

PONTIFICIA UNIVERSIDAD CATÓLICA DEL ECUADOR

FACULTAD DE CIENCIAS EXACTAS, NATURALES Y AMBIENTALES

CARRERA DE CIENCIA DE DATOS

Aplicación de modelos estadísticos y de aprendizaje automático en la predicción de series de tiempo financieras

Monografía previa a la obtención del título de: Ingeniería en Ciencia de Datos

Isaac Eduardo Porras Conde

Quito, 2026

DEDICATORIA

A mis padres, por su apoyo incondicional a lo largo de mi vida y de mi formación académica. Todo lo que he logrado tiene su origen en el esfuerzo y el amor que me han dado.

TABLA DE CONTENIDOS

Contenido

1. RESUMEN 4
2. ABSTRACT 5
3. INTRODUCCIÓN 6
4. OBJETIVOS 10
4.1. OBJETIVO GENERAL 10
4.2. OBJETIVOS ESPECÍFICOS 10
5. FUNDAMENTO TEÓRICO 11
5.1. SERIES DE TIEMPO FINANCIERAS 11
5.2. MODELOS ESTADÍSTICOS TRADICIONALES PARA SERIES DE TIEMPO FINANCIERAS 13
5.2.1. MODELOS LINEALES PARA LA PREDICCIÓN DE LA MEDIA 13
5.2.2. MODELADO DE LA VOLATILIDAD Y VARIANZA CONDICIONAL 15
5.2.3. LIMITACIONES DE LOS ENFOQUES PARAMÉTRICOS LINEALES 18
5.3. MODELOS DE APRENDIZAJE AUTOMÁTICO Y DEEP LEARNING APLICADOS A SERIES DE TIEMPO FINANCIERAS 19
5.3.1. FUNDAMENTOS NO LINEALES Y PROCESAMIENTO SECUENCIAL 19
5.3.2. ARQUITECTURAS PROFUNDAS AVANZADAS PARA DINÁMICAS TEMPORALES 21
5.3.3. DESAFÍOS PRÁCTICOS Y EL PROBLEMA DE LA "CAJA NEGRA" 24
5.4. ENFOQUES HÍBRIDOS EN LA PREDICCIÓN DE SERIES DE TIEMPO FINANCIERAS 26
5.4.1. FUNDAMENTOS Y JUSTIFICACIÓN DE LA METODOLOGÍA HÍBRIDA 26
5.4.2. ESTRATEGIAS DE DESCOMPOSICIÓN Y MODELADO SECUENCIAL 28
5.4.3. ARQUITECTURAS DESTACADAS Y EVIDENCIA EMPÍRICA DE RENDIMIENTO 31
5.5. ANÁLISIS CRÍTICO SOBRE LOS MODELOS HÍBRIDOS 36
5.5.1. LA BRECHA ENTRE EL DESEMPEÑO ESTADÍSTICO Y LA UTILIDAD ECONÓMICA 36
5.5.2. EL DILEMA ENTRE PRECISIÓN PREDICTIVA E INTERPRETABILIDAD 38
5.5.3. COSTOS OPERATIVOS Y RIESGOS METODOLÓGICOS 40
6. CONCLUSIONES 43
7. RECOMENDACIONES 45
8. REFERENCIAS BIBLIOGRÁFICAS 47

1. RESUMEN

La predicción de series de tiempo financieras es una tarea compleja debido a que estos datos combinan componentes lineales y no lineales de forma simultánea, presentan alta volatilidad, no estacionariedad y son sensibles a factores económicos, políticos y psicológicos. Esta monografía analiza la eficacia predictiva de los modelos híbridos en comparación con los enfoques estadísticos y de aprendizaje automático utilizados de forma individual, mediante una revisión bibliográfica de la literatura científica reciente. El desarrollo teórico aborda tres paradigmas: los modelos estadísticos clásicos como ARIMA y la familia GARCH, útiles para capturar estructuras lineales y volatilidad pero limitados por sus supuestos rígidos; los modelos de Machine Learning y Deep Learning como LSTM, DBN y SNN, superiores en la captura de estructuras no lineales pero propensos al sobreajuste, costosos computacionalmente y difíciles de interpretar; y los modelos híbridos, que integran ambos enfoques de forma secuencial o mediante descomposición de señales para aprovechar las fortalezas de cada uno. Los estudios analizados muestran que arquitecturas como ARIMA-LSTM, ARIMA-LSSVM y ensambles con Filtro de Kalman superan de forma consistente a los modelos individuales en métricas de error como RMSE y MAE. Sin embargo, uno de los hallazgos más relevantes es que una menor tasa de error estadístico no garantiza una mayor rentabilidad económica, ya que predecir la dirección del movimiento del precio suele ser más valioso que predecir su valor exacto. Se concluye que el paradigma híbrido es el enfoque más completo para este tipo de datos, siempre que se evalúe bajo criterios que incluyan precisión direccional, interpretabilidad y solidez metodológica, y no únicamente métricas de error estándar.

Palabras clave: aprendizaje automático, aprendizaje profundo, modelos híbridos, predicción financiera, series de tiempo

2. ABSTRACT

Forecasting financial time series is a complex task given that this type of data simultaneously combines linear and non-linear components, exhibits high volatility and non-stationarity, and is sensitive to economic, political and psychological factors. This monograph analyzes the predictive effectiveness of hybrid models compared to statistical and machine learning approaches used individually, through a bibliographic review of recent scientific literature. The theoretical framework addresses three paradigms: classical statistical models such as ARIMA and the GARCH family, which are useful for capturing linear structure and volatility clustering but constrained by rigid parametric assumptions; Machine Learning and Deep Learning models such as LSTM, DBN and SNN, which are superior at capturing non-linear dynamics but prone to overfitting, computationally demanding and lacking interpretability; and hybrid models, which integrate both approaches sequentially or through signal decomposition to leverage the strengths of each. The analysis of empirical evidence shows that architectures such as ARIMA-LSTM, ARIMA-LSSVM and Kalman Filter ensembles consistently outperform individual models on error metrics such as RMSE and MAE. However, one of the most relevant findings is that a lower statistical error does not guarantee higher economic returns, as predicting the direction of price movement tends to be more valuable than predicting its exact value. The study concludes that the hybrid paradigm is the most complete approach for this type of data, provided it is evaluated under criteria that include directional accuracy, interpretability and methodological rigor, and not solely standard error metrics.

Keywords: deep learning, financial forecasting, hybrid models, machine learning, time series

3. INTRODUCCIÓN

La predicción precisa de las series de tiempo financieras tales como precios de acciones o tasas de cambio es una tarea esencial para que los inversionistas e instituciones gestionen riesgos y optimicen sus estrategias (Zhang et al., 2023). Sin embargo, obtener pronósticos fiables es muy desafiante debido a la naturaleza compleja e incierta de estos datos (Zhang et al., 2023). Las series financieras reales se caracterizan por presentar comportamientos no estacionarios como volatilidad y no linealidad, propiedades que representan obstáculos importantes para un modelado preciso (Zhang et al., 2023). Este desafío se ve reforzado por la Hipótesis del Mercado Eficiente (EMH), que postula que los precios reflejan toda la información disponible y, por lo tanto, siguen un patrón parecido a un recorrido aleatorio, lo que limita la capacidad de predicción (Henrique et al., 2023).

Tradicionalmente, el modelado de series de tiempo se basa en métodos estadísticos que asumen linealidad, como ARIMA o GARCH, pero estos han demostrado ser insuficientes al no poder capturar los patrones no lineales de datos reales (Zhang, 2003). La evolución hacia los modelos de aprendizaje profundo como las redes LSTM (*Long Short Term Memory*) ha marcado un avance importante, pues tienen la capacidad de adaptarse a patrones ocultos, superando en muchos casos a los enfoques tradicionales (Selvin et al., 2017). Sin embargo, existe una discusión en la literatura en cuanto a su utilidad práctica real (Henrique et al., 2023). Al simular escenarios de trading que imitan condiciones de mercado, donde se pronostica la dirección de cierre del día siguiente, modelos básicos de *Machine Learning* y *Deep Learning* han demostrado producir resultados similares a una adivinación aleatoria (Henrique et al., 2023). Esta incapacidad de los modelos básicos para superar el umbral aleatorio aporta más evidencia a favor de la EMH y resalta la falta de utilidad práctica de estos modelos para proporcionar información en la toma de decisiones (Henrique et al., 2023).

En este contexto, la literatura reciente ha mostrado un crecimiento en el interés por mejorar el rendimiento de las técnicas predictivas para series temporales financieras. Especialmente debido a que estos datos son no estacionarios, ruidosos, incompletos y con propiedades estadísticas que cambian en el tiempo (Nóbrega et al., 2014; Reid et al., 2014). Frente a este escenario los modelos de *Deep Learning* han ganado protagonismo por su capacidad de capturar patrones no lineales, su naturaleza no paramétrica y su menor sensibilidad a una mala especificación del modelo (Shen et al., 2015). Estos esfuerzos han impulsado el desarrollo de arquitecturas especializadas, como las *Deep Belief Networks* (DBN) mejoradas con *Continuous Restricted Boltzmann Machines* (CRBMs), que han mostrado mayor precisión y estabilidad en el pronóstico de tasas de cambio frente a métodos tradicionales (Shen et al., 2015). Asimismo, la literatura ha resaltado la importancia de adaptar modelos preentrenados a datos nuevos mediante estrategias de *Transfer Learning* que permiten extender el conocimiento previo sin acceder a los datos originales (Shabani et al., 2023).

De manera complementaria, varias metodologías avanzadas se han enfocado en aprovechar la dinámica temporal y las combinaciones de modelos para maximizar tanto el rendimiento predictivo como el económico. Las *Spiking Neural Networks* (SNN), y particularmente la *Polychronous Spiking Network* (PSN), han mostrado ser adecuadas para capturar la complejidad temporal y no lineal de las series financieras al codificar la información en tiempos de activación (Reid et al., 2014). Estas redes han superado a los modelos tradicionales en métricas económicas como el Retorno Anualizado. Por otro lado, múltiples estudios confirman que la combinación de pronósticos individuales puede ofrecer mejores resultados que los modelos aplicados de manera aislada (Nóbrega et al., 2014). Un ejemplo exitoso es la combinación de *Extreme Learning Machine* (ELM) y *Support Vector Regression* (SVR) con un filtro de Kalman, que mejoró tanto el rendimiento estadístico como el comercial en estrategias de *trading* estadístico (Nóbrega et al., 2014). Adicionalmente, técnicas de *Machine Learning* aplicadas al pronóstico de criptomonedas han mostrado alta eficacia, incluso utilizando únicamente valores

pasados de la serie como predictores, lo que resalta la flexibilidad de estos algoritmos frente a las limitaciones de los modelos tradicionales (Derbentsev et al., 2019).

Debido a la insuficiencia de los modelos tradicionales y la falta de poder predictivo práctico de las arquitecturas de *Machine Learning*/*Deep Learning* en el pronóstico a corto plazo, surge la necesidad de buscar arquitecturas más robustas y fiables que logren resultados superiores en términos de predictibilidad y significancia estadística (Henrique et al., 2023). Esto requiere considerar enfoques híbridos capaces de capturar tanto la estructura lineal y los patrones no lineales de las series financieras, dado que estudios recientes muestran que la combinación de modelos econométricos como GARCH con técnicas de aprendizaje automático como SVR puede mejorar la precisión del pronóstico frente al uso de un único modelo por separado (Aldeki, 2025). A partir de esta discusión surge la pregunta: ¿En qué medida los modelos híbridos pueden aportar un desempeño superior en la predicción de series de tiempo financieras frente a los enfoques estadísticos y de aprendizaje automático utilizados de manera individual?

