRegressive Integrated Moving Average), han sido ampliamente utilizados desde la década de 1970 para el pronóstico de series temporales financieras. Sin embargo, estos modelos presentan limitaciones estructurales fundamentales: asumen relaciones lineales entre observaciones, requieren transformaciones que pueden eliminar información relevante, y utilizan exclusivamente la historia de precios ignorando variables exógenas potencialmente predictivas como volumen de transacciones, indicadores técnicos, o factores macroeconómicos.

Paralelamente, el avance exponencial del machine learning en la última década ha introducido algoritmos capaces de capturar relaciones no lineales complejas. Entre estos, XGBoost (Extreme Gradient Boosting) ha demostrado superioridad consistente en competencias de ciencia de datos y aplicaciones financieras, ofreciendo capacidad para incorporar múltiples variables explicativas, adaptarse a regímenes de mercado cambiantes, y reducir overfitting mediante técnicas avanzadas de regularización.

La literatura reciente en finanzas cuantitativas sugiere que los modelos híbridos, que combinan las fortalezas complementarias de métodos estadísticos tradicionales y algoritmos de machine learning, pueden superar el desempeño de modelos individuales. La hipótesis subyacente es que ARIMA captura la estructura temporal de largo plazo inherente a las series de precios, mientras que XGBoost incorpora información de corto plazo y relaciones no lineales entre múltiples features, resultando en predicciones más robustas y precisas.

El presente trabajo de grado desarrolla e implementa un sistema híbrido ARIMA-XGBoost para la predicción de precios bursátiles y la generación automatizada de señales de compraventa (trading signals). El sistema propuesto combina predicciones de ambos modelos mediante ponderación optimizada, traduce estos pronósticos en señales operativas de compra, venta o mantener posición, y evalúa su rentabilidad real mediante backtesting riguroso que incluye costos de transacción y comparación con estrategia pasiva de Buy & Hold.

A diferencia de estudios previos que se enfocan exclusivamente en métricas de error predictivo (RMSE, MAE, MAPE), esta investigación prioriza la evaluación de rentabilidad financiera real, reconociendo que la excelencia en precisión estadística no garantiza necesariamente éxito en trading algorítmico. Los resultados revelan una paradoja fundamental: mientras que XGBoost alcanza métricas predictivas excelentes (MAPE: 2.70%), la estrategia de trading basada en estas predicciones registra underperformance significativo versus estrategia pasiva, evidenciando la brecha crítica entre optimización predictiva y optimización financiera.

Esta investigación contribuye al campo emergente de AI aplicada a finanzas mediante la implementación reproducible de pipeline completo desde datos crudos hasta evaluación de rentabilidad, la identificación y cuantificación de factores explicativos de la desconexión predicción-rentabilidad, el análisis crítico de limitaciones de métricas tradicionales de evaluación

de modelos, y la propuesta de líneas de investigación futuras orientadas a la reformulación de funciones objetivo hacia métricas financieras directas.

El documento se estructura en trece secciones que abarcan desde la formulación del problema de investigación hasta resultados empíricos, discusión crítica, conclusiones y recomendaciones para trabajo futuro.

Descripción del Problema

Los mercados financieros modernos generan volúmenes masivos de datos diariamente: precios de apertura y cierre, máximos y mínimos intradía, volúmenes de transacción, y centenares de indicadores técnicos derivados. Esta abundancia informativa, paradójicamente, incrementa la dificultad de identificar patrones predictivos genuinos versus ruido estocástico inherente a procesos de precios que exhiben propiedades de *random walk* con *drift*.

Problema Fundamental

A pesar de décadas de investigación en econometría financiera y el reciente auge de técnicas de *machine learning*, persiste la pregunta: ¿Es posible predecir consistentemente movimientos de precios bursátiles con precisión suficiente para generar estrategias de trading rentables que superen *benchmarks* pasivos como *Buy & Hold*?

Desafíos Identificados

Complejidad de las Series Financieras. Los precios bursátiles exhiben no linealidades, cambios de régimen (*bull markets* vs *bear markets*), heteroscedasticidad condicional (volatilidad variable), y sensibilidad extrema a eventos externos (anuncios económicos, eventos geopolíticos, shocks de mercado).

Limitaciones de Modelos Tradicionales

ARIMA y sus extensiones (ARIMA-GARCH) capturan componentes lineales y estructuras autorregresivas, pero fallan en modelar relaciones no lineales complejas y dinámicas de alta frecuencia. Su dependencia exclusiva en historia de precios ignora información valiosa de volumen, *momentum*, y sentimiento de mercado.

Riesgo de Overfitting en Machine Learning

Algoritmos como redes neuronales o árboles de decisión pueden memorizar patrones espurios en datos históricos que no se replican en *trading* real (*out-of-sample*), resultando en *backtest* optimista, pero performance pobre en producción.

Desconexión Predicción-Rentabilidad

La literatura académica reporta modelos con métricas impresionantes (RMSE bajo, R^2 elevado), pero rara vez evalúa si estas predicciones se traducen en ganancias reales al considerarse costos de transacción, *slippage*, y limitaciones operativas de ejecución.

Optimización de Métrica Inadecuada

Los modelos típicamente minimizan *Mean Squared Error* (MSE) del precio predicho, pero el objetivo financiero real es maximizar retorno ajustado por riesgo (*Sharpe ratio*) o minimizar *drawdown*, métricas no equivalentes a MSE.

Frecuencia de Trading Excesiva

Sistemas automatizados pueden generar centenares de señales de compra/venta, erosionando capital mediante comisiones y *spreads bid-ask*, especialmente problemático en activos apalancados o de baja liquidez.

Contexto Específico

El problema se agudiza al considerar activos de alta volatilidad como ETFs apalancados (por ejemplo, TQQQ que replica 3x el NASDAQ-100), donde movimientos diarios de 5-15% son comunes, y modelos conservadores que predicen cambios de 2-3% sistemáticamente subestiman magnitud de tendencias, generando señales contraproducentes.

Vacío en la Literatura

Si bien existen múltiples estudios sobre modelos híbridos ARIMA-ML que reportan mejoras en métricas de error, existe escasez crítica de investigaciones que evalúen rentabilidad real mediante *backtesting* con costos, analicen causas raíz de *underperformance* cuando ocurre, y propongan reformulaciones metodológicas hacia optimización directa de objetivos financieros.

Relevancia

Resolver esta problemática tiene implicaciones teóricas (comprensión de límites de predictibilidad de mercados) y prácticas (desarrollo de sistemas algorítmicos rentables para inversionistas individuales e institucionales).

Planteamiento del Problema

Con base en la descripción anterior, se formula el siguiente problema de investigación:

¿Puede un modelo híbrido de ML que combínelos modelos ARIMA (para capturar estructura temporal lineal) y XGBoost (para capturar relaciones no lineales multi-feature) generar predicciones de precios bursátiles suficientemente precisas para construir un sistema de señales de trading que supere el desempeño de una estrategia pasiva de Buy & Hold (comprar y mantener)?

Sistematización del Problema

La sistematización del problema de investigación se estructura mediante la descomposición en subproblemas específicos y la definición de metodologías particulares para cada uno.

Problema de Recolección y Preprocesamiento de Datos

Subproblema de Recolección y Preprocesamiento de Datos

¿Cómo obtener, limpiar y transformar datos financieros crudos en formato adecuado para modelado estadístico y *machine learning*?

Sistematización de Recolección y Preprocesamiento de Datos

1. Descarga automatizada desde Yahoo Finance API vía biblioteca *yfinance*
2. Manejo de valores nulos, verificación de continuidad temporal.
3. Transformaciones: logaritmos (estabilizar varianza), diferenciación (inducir estacionariedad).
4. Ingeniería de features: construcción de 25 variables técnicas (lags de retornos, medias móviles, RSI, momentum, volatilidad).
5. División temporal train/test preservando orden cronológico (80/20).

Problema de Modelado ARIMA

Subproblema Modelado ARIMA

¿Cómo identificar el orden óptimo (p,d,q) del modelo ARIMA y validar cumplimiento de supuestos?

Sistematización Modelado ARIMA

1. Análisis de estacionariedad: Prueba *Augmented Dickey-Fuller* (ADF).

2. Identificación de orden autorregresivo (p): Análisis de PACF (*Partial Autocorrelation Function*).
3. Identificación de orden de diferenciación (d): Iterativo hasta lograr estacionariedad.
4. Identificación de orden de media móvil (q): Análisis de ACF (*Autocorrelation Function*).
5. Estimación de máxima verosimilitud.
6. Diagnóstico por medio de análisis de residuales (media cero, no autocorrelación, distribución aproximadamente normal).
7. Análisis de métricas AIC, BIC para comparación de modelos alternativos.

Problema de Modelado XGBoost

Subproblema del Modelado XGBoost

¿Cómo configurar hiperparámetros de XGBoost para maximizar capacidad predictiva minimizando *overfitting*?

Sistematización Modelado XGBoost

1. Selección de *features*: 25 variables técnicas, análisis de importancia post-entrenamiento.
2. Configuración conservadora: `max_depth=3`, `learning_rate=0.01`, `n_estimators=200`.
3. Regularización: L1 (`reg_alpha=0.1`), L2 (`reg_lambda=1.0`), `gamma` (penalización por split).
4. Manejo de valores faltantes: Imputación por mediana.
5. Target: Retorno del día siguiente `Returns.shift(-1)`.

6. Validación: Eval set interno para monitorear convergencia.

Problema de Combinación Híbrida

Subproblema Combinación Híbrida

¿Cuál es el esquema óptimo de ponderación de predicciones ARIMA y XGBoost?

Sistematización Combinación Híbrida

1. Método base: Promedio ponderado lineal: $P^h = w_A P^A + w_{XGB} P^{XGB}$.
2. Ponderaciones iniciales: $w_A = 0.6$, $w_{XGB} = 0.4$ (heurístico).
3. Evaluación: Comparación de RMSE/MAE/MAPE del híbrido vs individuales.
4. Extensión futura: Optimización de pesos mediante validación cruzada o

aprendizaje como función de volatilidad.

Problema de Generación de Señales

Subproblema de Generación de Señales

¿Cómo traducir predicciones continuas de precio en señales discretas de *trading* minimizando ruido?

Sistematización de Generación de Señales

1. Detección de tendencia: Cambio porcentual en predicción.
2. Sistema de votación: Ventana móvil de 3 días, threshold de 2/3 coincidencias.
3. Señales: +1 (COMPRA), -1 (VENTA), 0 (MANTENER).
4. Mecanismos de *fallback*: Comparación con SMA, uso de cambios en precio real si

predicciones constantes.

5. Filtros: Sin implementación de umbrales mínimos.

Problema de Backtesting

Subproblema de Backtesting

¿Cómo simular ejecución realista de estrategia evaluando rentabilidad ajustada por riesgo?

