Predicción del Precio del Bitcoin mediante Deep Learning

Yenny Bibiana Cadavid Rivillas

Trabajo de grado presentado para optar al título de

Magíster en Métodos Cuantitativos para Economía y Finanzas

Director:
Ph.D Luis Miguel Jiménez Gómez

Universidad de Antioquia
Facultad de Ciencias Económicas
Métodos Cuantitativos para Economía y Finanzas
Medellín, Colombia
2025

Cita	(Cadavid Rivilla, 2026)
Referencia **Estilo APA 7 (2020)**	Cadavid Rivillas, Y. B. (2026). Predicción del precio del Bitcoin mediante Deep Learning. [Tesis de maestría]. Universidad de Antioquia, Medellín, Colombia.

Maestría en Métodos Cuantitativos para Economía y Finanzas

UNIVERSIDAD DE ANTIOQUIA
Vicerrectoría de Docencia

Sistema de Bibliotecas

Biblioteca Carlos Gaviria Díaz.

Repositorio Institucional: http://bibliotecadigital.udea.edu.co

Universidad de Antioquia - www.udea.edu.co

Dedicatoria

A quienes me acompañaron en este proceso, incluso en silencio.
Su apoyo, constante aunque discreto, ha dejado una huella que valoro profundamente.

Resumen

El presente trabajo aborda uno de los principales desafíos en el ámbito financiero y tecnológico: la predicción del precio del Bitcoin. El creciente protagonismo de las criptomonedas en los mercados globales, junto con su elevada volatilidad y la limitada capacidad predictiva de los métodos tradicionales, ha motivado el uso de herramientas avanzadas de análisis, particularmente aquellas basadas en inteligencia artificial.

Como punto de partida, se estimaron procedimientos SARIMA como línea base, y se modeló la serie temporal del Bitcoin utilizando inicialmente la serie original, en concordancia con la práctica predominante en la literatura. No obstante, de manera complementaria se implementó el procedimiento de Box-Jenkins para el preprocesamiento de la serie, con el objetivo de facilitar el aprendizaje de las redes neuronales al transformar patrones altamente complejos en estructuras más simples y estables.

Adicionalmente, para fortalecer el proceso de selección del mejor modelo, se generaron miles de series sintéticas a partir de los residuales, las cuales fueron evaluadas mediante inspección visual y la prueba de Kolmogorov-Smirnov de dos muestras. Este análisis incorporó explícitamente el comportamiento de los residuales, dado que constituyen la base para la simulación, y permitió orientar la evaluación hacia la capacidad predictiva de los modelos más que únicamente hacia su ajuste dentro de la muestra.

Los resultados evidencian que el modelo con mejor desempeño global corresponde a una arquitectura LSTM aplicada sobre la serie transformada. Si bien la LSTM ajustada sobre la serie sin transformar presentó un desempeño satisfactorio dentro de la muestra, su capacidad predictiva resultó limitada al analizar la generación de series sintéticas, lo que confirma la importancia del preprocesamiento para mejorar la estabilidad y robustez de los pronósticos.

Palabras clave: Bitcoin, Predicción, Redes neuronales, LSTM, Series sintéticas.

Abstract

This study addresses one of the main challenges in the financial and technological fields: the prediction of Bitcoin prices. The growing prominence of cryptocurrencies in global markets, together with their high volatility and the limited predictive capacity of traditional methods, has driven the adoption of advanced analytical tools, particularly those based on artificial intelligence.

As a starting point, SARIMA procedures were implemented as a baseline**,** and the Bitcoin time series was initially modeled using the original data, in accordance with the predominant practice in the literature. Additionally, the Box-Jenkins procedure was applied for time series preprocessing, with the aim of facilitating neural network learning by transforming highly complex patterns into simpler and more stable structures.

Furthermore, to strengthen the model selection process, thousands of synthetic series were generated from the residuals and evaluated through visual inspection and the two-sample Kolmogorov-Smirnov test. This analysis explicitly incorporated the behavior of the residuals, as they constitute the basis for the simulation process, and allowed the evaluation to focus on predictive performance rather than solely on in-sample fit.

The results indicate that the model with the best overall performance corresponds to an LSTM architecture applied to the transformed series. Although the LSTM fitted to the untransformed series showed satisfactory in-sample performance, its predictive capacity was limited when analyzing the generation of synthetic series, confirming the importance of preprocessing to improve the stability and robustness of forecasts.

Keywords: Bitcoin, Forecasting, Neural Networks, LSTM, Synthetic Series

Tabla de Contenido

Lista de figuras

Lista de tablas

Introducción

En los últimos años, el creciente interés por las criptomonedas, en particular por el Bitcoin, ha impulsado un notable aumento en la producción científica orientada a comprender su comportamiento en los mercados financieros. La elevada volatilidad de este activo, su creciente adopción global y la ausencia de un respaldo institucional centralizado generan un entorno de alta incertidumbre, lo que ha motivado el desarrollo de modelos capaces de mejorar la precisión de los pronósticos de precios (Lahmiri & Bekiros, 2019; Patel et al., 2020; Oyedele et al., 2023). Este escenario ha favorecido la incorporación de técnicas avanzadas de modelación, especialmente aquellas basadas en inteligencia artificial y aprendizaje profundo.

El Bitcoin, como activo digital descentralizado, presenta características que lo diferencian sustancialmente de los instrumentos financieros tradicionales. Su dinámica de precios no está gobernada exclusivamente por fundamentos macroeconómicos, sino por una combinación de factores especulativos, expectativas de mercado, adopción tecnológica y eventos externos, lo cual induce comportamientos altamente no lineales, cambios estructurales abruptos y episodios recurrentes de alta volatilidad (Yao et al., 2018; Tanwar et al., 2021). En este contexto, los modelos estadísticos clásicos, como los procedimientos SARIMA, si bien resultan útiles para capturar dependencias lineales y patrones estacionales, presentan limitaciones para representar adecuadamente estas dinámicas complejas (Kim & Won, 2018).

Como respuesta a estas limitaciones, el uso de arquitecturas de aprendizaje profundo ha ganado protagonismo en la predicción de series temporales financieras. Modelos como las redes neuronales recurrentes (RNN), las redes de memoria a largo corto plazo (LSTM) y las unidades recurrentes con compuertas (GRU) han demostrado una mayor capacidad para capturar dependencias temporales de corto y largo plazo, así como patrones no lineales presentes en datos altamente dinámicos (Wu et al., 2019; Fleischer et al., 2022; Nasirtafreshi, 2022). Diversos estudios evidencian que estas arquitecturas superan en muchos casos a los enfoques tradicionales en términos de desempeño predictivo, especialmente en mercados caracterizados por alta volatilidad y estructuras cambiantes (Livieris et al., 2020; Kang et al., 2022; Ferdiansyah et al., 2023).

No obstante, pese al crecimiento sostenido de la literatura, se observa una marcada heterogeneidad metodológica en cuanto a la selección de arquitecturas, esquemas de validación, métricas de desempeño y tratamiento estadístico de las series. En particular, una proporción significativa de los estudios evalúa el desempeño exclusivamente dentro del conjunto de prueba, sin extender la validación hacia horizontes temporales posteriores ni analizar rigurosamente las propiedades estadísticas de los errores y de las trayectorias simuladas (Pintelas et al., 2020; Livieris et al., 2020). Asimismo, aunque algunos trabajos incorporan análisis descriptivos de residuales, rara vez se examina formalmente su homocedasticidad, autocorrelación y consistencia distributiva, aspectos clave para garantizar la estabilidad del pronóstico.

En este contexto, surge la necesidad de sistematizar el conocimiento existente mediante una revisión estructurada de la literatura que permita identificar patrones comunes, brechas metodológicas y tendencias emergentes en el uso de técnicas predictivas aplicadas al Bitcoin. Este proceso facilita la comparación de enfoques bajo criterios homogéneos, contribuye a la identificación de buenas prácticas replicables y permite establecer fundamentos sólidos para el desarrollo de esquemas metodológicos más robustos y transparentes.

Por tanto, la presente investigación tiene como objetivo evaluar métodos de Deep Learning orientados a minimizar el error predictivo y mejorar la capacidad de pronóstico por fuera de la muestra en la serie de tiempo del Bitcoin. Para ello, se analizan y comparan arquitecturas del tipo RNA, RNN, LSTM y GRU, considerando tanto el desempeño predictivo tradicional como el comportamiento estadístico de los residuales y la capacidad de generación de series sintéticas mediante simulación Montecarlo. Este enfoque permite no solo medir la precisión de los modelos, sino también evaluar su idoneidad para reproducir trayectorias futuras plausibles bajo escenarios de alta incertidumbre, fortaleciendo así la validez de los resultados obtenidos.

El estudio analiza la dinámica del precio del Bitcoin utilizando información semanal correspondiente al periodo comprendido entre el 01 de noviembre de 2020 y el 03 de noviembre de 2025, para un total de 262 observaciones. Los datos se dividieron en una proporción de 70% para entrenamiento y 30% para prueba. El conjunto de entrenamiento abarca el periodo comprendido entre el 01 de noviembre de 2020 y el 28 de abril de 2024, correspondiente a 183 observaciones. Por su parte, el conjunto de prueba está compuesto por 79 observaciones, que cubren el periodo entre el 05 de mayo de 2024 y el 02 de noviembre de 2025, siendo esta última la fecha final del conjunto de evaluación. Adicionalmente, se planteó un escenario fuera de la muestra de 52 semanas, comprendido entre el 03 de noviembre de 2025 y el 02 de noviembre de 2026. Para este horizonte temporal se generaron series sintéticas de forma probabilística, las cuales no se interpretan como pronósticos puntuales, sino como escenarios futuros posibles derivados de la dinámica estocástica capturada por los modelos.

El documento se organiza de la siguiente manera. En el Capítulo 1 se presenta el planteamiento del problema, acompañado de la revisión de antecedentes y su correspondiente discusión crítica. El Capítulo 2 expone la justificación de la investigación, mientras que el Capítulo 3 establece los objetivos que orientan el desarrollo del estudio. En el Capítulo 4 se formula la hipótesis de trabajo que guía el análisis empírico. Posteriormente, el Capítulo 5 desarrolla el marco teórico, en el cual se presentan los fundamentos conceptuales y metodológicos que sustentan la investigación. El Capítulo 6 describe la metodología empleada, incluyendo los enfoques de modelación basados en series temporales y técnicas de aprendizaje profundo, así como los procedimientos utilizados para la optimización de los modelos, su evaluación y la generación de escenarios mediante simulación. El Capítulo 7 presenta los resultados obtenidos a partir de la aplicación de los modelos propuestos, mientras que el Capítulo 8 desarrolla la discusión de dichos resultados a la luz de los objetivos de investigación y del marco conceptual adoptado. Finalmente, los Capítulos 9 y 10 presentan las conclusiones del estudio, las principales limitaciones identificadas durante el proceso de investigación y las recomendaciones orientadas a futuras líneas de investigación.

1. Planteamiento del problema

A partir del análisis sistemático de la literatura reciente, se identificaron varios vacíos metodológicos y limitaciones recurrentes que evidencian la necesidad de profundizar en el desarrollo y evaluación de modelos predictivos aplicados a series temporales de criptomonedas.

En primer lugar, una proporción significativa de los estudios revisados no reporta de manera explícita la aplicación de procedimientos de preprocesamiento o transformaciones estadísticas sobre las series de tiempo antes de su modelación. En la mayoría de los casos, se infiere que las series son utilizadas en su forma original, sin evaluar aspectos fundamentales como estacionariedad, heterocedasticidad, normalización, estabilización de varianza o reducción de

ruido. Esta omisión puede afectar la estabilidad del entrenamiento de los modelos, la calidad de las predicciones y la interpretación de los resultados.

En segundo lugar, las métricas de desempeño reportadas en la literatura se concentran principalmente en la evaluación del ajuste dentro de la muestra, utilizando indicadores visuales y métricas de error que sugieren un buen desempeño aparente de los modelos. Sin embargo, son escasos los estudios que evalúan rigurosamente la capacidad real de pronóstico, lo cual limita la validez externa de los resultados y su aplicabilidad en contextos reales de toma de decisiones.

Adicionalmente, aunque algunos trabajos afirman realizar evaluaciones por fuera de la muestra, en la práctica estas evaluaciones suelen corresponder únicamente a particiones del conjunto de datos original destinadas al conjunto de prueba. No se identificaron estudios que implementen esquemas de pronóstico verdaderamente fuera de la muestra, utilizando períodos futuros no contenidos en la base de datos empleada para el entrenamiento y validación del modelo, lo cual restringe la evaluación de la capacidad de generalización frente a escenarios reales y dinámicos del mercado.

Asimismo, la mayoría de las investigaciones se enfoca prioritariamente en la etapa de modelación y comparación de arquitecturas, relegando el análisis detallado de los pronósticos generados y la evaluación sistemática de su desempeño predictivo en horizontes futuros. Esta orientación hacia el ajuste del modelo, más que hacia la calidad del pronóstico, limita la utilidad práctica de los resultados obtenidos.

Finalmente, aunque la literatura reporta una amplia diversidad de metodologías basadas en Machine Learning y Deep Learning, no se evidencia un consenso ni un protocolo metodológico estandarizado que permita maximizar el aprovechamiento de estas técnicas en el proceso de modelado de series temporales financieras. La ausencia de lineamientos claros sobre selección de arquitecturas, preprocesamiento, validación y evaluación dificulta la replicabilidad y la comparación objetiva entre estudios.

En conjunto, estos vacíos evidencian la necesidad de desarrollar un enfoque metodológico integral que incorpore un adecuado tratamiento de la serie de tiempo, una evaluación rigurosa del pronóstico fuera de la muestra y una utilización sistemática de técnicas de aprendizaje profundo, contribuyendo así al fortalecimiento de la confiabilidad y aplicabilidad de los modelos predictivos en el mercado de las criptomonedas.

En este trabajo se optó por pronosticar directamente los precios del Bitcoin, siguiendo enfoques similares en la literatura, y se aplicó un preprocesamiento que garantizara la estacionariedad de los residuales según Box-Jenkins. Esto permitió generar series sintéticas fuera de la muestra con propiedades estadísticas consistentes con la serie original, asegurando la validez de los pronósticos y manteniendo coherencia con el objeto de estudio.

1.1 Antecedentes

La presente investigación se sustenta en una revisión sistemática de literatura académica, orientada a identificar los principales enfoques metodológicos, tendencias de investigación y aportes recientes en la predicción de precios de criptomonedas, con énfasis en el Bitcoin. Para este propósito, se diseñó una ecuación de búsqueda basada en palabras clave representativas

del objeto de estudio, la cual permitió filtrar publicaciones relevantes en bases de datos especializadas, particularmente en Scopus. La ecuación empleada fue la siguiente:

(Cryptocurrency AND prediction OR forecasting AND deep learning OR neural networks) AND (Hybrid models AND financial time SERIES OR volatility) AND (Out of sample forecasting)

Esta formulación facilitó la identificación de trabajos relacionados con técnicas de aprendizaje profundo, análisis de series temporales financieras y procedimientos de validación predictiva por fuera de la muestra, con un enfoque integral y actualizado del estado del arte.

La consulta inicial en Scopus arrojó un total de 159 documentos**,** sin restricción por área temática, tipo de documento o año de publicación, lo que permitió obtener una visión amplia del panorama investigativo asociado al tema. Con el fin de analizar la estructura y distribución de esta producción científica, los resultados fueron organizados según diferentes criterios bibliométricos.

En este contexto, la Tabla 1 presenta la distribución de los documentos por área temática, permitiendo identificar las disciplinas que concentran mayor producción científica en el estudio de la predicción de criptomonedas y evidenciar el carácter multidisciplinario del campo, donde convergen áreas como las ciencias computacionales, la ingeniería, las finanzas y las matemáticas aplicadas.

Tabla 1. *Distribución por área temática*

Área Temática	Porcentaje de publicaciones
Ciencias de la computación	27.6 %
Economía, Econometría y Finanzas	19.4 %
Ingeniería	12.4 %
Negocios, Administración y Contabilidad	11.5 %
Matemáticas	8.5 %
Ciencias Sociales	5.8 %
Ciencias para la Toma de Decisiones	4.5 %
Ciencia de los Materiales	3.0 %
Física y Astronomía	1.8%
Energía	1 %
Otros	4.5 %

Fuente. Elaboración propia a partir de registros obtenidos en Scopus.

Como se observa en la Tabla 1, se evidencia una clara preponderancia de los campos de la economía, las ciencias de la computación, la ingeniería y la administración, lo cual sugiere que la investigación en predicción de precios de criptomonedas presenta una marcada orientación interdisciplinaria. Esta convergencia refleja la integración entre el análisis financiero, los métodos computacionales y los enfoques de gestión, resaltando la necesidad creciente de abordar este fenómeno desde una perspectiva integral que combine fundamentos económicos con tecnologías avanzadas de modelado y procesamiento de datos.

Por otra parte, aunque el área de ciencia de la toma de decisiones, que constituye el eje principal de este estudio representa una participación relativamente reducida (4,5 %), su relevancia es significativa, dado que estos trabajos suelen abordar problemáticas con un enfoque estratégico, orientado al soporte de decisiones en contextos de alta incertidumbre. Asimismo, se observa una participación moderada de otras áreas afines, como las ciencias exactas, lo que evidencia el carácter multidisciplinario del campo de estudio.

Con el fin de analizar la evolución temporal de la producción científica en esta línea de investigación, la Tabla 2 presenta la distribución de los documentos publicados por año. Este análisis permite identificar tendencias de crecimiento, periodos de mayor actividad investigativa y posibles incrementos asociados al auge del mercado de las criptomonedas y al avance de las técnicas de aprendizaje profundo aplicadas a series temporales financieras.

Tabla 2. *Documentos publicados por año*

Año	Publicaciones
2020	3
2021	7
2022	12
2023	22
2024	37
2025	31

Fuente. Elaboración propia a partir de registros obtenidos en Scopus.

Como se observa en la Tabla 2**,** a partir del año 2023 se evidencia una aceleración significativa en el volumen de publicaciones, lo que indica una intensificación del interés académico por la predicción de precios de criptomonedas mediante técnicas de aprendizaje profundo y modelos avanzados. Esta tendencia se consolida en 2024, alcanzando un máximo de 37 estudios publicados, lo cual refleja la creciente relevancia y madurez del campo de investigación.

Adicionalmente, en lo corrido del año 2025, el número de publicaciones registradas prácticamente iguala el total del año anterior, lo que indica que el dinamismo investigativo no solo se mantiene, sino que continúa fortaleciéndose. Este comportamiento confirma que el análisis predictivo de criptomonedas sigue siendo un tema de alto impacto científico, impulsado tanto por el avance de las capacidades computacionales como por la persistente volatilidad y complejidad de los mercados digitales.

Con el propósito de identificar los investigadores que han contribuido de manera más significativa al desarrollo de este campo, la Tabla 3 presenta los principales autores según el número de publicaciones registradas en la muestra analizada. Este análisis permite reconocer líderes académicos, líneas consolidadas de investigación y posibles referentes teóricos y metodológicos que han orientado el avance de los modelos predictivos aplicados a criptomonedas.

Tabla 3. *Principales autores*

Autores	Publicaciones
Xu, X.	12
Jin, B.	8
Zhang, Y.	5
Behera, S.	2
Będowska-Sójka, B.	2
Hao, J.	2
Ko, H.	2
Lee, J.	2
Li, J.	2

Fuente. Elaboración propia a partir de registros obtenidos en Scopus.

Como se observa en la Tabla 3, Xu, X., Jin, B. y Zhang, Y. concentran una parte significativa de la producción científica, lo que evidencia continuidad investigativa y una consolidación progresiva del campo en el desarrollo de modelos predictivos para criptomonedas.

Desde una perspectiva territorial, la Tabla 4 muestra los países o territorios con mayor número de publicaciones en el área, lo que permite identificar los principales polos de generación de conocimiento y la distribución geográfica.

Tabla 4. *Países o territorios con más publicaciones*

País	Publicaciones
China	32
Estados Unidos	21
Francia	10
India	9
Reino Unido	8
Australia	6
Irán	6
Turquía	6
Alemania	5
Polonia	5

Fuente. Elaboración propia a partir de registros obtenidos en Scopus.

La tabla 4 muestra que China lidera el campo temático, mientras que Estados Unidos se perfila como un competidor académico. Regiones como Europa y Asia aportan marcos alternativos o aplicaciones regionales. La investigación se extiende a regiones como Europa, Asia y otras más lejanas agregando valor.

En continuidad con el análisis bibliométrico, la Tabla 5 presenta la distribución de los documentos según su tipología, tales como artículos, ponencias, revisiones y otros formatos de publicación. Este indicador permite evaluar el grado de madurez de la producción académica, así como los canales predominantes de difusión del conocimiento en el área de predicción de criptomonedas y aprendizaje profundo.

Tabla 5. *Distribución por tipo de documento*

Tipo de documento	Porcentaje de publicaciones
Artículos	71.7 %
Libros	10.1 %
Artículos de revisión	8.2 %
Capítulos de libros	1.9 %
Ponencias	8.1 %

Fuente. Elaboración propia a partir de registros obtenidos en Scopus.

El conjunto de publicaciones mostrados en la Tabla 5, los artículos en revistas científicas constituyen el formato predominante, al representar más del 71.7 % del total, lo que refleja que la comunidad académica privilegia este medio como principal canal para la difusión de hallazgos científicos validados por procesos de arbitraje. En segundo lugar, los libros presentan una participación significativa del 10.1 %**,** generalmente asociados a desarrollos conceptuales, compilaciones metodológicas o estudios de mayor alcance. Posteriormente, los artículos de revisión, con una proporción cercana al 8.2 %**,** cumplen un papel fundamental en la sistematización del estado del arte y en la orientación de futuras líneas de investigación.

A partir de los 159 documentos identificados en la búsqueda inicial, que incluyen artículos, libros y otros tipos de publicaciones, se procedió a filtrar exclusivamente los artículos científicos, obteniéndose un subconjunto de 112 trabajos para evaluación preliminar. No obstante, únicamente 30 artículos estuvieron disponibles en texto completo para su análisis detallado, dado que una parte significativa no contaba con acceso directo o presentaba limitaciones de disponibilidad, pese a su aparente relevancia temática. Posteriormente, estos 30 artículos fueron sometidos a un proceso de depuración metodológica y temática**,** con el fin de seleccionar únicamente aquellos estudios que guardaran una relación directa con los objetivos de la presente investigación. Si bien una proporción importante de los trabajos empleaba enfoques afines como redes neuronales, modelos de aprendizaje automático y esquemas de validación predictiva, muchos de ellos se orientaban a contextos distintos, tales como la modelación de dinámicas caóticas en índices bursátiles tradicionales, la predicción de precios en mercados de materias primas o la estimación de capitalizaciones bursátiles en mercados accionarios específicos. Dado que estos enfoques, aunque metodológicamente valiosos, no abordaban de manera directa la predicción de precios de criptomonedas, se optó por conservar únicamente diez artículos que presentan un alineamiento claro tanto en términos metodológicos como en el objeto de estudio.

La Tabla 6 sintetiza las principales características de estas investigaciones, como autores, año de publicación, tipo de datos utilizados, metodología empleada, horizonte de predicción y principales hallazgos, lo cual facilita la identificación de patrones metodológicos, tendencias dominantes y vacíos de investigación.

