



**Universidad  
Europea**

**UNIVERSIDAD EUROPEA DE MADRID**

**ESCUELA DE ARQUITECTURA, INGENIERÍA Y DISEÑO**

**MÁSTER UNIVERSITARIO EN ANÁLISIS DE DATOS MASIVOS**

**TRABAJO FIN DE MÁSTER**

**ESTIMACION DEL MERCADO DE  
CRIPTODIVISA EN BANCA DE INVERSION**

**XAVIER ELIAS MORA VELASQUEZ**

**Dirigido por**

**CARLOS WOLFRAM ROZAS RODRIGUEZ**

**CURSO 2025 - 2026**

**TÍTULO:** ESTIMACION DEL MERCADO DE CRIPTODIVISA EN BANCA DE INVERSION

**AUTOR:** XAVIER ELIAS MORA VELASQUEZ

**TITULACIÓN:** MÁSTER UNIVERSITARIO EN ANÁLISIS DE DATOS MASIVOS

**DIRECTOR DEL PROYECTO:** CARLOS WOLFRAM ROZAS RODRIGUEZ

**FECHA:** Julio de 2025

## RESUMEN

El presente Trabajo de Fin de Máster tiene como objetivo el desarrollo e implementación de modelos predictivos para estimar el comportamiento del precio de Bitcoin en el mercado de criptomonedas. Para ello, se emplean técnicas de análisis de datos, procesamiento de información financiera y algoritmos de aprendizaje automático.

El proyecto se basa en el uso de dos fuentes de datos principales. En primer lugar, el conjunto de datos CryptoLM-Bitcoin-BTC-USDT, disponible en la plataforma Hugging Face, que proporciona información histórica detallada sobre los precios de Bitcoin frente al USDT. Este dataset incluye múltiples indicadores técnicos, tales como medias móviles (MA), índice de fuerza relativa (RSI), bandas de Bollinger, MACD, entre otros, con una frecuencia de actualización de 3 minutos (y un retraso de 1 minuto), lo que permite trabajar con información casi en tiempo real. En segundo lugar, se integran variables macroeconómicas procedentes del World Economic Outlook (WEO) del Fondo Monetario Internacional (FMI), con el objetivo de enriquecer el análisis y evaluar el impacto de factores económicos globales sobre el comportamiento del mercado de criptomonedas.

A partir de esta información, se entrenan y comparan distintos modelos de aprendizaje automático, evaluando su rendimiento mediante métricas como el error cuadrático medio (RMSE), el error absoluto medio (MAE) y el coeficiente de determinación ( $R^2$ ). Los resultados obtenidos permiten analizar la capacidad predictiva de los modelos desarrollados y su posible aplicabilidad en la toma de decisiones de inversión, contribuyendo a la gestión del riesgo y la optimización de estrategias en un entorno altamente volátil.

**Palabras clave:** Bitcoin, Criptomonedas, Modelos predictivos, Aprendizaje automático, Indicadores técnicos, Datos macroeconómicos, Gestión de riesgos, Volatilidad

## ABSTRACT

The present Master's Thesis aims to develop and implement predictive models to estimate the behavior of Bitcoin prices in the cryptocurrency market. To achieve this, data analysis techniques, financial information processing, and machine learning algorithms are employed.

The project is based on two main data sources. First, the CryptoLM-Bitcoin-BTC-USDT dataset, available on the Hugging Face platform, which provides detailed historical information on Bitcoin prices against USDT. This dataset includes multiple technical indicators, such as moving averages (MA), relative strength index (RSI), Bollinger Bands, MACD, among others, with a refresh frequency of 3 minutes (and a delay of 1 minute), allowing for near real-time analysis. Second, macroeconomic variables from the International Monetary Fund's (IMF) World Economic Outlook (WEO) are integrated to enrich the analysis and evaluate the impact of global economic factors on the behavior of the cryptocurrency market.

Using this information, various machine learning models are trained and compared, evaluating their performance through metrics such as Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and the coefficient of determination ( $R^2$ ). The results obtained enable an analysis of the predictive capabilities of the developed models and their potential applicability in investment decision-making, contributing to risk management and the optimization of strategies in a highly volatile environment.

**Keywords:** Bitcoin, Cryptocurrency, Predictive models, Machine learning, Technical indicators, Macroeconomic data, Risk management, Volatility

# Índice general

<b>1. INTRODUCCION</b>	<b>8</b>
1.1. Contexto y justificación	8
1.2. Planteamiento del problema	8
1.3. Objetivos del proyecto	8
1.4. Estructura de la memoria	9
<b>2. ANTECEDENTES / ESTADO DEL ARTE</b>	<b>12</b>
2.1. Estado del arte	12
2.2. Contexto y justificación	13
2.3. Planteamiento del problema	14
<b>3. OBJETIVOS</b>	<b>16</b>
3.1. Objetivos generales	16
3.2. Objetivos específicos	16
<b>4. DESARROLLO DEL PROYECTO</b>	<b>17</b>
4.1. Descripción de la solución	17
4.2. Marco Teórico	17
4.2.1. El mercado de criptomonedas y su complejidad predictiva	17
Estrategias de <i>trading</i>	18
Estrategias de <i>hedging</i>	18
4.2.2. Series temporales y aprendizaje profundo	19
Long Short-Term Memory (LSTM)	20
Temporal Convolutional Networks (TCN)	20
Transformers	21
4.2.3. Selección de variables y mecanismos de atención	21
4.2.4. Variables Macroeconómicas	23
Crecimiento Económico	23
Empleo	24
Inflación	24
Cuenta Corriente	25
4.2.5. Variable objetivo	25
4.3. Metodología	26
4.3.1. Fuentes y descripción de los datos	26
Dataset de mercado de criptomoneda	26
Dataset de World Economic Outlook del IMF	28
4.3.2. Análisis Exploratorio de los datos	29
Histograma	29
BoxPlot	29
Matriz de correlaciones	30
4.3.3. Preprocesamiento de datos	31
Combinación de datos macro económicos y de criptodivisas	31

Filtrado de datos . . . . .	31
Imputación de datos . . . . .	31
Normalización y escalado . . . . .	31
Ventanas temporales y horizonte de predicción . . . . .	32
4.3.4. Selección de variables mediante mecanismos de atención . . . . .	32
4.3.5. Arquitecturas de modelo . . . . .	33
Long Short-Term Memory (LSTM) . . . . .	33
Temporal Convolutional Network (TCN) . . . . .	34
Transformer . . . . .	35
4.3.6. Entrenamiento y validación . . . . .	35
Curvas de entrenamiento y validación . . . . .	36
4.3.7. Planificación . . . . .	38
4.3.8. Plataforma tecnológica . . . . .	39
4.3.9. Presupuesto de entrenamiento . . . . .	40
<b>5. Resultados</b>	<b>41</b>
5.1. Métricas de precisión . . . . .	41
5.2. Serie real versus predicción . . . . .	41
5.3. Síntesis . . . . .	42
<b>6. Discusión</b>	<b>43</b>
6.1. Síntesis de resultados . . . . .	43
6.2. Comparación con el estado del arte . . . . .	43
6.3. Limitaciones del estudio . . . . .	43
6.3.1. Datos y alcance . . . . .	43
6.3.2. Tecnología y cómputo . . . . .	43
6.4. Cambios y adaptación de la metodología . . . . .	44
6.5. Implicaciones prácticas . . . . .	44
6.6. Conclusión . . . . .	44
6.7. Trabajo futuro . . . . .	44
<b>Bibliografía</b>	<b>46</b>

## Índice de Figuras

4.1. Evolución de histórica del Bitcoin. Fuente: TradingView, 2025 . . . . .	18
4.2. Arquitectura general de LSTM. Fuente: Hochreiter y Schmidhuber, 1997 . . . . .	20
4.3. Estructura general de una red TCN . . . . .	21
4.4. Visualización del modelo Transformer y su mecanismo interno. . . . .	21
4.5. Arquitectura con mecanismo de atención dual: temporal y por variable. Fuente: Qin et al., 2017 . . . . .	22
4.6. Mecanismo de atención enfocado en selección de variables. Fuente: Zhou et al., 2021 . . . . .	23
4.7. Producto Interno Bruto. Fuente: Elaboración propia . . . . .	24
4.8. Tasa de Desempleo. Fuente: Elaboración propia . . . . .	24
4.9. Índice de Precios al Consumidor. Fuente: Elaboración propia . . . . .	25
4.10. Saldo de cuenta corriente. Fuente: Elaboración propia . . . . .	25
4.11. Histograma del precio de cierre diario de Bitcoin. Fuente: Propia . . . . .	29
4.12. Boxplot del volumen de transacciones diarias. Fuente: Propia . . . . .	30
4.13. Matriz de correlaciones de variables técnicas. Fuente: Propia . . . . .	30
4.14. Ventana deslizante. Fuente: Chen et al., 2021 . . . . .	32
4.15. Representación de series temporales mediante codificación posicional en modelos Transformer. Fuente: Grigsby et al., 2021 . . . . .	33
4.16. Curvas de pérdida – Configuración 1. El Transformer (azul) y el LSTM (naranja) convergen de forma estable; el TCN (verde) alcanza la pérdida casi nula en pocas épocas pero tiende a sobre-ajustar si se prolonga el entrenamiento. . . . .	37
4.17. Curvas de pérdida – Configuración 2. El Transformer reduce la pérdida de validación de forma más lenta pero alcanza finalmente el valor mínimo; el LSTM se estanca ligeramente por encima. . . . .	37
4.18. Curvas de pérdida – Configuración 3. Se observa un descenso monótono de la pérdida en todos los modelos; el Transformer mantiene la brecha de generalización más pequeña. . . . .	37
4.19. Curvas de pérdida – Configuración 4. El LSTM muestra oscilaciones en la pérdida de validación; el TCN converge rápido pero ofrece la peor generalización. . . . .	38
4.20. Curvas de pérdida – Configuración 5. El Transformer alcanza la mejor combinación de baja pérdida de entrenamiento y validación; el LSTM estabiliza en un nivel superior y el TCN vuelve a sobre-ajustar. . . . .	38
4.21. Cronograma del proyecto TFM con flechas ortogonales y sin texto superpuesto. . . . .	39
5.1. Precio de cierre real (azul) y predicho (naranja) por el Transformer durante los últimos 90 días del conjunto de prueba. . . . .	41