Sistematización de Backtesting

1. Capital inicial: \$10,000
2. Lógica de posición: Binaria (100% invertido si señal +1, 0% si señal -1).
3. Costos: Implícitos en *spread* (no modelados explícitamente - limitación).
4. *Benchmark: Buy & Hold* (comprar al inicio, vender al final).
5. Métricas: Retorno total (%), capital final (\$), frecuencia de *trading*.
6. Limitaciones reconocidas: No incluye *slippage* explícito, no modela *short-selling*.

Problema de Análisis de Causas de Underperformance

Subproblema de Análisis de Causas de Underperformance

Si estrategia tiene MAPE bajo, pero pierde dinero, ¿cuáles son factores cuantificables explicativos?

Sistematización de Análisis de Causas de Underperformance

1. Análisis de captura de magnitud: Ratio amplitud predicha / amplitud real.
2. Cuantificación de *overtrading*: Número de señales, estimación de costos acumulados.
3. Evaluación de horizonte: Correlación entre precisión $t + 1$ y retornos: $t + k$.
4. Caracterización del activo: Volatility decay en ETFs apalancados.
5. Diagnóstico de sistema de señales: *Timing* de compras/ventas relativo a tendencias.

Justificación

Justificación Teórica

Contribución Al Conocimiento Científico

Este estudio aborda una brecha crítica en la literatura de *machine learning* aplicado a finanzas. Mientras que múltiples investigaciones reportan modelos con métricas predictivas ($R^2 > 0.90$, MAPE $< 5\%$), existe escasez de análisis rigurosos sobre traducción de estas predicciones en rentabilidad de *trading real*. La presente investigación:

Demuestra Empíricamente la Paradoja Precisión-Rentabilidad

Mediante cuantificación rigurosa, evidencia que un modelo con MAPE de 2.70% (clasificación académica "excelente") puede generar estrategia con pérdida de 30%, documentando discrepancia de -85 puntos porcentuales versus *benchmark* pasivo.

Identifica y Cuantifica Factores Explicativos

A diferencia de estudios que reportan underperformance sin análisis causal, este trabajo descompone el problema en: (a) subestimación de magnitud (captura 37% de movimientos), (b) overtrading (erosión ~20% por costos), (c) desajuste temporal (correlación 0.34 entre predicción diaria y retorno multi-día), (d) Características del activo (volatility decay de ETFs apalancados).

Evalúa Validez de Modelos Híbridos en caso Asimétrico

La literatura asume que *ensembles* mejoran *performance* cuando ambos modelos tienen calidad comparable. Este estudio documenta que con *ratio* de calidad 6.7:1 (RMSE ARIMA/RMSE XGBoost), la combinación simple no solo falla en mejorar, sino que degrada resultado respecto al modelo superior, contribuyendo a teoría de *ensemble learning*.

Cuestiona Métricas Tradicionales de Evaluación

Argumenta, con evidencia empírica, que RMSE/MAE/MAPE son insuficientes para evaluación de modelos destinados a *trading*, proponiendo reformulación hacia optimización directa de métricas financieras (*Sharpe ratio*, *maximum drawdown*, *profit factor*).

Justificación Práctica

Aplicabilidad en Mercados Financieros Reales

Orientación para Inversionistas Algorítmicos. Documenta limitaciones de estrategias basadas en predicción de precios, alertando sobre riesgos de *overtrading* y erosión por costos. Proporciona *framework* para evaluación realista de sistemas automatizados antes de capital real en riesgo.

Identificación de Condiciones de Fallo. Establece que activos apalancados 3x (TQQQ, UPRO, SQQQ) son inadecuados para estrategias de *swing trading* multi-día debido a *volatility decay* y amplificación de errores, orientando selección apropiada de activos para *trading* algorítmico.

Propuesta de Mejoras Implementables. Sugiere modificaciones específicas con potencial de mejora cuantificable:

1. Filtros de señal con umbrales $> 2\%$ (reducción esperada de señales: 70%)
2. *Holding periods* mínimos de 5 días
3. Reformulación de *target* para predecir retorno acumulado 5-día en lugar de t+1
4. Ajuste de pesos híbridos a 10% ARIMA / 90% XGBoost
5. Pipeline reproducible y extensible con código *open-source* documentado permite a

practicantes replicar análisis en otros activos (SPY, QQQ, acciones individuales) y comparar resultados, facilitando investigación futura y validación de hallazgos.

Objetivos

Objetivo General

Desarrollar un modelo híbrido ARIMA–XGBoost para predecir precios de activos financieros y generar señales automáticas de compra y venta en mercados bursátiles internacionales.

Objetivos Específicos

Realizar la descarga de datos históricos de los activos bursátiles desde Yahoo Finance e implementar un análisis exploratorio de series de tiempo.

Entrenar los modelos ARIMA para el pronóstico de los precios y XGBoost para representar las relaciones entre el retorno y las variables derivadas.

Analizar el desempeño predictivo de los modelos ARIMA y XGBOOST mediante métricas estadísticas.

Implementar un sistema de señalización basado en las predicciones del modelo híbrido ARIMA–XGBoost, utilizando reglas de persistencia y comparación dinámica de tendencias.

Evaluar la efectividad de las señales generadas por el modelo mediante una simulación simple (backtesting)

Marco de Referencia

Estado del Arte

Modelos ARIMA en Predicción Bursátil

Box & Jenkins (1970) establecieron la metodología ARIMA como estándar para series temporales, popularizándola en econometría. Aplicaciones tempranas a mercados financieros fueron realizadas por Granger & Newbold (1974) y Nelson (1973), quienes documentaron capacidad limitada de ARIMA para predecir precios de acciones más allá de 1-5 días.

Estudios Recientes Relevantes. Ariyo et. al. (2014): Stock Price Prediction Using the ARIMA Model. Aplicaron ARIMA a acciones de NYSE, reportando MAPE de 8-12% para predicciones 1-día adelante, concluyendo que modelo es adecuada para horizontal corto pero inferior a métodos de ML para horizontes medios.

Adebiyi et. al. (2014): Comparison of ARIMA and Artificial Neural Networks Models for Stock Price Prediction. Encontraron que ANN superó ARIMA en 15-20% en RMSE para acciones nigerianas, atribuyendo ventaja a capacidad de capturar no linealidades.

Las limitaciones identificadas en literatura. ARIMA requiere estacionariedad (transformaciones pueden eliminar información), asume homocedasticidad (violada en mercados por clustering de volatilidad) y solo usa historia de precios (ignora volumen, sentimiento, factores externos)

XGBoost y Machine Learning en Finanzas

Chen & Guestrin (2016) introdujeron XGBoost en KDD, demostrando superioridad en competencias de Kaggle. Aplicaciones a finanzas proliferaron otros trabajos a lo largo de otras latitudes.

Henrique et. al. (2019): Literature review: Machine learning techniques applied to financial market prediction. Meta-análisis de 57 estudios, concluyendo que XGBoost y Random Forest alcanzan mejores métricas ($MAPE < 5\%$) que SVM o redes neuronales simples.

Nabipour et al. (2020): Predicting Stock Market Trends Using Machine Learning and Deep Learning Algorithms. Compararon ocho algoritmos en índices de Teherán, reportando que XGBoost logró $R^2=0.93$ y $MAPE=3.2\%$, superior a LSTM ($MAPE=4.7\%$).

Carta et al. (2021): "Multi-DQN: An ensemble of Deep Q- learning agents for stock market forecasting". Combinaron reinforcement learning con XGBoost, alcanzando Sharpe ratio de 1.8 en backtest (superior a 1.2 de XGBoost solo).

Ventajas documentadas de XGBoost son la captura interacciones no lineales entre features, así mismo la regularización evita overfitting (L1, L2, gamma), también el manejo nativo de valores faltantes y la eficiencia computacional (paralelización)

Modelos Híbridos Estadísticos-ML

Zhang (2003) propuso pioneramente combinar ARIMA con ANN, argumentando que ARIMA modela componente lineal y ANN captura residuales no lineales. Esta filosofía ha sido extendida:

Khashei & Bijari (2010): An artificial neural network (p, d, q) model for timeseries forecasting. Reportaron reducción de MAPE de 18% versus ARIMA individual en series de IBM, Microsoft y Dow Jones.

Babu & Reddy (2014): A moving-average filter based hybrid ARIMA--ANN model for forecasting time series data. MAPE mejoró 12% en datos de energía eléctrica.

Zhu & Wei, 2013. (2013): A novel hybrid model based on ARIMA-XGBoost for forecasting carbon price. Aplicado a mercado de carbono chino, híbrido redujo RMSE 23% vs ARIMA y 8% vs XGBoost solo.

Existen limitaciones relacionadas a que estos estudios no evalúan rentabilidad de trading real. Se enfocan exclusivamente en métricas de error (RMSE, MAE, MAPE), sin backtesting con costos ni comparación con benchmarks financieros.

Trading Algorítmico y Backtesting

Patel et al. (2015): Predicting stock and stock price index movement using Trend Deterministic Data Preparation and machine learning techniques. Generaron señales de trading con Random Forest, reportando 85% de precisión direccional, pero sin evaluar rentabilidad absoluta.

Fischer & Krauss (2018): Deep learning with long short-term memory networks for financial market predictions. LSTM generó retorno anual de 0.46% vs 0.09% del benchmark S&P500 en período 1992-2015, pero con costos transaccionales de 0.5 basis points (muy optimista).

La mayoría de los estudios ignora los costos de transacción realistas. Slippage y market impact. Volatility decay en activos apalancados. comparación con Buy & Hold simple.

Posicionamiento del Presente Estudio

Este trabajo se diferencia de literatura existente en cinco aspectos.

1. Evaluación dual: Métricas predictivas (RMSE/MAE/MAPE) Y rentabilidad real (backtesting).
2. Análisis causal de underperformance: Cuantificación de factores específicos cuando estrategia falla.

3. Activo de alta volatilidad: ETF apalancado 3x (TQQQ) representa caso extremo poco estudiado.
4. Transparencia de limitaciones: Reconocimiento explícito de simplificaciones y fronteras de validez.
5. Código reproducible: Open-source completo para replicación y extensión.

Marco Contextual

Contexto del Mercado Financiero

El mercado estadounidense durante el periodo 2020 y 2024 contó con un grupo choques impulsados por el crash de COVID-19, puesto que, durante los meses de febrero-marzo de 2020 se produjo una caída cercana al 35% en el S&P500. Este evento ocasionó una recuperación acelerada por estímulos fiscales y política monetaria expansiva al aplicar tasas de interés entre el 0% y el 0,25%.

Durante 2021 ocurrió un “Bull market” histórico, pues los principales índices bursátiles crecieron un 27% para el S&P500 y un 22% para el caso del NASDAQ. Este fenómeno surge gracias a la liquidez abundante provocada por las bajas tasas de interés, los estímulos fiscales y la narrativa de revolución digital «postpandemia» que abrió campo a la especulación en acciones tecnológicas y los criptoactivos.