Tabla 6. *Investigaciones recientes que emplean pronósticos por fuera de la muestra en criptomonedas*

Autores (año)	**Criptomoneda modelada**	**Frecuencia y Ventana de tiempo**	**Variable de entrada**	**Objetivo**	**Modelo de Deep Learning**	**Conclusiones**	**Evaluación de los residuales**	**Pronóstico por fuera de la muestra**	**Métrica utilizada por fuera de la muestra**
Zhou et al. (2025)	Bitcoin, Ethereum, Litecoin y Ripple.	Diaria. 5.5 años	Volatilidad Real de los futuros de: S&P 500, de los bonos del Tesoro de EE. UU a 30 años, del índice del dólar de ICE y los futuros del oro de COMEX para representar la volatilidad en los mercados de acciones, bonos, divisas y metales preciosos. Además, los futuros de gas natural y los del petróleo crudo ligero para representar el mercado energético. Índice de	Volatilidad Realizada de los futuros de las criptomonedas.	Red neuronal gráfica multiescala evolutiva (EMGNN) 9 en total (AR, ARIMAX, HAR, GPR, RF, LightGBM, LSTM, LSTNet, MTGNN.	La EMGNN ofrece un rendimiento sobresaliente y robusto. Las criptomonedas no son un activo único totalmente aislado del sistema financiero.	No	Si. En el conjunto de test.	MAFE, MSFE

			riesgo geopolítico e índice de incertidumbre económica.						
Cakici et al. (2024)	574 criptomonedas	Diaria. 5 años	Datos de varias bases de datos. 34 características representadas en liquidez, volatilidad, retornos pasados, distribución, entre otros.	Rendimientos transversales de las criptomonedas.	10 modelos de aprendizaje automático (OLS, PLS, LASSO, ENET, SVM, GBRT, RF, NN1 y NN2, COMB.	Los árboles y las redes neuronales resultan especialmente eficaces.	No	No	R^2
A. Lopéz (2023)	Bitcoin	Diaria. 7 años	Precios de cierre	Predicción del precio de cierre del Bitcoin.	Tres tipos de redes neuronales profundas y un modelo comparativo (FFNN, CNN, RNN)	Los patrones temporales resultan ser m´as representativos que los espaciales al estimar el comportamiento del precio de cierre futuro de Bitcoin.	No	No	MAE, RMSE, MedAE, MAPE, SMAPE, EA.
Magner & Hardy (2022)	13 criptomonedas	Diaria. 5 años	Precios de cierre de las 13 principales criptomonedas	Analizar la predictibilidad en el mercado	Comparación del paseo aleatorio (RW) y el paseo	Los modelos aplicados demostraron ser más eficientes	No	Si. En el conjunto de test.	Prueba t abarcativa (ENC-t)

			(BNB, USDC, USDT, XRP, SOL, MATIC, LTC, WVTC, VOGE, VUSD, ADAS, LUNA, USDT, BITCOIN, ETHEREUM)	de las criptomonedas.	aleatorio sin deriva (DRW) con dos modelos econométricos.	frente al rendimiento del paseo aleatorio.			
Kim et al.(2025)	Bitcoin	Diaria.7.5 años	Precios de cierre	Mejorar la predicción del precio del Bitcoin.	CNN-LSTM combinados con indicadores de mercado (índices bursátiles, tipos de cambio, materias primas, tipos de interés, junto con indicadores derivados como la volatilidad y la MDD)	Los resultados obtenidos sugieren que la estrategia implementada ofrece potencial para una inversión estable en Bitcoin.	No	No	Exactitud y precisión.
Akgun & Gulay (2025)	Bitcoin, Ethereum y Binance Coin.	Dos conjuntos de datos para cada criptomoneda: uno con datos diarios y otro con datos de	Precios de cierre	Modelado y previsión de la volatilidad de los rendimientos de las tres principales criptomonedas.	11 modelos GARCH y 3 modelos de Deep Learning (ANN, LSTM, CNN)	Los resultados muestran que los modelos de aprendizaje profundo superan a los modelos de tipo GARCH en los	No	Si. En el conjunto de test.	MSE, HMSE, MAPE, MASE, QLIKE.

		cinco minutos. 4.4 años				pronósticos de volatilidad.			
Nadarajah et al. (2025)	Bitcoin, Ethereum y Litecoin.	Diaria. 7 años	Precios de cierre	Estudio comparativo empírico de modelos de volatilidad.	Modelos GARCH y ANFIS	El aprendizaje conjunto basado en árboles proporciona una mejor precisión de pronóstico. Mientras que el rendimiento de algunos modelos de volatilidad de tipo GARCH es relativamente cercano al del mejor modelo tanto en las muestras de entrenamiento como en las de prueba.	Si. Box–Ljung test ARCH-LM test	Si. Dentro y fuera del conjunto de test.	RMSE, RSE
AlMadany et al (2024)	10 criptomonedas	Diaria. 5 años	Precios de cierre	Pronosticar la rentabilidad.	Modelos ARMA, GARCH, EGARCH, LSTM	Los modelos exhiben alta precisión. Los modelos híbridos EGARCH-LSTM o GARCH-LSTM demuestran una precisión ligeramente mejor en comparación	No	Si. En el conjunto de test.	MAE, RMSE, MSE

						con los otros modelos.			
Berger & Koubova (2024)	Bitcoin	Diaria. 8.7 años	Precios de cierre	Pronóstico del rendimiento	Modelos ARMA y GARCH	Las arquitecturas de aprendizaje profundo como las capas complejas, como LSTM, no aumentan la precisión de los pronósticos diarios. En concreto, una red neuronal recurrente simple constituye una opción para pronosticar series de retorno diario.	Si	No	RMSE, MAE
Erfanian et al (2022)	Bitcoin	Mensual. 8 años	Precios de cierre	Predicción del precio	Enfoques comparativos, que incluyen MCO, RVS y MLP	SVR es superior a otros modelos de aprendizaje automático y tradicionales.	No	No	R^2, Pearson's r, RMSE

Fuente. *Elaboración propia*

Como se sintetiza en la Tabla 6**,** los estudios seleccionados evidencian una tendencia creciente hacia el uso de modelos basados en aprendizaje profundo y enfoques híbridos para la predicción de precios en criptomonedas, con especial énfasis en arquitecturas recurrentes como LSTM y GRU, así como en la incorporación de esquemas de validación por fuera de la muestra para evaluar la capacidad de generalización de los modelos. De manera consistente, las investigaciones reportan mejoras significativas en precisión frente a métodos estadísticos tradicionales, aunque también destacan limitaciones asociadas a la volatilidad extrema del mercado, la sensibilidad a los hiperparámetros y la disponibilidad de datos. Este panorama confirma la relevancia y actualidad del problema abordado, al tiempo que pone en evidencia la necesidad de continuar explorando metodologías robustas que integren simulación, validación rigurosa y análisis de incertidumbre.

1.2 Discusión de antecedentes

A partir de la depuración de artículos seleccionados y sistematizados en la Tabla 6, se realiza un análisis crítico de las principales tendencias metodológicas, enfoques de modelación, variables empleadas y estrategias de validación aplicadas en la predicción de las criptomonedas. Esta discusión permite no solo sintetizar los hallazgos más relevantes, sino también identificar convergencias, divergencias, fortalezas y vacíos existentes en la literatura reciente.

El mercado de las criptomonedas ha experimentado un crecimiento exponencial durante la última década, consolidándose como un activo financiero de creciente interés tanto para inversionistas como para la comunidad académica. Esta evolución ha impulsado el desarrollo de herramientas predictivas cada vez más sofisticadas, orientadas a estimar variables clave como el precio, el rendimiento y la volatilidad de estos activos digitales. En este contexto, la literatura especializada ha explorado un amplio espectro de enfoques metodológicos, que abarcan desde modelos econométricos tradicionales hasta técnicas avanzadas de aprendizaje profundo (Deep Learning). Los estudios analizados, publicados entre 2022 y 2025, abordan la modelación y predicción del comportamiento de las criptomonedas utilizando diferentes metodologías, conjuntos de variables y horizontes temporales. El análisis comparativo de estos trabajos permite comprender el estado actual del campo, así como los principales desafíos que aún persisten.

Una primera coincidencia entre los trabajos revisados es el uso predominante de datos con frecuencia diaria, con ventanas temporales que, en general, abarcan entre cuatro y ocho años. Esta elección responde a la elevada volatilidad del mercado de criptomonedas, que exige modelos capaces de capturar dinámicas de corto plazo y adaptarse a cambios abruptos en los precios. Estudios como los de Zhou et al. (2025), Kim et al. (2025), Akgun y Gulay (2024) y AlMadany et al. (2025) emplean series temporales diarias para analizar el comportamiento del Bitcoin, Ethereum, Litecoin y otras criptomonedas de alta capitalización.

En cuanto a las variables objetivo, se observa una dualidad recurrente en la literatura: la predicción del precio de cierre y la estimación de la volatilidad, ya sea realizada o implícita. Este patrón se evidencia en investigaciones como las de López (2023) y Kim et al. (2025), centradas en el precio del Bitcoin, así como en los trabajos de Zhou et al. (2025) y Akgun y Gulay (2024), enfocados en la modelación de la volatilidad.

Desde el punto de vista metodológico, destaca la integración de técnicas de Machine Learning y Deep Learning con modelos econométricos tradicionales. Se aprecia una tendencia hacia la hibridación de enfoques, combinando la interpretabilidad de modelos clásicos como ARMA, GARCH o EGARCH, con la capacidad de las redes neuronales profundas como LSTM y CNN para capturar relaciones no lineales y patrones complejos. Esta convergencia metodológica refleja un reconocimiento generalizado de las limitaciones de los enfoques unidimensionales en el análisis de activos digitales.

No obstante, los estudios analizados presentan diferencias significativas en cuanto al alcance, las variables explicativas, el número y tipo de criptomonedas consideradas, así como las técnicas empleadas. En términos de cobertura, algunos trabajos se concentran exclusivamente en el Bitcoin, como Berger y Koubova (2023), Erfanian et al. (2022) y Kim et al. (2025), mientras que otros amplían el análisis a múltiples criptomonedas, llegando hasta 574 activos digitales en el estudio de Cakici et al. (2022), lo que permite una visión más integral del mercado.

Respecto a las variables explicativas, la literatura muestra una clara segmentación. Por un lado, algunos estudios utilizan únicamente precios históricos, lo cual puede resultar limitado en un entorno influenciado por factores externos. Por otro lado, investigaciones como la de Zhou et al. (2025) incorporan un amplio conjunto de variables exógenas, incluyendo indicadores de mercados tradicionales (acciones, bonos, metales preciosos y energía), así como índices de riesgo geopolítico e incertidumbre económica. Este enfoque multivariado permite capturar la influencia de factores macroeconómicos y mejorar potencialmente la capacidad predictiva de los modelos.

Las metodologías empleadas también presentan una alta diversidad. Algunos estudios privilegian modelos econométricos consolidados, como ARMA-GARCH (Berger y Koubova, 2023; AlMadany et al., 2025), mientras que otros adoptan arquitecturas avanzadas de aprendizaje profundo, como CNN-LSTM (Kim et al., 2025) o redes neuronales gráficas multiescala (Zhou et al., 2025). Esta heterogeneidad evidencia que el campo aún se encuentra en una etapa exploratoria, sin un consenso definitivo sobre el enfoque más eficiente para la predicción de criptomonedas.

Entre las principales fortalezas identificadas se encuentran la innovación metodológica y la riqueza en la construcción de variables. El trabajo de Zhou et al. (2025), por ejemplo, destaca por el uso de una red neuronal gráfica multiescala evolutiva (EMGNN), integrando variables macroeconómicas de distintos sectores para modelar las interacciones complejas entre mercados financieros tradicionales y criptomonedas. Asimismo, Cakici et al. (2022) sobresalen por la amplitud de su muestra y la incorporación de 34 características relacionadas con liquidez, retornos pasados, volatilidad y distribución, lo que fortalece la robustez de sus conclusiones. Otros estudios, como Nadarajah et al. (2025), aportan rigor estadístico mediante pruebas de diagnóstico sobre los residuales (Box-Ljung y ARCH-LM), asegurando la validez de los supuestos y la confiabilidad de los resultados.

A pesar de estos avances, persisten debilidades relevantes, particularmente en la validación externa y la capacidad de generalización de los modelos. Un número considerable de investigaciones no incorpora pronósticos verdaderamente fuera de la muestra, lo que limita la evaluación del desempeño ante datos no observados, como ocurre en Cakici et al. (2022) y López (2023). La mayoría de los estudios realiza

validaciones sobre conjuntos de prueba derivados de una partición de la base de datos original, como en Zhou et al. (2025), Magner y Hardy (2022), Akgun y Gulay (2024) y AlMadany et al. (2025). Por su parte, Nadarajah et al. (2025) amplían el alcance al considerar evaluaciones tanto dentro como fuera del conjunto de prueba, aproximándose a un ejercicio de validación más exigente.

Adicionalmente, muchos trabajos se concentran en una sola criptomoneda, generalmente el Bitcoin, y emplean exclusivamente precios históricos, sin incorporar variables externas como indicadores sociales, noticias financieras o eventos regulatorios. También se evidencia una falta de estandarización en las métricas de evaluación, lo que dificulta la comparación directa entre resultados. Mientras algunos autores utilizan métricas de error (RMSE, MAE, MAPE), otros recurren a medidas de correlación, precisión o pruebas estadísticas, impidiendo establecer conclusiones homogéneas sobre el desempeño relativo de los modelos.

A partir de esta revisión se identifican vacíos relevantes en la literatura reciente. En primer lugar, se observa una limitada validación verdaderamente fuera de la muestra, lo que restringe la evaluación de la capacidad de generalización en escenarios reales y dinámicos. En segundo lugar, la ausencia de protocolos estandarizados para la evaluación del desempeño dificulta la replicabilidad y comparación entre estudios. Abordar estas limitaciones requiere el desarrollo de marcos metodológicos homogéneos, el uso de ventanas temporales móviles y la adopción de métricas comunes que fortalezcan la robustez empírica de futuras investigaciones.

2. Justificación

El estudio de la dinámica y predicción de precios en los mercados de criptomonedas se ha consolidado como una línea de investigación relevante dentro de las finanzas cuantitativas y la ciencia de datos. La creciente adopción de estos activos digitales, junto con su elevada volatilidad y sensibilidad a múltiples factores económicos y tecnológicos, ha motivado el desarrollo de diversos enfoques analíticos orientados a comprender y anticipar su comportamiento. En particular, el Bitcoin se ha convertido en el principal referente dentro de este ecosistema, concentrando gran parte del interés académico y financiero debido a su capitalización de mercado, disponibilidad de información histórica y papel como activo representativo del mercado de criptomonedas.

A pesar del avance metodológico observado en la literatura reciente, el análisis de antecedentes evidencia que aún persisten diversas limitaciones en el desarrollo y evaluación de modelos predictivos aplicados a series temporales de criptomonedas. Entre estas se destacan la ausencia de procedimientos explícitos de preprocesamiento de las series de tiempo, la concentración de evaluaciones en el ajuste dentro de la muestra y la limitada implementación de esquemas de validación verdaderamente fuera de la muestra. Estas limitaciones reducen la posibilidad de evaluar adecuadamente la capacidad de generalización de los modelos y dificultan la comparación sistemática entre diferentes enfoques metodológicos.

En este contexto, surge la necesidad de desarrollar investigaciones que contribuyan al fortalecimiento metodológico en el análisis predictivo de criptomonedas, incorporando procedimientos más rigurosos en el tratamiento de los datos, la construcción de modelos y la evaluación del desempeño predictivo. La presente investigación se orienta

precisamente a abordar estas limitaciones mediante la implementación de un enfoque que integra técnicas de análisis de series temporales con modelos de aprendizaje profundo, incorporando además procedimientos de preprocesamiento que permitan garantizar propiedades estadísticas adecuadas en la serie analizada.

La elección del Bitcoin como objeto de estudio responde a su relevancia dentro del mercado de las criptomonedas y a la disponibilidad de información histórica que permite realizar análisis cuantitativos consistentes. Asimismo, el enfoque metodológico propuesto busca aportar al campo de la modelación de series temporales financieras mediante la incorporación de esquemas de evaluación más exigentes, incluyendo la generación de escenarios sintéticos fuera de la muestra que preserven las propiedades estadísticas del proceso generador de datos.

En términos académicos, el aporte de esta investigación radica en la propuesta de un marco metodológico que articula técnicas de preprocesamiento, modelación mediante aprendizaje profundo y procedimientos de validación orientados a mejorar la confiabilidad de los pronósticos en mercados caracterizados por alta incertidumbre. De esta manera, el estudio contribuye al desarrollo de herramientas analíticas más robustas para el análisis predictivo de criptomonedas, así como al fortalecimiento de la investigación en el campo de las finanzas computacionales y la ciencia para la toma de decisiones.

3. Objetivos

3.1 Objetivo general

Evaluar métodos de Deep Learning que minimicen el error predictivo y mejoren la capacidad de realizar pronósticos por fuera de la muestra en la serie de tiempo del Bitcoin.

3.2 Objetivos específicos

- Identificar los principales algoritmos de Deep Learning para predecir series de tiempo.
- Aplicar los modelos identificados a la serie de tiempo del Bitcoin.
- Validar el ajuste de los modelos de Deep Learning.
- Evaluar la capacidad predictiva de los modelos por fuera de la muestra.

4. Hipótesis

4.1 Hipótesis de trabajo

H_1: La aplicación de modelos de aprendizaje profundo, particularmente arquitecturas LSTM, combinada con un adecuado preprocesamiento de la serie temporal y una evaluación rigurosa del pronóstico verdaderamente fuera de la muestra, mejora significativamente la precisión predictiva del precio del Bitcoin en comparación con enfoques basados únicamente en ajustes dentro de la muestra o modelos tradicionales.

H_0: La aplicación de modelos de aprendizaje profundo con preprocesamiento y evaluación fuera de la muestra no produce mejoras significativas en la precisión

predictiva del precio del Bitcoin frente a enfoques tradicionales o evaluaciones únicamente dentro de la muestra.

5. Marco teórico

El análisis y modelación de series temporales financieras constituye una herramienta fundamental para comprender la dinámica de los mercados y apoyar procesos de pronóstico y toma de decisiones. En términos generales, una serie temporal se define como una secuencia de observaciones registradas en intervalos sucesivos de tiempo, cuyo análisis permite identificar patrones, dependencias temporales y posibles tendencias en el comportamiento de una variable. En el ámbito financiero, estas series suelen representar variables como precios de activos, rendimientos o volatilidad, las cuales se caracterizan por presentar alta variabilidad, dependencia temporal y posibles cambios estructurales en su dinámica.

Las series financieras presentan propiedades estadísticas particulares que deben ser consideradas durante el proceso de modelación. Entre estas propiedades se encuentra la dependencia temporal, que implica que el valor de una observación puede estar relacionado con valores pasados de la misma serie. Esta característica es fundamental para el desarrollo de modelos predictivos, ya que permite utilizar información histórica para estimar el comportamiento futuro del sistema.

Un concepto central en el análisis de series temporales es la estacionariedad, que se refiere a la estabilidad de las propiedades estadísticas de la serie a lo largo del tiempo. Una serie estacionaria mantiene constantes su media, su varianza y su estructura de autocorrelación. Este supuesto es especialmente relevante en numerosos modelos estadísticos utilizados para el análisis de series temporales, ya que la presencia de tendencias o cambios en la varianza puede generar estimaciones inestables o pronósticos poco confiables. En el caso de las series financieras, es común que los precios presenten comportamientos no estacionarios, lo que requiere la aplicación de procedimientos de transformación o diferenciación que permitan aproximar la serie a una estructura estadística más adecuada para su modelación.

Con el fin de mejorar las propiedades estadísticas de una serie temporal, es frecuente aplicar transformaciones matemáticas orientadas a estabilizar la varianza o reducir la heterocedasticidad. Entre las transformaciones más utilizadas se encuentra la transformación de Box-Cox, la cual permite modificar la escala de la serie mediante un parámetro que ajusta su distribución. La aplicación de este tipo de transformaciones contribuye a mejorar el cumplimiento de los supuestos estadísticos requeridos por diversos modelos y facilita el proceso de estimación y pronóstico. Posteriormente, los valores pueden recuperarse en su escala original mediante la transformación inversa correspondiente.

En el ámbito de la modelación estadística de series temporales, uno de los enfoques más utilizados es la metodología Box-Jenkins, la cual proporciona un procedimiento sistemático para identificar, estimar y validar modelos autorregresivos integrados de media móvil. Estos modelos, conocidos como ARIMA, permiten representar la dinámica de una serie temporal a partir de la combinación de componentes autorregresivos, diferenciaciones y términos de media móvil. El objetivo de este enfoque es capturar la dependencia temporal presente en la serie y generar pronósticos basados en la estructura

estadística observada en los datos históricos. Cuando las series temporales presentan patrones repetitivos asociados a ciclos temporales específicos, es posible extender estos modelos mediante la incorporación de componentes estacionales. En este caso se utilizan los modelos SARIMA, los cuales permiten modelar simultáneamente la dinámica no estacional y estacional de la serie.

Además de los enfoques estadísticos tradicionales, el desarrollo reciente de técnicas de aprendizaje automático y aprendizaje profundo ha ampliado las posibilidades de modelación de series temporales. Estas metodologías se caracterizan por su capacidad para capturar relaciones no lineales complejas y aprender patrones a partir de grandes volúmenes de datos. Dentro de este grupo se encuentran las redes neuronales artificiales, las cuales procesan la información mediante capas de nodos interconectados que ajustan sus parámetros durante el proceso de entrenamiento.

En el caso de datos secuenciales, como las series temporales, resultan especialmente relevantes las redes neuronales recurrentes, diseñadas para incorporar información proveniente de observaciones anteriores dentro del proceso de aprendizaje. Estas arquitecturas permiten modelar dependencias temporales al mantener estados internos que almacenan información del pasado. Entre las variantes más utilizadas se encuentran las arquitecturas Long Short-Term Memory (LSTM) y Gated Recurrent Unit (GRU), las cuales incorporan mecanismos de compuertas que regulan el flujo de información dentro de la red y permiten capturar dependencias temporales de largo plazo, reduciendo problemas asociados al desvanecimiento del gradiente durante el entrenamiento.

Por otra parte, en el análisis de fenómenos financieros caracterizados por un alto grado de incertidumbre, es común utilizar métodos de simulación probabilística para explorar posibles trayectorias futuras del sistema. Entre estos métodos destaca la simulación Montecarlo, la cual consiste en generar múltiples escenarios posibles a partir de un modelo estimado mediante la incorporación de componentes aleatorios. Este enfoque permite representar la variabilidad inherente a los procesos estocásticos y analizar diferentes realizaciones posibles del comportamiento futuro de una serie temporal.

En este tipo de simulaciones, la variabilidad del proceso suele representarse mediante los residuales del modelo, los cuales capturan la parte no explicada por la estructura determinística del sistema. Al incorporar residuales aleatorios al valor medio estimado por el modelo, es posible generar trayectorias sintéticas que reproducen la variabilidad observada en los datos históricos. Estas trayectorias no deben interpretarse como predicciones puntuales, sino como escenarios probabilísticos que reflejan diferentes realizaciones posibles del proceso generador de datos.