## Índice de Tablas

2.1. Métodos recientes para la predicción del precio de Bitcoin . . . . .	13
4.1. Descripción del dataset CryptoLM-Bitcoin-BTC-USDT . . . . .	27
4.2. Descripción del Dataset de World Economic Outlook del IMF . . . . .	28
4.3. Comparación de configuraciones de entrenamiento y parámetros de los modelos. . . . .	36
4.4. Presupuesto de entrenamiento (coste eléctrico y uso proporcional de equipo). . . . .	40
5.1. Errores sobre el conjunto de prueba. . . . .	41



# Capítulo 1. INTRODUCCION

## 1.1 Contexto y justificación

En los últimos años, el mercado de las criptomonedas ha experimentado un crecimiento exponencial, tanto en términos de capitalización como en su adopción por parte de inversores individuales e institucionales. Entre ellas, Bitcoin se ha consolidado como el activo digital más representativo, convirtiéndose en un instrumento de inversión ampliamente negociado a nivel global.

No obstante, su alta volatilidad, influenciada por factores técnicos, especulativos y macroeconómicos, dificulta la toma de decisiones informadas respecto a cuándo comprar o vender. Esta incertidumbre representa un desafío significativo para quienes buscan optimizar sus estrategias de inversión, minimizar riesgos o identificar oportunidades en tiempo real.

En este contexto, el presente proyecto surge con el propósito de explorar el potencial de las técnicas de análisis de datos y aprendizaje automático en la predicción del precio de Bitcoin, con especial énfasis en su aplicación práctica a la toma de decisiones de compra y venta. Se busca así contribuir al desarrollo de soluciones tecnológicas que mejoren la capacidad de anticipación en un entorno de mercado complejo y dinámico.

Este trabajo se enmarca dentro del análisis predictivo, una de las ramas más relevantes de la ciencia de datos, con aplicaciones directas en los campos de las finanzas, la gestión del riesgo y los sistemas de soporte a decisiones automatizadas.

## 1.2 Planteamiento del problema

El mercado de criptomonedas, y en particular Bitcoin, se caracteriza por una alta volatilidad y un comportamiento difícil de anticipar mediante métodos tradicionales de análisis financiero. Esta incertidumbre complica la toma de decisiones estratégicas, especialmente en lo que respecta a determinar cuándo comprar o vender este activo digital.

Ante este escenario, surge la siguiente pregunta motriz:

¿Es posible desarrollar un modelo predictivo basado en técnicas de aprendizaje automático que permita estimar con precisión el precio futuro de Bitcoin, incorporando indicadores técnicos y variables macroeconómicas, con el fin de apoyar la toma de decisiones de compra y venta?

Este proyecto se enmarca dentro de una investigación de carácter aplicado, con orientación científico-técnica, que busca evaluar la eficacia de diversos algoritmos de machine learning en la predicción del comportamiento del precio de Bitcoin. La comparación de modelos se realiza bajo métricas de desempeño cuantitativas, con el objetivo de identificar herramientas predictivas que puedan aportar valor en contextos reales de inversión y gestión del riesgo.

## 1.3 Objetivos del proyecto

El objetivo de este trabajo es desarrollar un modelo predictivo que permita estimar con precisión el precio futuro de Bitcoin (BTC/USDT) mediante el uso de técnicas de análisis de datos,

procesamiento de información financiera y algoritmos de aprendizaje automático. El propósito central es ofrecer una herramienta que sirva de apoyo en la toma de decisiones de compra y venta, en un mercado altamente volátil y difícil de anticipar.

Para alcanzar este objetivo, se integrarán y procesarán datos provenientes de dos fuentes principales: el conjunto de datos CryptoLM-Bitcoin-BTC-USDT, que proporciona información técnica detallada y actualizada sobre el comportamiento histórico de Bitcoin; y las variables macroeconómicas extraídas del informe World Economic Outlook (WEO) del Fondo Monetario Internacional (FMI), que aportan contexto sobre factores económicos globales. A partir de esta información, se entrenarán y evaluarán distintos modelos de machine learning, utilizando métricas como RMSE, MAE y  $R^2$  para comparar su rendimiento.

Finalmente, se analizará la viabilidad del modelo con mejor desempeño en situaciones reales de inversión, considerando tanto su capacidad predictiva como su aplicabilidad práctica para optimizar estrategias de entrada y salida en el mercado de Bitcoin.

## 1.4 Estructura de la memoria

El desarrollo del trabajo se llevó a cabo siguiendo una metodología estructurada basada en análisis de datos y técnicas de aprendizaje automático aplicadas a series temporales. A continuación, se describen las principales etapas que conforman la estructura del proyecto:

### 1. Importación de librerías y configuración del entorno

Se preparó el entorno de trabajo mediante la incorporación de librerías especializadas para el tratamiento de datos, visualización, modelado predictivo y evaluación de modelos. Entre las herramientas utilizadas se incluyen frameworks para aprendizaje profundo, métricas de validación y utilidades para la gestión de series temporales.

### 2. Carga y exploración de datos

Se integraron dos fuentes de información perfectamente sincronizadas a frecuencia diaria:

- a) *CryptoLM-Bitcoin-BTC-USDT* Contiene  $\sim 5,7$  millones de velas de 1 minuto del par BTC/USDT entre el 1 ene 2015 y el 30 jun 2025. Cada registro incluye *timestamp*, *open*, *high*, *low*, *close*, *volume* y 15 indicadores técnicos pre-calculados ( $MA_{20}$ ,  $MA_{50}$ ,  $MA_{200}$ , RSI, %K, %D, ADX, ATR, pendiente de línea de tendencia, MACD, Signal, Histogram, BL\_Upper/Lower, MN\_Upper/Lower).
- b) *World Economic Outlook (WEO) – FMI*. Aporta variables macroeconómicas (PIB real %, CPI %, tasa de desempleo, cuenta corriente % PIB) para 2015–2024. Las series anuales/trimestrales se interpolaron a frecuencia mensual y posteriormente diaria para su alineación temporal con el precio de Bitcoin.

### 3. Preprocesamiento de datos

Esta fase incluyó:

- Limpieza de datos y tratamiento de valores faltantes.
- Conversión y alineación temporal entre las distintas fuentes.
- Normalización de variables y reducción de ruido.

- Ingeniería de características a partir de los indicadores técnicos y económicos.

#### 4. **Análisis exploratorio**

Para contextualizar los datos, se inició con una inspección global de las series.

*En primer lugar*, la Figura 4.11 pone de manifiesto tres rangos distintos en los precios de cierre: valores por debajo de 20 000 USDT, un segundo bloque entre 40 000–60 000 USDT y, finalmente, un conjunto de máximos extremos. Esta distribución asimétrica confirma la existencia de cambios de régimen que deberán capturarse con modelos flexibles.

*En segundo lugar*, el diagrama de caja del volumen (Figura 4.12) exhibe una mediana de 35 000 BTC, acompañada de valores atípicos superiores a 400 000 BTC. Para evitar que dichos picos distorsionen el entrenamiento, se aplicó un escalado robusto que atenúa su impacto.

*Por último*, la matriz de correlaciones (Figura 4.13) revela una fuerte colinealidad entre los precios y sus medias móviles, mientras que el volumen mantiene una asociación débil con las demás variables. Este hallazgo justifica el uso de mecanismos de atención para seleccionar características y reducir la redundancia.

*En síntesis*, la volatilidad no lineal respalda la elección de modelos LSTM, TCN y Transformer; los outliers requieren preprocesado robusto, y la colinealidad aconseja depurar variables redundantes antes del modelado.

#### 5. **Preparación de los datos para modelado**

Se estructuraron las secuencias de entrada y salida para los modelos de series temporales. Asimismo, se dividieron los datos en conjuntos de entrenamiento, validación y prueba, respetando su naturaleza secuencial.

#### 6. **Entrenamiento de modelos predictivos**

Se implementaron tres enfoques basados en redes neuronales profundas:

- TimeSeriesTransformer, con arquitectura de atención,
- LSTM (Long Short-Term Memory), diseñada para capturar dependencias de largo plazo,
- TCN (Temporal Convolutional Network), basada en convoluciones causales. Los modelos fueron entrenados bajo condiciones controladas, aplicando criterios de parada temprana para evitar sobreajuste.

#### 7. **Evaluación y comparación de modelos**

El desempeño de cada modelo se evaluó utilizando métricas estándar como:

- Error cuadrático medio (RMSE).
- Error absoluto medio (MAE).
- Coeficiente de determinación ( $R^2$ ).

Se generaron visualizaciones para comparar las predicciones con los valores reales y analizar la precisión de los modelos a lo largo del tiempo.

#### 8. **Análisis de resultados y selección del modelo óptimo**

Se identificó el modelo con mejor rendimiento global, evaluando no solo su precisión estadística, sino también su robustez y utilidad práctica para decisiones de compra y venta de Bitcoin.