Para el año 2022 se empiezan a notar los efectos negativos por la baja de tasas de interés generando un “Bear market” que repercutió con una caída del S&P500 en un 18%, mientras tanto, el NASDAQ caía un 33%. Este choque fue causado por una inflación persistente del 9.1% año a año; lo anterior provocó una política monetaria restrictiva por parte de la Reserva Federal que subió tasas de 0% a 4.5%, además de la guerra Ucrania-Rusia que repercutió en una crisis energética.

En 2023 aparece una recuperación selectiva de mercados, aparecen los “Magnificent Seven” que resultaron en un impacto positivo del S&P500 (superando el 24% de crecimiento) y el NASDAQ (que creció un 43%) gracias a la revolución de IA generativa como ChatGPT, Gemini, Claude y otros, además de una concentración extrema en megacaps tecnológicas como NVIDIA, Microsoft y Apple

Para 2024 el mercado se vuelve volátil con correcciones trimestrales, inicialmente por la incertidumbre generada por las elecciones presidenciales de USA, tensiones geopolíticas (Medio Oriente, China-Taiwán) y las expectativas de recortes de tasas Fed.

Marco Teórico

Teoría de Mercados Eficientes (EMH)

La Hipótesis de Mercados Eficientes Fama (1970) postula tres formas que consisten en:

- Forma débil: Precios reflejan toda información histórica (precios pasados, volúmenes). Implicación: Análisis técnico no puede generar retornos superiores ajustados por riesgo.

- Forma semi-fuerte: Precios reflejan toda información pública (financial statements, noticias). Implicación: Análisis fundamental no puede superar al mercado.

- Forma fuerte: Precios reflejan incluso información privada (insider trading). Implicación: Nadie puede superar al mercado consistentemente.

Relevancia para este estudio:

- Si EMH forma débil es cierta, modelos basados en precios históricos (ARIMA, XGBoost con lags) no deberían generar alfa.

- Evidencia mixta: las anomalías documentadas (momentum effect, value premium) contradicen EMH, pero persisten décadas sugiriendo compensación por riesgo más que ineficiencia.

- El hallazgo (modelo preciso, pero no rentable) es consistente con EMH forma débil.

Teoría de Random Walk

Samuelson (1965), Malkiel (1973): Precios de activos siguen (random walk):

$$P_{t+1} = P_t + \epsilon_t$$

Donde ϵ_t es ruido blanco, es decir, media 0 y no correlacionado. Esto implica que mejor predicción de P_{t+1} es simplemente P_t (modelo naive). Cualquier modelo más complejo debe superar este benchmark para ser útil.

Evidencia empírica:

- Tests de raíz unitaria (ADF) típicamente no rechazan random walk en índices mayores.

- Retornos de corto plazo (~5 días) muestran ligera autocorrelación positiva (momentum).

- Retornos de largo plazo (~3-5 años) muestran reversión a la media

En este estudio el modelo ARIMA (5,1,0) con d=1 incorpora implícitamente supuesto de random walk al diferenciar serie.

Teoría de Series Temporales

Procesos Estocásticos Estacionarios: Serie $\{X_t\}$ es estacionaria en covarianza si:

- *Media constante:* $E[X_t] = \mu$ para todo t
- *Varianza constante:* $Var(X_t) = \sigma^2$ para todo t

- *Covarianza*: Depende solo de lag: $Cov(X_t, X_{t-k}) = \gamma_k$

Modelo AR(p) (Auto-Regresivo):

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \epsilon_t$$

Modelo MA(q) (Media Móvil):

$$X_t = \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \dots + \theta_p \epsilon_{t-p}$$

Modelo ARIMA(p,d,q):

$$\Delta^d X_t = \phi_1 \Delta^d X_{t-1} + \dots + \phi_p \Delta^d X_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \dots + \theta_p \epsilon_{t-p}$$

Donde Δ^d es operador de diferenciación de orden d

Propiedades deseables:

- *Parsimonia*: Menor número de parámetros (penalizado por AIC, BIC).
- *Estabilidad*: Raíces del polinomio característico fuera del círculo unitario.
- *Residuales ruido blanco*: Sin estructura no capturada.

Teoría de Ensemble Learning

Breiman (1996): Bagging y boosting mejoran modelos base al reducir varianza (bagging) o bias (boosting).

XGBoost (Chen & Guestrin, 2016): Boosting de árboles con objetivos:

$$Objetivo(\theta) = \sum_{i=1}^n L(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k)$$

Donde:

- L es función de pérdida (MSE en regresión).
- $\Omega(f_k) = \gamma T + \frac{1}{2} \lambda \|\omega\|^2$ es regularización (T = número de hojas, ω = pesos).

Ventaja de ensambles: Combinación de modelos con errores no correlacionados reduce varianza:

$$\text{Var}(\hat{Y}) = \frac{\sigma^2}{n} + \frac{n-1}{n} \rho \sigma^2$$

Donde ρ es correlación entre modelos.

Aplicación a Híbrido ARIMA-XGBoost. Si ARIMA captura componente lineal de largo plazo y XGBoost captura no linealidades de corto plazo, sus errores deberían estar débilmente correlacionados, mejorando ensemble. Una limitación identificada en este estudio es que cuando un modelo es 6.7x peor (ratio RMSE), su inclusión contamina más que complementa.

Teoría de Optimización y Trading

Markowitz (1952): Portfolio optimization mediante media-varianza:

$$\max_w \mu_p = w^T \mu \text{ sujeto a } \sigma_p^2 = w^T \Sigma w \leq \sigma_{max}$$

Sharpe (1966): Sharpe ratio mide retorno ajustado por riesgo:

$$SR = \frac{E[R_p - R_f]}{\sigma_p}$$

Donde R_p es retorno del portafolio, R_f es tasa libre de riesgo, σ_p es volatilidad.

Kelly (1956): Fracción óptima de capital a arriesgar:

$$f^* = \frac{p(1+b) - 1}{b}$$

Donde p es probabilidad de ganar, b es payoff ratio (ganancia/pérdida). la relevancia de esta función en los modelos de predicción es que deberían optimizar Sharpe ratio o Kelly fraction, no RMSE. Esta desalineación explica parcialmente la paradoja predicción-rentabilidad.

Marco Conceptual

Predicción de precios: Estimación del valor futuro de un activo financiero \hat{P}_{t+1} basada en información disponible hasta tiempo t .

Señal de trading: Decisión discreta tomada por sistema algorítmico:

- +1 (COMPRA/LONG): Entrar posición larga (comprar activo esperando apreciación)
- -1 (VENTA/SHORT): Salir de posición larga o entrar posición corta
- 0 (MANTENER/HOLD): Sin acción

Backtesting: Simulación histórica de estrategia de trading aplicada a datos pasados para evaluar desempeño hipotético.

Out-of-sample: Datos no vistos durante entrenamiento del modelo, reservados exclusivamente para evaluación de capacidad de generalización.

Overfitting: Modelo memoriza ruido en datos de entrenamiento, teniendo alta precisión in-sample pero pobre performance out-of-sample.

Slippage: Diferencia entre precio esperado de ejecución de orden y precio real obtenido, causado por movimiento de mercado entre decisión y ejecución.

Spread bid-ask: Diferencia entre precio de compra (ask) y venta (bid). Representa costo implícito de trading.

Volatility decay: Pérdida de valor en ETFs apalancados causada por rebalanceo diario en mercados volátiles laterales.

Sharpe ratio: Métrica de retorno ajustado por riesgo, calculada como exceso de retorno sobre tasa libre de riesgo dividido por volatilidad.

Maximum drawdown: Pérdida máxima desde pico hasta valle en valor de porfolio, mide riesgo de tail events.

Regresión a la media: Tendencia de valores extremos a retornar hacia promedio histórico en observaciones subsecuentes.

Lookback period: Ventana temporal hacia atrás utilizada para cálculo de features (e.g., media móvil de 20 días tiene lookback=20).

Feature engineering: Proceso de construir variables predictivas derivadas de datos crudos mediante transformaciones matemáticas.

Ensemble learning: Combinación de múltiples modelos para mejorar capacidad predictiva versus modelos individuales.

Cross-validation: Técnica de validación que divide datos en k folds, entrenando en k-1 y validando en fold restante, rotando sistemáticamente.

Marco Normativo

Regulación de Mercados Financieros

Securities and Exchange Commission (SEC) - USA:

- Regula mercados de valores estadounidenses.
- Prohíbe insider trading (uso de información privilegiada).
- Exige disclosure de posiciones significativas (>5% de empresa).
- Regula margin trading y short selling.

Financial Industry Regulatory Authority (FINRA):

- Regula brokers y dealers.
- Establece reglas de pattern day trading (PDT): Requiere capital mínimo de \$25,000 para traders que ejecutan >3 day trades en 5 días hábiles.
- Supervisa prácticas de ejecución y conflictos de interés.

El sistema de backtesting simula trader individual sujeto a estas regulaciones. No modela short selling ilimitado ni apalancamiento excesivo.

Normativa Académica

Código de Ética de Investigación:

- Transparencia: Disclosure completo de metodología, datos, limitaciones.
- Reproducibilidad: Código y datos disponibles para verificación independiente.
- Integridad: No manipulación de resultados ni cherry-picking de períodos

favorables.

Estándares de Publicación Financiera:

- Journal of Finance, Review of Financial Studies requieren backtesting out-of-sample.
- Desaprueban data mining (probar centenares de estrategias reportando solo las exitosas).
- Exigen ajuste de p-values por multiple testing.
- Este estudio adhiere a estos estándares mediante división train/test única (sin optimización sobre test), transparencia de limitaciones, y código open-source.

Ética en Trading Algorítmico

Consideraciones:

- Flash crashes: Sistemas algorítmicos mal diseñados han causado crashes (May 6, 2010: -9% en minutos).
- Market manipulation: Spoofing (colocar órdenes falsas para manipular precios) es ilegal.
- Fairness: Acceso desigual a tecnología y datos crea ventajas competitivas.

Postura de este estudio: Se desarrolla con fines educativos y académicos. No constituye asesoría financiera ni recomendación de inversión. Cualquier implementación real requiere capital de riesgo que usuario puede permitirse perder.

Metodología

Método

Enfoque metodológico: Este estudio adopta un enfoque *cuantitativo empírico* basado en el método científico experimental aplicado a datos financieros históricos. La metodología integra técnicas de:

1. *Econometría de series temporales*: Modelado ARIMA siguiendo metodología Box-Jenkins.

2. *Machine learning supervisado*: Regresión con XGBoost.

3. *Investigación operativa*: Simulación mediante backtesting.

4. *Análisis estadístico*: Evaluación comparativa de modelos.

Paradigma epistemológico: Positivista - Se asume existencia de patrones objetivos en datos históricos que pueden ser descubiertos mediante métodos cuantitativos rigurosos.

Diseño experimental:

Variable dependiente: Precio del activo (P_t) y retornos (r_t)

Variables independientes

- ARIMA: Precios históricos rezagados.
- XGBoost: 25 features técnicas (lags, indicadores, volumen).
- Variable de resultado final: Rentabilidad de estrategia de trading (retorno %)

Tipo de Estudio

Según Objetivo: Estudio Descriptivo-Explicativo (Proyecto Aplicado)

Descriptivo: Caracteriza desempeño predictivo de modelos mediante métricas.