Finalmente, cuando se generan series sintéticas mediante simulación, resulta necesario verificar que estas reproduzcan adecuadamente las propiedades estadísticas de la serie original. Para este propósito se utilizan pruebas estadísticas de comparación de distribuciones, como la prueba de Kolmogórov-Smirnov de dos muestras, la cual permite evaluar si dos conjuntos de datos provienen de una misma distribución mediante la comparación de sus funciones de distribución acumulada. La aplicación de este tipo de pruebas contribuye a validar que las series simuladas mantengan características estadísticas consistentes con la serie histórica.

En conjunto, estos fundamentos teóricos proporcionan el marco conceptual necesario para el análisis, modelación y simulación de series temporales financieras, integrando enfoques estadísticos tradicionales con metodologías modernas de aprendizaje profundo y técnicas de simulación probabilística.

6. Metodología

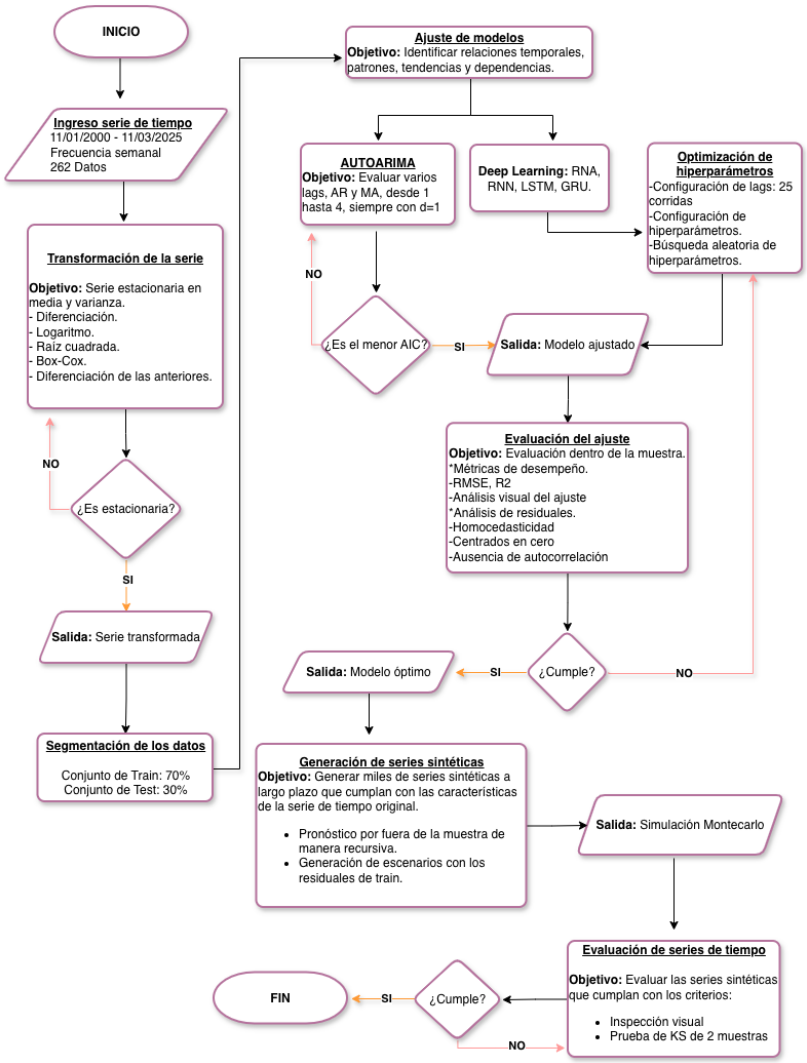
El presente capítulo describe el procedimiento metodológico seguido para el modelado, evaluación y simulación de la serie temporal del precio del Bitcoin. La metodología se estructura en cinco etapas principales: (i) recopilación y preprocesamiento de los datos, (ii) modelación mediante enfoques estadísticos tradicionales, (iii) modelación mediante técnicas de Deep Learning (DL), (iv) evaluación integral del desempeño de los modelos y (v) generación de series sintéticas mediante simulación. Esta estructura permite comparar de forma sistemática el desempeño de modelos clásicos y modelos basados en aprendizaje profundo, así como analizar su capacidad para reproducir la dinámica temporal de la serie original.

En el marco de este procedimiento metodológico, la información utilizada se organiza en tres conjuntos de datos diferenciados: el conjunto de entrenamiento, utilizado para estimar los parámetros de los modelos; el conjunto de prueba, correspondiente a una partición del período observado que no participa en el entrenamiento y se emplea para evaluar el desempeño predictivo; y los escenarios futuros simulados, generados mediante simulación Montecarlo a partir del modelo estimado, los cuales representan trayectorias probabilísticas posteriores al período observado.

El conjunto de entrenamiento abarca el período comprendido entre el 01 de noviembre de 2020 y el 28 de abril de 2024, correspondiente a un total de 183 observaciones. Por su parte, el conjunto de prueba está compuesto por 79 observaciones, que cubren el intervalo entre el 05 de mayo de 2024 y el 02 de noviembre de 2025, fecha que marca el final del período de evaluación. Adicionalmente, se considera un horizonte de simulación de 52 semanas, comprendido entre el 03 de noviembre de 2025 y el 02 de noviembre de 2026, el cual se utiliza para generar trayectorias sintéticas mediante simulación Montecarlo a partir de los modelos estimados.

La Figura 1 presenta de manera esquemática el proceso metodológico seguido en este estudio para el análisis y modelación de la serie temporal.

Figura 1. *Esquema general para el pronóstico de la serie de tiempo del BTC*



El diagrama de la Figura 1 resume la secuencia de etapas implementadas, así como los mecanismos de validación y retroalimentación utilizados, permitiendo una comprensión general del flujo de trabajo adoptado. A partir de este esquema, el procedimiento se organiza en tres fases principales, que se describen a continuación.

En la primera fase se realiza la recopilación de los datos provenientes de Yahoo Finance, correspondientes al período comprendido entre el 1 de noviembre de 2020 y el 3 de noviembre de 2025. La frecuencia de muestreo es semanal, lo que da como resultado un total de 262 observaciones.

Posteriormente, se aplican diversas transformaciones con el propósito de garantizar la estacionariedad en media y varianza. Entre las transformaciones consideradas se incluyen: diferenciación simple, logaritmo natural, raíz cuadrada, transformación Box-Cox y la diferenciación sucesiva de cada una de estas transformaciones. Si la serie resultante cumple con los criterios de estacionariedad, se adopta como serie transformada; en caso contrario, el proceso retorna al paso anterior para aplicar una transformación adicional.

Una vez obtenida la serie estacionaria, se procede a la partición de los datos en conjuntos de entrenamiento y prueba, asignando el 70% de las observaciones al primero y el 30% al segundo. Se optó por dicha partición debido a la limitada cantidad de datos disponibles, asegurando así un número suficiente de observaciones para el conjunto de prueba. En la literatura se ha identificado que los modelos de redes neuronales son sensibles a la composición del conjunto de entrenamiento, especialmente cuando las series no se transforman para garantizar estacionariedad, lo que puede generar diferencias en los patrones entre entrenamiento y prueba y afectar la capacidad predictiva. En esta tesis, al transformar las series a estacionarias, se asegura que los patrones observados en el conjunto de entrenamiento se mantengan consistentes en el conjunto de prueba, reduciendo la sensibilidad del modelo a la partición de los datos y fortaleciendo la validez del ajuste y de los pronósticos.

El período de análisis no se extendió a años anteriores al 2020, esto debido a que a partir de la pandemia del COVID-19, el comportamiento del mercado de las criptomonedas y particularmente del Bitcoin experimentó cambios significativos en su dinámica y volatilidad. La inclusión de datos más antiguos introducía patrones estructurales distintos que dificultaban la obtención de una serie estacionaria, incluso después de aplicar transformaciones estadísticas. Por esta razón, se optó por trabajar con un período más reciente, caracterizado por una dinámica más homogénea en términos de variabilidad.

La segunda fase corresponde al ajuste de los modelos de pronóstico. En primer lugar, se aplica el procedimiento AUTOARIMA como método de identificación automática del modelo ARIMA, con el fin de evaluar diferentes combinaciones de rezagos para los componentes autorregresivos (AR) y de medias móviles (MA), variando entre 1 y 4, y manteniendo el parámetro de diferenciación en d=1. El modelo óptimo se selecciona con base en el menor valor del criterio de información de Akaike (AIC).

De manera complementaria, se implementan modelos de aprendizaje profundo, incluyendo Redes Neuronales Artificiales (RNA), Redes Neuronales Recurrentes (RNN), Memoria a Largo y Corto Plazo (LSTM) y Unidades Recurrentes con Compuertas (GRU). Para todos los modelos se realiza la definición de la estructura de rezagos, la configuración de hiperparámetros y la aplicación de un procedimiento de búsqueda aleatoria para su optimización. Los modelos resultantes se evalúan dentro de la muestra mediante métricas de desempeño como la raíz del error cuadrático medio (RMSE) y el coeficiente de determinación (R^2), además de un análisis visual del ajuste.

Asimismo, se examinan los residuales con el fin de verificar el cumplimiento de los supuestos fundamentales: homocedasticidad, centrado en cero y ausencia de autocorrelación. Si estas condiciones se satisfacen, el modelo se considera óptimo; de lo contrario, se retorna al proceso de ajuste del modelo.

La tercera y última fase consiste en la generación de series sintéticas a partir del modelo óptimo obtenido en la etapa anterior. El objetivo es producir miles de trayectorias sintéticas de largo plazo (52 semanas) que reproduzcan las características estadísticas de la serie original. Para ello, se emplea un esquema de pronóstico recursivo fuera de la muestra y la generación de escenarios mediante la utilización de los residuales del conjunto de entrenamiento, lo cual permite realizar una simulación Montecarlo.

Las series sintéticas generadas se evaluaron mediante un proceso que combinó inspección visual y la aplicación de la prueba de Kolmogórov-Smirnov de dos muestras (KS2). Esta prueba se utilizó para comparar directamente las observaciones semanales de cada serie sintética con la serie original, empleando los residuales previamente preprocesados para reducir la dependencia temporal inherente a los datos. De esta manera, fue posible verificar si las series simuladas conservaban la misma distribución estadística que la serie histórica, asegurando que se mantuvieran sus propiedades fundamentales.

Para aquellas series que no cumplían con los criterios establecidos por la prueba KS2, se procedía a generarlas nuevamente mediante el proceso de simulación, repitiendo este procedimiento de manera iterativa hasta obtener series sintéticas que presentaran un comportamiento estadístico consistente con la serie original. Durante la simulación, se utilizó la distribución empírica de los residuales con el objetivo de preservar las propiedades estadísticas observadas, incorporando al valor medio pronosticado por el modelo ajustado un residual aleatorio que reproduce la variabilidad inherente a la serie.

En consecuencia, cada serie sintética constituye un escenario probabilístico generado a partir de simulación Montecarlo. Por ello, los resultados obtenidos deben interpretarse como trayectorias sintéticas con carácter probabilístico, y no como un pronóstico puntual del comportamiento de la serie de precios. Este enfoque garantiza que los escenarios generados reflejen de manera realista la incertidumbre y las dinámicas estadísticas de la serie histórica, fortaleciendo la validez y robustez de los pronósticos obtenidos.

6.1 Modelos de series de tiempo: enfoque Box-Jenkins

Box et al. (2015) proponen una metodología basada en un proceso sistemático compuesto por las etapas de identificación, estimación, diagnóstico y validación del modelo. Un supuesto fundamental de este enfoque es que la serie de tiempo debe ser estacionaria en media y varianza; en caso contrario, se requiere aplicar transformaciones como la diferenciación para alcanzar la estacionariedad. Durante la etapa de identificación, se analizan las funciones de autocorrelación (ACF) y autocorrelación parcial (PACF) con el fin de determinar los órdenes autorregresivos y de medias móviles del modelo. Posteriormente, los parámetros se estiman mediante métodos de máxima verosimilitud, y se valida el modelo a través del análisis de los residuales, los cuales deben comportarse como ruido blanco, es decir, no presentar autocorrelación ni patrones sistemáticos. Finalmente, se emplean criterios de información como el AIC o el Criterio de Información Bayesiano (BIC) para comparar especificaciones alternativas y seleccionar el modelo más parsimonioso con adecuado poder predictivo.

6.1.1 Transformación de la serie de tiempo

En el análisis de series temporales financieras es común aplicar transformaciones a los datos originales con el objetivo de mejorar sus propiedades estadísticas, particularmente en términos de estabilidad de la varianza, reducción de asimetrías y aproximación a la estacionariedad. Estas transformaciones facilitan el proceso de modelación y contribuyen a mejorar el desempeño de los modelos predictivos. Entre las transformaciones más utilizadas en la literatura se encuentra la transformación de Box-Cox, propuesta por George Box y David Cox (1964), la cual permite estabilizar la varianza de la serie mediante un parámetro de transformación ajustable.

La transformación Box-Cox se define de la siguiente manera:

$$y^{(\lambda)} = \begin{cases} \dfrac{y^{\lambda} - 1}{\lambda}, si\ \lambda \neq 0 \\ \ln(y), si\ \lambda = 0 \end{cases}$$

Donde *y* representa la variable original de la serie temporal y λ es el parámetro de transformación que determina el grado de ajuste aplicado a los datos. El valor de λ suele estimarse mediante métodos de máxima verosimilitud, buscando el parámetro que mejor estabilice la varianza de la serie.

Una vez realizada la modelación sobre la serie transformada, es necesario aplicar la transformación inversa para interpretar los resultados en la escala original de los datos. La función inversa de la transformación Box-Cox se expresa como:

$$y = \begin{cases} (\lambda y^{(\lambda)} + 1)^{1/\lambda}, & si\ \lambda \neq 0 \\ exp(y^{(\lambda)}), & si\ \lambda = 0 \end{cases}$$

En el caso en que λ≠0, la recuperación de los valores originales se realiza mediante la expresión:

$$y = (\lambda y^{(\lambda)} + 1)^{\mathbf{1}/\lambda}$$

Por el contrario, cuando λ=0, la transformación corresponde al logaritmo natural, por lo que la inversión se realiza mediante la función exponencial.

El uso de este tipo de transformaciones es ampliamente documentado en la literatura de análisis de series temporales, particularmente en contextos donde las series presentan heterocedasticidad o varianza no constante. Autores como George Box, Gwilym Jenkins, Gregory Reinsel y Greta Ljung han destacado la importancia de estabilizar la varianza antes de aplicar modelos estadísticos, ya que esto contribuye a garantizar el cumplimiento de los supuestos del modelo y mejora la calidad de los pronósticos obtenidos.

En el contexto de la presente investigación, la aplicación de transformaciones constituye una etapa fundamental del proceso de preprocesamiento de la serie temporal. Este procedimiento permite reducir problemas asociados con la heterocedasticidad y facilita la posterior obtención de residuales aproximadamente estacionarios, condición necesaria para la estimación adecuada del modelo y para la generación de escenarios sintéticos mediante simulación.

6.1.2 Ajuste del procedimiento SARIMA

Los Modelos Autorregresivos Integrados de Media Móvil Estacional (SARIMA) constituyen una extensión del modelo ARIMA tradicional, al incorporar de manera explícita componentes estacionales en la estructura autorregresiva y de media móvil. De acuerdo con Box et al. (2015),un modelo SARIMA (p,d,q)(P,D,Q)s combina términos no estacionales y estacionales, permitiendo representar series que presentan patrones

periódicos recurrentes en el tiempo, una característica común en series económicas y financieras. Formalmente, el modelo se expresa mediante operadores de rezago como:

$$\phi(B)\Phi(B^s)(1-B)^d(1-B^s)^D Y_t = \theta(B)\Theta(B^s)\epsilon_t$$

Donde los polinomios $\phi(B)$ y $\theta(B)$ representan los componentes autorregresivos y de media móvil no estacionales, respectivamente, mientras que $\Phi(B^s)$ y $\Theta(B^s)$ corresponden a los componentes estacionales con período s. Las diferenciaciones de orden d y D permiten eliminar la no estacionariedad tanto en el nivel como en la componente estacional de la serie.

La validación del modelo se realiza mediante el análisis de los residuales, los cuales deben comportarse como un proceso de ruido blanco, es decir, presentar media cero, varianza constante y ausencia de autocorrelación significativa. Este análisis garantiza que la estructura temporal, incluida la estacionalidad, ha sido adecuadamente capturada y que el modelo es apropiado para fines de inferencia y pronóstico.

6.1.3 Selección del modelo mediante AUTOARIMA

Para la identificación automática de los parámetros del SARIMA se emplea el procedimiento AUTOARIMA, el cual evalúa múltiples combinaciones de rezagos para los componentes autorregresivos y de medias móviles, tanto estacionales como no estacionales, seleccionando el modelo óptimo con base en el menor valor de AIC. Este enfoque permite automatizar la etapa de identificación dentro del procedimiento Box-Jenkins, reduciendo la subjetividad en la selección de parámetros.

6.2 Métodos de Deep Learning

Los métodos de DL constituyen una clase de modelos de aprendizaje automático basados en redes neuronales artificiales con múltiples capas, capaces de modelar relaciones altamente no lineales y patrones complejos presentes en los datos. A diferencia de los modelos estadísticos tradicionales, estos enfoques aprenden representaciones jerárquicas directamente a partir de la información histórica, lo que los hace especialmente adecuados para la predicción de series de tiempo no estacionarias y altamente volátiles.

En esta investigación se emplean distintas arquitecturas, incluyendo RNA, RNN y variantes avanzadas como LSTM y GRU, con el fin de capturar tanto relaciones no lineales como dependencias temporales de corto y largo plazo presentes en la serie de precios del Bitcoin.

6.2.1 Red Neuronal Artificial

Según Haykin (1999), una red neuronal artificial está compuesta por neuronas que calculan una combinación lineal de las entradas seguida de una función de activación no lineal. Matemáticamente, la salida de la neurona j en una capa se obtiene a partir de las siguientes expresiones:

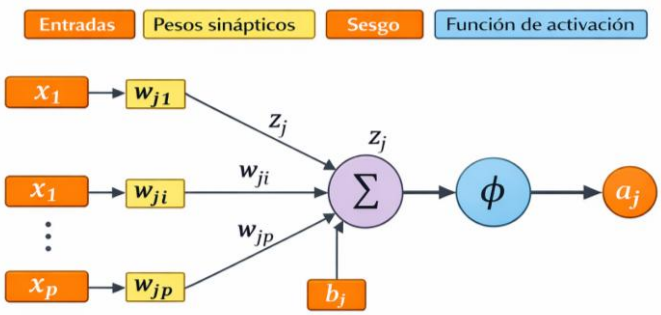
$$z_j = \sum_{i=1}^{p} w_{ji} x_i + b_j$$

$$a_j = \phi(z_j)$$

Donde x_i representa la i-ésima variable de entrada, w_{ji} es el peso asociado a la conexión entre la entrada *i* y la neurona *j*, b_j es el término de sesgo (*bias*), *p* es el número de entradas, z_j corresponde a la activación neta o potencial interno de la neurona y $\phi(\cdot)$ denota la función de activación no lineal, que puede ser sigmoide, tangente hiperbólica, ReLU o variantes modernas como ELU, dependiendo del diseño de la arquitectura. La inclusión de funciones de activación no lineales es esencial para que la red pueda aproximar relaciones no lineales complejas entre las variables de entrada y la salida del modelo (Goodfellow, Bengio y Courville, 2016).

A partir de las ecuaciones que describen el funcionamiento de la neurona artificial, la Figura 2 presenta un esquema general de la arquitectura de una RNA, ilustrando el flujo de información desde las variables de entrada, la combinación lineal mediante los pesos sinápticos y el sesgo, hasta la aplicación de la función de activación que produce la salida de cada neurona.

Figura 2. *Estructura de la red Neuronal RNA*



Fuente: *Elaboración propia*

En el esquema de la Figura 2 se observa cómo las salidas de una capa sirven como entradas para la capa siguiente, permitiendo que la red construya representaciones jerárquicas de los datos. Este mecanismo de propagación hacia adelante (*forward propagation*) es el que posibilita que la red modele relaciones no lineales entre las variables, mientras que los parámetros w_{ji} y b_{jb} se ajustan durante el proceso de entrenamiento mediante algoritmos de optimización basados en gradiente, con el objetivo de minimizar el error de predicción. De esta forma, la RNA aprende una función de aproximación flexible capaz de capturar patrones complejos presentes en la serie temporal.

6.2.2 Red Neuronal Recurrente

Patel et al. (2020) describen que una RNN es una arquitectura de aprendizaje profundo diseñada para modelar dependencias temporales en los datos. A diferencia de las redes feedforward, en las que las observaciones de entrada se procesan de manera independiente, las RNN incorporan conexiones recurrentes que permiten que la información de estados previos influya en el procesamiento actual. Esto se logra mediante un estado oculto**,** el cual actúa como una memoria que captura la dinámica secuencial de la serie. El comportamiento de una neurona en una RNN puede describirse mediante las ecuaciones:

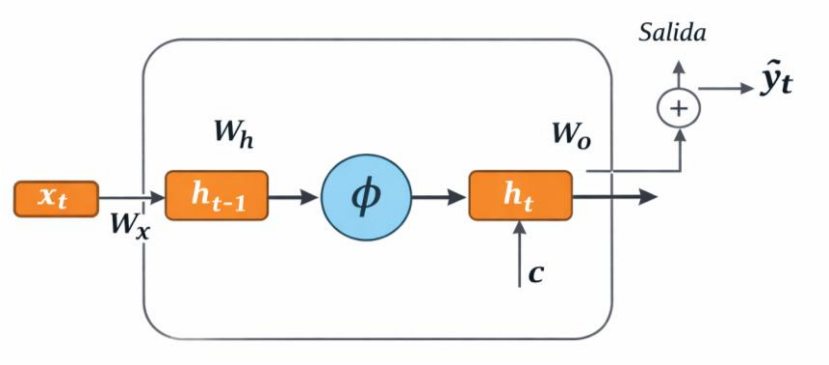
$$h_t = \phi(W_x x_t + W_h h_{t-1} + b)$$

$$y_t = W_o h_t + c$$

Donde x_t es el vector de entrada en el instante *t*, h_t representa el estado oculto, W_x, W_h y W_o son las matrices de pesos asociadas a la entrada, a la recurrencia y a la salida, respectivamente, *b* y *c* son términos de sesgo, y Ø es una función de activación no lineal, típicamente *tanh* o *ReLU* (Patel et al., 2020).

En extensión al modelo de red neuronal artificial tradicional, la Figura 3 presenta la estructura de una RNN, en la cual se incorporan conexiones de retroalimentación que permiten que la información de estados anteriores influya en el procesamiento del estado actual, habilitando así el modelado explícito de dependencias temporales en datos secuenciales.

Figura 3.*Estructura de la red Neuronal RNN*



Fuente: *Elaboración propia.*

En el esquema de la Figura 3 se ilustra cómo, en cada instante de tiempo *t*, la RNN recibe como entradas tanto el vector de observaciones x_t como el estado oculto del periodo previo h_{t-1}. Estos componentes se combinan mediante transformaciones lineales y una función de activación no lineal para generar el nuevo estado oculto h_t, el cual resume la información relevante acumulada hasta ese momento. Posteriormente, el estado oculto se proyecta hacia la capa de salida para producir la predicción correspondiente.

6.2.3 Long Short-Term Memory

La red LSTM es una arquitectura recurrente diseñada para superar las limitaciones de las RNN convencionales en la captura de dependencias de largo plazo. Para ello, introduce una estructura de memoria interna controlada mediante compuertas que regulan el flujo de información a lo largo del tiempo.