Este proyecto de investigación es relevante debido a la complejidad de las series de tiempo financieras y al papel cada vez más notorio que han tomado los modelos híbridos en los últimos años. Pronosticar precios o índices financieros es difícil porque los mercados siguen dinámicas no lineales, cambiantes y ruidosas (Jaramillo et al., 2017). Los métodos clásicos de predicción, como los modelos de Box-Jenkins, suelen quedarse cortos frente a estas características (Jaramillo et al., 2017), por lo cual el interés académico y práctico se ha inclinado hacia técnicas capaces de manejar comportamientos no lineales (Zaffar y Hussain, 2022). En este sentido, los modelos de *Machine Learning* y *Deep Learning* son más adecuados para manejar la naturaleza caótica de los datos del mercado de valores, y generalmente superan a los modelos estadísticos y econométricos (Ampomah et al., 2020). Sin embargo, dado que los datos de series de tiempo financieras contienen componentes tanto lineales como no lineales, la predicción basada únicamente en modelos individuales no suele ser fiable (Xiao y

Su, 2022). Por ello, la combinación de múltiples métodos se vuelve una estrategia necesaria para alcanzar mejores niveles de precisión (Xiao y Su, 2022).

Más concretamente, la literatura reciente demuestra que los modelos híbridos pueden procesar eficazmente tanto la parte lineal como la no lineal de los datos (Xiao y Su, 2022). Por ejemplo, el modelo híbrido ARIMA-LSTM ha demostrado ser factible y mejora la precisión frente al desempeño de los modelos por separado (Xiao y Su, 2022). Este enfoque se alinea con el hallazgo de que los modelos que incluyen algún tipo de mejora o combinación con otro modelo resultan en una ganancia en la precisión de la predicción (Jaramillo et al., 2017). Además, aunque la precisión es importante, la interpretabilidad también es un componente crucial en las predicciones financieras que a menudo se pasa por alto (El Majzoub et al., 2023). Por esta razón, analizar cómo distintos modelos híbridos pueden equilibrar precisión y capacidad de interpretación aporta un enfoque necesario.

Finalmente, este trabajo contribuye al entendimiento de un tema clave para la toma de decisiones en los mercados financieros, un campo donde el uso de modelos híbridos ha mostrado resultados prometedores (Ampomah et al., 2020). Este trabajo busca aportar una visión crítica y fundamentada sobre el paradigma híbrido, útil tanto para investigadores como para quienes diseñan estrategias de inversión basadas en modelos cuantitativos.

4. OBJETIVOS

4.1. OBJETIVO GENERAL

Analizar la eficacia predictiva de los modelos híbridos en comparación con los enfoques estadísticos y de aprendizaje automático utilizados de forma individual en series de tiempo financieras, mediante una revisión sistemática de la literatura, con el fin de identificar su valor teórico y las ventajas que pueden aportar en el análisis y predicción de este tipo de datos.

4.2. OBJETIVOS ESPECÍFICOS

4.2.1. Describir los fundamentos teóricos y metodológicos de los modelos estadísticos, de aprendizaje automático y de los modelos híbridos más utilizados en la predicción de series de tiempo financieras.

4.2.2. Analizar los resultados reportados en investigaciones previas sobre el desempeño de los modelos híbridos frente a los modelos individuales, considerando las métricas de error y precisión empleadas en distintos contextos financieros.

4.2.3. Evaluar las principales ventajas y limitaciones de los modelos híbridos identificadas en la bibliografía, para formular un criterio sólido sobre su aplicación en el manejo de datos financieros reales.

5. FUNDAMENTO TEÓRICO

5.1. SERIES DE TIEMPO FINANCIERAS

Una serie de tiempo financiera se define como una secuencia finita de datos, como precios de activos o tasas de cambio, registrados e indexados de manera cronológica a lo largo de un período determinado (Zhang et al., 2023). En el ámbito del mercado de capitales, estas series reflejan la evolución de valores observados en intervalos temporales equitativos, que pueden abarcar desde frecuencias anuales o mensuales hasta datos diarios, horarios o incluso a nivel de ticks (minuto a minuto) en el comercio de alta frecuencia (Zhang et al., 2023). Desde una perspectiva analítica, estas series pueden abordarse de forma univariada, cuando consisten exclusivamente en las observaciones históricas de una variable objetivo (como el precio de cierre diario), o de forma multivariada, cuando integran simultáneamente otras variables exógenas con relaciones complejas, tales como el volumen de negociación, indicadores macroeconómicos o puntuaciones de sentimiento del mercado (Zhang et al., 2023).

A diferencia de otros conjuntos de datos, las series de tiempo financieras presentan características intrínsecas que dificultan su modelado. La primera de ellas es la no estacionariedad, es decir, la variabilidad de sus propiedades estadísticas (como la media y la varianza) a lo largo del tiempo (Zhang et al., 2023). Esta no estacionariedad es el resultado de interacciones complejas en los ciclos comerciales y económicos, como, por ejemplo, fluctuaciones estacionales en la demanda que generan efectos en cadena sobre los precios y las tasas de cambio (Reid et al., 2014). Además de los cambios de tendencia y estacionalidad, los datos financieros se caracterizan por una marcada no linealidad, volatilidad, heteroscedasticidad condicional y autocorrelación (Zhang et al., 2023). Los mercados operan como sistemas dinámicos que son caóticos, ruidosos y de memoria larga (Xiao y Su, 2022). Los precios no solo reaccionan a las leyes de la oferta y la demanda, sino que también son altamente

susceptibles a factores políticos, económicos, sociales e incluso psicológicos, lo que a menudo resulta en series de datos incompletas o con alta incertidumbre (Ampomah et al., 2020).

La dificultad en el modelado de estos datos se resalta teóricamente por la Hipótesis del Mercado Eficiente (EMH, por sus siglas en inglés). Propuesta originalmente por Fama (1970), la EMH sostiene que los precios de los activos reflejan instantáneamente toda la información disponible en el mercado (Zhang et al., 2023). La eficiencia del mercado se clasifica en tres formas distintas: débil, semi-fuerte y fuerte, cada una definida por el grado y tipo de información que ya está internalizada en los precios (Mao et al., 2026). Bajo este paradigma, cualquier intento de desarrollar estrategias comerciales rentables basadas en información histórica no podría seguir el ritmo de las fluctuaciones de los precios, haciendo que los movimientos del mercado se comporten de manera similar a un recorrido aleatorio (random walk) (Zhang et al., 2023). Este fenómeno se conoce en la literatura como el Dilema del Recorrido Aleatorio (RWD), el cual postula que los retornos son generados por procesos estocásticos que impiden un pronóstico consistente utilizando métodos convencionales (Henrique et al., 2023).

No obstante, a pesar de los sólidos argumentos de la EMH, la literatura empírica contemporánea sugiere que los mercados reales no son perfectamente eficientes y contienen patrones ocultos que pueden ser explotados. Diversos estudios han documentado la existencia de anomalías en el mercado de capitales que superan consistentemente el rendimiento aleatorio mediante el uso de señales predictivas (Shen et al., 2018). Esta predictibilidad se explica en gran medida a través de la lente de las finanzas conductuales, la cual argumenta que los sesgos cognitivos, los factores emocionales y las influencias sociales de los inversores crean ineficiencias sistemáticas en la fijación de precios, desafiando especialmente la forma semi-fuerte de la eficiencia del mercado (Mao et al., 2026).

Los métodos estadísticos tradicionales, al estar basados en supuestos lineales, no logran capturar adecuadamente los cambios abruptos ni las dinámicas no lineales de los mercados reales. Por eso, la investigación ha evolucionado hacia el aprendizaje automático y el aprendizaje profundo. Estos enfoques avanzados, al ser aproximadores de funciones no lineales, demuestran una capacidad superior para descifrar las relaciones ocultas y las dependencias a largo plazo dentro de la turbulencia de las series de tiempo financieras, permitiendo así pronósticos más precisos y rentables (Henrique et al., 2023; Jaramillo et al., 2017).

5.2. MODELOS ESTADÍSTICOS TRADICIONALES PARA SERIES DE TIEMPO FINANCIERAS

5.2.1. MODELOS LINEALES PARA LA PREDICCIÓN DE LA MEDIA

Históricamente, el intento de "adivinar el precio de mañana" en los mercados financieros se ha basado en la estadística y la econometría clásica. Por ello, la predicción de series de tiempo se ha abordado mediante modelos econométricos lineales, los cuales asumen que el valor futuro de una variable puede estimarse directamente como una función lineal de sus observaciones pasadas y de sus errores aleatorios históricos (Zhang, 2003). Esta metodología establece la base de cómo los analistas han intentado capturar el retorno medio o la tendencia futura de un activo antes de la adopción popular y masiva del aprendizaje automático.

La construcción de este paradigma predictivo recae en dos componentes fundamentales. Por un lado, tenemos el modelo Autorregresivo, denotado como AR(p), el cual dice que el valor actual de la serie depende linealmente de sus p valores previos o rezagos, operando bajo la premisa de que los precios del pasado contienen información estructuralmente válida para predecir el futuro (Zaffar y Hussain, 2022). Por otro lado, se encuentra el modelo de Media Móvil, o MA(q), que aborda la predicción desde una perspectiva distinta: estima el valor actual de la serie no a partir de los precios

pasados en sí, sino como una combinación lineal de los términos de error aleatorio observados en los q periodos anteriores (Zaffar y Hussain, 2022).

La fusión de ambos enfoques da lugar al modelo ARMA(p,q), el cual fue diseñado para capturar dinámicas más complejas al combinar la dependencia sobre los valores pasados y sobre los errores pasados de manera simultánea (Zaffar y Hussain, 2022). No obstante, la familia de modelos ARMA tiene un prerrequisito matemático para poder funcionar: requiere que la serie de tiempo sea estacionaria, lo que significa que sus propiedades estadísticas, como la media y la estructura de autocorrelación, deben permanecer constantes a lo largo del tiempo (Cappello et al., 2025; Zhang, 2003).

En el contexto financiero del mundo real, una característica persistente en los precios de las acciones, los tipos de cambio o los índices bursátiles es precisamente su naturaleza no estacionaria, la cual se manifiesta frecuentemente a través de tendencias marcadas, ciclos variables y cambios de régimen (Zhang, 2003). Para solucionar esta limitación e introducir la estacionariedad requerida en los datos, se agrega el componente Integrado (I), dando origen al modelo Autorregresivo Integrado de Media Móvil, comúnmente conocido como ARIMA(p,d,q) (Zhu y Wei, 2012). El nuevo parámetro d representa el número de diferencias o restas sucesivas que se debe aplicar a la serie original (por ejemplo, calcular el retorno restando el precio de ayer al precio de hoy) hasta lograr que la serie pierda su tendencia y estabilice su varianza, volviéndose estacionaria (Cappello et al., 2025).