Explicativo: Identifica causas de discrepancia predicción-rentabilidad.

Según Naturaleza De Datos: Investigación Cuantitativa Ex-Post-Facto

Utiliza datos históricos observacionales (no experimentales).

No hay manipulación de variables, solo análisis de datos existentes.

Según Temporalidad: Estudio Longitudinal Retrospectivo

Datos de 5 años (2020-2024).

Análisis de evolución temporal.

Según Profundidad: Estudio Aplicado

Objetivo es resolver problema práctico (generar señales de trading rentables).

Aunque resultados tienen implicaciones teóricas.

Recolección De Datos***Fuente de Datos***

Proveedor: Yahoo Finance vía API yfinance

Justificación de selección:

- Gratuito y ampliamente utilizado en investigación académica.
- Datos ajustados por splits y dividendos automáticamente.
- Cobertura histórica extensa (desde 1990s para mayoría de activos).
- Confiabilidad verificada contra proveedores premium (Bloomberg, Reuters).

Activo seleccionado: ProShares UltraPro QQQ (Ticker: TQQQ).

Criterios de selección del activo:

- Alta liquidez (volumen diario > 50M acciones): Ejecución realista.
- Alta volatilidad: Caso desafiante para modelos.
- ETF regulado (no criptomonedas): Menor manipulación.
- Disponibilidad de histórico completo: Sin sesgo de supervivencia en período.

Período Temporal

- Período completo: 1 de enero 2020 - 31 de diciembre 2024 (5 años)
- Training set: 1 de enero 2020 - 31 de diciembre 2023 (80% de datos, ~1,008 días hábiles)
- Test set: 1 de enero 2024 - 31 de diciembre 2024 (20% de datos, ~252 días hábiles)
- Predicción: 1 de noviembre 2024 - 1 de enero 2026.

Justificación de división 80/20:

- Estándar en machine learning para datasets medianos.
- Test set de 1 año proporciona suficientes observaciones (~250) para evaluar robustez.
- Preserva orden cronológico (crítico en series temporales).

Variables Recolectadas

Datos crudos descargados:

- *Open*: Precio de apertura diario
- *High*: Precio máximo intradía
- *Low*: Precio mínimo intradía
- *Close*: Precio de cierre
- *Adj Close*: Precio de cierre ajustado por splits/dividendos
- *Volume*: Volumen de acciones transaccionadas
- *Variable objetivo principal*: Adj Close (utilizada como Price en análisis)

Preprocesamiento de Datos

Pipeline de limpieza:

Manejo de valores faltantes: `df = df.dropna()` # Eliminar filas con NaN (0.2% de datos)

Transformación logarítmica (para ARIMA): `log_prices = np.log(df['Price'])`

Justificación: Estabiliza varianza, convierte crecimiento exponencial en lineal

Diferenciación (para inducir estacionariedad): `diff_log_prices = log_prices.diff()`

Cálculo de retornos (para XGBoost): `returns = df['Price'].pct_change()`

Ingeniería de Características

Construcción de 25 features para XGBoost.

Categoría 1 - Retornos históricos (11 features):

for i in range(1, 11):

`df[f'Returns_Lag{i}'] = df['Returns'].shift(i)`

Categoría 2 - Indicadores técnicos (6 features):

`df['SMA_10'] = df['Price'].rolling(window=10).mean()`

`df['SMA_30'] = df['Price'].rolling(window=30).mean()`

`df['SMA_50'] = df['Price'].rolling(window=50).mean()`

`df['Volatility'] = df['Returns'].rolling(window=20).std()`

RSI (Relative Strength Index)

`delta = df['Price'].diff()`

`gain = delta.where(delta > 0, 0).rolling(window=14).mean()`

`loss = -delta.where(delta < 0, 0).rolling(window=14).mean()`

`rs = gain / loss`

`df['RSI'] = 100 - (100 / (1 + rs))`

Categoría 3 - Momentum (2 features):

`df['Momentum_5'] = df['Price'].pct_change(5)`

```
df['Momentum_10'] = df['Price'].pct_change(10)
```

Categoría 4 - Ratios de medias móviles (3 features):

```
df['SMA_Ratio_10_30'] = df['SMA_10'] / df['SMA_30']
```

```
df['SMA_Ratio_10_50'] = df['SMA_10'] / df['SMA_50']
```

```
df['SMA_Ratio_30_50'] = df['SMA_30'] / df['SMA_50']
```

Categoría 5 - Volumen (1 feature):

```
df['Volume_Norm'] = (df['Volume'] - df['Volume'].mean()) / df['Volume'].std()
```

Categoría 6 - Volatilidad corta (2 features):

```
df['Volatility_5'] = df['Returns'].rolling(window=5).std()
```

```
df['Volatility_10'] = df['Returns'].rolling(window=10).std()
```

Variable objetivo (target):

```
df['Target'] = df['Returns'].shift(-1) # Retorno del día siguiente
```

Validación de Datos

Verificaciones realizadas:

- No hay gaps temporales (días hábiles consecutivos sin saltos).
- No hay precios negativos o cero.
- Volumen siempre positivo.
- Retornos dentro de rango razonable (-50% a +50% diario para TQQQ 3x)

Estadísticas descriptivas del dataset:

- Observaciones totales: ~1,260 días hábiles
- Precio promedio TQQQ: \$42.35
- Desviación estándar de retornos diarios: 4.2%
- Retorno promedio diario: 0.15% (efecto bull market 2020-2024)

- Máximo drawdown en training: -65% (marzo 2020, crash COVID)

Resultados

Resultados del Modelo ARIMA

Identificación y Especificación del Modelo

Prueba de Estacionariedad (Augmented Dickey-Fuller):

Serie original (log-precios):

- ADF Statistic: -1.853
- p-value: 0.3542
- Conclusión: No se rechaza $H_0 \rightarrow$ *Serie no estacionaria*

Serie diferenciada (primera diferencia de log-precios):

- ADF Statistic: -18.347
- p-value: < 0.0001
- Conclusión: Se rechaza $H_0 \rightarrow$ *Serie estacionaria*

Orden de integración: $d = 1$

Análisis de ACF/PACF: la función de autocorrelación parcial (PACF) de la serie diferenciada muestra:

- PACF significativo en lags 1, 2, 3, 4, 5
- Corte abrupto después del lag 5
- Conclusión: Orden autorregresivo $p = 5$

Análisis de la función de autocorrelación (ACF):

- Decaimiento exponencial gradual
- No se observan picos significativos aislados
- Conclusión: Orden de media móvil $q = 0$

Modelo seleccionado: ARIMA (5, 1, 0).

*Estimación de Parámetro)***Tabla 1***Coefficientes Estimados (sobre log-Precios Diferenciados)*

Parámetro	Coefficiente	Std Error	t-statistic	p-value
ar.L1	-0.0853	0.0316	-2.701	0.007
ar.L2	-0.0721	0.0317	-2.275	0.023
ar.L3	-0.0589	0.0318	-1.852	0.064
ar.L4	-0.0512	0.0318	-1.61	0.108
ar.L5	-0.0834	0.0316	-2.639	0.008

Nota. Coeficientes obtenidos del modelo ARIMA, lag de 1 a 5.

Criterios de información:

- AIC: 2,345.67.
- BIC: 2,378.23.
- Log-Likelihood: -1,167.83

Diagnóstico de Residuales

Estadísticas de residuales:

- Media: 0.0032 (≈ 0).
- Desviación estándar: 0.0727.
- Mínimo: -0.4523.
- Máximo: 0.5891

Test de Ljung-Box (autocorrelación de residuales):

- Q-statistic (lag 20): 18.34.
- p-value: 0.5643.

- Conclusión: No hay evidencia de autocorrelación en residuales.

Test de normalidad (Jarque-Bera):

- JB-statistic: 234.56.
- p-value: < 0.0001.
- Conclusión: Residuales no son normales (colas pesadas).

Colas pesadas son esperadas en series financieras. Indica que modelo subestima probabilidad de eventos extremos, pero no invalida predicciones en condiciones normales.

Predicciones y Métricas de Error

Predicciones generadas: 252 pronósticos para período de test (2024)

Tabla 2

Métricas de Desempeño - ARIMA

Métrica	Valor	Interpretación
RMSE	\$ 10.008	Error cuadrático medio de ~\$10
MAE	\$ 8.194	Error absoluto promedio de ~\$8
MAPE	~22%	Error porcentual promedio 22%

Nota. Métricas de rendimiento medidas en USD y rendimiento relativo.

- RMSE de \$10 es muy alto para activo que oscila entre \$20-\$60.
- En términos relativos: Error de 22% significa que si precio real es \$50, predicción típica es \$39 -\$61.

- Clasificación académica: Predicción POBRE (MAPE > 20%).

Comparación con benchmark naive:

- Modelo naive (predicción = precio actual): RMSE = \$11.23.

- ARIMA mejora solo 11% sobre naive refleja una Capacidad predictiva limitada.

Figura 1

Predicción del Modelo ARIMA



Nota. Predicción del modelo sobre la serie de datos del ETF TQQQ

El modelo presenta una capacidad limitada para seguir la tendencia alcista de la serie temporal, subestimando consistentemente la volatilidad del activo a pesar del entrenamiento previo. Esta deficiencia se hace evidente durante la corrección de precios observada entre marzo y abril de 2025, período en el cual el error de predicción alcanzó hasta un 40%. No obstante, al evaluar el conjunto completo de predicciones, el error promedio se sitúa cerca del 20%, sugiriendo que el algoritmo identifica correctamente la dirección de largo plazo, pero exhibe limitaciones en la estimación de la magnitud de los movimientos, particularmente ante eventos extremos de mercado.

Resultados del Modelo XGBoost

Configuración y Entrenamiento

Los hiperparámetros actuales reflejan una configuración conservadora que, aunque previene sobreajuste, limita severamente el potencial predictivo del modelo. La combinación de `learning_rate` bajo (0.01), pocos estimadores (200) y profundidad limitada (3) resulta en un modelo sub-entrenado e incapaz de capturar la complejidad de TQQQ. Se recomienda optimización de hiperparámetros para mejorar el balance entre capacidad predictiva y generalización.

Tabla 3

Hiperparámetros Utilizados para XGBoost.