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f)$$

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i)$$

$$\tilde{c}_t = \tanh(W_g x_t + U_g h_{t-1} + b_g)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t$$

$$o_t = \sigma(W_0 x_t + U_0 h_{t-1} + b_0)$$

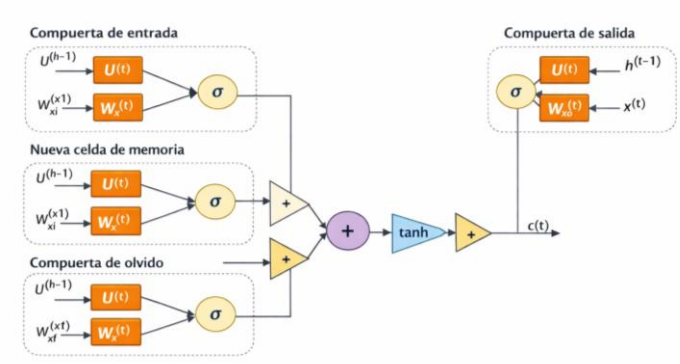
$$h_t = O_t \odot \tanh(C_t)$$

Las redes LSTM incorporan un mecanismo de memoria basado en tres compuertas que regulan el flujo de información dentro de la celda, lo cual permite capturar dependencias temporales de largo plazo (Ammer & Aldhyani, 2022). En este sentido, Ammer y Aldhyani (2022) destacan que la compuerta de olvido (f_t) controla qué parte del estado previo debe conservarse, la compuerta de entrada (i_t) determina cuánta información nueva se incorpora al estado de la celda, y la compuerta de salida (o_t) regula la información que se transfiere al estado oculto.

El estado de memoria se actualiza mediante una combinación del estado previo y un estado candidato, el cual se calcula a partir de la entrada actual y del estado oculto anterior, utilizando funciones de activación no lineales. En este proceso intervienen matrices de pesos asociadas a las entradas y a los estados previos, así como términos de sesgo, y se emplean funciones de activación sigmoide y tangente hiperbólica para modelar relaciones no lineales. Este mecanismo permite a la red conservar información relevante durante largos horizontes temporales, superando las limitaciones de las redes recurrentes convencionales (Ammer & Aldhyani, 2022).

La Figura 4 presenta la estructura de una red LSTM, la cual incorpora un mecanismo de memoria interna regulado por compuertas que controlan de manera explícita el flujo de información a lo largo del tiempo.

Figura 4. *Estructura de la red Neuronal LSTM*



Fuente: Adaptado de *Ammer y Aldhyani (2022).*

En la Figura 4 se observa que la LSTM mantiene un estado de celda C_t que actúa como una memoria de largo plazo, mientras que el estado oculto h_t representa la salida a corto plazo del sistema. La actualización de la memoria está gobernada por tres compuertas: la compuerta de olvido, que determina qué información previa debe descartarse; la compuerta de entrada, que controla la incorporación de nueva información relevante; y la compuerta de salida, que regula qué parte del contenido de la celda se transmite como

salida. Este diseño permite preservar gradientes estables durante el entrenamiento y facilita el aprendizaje de relaciones temporales complejas, lo que hace a las LSTM especialmente adecuadas para series financieras caracterizadas por cambios estructurales y alta volatilidad.

6.2.4 Gated Recurrent Unit

La red GRU es una variante de las redes neuronales recurrentes diseñada para capturar dependencias temporales de largo plazo con una arquitectura más simple que la LSTM. A diferencia de esta última, la GRU combina el estado oculto y el estado de memoria en una única representación, reduciendo el número de parámetros y, por ende, la complejidad computacional del modelo.

De acuerdo con Patel et al.(2020) la GRU emplea dos compuertas principales: la compuerta de actualización, que controla cuánto de la información pasada se conserva, y la compuerta de reinicio, que determina en qué medida el estado previo debe ser ignorado al incorporar nueva información. Esta estructura permite un equilibrio entre retención de memoria y adaptación a nuevos patrones, siendo especialmente eficiente en problemas de series de tiempo con dependencias temporales relevantes. Su formulación matemática es la siguiente:

$$z_t = \sigma(W_{zx_t} + U_z h_{t-1} + b_z)$$

$$r_t = \sigma(W_{rx_t} + U_r h_{t-1} + b_r)$$

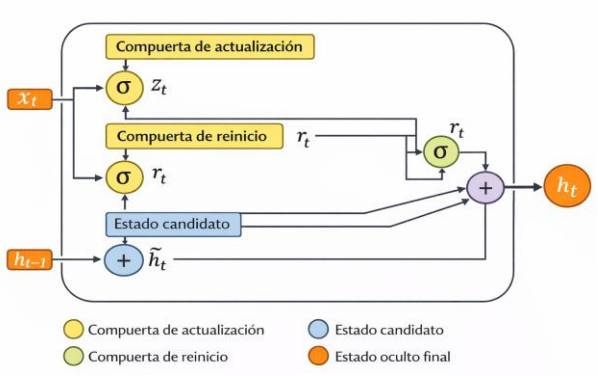
$$\tilde{h}_t = \tanh(W_{hx_t} + U_h(r_t \odot h_{t-1}) + b_h)$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t$$

Donde z_t es la compuerta de actualización, r_t la compuerta de reinicio, h_t el estado oculto, x_t la entrada en el tiempo t, $\sigma(\cdot)$ la función sigmoide y $\odot$ denota el producto elemento a elemento.

La Figura 5 muestra la arquitectura de la red GRU, la cual simplifica el mecanismo de memoria mediante la integración de compuertas que combinan el control del estado previo y la actualización de la información relevante.

Figura 5. Estructura de la red Neuronal GRU



Fuente: *Elaboración propia.*

En la arquitectura de la Figura 5, el estado oculto h_t cumple simultáneamente el papel de memoria y salida del sistema. La compuerta de actualización regula el grado en que el estado previo se conserva, mientras que la compuerta de reinicio controla cuánta información pasada se descarta al calcular el nuevo contenido candidato.

Las arquitecturas descritas RNA, RNN, LSTM y GRU presentan diferentes capacidades para modelar relaciones no lineales y dependencias temporales en series de tiempo. Mientras que las RNA permiten capturar relaciones estáticas complejas entre variables, las arquitecturas recurrentes incorporan información secuencial, siendo las LSTM y GRU particularmente adecuadas para modelar dependencias de largo plazo.

Siguiendo a Wu et al. (2019), las redes LSTM han demostrado ser especialmente adecuadas para la predicción de precios de criptomonedas debido a su capacidad para modelar dependencias temporales de corto y largo plazo en contextos altamente volátiles y no estacionarios. Su estructura de memoria interna permite mitigar el problema de desaparición del gradiente y capturar relaciones no lineales complejas, superando el desempeño de modelos tradicionales y de RNN convencionales en tareas de predicción financiera.

6.3 Optimización de hiperparámetros

Durante el proceso de diseño e implementación de los modelos de aprendizaje profundo, se identificaron y ajustaron diversos hiperparámetros clave que afectan tanto la arquitectura con el rendimiento del modelo. Entre ellos se incluyeron el tipo de Red neuronal utilizada (RNA, RNN, LSTM o GRU), la bidireccionalidad de las capas, el número de capas ocultas y de unidades por capa, así como la función de activación. También se configuraron parámetros relacionados con el entrenamiento, como la cantidad de épocas, el tamaño del lote, la proporción de rezagos temporales considerados (lags) y el tipo de optimizador. La regularización se aplicó mediante Dropout, y se utilizó detención temprana (early stopping) para prevenir el sobreajuste.

La Tabla 7 presenta un resumen de los principales hiperparámetros considerados en el diseño y ajuste de los modelos, junto con sus respectivas configuraciones y rangos de exploración.

Tabla 7. *Hiperparámetros considerados en el diseño del modelo*

Hiperparámetro	Valores configurables
RNA, RNN, LSTM, GRU	
Capas ocultas	1, 2 y 3
Neuronas por capa	5, 10, 12, 15, 20, 25 y 30
Funciones de activación	ReLU, tanh, selu, elu y leaky_relu
Optimizadores	Adam, RMSprop
Epocas	300
Tamaño del batch	6, 12, 24, 32 y 64

Fuente. *Elaboración propia.*

La Tabla 7 muestra el proceso de diseño y optimización de los modelos, contemplando la evaluación de diferentes arquitecturas de redes neuronales, incluyendo RNA, RNN, LSTM y GRU, con el fin de identificar la configuración más adecuada para la modelación de series temporales. Se consideraron entre 1 y 3 capas ocultas, con un número de

neuronas por capa que varía entre 5, 10, 12, 15, 20, 25 y 30, permitiendo explorar distintos niveles de complejidad y capacidad de representación.

En cuanto a las funciones de activación, se evaluaron ReLU, tanh, selu y leaky Relu, con el objetivo de introducir no linealidad y favorecer la estabilidad del proceso de entrenamiento. Para la optimización de los pesos se emplearon los algoritmos Adam y RMSprop, reconocidos por su eficiencia en problemas de aprendizaje profundo. El entrenamiento se realizó durante un máximo de 300 épocas y se analizaron distintos tamaños de lote (6, 12, 24, 32 y 64) para equilibrar la estabilidad numérica y el costo computacional. Este esquema de búsqueda de hiperparámetros permitió evaluar de manera sistemática el impacto de la arquitectura y de las configuraciones de entrenamiento en el desempeño predictivo, facilitando la selección de modelos capaces de capturar las dinámicas complejas y no lineales presentes en las series temporales analizadas.

6.4 Evaluación de los modelos

En esta sección se evalúa el desempeño y la adecuación estadística de los modelos estimados mediante un conjunto de criterios complementarios. En primer lugar, se emplean métricas cuantitativas de desempeño, específicamente la raíz del error cuadrático medio (RMSE) y el coeficiente de determinación (R^2), con el fin de medir la precisión predictiva y la capacidad explicativa de los modelos tanto en entrenamiento como en validación. En segundo lugar, se realiza un análisis de los residuales para verificar el cumplimiento de supuestos estadísticos fundamentales, en particular la homocedasticidad, mediante las pruebas de Breusch-Pagan y White, lo cual permite evaluar la estabilidad de la varianza de los errores y la correcta especificación del modelo. Finalmente, se incorpora la generación de series sintéticas como una estrategia adicional para analizar el comportamiento dinámico de los modelos y su capacidad para reproducir las propiedades estadísticas y temporales de la serie original, proporcionando una evaluación más integral de su desempeño en escenarios simulados.

6.4.1 Métricas de desempeño

El desempeño de los modelos fue evaluado utilizando como métricas principales la raíz del error cuadrático medio (RMSE) y el coeficiente de determinación R^2. Estos fueron empleados tanto durante el entrenamiento como en prueba. Midiendo así la magnitud promedio de los errores entre los valores predichos y los valores reales, así como también mide la precisión de un modelo para explicar la variabilidad de los datos observados.

$$RMSE = \sqrt{1/n \sum_{i=1}^{n} (\hat{y}_t - y_t)^2}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (\hat{y}_t - y_t)^2}{\sum_{i=1}^{n} (\hat{y}_t - \bar{y})^2}$$

Posteriormente, el RMSE permite comparar directamente la precisión de distintos modelos en la misma escala de la variable objetivo, siendo preferibles valores más bajos,

ya que indican una menor discrepancia entre los valores observados y los predichos. Por su parte, el coeficiente R^2 toma valores entre $(-\infty,1]$, donde valores cercanos a 1 indican un alto nivel de explicación de la variabilidad de los datos, mientras que valores cercanos a cero o negativos sugieren un desempeño similar o inferior al de un modelo que utiliza únicamente la media como predicción. En conjunto, ambas métricas ofrecen una evaluación complementaria del desempeño predictivo, permitiendo analizar tanto la magnitud del error como la capacidad explicativa de los modelos estimados.

6.4.2 Análisis de residuales

De acuerdo con la metodología Box-Jenkins, el análisis de residuales representa la fase de diagnóstico del modelo y es esencial para verificar su correcta especificación. Una vez completadas las etapas del procedimiento SARIMA, los residuales deben comportarse como un proceso de ruido blanco, lo que implica que no deben contener información sistemática adicional que pueda ser modelada (Box, Jenkins, Reinsel & Ljung, 2015). En este sentido, para esta tesis se requiere que los residuales de cada modelo cumplan con las siguientes propiedades según Box-Jenkins: estar centrados en cero, ser homocedásticos y no presentar autocorrelación serial, condiciones que aseguran que la dinámica temporal de la serie ha sido adecuadamente capturada por el modelo y que no existe estructura temporal remanente que pueda ser aprovechada para mejorar el ajuste (Box et al., 2015; Brockwell & Davis, 2016).

La ausencia de autocorrelación en los residuales se evalúa mediante las funciones de ACF y PACF, así como a través de pruebas estadísticas globales como la prueba de Ljung-Box (Ljung & Box, 1978). Si los coeficientes de autocorrelación se encuentran dentro de los intervalos de confianza, se concluye que no existe dependencia temporal remanente. Asimismo, aunque la normalidad de los residuales no es un supuesto estricto para la estimación de los procedimientos ARIMA/SARIMA, una distribución aproximadamente normal facilita la inferencia y la obtención de intervalos de confianza para los pronósticos (Brockwell & Davis, 2016).

En el contexto de esta tesis, la normalidad de los residuales no se exigió como requisito obligatorio, debido a que la generación de series sintéticas se realizó mediante simulación Montecarlo utilizando la distribución empírica de los residuales obtenidos. Este enfoque permite introducir variabilidad realista en las simulaciones, preservando la estructura estocástica observada en la serie original sin imponer una forma de distribución específica (Kleijnen, 1995; Law & Kelton, 2000). Al utilizar la distribución empírica de los residuales, se generan múltiples trayectorias sintéticas que reflejan fielmente la dispersión y heterogeneidad presentes en los datos originales, lo cual es particularmente útil cuando los residuales no se ajustan perfectamente a la normalidad, pero sí son representativos del comportamiento del proceso subyacente.

En consecuencia, cuando los residuales cumplen con las propiedades de estar centrados en cero, ser homocedásticos y no presentar autocorrelación serial, y se utilizan de manera empírica en simulaciones Montecarlo, se considera que el modelo es estadísticamente adecuado y válido tanto para el análisis como para la generación de series sintéticas y la predicción de la serie temporal.

6.4.2.1 Prueba de homocedasticidad Breusch-Pagan

La prueba de Breusch-Pagan (BP) evalúa la hipótesis de varianza constante de los errores (homocedasticidad) en un modelo de regresión. Su objetivo es detectar si la varianza de los residuales depende linealmente de una o más variables explicativas. De acuerdo con Breusch y Pagan (1979), el estadístico de prueba se define como:

$$\hat{\varepsilon}_t^2 = \alpha_0 + \sum_{i=1}^{k} \alpha_i z_{it} + u_t$$

Donde $\hat{\epsilon}^2_t$ representa el cuadrado del residual en el tiempo *t*, z_{it} son las variables explicativas del modelo original y u_t es el término de error de la regresión auxiliar. La presencia de heterocedasticidad se evalúa a partir del estadístico:

$$LM = nR^2 \sim \chi^2(k)$$

el cual sigue una distribución Chi-cuadrado con *k* grados de libertad. Valores elevados del estadístico indican que la varianza de los errores depende sistemáticamente de las variables explicativas, lo que sugiere heterocedasticidad en el modelo.

6.4.2.2 Prueba de homocedasticidad White

La prueba de White constituye un contraste general de heterocedasticidad que no impone una forma funcional específica para la varianza de los errores y permite detectar dependencias tanto lineales como no lineales con respecto a las variables explicativas (White, 1980). La prueba se implementa estimando una regresión auxiliar donde el cuadrado de los residuales del modelo principal se expresa como función de las variables originales, sus cuadrados y sus interacciones:

$$\hat{u}_i^2 = \alpha_0 + \sum_{j=1}^{k} \alpha_j z_{ij} + \sum_{j=1}^{k} \sum_{l=j}^{k} \alpha_{jl} z_{ij} z_{il} + v_i$$

Donde $\hat{u}_i^2$ son los residuales al cuadrado del modelo original para la observación *i*, a_0 es el término constante de la regresión auxiliar, z_{ij} representan las variables explicativas, $\sum_{j=1}^{k} \alpha_j z_{ij}$ captura la posible relación lineal entre la varianza del error y cada variable explicativa, $\sum_{j=1}^{k} \sum_{l=j}^{k} \alpha_{jl} z_{ij} z_{il}$ incluye los términos cuadráticos e interacciones entre variables explicativas, permitiendo detectar formas no lineales de heterocedasticidad, es decir, cuando la varianza cambia según combinaciones o potencias de las variables, a_{jl} representa los coeficientes asociados a los términos de interacción y cuadrados. Su significancia conjunta es clave para detectar heterocedasticidad general y u_i es el término de error de la regresión auxiliar, que recoge la parte de $\hat{u}_i^2$ no explicada por las variables incluidas.

A partir de esta regresión auxiliar se calcula el estadístico de contraste de White:

$$LM = nR^2 \sim \chi^2(m)$$

Donde *n* es el tamaño de la muestra, R^2 el coeficiente de determinación de la regresión auxiliar. Bajo la hipótesis nula de homocedasticidad, este estadístico sigue una distribución Chi-cuadrado con *m* grados de libertad iguales al número de regresores de dicha regresión, permitiendo evaluar formalmente la constancia de la varianza de los errores.

6.5 Generación de series sintéticas

La generación recursiva de series sintéticas consiste en producir valores futuros de una serie temporal utilizando un modelo estimado y realimentando cada pronóstico como insumo para el siguiente paso. En este enfoque, el modelo se entrena inicialmente con los datos históricos y, para cada horizonte futuro, se calcula un pronóstico al que se le añade un residual tomado aleatoriamente de los errores obtenidos durante el entrenamiento. Este residual introduce variabilidad realista, preservando la estructura estocástica observada en la serie original. El valor pronosticado más el residual se incorpora luego como nuevo dato para generar el siguiente punto, permitiendo construir trayectorias sintéticas que reflejan tanto la dinámica del modelo como la incertidumbre inherente al proceso temporal.
La generación de series sintéticas se realiza mediante simulaciones Montecarlo, asumiendo que los errores son independientes e idénticamente distribuidos, con media cero y varianza constante, de acuerdo con los supuestos de ruido blanco del enfoque Box-Jenkins. En este estudio se generaron 5.000 trayectorias sintéticas con el fin de aproximar la distribución futura de la serie y evaluar de manera robusta la variabilidad de los pronósticos a múltiples horizontes.

La generación de series sintéticas se realiza mediante simulaciones Montecarlo, donde los errores se asumen independientes e idénticamente distribuidos, con media cero y varianza constante. La ecuación del modelo es:

$$y_t^{(s)} = \hat{y}_t + \varepsilon_t^{(s)}$$

Donde $y_t{}^{(s)}$ corresponde al valor simulado de la serie en el tiempo *t* para la simulación *s*, $\hat{y}_t$ es el valor pronosticado por el modelo ajustado a los datos históricos (seleccionado mediante el procedimiento SARIMA o el modelo de aprendizaje profundo correspondiente), y $\varepsilon_t{}^{(s)}$ el residual extraído de la distribución empírica de los errores del modelo.

El análisis de las series sintéticas generadas se inicia mediante una inspección visual**,** con el fin de verificar que las trayectorias simuladas reproduzcan los principales patrones observados en la serie original, tales como tendencia, variabilidad y comportamiento temporal general. Esta evaluación preliminar permite detectar discrepancias estructurales antes de aplicar contrastes estadísticos formales, práctica común en estudios de simulación de series de tiempo (Box & Jenkins, 1994; Hyndman & Athanasopoulos, 2021).

Posteriormente, se emplea la prueba de Kolmogórov-Smirnov de dos muestras (KS2) para comparar la distribución empírica de la serie original con la de las series sintéticas. Esta prueba no paramétrica evalúa la hipótesis nula de que ambas muestras provienen de la misma distribución, a partir del estadístico:

$$D = Sup|F_1(x) - F_2(x)|. F_1(x) \text{ y } F_2(x)$$

Donde $F_1(x)$ y $F_2(x)$ representan las funciones de distribución acumulada empíricas de cada muestra (Kolmogorov, 1933; Smirnov, 1948). El no rechazo de la hipótesis nula indica que las series sintéticas reproducen adecuadamente las propiedades estadísticas de la serie original.

En conjunto, la aplicación de métricas de desempeño, el análisis de residuales y la evaluación mediante series sintéticas permite una valoración integral de la calidad de los modelos, no solo desde el punto de vista predictivo, sino también en términos de consistencia estadística y capacidad para replicar la dinámica temporal de la serie original. Este enfoque combinado fortalece la validez de las conclusiones derivadas del proceso de modelación, al reducir el riesgo de seleccionar modelos con buen ajuste aparente pero con deficiencias estructurales en sus supuestos. Los resultados obtenidos a partir de estos criterios se presentan y discuten en la siguiente sección, donde se comparan sistemáticamente los modelos tradicionales y los basados en Deep Learning en términos de precisión, estabilidad y capacidad de generalización.

7. Resultados

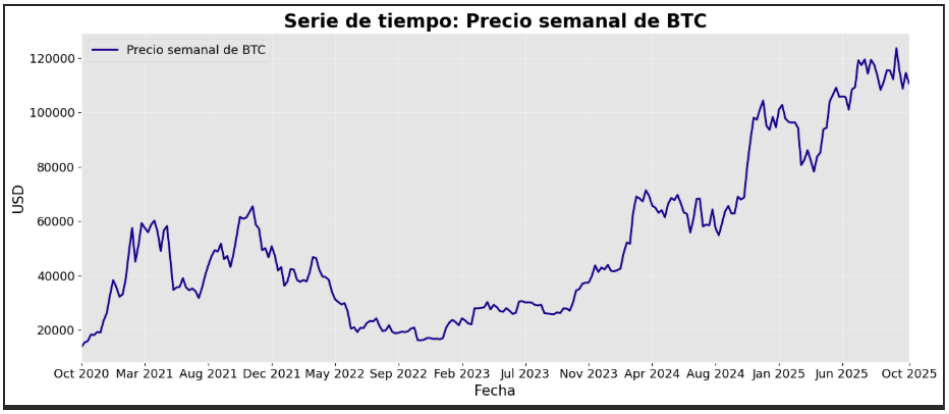
En este capítulo se presentan y analizan los resultados obtenidos a partir de la aplicación de los modelos de series de tiempo y de aprendizaje profundo descritos en el capítulo metodológico, utilizando la serie semanal de precios del Bitcoin para el período comprendido entre el 01 de noviembre de 2020 y el 03 de noviembre de 2025, con un total de 262 observaciones. El análisis contempla, en primer lugar, la exploración del comportamiento temporal de la serie y la evaluación de sus propiedades estadísticas; posteriormente, se examinan los resultados del ajuste de los modelos SARIMA y de las distintas arquitecturas de DL (RNA, RNN, LSTM y GRU), así como su desempeño predictivo mediante métricas cuantitativas y análisis de residuales. Finalmente, se presentan los resultados asociados a la generación de series sintéticas mediante simulación Montecarlo, con el fin de evaluar la capacidad de los modelos para reproducir la dinámica y la variabilidad observadas en la serie original. En conjunto, este capítulo permite comparar de manera integral el desempeño de los enfoques tradicionales y los basados en aprendizaje profundo, aportando evidencia empírica sobre su idoneidad para la modelación y predicción de precios de criptomonedas.