9. **Conclusiones y recomendaciones** Finalmente, se presentan las conclusiones del estudio, destacando las contribuciones del modelo desarrollado, sus limitaciones y propuestas para trabajos futuros que busquen ampliar el alcance o mejorar la capacidad predictiva del sistema.

## Capítulo 2. ANTECEDENTES / ESTADO DEL ARTE

### 2.1 Estado del arte

En los últimos años, la investigación académica ha mostrado un notable interés en el desarrollo de modelos predictivos basados en inteligencia artificial para el análisis del mercado de criptomonedas, especialmente en la predicción de precios de Bitcoin. Debido a su alta volatilidad y complejidad, numerosos estudios han propuesto el uso de técnicas avanzadas de aprendizaje automático (machine learning) y aprendizaje profundo (deep learning), superando en muchos casos a los enfoques estadísticos tradicionales como ARIMA.

Entre los modelos más destacados se encuentran las redes neuronales híbridas que combinan arquitecturas convolucionales (CNN) con LSTM. Omole y Enke (2024) propusieron un modelo CNN-LSTM que, al integrar la técnica de selección de características Boruta, logró una precisión del 82.44 % en la predicción de la dirección del precio de Bitcoin. Este enfoque fue comparado con modelos ARIMA y TCN (Temporal Convolutional Network), superándolos en términos de precisión y robustez durante períodos de alta volatilidad.

En una línea similar, Chen et al. (2024) desarrollaron un modelo híbrido que combina CNN, BiLSTM y mecanismos de atención (CNN-BiLSTM-AM). Este modelo demostró mejoras significativas respecto a las arquitecturas convencionales, logrando capturar tanto patrones locales como dependencias a largo plazo en series temporales de criptomonedas. Los resultados obtenidos mostraron un rendimiento superior en métricas como RMSE y MAE, reforzando la utilidad de modelos híbridos con componentes de atención en contextos financieros no lineales.

Por otra parte, una revisión sistemática realizada por Wu et al. (2024) evaluó comparativamente modelos basados en Transformer, LSTM y CNN, concluyendo que los modelos multivariantes, especialmente los CNN-LSTM, ofrecen ventajas sustanciales en la predicción a múltiples pasos frente a técnicas tradicionales. Esta revisión destaca también que el uso de datos multifuente —incluyendo indicadores técnicos y macroeconómicos— incrementa el poder predictivo de los modelos.

En cuanto a técnicas más clásicas de machine learning, algoritmos como XGBoost también han sido aplicados con éxito al mercado de criptodivisas. Modelos entrenados con indicadores técnicos como RSI, MACD y medias móviles lograron resultados competitivos, aunque generalmente inferiores a los modelos de deep learning en escenarios con alta complejidad no lineal.

Finalmente, Livieris et al. (2021) presentaron un modelo de red neuronal profunda con múltiples entradas (MICDL) que procesa de forma independiente distintas series de precios de criptomonedas, lo que permitió mejorar la generalización y reducir el sobreajuste. En comparación con redes LSTM simples, este enfoque obtuvo mejoras tanto en precisión como en estabilidad de las predicciones.

En conjunto, estos estudios sugieren que los modelos híbridos basados en CNN, LSTM y mecanismos de atención son actualmente los más prometedores para la predicción de precios de criptomonedas. Además, la incorporación de variables macroeconómicas y el uso de técnicas de selección de características contribuyen a mejorar significativamente el rendimiento predictivo, especialmente en horizontes de corto a mediano plazo.  
Autor desconocido, 2024; Chen et al., 2024; Livieris et al., 2021; Omole y Enke, 2024; Wu et al., 2024

Autor / Año	Modelo	Datos	Horiz.	Métrica / Resultado
Omole y Enke (2024)	CNN–LSTM + Boruta	BTC/USD (Binance)	1 d	Precisión direccional (DA) = 82.44 %
Chen et al. (2024)	CNN–BiLSTM–AM	BTC/USD + indicadores	1 d	RMSE y MAE 15 % menores que LSTM
Wu et al. (2024)	Comparativo (Transformer, LSTM, CNN)	Varios exchanges	Multi	CNN–LSTM destaca en predicción multipasos
Autor desconocido (2024)	XGBoost + indicadores técnicos	BTC/USD	1 d	MAE competitivo; inferior a DL en alta volatilidad
Livieris et al. (2021)	MICDL (multi-input CNN–LSTM)	BTC/USD (Kraken)	1 d	RMSE 15 % menor y mayor estabilidad que LSTM simple

Tabla 2.1. Métodos recientes para la predicción del precio de Bitcoin

## 2.2 Contexto y justificación

El mercado de criptomonedas ha irrumpido con fuerza en la última década, transformándose en uno de los sectores financieros con mayor dinamismo y crecimiento. Desde la creación de Bitcoin en 2009, y la posterior aparición de Ethereum y otros activos digitales, este ecosistema ha evolucionado desde ser un nicho tecnológico a convertirse en un mercado de capitalización multimillonaria con una creciente participación de inversores particulares, empresas fintech y grandes instituciones financieras.

A diferencia de los mercados financieros tradicionales, las criptomonedas se caracterizan por una operativa ininterrumpida las 24 horas del día, los 7 días de la semana, y por una notable volatilidad, con variaciones bruscas en los precios que pueden producirse en intervalos de tiempo muy reducidos. Esta alta volatilidad, combinada con la ausencia de regulación estandarizada y la influencia de factores técnicos, especulativos y sociales, dificulta la predicción fiable del comportamiento de estos activos utilizando modelos estadísticos clásicos.

En este contexto, el análisis de datos masivos (Big Data) y las técnicas de inteligencia artificial, particularmente el aprendizaje automático (machine learning), se consolidan como herramientas prometedoras para abordar este tipo de problemas complejos y no lineales. La posibilidad de entrenar modelos predictivos con datos históricos de precios y una amplia variedad de indicadores técnicos permite generar estimaciones más ajustadas, contribuyendo a una toma de decisiones informada y a una gestión del riesgo más eficiente.

El presente proyecto se inscribe dentro de esta línea de trabajo, aplicando distintas técnicas

de análisis de datos, preprocesado y modelado predictivo sobre el comportamiento de Bitcoin y Ethereum, dos de las criptomonedas con mayor peso en el mercado, mediante el uso de un dataset especializado y actualizado en tiempo casi real.

El interés académico, científico y práctico por el estudio de modelos predictivos en mercados financieros ha crecido de forma significativa en los últimos años, especialmente en lo relativo a activos de alta volatilidad como las criptomonedas. La capacidad de anticipar movimientos de precio, aunque sea de forma aproximada, puede suponer una ventaja competitiva importante tanto para inversores como para desarrolladores de soluciones automatizadas de trading y gestión de carteras.

En este sentido, este proyecto resulta pertinente por varias razones:

**Innovación tecnológica:** Explora la aplicabilidad de técnicas de machine learning en un entorno emergente y complejo, como es el mercado de criptomonedas, donde la alta volatilidad y la dinámica de los precios requieren enfoques de modelado avanzados y adaptativos.

**Valor académico y profesional:** Aporta conocimiento al campo de la ciencia de datos aplicada a finanzas, con potencial de extrapolación a otros mercados volátiles o activos financieros de características similares.

**Utilidad práctica:** Los resultados pueden contribuir a mejorar la toma de decisiones en estrategias de inversión y gestión de riesgos, tanto para perfiles institucionales como para pequeños inversores interesados en las criptomonedas.

**Aprovechamiento de recursos actuales:** Utiliza un dataset actualizado y de alta frecuencia disponible en la plataforma Hugging Face, lo que permite trabajar con información cercana al tiempo real, optimizando la relevancia y aplicabilidad de los modelos generados.

Además, el proyecto permite analizar las limitaciones y desafíos asociados al uso de algoritmos de aprendizaje automático en entornos donde factores externos (noticias, regulaciones, sentimiento de mercado) pueden impactar significativamente en el precio, algo que no siempre queda reflejado en los datos históricos disponibles.

## **2.3 Planteamiento del problema**

Bitcoin, como principal exponente del mercado de criptomonedas, presenta una dinámica de precios altamente volátil que desafía los métodos tradicionales de análisis financiero. Esta volatilidad dificulta la identificación de patrones consistentes y, en consecuencia, complica la toma de decisiones estratégicas en torno a la compra y venta del activo. Aunque los inversores cuentan con herramientas técnicas para el análisis de tendencias, su eficacia se ve limitada en entornos tan inestables como el cryptoactivo.

En los últimos años, se ha observado un avance significativo en la aplicación de modelos de inteligencia artificial al ámbito financiero, en particular en el uso de algoritmos de aprendizaje automático y profundo para la predicción de series temporales. Estudios recientes han mostrado que enfoques como CNN-LSTM, XGBoost o arquitecturas híbridas con mecanismos de atención superan con creces a modelos tradicionales como ARIMA, gracias a su capacidad para capturar relaciones no lineales, secuenciales y multivariadas.

No obstante, muchas de estas investigaciones se han centrado exclusivamente en indicadores técnicos, obviando el papel que pueden desempeñar las variables macroeconómicas —como tasas de interés, crecimiento económico o expectativas de inflación— en la formación del precio del Bitcoin, especialmente a medida que este activo se inserta en dinámicas financieras globales.

A partir de esta brecha, se plantea la siguiente pregunta de investigación:

¿Es posible construir un modelo predictivo basado en técnicas de machine learning que, al combinar indicadores técnicos y variables macroeconómicas, logre estimar con mayor precisión el precio futuro de Bitcoin y proporcione una base robusta para decisiones de compra y venta?