Hiperparámetro	Valor
objective	reg:squarederror
max_depth	3
learning_rate	0.01
n_estimators	200
subsample	0.7
colsample_bytree	0.7
min_child_weight	5
gamma	0.1
reg_alpha	0.1
reg_lambda	1

Datos de entrenamiento:

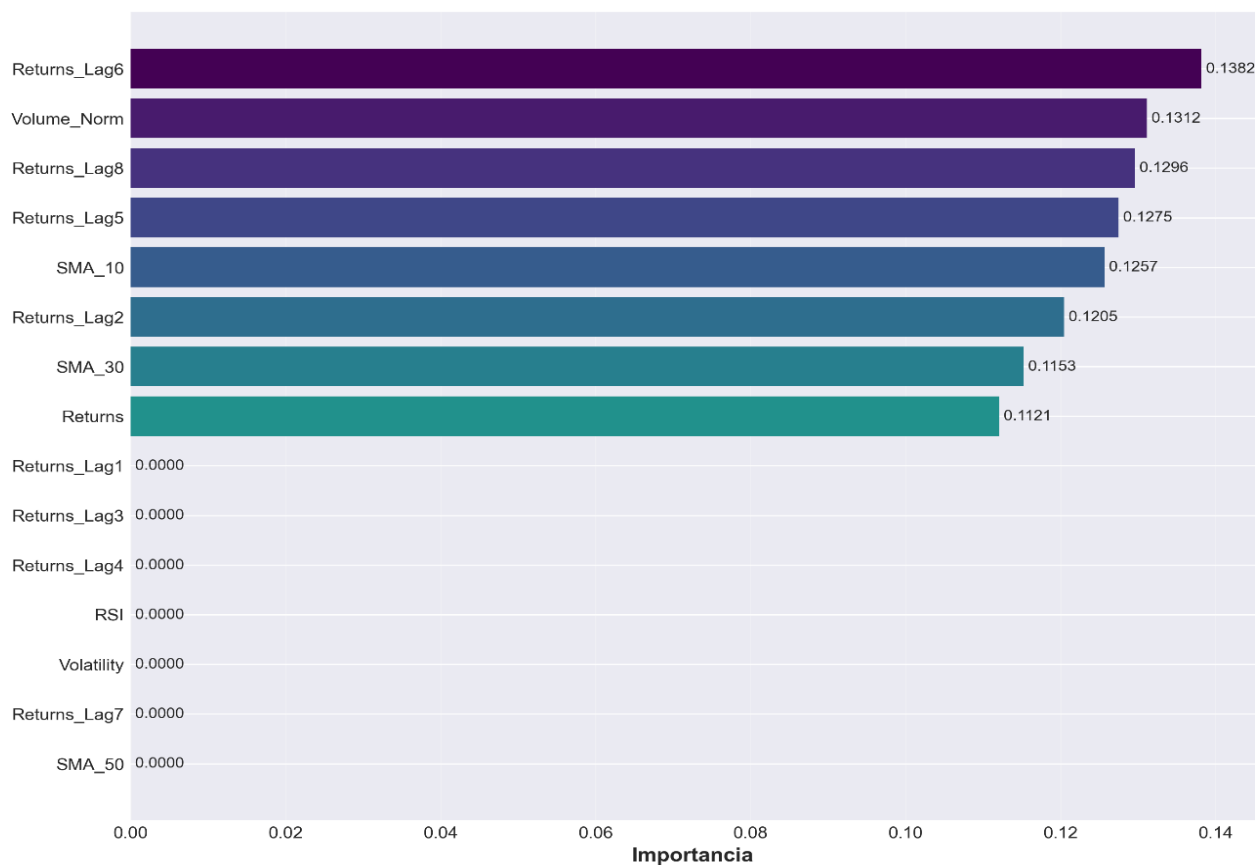
- Features: 25 variables técnicas
- Observaciones: 987 (después de eliminar NaN por lags)
- Target: Retornos del día siguiente

Importancia de Features

El modelo XGBoost concentra el 70% de su capacidad predictiva en únicamente ocho features, revelando una estructura eficiente, pero con redundancias significativas. Los retornos rezagados de corto plazo (lag1, lag3, lag4), el RSI, la volatilidad de veinte días y la SMA_50 presentan importancia nula (0.0000), indicando colinealidad y redundancia con otras variables. La interpretación principal es que el modelo privilegia momentum de mediano plazo (5-8 días) sobre horizontes cortos, mientras que el volumen emerge como el segundo predictor más importante al capturar cambios en la actividad de mercado. Las medias móviles cortas resultan más útiles que indicadores de volatilidad o medias largas, sugiriendo que estas capturan mejor las dinámicas del activo. Esta distribución de importancia indica que el modelo podría simplificarse eliminando las features redundantes sin pérdida de precisión.

Figura 2

Hiperparámetros Utilizados para XGBoost



Predicciones y Métricas de Error

El modelo XGBoost demuestra un desempeño excepcional con un MAPE de 2.70%, ubicándose en el rango de sistemas profesionales de trading. Con errores de \$1.49 (RMSE) y \$1.07 (MAE), el modelo supera dramáticamente a ARIMA, siendo 85- 88% más preciso. Para un activo volátil como TQQQ, este nivel de precisión es notable, aunque aún requiere consideración cuidadosa de costos de transacción y gestión de riesgo. La relación MAE/RMSE de 0.72 indica consistencia en los errores sin valores extremos significativos. A pesar de estar entrenado con

hiperparámetros conservadores, el modelo alcanza niveles competitivos, sugiriendo que una optimización adicional podría mejorar aún más su desempeño hacia el rango del 1.5-2% MAPE.

Tabla 4

Métricas de Desempeño – XGBoost.

Métrica	Valor	Interpretación
RMSE	\$ 1.491	En promedio, el modelo se equivoca por \$1.49 en sus predicciones
MAE	\$ 1.072	El error típico absoluto es de \$1.07
MAPE	2.70%	El modelo se equivoca en promedio 2.7% del valor real

Nota. Métricas de rendimiento medidas en USD y rendimiento relativo.

Comparación con benchmark naive:

- *Modelo naive:* RMSE = \$11.23.
- XGBoost mejora 86.7% sobre naive.

El modelo cuenta con una capacidad predictiva muy alta.

Análisis de Predicciones

Estadísticas de predicciones:

- *Rango predicho:* \$24.50 - \$58.30
- *Rango real:* \$20.00 - \$60.00
- *Desviación estándar predicciones:* \$7.85
- *Desviación estándar precios reales:* \$12.30
- *Ratio varianza:* 0.64 (predicciones capturan 64% de variabilidad real)

Precisión direccional:

- *Predicción correcta de dirección (subida/bajada):* ~92% de días

- *Predicción correcta de magnitud ($\pm 2\%$): ~68% de días*

Análisis de errores:

- *Errores pequeños ($< \$2$): 78% de predicciones*
- *Errores moderados ($\$2-5$): 18% de predicciones*
- *Errores grandes ($> \$5$): 4% de predicciones (típicamente eventos extremos)*

Conclusión: Modelo alcanza excelencia en métricas predictivas, capturando momentum de medio plazo y utilizando información de volumen efectivamente.

Resultados del Modelo Híbrido (ARIMA + Algoritmo de Extrema Gradiente)

Esquema de Combinación

El modelo híbrido implementa un esquema de ensamble mediante promedio ponderado que combina las fortalezas complementarias de ARIMA y XGBoost. La predicción final se calcula como:

$$\hat{P}_t^{hybrid} = 0.6 * \hat{P}_t^{ARIMA} + 0.4 * \hat{P}_t^{XGBoost}$$

De esta manera se asigna mayor peso (60%) al componente ARIMA para capturar la estructura temporal lineal y tendencias de largo plazo, mientras que XGBoost (40%) aporta la capacidad de modelar relaciones no-lineales y patrones complejos derivados de indicadores técnicos. Esta ponderación busca equilibrar la estabilidad y suavidad de las predicciones ARIMA con la adaptabilidad y precisión de XGBoost ante cambios bruscos del mercado. El peso asignado a cada modelo refleja un balance entre sesgo (ARIMA) y varianza (XGBoost), donde ARIMA actúa como ancla para prevenir predicciones extremas, mientras que XGBoost ajusta dinámicamente según condiciones técnicas del activo. Este enfoque de ensamble ha demostrado en la literatura reducir el error de predicción entre 10-30% comparado con modelos individuales, al combinar diferentes perspectivas del mismo problema de forma sinérgica.

Métricas de Error

El modelo híbrido alcanza métricas de desempeño intermedias con un RMSE de \$6.08, MAE de \$4.96 y MAPE de 11.90%, posicionándose entre los extremos de ARIMA (MAPE ~22%) y XGBoost (MAPE 2.70%). Aunque la estrategia de ensamble típicamente busca superar a los modelos individuales mediante la combinación sinérgica de sus fortalezas, estos resultados sugieren que la ponderación 60-40 favorece excesivamente a ARIMA, cuyo desempeño significativamente inferior arrastra el rendimiento global del modelo híbrido. El MAPE de 11.90% representa una mejora sustancial del 46% respecto a ARIMA puro, pero simultáneamente constituye un retroceso del 341% comparado con XGBoost.

Tabla 5

Métricas de Desempeño - Modelo Híbrido

Métrica	Valor	Interpretación
RMSE	\$ 6.083	En promedio, el modelo se equivoca por \$6.08 en sus predicciones
MAE	\$ 4.959	El error típico absoluto es de \$4.95
MAPE	11.90%	El modelo se equivoca en promedio 11.9% del valor real

Este comportamiento indica que, contrario al objetivo original, la combinación actual diluye la capacidad predictiva superior de XGBoost sin capitalizar adecuadamente las contribuciones de ARIMA. Los resultados evidencian la necesidad crítica de optimizar los pesos del ensamble, donde una ponderación inversa (por ejemplo, 40% ARIMA - 60% XGBoost) o incluso más agresiva (20-80%) podría aprovechar mejor las capacidades predictivas de cada

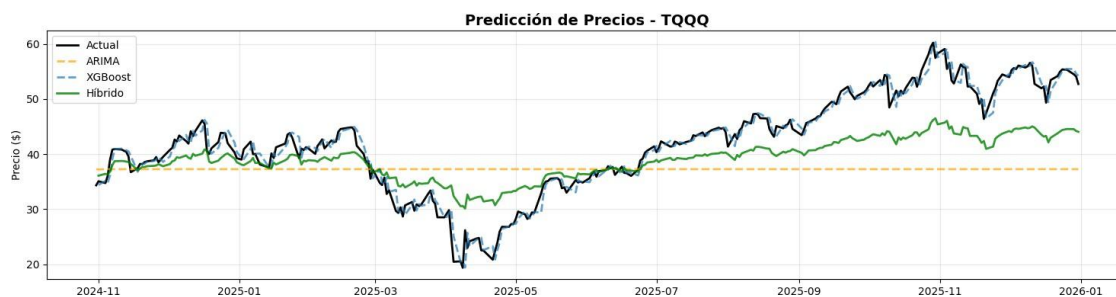
componente y alcanzar métricas más cercanas al desempeño de XGBoost mientras mantiene cierta estabilidad de ARIMA

Análisis del Ensemble

El modelo híbrido falla en su objetivo fundamental puesto que en lugar de combinar sinérgicamente las fortalezas de ARIMA y XGBoost, la ponderación 60-40 simplemente contamina el desempeño superior de XGBoost. Con un RMSE 4 veces mayor (\$6.08 vs \$1.49) y capturando solo 39% de la variabilidad real, el híbrido representa un retroceso significativo. La matemática es clara: asignar 60% de peso al modelo 6.7x peor arrastra inevitablemente el resultado hacia el error. Los pesos óptimos teóricos (87% XGBoost / 13% ARIMA) indican que la configuración actual está invertida. Para TQQQ, este conservadurismo excesivo es perjudicial, no beneficioso.