Los modelos evaluados utilizaron el mismo conjunto de datos de entrenamiento y prueba. Esta uniformidad garantiza que las comparaciones entre modelos sean consistentes y que los análisis de desempeño, así como la generación de series sintéticas, sean directamente comparables sin sesgos derivados de diferencias en los datos utilizados. Los pronósticos obtenidos sobre la serie transformada mediante Box-Cox fueron posteriormente re-expresados en la escala original aplicando la transformación inversa asociada al parámetro λ estimado.

La Figura 6 ilustra la evolución temporal de la serie, en la cual se identifican episodios de alta volatilidad, fases de crecimiento acelerado y correcciones abruptas, características típicas de los mercados de criptomonedas. Estas fluctuaciones pueden asociarse a factores macroeconómicos globales, decisiones de política monetaria, innovaciones tecnológicas y cambios en los marcos regulatorios, los cuales influyen significativamente en la dinámica del precio. La presencia de variabilidad elevada, posibles rupturas estructurales

y comportamientos no lineales sugiere que la serie presenta una dinámica compleja, lo que justifica la utilización de modelos capaces de capturar tanto dependencias temporales como relaciones no lineales.

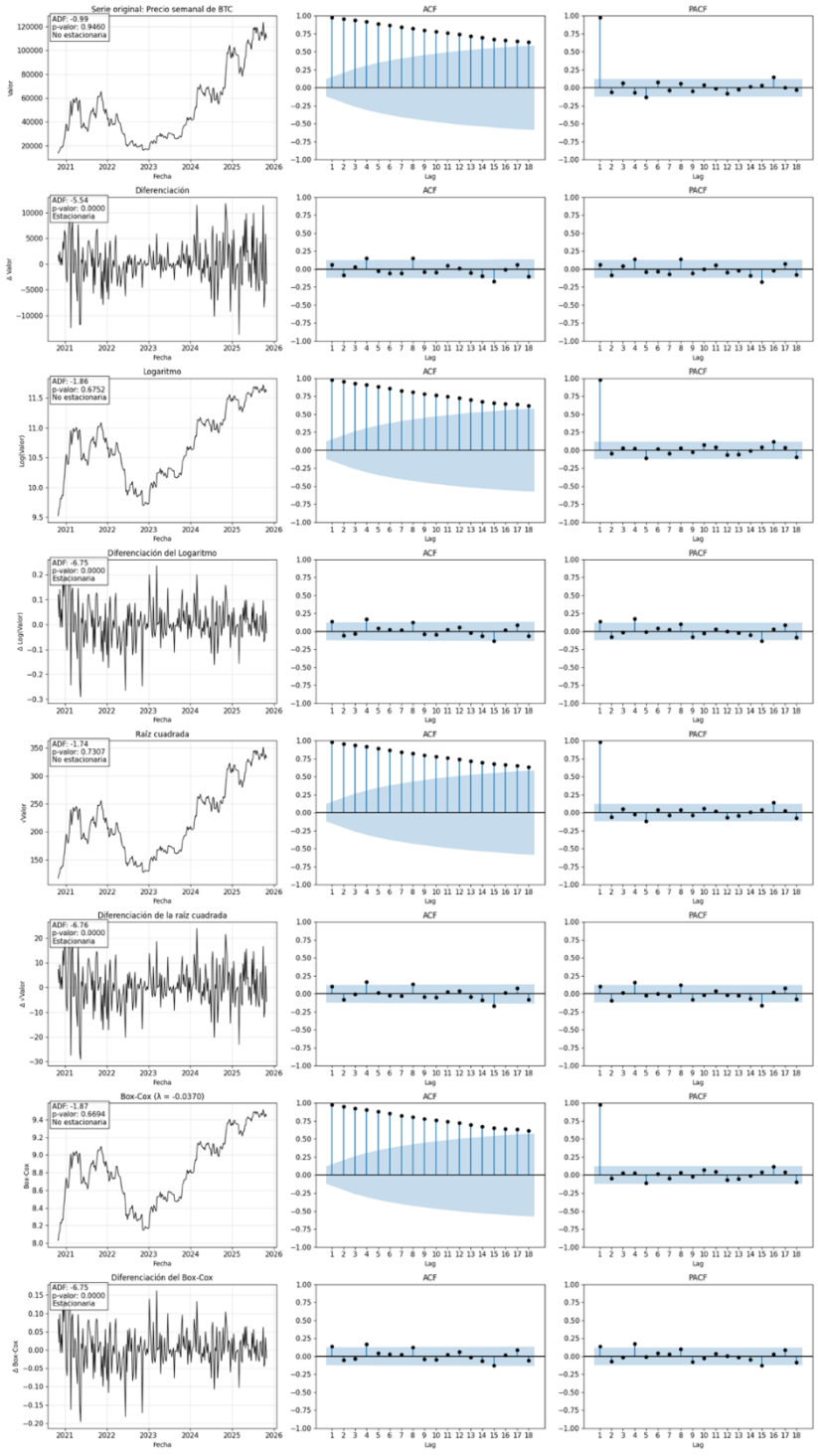
Figura 6. *Precios semanales del BTC*



La Figura 6 en concordancia con lo anterior, el análisis exploratorio y las pruebas estadísticas indican que la serie original no cumple con el supuesto de estacionariedad, al presentar cambios sistemáticos en la media y una varianza no constante a lo largo del tiempo. Estas características limitan la aplicación directa de modelos basados en supuestos clásicos y pueden afectar el desempeño predictivo si no se tratan adecuadamente. Por esta razón, se aplicaron transformaciones orientadas a estabilizar tanto la media como la varianza, incluyendo diferenciación y transformaciones funcionales como logaritmo, raíz cuadrada y Box-Cox, así como combinaciones de estas, hasta obtener una representación de la serie que satisficiera los criterios de estacionariedad requeridos para el modelado econométrico y el entrenamiento de los modelos de aprendizaje profundo.

Como parte del proceso de identificación de la estructura temporal, el análisis de las funciones ACF y PACF se empleó como herramienta diagnóstica para evaluar la persistencia temporal y orientar la selección del orden de los modelos. En este marco, la Figura 7 presenta la serie temporal del precio semanal de Bitcoin junto con las transformaciones consideradas: diferenciación, logaritmo, raíz cuadrada y Box-Cox y sus respectivas funciones ACF y PACF. Estas representaciones permiten comparar el grado de estacionariedad alcanzado bajo cada transformación y facilitan la selección de la forma más adecuada de la serie para el posterior ajuste de los modelos SARIMA y de las arquitecturas de aprendizaje profundo empleadas en el estudio.

Figura 7. *Serie, ACF, PACF, transformaciones y diferenciaciones*

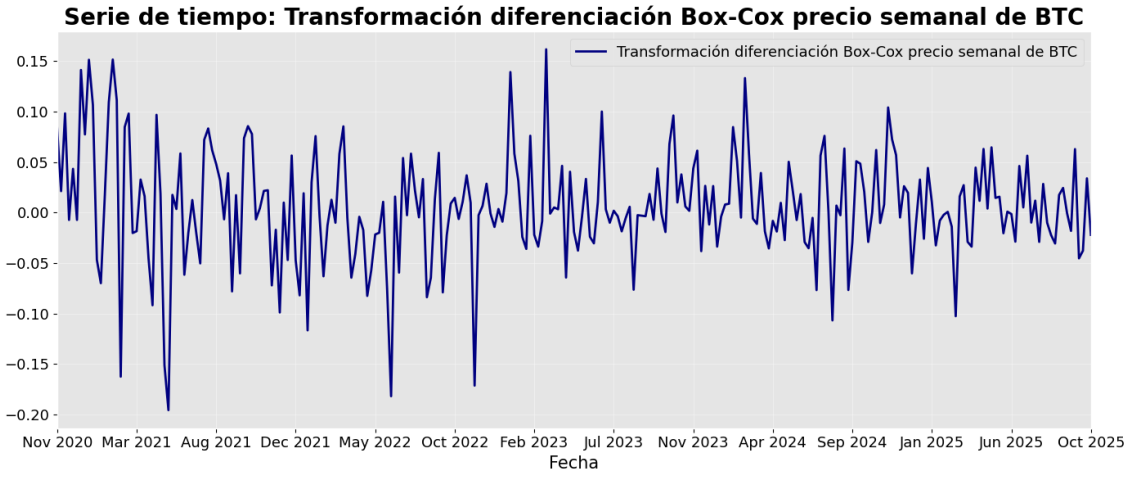


Como se observa en la Figura 7, la serie original del precio semanal del Bitcoin presenta una marcada tendencia creciente y episodios de alta volatilidad, lo que confirma su carácter no estacionario, evidenciado también por las funciones ACF y PACF, las cuales muestran persistencia significativa en múltiples rezagos. Al aplicar transformaciones funcionales como logaritmo, raíz cuadrada y Box-Cox, se logra una estabilización parcial de la varianza; sin embargo, la tendencia subyacente persiste, lo que indica que dichas transformaciones por sí solas no son suficientes para alcanzar la estacionariedad. En contraste, al aplicar la primera diferenciación tanto sobre la serie original como sobre las

series transformadas, se elimina la tendencia y se obtiene una serie que oscila alrededor de una media aproximadamente constante, con autocorrelaciones que decaen rápidamente hacia valores cercanos a cero. Entre las alternativas evaluadas, la combinación de transformación Box-Cox seguida de diferenciación muestra el comportamiento más cercano a un proceso estacionario, por lo que se selecciona como la base para el ajuste del procedimiento SARIMA y para el entrenamiento de las arquitecturas de aprendizaje profundo.

Con el fin de ilustrar con mayor claridad las propiedades estadísticas de la serie finalmente seleccionada, la Figura 8 presenta la serie resultante de aplicar la transformación Box-Cox seguida de la diferenciación. En esta figura se evidencia una dinámica más estable, sin tendencia sistemática y con fluctuaciones alrededor de una media constante, lo que confirma que el proceso resultante cumple de manera más adecuada los criterios de estacionariedad requeridos para el análisis econométrico y la modelación predictiva. Esta serie transformada constituye, por tanto, la entrada definitiva utilizada en los experimentos de ajuste y evaluación de los modelos desarrollados en el estudio.

Figura 8. *Transformación y diferenciación Box-Cox precio semanal del Bitcoin*

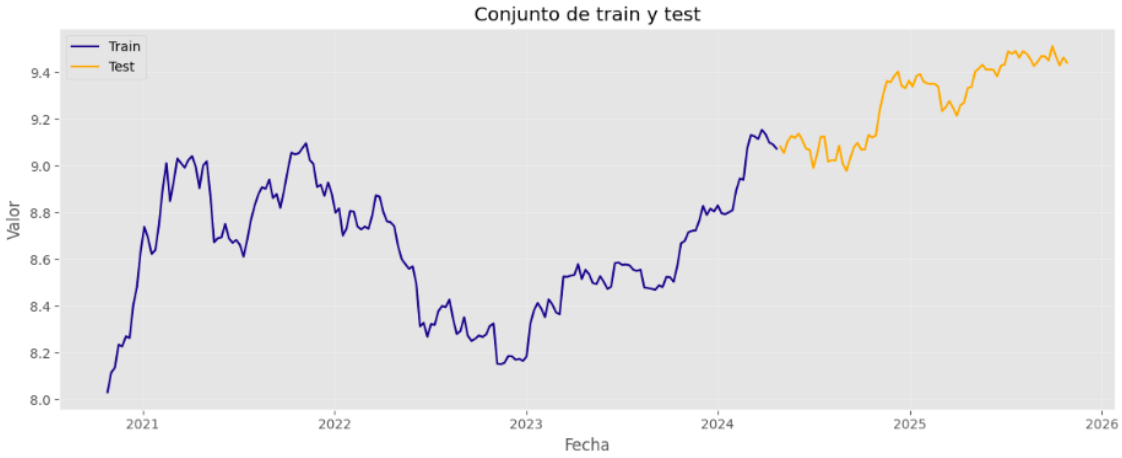


Esta combinación metodológica mostrada en la Figura 8, permite atenuar la heterocedasticidad inherente a la serie original y eliminar patrones deterministas de tendencia, favoreciendo una estructura de dependencia temporal más simple. No obstante, como se observa en la figura, antes de octubre de 2022 la serie transformada presenta una mayor amplitud en sus fluctuaciones, lo cual refleja la elevada volatilidad característica del mercado del Bitcoin durante ese período. A pesar de ello, la serie obtenida mediante la transformación Box-Cox y la diferenciación constituye la representación más cercana a la estacionariedad que fue posible alcanzar, manteniendo al mismo tiempo la información dinámica relevante del proceso. Asimismo, los patrones temporales resultan considerablemente menos complejos que en la serie original, lo que facilita la identificación de estructuras de dependencia y mejora la estabilidad del ajuste de los modelos considerados.

A continuación, la Figura 9 presenta la segmentación de la serie de tiempo en su escala original, correspondiente a la división en conjuntos de entrenamiento (70%) y prueba

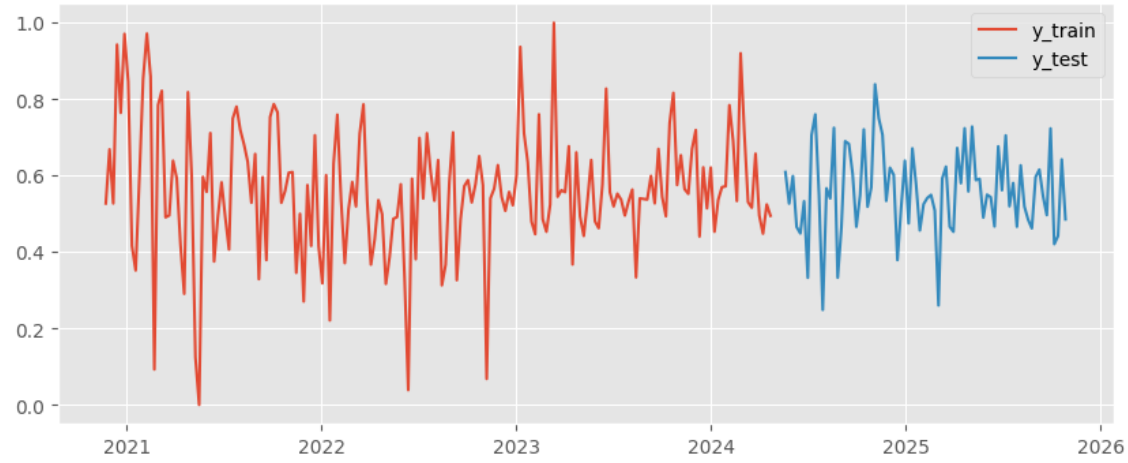
(30%), realizada respetando el orden temporal de las observaciones con el fin de simular un escenario real de predicción fuera de la muestra.

Figura 9. *Segmentación temporal del conjunto de datos para la predicción del precio del bitcoin mediante transformación Box-Cox*



Si bien esta partición presentada en la Figura 9 se efectúa sobre la serie sin transformar, el entrenamiento y prueba de los modelos se realiza utilizando la serie previamente transformada, garantizando así que ambos subconjuntos presenten propiedades estadísticas comparables y cumplan con los supuestos requeridos para el modelado econométrico y el entrenamiento de las arquitecturas de aprendizaje profundo. En este contexto, la Figura 10 presenta la segmentación en conjuntos de entrenamiento y prueba de la serie transformada, incluyendo la diferenciación de primer orden.

Figura 10. *Conjunto de train y test de la serie transformada incluyendo diferenciación de primer orden*



La Figura 10 de la serie transformada se observa que los patrones presentes en el conjunto de prueba son altamente consistentes con los del conjunto de entrenamiento, lo que facilita el reconocimiento de regularidades temporales por parte de los métodos de aprendizaje profundo. Asimismo, el rango de valores es comparable en ambos subconjuntos, reduciendo la presencia de cambios estructurales que puedan afectar la capacidad de generalización de los modelos. En contraste, en la Figura 5, donde únicamente se aplica la transformación Box-Cox, el conjunto de prueba presenta tanto

niveles como patrones distintos a los observados en el conjunto de entrenamiento, lo que introduce un mayor grado de dificultad para el aprendizaje de las dinámicas subyacentes por parte de las arquitecturas de aprendizaje profundo y puede deteriorar el desempeño predictivo.

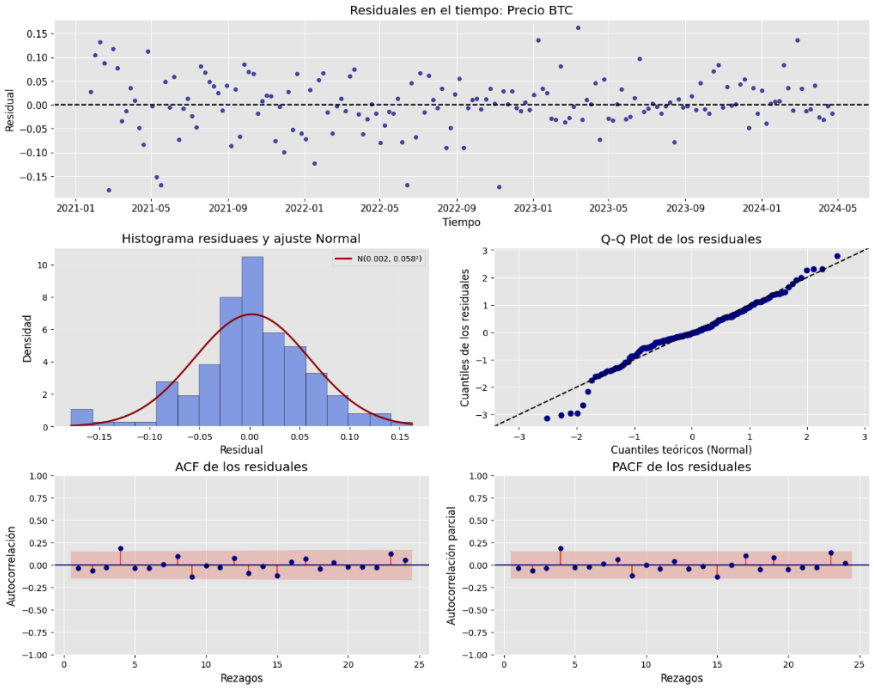
7.1 Resultados del procedimiento SARIMA

En la fase de identificación se evaluaron trece estructuras SARIMA, con el objetivo de seleccionar aquella que representara de forma más adecuada la dinámica temporal de la serie analizada. La comparación entre los modelos se realizó mediante el criterio de información de Akaike (AIC), el cual permite balancear la calidad del ajuste con la complejidad del modelo, favoreciendo estructuras parsimoniosas.

Entre las especificaciones consideradas, el procedimiento SARIMA(4, 1, 0)(0, 1, 1, 4) presentó el menor valor del AIC (-523.479), por lo que fue seleccionado como la estructura óptima para la modelación de la serie.

La Figura 11 presenta los principales diagnósticos gráficos del modelo estimado, orientados a evaluar la adecuación del ajuste y a verificar el cumplimiento de los supuestos fundamentales de la metodología Box-Jenkins, en particular la ausencia de autocorrelación en los residuales, la homocedasticidad y la normalidad aproximada de los errores.
En esta especificación, la dinámica del proceso queda explicada exclusivamente por un término de medias móviles, el cual captura la dependencia serial presente en los errores, mientras que el parámetro de diferenciación d=1 garantiza la estacionariedad en media de la serie original. El modelo seleccionado fue estimado utilizando el conjunto de entrenamiento y posteriormente evaluado sobre el conjunto de prueba, con el fin de analizar su capacidad de generalización y su desempeño predictivo fuera de la muestra.

Figura 11. *Análisis de residuales del procedimiento SARIMA*



A partir del análisis gráfico de los residuales del procedimiento SARIMA mostrado en la Figura 11, se observa que los errores se encuentran aproximadamente centrados en cero a lo largo del tiempo, sin evidencia de tendencias sistemáticas ni patrones estacionales persistentes, lo que sugiere que el modelo captura adecuadamente la dinámica media de la serie (Figura 6a). Sin embargo, la dispersión de los residuales no es completamente constante, observándose subperíodos con mayor variabilidad, lo cual sugiere la posible presencia de heterocedasticidad. El histograma de los residuales (Figura 6b), junto con el ajuste de la distribución normal, muestra una forma aproximadamente simétrica, aunque con desviaciones en las colas; este comportamiento se confirma en el gráfico Q-Q, donde los cuantiles extremos se apartan de la recta teórica, indicando colas más pesadas que las esperadas bajo el supuesto de normalidad (Figura 6c). En cuanto a la dependencia serial, las funciones ACF y PACF de los residuales indican ausencia de autocorrelación serial significativa de los rezagos analizados, lo que indica ausencia de autocorrelación remanente y, por tanto, una adecuada captura de la estructura temporal por parte del modelo seleccionado, aunque la dinámica de la volatilidad no es modelada explícitamente. (Figura 6d y 6e).

Este diagnóstico visual no evidencia de manera marcada cambios sistemáticos en la varianza a lo largo del tiempo; sin embargo, los contrastes formales de homocedasticidad indican que la varianza no es constante, ya que las pruebas de Breusch-Pagan y White arrojan valores *p* inferiores al nivel de significancia convencional ($\alpha = 0.05$), lo que sugiere la presencia de heterocedasticidad en los residuales del modelo obtenido mediante el procedimiento SARIMA, aunque de forma moderada desde el punto de vista gráfico. Adicionalmente, los residuales se encuentran centrados alrededor de cero y presentan una distribución leptocúrtica, con mayor concentración de valores cercanos a cero, ligera asimetría hacia el lado negativo y una cola izquierda más pesada, características típicas de series financieras.

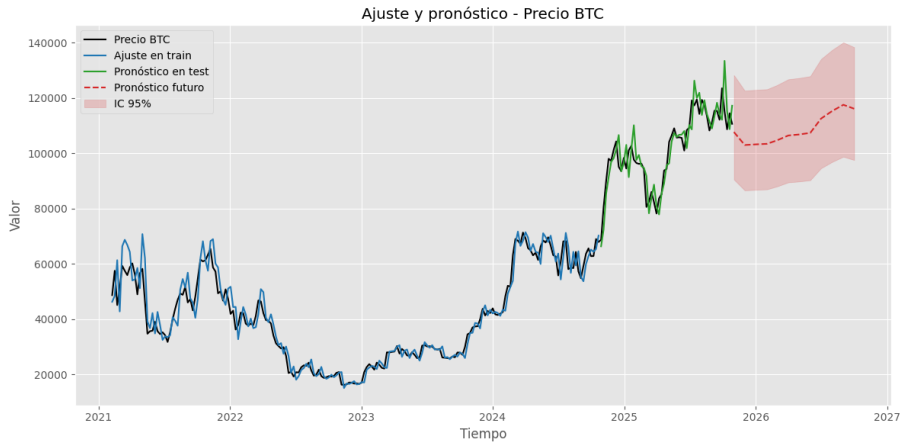
Adicionalmente, la validación de las series sintéticas generadas a partir de los residuales del conjunto de entrenamiento mediante la prueba de KS de dos muestras indica que ninguna de las trece simulaciones cumple el criterio de similitud estadística con la serie original ($p > 0.05$). En consecuencia, no se satisface el requisito mínimo para la aceptación de las trayectorias simuladas, por lo que se descarta la continuación del procedimiento de simulación Montecarlo. Este resultado evidencia que el modelo no reproduce adecuadamente la distribución empírica ni la variabilidad observada en los datos reales.