Este trabajo se orienta a una investigación aplicada con enfoque científico y tecnológico, cuyo objetivo es comparar el desempeño de distintos modelos predictivos sobre datos multivariantes. A través del uso de métricas de evaluación cuantitativas, se busca identificar cuál de los enfoques proporciona resultados más consistentes y útiles para su aplicación en contextos reales de inversión y gestión del riesgo.



## Capítulo 3. OBJETIVOS

### 3.1 Objetivos generales

Diseñar un modelo predictivo utilizando técnicas avanzadas de aprendizaje automático que estime el precio futuro de Bitcoin, integrando tanto indicadores técnicos como variables macroeconómicas, con el objetivo de servir como base para que las aplicaciones de banca digital en Ecuador puedan, en un futuro, permitir a sus usuarios identificar los momentos más adecuados para invertir, mantener o vender sus activos en Bitcoin.

### 3.2 Objetivos específicos

- Definir la arquitectura del modelo predictivo utilizando algoritmos avanzados de aprendizaje automático adecuados para series temporales, como LSTM y TimeSeriesTransformer.
- Seleccionar datasets relevantes, incluyendo datos históricos de precios de Bitcoin e indicadores macroeconómicos, aplicando criterios de calidad, actualización y representatividad.
- Realizar un análisis estadístico descriptivo de los datos, considerando medidas de tendencia central, dispersión y detección de valores atípicos.
- Explorar visualmente los datos mediante herramientas gráficas como histogramas y diagramas de caja, con el fin de identificar patrones y anomalías relevantes.
- Preprocesar los datos aplicando técnicas de limpieza, imputación, normalización y generación de nuevas variables que mejoren el desempeño del modelo.
- Entrenar y validar diferentes modelos predictivos, evaluando su rendimiento a través de métricas cuantitativas como MAE, RMSE y  $R^2$ .
- Analizar e interpretar los resultados obtenidos, identificando las variables más influyentes y la robustez del modelo frente a distintos escenarios.
- Proponer un esquema de implementación del modelo como herramienta de análisis en plataformas digitales, orientado a respaldar decisiones de inversión en criptomonedas en el contexto ecuatoriano.

## Capítulo 4. DESARROLLO DEL PROYECTO

### 4.1 Descripción de la solución

En este capítulo se describe detalladamente el proceso de desarrollo del modelo predictivo aplicado al mercado de criptomonedas, cuyo objetivo principal es estimar el precio de cierre de Bitcoin. El enfoque metodológico combina técnicas de análisis de series temporales con métodos avanzados de aprendizaje automático, integrando tanto información histórica del mercado de criptomonedas como variables macroeconómicas relevantes a nivel nacional y global.

El desarrollo del proyecto se ha estructurado en varias fases: exploración y análisis de los datos, preprocesamiento, selección de variables, diseño del modelo, entrenamiento y validación de resultados. Cada etapa ha sido diseñada para garantizar la construcción de un modelo robusto, capaz de capturar patrones complejos y ofrecer predicciones precisas bajo distintas condiciones de mercado.

Durante el proceso se han aplicado técnicas de transformación y normalización de variables, selección automatizada de características mediante mecanismos de atención, y ajuste de hiperparámetros, todo ello evaluado a través de métricas estándar de rendimiento.

Además del objetivo netamente predictivo, se ha priorizado la interpretabilidad del modelo, con el fin de identificar qué variables —tanto técnicas como macroeconómicas— influyen en mayor medida en la evolución del precio de los cryptoactivos.

En las secciones siguientes se detalla de forma estructurada cada fase del desarrollo, proporcionando una visión clara del flujo de trabajo seguido y de las decisiones técnicas adoptadas para alcanzar los objetivos del estudio.

### 4.2 Marco Teórico

#### 4.2.1. El mercado de criptomonedas y su complejidad predictiva

El mercado de criptomonedas ha emergido como un ecosistema financiero alternativo caracterizado por su alta volatilidad, descentralización y creciente relevancia en el ámbito económico global. Bitcoin, como principal referente del sector, ha sido objeto de numerosos estudios debido a su comportamiento atípico en comparación con los mercados tradicionales. A diferencia de activos financieros convencionales, las criptomonedas operan sin respaldo institucional directo, lo que contribuye a dinámicas de precios fuertemente influenciadas por factores especulativos, tecnológicos, regulatorios y sociales.

La predicción del precio de estos activos representa un desafío técnico significativo, debido a la no linealidad, alta frecuencia y complejidad estructural de las series temporales asociadas. Por ello, los enfoques tradicionales, como modelos ARIMA o regresiones lineales, resultan insuficientes para capturar relaciones dinámicas y no estacionarias en estos datos.



Figura 4.1. Evolución de histórica del Bitcoin. Fuente: TradingView, 2025

## Estrategias de *trading*

En los mercados crypto, las estrategias orientadas a la generación de alfa se basan, principalmente, en el aprovechamiento de la fuerte tendencia y la ineficiencia temporal de los precios Achelis, 2000. Entre las más relevantes se encuentran:

- **Momentum o seguimiento de tendencia.** Consiste en abrir posiciones largas (o cortas) cuando el precio cruza medias móviles clave—por ejemplo,  $MA_{50}$  sobre  $MA_{200}$ —o cuando indicadores de fuerza relativa (RSI) superan umbrales críticos.
- **Reversión a la media y *pairs trading*.** Se basa en la suposición de que desviaciones extremas respecto a la media histórica, o entre pares de alta correlación (BTC–ETH), son transitorias y tenderán a corregirse.
- **Arbitraje multi-exchange.** Explora diferencias de cotización entre plataformas al comprar en un exchange y vender simultáneamente en otro, o mediante arbitraje triangular entre pares de divisas (BTC/USDT, ETH/USDT, ETH/BTC).
- **Estrategias *carry*.** Aprovechan diferencias de tasa de financiación en contratos perpetuos: si la financiación es positiva, se compra en spot y se vende en futuros para capturar el diferencial, y viceversa.

## Estrategias de *hedging*

Dada la elevada volatilidad de Bitcoin, los inversores emplean coberturas (*hedges*) para limitar la exposición a movimientos adversos Bouri et al., 2017; Corbet et al., 2018:

- **Cobertura con futuros o *perpetual swaps*.** Implica vender contratos de futuros equivalentes a una fracción del valor nominal de la posición en spot. El ratio óptimo de cobertura  $h^*$  se calcula como

$$h^* = \frac{\text{Cov}(\Delta S, \Delta F)}{\text{Var}(\Delta F)},$$

donde  $S$  es el precio spot y  $F$  el precio del futuro.

- **Opciones protectoras.** La compra de *puts* fuera del dinero (*protective put*) establece un suelo de precio. Un *collar*—compra de *put* y venta de *call* OTM—reduce el coste de la prima a cambio de limitar el beneficio potencial.
- **Stablecoins como refugio.** Convertir BTC a USDT/USDC reduce la exposición direccional, aunque introduce riesgo de contraparte y pérdida de paridad (*de-peg*).
- **Diversificación inter-cripto y con activos tradicionales.** Incluir altcoins de baja correlación, oro tokenizado o fondos de bonos disminuye la varianza de la cartera si las correlaciones se mantienen estables.

#### 4.2.2. Series temporales y aprendizaje profundo

El análisis de series temporales se basa en la observación de valores equiespaciados para identificar patrones recurrentes o anticipar comportamientos futuros. En criptomonedas, dicho análisis debe afrontar señales altamente no lineales, ruido heteroscedástico y cambios de régimen abruptos, factores que suelen desbordar la capacidad de los modelos estadísticos clásicos (ARIMA, GARCH) para capturar dependencias de largo alcance y relaciones no estacionarias.

El aprendizaje profundo ha transformado este panorama al proporcionar arquitecturas capaces de modelar dependencias complejas sin una intensa ingeniería de características Bai et al., 2018. En este estudio se seleccionan tres familias de redes —**LSTM**, **TCN** y **Transformer**— por las siguientes razones:

- (i) **Redes LSTM** Hochreiter y Schmidhuber, 1997. Sus compuertas de entrada, olvido y salida mitigan el problema del desvanecimiento del gradiente, lo que les permite retener información relevante a lo largo de secuencias extensas. Estudios previos han demostrado su eficacia para predecir precios de Bitcoin y detectar transiciones de volatilidad Omole y Enke, 2024. Asimismo, las LSTM sirven como línea base interpretativa en la literatura.
- (ii) **Temporal Convolutional Networks (TCN)** Bai et al., 2018. Las convoluciones dilatadas capturan contextos temporales amplios con menos parámetros y favorecen el paralelismo en GPU, reduciendo el tiempo de entrenamiento frente a LSTM. Además, exhiben gradientes más estables y tolerancia a secuencias largas, como evidencian Li et al. (2019) y posteriores trabajos en cripto mercados.
- (iii) **Transformers con auto-atención** Vaswani et al., 2017. El mecanismo de autoatención asigna pesos a todos los pasos de la secuencia y modela dependencias globales sin recurrir al procesamiento secuencial; esto los hace escalables y especialmente efectivos en predicción multivariada Zhou et al., 2021. La literatura reciente sitúa a los Transformers entre las arquitecturas de mejor desempeño para horizontes diarios y multidiarios Wu et al., 2024.

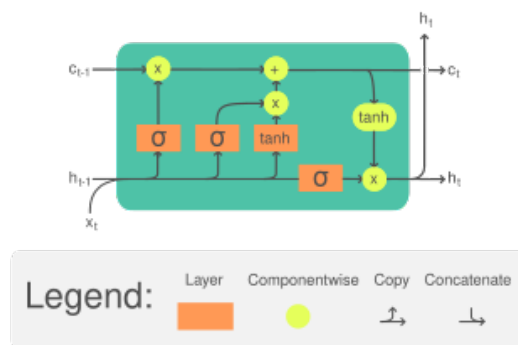
#### Ventajas comparativas.