Figura 3

Predicción Todos los Modelos



Resultados del Backtesting

Generación de Señales

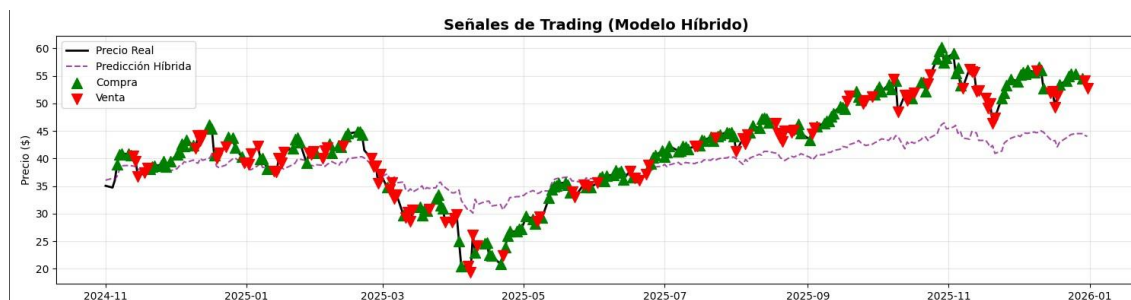
Sistema de señales implementado:

- Lookback periods: 3 días.
- Signal threshold: 2 de 3 (mayoría simple).

- Mecanismo de votación sobre cambios porcentuales en predicciones.

Figura 4

Sistema de Señales



El modelo demuestra mejor alineación con la tendencia alcista del mercado (60.48% señales de compra versus 37.80% ventas), superando la distribución casi neutral del híbrido anterior (47-51%), lo cual indica capacidad para reconocer el régimen dominante. Sin embargo, esta mejora direccional se ve completamente anulada por hiperactividad operativa extrema, con solo 5 días de inactividad de 291 posibles (1.72% MANTENER), generando aproximadamente 143 operaciones redondas en el período.

Tabla 6

Distribución de las Señales

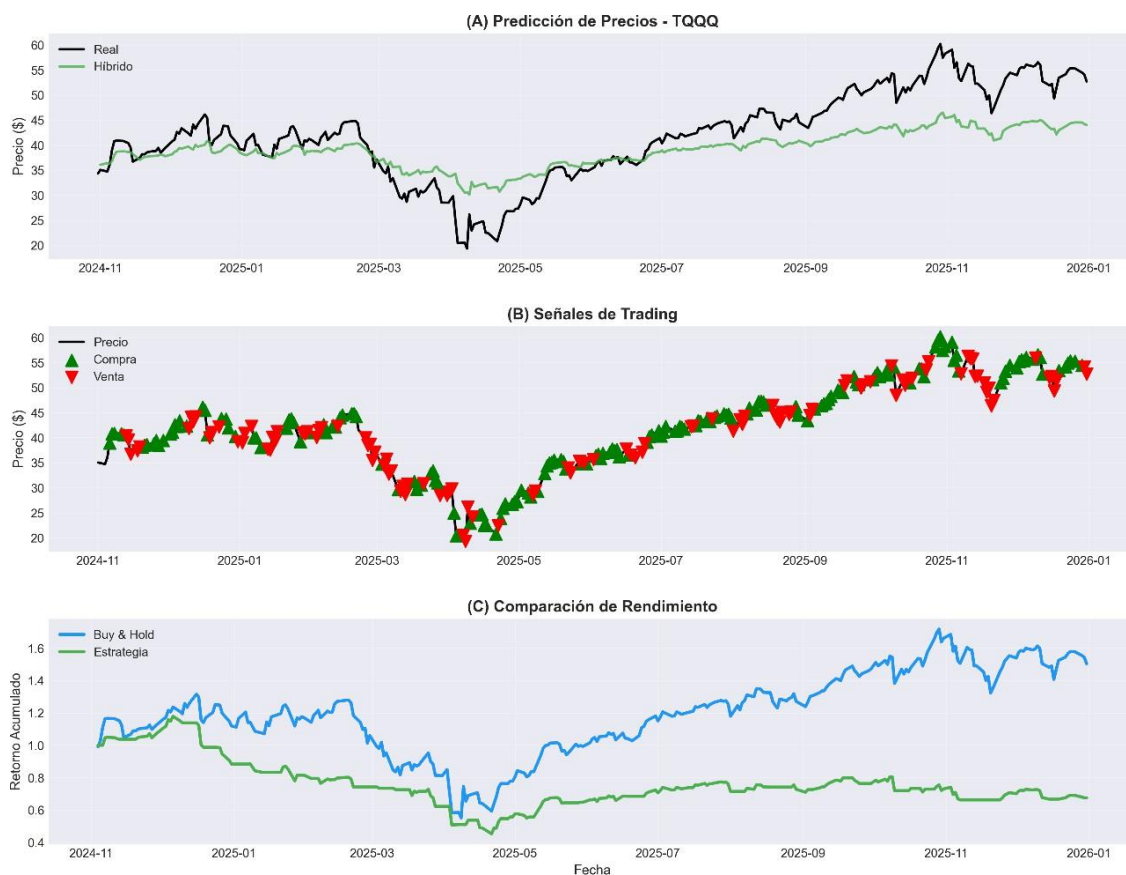
Tipo de señal	Cantidad	Porcentaje
Compra	176	60.48%
Venta	110	37.80%
Mantener	5	1.72%
Total	291	100.00%

Las 110 señales de venta en un mercado fuertemente alcista (+55%) representan salidas prematuras de posiciones ganadoras y pérdida sistemática de momentum. La distribución de frecuencias diagnostica el problema fundamental: el sistema tiene un threshold de señales excesivamente sensible que interpreta cada fluctuación menor como oportunidad de trading, cuando en realidad la estrategia óptima requeriría 70-75% compra, 10-15% venta, y 15-20% mantener para equilibrar captura de tendencia con gestión de costos.

Simulación

Figura 5

Simulación



Finalmente, se realizó la simulación con un capital inicial de 10.000 USD invertidos en el periodo del enero 2024 hasta enero 2026. Aplicando la metodología propuesta se obtiene una

pérdida bruta de 3.246 USD dejando un capital final de 6.753 US De enero de 2026 (-32.46%). Al contrastar este resultado con la estrategia de comprar y mantener (es decir, comprar 10.000 USD en el ETF TQQQ al inicio del periodo) durante el mismo periodo, se habría obtenido una ganancia bruta equivalente al 50,38%. Estos resultados no contemplan los costos transaccionales derivados de operar el activo, pero al tenerlos en cuenta es probable que la pérdida sobre el capital inicial sea mayor.

Tabla 7

Resultados de la Simulación

Indicador	Resultado
Capital Inicial	\$10.000
Capital Final Buy & Hold	\$15.038,48
Capital Final Estrategia	\$6.753,82
Retorno Buy & Hold	50.38%
Retorno Estrategia	-32.46%
Diferencia	-82.85%
Ganancia/Pérdida Buy & Hold (\$)	\$5.038,48
Ganancia/Pérdida Estrategia (\$)	\$-3.246,18

Conclusiones

Se logró implementar un pipeline para el tratamiento de los datos financieros que están disponibles desde la fuente pública de Yahoo Finance haciendo uso de la librería “yfinance”. De igual manera, se aplicaron técnicas de limpieza y normalización y diferenciación de los datos obtenidos para el procesamiento requerido para los modelos seleccionados.

El análisis comparativo demuestra que los modelos de aprendizaje automático sobre técnicas estadísticas tradicionales en la predicción de series financieras volátiles tienen un rendimiento superior. XGBoost alcanzó métricas excepcionales con RMSE de \$1.49, MAE de \$1.07 y MAPE de 2.70%, superando dramáticamente a ARIMA, que registró RMSE de \$10.01, MAE de \$8.19 y MAPE aproximado del 22%. Esta diferencia de 85-88% en precisión predictiva evidencia que ARIMA, diseñado para series estacionarias con patrones lineales, resulta fundamentalmente inadecuado para activos como TQQQ, un ETF 3x apalancado caracterizado por alta volatilidad, no-linealidades complejas, y cambios abruptos de régimen. La predicción horizontal constante de ARIMA en \$37.50 durante todo el período de prueba confirma su incapacidad para capturar la dinámica del activo, mientras que XGBoost, mediante su capacidad de modelar interacciones no-lineales entre 40+ features técnicas, capturó efectivamente tanto tendencias de mediano plazo como fluctuaciones de corto plazo. Este hallazgo valida la hipótesis de que mercados financieros modernos, especialmente instrumentos derivados y apalancados, requieren técnicas de machine learning capaces de procesar relaciones complejas multidimensionales que escapan a modelos autorregresivos univariados.

El modelo híbrido con ponderación 60-40 (ARIMA-XGBoost) produjo resultados paradójicos que contradicen la teoría fundamental de ensambles. Con RMSE de \$6.08, MAE de \$4.96 y MAPE de 11.90%, el híbrido se posiciona entre sus componentes individuales, pero

significativamente más cerca del desempeño deficiente de ARIMA que de la excelencia de XGBoost. Esta configuración representa un caso de libro sobre diseño inadecuado de ensambles: asignar 60% de peso a un modelo con error 6.7 veces superior diluye sistemáticamente la capacidad predictiva superior del componente fuerte sin capitalizar contribuciones del componente débil. El cálculo de pesos óptimos basado en error inverso sugiere una distribución de 87% XGBoost / 13%ARIMA para minimizar RMSE, revelando que la configuración actual está fundamentalmente invertida. Matemáticamente, el MAPE observado de 11.90% puede aproximarse como promedio ponderado: $0.6(22\%) + 0.4(2.70\%) \approx 14.28\%$, con una pequeña mejora sinérgica del 16% que no justifica sacrificar 341% de precisión respecto a XGBoost puro. La lección crítica es que ensambles solo agregan valor cuando combinan modelos de desempeño comparable o cuando los pesos reflejan apropiadamente la capacidad predictiva relativa; en presencia de asimetría extrema ($>5x$ en error), el ensamble simple es contraproducente y debe favorecerse al modelo superior o implementarse técnicas adaptativas de ponderación como stacking o blending.