En conjunto, los resultados indican que el procedimiento SARIMA es adecuado para capturar la dependencia temporal y la dinámica promedio de la serie, pero presenta limitaciones para modelar la heterocedasticidad, los eventos extremos y las estructuras no lineales propias del mercado de las criptomonedas. Esta evidencia justifica la exploración de enfoques alternativos de modelación, en particular arquitecturas de aprendizaje profundo, que permiten capturar patrones más complejos y dinámicas de volatilidad variables, con el fin de mejorar la representación estadística y el desempeño predictivo del proceso analizado.

Desde una perspectiva de evaluación predictiva, resulta igualmente relevante examinar el comportamiento del modelo en términos de ajuste dentro de la muestra y capacidad de pronóstico fuera de la muestra. En este sentido, la Figura 12 presenta el ajuste y el pronóstico del procedimiento SARIMA$(4, 1, 0)(0, 1, 1, 4)$ en la escala original de la serie,

mostrando el comportamiento observado junto con los valores ajustados en el conjunto de entrenamiento y las proyecciones por fuera de la muestra para el conjunto de prueba, incluyendo sus respectivos intervalos de confianza al 95%. Esta representación permite evaluar visualmente la capacidad del modelo para seguir la trayectoria general de la serie, así como la incertidumbre asociada a las predicciones.

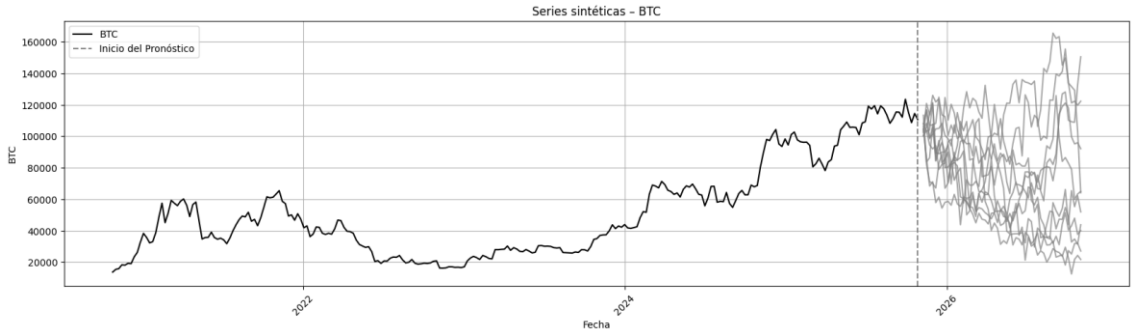
Figura 12. *Ajuste y pronóstico del precio del BTC*



Los resultados presentados en la Figura 12 evidencian un desempeño adecuado del modelo tanto dentro como fuera de la muestra. En el conjunto de entrenamiento, el modelo alcanza un RMSE de 4272.8600 y un R^2 de 0.9099, lo que indica una alta capacidad para reproducir la dinámica general de la serie observada. De manera consistente, en el conjunto de prueba se obtiene un RMSE de 5565. 5279 y un R^2 de 0.9286 lo que sugiere una adecuada capacidad de generalización y estabilidad predictiva, pese al incremento esperado del error al extrapolar fuera de la muestra. En conjunto, estos resultados indican que el SARIMA captura eficazmente la tendencia y las variaciones de mediano plazo del precio del Bitcoin, aunque, como se evidenció en el análisis de residuales, persisten limitaciones para representar completamente la dinámica de la volatilidad.

La Figura 13 presenta la serie histórica del precio de Bitcoin junto con las trayectorias sintéticas generadas mediante simulación Montecarlo a partir del modelo ajustado. A partir del punto de corte correspondiente al inicio del período de pronóstico, se muestran múltiples realizaciones futuras del proceso, las cuales permiten analizar no solo la evolución esperada de la serie, sino también la dispersión y la incertidumbre asociadas a los valores proyectados a diferentes horizontes temporales.

Figura 13. *Ejemplo 10 sintéticas SARIMA(4, 1, 0)(0, 1, 1, 4)*



En la Figura 13 se observa el comportamiento de las trayectorias sintéticas generadas mediante simulación Montecarlo a partir del modelo ajustado, superpuestas a la serie histórica transformada. Las trayectorias simuladas presentan una estructura de dispersión que se amplía progresivamente a medida que aumenta el horizonte de pronóstico, configurando un patrón visual similar a un embudo. Este comportamiento es consistente con la propagación de la incertidumbre inherente a los procesos de predicción recursiva, ya que los errores introducidos en cada paso se acumulan y se transmiten a los pronósticos posteriores.

En los primeros períodos de simulación, las trayectorias permanecen relativamente concentradas alrededor del valor esperado del modelo, reflejando una menor incertidumbre a corto plazo. Sin embargo, conforme se avanza en el tiempo, la variabilidad entre trayectorias aumenta, lo que se manifiesta en una mayor dispersión de los valores simulados. Este fenómeno es característico de los modelos estocásticos cuando se emplean simulaciones Montecarlo y constituye un resultado esperado desde el punto de vista teórico, dado que cada trayectoria incorpora una secuencia distinta de residuales aleatorios. Adicionalmente, las trayectorias sintéticas preservan el nivel y la tendencia general del proceso modelado, lo que indica que la dinámica media de la serie ha sido adecuadamente capturada por el modelo. No obstante, la creciente dispersión evidencia que la incertidumbre asociada al pronóstico se incrementa con el horizonte temporal, lo cual es particularmente relevante en series financieras como el precio de Bitcoin, donde la volatilidad puede amplificarse en el largo plazo.

La Figura 14 presenta un resumen mensual de la distribución de las 5.000 series sintéticas generadas mediante simulación Montecarlo, representado a través de diagramas de caja y bigotes (boxplots) para cada período del horizonte de pronóstico. Esta representación permite analizar la evolución temporal de la mediana, la dispersión y la presencia de valores extremos en los escenarios simulados, complementando el análisis visual de las trayectorias individuales mostrado en la figura anterior.

Figura 14. *Boxplot mensual 5000 series sintéticas*

El comportamiento de la variabilidad mostrado en la Figura 14 evidencia que, a medida que avanza el horizonte de pronóstico, la dispersión de los valores simulados aumenta progresivamente, lo cual se refleja en el ensanchamiento de las cajas y la extensión de los bigotes. Este comportamiento es consistente con el patrón en forma de embudo observado en las trayectorias sintéticas individuales, confirmando que la incertidumbre se amplifica conforme se acumulan los errores en el proceso de simulación recursiva. La mediana de las distribuciones se mantiene relativamente estable y cercana a la trayectoria esperada del modelo, lo que indica que la dinámica media es preservada en las simulaciones.

Asimismo, se observa una asimetría en algunas distribuciones mensuales y la presencia de valores extremos, particularmente en los períodos más alejados del punto de origen del pronóstico, lo que refleja la influencia de residuales con colas pesadas, característica típica de series financieras. En conjunto, este análisis confirma que las simulaciones Montecarlo no solo reproducen la tendencia central estimada por el modelo, sino que también capturan la variabilidad y el riesgo asociados al comportamiento futuro del precio de Bitcoin, proporcionando un marco adecuado para el análisis de escenarios y la evaluación de la incertidumbre en los pronósticos.

7.2 Resultados de los modelos de Deep Learning

En esta sección se consideran dos enfoques diferenciados para el ajuste de los modelos de Deep Learning. El primero sigue el procedimiento comúnmente reportado en la literatura, en el cual los modelos se entrenan directamente sobre la serie original, sin aplicar transformaciones previas ni verificar formalmente la condición de estacionariedad. Este enfoque prioriza la capacidad de los modelos de aprendizaje profundo para capturar patrones complejos de manera flexible, independientemente de los supuestos clásicos de las series de tiempo.

El segundo enfoque se fundamenta en las recomendaciones de la metodología Box-Jenkins y consiste en ajustar los modelos de Deep Learning sobre series previamente transformadas para garantizar la estacionariedad. En este caso, se aplican transformaciones orientadas a estabilizar la varianza y eliminar componentes deterministas como la tendencia, con el fin de modelar un proceso temporal más estable. Este procedimiento busca evaluar si la incorporación de principios econométricos clásicos en el preprocesamiento de los datos contribuye a mejorar el desempeño predictivo y la robustez estadística de los modelos de Deep Learning.

7.2.1 Resultados de los modelos de Deep Learning serie de tiempo sin transformar

En esta sección se presenta el ajuste de los modelos de aprendizaje profundo utilizando la serie de tiempo en su escala original, sin aplicar transformaciones estadísticas. Este enfoque se adopta con el propósito de establecer una línea base de desempeño que permita comparar posteriormente el efecto de las transformaciones aplicadas a la serie, y así evaluar su impacto en la capacidad predictiva de las distintas arquitecturas. De esta manera, se busca analizar en qué medida el preprocesamiento de la serie contribuye a mejorar el aprendizaje de patrones temporales y la generalización de los modelos de aprendizaje profundo.

Tabla 8. *Hiperparámetros del mejor modelo*

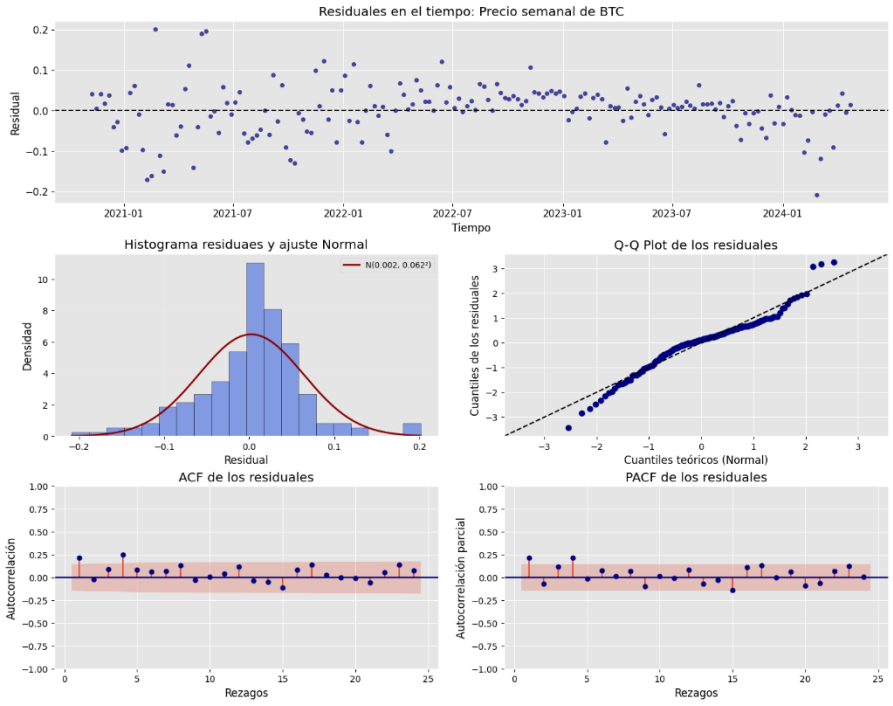
Hiperparámetro	Valores configurables
LSTM	
Capas ocultas	2
Neuronas por capa	25
Funciones de activación	Elu
Optimizadores	Adam
Número de rezagos (lags)	1
Tamaño de lote (batch size)	6

El modelo seleccionado corresponde a una arquitectura LSTM configurada con un único rezago como entrada, dos capas ocultas y 25 neuronas en cada capa, lo que permite

capturar dependencias temporales de corto y mediano plazo sin introducir una complejidad excesiva. Esta estructura favorece la parsimonia del modelo y contribuye a reducir el riesgo de sobreajuste, manteniendo al mismo tiempo una adecuada capacidad de representación de la dinámica temporal de la serie. La función de activación ELU se emplea debido a que permite salidas negativas, lo que favorece que las activaciones estén centradas alrededor de cero, mejorando la estabilidad del proceso de entrenamiento y acelerando la convergencia del algoritmo de optimización. El optimizador Adam se utiliza por su eficiencia y por su capacidad de adaptar dinámicamente la tasa de aprendizaje durante el entrenamiento, lo que contribuye a una convergencia más estable en presencia de ruido y no linealidades propias de series financieras. El entrenamiento se realizó utilizando un tamaño de lote de seis observaciones, lo que implica actualizaciones frecuentes de los pesos del modelo con un número reducido de muestras; este esquema introduce mayor variabilidad en las estimaciones del gradiente, actuando como un mecanismo de regularización implícita que puede favorecer la capacidad de generalización. Si bien el uso de un solo rezago podría limitar la captura de dependencias de largo plazo, esta restricción se ve parcialmente compensada por la memoria interna de la arquitectura LSTM. En conjunto, esta configuración de hiperparámetros explica el desempeño superior del modelo frente a las demás especificaciones evaluadas.

La Figura 15 muestra el diagnóstico de los residuales obtenido a partir del procedimiento aplicado a la serie de tiempo sin transformar, incorporando el análisis de la evolución temporal de los errores, el ajuste a la distribución normal y el análisis de autocorrelación. Este conjunto de herramientas gráficas permite evaluar si los residuales presentan un comportamiento compatible con un proceso de ruido blanco, requisito fundamental para considerar que la dinámica sistemática de la serie ha sido adecuadamente capturada por el procedimiento de modelado.

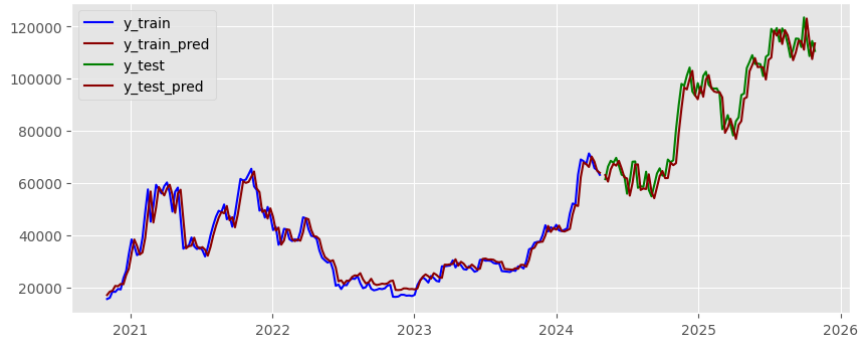
Figura 15. *Análisis de residuales de la serie de tiempo sin transformar*



En el gráfico de residuales en el tiempo (Figura 15a) se observa que los errores se encuentran centrados alrededor de cero, sin una tendencia sistemática evidente, lo que sugiere que no existe sesgo en los pronósticos. Sin embargo, se aprecia cierta variabilidad irregular a lo largo del periodo, con episodios de mayor dispersión, lo cual puede indicar la presencia de cambios en la volatilidad. El histograma de los residuales (Figura 15b) muestra una distribución aproximadamente simétrica alrededor de cero, aunque con una ligera leptocurtosis, reflejando una mayor concentración de valores cercanos al centro y colas algo más pesadas que las de una distribución normal. Esta característica también se evidencia en el gráfico Q–Q (Figura 15c), donde los cuantiles centrales se alinean razonablemente con la recta teórica, pero se observan desviaciones en las colas, especialmente en los cuantiles extremos, lo que sugiere que la normalidad no se cumple estrictamente. Por su parte, las funciones ACF y PACF (Figura 15d y 15e) de los residuales no presentan picos significativos fuera de los intervalos de confianza, lo que indica ausencia de autocorrelación serial remanente. En conjunto, estos resultados sugieren que el procedimiento captura adecuadamente la dependencia temporal de la serie, aunque persisten indicios de heterocedasticidad y desviaciones de la normalidad.

La Figura 16 presenta el ajuste del modelo sobre la serie de tiempo sin transformación, mostrando simultáneamente los valores observados y los valores estimados tanto en el conjunto de entrenamiento como en el de prueba. Esta representación permite evaluar visualmente la capacidad del modelo para reproducir la dinámica histórica de la serie y su desempeño fuera de la muestra.

Figura 16. *Ajuste de la serie sin transformación*



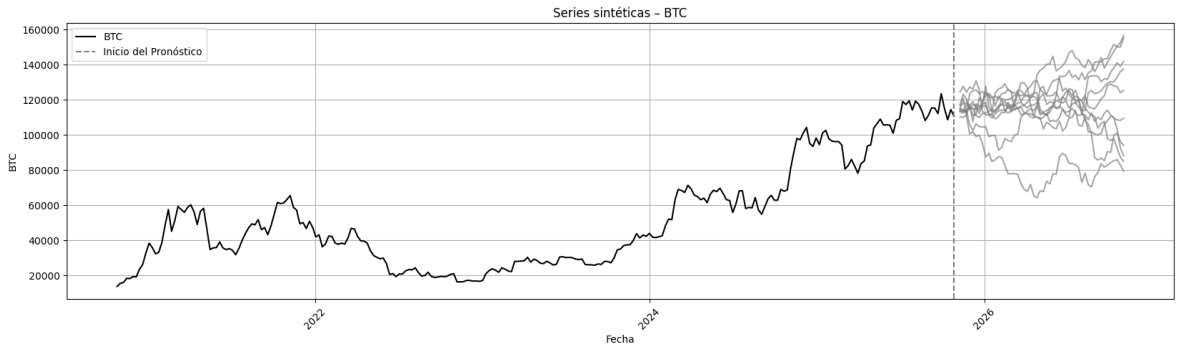
Como se muestra en la Figura 16, en el período de entrenamiento, el modelo logra seguir de manera razonable la tendencia general y las principales oscilaciones de la serie, lo que indica una adecuada capacidad de ajuste dentro de la muestra. Sin embargo, en el conjunto de prueba se observa un sesgo sistemático hacia valores inferiores, es decir, las predicciones tienden a situarse por debajo de los valores reales durante gran parte del horizonte de evaluación. Este comportamiento sugiere una subestimación persistente del nivel de la serie. Adicionalmente, alrededor del año 2023 se evidencia un desfase notable, donde el modelo ajusta por encima de la serie observada, constituyendo el mayor error puntual del período de prueba. Este cambio de signo en el sesgo indica dificultades para adaptarse oportunamente a transiciones en la dinámica de la serie, particularmente en fases de recuperación o aceleración del crecimiento. En conjunto, estos resultados evidencian que, aunque el modelo captura parcialmente la tendencia global, la ausencia de transformaciones previas limita su capacidad para representar adecuadamente los cambios estructurales y la variabilidad de la serie en el período fuera de la muestra.

De forma complementaria al análisis visual presentado en la Figura 16, se evaluó el desempeño del modelo mediante métricas cuantitativas tanto dentro como fuera de la muestra. El RMSE obtenido en el conjunto de entrenamiento fue de 3532.69, mientras que en el conjunto de prueba aumentó a 5418.62, lo que indica una pérdida de precisión en el pronóstico fuera de la muestra, consistente con el sesgo observado en el período de test. No obstante, los coeficientes de determinación fueron elevados en ambos conjuntos ($R^2 = 0.938$ y $R^2 = 0.932$) train y test respectivamente, lo que sugiere que el modelo explica una proporción considerable de la variabilidad total de la serie, aunque con errores relevantes en magnitud.

Adicionalmente, las pruebas de BP y White indican la presencia de heterocedasticidad en los residuales, lo que implica que la varianza de los errores no es constante a lo largo del tiempo. Este resultado es consistente con la naturaleza altamente volátil del mercado de criptomonedas y sugiere que, aunque el modelo captura adecuadamente la dinámica media de la serie, no logra representar de forma satisfactoria la estructura de la volatilidad. En conjunto, estos resultados refuerzan la evidencia de que el modelado directo sobre la serie sin transformar presenta limitaciones, lo que motiva la evaluación posterior de modelos entrenados sobre series transformadas con el fin de mejorar la estabilidad estadística y el desempeño predictivo.

La Figura 17 presenta un ejemplo de diez series sintéticas generadas a partir del modelo LSTM ajustado sobre la serie de tiempo sin transformación. Las trayectorias simuladas se construyen mediante un proceso recursivo, en el cual cada valor futuro se obtiene a partir del pronóstico del modelo al que se le adiciona un residual muestreado de forma aleatoria, permitiendo incorporar variabilidad estocástica alrededor de la trayectoria estimada.

Figura 17. *Ejemplo de 10 series sintéticas del modelo LSTM sin transformación*



En el gráfico de la Figura 17 se observa que las simulaciones conservan la tendencia y el nivel de variabilidad observados en la serie original, evidenciando coherencia dinámica y estabilidad en el proceso de generación. Asimismo, se aprecia que varias trayectorias tienden a mantenerse por debajo del nivel observado al final del período histórico, lo que es consistente con el sesgo de subestimación previamente identificado en el conjunto de prueba para el modelo sin transformación. Esta combinación de sesgo en el nivel y aumento de la dispersión sugiere que, si bien el LSTM reproduce la tendencia global, presenta limitaciones para representar adecuadamente la variabilidad y los cambios en la dinámica de la serie cuando se trabaja directamente con los datos sin transformar, lo cual refuerza la necesidad de aplicar transformaciones previas para estabilizar la varianza y mejorar la capacidad predictiva del modelo.

La representación de la Figura 18 permite evaluar la dispersión, la mediana y la presencia de valores extremos en cada periodo, proporcionando una visión comparativa de la evolución probabilística del comportamiento simulado.

Figura 18. *Boxplot mensual 5000 series sintéticas*

La Figura 18 muestra mediante los diagramas de caja que se evidencia una dispersión moderada y relativamente estable entre meses, mostrando coherencia interna en la variabilidad de las trayectorias simuladas. La mediana de cada boxplot se mantiene cercana al nivel central de las proyecciones, indicando ausencia de sesgos sistemáticos en la simulación. Asimismo, los rangos intercuartílicos reflejan una amplitud consistente, lo cual es indicativo de una adecuada representación de la incertidumbre inherente al proceso estocástico. La presencia de valores extremos en algunos meses sugiere escenarios de alta volatilidad, coherentes con la dinámica observada históricamente en el mercado de las criptomonedas. En conjunto, la figura valida que las series sintéticas preservan la estructura de variabilidad y dispersión del activo, aportando soporte empírico a la robustez del procedimiento de simulación y a su utilidad para el análisis de riesgo y escenarios futuros.

7.2.2 Resultados de los modelos de Deep Learning serie de tiempo estacionaria

En esta sección se aborda el ajuste de los modelos de Deep Learning empleando series de tiempo previamente transformadas para cumplir la condición de estacionariedad, en concordancia con los principios de la metodología Box-Jenkins. La Tabla 8 presenta los hiperparámetros correspondientes al mejor modelo de Deep Learning identificado en el análisis, los cuales definen su arquitectura y el esquema de entrenamiento utilizado. La selección de estos hiperparámetros responde a un criterio de desempeño predictivo y capacidad de generalización.