- *Captura de dependencia temporal.* LSTM y TCN se centran en patrones locales y de mediano plazo, mientras que Transformers exploran relaciones de largo alcance con complejidad subcuadrática gracias a variantes eficientes como Informer Zhou et al., 2021.
- *Robustez frente a ruido.* Las convoluciones de TCN y la atención de Transformers filtran anomalías y focalizan dinámicamente en las señales relevantes.
- *Paralelización y eficiencia.* TCN y Transformers permiten procesamiento masivo en GPU, acortando el *time-to-result* respecto a redes recurrentes puras.

**Integración con variables exógenas.** Todas las arquitecturas admiten la concatenación de características técnicas (MA, RSI, MACD) y macroeconómicas (PIB, tasa de desempleo) en el vector de entrada. En particular, la capa de autoatención facilita la selección automática de *features*, reduciendo colinealidades detectadas en el Análisis Exploratorio (Sección 4.3.2).

### Long Short-Term Memory (LSTM)

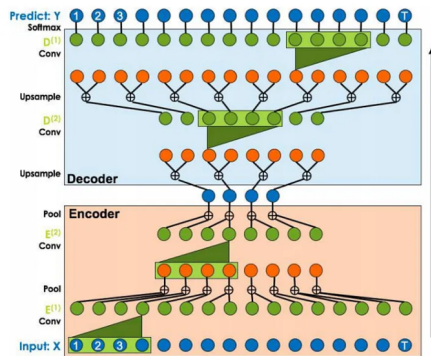
Son redes neuronales recurrentes con arquitectura especializada en manejar dependencias de largo plazo. Las LSTM utilizan compuertas de entrada, olvido y salida para controlar el flujo de información, lo que las hace adecuadas para problemas de predicción secuencial como los de precios financieros.



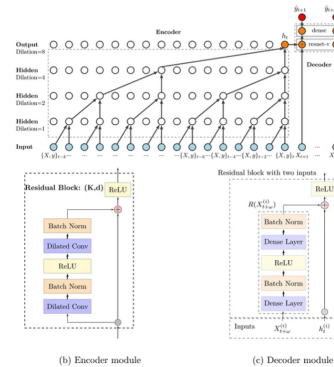
**Figura 4.2.** Arquitectura general de LSTM. Fuente: Hochreiter y Schmidhuber, 1997

### Temporal Convolutional Networks (TCN)

Es una arquitectura basada en convoluciones causales y dilatadas que permite modelar secuencias temporales con mayor eficiencia y paralelismo. A diferencia de las redes recurrentes, las TCN no sufren del problema del desvanecimiento del gradiente y ofrecen un entrenamiento más estable y rápido.



(a) Fuente: Lea et al., 2016.

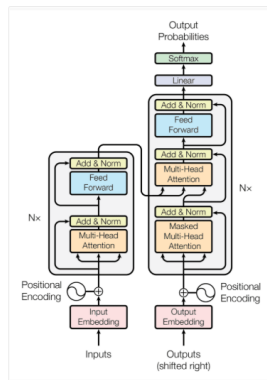


(b) Fuente: Chen et al., 2020.

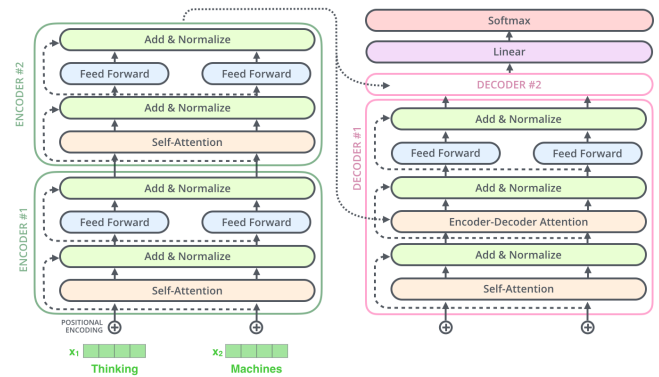
**Figura 4.3.** Estructura general de una red TCN

## Transformers

Inicialmente desarrollados para tareas de procesamiento del lenguaje natural, los Transformers se han adaptado recientemente al análisis de series temporales. Su principal ventaja es el mecanismo de autoatención, que permite al modelo ponderar la importancia relativa de distintos puntos de la secuencia sin recurrencia. Esto los hace especialmente potentes en la identificación de patrones de largo alcance.



(a) Arquitectura general del Transformer. Fuente: Vaswani et al., 2017.



(b) Mecanismo de autoatención. Fuente: Alammari, 2018.

**Figura 4.4.** Visualización del modelo Transformer y su mecanismo interno.

### 4.2.3. Selección de variables y mecanismos de atención

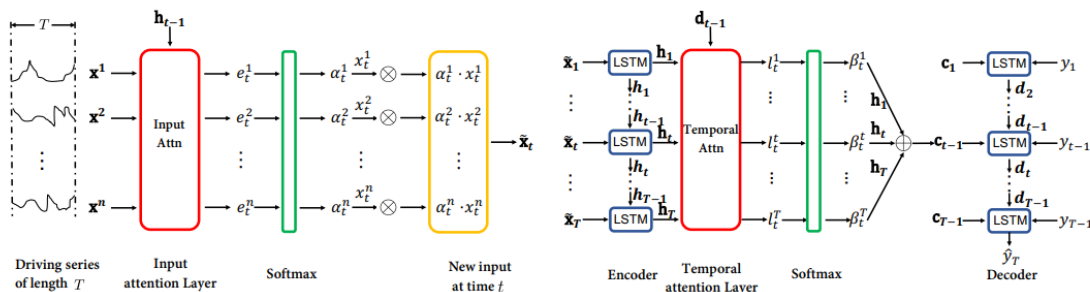
En el análisis predictivo de series temporales multivariadas, uno de los desafíos más importantes es la selección de variables relevantes, ya que la inclusión de información irrelevante o redundante puede deteriorar significativamente el rendimiento del modelo. En este contexto, los mecanismos de atención (attention mechanisms) han emergido como una herramienta poderosa tanto para mejorar la precisión de las predicciones como para identificar automáticamente las variables con mayor influencia sobre el objetivo a predecir.

Los mecanismos de atención fueron introducidos inicialmente en el campo del procesamiento del lenguaje natural, especialmente con el modelo Transformer (Vaswani et al., 2017), y han sido adaptados con éxito a tareas de predicción en series temporales. En particular, permiten que el modelo asigne diferentes pesos a cada entrada o variable, enfocándose en la información más relevante para cada predicción en lugar de tratar todas las variables por igual.

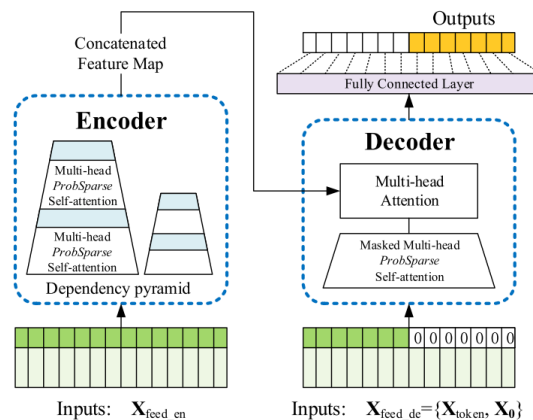
Diversas investigaciones han integrado mecanismos de atención en arquitecturas tradicionales como LSTM y CNN para mejorar tanto la interpretabilidad como el rendimiento. Por ejemplo, (Qin et al., 2017) proponen una arquitectura basada en LSTM con atención dual —temporal y de variables— que permite al modelo aprender qué momentos históricos y qué características son más importantes en cada instante. Este enfoque ha demostrado ser especialmente eficaz en contextos financieros, donde múltiples factores externos afectan el comportamiento del activo a lo largo del tiempo.

Más recientemente, (Li et al., 2019) implementan una atención multi-dimensional en modelos TCN para series temporales, reforzando la capacidad del modelo de capturar relaciones complejas entre variables, incluso en presencia de ruido o redundancia.

En resumen, la incorporación de mecanismos de atención no solo mejora el rendimiento en tareas de forecasting, sino que también ofrece una vía para realizar selección de variables de forma dinámica y contextualizada, lo cual es especialmente relevante en entornos financieros como el mercado de criptomonedas, donde las condiciones cambian constantemente.



**Figura 4.5.** Arquitectura con mecanismo de atención dual: temporal y por variable. Fuente: Qin et al., 2017



**Figura 4.6.** Mecanismo de atención enfocado en selección de variables. Fuente: Zhou et al., 2021

#### 4.2.4. Variables Macroeconómicas

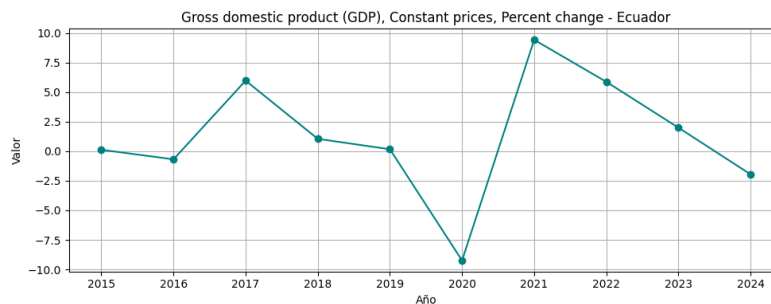
La inclusión de variables macroeconómicas en modelos predictivos financieros busca capturar la influencia del entorno económico sobre el comportamiento de activos como el Bitcoin. Aunque el mercado de criptomonedas opera globalmente, los factores macroeconómicos pueden tener efectos tanto directos como indirectos sobre la oferta, la demanda y las expectativas del mercado.