Para aplicar estos modelos en la práctica, el estándar de la industria estadística es la rigurosa metodología popularizada por Box y Jenkins en la década de 1970 (Cappello et al., 2025). Esta metodología proporciona un enfoque sistemático e iterativo que consta de pasos definidos para determinar la arquitectura óptima del modelo predictivo: el primer paso es la identificación del modelo para fijar el orden (p,d,q), apoyándose en el análisis visual y cuantitativo de la función de

autocorrelación (ACF) y la función de autocorrelación parcial (PACF) de los datos muestrales; el segundo paso es la estimación de los coeficientes del modelo, para lo cual se suelen emplear técnicas como la estimación de máxima verosimilitud; el tercer paso consiste en la verificación o diagnóstico del modelo, donde se examinan los residuos para confirmar que se comporten como ruido blanco, es decir, que sean independientes y carezcan de patrones remanentes no explicados; y finalmente, se utiliza el modelo ajustado para extrapolar y predecir los precios futuros de los activos (Cappello et al., 2025; Zhang, 2003; Zhu y Wei, 2012).

En términos generales, los modelos de la familia ARIMA se consolidan como las herramientas lineales más populares y estructuradas en el análisis predictivo de series de tiempo, y sirven como punto de referencia para comparar el desempeño de las técnicas modernas basadas en aprendizaje automático (Zhang, 2003).

5.2.2. MODELADO DE LA VOLATILIDAD Y VARIANZA CONDICIONAL

En el análisis de series de tiempo financieras, la predicción del precio o rendimiento exacto es una tarea bastante compleja y, en muchos escenarios, insuficiente. Por ello, predecir el riesgo o la incertidumbre a través de la volatilidad es a menudo tan o más importante que predecir el precio direccional de un activo (Aldeki, 2025). Dado que los modelos econométricos clásicos orientados a la media, como el ARIMA, asumen una varianza constante a lo largo del tiempo (homocedasticidad), resultan incapaces de capturar las fluctuaciones dinámicas del riesgo en el mercado real (Cappello et al., 2025). Como predecir el precio exacto del mañana es una tarea esquiva debido a la naturaleza caótica de los mercados financieros, los estadísticos desarrollaron modelos especializados no para adivinar el valor futuro, sino para estimar la magnitud o fuerza con la que se producirán los movimientos del mercado (Aldeki, 2025).

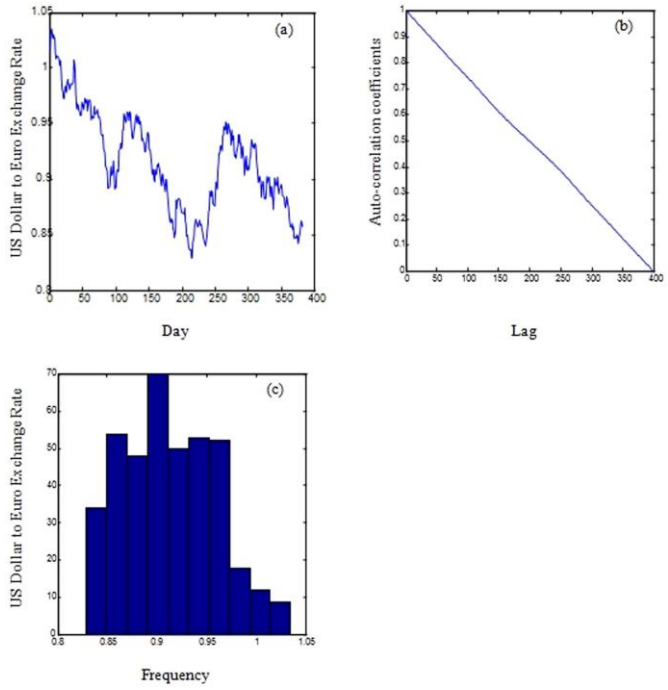


Figura 1. Tasa de cambio entre el dólar estadounidense y el euro (a), correlograma de la señal (b) e histograma de la tasa de cambio (c) (Reid et al., 2014).

Esta necesidad de modelar la varianza condicional surge de la observación de una característica omnipresente en los datos financieros conocida como "agrupamiento de volatilidad" (volatility clustering). El agrupamiento de volatilidad se define como la tendencia de que los periodos de alta turbulencia y grandes cambios en los precios sean seguidos por fluctuaciones igualmente grandes (de cualquier signo), mientras que los periodos de baja volatilidad tienden a ser seguidos por pequeños cambios (Aldeki, 2025). Para capturar y modelar esta varianza cambiante en el tiempo y superar la limitación de los modelos lineales que asumen varianzas estáticas, se introdujeron clases de modelos no lineales (Cappello et al., 2025).

El esfuerzo más significativo en esta dirección fue el modelo de Heteroscedasticidad Condicional Autorregresiva (ARCH, por sus siglas en inglés), propuesto originalmente por Engle en 1982. El modelo ARCH es un enfoque univariado, basado fundamentalmente en los rendimientos históricos de los activos, que asume que la varianza actual de los errores de un modelo no es constante,

sino que depende matemáticamente de los errores al cuadrado (o choques) de los periodos pasados (Aldeki, 2025). Al modelar la varianza de hoy en función de la magnitud de los choques recientes, el modelo ARCH permite replicar el fenómeno de los periodos de turbulencia agrupados (Aldeki, 2025).

Sin embargo, para capturar adecuadamente los datos financieros que presentan exceso de curtosis (colas pesadas en las distribuciones), el modelo ARCH a menudo requería una longitud de rezago demasiado grande para ser computacionalmente eficiente (Aldeki, 2025). Para resolver este inconveniente, Bollerslev desarrolló una extensión y generalización conocida como el modelo GARCH (Generalized ARCH) (Aldeki, 2025). El modelo GARCH mejora al ARCH original al incorporar no solo los errores al cuadrado pasados, sino también un número definido de varianzas condicionales rezagadas en su propia ecuación (Aldeki, 2025). Esto lo convierte en un modelo multivariado en el sentido de que actúa como un proceso autorregresivo de medias móviles para la varianza condicional, otorgándole una estructura más flexible para capturar la persistencia de la volatilidad a largo plazo (Aldeki, 2025). En la práctica y en la literatura financiera, el modelo GARCH es reconocido como el estándar tradicional por excelencia para evaluar y predecir los efectos de agrupamiento de volatilidad en índices bursátiles y tasas de cambio (Wang, 2024).

A pesar de su indudable impacto en la econometría y su utilidad para la fijación de precios de activos, la familia de modelos ARCH/GARCH presenta limitaciones. Estos modelos imponen suposiciones sobre la forma funcional del proceso de generación de datos y asumen distribuciones de probabilidad predefinidas para los términos de error (Aldeki, 2025; Zhang et al., 2023). Además, en su formulación estándar, el modelo GARCH es simétrico, lo que implica que un choque de mercado positivo (buenas noticias) o negativo (malas noticias) de la misma magnitud genera exactamente el mismo impacto en la estimación de la varianza condicional (Aldeki, 2025). Esta simetría les impide modelar adecuadamente el efecto de apalancamiento (leverage effect), una realidad en la cual las caídas de los precios tienden a incrementar la volatilidad de forma mucho más drástica que las subidas

(Aldeki, 2025). A causa de estas restricciones y un rendimiento que puede decaer durante periodos de cambio en el mercado, la investigación reciente ha comenzado a explorar enfoques de aprendizaje automático para superar las debilidades de la estadística tradicional en la predicción del riesgo (Aldeki, 2025; Zhang et al., 2023).

5.2.3. LIMITACIONES DE LOS ENFOQUES PARAMÉTRICOS LINEALES

Aunque modelos como ARIMA y la familia GARCH fueron los pilares de la econometría en el siglo pasado, hoy encuentran limitaciones claras al aplicarse en los mercados financieros modernos. El problema principal radica en su estructura paramétrica, la cual impone restricciones muy rígidas. En el caso de ARIMA, su mayor debilidad es que asume una relación lineal predefinida, bajo la idea de que los valores futuros son simplemente una combinación lineal de los datos pasados (Zhang, 2003). Además, estos modelos obligan a que los datos cumplan requisitos estadísticos, como transformar las series mediante diferenciación para forzar la estacionariedad antes de trabajar con ellas (Cappello et al., 2025). De igual forma, modelos de volatilidad como GARCH, diseñados para medir el riesgo, dependen de suposiciones fijas sobre la distribución de los errores y la forma de sus funciones, lo que limita su capacidad para adaptarse a la complejidad real del mercado (Aldeki, 2025).

Estas suposiciones, aunque teóricamente consistentes, no logran representar la naturaleza real de los datos financieros. En la práctica, los mercados de valores funcionan como sistemas dinámicos complejos, caracterizados por ser ruidosos, caóticos y cambiantes en el tiempo (Jaramillo et al., 2017). Los precios no se mueven de manera ordenada ni mantienen una varianza constante; por el contrario, responden a ciclos económicos y variaciones estacionales que modifican sus propiedades estadísticas continuamente (Reid et al., 2014). Esta complejidad implica que los activos financieros presentan, de manera simultánea, patrones lineales y no lineales (Zhu y Wei, 2012; Xiao y Su, 2022). Al final, estas dinámicas generan dependencias ocultas y cambios en la volatilidad que las funciones

lineales o las distribuciones predefinidas de los modelos tradicionales simplemente no pueden capturar.

Debido a esta falta de ajuste, cuando las series financieras contienen patrones no lineales complejos, modelos como ARIMA suelen fallar al intentar capturarlos (Zhu y Wei, 2012). Esto provoca que el error de predicción crezca significativamente bajo condiciones reales de mercado. Este escenario define el problema central del pronóstico actual: aunque la econometría clásica permitió filtrar tendencias básicas, la complejidad de los mercados modernos exige superar las suposiciones de los modelos lineales. Por ello, es necesaria la transición hacia algoritmos más flexibles, como el Aprendizaje Automático (Machine Learning) y el Aprendizaje Profundo (Deep Learning). Estas técnicas tienen la capacidad de adaptarse automáticamente a dependencias no lineales en los datos sin depender de supuestos sobre su distribución (Zhang et al., 2023; Aldeki, 2025). En última instancia, las limitaciones técnicas de los enfoques tradicionales son las que justifican el uso de inteligencia artificial para intentar descifrar el comportamiento del entorno financiero.

5.3. MODELOS DE APRENDIZAJE AUTOMÁTICO Y DEEP LEARNING APLICADOS A SERIES DE TIEMPO FINANCIERAS

5.3.1. FUNDAMENTOS NO LINEALES Y PROCESAMIENTO SECUENCIAL

Frente a las limitaciones que imponen los modelos estadísticos clásicos, el aprendizaje automático (Machine Learning) y el aprendizaje profundo (Deep Learning) se han ido consolidando como la alternativa más utilizada en el pronóstico financiero (Zhang et al., 2023). Lo que hace atractivos a estos enfoques es que pueden aprender directamente de los datos y capturar tanto patrones lineales como no lineales, sin que sea necesario que un experto defina de antemano la estructura del modelo (Zhang et al., 2023). A diferencia de métodos como ARIMA, los algoritmos de inteligencia artificial no

exigen que las series cumplan condiciones estadísticas específicas antes de ser procesadas (Derbentsev et al., 2019). Un ejemplo concreto es el de intentar predecir el índice S&P 500: con un modelo estadístico clásico, el analista necesita transformar los datos previamente para lograr que la serie sea estacionaria; con una red neuronal, en cambio, es posible ingresar los datos directamente o junto a indicadores técnicos, y el modelo se encarga de identificar las relaciones relevantes por su cuenta (Henrique et al., 2023).