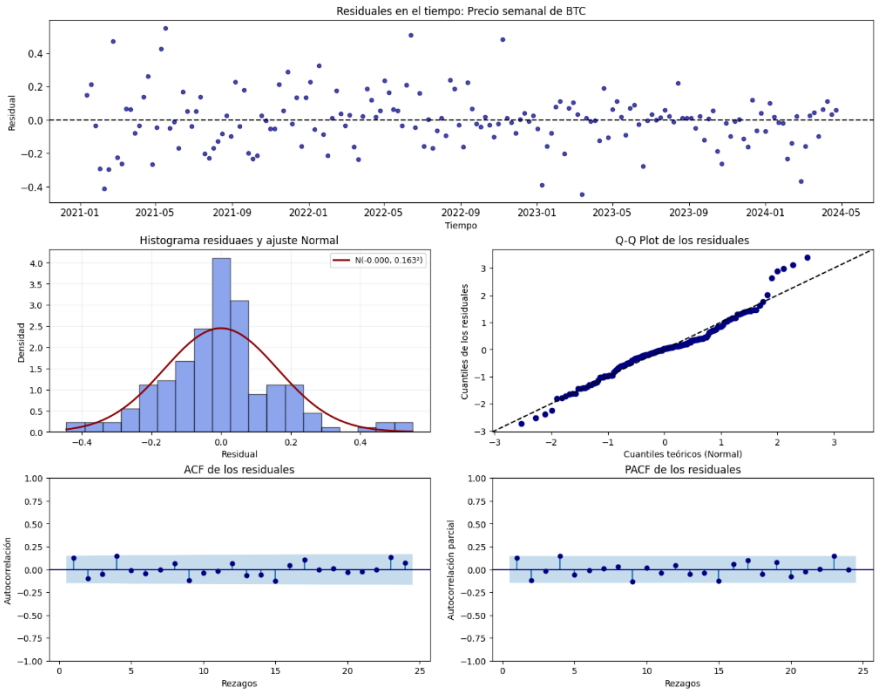
Tabla 9. *Hiperparámetros del mejor modelo*

Hiperparámetro	Valores configurables
LSTM	
Capas ocultas	2
Neuronas por capa	5
Funciones de activación	tanh
Optimizadores	Adam
Número de rezagos (lags)	9
Tamaño del batch	32

El modelo seleccionado se basa en una arquitectura LSTM con dos capas ocultas, lo que permite capturar dependencias temporales de corto y mediano plazo sin introducir una complejidad excesiva. Cada capa contiene cinco neuronas, configuración que favorece la parsimonia del modelo y reduce el riesgo de sobreajuste. La función de activación *tanh* se emplea por su idoneidad en redes recurrentes, al facilitar la modelación de relaciones no lineales y estabilizar el flujo de gradientes durante el entrenamiento. El optimizador Adam se utiliza por su eficiencia y capacidad de adaptación dinámica de la tasa de aprendizaje, contribuyendo a una convergencia más estable. Adicionalmente, el uso de nueve rezagos permite incorporar información temporal relevante del pasado reciente de la serie, mientras que un tamaño de batch de 32 observaciones equilibra la estabilidad del proceso de entrenamiento con la eficiencia computacional. En conjunto, esta configuración de hiperparámetros explica el desempeño superior del modelo frente a las demás especificaciones evaluadas.

La Figura 19 complementa la descripción del modelo seleccionado al presentar el análisis de los residuales asociados a la serie de tiempo transformada, lo que permite evaluar visualmente la idoneidad de la configuración adoptada. Este análisis constituye un paso fundamental para verificar la idoneidad del preprocesamiento realizado y para comprobar que la estructura estocástica remanente es consistente con los supuestos requeridos para la modelación y la inferencia estadística.

Figura 19. *Análisis de residuales de la serie de tiempo transformada*



En coherencia con la arquitectura LSTM parsimoniosa y el esquema de entrenamiento empleado, los diagnósticos gráficos presentados en la Figura 19 permiten verificar que la transformación aplicada logró estabilizar la serie y que los errores resultantes no presentan patrones sistemáticos remanentes. El análisis de los residuales en el tiempo (Figura 9a) evidencia un comportamiento aleatorio alrededor de cero, sin tendencias determinísticas, cambios estructurales visibles ni patrones persistentes, lo que sugiere que

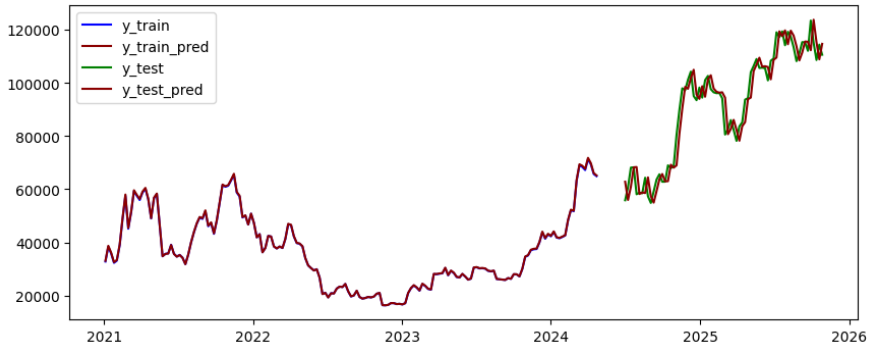
el modelo ha capturado adecuadamente la dinámica subyacente de la serie, dejando errores que pueden interpretarse como ruido blanco. Asimismo, no se observan conglomerados ni bandas de variabilidad creciente o decreciente, lo que respalda visualmente el supuesto de varianza constante.

El histograma de los residuales y el gráfico Q–Q (Figuras 19b y 19c) muestran una distribución aproximadamente simétrica y cercana a la normalidad, aunque con ligeras desviaciones en las colas, comportamiento habitual en series financieras caracterizadas por eventos extremos. En cuanto a la dependencia serial, las funciones de autocorrelación (ACF) y autocorrelación parcial (PACF) de los residuales (Figuras 19d y 19e) no presentan coeficientes significativamente distintos de cero en los rezagos analizados, lo que indica ausencia de autocorrelación remanente y confirma que la estructura temporal ha sido adecuadamente modelada.

Finalmente, los resultados de las pruebas formales de Breusch-Pagan y White confirman la homocedasticidad de los residuales, al no rechazarse la hipótesis nula de varianza constante. Este resultado es consistente con la evidencia gráfica observada en la Figura 19 y fortalece la robustez estadística del modelo, al indicar que los errores no presentan heterogeneidad sistemática a lo largo del tiempo. En conjunto, el análisis residual respalda que el modelo LSTM no solo presenta un buen desempeño predictivo, sino que también cumple con supuestos estadísticos fundamentales, lo que incrementa la fiabilidad de las inferencias y proyecciones derivadas del mismo.

En este contexto, la Figura 20 presenta el ajuste de la serie de tiempo transformada correspondiente al mejor modelo de Deep Learning seleccionado, permitiendo evaluar directamente su capacidad para reproducir la trayectoria observada y su desempeño tanto dentro como fuera de la muestra. La proximidad entre los valores observados y los valores ajustados en el conjunto de entrenamiento, así como la adecuada continuidad de las proyecciones en el conjunto de prueba, sugieren una buena capacidad de generalización. Este resultado es consistente con las métricas de desempeño reportadas y con el análisis de residuales, y refuerza la conclusión de que el modelo LSTM ofrece una representación más flexible y robusta de la dinámica del precio del Bitcoin frente a los enfoques lineales tradicionales.

Figura 20. *Ajuste de la serie transformada*



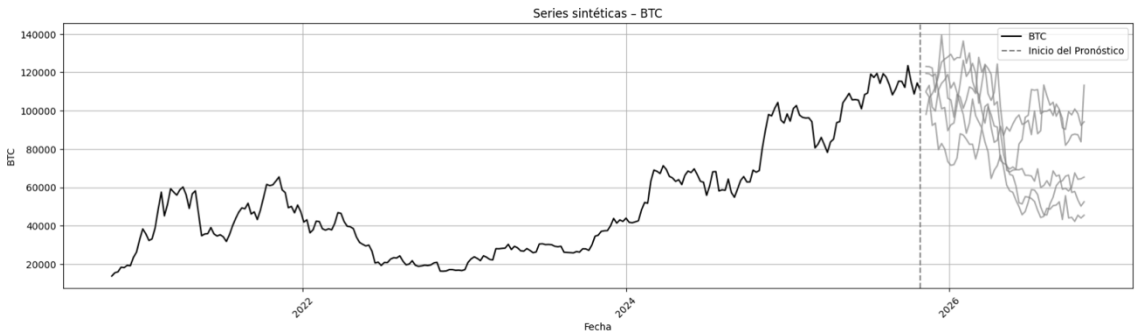
Los resultados asociados a la Figura 20 evidencian un ajuste muy preciso del modelo en el conjunto de entrenamiento, reflejado en un RMSE de 173.36 y un coeficiente de determinación R^2 de 0.9998, lo que indica que el modelo logra reproducir casi en su totalidad la variabilidad observada durante la fase de aprendizaje. En el conjunto de

prueba, aunque se registra un incremento en el error (RMSE = 5323.25), el valor de R^2 de 0.9307 confirma que el modelo mantiene una elevada capacidad explicativa fuera de la muestra, sugiriendo una adecuada generalización y ausencia de sobreajuste severo.

La diferencia entre los errores de entrenamiento y prueba es consistente con la mayor incertidumbre inherente al pronóstico fuera de la muestra en series financieras altamente volátiles; sin embargo, el nivel de ajuste alcanzado indica que el modelo conserva la capacidad de capturar la dinámica no lineal dominante de la serie. En conjunto, estos resultados respaldan la idoneidad del enfoque LSTM para modelar y predecir el comportamiento del precio del Bitcoin, superando las limitaciones observadas en el procedimiento SARIMA en términos de estabilidad residual y capacidad para representar estructuras complejas.

Con el fin de evaluar la capacidad del modelo LSTM no solo para realizar pronósticos puntuales, sino también para reproducir la dinámica estocástica de la serie original, se procedió a la generación de trayectorias sintéticas a partir del modelo entrenado. En este contexto, la Figura 21 presenta un ejemplo de cinco series sintéticas generadas por el modelo LSTM, las cuales permiten analizar visualmente si el modelo es capaz de preservar características fundamentales del proceso, tales como la volatilidad, la estructura temporal y la amplitud de las fluctuaciones observadas en la serie real.

Figura 21. *Ejemplo de 5 series sintéticas del modelo LSTM*



En el ejemplo de la figura 21 se observa una dinámica temporal similar a la de la serie original, caracterizada por una tendencia comparable y niveles de volatilidad coherentes con el comportamiento histórico del precio.

Con el objetivo de evaluar de manera agregada la distribución y la variabilidad temporal de las trayectorias simuladas, se generaron múltiples series sintéticas a partir del modelo LSTM y se analizaron sus propiedades estadísticas por período. En este contexto, la Figura 22 presenta el diagrama de cajas mensual correspondiente a 5000 series sintéticas, lo que permite comparar la dispersión, asimetría y presencia de valores extremos a lo largo del horizonte temporal, así como contrastar si la estructura de volatilidad estimada por el modelo es consistente entre diferentes meses.

Figura 22. *Boxplot mensual 5000 series sintéticas*

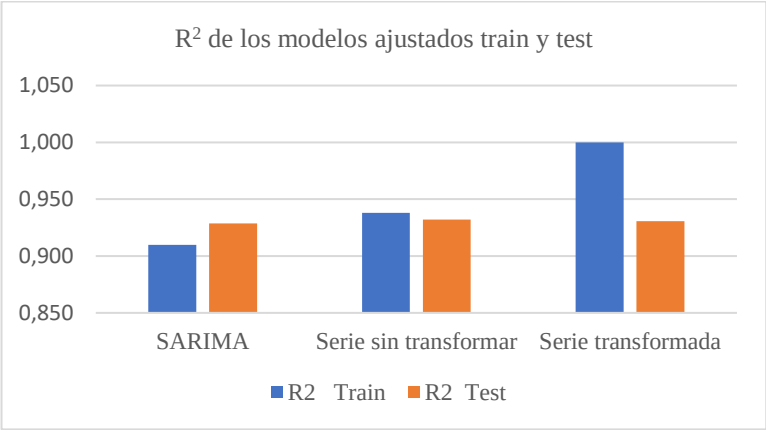
El diagrama de cajas presentado en la Figura 22, construido a partir de 5000 series sintéticas y agrupado en intervalos de cuatro semanas, permite analizar la evolución de la distribución de los precios simulados a lo largo del horizonte de proyección. En los primeros períodos se observa una tendencia bajista en la mediana de las distribuciones, seguida de una fase de estabilización alrededor del nivel de 80000 USD, lo que sugiere que el modelo converge hacia una región de equilibrio tras las fluctuaciones iniciales. A medida que avanza el horizonte temporal, se evidencia un incremento progresivo en la dispersión, reflejado en el ensanchamiento de las cajas y en la mayor extensión de los bigotes, lo cual indica un aumento en la incertidumbre y variabilidad de las trayectorias simuladas. No obstante, los valores extremos permanecen acotados, con máximos que no superan aproximadamente los 150000 USD, mostrando que, si bien el modelo incorpora mayor volatilidad en el largo plazo, no proyecta escenarios de crecimiento explosivo. En conjunto, estos resultados indican que el modelo LSTM es capaz de reproducir una dinámica realista de estabilización del nivel medio del precio junto con un aumento gradual de la incertidumbre, comportamiento coherente con la naturaleza altamente volátil pero acotada del mercado de criptomonedas en el horizonte analizado.

8. Discusión

En esta sección se discuten de manera integrada los resultados obtenidos a partir de los diferentes enfoques de modelado considerados, incluyendo el procedimiento SARIMA y los modelos de aprendizaje profundo entrenados tanto sobre la serie sin transformar como sobre la serie transformada. El objetivo de esta comparación es evaluar no solo el desempeño predictivo medido mediante métricas tradicionales, sino también la capacidad de cada enfoque para representar adecuadamente la dinámica estadística de la serie, así como su utilidad para la generación de escenarios sintéticos mediante simulación Montecarlo, aspecto central de esta investigación.

La Figura 23 presenta una comparación del desempeño predictivo de los modelos evaluados, medida a través del R^2 tanto en el conjunto de entrenamiento como en el conjunto de prueba, para el procedimiento SARIMA, el modelo LSTM ajustado sobre la serie sin transformar y el modelo LSTM ajustado sobre la serie transformada.

Figura 23. *R^2 de los modelos ajustados train y test*



En conjunto, los resultados presentados en la Figura 23 muestran que el procedimiento SARIMA alcanza un desempeño competitivo, con valores R^2 de 0.9099 en entrenamiento y 0.9286 en prueba, lo que indica una adecuada capacidad para capturar la dinámica media de la serie. No obstante, el análisis de residuales evidenció la presencia de heterocedasticidad, lo cual limita su idoneidad para aplicaciones que requieren una correcta representación de la volatilidad, como la generación de trayectorias sintéticas realistas.

Por su parte, el modelo LSTM entrenado sobre la serie sin transformar presenta un mejor ajuste dentro de la muestra con un R^2 de 0.9380 y un desempeño ligeramente superior en prueba con un R^2 de 0.9320, lo que sugiere una mayor capacidad para modelar relaciones no lineales. Sin embargo, los análisis gráficos y la simulación de series sintéticas evidenciaron sesgos sistemáticos en el nivel de las predicciones, así como una dispersión creciente de las trayectorias simuladas, lo que indica dificultades para representar de forma estable la variabilidad del proceso cuando la serie no ha sido previamente estabilizada.

En contraste, el modelo LSTM ajustado sobre la serie transformada exhibe un ajuste casi perfecto en el conjunto de entrenamiento con un R^2 de 0.9998, lo que refleja una elevada capacidad de representación de la dinámica interna del proceso bajo condiciones de varianza estabilizada. Aunque el valor del R^2 en prueba fue de 0.9307, comparable con el de los otros enfoques.

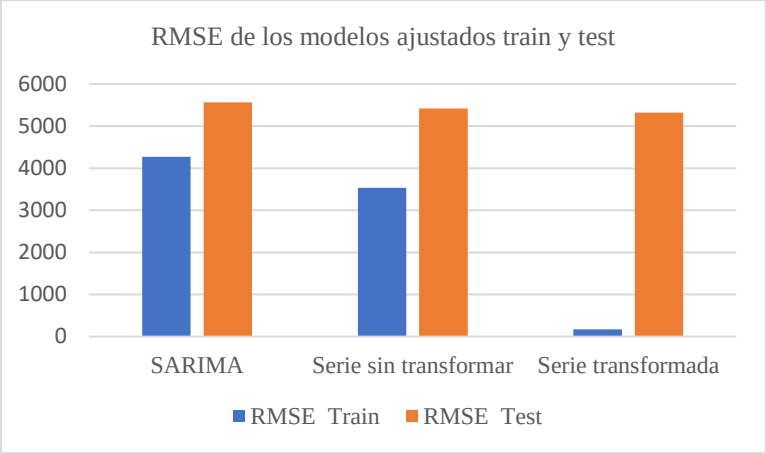
El análisis del R^2 evidencia que todos los enfoques considerados alcanzan niveles elevados de explicación de la variabilidad de la serie, con valores superiores al 90% tanto en entrenamiento como en prueba, lo que confirma la alta capacidad predictiva de los modelos evaluados. Sin embargo, las diferencias observadas entre el ajuste dentro de la muestra y el desempeño por fuera de ella, particularmente en el modelo LSTM sobre la serie transformada, ponen de manifiesto que un valor extremadamente alto de R^2 en entrenamiento no garantiza necesariamente una mejora proporcional en la capacidad de generalización. En este sentido, los resultados refuerzan la importancia de interpretar el R^2 como una medida complementaria y no exclusiva del desempeño del modelo, especialmente en contextos financieros caracterizados por alta volatilidad y posibles cambios estructurales. Así, aunque el modelo LSTM con serie transformada presenta un ajuste casi perfecto en entrenamiento, su desempeño en prueba resulta comparable al de los demás enfoques, lo que sugiere que el principal beneficio de la transformación no

radica únicamente en mejorar métricas globales de ajuste, sino en estabilizar la estructura estadística de la serie y favorecer una simulación más realista de escenarios futuros.

Con el fin de complementar el análisis basado en el R^2, a continuación se evalúa el desempeño de los modelos mediante el RMSE, métrica que cuantifica la magnitud promedio de los errores de predicción en las mismas unidades de la variable original. A diferencia del R^2, el RMSE permite una interpretación directa del error cometido por cada modelo, lo que resulta especialmente relevante en el contexto en análisis, donde desviaciones absolutas pueden tener implicaciones económicas significativas.

La Figura 24 presenta la comparación de los resultados del RMSE en los conjuntos de entrenamiento y prueba para el procedimiento SARIMA, el modelo LSTM ajustado sobre la serie sin transformar y el modelo LSTM ajustado sobre la serie transformada, permitiendo evaluar simultáneamente la capacidad de ajuste y la generalización de cada enfoque.

Figura 24. *RMSE de los modelos ajustados train y test*



Los resultados de la Figura 24 muestran que el SARIMA presenta valores de RMSE de 4272.86 en el conjunto de entrenamiento y de 5565.53 en el conjunto de prueba, los más altos entre los enfoques evaluados. Esto indica que, aunque el modelo logra capturar adecuadamente la estructura media y la dependencia temporal de la serie, las desviaciones absolutas entre los valores observados y los estimados son relativamente elevadas. Este resultado es consistente con las limitaciones de los modelos lineales para representar dinámicas complejas y cambios abruptos.

Por su parte, el modelo LSTM ajustado sobre la serie sin transformar presenta una mejora en el RMSE de entrenamiento 3532.69 respecto al procedimiento SARIMA, lo que refleja una mayor capacidad de ajuste dentro de la muestra, atribuible a la habilidad de las redes neuronales para capturar relaciones no lineales. No obstante, el RMSE en el conjunto de prueba fue de 5418.62 permaneciendo en un rango similar al del procedimiento SARIMA, lo que sugiere que, aunque el modelo aprende mejor la dinámica histórica, su capacidad de generalización sigue siendo limitada cuando la serie no ha sido previamente estabilizada mediante transformaciones estadísticas.

Finalmente, el modelo LSTM entrenado sobre la serie transformada exhibe un RMSE extremadamente bajo en el conjunto de entrenamiento 173.36, lo que evidencia un ajuste

casi perfecto de la dinámica interna de la serie bajo condiciones de varianza estabilizada. Sin embargo, en el conjunto de prueba el RMSE aumenta a 5323.25, ubicándose en niveles comparables a los otros enfoques. Este comportamiento indica que la transformación mejora sustancialmente la capacidad de representación dentro de la muestra, pero no se traduce en una reducción proporcional del error fuera de la muestra, lo cual puede estar asociado a la alta volatilidad del proceso y a la presencia de cambios estructurales difíciles de anticipar incluso con modelos no lineales.

En conjunto, el análisis del RMSE confirma que las transformaciones previas de la serie son determinantes para facilitar el proceso de aprendizaje y lograr un ajuste interno altamente preciso en los modelos de aprendizaje profundo. Sin embargo, las diferencias relativamente pequeñas en el error fuera de la muestra entre los distintos enfoques evidencian que la predicción puntual en mercados financieros continúa siendo un problema inherentemente complejo. En este sentido, el principal aporte de la transformación no radica únicamente en la reducción de métricas de error, sino en la mejora de la estabilidad estadística y de la calidad de las trayectorias simuladas, aspectos fundamentales para la generación de escenarios sintéticos y el análisis de incertidumbre, objetivos centrales de esta investigación.

Si bien las métricas de desempeño como el R^2 y el RMSE permiten evaluar la capacidad predictiva global de los modelos, estas no son suficientes para determinar si la estructura estadística de la serie ha sido adecuadamente representada. En particular, valores elevados de ajuste pueden coexistir con deficiencias en la modelación de la varianza, la distribución de los errores o la dependencia temporal remanente. Por esta razón, resulta necesario complementar la evaluación mediante un análisis detallado de los residuales, el cual permite verificar si los errores se comportan como un proceso aleatorio, condición fundamental para la validez del modelo y, en el contexto de esta investigación, para la generación confiable de series sintéticas mediante simulación Montecarlo.

El análisis de los residuales evidencia diferencias sustanciales en el comportamiento estadístico de los errores según el enfoque de modelado empleado. En el caso del procedimiento SARIMA, los residuales se encuentran aproximadamente centrados en cero y no presentan autocorrelación serial significativa, lo que indica que la estructura media y la dependencia temporal de la serie han sido adecuadamente capturadas. Sin embargo, tanto el diagnóstico gráfico como las pruebas formales evidencian que la varianza no es constante, con presencia de subperíodos de mayor dispersión, lo cual es consistente con fenómenos de heterocedasticidad típicos de series financieras. Adicionalmente, la distribución empírica de los residuales presenta colas más pesadas que la normal y ligera asimetría, reflejando la ocurrencia de eventos extremos que no son modelados explícitamente por el procedimiento.

Para el modelo LSTM ajustado sobre la serie sin transformar, los residuales también se encuentran centrados alrededor de cero y no muestran autocorrelación remanente, lo que sugiere una adecuada captura de la dinámica temporal. No obstante, persisten indicios de variabilidad irregular y colas relativamente pesadas en la distribución empírica, lo que indica que, aunque el modelo logra reducir el error medio respecto al SARIMA, aún no consigue estabilizar completamente la estructura de la varianza ni reproducir de forma consistente la dispersión observada en los datos reales.

En contraste, el modelo LSTM entrenado sobre la serie transformada presenta residuales claramente centrados en cero, con comportamiento aleatorio, ausencia de autocorrelación y varianza aproximadamente constante a lo largo del tiempo. Tanto los gráficos diagnósticos como las pruebas formales respaldan la homocedasticidad de los errores, lo que indica que la transformación previa permitió estabilizar la serie y facilitar el aprendizaje de su dinámica estocástica. Si bien se observan ligeras desviaciones en las colas, estas son consistentes con la presencia de eventos extremos propios de mercados financieros, pero sin afectar la estabilidad global de la distribución. En consecuencia, este enfoque no solo mejora la coherencia estadística de los residuales, sino que también proporciona una base más sólida para la generación de trayectorias sintéticas.