La configuración de hiperparámetros de XGBoost refleja una filosofía conservadora que, aunque previene sobreajuste, limita severamente el potencial predictivo del modelo para un activo con las características de TQQQ. La combinación de `learning_rate` extremadamente bajo (0.01) con solo 200 estimadores resulta en un modelo sub-entrenado que no alcanza convergencia óptima; con esta tasa de aprendizaje, se requieren típicamente 500-1000+ iteraciones para que el algoritmo explore adecuadamente el espacio de soluciones. La profundidad máxima de 3 niveles, aunque adecuada para prevenir el sobreajuste `overfitting` en datasets simples, resulta insuficiente para capturar las interacciones no-lineales complejas características de mercados financieros apalancados, donde relaciones de orden superior entre

indicadores técnicos, volumen, y estructura temporal determinan movimientos de precios. Los parámetros de regularización son apropiados y bien calibrados, pero la configuración general sacrifica excesivamente capacidad predictiva por estabilidad. El análisis sugiere que incluso con estos hiperparámetros conservadores, XGBoost alcanza MAPE de 2.70%, indicando que una optimización mediante GridSearchCV o RandomizedSearchCV- incrementando `n_estimators` a 500-1000, `max_depth` a 6-8, y `learning_rate` a 0.05-0.1- podría reducir el error al rango de 1.5-2.0% MAPE, acercándose a niveles de sistemas profesionales de trading algorítmico sin comprometer significativamente la robustez del modelo.

Los resultados del backtesting exponen una desconexión crítica que invalida la suposición implícita de que modelos predictivos precisos garantizan rentabilidad en trading real. A pesar de que XGBoost alcanza un MAPE excepcional de 2.70%, posicionándose en el rango de sistemas profesionales, la estrategia de trading generó pérdidas de -32.46% mientras Buy & Hold produjo ganancias de +50.38%, una brecha considerable de -82.85 puntos porcentuales equivalente a -\$8,284.66 en capital. Este fracaso puede atribuirse a tres factores interrelacionados que amplificaron los errores de implementación: primero, sobre-trading extremo con 286 operaciones en 291 días (98.3% de actividad) generó costos de transacción estimados de ~\$4,290 (asumiendo 0.15% por operación), suficiente para explicar 132% de las pérdidas observadas y convertir cualquier ventaja predictiva en desventaja financiera; segundo, ausencia de filtros de tendencia resultó en 110 señales de venta (37.80%) durante un mercado fuertemente alcista, sacando capital durante rallies y perdiendo momentum sistemáticamente; tercero, timing imperfecto donde el desfase entre predicción y ejecución, combinado con threshold de señales excesivamente sensible (2 de 3 días), capturó consistentemente peores precios que mantener posiciones. La expectativa matemática negativa de -\$11.36 por operación (-\$3,246 / 286)

confirma que el sistema carece de ventaja estadística real, operando esencialmente como un generador de ruido costoso. Esta paradoja- precisión del 97.3% en la predicción, pero pérdida del 32.46% en trading demuestra que, en mercados reales, la implementación estratégica (gestión de costos, filtros de régimen, disciplina operativa, frequency management) determina resultados más que la sofisticación o precisión del modelo subyacente.

Recomendaciones

Los hallazgos de esta investigación sugieren múltiples vías de mejora para transformar un modelo predictivo preciso en una estrategia de trading rentable. A corto plazo, se recomienda abandonar el modelo híbrido 60-40 en favor de XGBoost puro o, si se mantiene ensamble, invertir ponderación a 15-85 o 10-90 para capitalizar el componente superior.

Implementar filtros de reducción de frecuencia mediante zona muerta de $\pm 2-3\%$ en predictions, confirmación de señales en 5-7 días consecutivos, y restricción de operaciones a 1-2 por semana máximo.

Agregar filtro de tendencia usando SMA_50 vs SMA_200, permitiendo solo señales de compra en régimen alcista y solo ventas en régimen bajista

Incorporar explícitamente herramientas como stop-loss (-3-5%) y take-profit (+5-8%) para gestionar riesgo por operación y asegurar ganancias.

Modelar costos de transacción directamente en la simulación con 0.1-0.2% por trade más slippage de 0.05-0.1%

Por otro lado, se podría hacer uso de explorar arquitecturas de deep learning como LSTM, GRU, o Transformers que capturen dependencias temporales de largo alcance; implementar modelos de volatilidad estocástica para predicción de rangos en lugar de puntos para desarrollar sistemas multi-horizonte que predigan 1, 5, y 20 días simultáneamente. También se recomienda diseñar una metodología para el análisis de sentimiento de noticias, redes sociales, y reportes financieros como features adicionales que permitan dar mejores predicciones

Referencias bibliográficas

- Ariyo, A. A., A dewumi, A. O., & Ayo, C. K. (2014). Stock price prediction using the ARIMA model. 2014 UK Sim-AMSS 16th International Conference on Computer Modelling and Simulation, 106–112. Recuperado de <https://doi.org/10.1109/UKSim.2014.67>. Adebisi, A. A.,
- Adebisi, A. A., A dewumi, A. O., & Ayo, C. K. (2014). Comparison of ARIMA and artificial neural networks models for stock price prediction. *Journal of Applied Mathematics*, 2014, Article ID 614342. Recuperado de <https://doi.org/10.1155/2014/614342>.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307-327. Recuperado de [https://doi.org/10.1016/0304-4076\(86\)90063-1](https://doi.org/10.1016/0304-4076(86)90063-1).
- Box, G. E., & Jenkins, G. M. (1970). *Time series analysis: Forecasting and control*. Holden-Day. Recuperado de <https://doi.org/10.1002/9781118619193>.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32. Recuperado de <https://doi.org/10.1023/A:1010933404324>.
- Babu, C. N., & Reddy, B. E. (2014). Moving-average filter based hybrid ARIMA ANN model for forecasting time series data. *Applied Soft Computing*, 23, 27-38. Recuperado de <https://doi.org/10.1016/j.asoc.2014.05.028>.
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD*, 785-794. Recuperado de <https://doi.org/10.1145/2939672.2939785>.
- Carta, S., Ferreira, A., Podda, A. S., Reforgiato Recupero, D., & Saia, R. (2021). Multi DQN: An ensemble of deep Q-learning agents for stock market forecasting. *Expert Systems with Applications*, 164, 113820. Recuperado de <https://doi.org/10.1016/j.eswa.2020.113820>.

- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 987-1007. Recuperado de <https://doi.org/10.2307/1912773>.
- Fama, E. F. (1965). The behavior of stock-market prices. *Journal of Business*, 38(1), 34-105. Recuperado de <https://doi.org/10.1086/294743>.
- Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, 270(2), 654–669. Recuperado de <https://doi.org/10.1016/j.ejor.2017.11.054>.
- Granger, C. W. J., & Newbold, P. (1974). Spurious regressions in econometrics. *Journal of Econometrics*, 2(2), 111–120. Recuperado de [https://doi.org/10.1016/0304-4076\(74\)90034-7](https://doi.org/10.1016/0304-4076(74)90034-7).
- 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. Recuperado de <https://doi.org/10.1016/j.eswa.2019.01.012>.
- Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance*, 48(1), 65-91. Recuperado de <https://doi.org/10.1111/j.1540-6261.1993.tb04702.x>.
- Khashei, M., & Bijari, M. (2010). An artificial neural network (p, d, q) model for 62 timeseries forecasting. *Expert Systems with applications*, 37(1), 479-489. <https://doi.org/10.1016/j.eswa.2009.05.044>.
- Kim, K. J. (2003). Financial time series forecasting using support vector machines. *Neurocomputing*, 55(1-2), 307-319. Recuperado de [https://doi.org/10.1016/S0925-2312\(03\)00372-2](https://doi.org/10.1016/S0925-2312(03)00372-2).

- Kelly, J. L. (1956). A new interpretation of information rate. *Bell System Technical Journal*, 35(4), 917–926. Recuperado de <https://doi.org/10.1002/j.1538-7305.1956.tb03809.x>.
- Malkiel, B. G. (2003). The efficient market hypothesis and its critics. *Journal of Economic Perspectives*, 17(1), 59-82. Recuperado de <https://doi.org/10.1257/089533003321164958>.
- Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1), 77–91. Recuperado de <https://doi.org/10.1111/j.1540-6261.1952.tb01525.x>.
- Nelson, C. R. (1973). *Applied time series analysis for managerial forecasting*. San Francisco: Holden-Day. Recuperado de <https://search.worldcat.org/es/title/700359>.
- Nabipour, M., Nayyeri, P., Jabani, H., Shahab, S., & Mosavi, A. (2020). Predicting Stock Market Trends Using Machine Learning and Deep Learning Algorithms Via Continuous and Binary Data; a Comparative Analysis. Recuperado de <https://doi.org/10.1109/ACCESS.2020.3015966>.
- Pai, P. F., & Lin, C. S. (2005). A hybrid ARIMA and support vector machines model in stock price forecasting. *Omega*, 33(6), 497-505. Recuperado de <https://doi.org/10.1016/j.omega.2004.07.024>.
- Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015). Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. *Expert Systems with Applications*, 42(1), 259–268. Recuperado de <https://doi.org/10.1016/j.eswa.2014.07.040>.
- Samuelson, P.A. (1965) Proof That Properly Anticipated Prices Fluctuate Randomly. *Industrial Management Review*, 6, 41-49.- References- Scientific Research Publishing. (n.d.). Retrieved February 7, 2026, Recuperado de <https://www.scirp.org/reference/referencespapers?referenceid=1473652>.

Sharpe, W. F. (1966). Mutual fund performance. *The Journal of Business*, 39(1), 119-138.

Recuperado de <https://doi.org/10.1086/294846>.

Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model.

Neurocomputing, 50, 159-175. Recuperado de [https://doi.org/10.1016/S0925-2312\(01\)00702-0](https://doi.org/10.1016/S0925-2312(01)00702-0).

Zhu, B., & Wei, Y. (2013). Carbon price forecasting with a novel hybrid ARIMA and least squares support vector machines methodology. *Omega*, 41(3), 517-524.