Dentro de este paradigma, las Redes Neuronales Artificiales (ANN) son la base, y el Perceptrón Multicapa (MLP) es una de sus arquitecturas más conocidas (Kabir et al., 2025). En términos prácticos, estas redes funcionan como aproximadores de funciones no lineales: utilizan capas de entrada, ocultas y de salida junto a funciones de activación para transformar las variables de entrada y extraer patrones relevantes, lo que les da una buena capacidad de generalización (Kabir et al., 2025; Henrique et al., 2023; Zhang, 2003). Para entender cómo opera un MLP, pensemos en el caso de querer predecir si el precio de una acción como Amazon subirá o bajará al día siguiente. El modelo recibiría como entrada un conjunto de indicadores técnicos del mercado en ese momento, como el Índice de Fuerza Relativa (RSI) o las medias móviles, que representan una especie de "foto" del estado actual del mercado (Henrique et al., 2023). Con esa información, la red procesa todos los valores a través de sus capas ocultas, asignando mayor o menor peso a cada indicador según su relevancia, y produce una clasificación final: "alza" (UP) o "baja" (DOWN) (Henrique et al., 2023).

Sin embargo, arquitecturas como el MLP tienen una limitación importante cuando se trata de series de tiempo financieras: procesan todas las observaciones de entrada como si fueran independientes entre sí, ignorando el orden en que ocurrieron (Zhang et al., 2023). En finanzas esto es un problema, porque el precio de un activo hoy no es independiente de lo que pasó ayer o la semana pasada; la secuencia en que se dan los movimientos del mercado importa (Shen et al., 2018).

Para resolver esta limitación, se introdujeron las Redes Neuronales Recurrentes (RNN), diseñadas específicamente para trabajar con datos secuenciales (Zhang et al., 2023). A diferencia del MLP, las RNN incorporan conexiones que permiten que la información de pasos anteriores influya en el procesamiento del paso actual, lo que les da una especie de memoria sobre lo que ocurrió antes (Zhang et al., 2023; Shen et al., 2018). Retomando el ejemplo del pronóstico de precios: si se quiere predecir la rentabilidad de un activo usando los datos de los últimos 20 días, un MLP tomaría esos 20 días todos a la vez. Una RNN, en cambio, los procesa uno por uno en orden cronológico: analiza el día 1 y guarda un estado interno con esa información, luego analiza el día 2 combinándolo con lo que "recuerda" del día anterior, y así sucesivamente hasta el día 20 (Shen et al., 2018). De esta forma, las RNN logran capturar la evolución temporal de los precios de una manera que los modelos tradicionales no pueden.

5.3.2. ARQUITECTURAS PROFUNDAS AVANZADAS PARA DINÁMICAS TEMPORALES

Aunque las RNN estándar resolvieron el problema del procesamiento secuencial, tienen una limitación técnica importante conocida como el "desvanecimiento del gradiente". En términos simples, esto significa que a medida que la red procesa nuevos datos, la influencia de las observaciones más antiguas se va perdiendo, lo que hace que el modelo "olvide" información del pasado que puede ser relevante en los mercados financieros. Además, como cada predicción sirve de entrada para la siguiente, los errores se van acumulando con el tiempo (Kim et al., 2026). Para resolver esto, se desarrollaron las redes de Memoria a Corto y Largo Plazo (LSTM), que incorporan una celda de memoria interna controlada por tres compuertas: entrada, olvido y salida (Kabir et al., 2025; Kim et al., 2026). Un ejemplo concreto de cómo funciona esto: si el mercado reacciona de forma brusca a un evento económico, la compuerta de olvido descarta el ruido de las pequeñas fluctuaciones previas, mientras que la compuerta de entrada decide qué información es realmente importante guardar para identificar la tendencia (Kim et al., 2026). Esto hace que las LSTM sean más robustas frente al ruido y

a los datos faltantes, dos características habituales en las series financieras (Kabir et al., 2025). Como alternativa más ligera, las Unidades Recurrentes Cerradas (GRU) simplifican este mecanismo combinando las compuertas en dos: actualización y reinicio, logrando un procesamiento similar, pero con menor costo computacional.

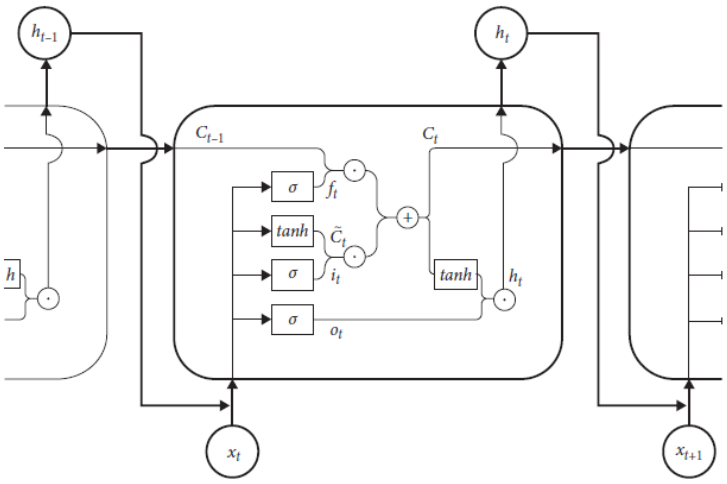


Figura 2. Estructura interna de las celdas en una cuadrícula neuronal de Memoria a Corto y Largo Plazo (LSTM) (Xiao y Su, 2022).

Más allá de las arquitecturas basadas en memoria, las Redes de Creencia Profunda (DBN) surgieron como una alternativa para manejar la alta complejidad y dimensionalidad de los datos financieros. A diferencia de las redes neuronales convencionales, que se entrenan de forma supervisada en un solo proceso, las DBN se construyen apilando múltiples bloques llamados Máquinas de Boltzmann Restringidas (RBM) y se entrenan de manera no supervisada, capa por capa (Shen et al., 2015). Lo que esto significa en la práctica es que la red primero aprende a identificar las características más relevantes de los datos y usa eso para inicializar sus parámetros antes de hacer cualquier predicción, lo que le da una mejor base de partida (Shen et al., 2015). Para trabajar con series financieras, se utiliza una variante llamada Máquinas de Boltzmann Restringidas Continuas (CRBM), que permite al modelo procesar valores continuos como tasas de cambio, en lugar de limitarse a datos binarios (Shen et al., 2015). Este proceso de preentrenamiento por etapas es clave porque reduce el

riesgo de que el modelo quede atrapado en soluciones subóptimas durante la optimización, un problema frecuente en los métodos de entrenamiento clásicos.

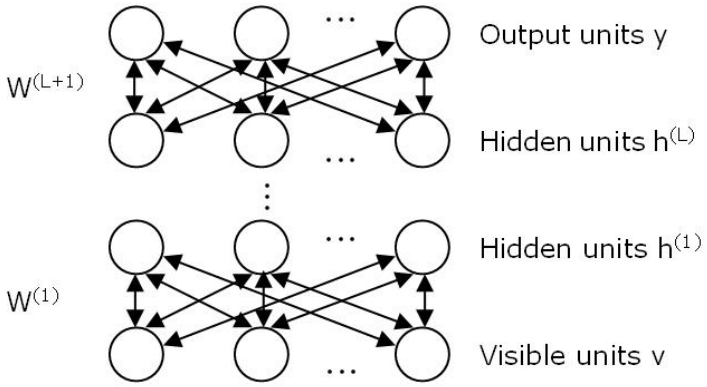


Figura 3. Arquitectura de una Red de Creencia Profunda (DBN) estructurada con múltiples capas ocultas (Shen et al., 2015).

Por otro lado, las Redes Neuronales de Picos (SNN) representan un enfoque distinto, especialmente útil para entornos de mercado volátiles y no estacionarios (Reid et al., 2014). A diferencia de modelos como el MLP o las LSTM, que procesan la información usando valores de activación continuos, las SNN codifican la información a través del momento exacto en que ocurren los disparos neuronales o spikes (Reid et al., 2014). Una arquitectura particular dentro de esta familia es la Red de Picos Policrónica (PSN), que usa reglas de aprendizaje inspiradas en la biología para detectar patrones a partir del retraso entre señales de entrada (Reid et al., 2014). En términos prácticos, ante un shock geopolítico que afecta los precios, una SNN no solo evalúa la magnitud del movimiento sino también el momento en que ocurrió la anomalía, lo que le permite reaccionar de forma más precisa a eventos abruptos. Esta capacidad para trabajar con información temporal discreta las hace adecuadas para capturar dinámicas de mercado que modelos basados en promedios tienden a suavizar o perder (Reid et al., 2014).

5.3.3. DESAFÍOS PRÁCTICOS Y EL PROBLEMA DE LA "CAJA NEGRA"

Aunque arquitecturas como LSTM, DBN o SNN han demostrado una mayor precisión predictiva y una capacidad notable para capturar dinámicas no lineales del mercado, su aplicación en contextos reales trae consigo una serie de compromisos técnicos y operativos que no pueden ignorarse. El aprendizaje automático y el aprendizaje profundo no son soluciones universales, sino herramientas probabilísticas que enfrentan limitaciones prácticas y que, si no se gestionan adecuadamente, pueden comprometer tanto la fiabilidad como la utilidad de los resultados.

El primer gran obstáculo es el sobreajuste (overfitting). Las redes neuronales son modelos muy flexibles con miles o incluso millones de parámetros, lo que les da una tendencia a memorizar los datos de entrenamiento en lugar de aprender patrones generalizables del mercado. En un entorno financiero ruidoso y caótico, esto significa que un modelo puede terminar ajustándose a fluctuaciones aleatorias del mercado, confundiéndolas con patrones reales (Gerlein et al., 2016). Un ejemplo concreto: si entrenamos un modelo para predecir el precio de una criptomoneda y este se sobreajusta, podría asociar erróneamente una caída del precio con un día de la semana específico simplemente porque esa coincidencia apareció en los datos históricos, lo que haría que fallara al operar con datos nuevos. Para reducir este riesgo y lograr que el modelo generalice bien, se necesitan grandes volúmenes de datos de entrenamiento (Zhang et al., 2023). Sin embargo, en finanzas esto genera una paradoja: usar datos demasiado antiguos, como precios de hace 20 años, puede ser igual de problemático, ya que el comportamiento del mercado y sus regulaciones cambian con el tiempo, haciendo que esa información ya no sea representativa del contexto económico actual.

El segundo desafío está directamente relacionado con el anterior: el alto costo computacional. A diferencia de modelos como ARIMA, que pueden estimarse con relativa rapidez usando hardware convencional, las redes neuronales profundas requieren ajustar una gran cantidad de parámetros de

forma iterativa durante el entrenamiento (Nóbrega et al., 2014). Este proceso demanda hardware especializado como Unidades de Procesamiento Gráfico (GPUs), además de un consumo considerable de tiempo y energía (Shabani et al., 2023). En contextos como el trading de alta frecuencia, donde los modelos necesitan reentrenarse constantemente para mantenerse al día con los nuevos datos del mercado, este costo deja de ser solo un problema técnico y puede llegar a afectar directamente la rentabilidad del sistema (Shabani et al., 2023).