En síntesis, mientras que los tres enfoques logran eliminar la autocorrelación remanente en los residuales, solo el modelo LSTM con serie transformada cumple simultáneamente con los criterios de centrado en cero, varianza aproximadamente constante y estabilidad distributiva, condiciones necesarias para una representación estadísticamente consistente del proceso generador de datos y para la simulación confiable de escenarios futuros.

A partir de la evaluación conjunta del desempeño predictivo y del comportamiento estadístico de los residuales, los resultados evidencian que no todos los modelos con buen ajuste en términos de métricas tradicionales son igualmente adecuados para la generación de escenarios futuros. Los resultados muestran comportamientos claramente diferenciados entre los enfoques evaluados. En el caso del procedimiento SARIMA, las trayectorias sintéticas presentan una elevada volatilidad y una dispersión considerable entre simulaciones, lo que refleja la incapacidad del modelo para estabilizar la varianza de los errores y reproducir de forma consistente la estructura estocástica del proceso. Este comportamiento es coherente con la heterocedasticidad identificada en los residuales y limita la utilidad de estas trayectorias para análisis prospectivos.

Para el modelo LSTM ajustado sobre la serie sin transformar, si bien las trayectorias simuladas siguen la tendencia general de la serie, se observa un incremento progresivo de la volatilidad a medida que avanza el horizonte de simulación, generando una forma de dispersión tipo “embudo”. Este patrón indica que la acumulación de errores en el esquema recursivo amplifica la variabilidad, lo que sugiere inestabilidad dinámica cuando la serie no ha sido previamente transformada para estabilizar su varianza.

En contraste, el modelo LSTM entrenado sobre la serie transformada produce trayectorias sintéticas más estables, con niveles de dispersión consistentes a lo largo del horizonte de simulación y patrones que reproducen de manera más fiel la dinámica observada en la serie original. La homocedasticidad de los residuales y la ausencia de autocorrelación remanente contribuyen a que los errores incorporados en la simulación mantengan una estructura estadística coherente, permitiendo generar escenarios más realistas y controlados.

Si bien en el estudio se evaluaron distintas arquitecturas de aprendizaje profundo, incluyendo RNA, RNN y redes GRU, los resultados detallados de estos modelos no se presentan en esta sección debido a que, de manera consistente, el modelo LSTM mostró un desempeño superior tanto en términos de ajuste como de estabilidad de las trayectorias simuladas. En particular, las arquitecturas alternativas no lograron reproducir adecuadamente la dinámica estadística de la serie ni generar series sintéticas plausibles mediante simulación Montecarlo, presentando comportamientos altamente inestables o

patrones que se alejaban significativamente de la estructura empírica observada. En consecuencia, el análisis se centra en el modelo LSTM como representante de los enfoques de aprendizaje profundo, al ser el único que combina un buen desempeño predictivo con una adecuada representación de la variabilidad y dependencia temporal de la serie, criterios fundamentales para los objetivos de esta investigación.

De acuerdo con la revisión de la literatura, la mayoría de los estudios que realizan pronósticos fuera de la muestra lo hacen únicamente dentro del conjunto de prueba, mientras que aquellos que proyectan sobre fechas futuras no llevan a cabo una evaluación rigurosa de la calidad del pronóstico ni de las propiedades estadísticas de los errores asociados. Asimismo, solo los trabajos de Nadarajah et al. (2025) y Berger y Koubova (2024) incorporan un análisis explícito de residuales; sin embargo, dicho diagnóstico se limita a una inspección descriptiva y no evalúa formalmente si los residuales cumplen con supuestos clave como homocedasticidad, ausencia de autocorrelación y adecuación de la distribución empírica, ni si estos son consistentes con una capacidad de pronóstico estable. En este sentido, el presente estudio amplía el enfoque metodológico al vincular explícitamente la evaluación predictiva fuera de la muestra con un diagnóstico estadístico integral de los residuales y con la validez de las simulaciones Montecarlo, fortaleciendo así la confiabilidad de los escenarios prospectivos derivados de los modelos.

Adicionalmente, la revisión de la literatura evidencia que las métricas de desempeño reportadas fuera de la muestra se calculan, en la práctica, exclusivamente sobre el conjunto de prueba extraído del mismo período histórico, sin considerar evaluaciones sobre horizontes temporales posteriores al rango de la base de datos. En este sentido, no se identifican estudios que validen el desempeño predictivo en escenarios genuinamente prospectivos ni que examinen de manera sistemática la calidad estadística de las trayectorias simuladas. En contraste, esta tesis evalúa explícitamente el comportamiento del modelo en fechas posteriores al conjunto de observación y somete cada serie sintética generada a contrastes de idoneidad estadística, lo que permite discriminar entre trayectorias plausibles y no plausibles desde el punto de vista empírico. Este enfoque no solo fortalece la evaluación del pronóstico, sino que aporta un criterio adicional de validación para aplicaciones de simulación y análisis de riesgo en contextos financieros altamente volátiles.

9. Conclusiones

En conjunto, los resultados de esta investigación evidencian que un tratamiento estadístico riguroso de la serie de tiempo, en particular mediante la metodología de Box-Jenkins y transformaciones orientadas a garantizar estacionariedad y homocedasticidad, no solo mejora el desempeño de los modelos clásicos, sino que también potencia significativamente la capacidad de ajuste y estabilidad de las arquitecturas de aprendizaje profundo. En particular, los modelos LSTM entrenados sobre la serie transformada presentan residuales más cercanos a un proceso de ruido blanco, mayor estabilidad de varianza y trayectorias simuladas estadísticamente plausibles, lo que se traduce en una representación más realista de la dinámica del mercado. Asimismo, este estudio demuestra que la selección de modelos basada exclusivamente en métricas tradicionales de desempeño, como RMSE o R^2, resulta insuficiente cuando el objetivo es la proyección de escenarios futuros, ya que un buen ajuste promedio no garantiza una adecuada reproducción de la distribución, la volatilidad ni la estructura temporal de la serie. En este sentido, la capacidad de generación de series sintéticas válidas emerge como un criterio

más exigente y coherente con el propósito del pronóstico, permitiendo evaluar simultáneamente el ajuste, la estabilidad y la plausibilidad estadística de las trayectorias futuras. Por tanto, esta tesis aporta evidencia de que la integración de principios de modelación estadística con técnicas de aprendizaje profundo, junto con una validación basada en simulación Montecarlo, constituye una estrategia metodológica más robusta para el análisis y la proyección de precios en mercados financieros altamente volátiles como el de las criptomonedas.

9.1 Cumplimiento de objetivos

En este capítulo se evalúa el grado de cumplimiento de los objetivos planteados en la investigación, con énfasis en la comparación sistemática de distintos enfoques de aprendizaje profundo aplicados a la serie de tiempo del precio del Bitcoin. En particular, se analiza en qué medida las arquitecturas de DL consideradas permiten minimizar el error predictivo y mejorar la capacidad de realizar pronósticos por fuera de la muestra, no solo en términos de métricas tradicionales de desempeño, sino también en su aptitud para reproducir la dinámica temporal y la variabilidad observada en los datos reales. Este análisis permite establecer de forma objetiva si los modelos propuestos cumplen con el propósito central de generar proyecciones futuras estadísticamente adecuadas.

A partir de los resultados presentados en los capítulos anteriores, surge la pregunta central de este apartado: ¿cómo se cumple en esta tesis el objetivo general de evaluar métodos de Deep Learning que minimicen el error predictivo y mejoren la capacidad de realizar pronósticos por fuera de la muestra en la serie de tiempo del Bitcoin? Para responder a este interrogante, se integran los hallazgos obtenidos en términos de desempeño predictivo (R^2 y RMSE), diagnóstico estadístico de los residuales y comportamiento de las series sintéticas generadas mediante simulación Montecarlo, considerando tanto el ajuste dentro de la muestra como la estabilidad y plausibilidad de las trayectorias proyectadas en horizontes futuros. Esta evaluación integral permite establecer de manera objetiva si los modelos de aprendizaje profundo analizados cumplen efectivamente con el propósito de mejorar la calidad del pronóstico en un contexto de alta volatilidad, más allá del simple ajuste histórico de los datos.

En este contexto, resulta pertinente examinar el cumplimiento del primer objetivo específico, formulado como: Identificar los principales algoritmos de Deep Learning para la predicción de series de tiempo. La respuesta se fundamenta en la revisión sistemática de la literatura y en la selección, implementación y comparación experimental de diversas arquitecturas representativas del aprendizaje profundo para datos secuenciales, lo que permite no solo identificar los modelos más utilizados en este campo, sino también evaluar su comportamiento relativo en el contexto particular de la serie del precio del Bitcoin.

De manera complementaria, se evalúa el cumplimiento del segundo objetivo específico, orientado a aplicar los modelos de Deep Learning identificados a la serie de tiempo del precio del Bitcoin**.** Este objetivo se satisface mediante la implementación sistemática de las arquitecturas seleccionadas, incluyendo RNA, RNN, LSTM y GRU sobre la serie histórica del Bitcoin, siguiendo un esquema de entrenamiento y validación con partición temporal de los datos en conjuntos de entrenamiento y prueba. Asimismo, se consideraron distintos escenarios de modelación, tanto sobre la serie original como sobre la serie transformada para garantizar propiedades estadísticas adecuadas, lo que permitió evaluar

el comportamiento de cada modelo bajo diferentes condiciones de estacionariedad y volatilidad. Este proceso aseguró una aplicación homogénea y comparable de los modelos, permitiendo analizar su capacidad para capturar la dinámica temporal del activo y su desempeño predictivo en contextos realistas de pronóstico.

A continuación, se evalúa el cumplimiento del tercer objetivo específico, orientado a validar el ajuste de los modelos de Deep Learning. En este sentido, se examina de qué manera la metodología desarrollada en la presente tesis permite verificar de forma rigurosa la calidad del ajuste alcanzado por las arquitecturas implementadas, tanto desde el punto de vista predictivo como desde el punto de vista estadístico. Este objetivo se alcanza mediante una estrategia de validación integral que combina métricas cuantitativas de desempeño predictivo, análisis gráfico del ajuste y un diagnóstico estadístico de los residuales. En particular, se evalúan métricas como el RMSE y el R^2 tanto dentro como fuera de la muestra, lo que permite contrastar la capacidad de generalización de los modelos. Adicionalmente, se analizan los residuales a través de su comportamiento temporal, su distribución empírica, las funciones de autocorrelación y autocorrelación parcial, así como pruebas formales de homocedasticidad, con el fin de verificar si los errores presentan propiedades compatibles con un proceso de ruido blanco. Esta validación no se limita al ajuste dentro de la muestra, sino que se extiende a la coherencia estadística de los errores y a la estabilidad del modelo, proporcionando así una evaluación robusta de la idoneidad de las arquitecturas de DL empleadas.

Finalmente, se analiza el cumplimiento del cuarto objetivo específico, orientado a evaluar la capacidad predictiva de los modelos por fuera de la muestra. En este marco, se examina cómo la estrategia metodológica adoptada en la tesis permite no solo contrastar el desempeño de los modelos en el conjunto de prueba, sino también valorar su comportamiento en horizontes temporales posteriores al período de estimación, mediante la generación y validación de series sintéticas, proporcionando así una evaluación más exigente y realista de la capacidad de pronóstico en contextos de alta volatilidad como el mercado del Bitcoin.

9.2 Limitaciones

A pesar de los resultados obtenidos, es importante reconocer algunas limitaciones inherentes al alcance de la investigación.

9.2.1 Enfoque basado en precios en lugar de retornos

El análisis se realizó directamente sobre la serie de precios del Bitcoin. Sin embargo, en la literatura financiera es común trabajar con retornos, debido a que estas transformaciones tienden a presentar mejores propiedades estadísticas, como mayor estabilidad de varianza y mayor cercanía a la estacionariedad. Aunque en este estudio se aplicaron transformaciones y procedimientos de preprocesamiento orientados a garantizar condiciones adecuadas para el modelamiento, el uso directo de precios puede limitar la comparación con algunos enfoques tradicionales de la econometría financiera.

9.2.2 Tamaño y frecuencia de la muestra

La investigación utilizó datos semanales correspondientes al periodo 2022-2025. Si bien esta frecuencia permite capturar tendencias generales del mercado y reducir ruido de alta

frecuencia, también implica una menor cantidad de observaciones en comparación con datos diarios. En contextos de aprendizaje profundo, un mayor volumen de datos puede contribuir a mejorar la capacidad de generalización de los modelos.

9.2.3 Uso de una única variable endógena

El proceso de modelación se realizó utilizando únicamente la serie histórica del precio de la criptomoneda**,** sin incorporar variables explicativas externas. En este sentido, el enfoque adoptado corresponde a un esquema univariado, donde la dinámica del sistema se explica exclusivamente a partir de la información contenida en la propia serie temporal. En esta investigación, diversos estudios incorporan múltiples variables exógenas provenientes de diferentes mercados financieros y de indicadores macroeconómicos con el fin de capturar de manera más completa la dinámica del mercado. Por ejemplo, algunos trabajos incluyen medidas de volatilidad de mercados financieros globales, como la volatilidad de los futuros del S&P 500, de los bonos del Tesoro de Estados Unidos a 30 años, del índice del dólar (ICE) y de los futuros del oro de COMEX, con el objetivo de representar la volatilidad en los mercados de acciones, bonos, divisas y metales preciosos. Asimismo, se incorporan variables relacionadas con el mercado energético, como los futuros de gas natural y petróleo crudo ligero, así como indicadores de riesgo macroeconómico como el índice de riesgo geopolítico y el índice de incertidumbre económica. Otros estudios amplían aún más el conjunto de información mediante el uso de múltiples características derivadas de los datos financieros, incluyendo variables asociadas con liquidez, volatilidad, retornos pasados y características de la distribución de los precios, llegando a utilizar decenas de variables explicativas en sus modelos. De igual manera, algunos trabajos emplean precios de cierre de múltiples criptomonedas (por ejemplo, BNB, XRP, SOL, MATIC, LTC, ADA, ETH, entre otras) para capturar posibles relaciones entre distintos activos del mercado de las criptomonedas.

En contraste, el presente estudio se centra en un conjunto más limitado de variables con el fin de evaluar de manera más directa la capacidad predictiva de los modelos aplicados al precio del Bitcoin, lo cual simplifica el análisis pero puede restringir la capacidad del modelo para capturar interacciones con otros mercados o activos financieros.

9.2.4 Evaluación centrada en simulación empírica de los residuales

La generación de escenarios futuros se basa en la simulación probabilística de los residuales del modelo, a partir de la distribución observada en los datos históricos. Este enfoque permite preservar ciertas propiedades estadísticas de la serie, pero no necesariamente garantiza la reproducción completa de todas las características estocásticas presentes en la dinámica del mercado.

En la literatura de series temporales financieras, es común asumir que los residuales siguen una distribución teórica específica como la normal o t-Student, con el fin de modelar de forma más explícita el comportamiento probabilístico del proceso generador de datos. En contraste, en este estudio la simulación se realiza utilizando la información probabilística derivada de los residuales estimados, lo que permite generar trayectorias futuras consistentes con la variabilidad observada en la serie, aunque sin imponer explícitamente una forma funcional teórica para su distribución. En consecuencia, los escenarios generados deben interpretarse como posibles trayectorias probabilísticas del

precio, consistentes con la distribución observada de los errores del modelo, y no como una representación completa de la estructura estocástica subyacente del mercado.

9.2.5 Alcance específico al mercado del Bitcoin

De acuerdo con la información reportada por CoinMarketCap (2025), el mercado de las criptomonedas está compuesto por múltiples activos digitales con dinámicas propias. Entre las criptomonedas con mayor capitalización de mercado se encuentran, además de Bitcoin, Ethereum, Tether, XRP, BNB, Solana, USD Coin, TRON, Dogecoin y Cardano, las cuales presentan estructuras de mercado, niveles de liquidez y mecanismos tecnológicos distintos. Cada uno de estos activos puede responder a factores específicos relacionados con su arquitectura tecnológica, su ecosistema de aplicaciones o su función dentro del mercado financiero digital. Por esta razón, aunque el estudio proporciona evidencia empírica sobre la capacidad predictiva de los modelos aplicados al precio del Bitcoin, los resultados deben interpretarse con cautela al extrapolarlos a otras criptomonedas.

En este sentido, futuras investigaciones podrían ampliar el alcance del análisis incorporando múltiples criptomonedas dentro del proceso de modelación, lo que permitiría evaluar si el desempeño de los modelos se mantiene consistente en diferentes activos dentro del mercado de criptoactivos.

10. Recomendaciones

Con base en las limitaciones identificadas en esta investigación, se recomienda que futuros estudios amplíen el análisis incorporando retornos u otras transformaciones de la serie de precios que favorezcan la estacionariedad y la estabilidad de la varianza, así como la utilización de datos con mayor frecuencia temporal o de periodos históricos más extensos, cuando la dinámica de mercado lo permita. Asimismo, resulta pertinente explorar modelos multivariados que integren variables exógenas provenientes de distintos mercados financieros, indicadores macroeconómicos y precios de otras criptomonedas, con el objetivo de capturar interacciones complejas y mejorar la precisión de los pronósticos. En cuanto a la generación de escenarios futuros, se sugiere evaluar alternativas de simulación que incorporen distribuciones teóricas de los residuales, así como análisis de eventos extremos o condiciones de estrés de mercado, con el fin de caracterizar más explícitamente la incertidumbre y la dinámica del mercado. Finalmente, ampliar el alcance del estudio a múltiples criptomonedas permitiría verificar la consistencia del desempeño de los modelos en distintos activos y fortalecer la aplicabilidad práctica de los resultados. En conjunto, estas líneas de investigación contribuirían a consolidar un marco metodológico más robusto y generalizable, fortaleciendo el conocimiento científico en la predicción de precios de criptomonedas.

Referencias

Akgun, O. B., & Gulay, E. (2025). Dynamics in Realized Volatility Forecasting: Evaluating GARCH Models and Deep Learning Algorithms Across Parameter Variations: OB Akgun, E. Guly. Computational Economics, 65(6), 3971-4013.

AlMadany, N. N., Hujran, O., Al Naymat, G., & Maghyereh, A. (2024). Forecasting cryptocurrency returns using classical statistical and deep learning techniques. International Journal of Information Management Data Insights, 4(2), 100251.

Berger, T., & Koubová, J. (2024). Forecasting Bitcoin returns: Econometric time series analysis vs. machine learning. Journal of Forecasting, 43(7), 2904-2916.

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). Time Series Analysis: Forecasting and Control (5th ed.). Hoboken, NJ: John Wiley & Sons.

Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. Journal of the Royal Statistical Society: Series B (Methodological), 26(2), 211–243.

Breusch, T. S., & Pagan, A. R. (1979). *A simple test for heteroscedasticity and random coefficient variation*. Econometrica, 47(5), 1287–1294.

Brockwell, P. J., & Davis, R. A. (2016). Introduction to Time Series and Forecasting (3rd ed.). Springer.

Cakici, N., Shahzad, S. J. H., Będowska-Sójka, B., & Zaremba, A. (2024). Machine learning and the cross-section of cryptocurrency returns. International Review of Financial Analysis, 94, 103244.

CoinMarketCap. (2025). Ranking de criptomonedas por capitalización de mercado. https://coinmarketcap.com/es/

Erfanian, S., Zhou, Y., Razzaq, A., Abbas, A., Safeer, A. A., & Li, T. (2022). Predicting bitcoin (BTC) price in the context of economic theories: A machine learning approach. Entropy, 24(10), 1487.

Ferdiansyah, et al. (2023). Hybrid gated recurrent unit bidirectional-long short-term memory model to improve cryptocurrency prediction accuracy. *IAES International Journal of Artificial Intelligence*, *12*(1).

Fleischer, J. P., von Laszewski, G., Theran, C., & Bautista, Y. J. P. (2022). Time series analysis of cryptocurrency prices using long short-term memory. *Algorithms*, *15*(7).

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

Haykin, S. (1999). *Neural networks: A comprehensive foundation* (2nd ed.). Prentice Hall.

Kang, C. Y., Lee, C. P., & Lim, K. M. (2022). Cryptocurrency price prediction with convolutional neural network and stacked gated recurrent unit. *Data, 7*(11).

Kim, H. Y., & Won, C. H. (2018). Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models. *Expert Systems with Applications, 103*, 25–37.

Kim, T., Jo, H., Choi, W., & Jang, B. G. (2025). Bitcoin Price Direction Forecasting and Market Variables. Journal of Futures Markets.

Kleijnen, J. P. C. (1995). Verification and validation of simulation models. European Journal of Operational Research, 82(1), 145–162

Lahmiri, S., & Bekiros, S. (2019). Cryptocurrency forecasting with deep learning chaotic neural networks. *Chaos, Solitons & Fractals, 118*, 35–40.

Livieris, I. E., Pintelas, E., Stavroyiannis, S., & Pintelas, P. (2020). Ensemble deep learning models for forecasting cryptocurrency time-series. *Algorithms, 13*(5).

Ljung, G. M., & Box, G. E. P. (1978). On a Measure of Lack of Fit in Time Series Models. Biometrika, 65(2), 297–303.

López-García, A. (2023). Análisis comparativo de redes neuronales profundas para la predicción del precio de mercado de bitcoin. Revista Electrónica de Comunicaciones y Trabajos de ASEPUMA, 24(1), 1-16.

M. A. Ammer and T. H. H. Aldhyani, "Deep Learning Algorithm to Predict Cryptocurrency Fluctuation Prices: Increasing Investment Awareness," Electron., vol. 11, no. 15, 2022.

M. M. Patel, S. Tanwar, R. Gupta, and N. Kumar, "A Deep Learning-based Cryptocurrency Price Prediction Scheme for Financial Institutions," J. Inf. Secur. Appl., vol. 55, 2020.

Magner, N., & Hardy, N. (2022). Cryptocurrency forecasting: More evidence of the Meese-Rogoff puzzle. Mathematics, 10(13), 2338.

Nadarajah, S., Mba, J. C., Rakotomarolahy, P., & Ratolojanahary, H. T. (2025). Ensemble Learning and an Adaptive Neuro-Fuzzy Inference System for Cryptocurrency Volatility Forecasting. Journal of Risk and Financial Management, 18(2), 52.

Nasirtafreshi, I. (2022). Forecasting cryptocurrency prices using recurrent neural network and long short-term memory. *Data & Knowledge Engineering, 139*.

Oyedele, A. A., et al. (2023). Performance evaluation of deep learning and boosted trees for cryptocurrency closing price prediction. *Expert Systems with Applications, 213*.

Patel, M. M., Tanwar, S., Gupta, R., & Kumar, N. (2020). A deep learning-based cryptocurrency price prediction scheme for financial institutions. *Journal of Information Security and Applications, 55*.

Tanwar, S., Patel, N. P., Sharma, G., & Davidson, I. E. (2021). Deep learning-based cryptocurrency price prediction scheme with inter-dependent relations. *IEEE Access, 9*, 138633–138646.

Wu, C.-H., Lu, C.-C., Ma, Y.-F., & Lu, R.-S. (2019). A new forecasting framework for bitcoin price with LSTM. In *IEEE International Conference on Data Mining Workshops*.

Yao, Y., et al. (2018). Predictive analysis of cryptocurrency price using deep learning. *International Journal of Engineering & Technology*.

Zhou, Y., Xie, C., Wang, G. J., Gong, J., & Zhu, Y. (2025). Forecasting cryptocurrency volatility: a novel framework based on the evolving multiscale graph neural network. *Financial Innovation*, *11*(1), 87.