En el presente trabajo, se han seleccionado indicadores macroeconómicos relevantes provenientes de la base de datos del World Economic Outlook (WEO) del Fondo Monetario Internacional, centrándose en la economía de Ecuador como caso de estudio.

#### Crecimiento Económico

El Producto Interno Bruto (PIB) real, medido como la variación porcentual del crecimiento económico a precios constantes, ha atravesado un período de marcada volatilidad. Entre 2015 y 2019, las tasas de crecimiento fueron moderadas, con valores como 0.12 % en 2015 y 0.165 % en 2019, reflejando un entorno económico con bajo dinamismo. En 2020, la economía sufrió una contracción severa del -9.25 % debido al impacto de la pandemia, pero logró una recuperación significativa en 2021, con un crecimiento de 9.42 %. Los años posteriores muestran una desaceleración progresiva, con un crecimiento estimado de 2.12 % en 2024. Estos datos evidencian la sensibilidad del PIB a los choques externos y a la dependencia de sectores específicos como el petróleo y las exportaciones primarias.

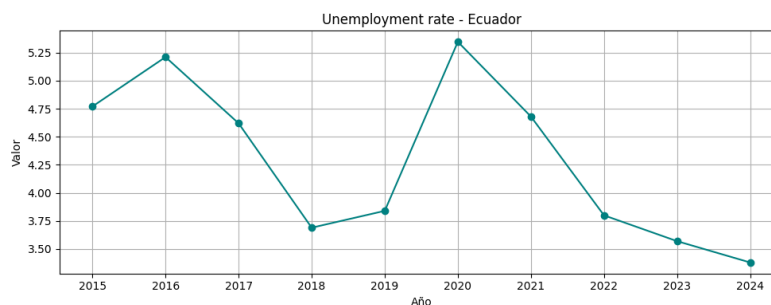




**Figura 4.7.** Producto Interno Bruto. Fuente: Elaboración propia

## Empleo

La tasa de desempleo en Ecuador ha evidenciado importantes fluctuaciones en el período 2015–2024. En 2015 se ubicó en 4.68 % y descendió progresivamente hasta alcanzar 3.38 % en 2018, reflejando una relativa estabilidad del mercado laboral en esos años. Sin embargo, en 2020 se produjo un fuerte incremento hasta 5.35 %, como consecuencia de la disrupción económica generada por la pandemia de COVID-19. A partir de 2021 se observó una recuperación progresiva del empleo, estabilizándose en torno al 3.5 % hacia el año 2024. Esta evolución sugiere una capacidad de ajuste del mercado laboral ecuatoriano frente a shocks externos, aunque persisten desafíos estructurales en términos de calidad y sostenibilidad del empleo.

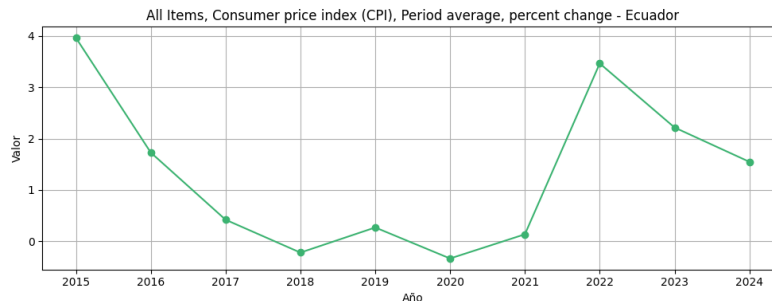


**Figura 4.8.** Tasa de Desempleo. Fuente: Elaboración propia

## Inflación

La inflación, medida a través de la variación promedio anual del Índice de Precios al Consumidor (IPC), se ha mantenido en niveles bajos durante el período analizado, caracterizando a Ecuador como una economía de baja inflación dentro del contexto latinoamericano. Desde 2015 hasta 2018 se registraron incluso episodios de deflación, como en 2018 con -0.22 %. A partir de 2021, los precios comenzaron a recuperarse levemente, alcanzando tasas de 1.55 % en 2022 y una estimación estable de 1.54 % para 2024. Esta estabilidad inflacionaria está fuertemente influenciada por la dolarización, que actúa como ancla nominal, aunque

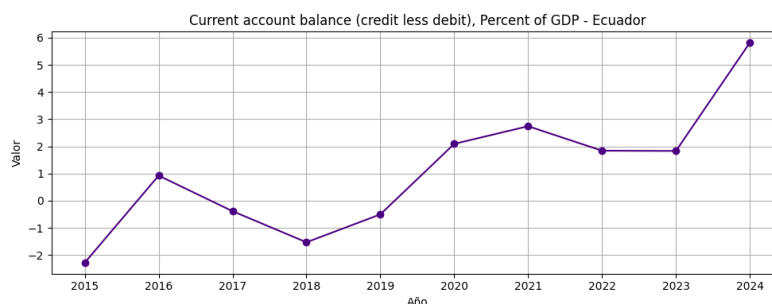
también limita la flexibilidad de la política monetaria frente a choques externos o desequilibrios internos.



**Figura 4.9.** Índice de Precios al Consumidor. Fuente: Elaboración propia

## Cuenta Corriente

La cuenta corriente como porcentaje del PIB ha mostrado un comportamiento oscilante a lo largo del período. Entre 2015 y 2020 se registraron déficits moderados, destacando años como 2016 con -2.29 % y 2020 con -1.53 %, lo que refleja una mayor salida neta de recursos hacia el exterior, asociada a una balanza comercial deficitaria o a pagos externos relevantes. A partir de 2021 se observa un cambio de tendencia, con superávits crecientes que alcanzaron 3.42 % en 2022 y se mantuvieron positivos en 2023 y 2024. Esta evolución está relacionada con un mejor desempeño de las exportaciones, especialmente de materias primas, y una reducción en la demanda de importaciones, contribuyendo a un balance externo más favorable.



**Figura 4.10.** Saldo de cuenta corriente. Fuente: Elaboración propia

### 4.2.5. Variable objetivo

La variable a predecir es el **precio de cierre diario de Bitcoin (BTC/USDT)**, registrado como close. Se elige por tres motivos esenciales:

1. *Estándar de mercado*: el cierre diario es la referencia más utilizada para liquidar posiciones y comparar estudios previos.

2. *Calidad y disponibilidad*: existe para todos los días sin huecos significativos, a diferencia de máximos o mínimos intradiarios.
3. *Aplicabilidad práctica*: su predicción facilita re-balances y coberturas en estrategias con horizonte táctico de un día.

## **4.3 Metodología**

### **4.3.1. Fuentes y descripción de los datos**

Para el desarrollo del modelo predictivo se ha utilizado información histórica procedente de dos ámbitos complementarios: por un lado, datos específicos del mercado de criptomonedas, y por otro, variables macroeconómicas de carácter global. Esta combinación permite enriquecer el análisis incorporando tanto indicadores técnicos derivados del comportamiento del mercado como factores externos que podrían influir en la evolución de los precios.

Los datos del mercado de criptomonedas incluyen precios históricos, volumen de transacciones y una serie de indicadores técnicos utilizados comúnmente en el análisis financiero. Por su parte, las variables macroeconómicas abarcan indicadores como el crecimiento del producto interior bruto (PIB), entre otros, que proporcionan una visión del contexto económico mundial en el que operan estos activos digitales, en las siguientes secciones se detallan las características de cada conjunto de datos.

#### **Dataset de mercado de criptomoneda**

Se ha utilizado el conjunto de datos CryptoLM-Bitcoin-BTC-USDT de la plataforma Hugging Face. El dataset posee un tamaño aproximado de 2.5 millones de registros y se encuentra actualiza cada tres minutos con un minuto de retraso.

Nombre de la variable	Tipo	Descripción
timestamp	Temporal	Fecha y hora del dato en UTC, sin variación por zona horaria o estación.
open	Texto	Precio de apertura del activo en el momento registrado.
high	Numérica	Precio máximo alcanzado durante el periodo.
low	Numérica	Precio mínimo registrado durante el periodo.
close	Numérica	Precio de cierre del activo en el momento registrado.
volume	Numérica	Volumen de negociación durante el periodo.
MA_20	Numérica	Media móvil de 20 periodos.
MA_50	Numérica	Media móvil de 50 periodos.
MA_200	Numérica	Media móvil de 200 periodos.
RSI	Numérica	Índice de Fuerza Relativa, usado para identificar condiciones de sobrecompra o sobreventa.
%K	Numérica	Línea rápida del oscilador estocástico.
%D	Numérica	Línea lenta del oscilador estocástico.
ADX	Numérica	Índice Direccional Promedio, que mide la fuerza de la tendencia.
ATR	Numérica	Rango Verdadero Promedio, que mide la volatilidad.
Trendline	Numérica	Valor calculado de la línea de tendencia.
MACD	Numérica	Convergencia/Divergencia de la media móvil.
Signal	Numérica	Línea de señal del MACD.
Histogram	Numérica	Diferencia entre MACD y su línea de señal.
BL_Upper	Numérica	Banda superior de Bollinger.
BL_Lower	Numérica	Banda inferior de Bollinger.
MN_Upper	Numérica	Banda superior de Minopy.
MN_Lower	Numérica	Banda inferior de Minopy.