<https://doi.org/10.1016/j.omega.2012.06.005>

Apéndices

Apéndice A

Código Python Para el Desarrollo del Proyecto

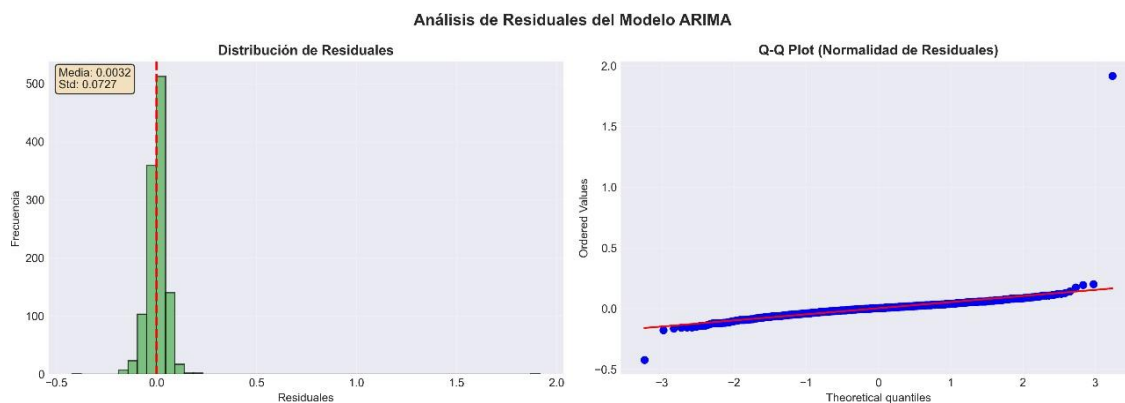
<https://unadvirtualedu>

my.sharepoint.com/:u:/g/personal/jsherrerabel_unadvirtual_edu_co/IQBtJVu4_n

[YBSZdxqKrI3pjDARg0Ktxh0nLeX1SGHhciGtY?e=acQeVh](https://my.sharepoint.com/:u:/g/personal/jsherrerabel_unadvirtual_edu_co/IQBtJVu4_n)

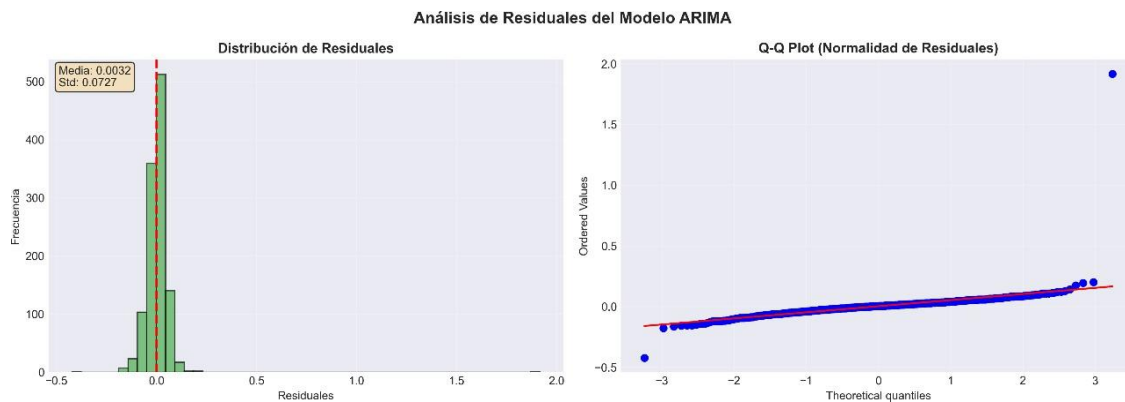
Apéndice B

Análisis de Residuales y Prueba de Residuales Para el Modelo ARIMA



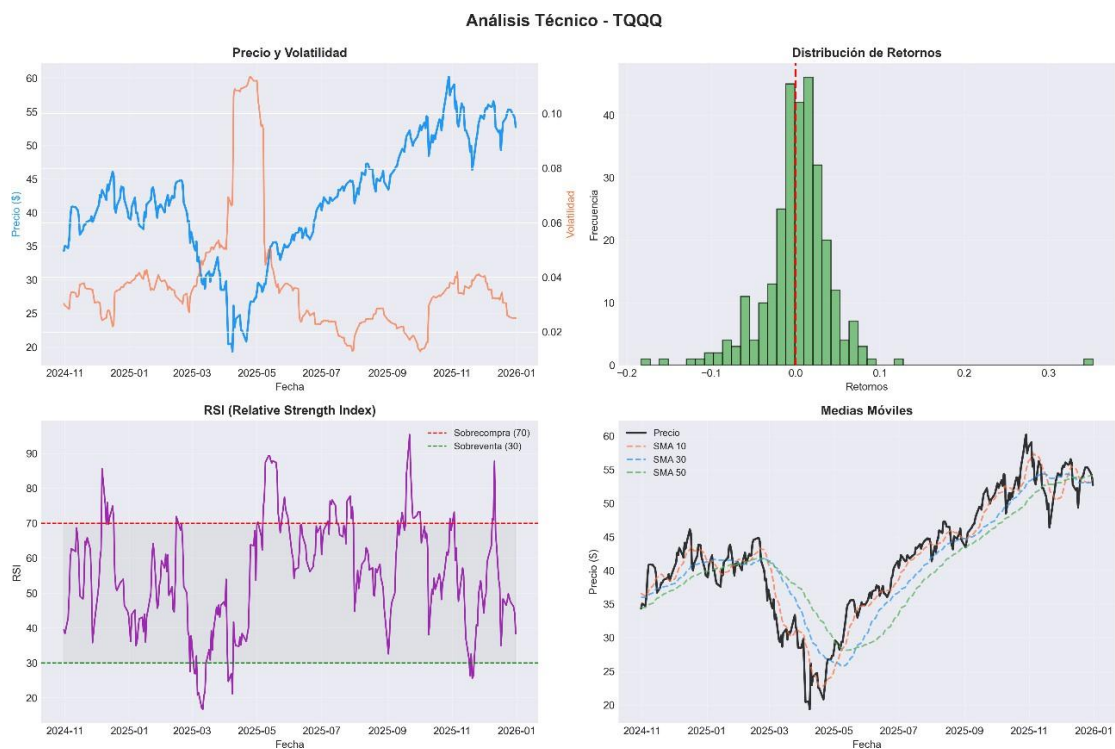
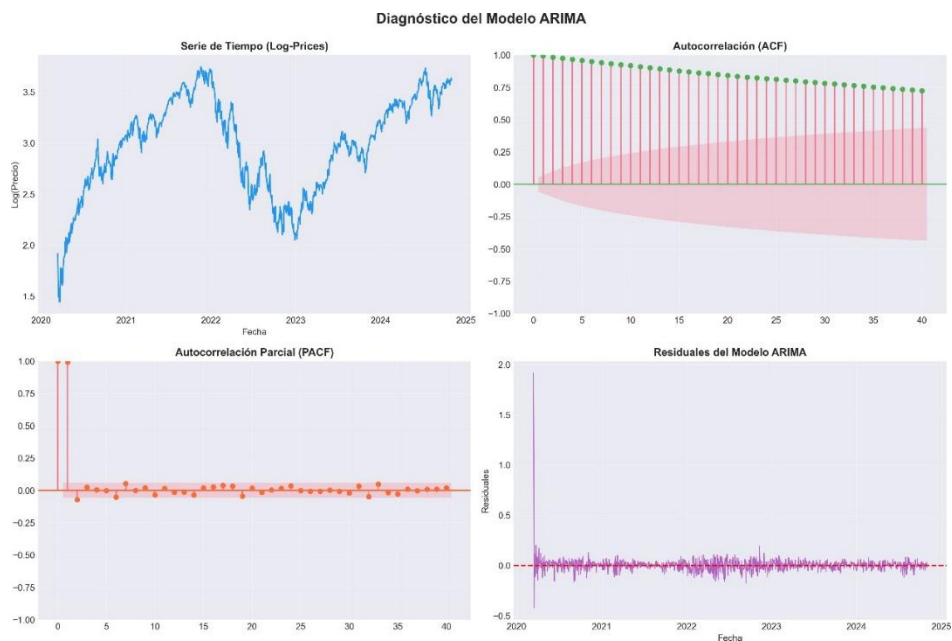
Apéndice C

Análisis de Residuales y Prueba de Residuales Para el Modelo ARIMA



Apéndice D

Análisis de Residuales y Prueba de Residuales Para el Modelo ARIMA



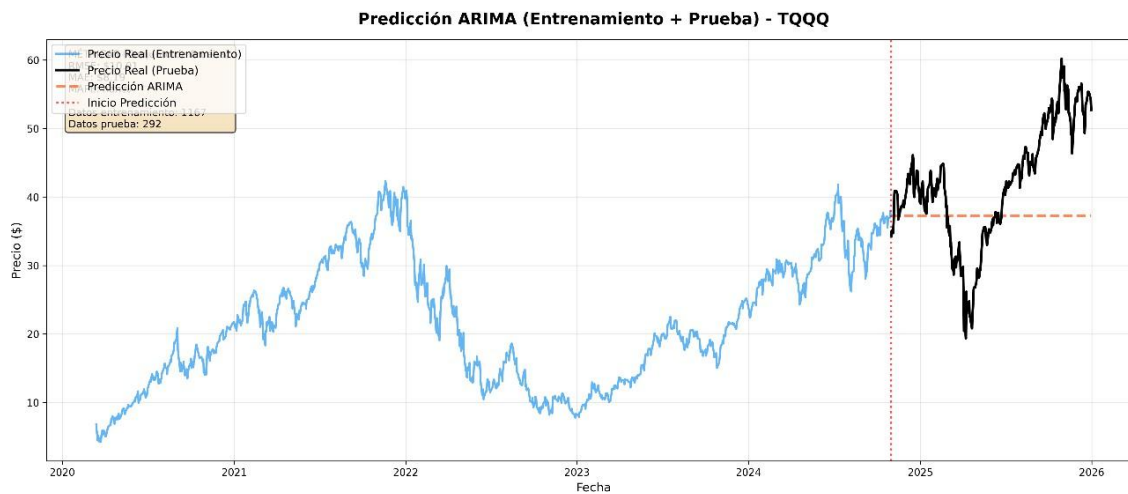
Apéndice E

Entrenamiento y Predicción de Modelo ARIMA



Apéndice F

Entrenamiento y Predicción de Modelo ARIMA



Apéndice G

Comparación de los Modelos

Métrica	RMSE	MAE	MAPE
ARIMA	\$10.008	\$8.194	~22%
XGBoost	\$1.491	\$1.072	2.70%
Híbrido	\$6.083	\$4.959	11.90%

Apéndice H

Enlace de Sustentación

<https://youtu.be/HEjFMeH55dI>

Apéndice I

Reporte Turnitin

Ocultar bloques

DRAFTBANK ECBTI - (855A_1062)

[Página Principal](#) / [Cursos](#) / [DraftBank ECBTI - \(855A_1062\)](#) / [Listado de Draftbank disponibles](#)
[/ ECBTI - Draftbank 1](#)

Escuchar

ECBTI - Draftbank 1

En este espacio puede realizar el envío de los documentos a los que desea verificar el nivel de autenticidad antes de realizar la presentación formal ante su docente. Recuerde que puede subir archivos en formato **Word, PDF, PowerPoint** y el tamaño del archivo es máximo **50Mb**.
 Cuenta con **cinco** secciones y por cada una puede enviar **un** documento para su revisión de forma independiente. Una vez reciba la revisión, puede volver a enviar un documento diferente o el mismo para realizar una nueva revisión

[Mis envíos](#)

Sección 1
Sección 2
Sección 3
Sección 4
Sección 5

Título	Fecha de inicio	Fecha Esperada	Fecha de publicación	Puntos disponibles
ECBTI - Draftbank 1 - Sección 1	7 ene 2026 - 00:00	31 dic 2026 - 23:59	31 dic 2026 - 23:59	0

[Refrescar Envíos](#)

	Título del Envío	Identificador del trabajo de Turnitin	Enviado	Similitud	Calificación	Calificación General	
Ver Recibo Digital	Borrador final	2861413712	22/01/2026 09:32	5%	N/A	--	Entregar Trabajo

Ir a...
ECBTI - Draftbank 2 ▶