Quizás la limitación más relevante para la adopción de estas tecnologías en la industria financiera es el problema de la "caja negra". A diferencia de la estadística clásica o la econometría, donde cada coeficiente tiene un significado claro e indica en qué magnitud y dirección una variable afecta al precio, las redes neuronales profundas no ofrecen esa transparencia (Zhang et al., 2023). Un modelo avanzado puede procesar miles de indicadores técnicos y generar una predicción con alta precisión, pero no es capaz de explicar por qué llegó a esa conclusión. Esto es un problema concreto en contextos institucionales: si un algoritmo recomienda liquidar una cartera de gran valor, los analistas de riesgo y gestores de fondos necesitan entender la lógica detrás de esa decisión para justificarla ante auditores, reguladores y directivos sin formación técnica (El Majzoub et al., 2023).

En conclusión, el panorama actual del pronóstico financiero presenta una disyuntiva clara. Los modelos de Deep Learning tienen una capacidad predictiva superior a la de los enfoques estadísticos clásicos para capturar la complejidad del mercado, pero esto tiene un costo: requieren grandes volúmenes de datos para no caer en sobreajuste, demandan un poder de cómputo considerable y, quizás lo más problemático en la práctica, producen decisiones automatizadas que no son posibles de auditar ni interpretar directamente.

5.4. ENFOQUES HÍBRIDOS EN LA PREDICCIÓN DE SERIES DE TIEMPO FINANCIERAS

5.4.1. FUNDAMENTOS Y JUSTIFICACIÓN DE LA METODOLOGÍA HÍBRIDA

El comportamiento de los precios bursátiles combina, al mismo tiempo, patrones con cierto grado de predictibilidad y dinámicas difíciles de capturar. Ante esta complejidad, la literatura ha llegado a un punto de acuerdo bastante extendido: no existe un único modelo predictivo que funcione como el "mejor" en todas las situaciones de mercado (Zhang, 2003). Como respuesta a este problema, tanto teórico como práctico, surge la metodología híbrida. Un modelo híbrido se define como la integración de dos o más algoritmos predictivos, generalmente combinando enfoques estadísticos clásicos con técnicas de inteligencia artificial, con el objetivo de aprovechar las fortalezas de cada uno y compensar sus limitaciones individuales (Zhang, 2003; Kabir et al., 2025). En lugar de exigirle a un solo modelo que capture todo el comportamiento del mercado, el enfoque híbrido divide el problema, asignando a cada componente la tarea para la cual fue diseñado (Zhu y Wei, 2012).

La justificación técnica de esta integración parte de una característica de los datos financieros: las series de tiempo del mundo real raramente son puramente lineales o puramente no lineales, sino que contienen ambos componentes de forma simultánea (Zhang, 2003; Zhu y Wei, 2012). Para entender esto, basta con observar el comportamiento de un índice bursátil: por un lado, suele mostrar una tendencia a largo plazo relativamente estable, impulsada por factores macroeconómicos como el crecimiento económico o la inflación. Pero sobre esa tendencia actúan fuerzas no lineales, como los agrupamientos de volatilidad, los ciclos económicos, los shocks por eventos geopolíticos y el ruido generado por la psicología colectiva de los inversores (Fenghua et al., 2014; Aldeki, 2025). Desde una perspectiva matemática, el paradigma híbrido aborda esta dualidad asumiendo que el precio de un activo en un momento dado puede descomponerse en dos partes: un componente lineal subyacente

y un componente no lineal residual que el primer modelo no logra capturar (Zhang, 2003; Zhu y Wei, 2012).

Reconocida esta naturaleza dual, se entiende mejor por qué los modelos evaluados de forma individual presentan limitaciones importantes. Los modelos estadísticos tradicionales como ARIMA o GARCH son útiles para capturar la estructura lineal de autocorrelación o proyectar una varianza base, pero su rendimiento se deteriora considerablemente cuando se enfrentan a relaciones no lineales, precisamente porque sus supuestos son demasiado rígidos para adaptarse a ese tipo de dinámicas (Zhu y Wei, 2012; Aldeki, 2025). En el otro extremo, los modelos de Machine Learning y Deep Learning son muy eficaces para capturar patrones no lineales complejos, pero cuando se los aplica a estructuras predominantemente lineales tienden a producir resultados inestables o subóptimos (Zhang, 2003). A esto se suma su propensión al sobreajuste, ya que su alta flexibilidad puede llevarlos a memorizar el ruido del mercado en lugar de aprender patrones generalizables, además de la ya mencionada limitación de operar como "cajas negras" con baja interpretabilidad (Zhang et al., 2023).

Queda claro entonces que ni la estadística clásica ni la inteligencia artificial ofrecen por sí solas una solución completa para el pronóstico financiero, lo que hace necesaria una combinación estructurada de ambos enfoques. Al combinar ambos enfoques, por ejemplo, usando un modelo ARIMA para capturar la tendencia lineal y luego una red neuronal o una Máquina de Vectores de Soporte para modelar los residuos no lineales, se construye un sistema predictivo más completo que aprovecha lo mejor de cada parte (Zhang, 2003; Kabir et al., 2025). Esta integración no solo mejora la precisión del pronóstico, sino que también aporta mayor robustez frente a los cambios de régimen del mercado, lo que en la práctica se traduce en una toma de decisiones más confiable y estable (Zhu y Wei, 2012; Zhang et al., 2023).

5.4.2. ESTRATEGIAS DE DESCOMPOSICIÓN Y MODELADO SECUENCIAL

El enfoque híbrido en la predicción de series de tiempo financieras parte de una lógica bastante directa: en lugar de exigirle a un solo algoritmo que maneje toda la complejidad del mercado, se propone una colaboración secuencial entre modelos. La idea es que cada componente del sistema se encargue de la parte del problema para la cual está mejor preparado: primero, un modelo econométrico clásico se encarga de identificar y modelar la estructura lineal de la serie, y luego un modelo de aprendizaje automático toma los residuos de ese primer ajuste para capturar los patrones no lineales que quedaron sin explicar.

El núcleo de esta integración es el paradigma de descomposición de errores. Este enfoque parte del supuesto de que una serie de tiempo financiera (Y_t) no es puramente lineal ni puramente no lineal, sino que contiene ambos tipos de patrones de forma simultánea, es decir, un componente lineal (L_t) y un componente no lineal (N_t) superpuestos (Zhu y Wei, 2012). Para capturar esta naturaleza dual sin que un solo modelo tenga que lidiar con ambas partes a la vez, el proceso de pronóstico se estructura en tres fases secuenciales que se describen a continuación:

- **Fase 1 (Filtrado Lineal):** En esta primera etapa se aplica un modelo estadístico, típicamente de la familia ARIMA, sobre la serie de tiempo original Y_t. El objetivo es que este modelo actúe como un filtro, capturando la estructura de autocorrelación lineal y las tendencias predecibles presentes en los datos para generar un pronóstico inicial $\hat{L}_t$ (Zhang, 2003).

- **Fase 2 (Extracción de Residuos):** Dado que el modelo lineal no puede capturar toda la dinámica del mercado, sus predicciones tendrán un margen de error. En esta fase se calcula ese error o "residuo" e_t, que no es más que la diferencia entre el valor real de la serie y el pronóstico lineal obtenido en la fase anterior $(e_t = Y_t - \hat{L}_t)$ (Zhang, 2003). Este paso es clave

en la metodología híbrida: al aislar la parte lineal, los residuos resultantes concentran la información no lineal que el primer modelo no logró explicar, incluyendo la volatilidad y las dependencias ocultas en los datos.

- **Fase 3 (Modelado No Lineal):** En esta última etapa se aplican algoritmos de Machine Learning o Deep Learning tomando los residuos e_t como variables de entrada. Al entrenar el modelo exclusivamente sobre los errores del primer ajuste, este puede concentrarse en identificar las relaciones no lineales que el modelo estadístico no logró capturar (Zhu y Wei, 2012). El pronóstico final del sistema híbrido se obtiene sumando ambas predicciones: la lineal de la primera fase y la no lineal de esta tercera fase ($\hat{Y}_t = \hat{L}_t + \hat{N}_t$) (Zhang, 2003).

Para entender cómo esta lógica secuencial se aplica en la práctica, es útil revisar cómo la literatura ha implementado estas estrategias. El trabajo de Zhang (2003) es el referente más citado en este enfoque: en su investigación desarrolló la arquitectura híbrida aditiva ARIMA-ANN y demostró empíricamente que usar una red neuronal artificial para modelar los residuos de un modelo ARIMA producía mejores resultados que ambos modelos operando por separado. Su conclusión principal fue que separar las tareas de reconocimiento de patrones no solo mejora la precisión, sino que también reduce el riesgo de sobreajuste al evitar que un único modelo intente capturarlo todo a la vez.

Esta misma lógica fue retomada y extendida por Zhu y Wei (2012) en el contexto de la predicción de precios en el mercado de carbono. En su investigación, los autores siguieron las tres fases descritas: aplicaron ARIMA para el filtrado lineal, extrajeron los residuos, y en la tercera fase introdujeron una variación interesante al reemplazar las redes neuronales clásicas por Máquinas de Vectores de Soporte de Mínimos Cuadrados (LSSVM), optimizadas mediante algoritmos de enjambre de partículas (PSO). Sus resultados confirmaron que aislar matemáticamente el residuo en la Fase 2

permite que el componente de Machine Learning se concentre en la no linealidad con mayor efectividad, logrando una precisión superior a la de modelos individuales como ANN o SVM estándar.

Vale la pena destacar que la evolución del campo ha llevado el concepto de descomposición más allá del enfoque secuencial tradicional. Mientras que trabajos como los de Zhang (2003) y Zhu y Wei (2012) descomponen la serie a partir de los errores del modelo lineal, investigaciones más recientes como las de Zhang et al. (2023) y Kim et al. (2026) aplican la descomposición directamente sobre la señal original. Estos trabajos utilizan estrategias como la Descomposición Empírica de Modos por Conjuntos Completa (CEEMD) o la transformación Wavelet, que en lugar de extraer un residuo lineal, separan la serie de precios en múltiples sub-series de distintas frecuencias, conocidas como Funciones de Modo Intrínseco (IMFs), agrupando de forma separada el ruido, la estacionalidad y la tendencia. Sobre cada uno de estos componentes se aplican luego arquitecturas profundas como LSTM o Crossformer de manera paralela, lo que permite que cada modelo se especialice en una frecuencia específica de la dinámica del mercado.

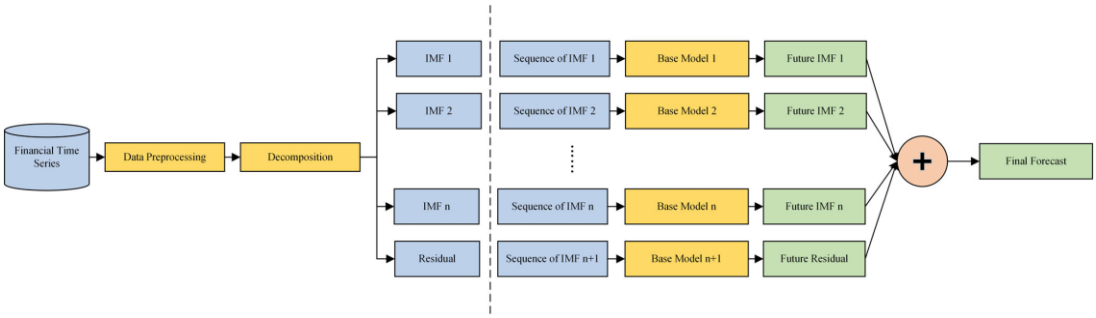


Figura 4. Ensamble por descomposición aplicado al pronóstico de precios financieros (Zhang et al., 2023).