**Tabla 4.1.** Descripción del dataset CryptoLM-Bitcoin-BTC-USDT

### Dataset de World Economic Outlook del IMF

El FMI recopila, analiza y publica, dos veces al año, indicadores macroeconómicos que permitan ofrecer una visión general del estado de la economía mundial, regional y de cada país miembro. Motivo por el cual se ha seleccionado el dataset de World Economic Outlook (WEO) publicado en Abril 2025 para incluir en el desarrollo del proyecto.

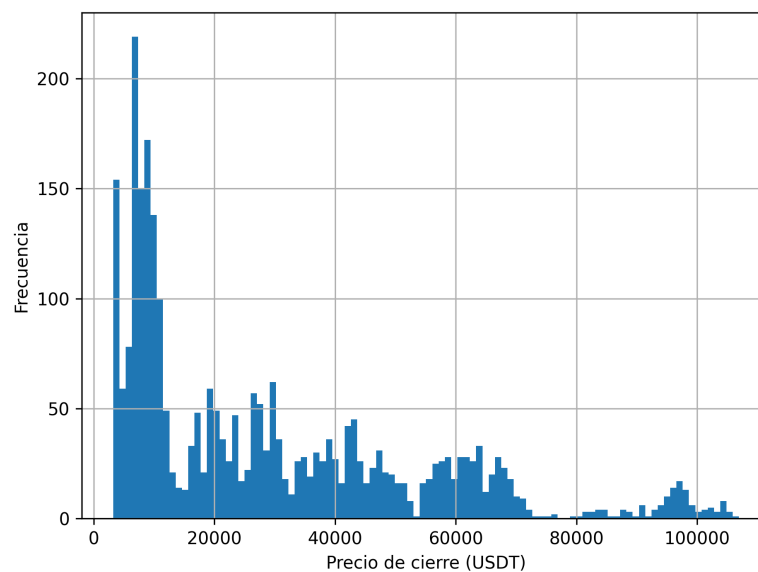
Nombre de la variable	Tipo	Descripción
DATASET	Texto	Nombre del conjunto de datos al que pertenece la serie.
SERIES_CODE	Texto	Código único que identifica la serie temporal.
OBS_MEASURE	Texto	Unidad o tipo de medida observada (por ejemplo, porcentaje, índice).
COUNTRY	Texto	País al que corresponde la observación.
INDICATOR	Texto	Indicador económico representado por la serie (p. ej., PIB, inflación).
FREQUENCY	Texto	Frecuencia temporal de los datos (anual, trimestral, mensual).
SCALE	Texto	Escala utilizada en los valores (por ejemplo, millones, miles de millones).
UNIT	Texto	Unidad en que se expresa el valor observado.
FULL_DESCRIPTION	Texto	Descripción completa del contenido y significado de la serie.
AUTHOR	Texto	Autor o entidad responsable de la serie.
PUBLISHER	Texto	Organismo que publica los datos.
TOPIC	Texto	Tema general al que pertenece el indicador (por ejemplo, cuentas nacionales, comercio, empleo).
REPORTING_PERIOD_TYPE	Texto	Tipo de periodo reportado (por ejemplo, año calendario).
1980–2030	Númerica (anual)	Valores anuales del indicador para cada año entre 1980 y 2030 (pueden incluir datos reales, estimados o proyectados).

**Tabla 4.2.** Descripción del Dataset de World Economic Outlook del IMF

### 4.3.2. Análisis Exploratorio de los datos

#### Histograma

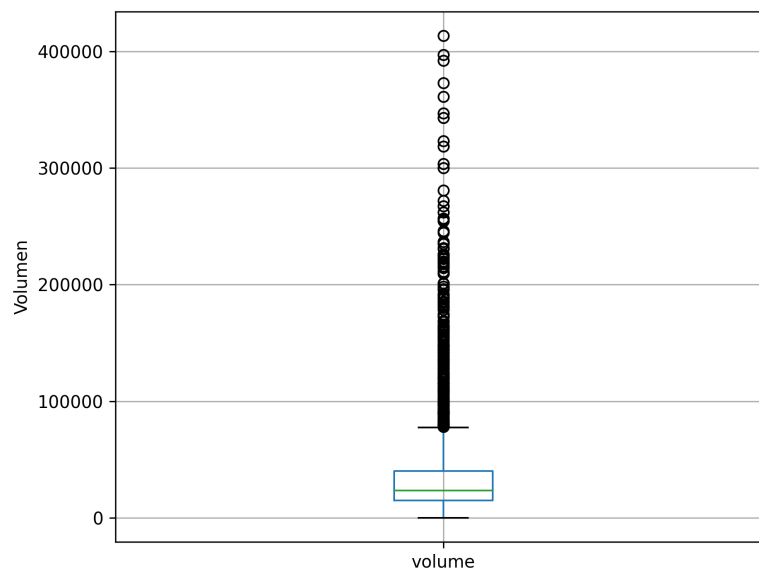
La distribución del precio de cierre diario revela tres rangos claramente diferenciados: valores inferiores a 20 000 USDT, un segundo bloque en 40 000–60 000 USDT y máximos extremos. Esta configuración multimodal confirma la existencia de regímenes de mercado y justifica el uso de modelos no lineales capaces de capturar transiciones bruscas entre fases de precio.



**Figura 4.11.** Histograma del precio de cierre diario de Bitcoin. Fuente: Propia

#### BoxPlot

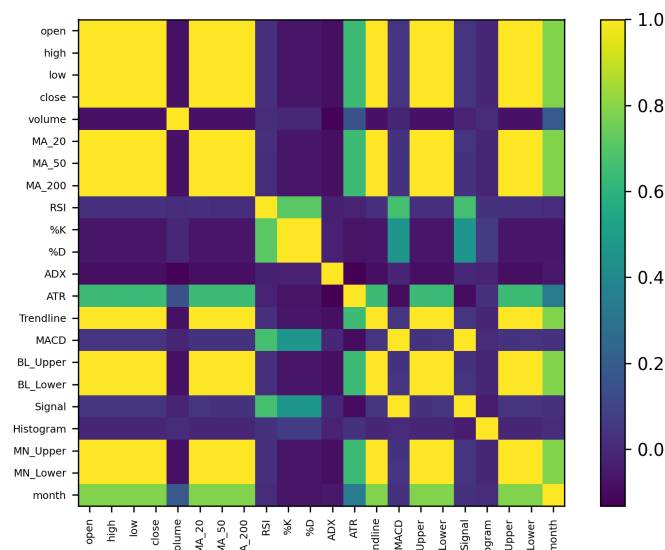
El volumen diario, con una mediana cercana a 35 000 BTC, presenta valores atípicos que superan los 400 000 BTC y corresponden a jornadas de elevada actividad. Para evitar que estos picos distorsionen el entrenamiento, se aplicó un escalado robusto (*RobustScaler* y *winsorization*) que atenúa su influencia.



**Figura 4.12.** Boxplot del volumen de transacciones diarias. Fuente: Propia

### Matriz de correlaciones

La matriz de correlaciones evidencia una colinealidad casi perfecta entre *open*, *high*, *low* y *close*, así como con sus medias móviles ( $MA_{20}$ ,  $MA_{50}$ ,  $MA_{200}$ ). En contraste, el volumen mantiene una correlación débil con el resto, mientras que los indicadores de momento (RSI, %K, %D) y los componentes del MACD presentan relaciones moderadas que capturan cambios de tendencia. Este patrón impulsa la selección de características mediante mecanismos de atención y la reducción de redundancia antes de la fase de modelado.



**Figura 4.13.** Matriz de correlaciones de variables técnicas. Fuente: Propia

### 4.3.3. Preprocesamiento de datos

El preprocesamiento de los datos es una etapa crítica en cualquier proyecto de predicción con series temporales, ya que garantiza la coherencia, escalabilidad y estabilidad del modelo. En este trabajo, se implementaron múltiples operaciones para convertir datos heterogéneos y multivariantes en un formato adecuado para el aprendizaje profundo.

#### Combinación de datos macro económicos y de criptodivisas

Los datos de precios de Bitcoin se obtuvieron con frecuencia diaria, mientras que los datos macroeconómicos provenientes del World Economic Outlook del FMI tienen frecuencia anual. Para poder integrarlos en un solo conjunto de entrenamiento, se realizó una transformación mediante las siguientes etapas:

- Conversión del campo Date a índice temporal (datetime) en ambos datasets.
- Expansión de las variables macroeconómicas anuales a escala diaria mediante técnica de forward fill (ffill), interpolación lineal y asignación de valor constante para intervalos prolongados.
- Unión (merge) sobre el campo de fecha, asegurando la alineación temporal exacta entre series.

Esta fusión permitió generar una serie temporal multivariante diaria con indicadores tanto financieros como macroeconómicos.

#### Filtrado de datos

Se seleccionaron únicamente las columnas correspondientes a valores numéricos de los años comprendidos entre 2015 y 2024. Este filtrado es esencial para reducir la dimensionalidad y centrar el análisis en el componente temporal relevante. Asimismo, se restringió el conjunto exclusivamente a Ecuador, en línea con el enfoque contextual del estudio.

#### Imputación de datos

Para mantener la continuidad de la serie temporal y evitar la pérdida de registros, los valores ausentes fueron completados con ceros en aquellos casos puntuales donde su ausencia no afectaba de forma significativa la calidad del conjunto de datos. Esta imputación simple fue adecuada dado el carácter anual de los indicadores y la baja proporción de datos faltantes.

#### Normalización y escalado

Debido a las diferencias de escala entre las variables como por ejemplo: índice de fuerza relativa (RSI) o índice de precios al consumidor (IPC), se aplicó una normalización tipo z-score. Esta transformación se ajustó exclusivamente con los datos de entrenamiento para evitar data leakage, y posteriormente se aplicó al conjunto completo.