Por su parte, el trabajo reciente de Cappello et al. (2025) reconoce el impacto histórico del modelo de filtrado de residuos de Zhang (2003), pero propone una variante integradora diferente. En lugar de modelar los errores secuencialmente, su metodología entrena un modelo ARIMA y modelos

avanzados (LSTM/GRU) por separado, para luego fusionar sus salidas mediante una optimización de combinación convexa que pondera dinámicamente qué modelo tiene menor error en un momento dado del tiempo.

En conjunto, estos enfoques muestran que el paradigma de descomposición ya sea separando los errores de forma secuencial o dividiendo la serie en componentes de distintas frecuencias, es el principio metodológico central que permite abordar la complejidad de las series financieras de manera estructurada, asignando a cada modelo la parte del problema para la que está mejor preparado.

5.4.3. ARQUITECTURAS DESTACADAS Y EVIDENCIA EMPÍRICA DE RENDIMIENTO

La evidencia empírica en la literatura reciente muestra de forma consistente que los modelos híbridos superan a los enfoques individuales en el pronóstico financiero, lo que ha llevado esta idea del plano teórico a convertirse en un resultado bien respaldado por datos y métricas de error concretas. Cuando se combinan modelos econométricos, algoritmos de aprendizaje automático y arquitecturas profundas de forma estructurada, el sistema resultante tiende a producir mejores resultados que cualquiera de sus componentes por separado. El análisis de las métricas de error reportadas en distintas investigaciones permite observar este patrón con claridad: el ensamblaje de modelos es una estrategia más robusta y competitiva para enfrentar la volatilidad y complejidad de los mercados financieros modernos.

A continuación, se describen los ensamblajes clave documentados en la literatura reciente, detallando cómo su arquitectura resolvió problemas específicos del mercado y la evidencia cuantitativa de su éxito:

- **Integración Econométrica con Vectores de Soporte Optimizados (ARIMA-LSSVM):** En su investigación sobre la predicción de precios del carbono, Zhu y Wei (2012) propusieron un modelo híbrido que utiliza ARIMA para filtrar la tendencia lineal, combinado con Máquinas de Vectores de Soporte de Mínimos Cuadrados (LSSVM) optimizadas por un algoritmo de enjambre de partículas (PSO) para modelar los residuos no lineales. El hallazgo más impactante de este estudio no fue solo la reducción del error absoluto, sino su gran capacidad de pronóstico direccional. Mientras que los modelos individuales (ARIMA o SVM) fallaban en detectar los puntos de inflexión del mercado, la arquitectura híbrida alcanzó un estadístico direccional (D_{stat}) del 97.98% en la predicción fuera de muestra, demostrando que separar el ruido lineal permite a la inteligencia artificial acertar casi a la perfección en la dirección de la volatilidad.

- **Combinaciones Convexas de Modelos Lineales y Memoria Profunda (ARIMA-LSTM/GRU):** Para abordar el problema de la excesiva parametrización, Cappello et al. (2025) desarrollaron un modelo híbrido basado en la combinación convexa de las predicciones de un modelo ARIMA y redes recurrentes profundas (LSTM y GRU). Al aplicar este enfoque sobre las acciones de Unicredit SpA y la tasa de cambio EUR/USD, el sistema logró compensar las debilidades de cada componente. La combinación híbrida con LSTM mejoró al modelo ARIMA puro en aproximadamente un 4.8% y al modelo LSTM puro en un 5.1% en términos de Error Absoluto Medio (MAE). Estos resultados sugieren que no es necesario descartar la estadística tradicional, sino integrarla de forma estructurada con el aprendizaje profundo para reducir los márgenes de error de manera efectiva.

- **Mecanismos de Atención en Entornos Híbridos (LSTM-mTrans-MLP):** Para mercados de alta volatilidad como el de las criptomonedas, Kabir et al. (2025) introdujeron una arquitectura que combina redes LSTM para capturar dependencias temporales a distintos plazos, un modelo

Transformer modificado que utiliza mecanismos de auto-atención para identificar qué partes de la serie son más relevantes globalmente, y un Perceptrón Multicapa (MLP) en la etapa de salida. Al evaluar el precio de cierre de Bitcoin, un activo conocido por su comportamiento errático, el modelo híbrido mostró una mejora considerable frente a sus componentes individuales: registró un Error Cuadrático Medio (RMSE) de 288.34 y un Error Absoluto Medio (MAE) de 186.23, frente al RMSE de 476.84 que obtenía la red LSTM operando de forma individual.

- **Descomposición de Señales y Agrupación de Patrones (PGFF / CEEMD-Crossformer):** Con el objetivo de reducir el impacto del ruido en los precios de metales preciosos, Kim et al. (2026) aplicaron un enfoque basado en la Descomposición Empírica de Modos por Conjuntos Completa (CEEMD) integrada con una arquitectura Crossformer (PGFF). A diferencia de procesar cada componente intrínseco (IMF) de forma individual, su metodología agrupa aquellos con patrones de autocorrelación similares antes de modelarlos, lo que permitió reducir el error MAE en un 60.33% para pronósticos a un paso en el mercado del paladio, superando de forma clara a los enfoques de descomposición estándar.

- **Ensamblajes de Regresión con Filtros de Estado (ELM-SVR + Kalman):** Desde la perspectiva del arbitraje estadístico, Nóbrega y Oliveira (2014) combinaron las predicciones de una Máquina de Aprendizaje Extremo (ELM) y un SVR mediante un Filtro de Kalman, que ajusta el modelo de forma recursiva a medida que ingresan nuevas observaciones. En una simulación de trading institucional, esta combinación no solo redujo el error estadístico en términos de RMSE, sino que también mostró resultados relevantes en términos prácticos: logró un retorno anualizado del 26.13% con un ratio de Sharpe de 5.29, métricas que aproximadamente duplicaban el desempeño de los modelos operando de forma individual.

La **Tabla 1** resume de forma comparativa los resultados reportados por los estudios analizados:

Tabla 1. Comparativa Empírica de Rendimiento: Modelos Individuales vs. Arquitecturas Híbridas

Autor(es) y Año	Mercado / Activo	Modelo Individual o Base	Arquitectura Híbrida Propuesta	Resultados y Mejoras Empíricas Clave
Zhu & Wei (2012)	Precios del Carbono (ECX)	ARIMA / LSSVM	**ARIMA-LSSVM (optimizado con PSO)**	Superioridad en predicción direccional (D_{stat} del 97.98%); reducción significativa de RMSE y sobreajuste frente a modelos individuales.
Cappello et al. (2025)	Acciones (Unicredit SpA)	ARIMA / LSTM puro	**ARIMA-LSTM (Combinación Convexa)**	Reducción del error MAE en ~4.8% vs. ARIMA y ~5.1% vs. LSTM puro; menor error cuadrático global.
Kabir et al. (2025)	Criptomonedas (Bitcoin)	LSTM puro	**LSTM-mTrans-MLP**	Reducción del RMSE de 476.84 (modelo base) a 288.34 (híbrido). Mejora en precisión direccional (MAPE de 0.0268).
Kim et al. (2026)	Metales Preciosos (Paladio)	CEEMD-LSTM	**PGFF (CEEMD-Crossformer agrupado)**	Disminución del MAE en 60.33% (a un paso temporal); captura óptima de tendencias de baja volatilidad.
Nóbrega & Oliveira (2014)	Arbitraje Estadístico de Pares	ELM / SVR	**ELM + SVR con Filtro de Kalman**	Retorno anualizado del 26.13% (Sharpe 5.29) en datos fuera de muestra; mejora frente a modelos individuales.

Mao et al. (2026)	Log-retornos (Sector Tecnológico)	VARMA / Multi-LSTM	**BERTFSIE-VARMA-LSTM (Multimodal)**	Integración de análisis de sentimiento (BERT); RMSE de 16.46 frente a 21.24 del VARMA tradicional.

El análisis de esta evidencia permite consolidar tres ventajas del paradigma híbrido:

1. Al dividir la tarea de aprendizaje entre la captura de tendencias lineales y el modelado de residuos no lineales, los modelos híbridos logran reducir los márgenes de error a niveles que los modelos paramétricos o las redes neuronales estándar difícilmente alcanzan operando de forma individual.

2. La incorporación de filtros recursivos como el Filtro de Kalman, o mecanismos de ponderación dinámica como los mecanismos de atención de los Transformers o las combinaciones convexas, permite que los modelos híbridos se adapten a nuevos periodos de volatilidad o situaciones de crisis sin necesidad de ser reentrenados desde cero.

3. El uso combinado de arquitecturas reduce el riesgo de que modelos de alta complejidad, como las redes profundas, terminen memorizando el ruido del mercado en lugar de aprender patrones. Al entregarles residuos previamente filtrados o descompuestos por un modelo precursor, se restringe el espacio de aprendizaje de la red neuronal a la información realmente relevante, lo que contribuye a construir sistemas más fiables y con mayor potencial de aplicación práctica.

5.5. ANÁLISIS CRÍTICO SOBRE LOS MODELOS HÍBRIDOS

5.5.1. LA BRECHA ENTRE EL DESEMPEÑO ESTADÍSTICO Y LA UTILIDAD ECONÓMICA

En la literatura existe un acuerdo bastante extendido sobre las ventajas matemáticas de los enfoques combinados. La arquitectura híbrida, al separar la serie de tiempo en un componente lineal modelado por técnicas estadísticas y un componente no lineal o residual capturado por redes neuronales o algoritmos de vectores de soporte, logra en general reducir el error de predicción frente a modelos individuales. Sin embargo, esto no siempre se traduce en una herramienta útil para generar rentabilidad en el mercado real. Dicho de otra forma, que un modelo tenga un error estadístico bajo no garantiza que sea capaz de tomar decisiones de trading que resulten en ganancias concretas, y esta brecha entre el desempeño estadístico y la utilidad económica es uno de los puntos más debatidos en la literatura reciente.

Uno de los problemas menos evidentes en la evaluación de modelos financieros es que minimizar el error estadístico puede ser engañoso. Métricas como el Error Cuadrático Medio (RMSE) o el Error Absoluto Medio (MAE) son los estándares habituales en ciencia de datos, y penalizan la magnitud promedio del error, pero no tienen en cuenta el momento exacto en que este ocurre. En la práctica, un modelo puede obtener un RMSE muy bajo simplemente replicando el precio del día anterior o siguiendo la tendencia con un pequeño retraso, porque los cambios de precio diarios suelen ser pequeños en términos absolutos. El problema es que un modelo así, aunque parezca preciso sobre el papel, puede estar reaccionando sistemáticamente un día tarde a las caídas del mercado o fallando en identificar el punto exacto donde una tendencia cambia de dirección. En el trading cuantitativo y el arbitraje institucional, anticipar correctamente la dirección del movimiento suele ser mucho más valioso que acertar con precisión el valor del precio.

Diversos estudios han documentado esta brecha entre desempeño estadístico y utilidad económica. Un ejemplo concreto y bien documentado es el de Gerlein et al. (2016), quienes evaluaron distintos algoritmos de clasificación en simulaciones de trading sobre el mercado Forex, encontrando resultados que ilustran claramente esta paradoja. El clasificador K-Star obtuvo una precisión predictiva general de apenas 49.30%, un valor estadísticamente inferior al azar, pero generó un retorno económico acumulado del 130.74%. Esto ocurrió porque el modelo logró identificar correctamente las operaciones direccionales de mayor magnitud. En contraste, el clasificador JRip dentro del mismo estudio alcanzó una precisión estadística superior del 52.32%, pero su rentabilidad fue considerablemente menor, con un retorno acumulado de apenas 12.46%. Este resultado muestra que una mayor precisión estadística no solo no garantiza una mejor rentabilidad, sino que en ciertos casos puede coincidir con un desempeño económico notablemente inferior si el modelo falla en identificar los momentos más relevantes del mercado.

Nóbrega y Oliveira (2014) documentaron esta misma desconexión en simulaciones de arbitraje estadístico de pares. En su estudio, la Regresión de Vectores de Soporte (SVR) mostró un mejor ajuste estadístico inicial, con el RMSE más bajo entre los modelos evaluados, frente a la Máquina de Aprendizaje Extremo (ELM). Operando de forma individual, ambos modelos generaron retornos anualizados similares: 21.32% para el SVR y 20.18% para el ELM. Sin embargo, al combinar dinámicamente sus señales predictivas mediante un ensamble con Filtro de Kalman, el sistema híbrido logró estabilizar el error estadístico, reducir el drawdown máximo y mejorar de forma notable la rentabilidad práctica. Este modelo alcanzó un retorno anualizado del 26.13% con un ratio de Sharpe de 5.29, lo que confirma que el mejor desempeño financiero ajustado al riesgo no provino del modelo con el RMSE más bajo, sino del sistema que supo ponderar y combinar las señales predictivas de forma adaptativa a lo largo del tiempo.

La lección central de este análisis es que la investigación en predicción de series de tiempo financieras necesita ampliar sus criterios de evaluación. Enfocarse únicamente en minimizar el error estadístico mediante arquitecturas cada vez más complejas no es suficiente; la evaluación de los modelos debe incorporar también métricas de precisión direccional y pruebas basadas en la rentabilidad económica real (Gerlein et al., 2016; Nóbrega y Oliveira, 2014). En los mercados financieros, un pronóstico que sea numéricamente cercano al valor real pero que se equivoque de forma consistente en la dirección del movimiento tiene poco valor práctico, porque es precisamente esa dirección la que determina si una estrategia de trading genera o pierde dinero.

5.5.2. EL DILEMA ENTRE PRECISIÓN PREDICTIVA E INTERPRETABILIDAD

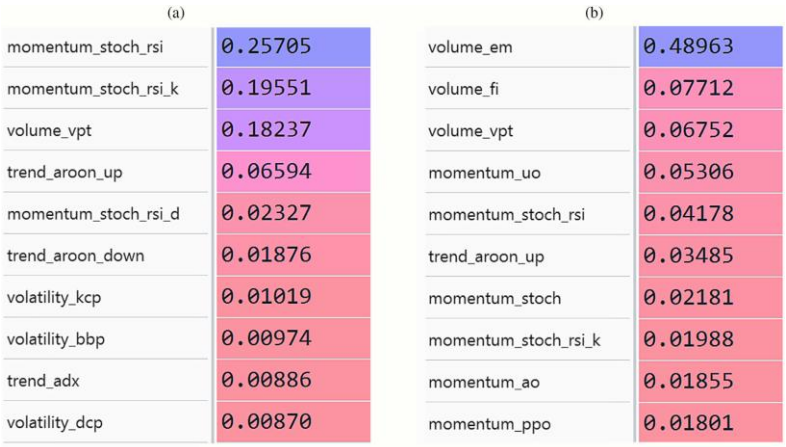
Uno de los aspectos menos discutidos en la literatura sobre modelos híbridos es el costo que tiene la complejidad en términos de interpretabilidad. Los modelos econométricos clásicos como la familia GARCH, aunque fueron diseñados específicamente para modelar la varianza condicional, tienen una limitación importante: imponen suposiciones rígidas sobre la forma de los datos y asumen distribuciones de probabilidad para los términos de error, lo que restringe su capacidad de adaptación cuando el mercado se comporta de forma atípica. Integrar algoritmos de aprendizaje automático, como la Regresión de Vectores de Soporte o redes neuronales, con estos modelos estadísticos permite capturar dinámicas no lineales como el agrupamiento de volatilidad o la asimetría del mercado, sin estar atados a supuestos tan estrictos sobre la distribución de los datos ni a una varianza constante (Aldeki, 2025).

Sin embargo, esta capacidad de adaptación en entornos volátiles tiene un costo que no siempre se discute con suficiente atención: la pérdida de explicabilidad. Un modelo híbrido complejo o un ensamble profundo puede acertar en la dirección futura del mercado, pero si un analista de riesgos no puede explicar ante una junta directiva o un ente regulador por qué el modelo tomó esa

decisión, su utilidad en el ámbito financiero queda comprometida. En la gestión de carteras, las normativas de cumplimiento y los marcos de control de riesgo exigen que las decisiones de inversión sean transparentes y auditables. Las redes neuronales profundas operan como "cajas negras": procesan cientos o miles de variables a través de capas ocultas y producen un resultado de alta precisión, pero no permiten rastrear la lógica que llevó al modelo a generar, por ejemplo, una señal de venta durante un episodio de pánico en el mercado (El Majzoub et al., 2023).

La investigación reciente ha comenzado a cuestionar si la precisión algorítmica debe ser el único criterio para evaluar y adoptar un modelo predictivo. En esta línea, El Majzoub et al. (2023) desarrollaron un estudio en el mercado de criptomonedas, un entorno de alta volatilidad, evaluando diversos modelos de Machine Learning con especial atención a su interpretabilidad. Sus resultados mostraron que las herramientas más complejas y automatizadas, como Google AutoML, lograron la mayor precisión general y los retornos económicos más altos, pero ofrecían prácticamente ninguna posibilidad de personalización y operaban como cajas negras sin capacidad de explicación.

Frente a este problema, la evidencia sugiere que ceder un pequeño margen de precisión a cambio de mayor transparencia es una decisión razonable y justificada en contextos institucionales. En el mismo estudio, El Majzoub et al. (2023) encontraron que algoritmos menos complejos como los Árboles de Decisión y los Bosques Aleatorios ofrecieron resultados competitivos en términos de rentabilidad, pero con una ventaja concreta: permitieron extraer análisis de importancia de variables, lo que da una idea clara de qué factores están influyendo en cada predicción. Al aplicar técnicas de poda sobre los árboles de decisión, los investigadores generaron diagramas lógicos visuales que reducían el sobreajuste y hacían que la lógica del modelo fuera legible para gestores sin formación técnica. En una línea similar, modelos de aprendizaje profundo como N-BEATS han mostrado que es posible mantener una precisión alta sin sacrificar completamente la interpretabilidad, al entregar descomposiciones explícitas de la tendencia y la estacionalidad de la serie temporal.



(a)

momentum_stoch_rsi	0.25705
momentum_stoch_rsi_k	0.19551
volume_vpt	0.18237
trend_aroon_up	0.06594
momentum_stoch_rsi_d	0.02327
trend_aroon_down	0.01876
volatility_kcp	0.01019
volatility_bbp	0.00974
trend_adx	0.00886
volatility_dcp	0.00870

(b)

volume_em	0.48963
volume_fi	0.07712
volume_vpt	0.06752
momentum_uo	0.05306
momentum_stoch_rsi	0.04178
trend_aroon_up	0.03485
momentum_stoch	0.02181
momentum_stoch_rsi_k	0.01988
momentum_ao	0.01855
momentum_ppo	0.01801

Figura 5. Análisis de importancia de las variables para algoritmos de clasificación (a) y de regresión (b) (El Majzoub et al., 2023).

Por lo tanto, la evidencia revisada sugiere que el modelo más valioso no es necesariamente el más complejo o el que minimiza el error con mayor precisión. Para que los sistemas híbridos sean herramientas realmente útiles en la toma de decisiones financieras, deben encontrar un equilibrio entre su capacidad para capturar dinámicas no lineales y la posibilidad de que los analistas puedan entender y auditar sus predicciones, de manera que el juicio humano siga teniendo un rol activo en la evaluación de los riesgos asumidos.

5.5.3. COSTOS OPERATIVOS Y RIESGOS METODOLÓGICOS

La implementación de arquitecturas profundas y modelos híbridos en el pronóstico financiero conlleva una carga que frecuentemente se subestima en los entornos académicos: el costo operativo y computacional. Como señalan Shabani et al. (2023), a diferencia de los modelos econométricos paramétricos, la optimización de redes neuronales profundas exige un alto número de iteraciones y depende del uso de hardware especializado, lo que implica un consumo considerable de tiempo incluso en condiciones favorables. Los mismos autores advierten que las características cambiantes del mercado financiero pueden volver obsoleto un sistema basado en datos con rapidez, lo que obliga

a reentrenarlo y actualizarlo de forma frecuente con información reciente. Este ciclo de mantenimiento continuo se traduce en un alto consumo de tiempo y energía que, en ciertos contextos, puede reducir o incluso anular la rentabilidad del sistema. Si el costo de mantener y actualizar el modelo algorítmico se acerca o supera la ventaja predictiva que este ofrece, su adopción deja de estar justificada tanto desde el punto de vista técnico como económico.

Más allá de los costos de infraestructura, el riesgo más relevante para la viabilidad de estos modelos está en las fallas metodológicas durante su validación. En este sentido, no todos los resultados reportados en la literatura académica son tan sólidos como parecen. Henrique et al. (2023) advierten sobre la necesidad de implementar mecanismos estrictos de validación cruzada para reducir los efectos del data snooping, un sesgo que ocurre cuando el proceso de selección y ajuste del modelo se repite tantas veces sobre los mismos datos que los resultados terminan siendo optimistas por construcción. Muchos de los niveles de precisión casi perfectos reportados en publicaciones recientes son, en realidad, producto de este sesgo o de fugas de información (data leakage), que ocurren cuando el algoritmo utiliza, directa o indirectamente, información futura que no estaría disponible en un escenario de predicción en tiempo real.

En esta misma línea, Zhang et al. (2023) señalan que este error metodológico representa una de las principales debilidades de los modelos de ensamble por descomposición, como EMD o CEEMD, indicando que estos son especialmente susceptibles a la exposición de características futuras y a la fuga de datos durante el preprocesamiento. El problema ocurre cuando los investigadores aplican la descomposición sobre la totalidad de la serie histórica antes de separar los datos en conjuntos de entrenamiento y prueba. Al proceder de esta forma, el modelo accede de manera indirecta a información futura durante su fase de preparación, lo que infla artificialmente las métricas de precisión y compromete su capacidad de generalización cuando se despliega en condiciones reales de mercado.

Como conclusión de esta sección, el análisis de la literatura muestra que la calidad de un modelo predictivo no es directamente proporcional a su complejidad o al número de capas y parámetros que lo componen. El mejor modelo no es necesariamente el más sofisticado ni el que reporta el menor error en un entorno controlado, sino aquel que logra adaptarse de forma incremental a la dinámica del mercado sin memorizar el ruido histórico. En definitiva, evaluar un modelo híbrido solo por su error estadístico no es suficiente. Lo que realmente importa es si funciona bien con datos nuevos, si su costo operativo es justificable y si sus decisiones pueden explicarse a quienes asumen el riesgo.

6. CONCLUSIONES

1. Los modelos estadísticos tradicionales como ARIMA y la familia GARCH establecieron las bases del análisis predictivo financiero y siguen siendo útiles para capturar estructuras lineales y agrupamientos de volatilidad. Sin embargo, sus supuestos rígidos, como la exigencia de estacionariedad, la linealidad y la varianza constante, los limitan considerablemente frente a la complejidad real de los mercados. Dado que los precios de los activos responden a dinámicas caóticas, asimétricas y cambiantes, estos modelos tienden a incrementar su error de predicción en escenarios de alta volatilidad o ante cambios abruptos, lo que hace que su uso en solitario sea insuficiente para el pronóstico financiero en la actualidad.

2. Los modelos de Machine Learning y Deep Learning, especialmente arquitecturas como LSTM y GRU, han demostrado una capacidad superior para capturar dependencias temporales complejas sin necesidad de supuestos estadísticos previos sobre los datos. No obstante, esta mayor precisión viene acompañada de limitaciones: operan como cajas negras cuya lógica interna es difícil de explicar, demandan grandes volúmenes de datos y requieren hardware especializado con un costo computacional considerable. En contextos institucionales donde las decisiones de inversión deben ser justificables, esta falta de interpretabilidad es un obstáculo real. Por ello, algoritmos más transparentes como los Bosques Aleatorios, que permiten identificar qué variables influyen en cada predicción, pueden ser preferibles, aunque sacrifiquen un pequeño margen de precisión.

3. La revisión bibliográfica confirma que la combinación de métodos es el enfoque más robusto para la predicción de series de tiempo financieras. Al reconocer que estas series contienen componentes tanto lineales como no lineales de forma simultánea, los modelos híbridos asignan a cada parte del problema el algoritmo más adecuado: la estadística clásica se encarga

de la tendencia lineal, y las redes neuronales se concentran en capturar los patrones no lineales residuales. Este diseño reduce el riesgo de sobreajuste y logra niveles de error que los modelos individuales difícilmente alcanzan, como lo demuestran estudios como los de Zhang (2003), Zhu y Wei (2012), Cappello et al. (2025) y Kabir et al. (2025).

4. Uno de los hallazgos más importantes de esta revisión es que reducir el error estadístico no equivale a obtener mejores retornos en el mercado real. Un modelo puede seguir la tendencia de los precios con gran precisión y aun así fallar en identificar los momentos clave donde el mercado cambia de dirección, que son precisamente los que generan o destruyen valor en una estrategia de trading. La evidencia de Gerlein et al. (2016) y Nóbrega y Oliveira (2014) muestra que predecir correctamente la dirección del movimiento, si el precio sube o baja, tiene un impacto mucho más directo en la rentabilidad que predecir el valor exacto del activo. Esto implica que los criterios de evaluación de estos modelos deben incluir métricas económicas reales y no limitarse al RMSE o el MAE.

5. Finalmente, una parte considerable de los resultados de alta precisión reportados en la literatura reciente está inflada por fallas metodológicas, principalmente por data snooping y fugas de información durante el preprocesamiento, donde el modelo accede de forma indirecta a datos futuros que no estarían disponibles en un escenario real. El mejor modelo no es el más complejo, sino el que mantiene su desempeño al ser evaluado con datos estrictamente fuera de muestra y bajo condiciones que simulen el mercado real. A esto se suma que el despliegue de arquitecturas profundas implica costos computacionales y tiempos de mantenimiento que pueden reducir los márgenes de rentabilidad. Por ello, la adopción de estos sistemas debe considerar también su viabilidad operativa y no solo su rendimiento en entornos controlados.

7. RECOMENDACIONES

1. Los futuros estudios deberían incorporar simulaciones que midan la precisión direccional y la rentabilidad real del modelo, y no limitarse a reportar RMSE o MAE. Idealmente, estas simulaciones deberían considerar también costos de transacción y otros factores del mercado real, ya que un modelo puede tener un error bajo en papel, pero ser poco útil en la práctica si no logra identificar los momentos clave del mercado.

2. Arquitecturas basadas en mecanismos de atención como los Transformers han mostrado ventajas sobre las LSTM para capturar dependencias a largo plazo, y su integración en esquemas híbridos es un área con potencial. Asimismo, combinar estos modelos con técnicas de análisis de sentimiento de noticias o redes sociales podría mejorar la respuesta del sistema ante eventos externos que los datos históricos de precios no anticipan.

3. Dado que la falta de interpretabilidad es una de las limitaciones más relevantes identificadas en esta revisión, se recomienda que futuros estudios no solo busquen maximizar la precisión, sino también explorar formas de hacer que la lógica del modelo sea más comprensible. Herramientas como los valores SHAP o el análisis de importancia de variables son un punto de partida útil para avanzar en esta dirección sin sacrificar demasiado rendimiento.

4. Los modelos que utilizan técnicas de descomposición de señales como EMD o CEEMD son especialmente vulnerables a la fuga de información si no se aplican correctamente. Se recomienda que las investigaciones futuras garanticen que la descomposición se realice dentro de cada ventana de entrenamiento de forma estricta, evitando que el modelo acceda de forma indirecta a datos futuros durante el preprocesamiento, lo que infla artificialmente los resultados reportados.

5. La mayoría de los estudios revisados trabajan con datos diarios y modelos estáticos que requieren reentrenamiento completo cuando el mercado cambia. Sería valioso explorar enfoques de aprendizaje en línea que actualicen los parámetros del modelo de forma incremental a medida que llegan nuevos datos, reduciendo así los costos de mantenimiento y la obsolescencia del sistema en entornos de alta frecuencia.

8. REFERENCIAS BIBLIOGRÁFICAS

Ampomah, E. K., Qin, Z., Nyame, G. y Botchey, F. E. (2020). Stock market decision support modeling with tree-based AdaBoost ensemble machine learning models. *Informatica*, 44(4). https://doi.org/10.31449/inf.v44i4.3159

Aldeki, R. G. (2025). Predicting Financial Market Volatility with Modern Model and Traditional Model. *Finance Theory And Practice*, 29(2), 154-165. https://doi.org/10.26794/2587-5671-2025-29-2-154-165

Cappello, C., Congedi, A., De Iaco, S. y Mariella, L. (2025). Traditional Prediction Techniques and Machine Learning Approaches for Financial Time Series Analysis. *Mathematics*, 13(3), 537. https://doi.org/10.3390/math13030537

Derbentsev, V., Velykoivanenko, H. y Datsenko, N. (2019). Machine learning approach for forecasting cryptocurrencies time series. *Neuro-Fuzzy Modeling Techniques In Economics*, 8, 65-93. https://doi.org/10.33111/nfmte.2019.065

Gerlein, E. A., McGinnity, M., Belatreche, A., y Coleman, S. (2016). Evaluating machine learning classification for financial trading: An empirical approach. *Expert Systems With Applications*, 54, 193-207. https://doi.org/10.1016/j.eswa.2016.01.018

Henrique, B. M., Sobreiro, V. A. y Kimura, H. (2023). Practical machine learning: Forecasting daily financial markets directions. *Expert Systems With Applications*, 233, 120840. https://doi.org/10.1016/j.eswa.2023.120840

Jaramillo, J., Velasquez, J. D. y Franco, C. J. (2017). Research in Financial Time Series Forecasting with SVM: Contributions from Literature. *IEEE Latin America Transactions*, 15(1), 145-153. https://doi.org/10.1109/tla.2017.7827918

Kabir, M. R., Bhadra, D., Ridoy, M. y Milanova, M. (2025). LSTM–Transformer-Based Robust Hybrid Deep Learning Model for Financial Time Series Forecasting. *Sci*, 7(1), 7. https://doi.org/10.3390/sci7010007

Majzoub, A. E., Rabhi, F. A. y Hussain, W. (2023). Evaluating interpretable machine learning predictions for cryptocurrencies. *Intelligent Systems In Accounting, Finance And Management/Intelligent Systems In Accounting, Finance & Management*, 30(3), 137-149. https://doi.org/10.1002/isaf.1538

Mao, S., Meng, L., Barbu, A. y Niu, X. (2026). Is the financial market efficient? An advanced time series analysis. *Financial Innovation*, 12(1). https://doi.org/10.1186/s40854-025-00856-y

Nóbrega, J. P. y Oliveira, A. L. I. (2014). A combination forecasting model using machine learning and Kalman filter for statistical arbitrage. *In 2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (pp. 1294–1299). IEEE. https://doi.org/10.1109/SMC.2014.6974093

Kim, D., Lee, J. y Kim, H. (2026). Pattern-guided forecasting framework for metal price prediction with grouping decomposed series. *Financial Innovation*, 12(1). https://doi.org/10.1186/s40854-025-00833-5

Reid, D., Hussain, A. J. y Tawfik, H. (2014). Financial time series prediction using spiking neural networks. *PLoS ONE*, 9(8), e103656. https://doi.org/10.1371/journal.pone.0103656

Selvin, S., Vinayakumar, R., Gopalakrishnan, E. A., Menon, V. K. y Soman, K. P. (2017). Stock price prediction using LSTM, RNN and CNN-sliding window model [IEEE]. En 2017 *International Conference on Data Mining and Big Data (CBDCom)* (pp. 1643-1647). https://doi.org/10.1109/icacci.2017.8126078

Shabani, M., Tran, D. T., Kanniainen, J., Iosifidis, A., Shabani, M., Tran, D. T., Kanniainen, J. y Iosifidis, A. (2023). Augmented bilinear network for incremental multi-stock time-series classification. *Pattern Recognition*, 141, 109604. https://doi.org/10.1016/j.patcog.2023.109604

Shen, G., Tan, Q., Zhang, H., Zeng, P. y Xu, J. (2018). Deep Learning with Gated Recurrent Unit Networks for Financial Sequence Predictions. *Procedia Computer Science*, 131, 895-903. https://doi.org/10.1016/j.procs.2018.04.298

Shen, F., Chao, J., Zhao, J., Shen, F., Chao, J. y Zhao, J. (2015). Forecasting exchange rate using deep belief networks and conjugate gradient method. *Neurocomputing*, 167, 243-253. https://doi.org/10.1016/j.neucom.2015.04.071

Fenghua, W., Jihong, X., Zhifang, H. y Xu, G. (2014). Stock Price Prediction based on SSA and SVM. *Procedia Computer Science*, 31, 625-631. https://doi.org/10.1016/j.procs.2014.05.309

Xiao, D. y Su, J. (2022). Research on Stock Price Time Series Prediction Based on Deep Learning and Autoregressive Integrated Moving Average. *Scientific Programming*, 2022,? 1-12. https://doi.org/10.1155/2022/4758698

Zaffar, A. y Hussain, S. M. A. (2022). Modeling and prediction of KSE – 100 index closing based on news sentiments: an applications of machine learning model and ARMA (p, q) model. *Multimedia Tools And Applications*, 81(23), 33311-33333. https://doi.org/10.1007/s11042-022-13052-2

Zhang, C., Sjarif, N. N. A. y Ibrahim, R. (2023). Deep learning models for price forecasting of financial time series: A review of recent advancements: 2020–2022. *WIREs Data Mining and Knowledge Discovery*, 14(1), e1519. https://doi.org/10.1002/widm.1519

Zhang, G. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, 50, 159-175. https://doi.org/10.1016/s0925-2312(01)00702-0

Zhu, B. y Wei, Y. (2012). 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