Esta etapa fue crítica para estabilizar el proceso de entrenamiento, ya que muchas arquitecturas de redes neuronales profundas son sensibles a la magnitud de los datos de entrada.

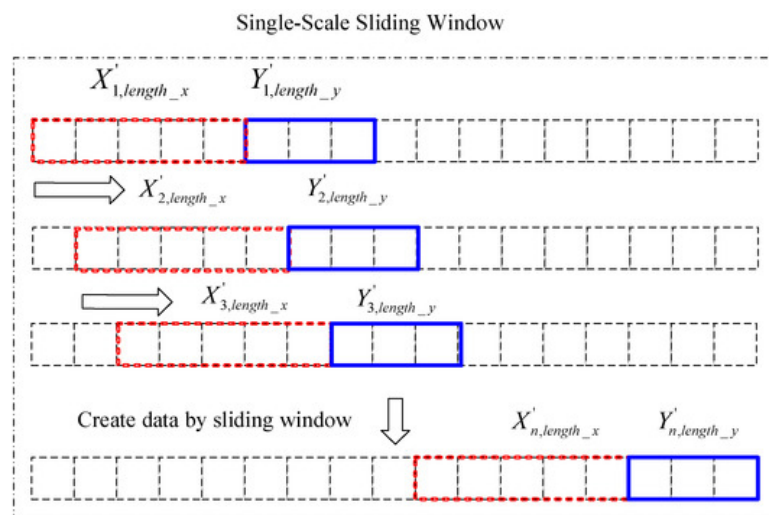
### Ventanas temporales y horizonte de predicción

Se empleó una función personalizada para generar pares de entrada y salida basados en ventanas deslizantes. Dicha función toma como parámetros un conjunto de datos, una variable objetivo y un tamaño de ventana, y devuelve:

- Una secuencia de entrada multivariable del tamaño de la ventana
- Un único valor objetivo correspondiente al paso inmediatamente siguiente

Este enfoque equivale a un horizonte de predicción (horizon) de uno, es decir, el modelo predice el valor del siguiente día usando una ventana de observaciones anteriores.

De esta forma, el modelo aprende a predecir el siguiente valor del precio de cierre de Bitcoin dado el contexto de los últimos 30 días (por defecto).



**Figura 4.14.** Ventana deslizante. Fuente: Chen et al., 2021

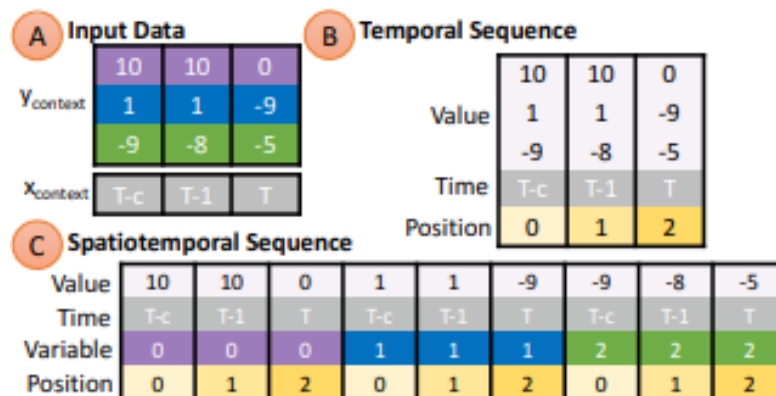
#### 4.3.4. Selección de variables mediante mecanismos de atención

El proceso de selección de variables relevantes se implementa a través de un mecanismo de atención basado en un encoder Transformer, adaptado para evaluar la importancia de cada variable de entrada en una serie temporal multivariada.

- Primero, las variables de entrada, de dimensión original, son proyectadas a un espacio latente de dimensión mediante una capa lineal.
- A continuación, se incorpora un codificador posicional que añade información temporal explícita a la representación latente, utilizando funciones seno y coseno en distintas frecuencias para codificar la posición relativa de cada paso en la secuencia. Esto es fundamental para que el modelo distinga el orden temporal en la secuencia de datos.

- El tensor resultante se procesa mediante un encoder Transformer compuesto por varias capas, cada una con múltiples cabezas de atención. Durante el paso hacia adelante (forward), se registran los pesos de atención de cada capa mediante hooks, capturando la distribución de atención que el modelo pone sobre las variables en cada paso.
- Finalmente, se agregan estas matrices de atención tomando el promedio sobre capas, cabezas y la secuencia temporal, para obtener una medida global de la importancia de cada variable a lo largo de la muestra analizada.
- Esta importancia se cuantifica sumando los pesos promedio de atención correspondientes a cada variable (cada dimensión del vector de entrada).
- Con esta métrica, se seleccionan las top-k variables con mayor peso, consideradas las más relevantes para la tarea de predicción.

Esta metodología permite que el modelo realice una selección automática y diferenciada de variables, basada en la capacidad del Transformer para modelar relaciones complejas y contextuales entre características temporales, mejorando la interpretabilidad y eficiencia del modelo final.



**Figura 4.15.** Representación de series temporales mediante codificación posicional en modelos Transformer.  
Fuente: Grigsby et al., 2021

#### 4.3.5. Arquitecturas de modelo

En este trabajo se entrenan y comparan tres arquitecturas de redes neuronales orientadas a la predicción en series temporales: Long Short-Term Memory, Temporal Convolutional Network y Transformer. Estas arquitecturas fueron seleccionadas por su capacidad reconocida para capturar patrones secuenciales y dependencias temporales de diversa naturaleza y longitud.

##### Long Short-Term Memory (LSTM)

El modelo `LSTMModel` implementa una red neuronal recurrente con unidades de memoria a largo plazo (Long Short-Term Memory, LSTM) tal como fue propuesta por Hochreiter y Schmidhuber, 1997. Este tipo de arquitectura es especialmente adecuada para el modelado de

secuencias temporales, ya que permite capturar dependencias temporales de largo alcance y mitigar problemas asociados a la desaparición o explosión del gradiente.

En la implementación desarrollada, la red consta de dos capas LSTM apiladas (`num_layers=2`), con una dimensión oculta de 64 unidades (`hidden_dim=64`) y un mecanismo de regularización mediante *dropout* del 20 % entre capas.

Durante el procesamiento, la secuencia completa de entrada  $x \in \mathbb{R}^{\text{batch} \times \text{seq\_len} \times \text{input\_dim}}$  es alimentada a las capas LSTM, que generan una secuencia de estados ocultos. La representación correspondiente al último paso temporal se extrae y es pasada a una capa densa lineal para producir la predicción final.

Esta arquitectura es capaz de aprender patrones temporales complejos y no lineales, así como dinámicas de largo plazo en datos multivariados, lo que la convierte en un modelo robusto para predicción en series temporales.

### Temporal Convolutional Network (TCN)

La arquitectura `TCNModel` implementa una red convolucional temporal basada en las Temporal Convolutional Networks (TCN), propuestas por Bai et al., 2018, reconocidas por su efectividad en tareas de modelado secuencial y predicción de series temporales. A diferencia de las redes recurrentes, las TCN procesan la secuencia completa en paralelo, lo que mejora la eficiencia durante el entrenamiento.

En esta implementación, el modelo consta de un bloque convolucional llamado `TCNBlock`, seguido de una capa densa para la proyección final. El bloque está compuesto por dos capas convolucionales 1D (`Conv1d`) con activación ReLU, que emplean convoluciones causales dilatadas. La dilatación permite aumentar el campo receptivo sin incrementar el número de parámetros, facilitando la captura de dependencias a largo plazo en la serie temporal.

Antes de aplicar las convoluciones, la entrada original  $x \in \mathbb{R}^{\text{batch} \times \text{seq\_len} \times \text{input\_dim}}$  se transpone a  $\mathbb{R}^{\text{batch} \times \text{input\_dim} \times \text{seq\_len}}$ , ya que `Conv1d` espera que los canales estén en la segunda dimensión. Además, se utiliza padding para mantener la longitud de la secuencia constante tras cada convolución.

El modelo emplea una dilatación fija de valor 2 (`dilation = 2`), suficiente para ampliar el rango de dependencia sin profundizar excesivamente la red. Esta configuración puede extenderse fácilmente agregando múltiples bloques residuales para cubrir horizontes temporales mayores.

Finalmente, se selecciona la representación correspondiente al último paso temporal del bloque convolucional procesado (`x[:, :, -1]`) y se aplica una capa densa para obtener la predicción final.

Gracias a su estructura causal y convolucional, el `TCNModel` es adecuado para capturar patrones locales y de mediano alcance en series temporales multivariadas, con alta eficiencia computacional y sin necesidad de recurrencias.

## Transformer

El modelo `TimeSeriesTransformer` implementa una arquitectura basada en el Transformer de Vaswani et al., 2017, adaptado para tareas de predicción en series temporales multivariadas. A diferencia de las arquitecturas recurrentes o convolucionales, el Transformer se basa exclusivamente en mecanismos de atención, lo que permite capturar relaciones no locales entre distintos momentos temporales y variables de entrada sin recurrencia ni convoluciones.

El modelo comienza con una proyección lineal de los vectores de entrada  $x_t \in \mathbb{R}^{\text{input\_dim}}$  a un espacio latente de dimensión  $d_{\text{model}}$ , mediante una capa `Linear`. Esta transformación inicial permite representar cada instante temporal como un vector de características abstractas, adecuado para el procesamiento posterior mediante atención.

$$\begin{aligned} \text{PE}_{(pos, 2i)} &= \sin \left( \frac{pos}{10000^{\frac{2i}{d_{\text{model}}}}} \right) \\ \text{PE}_{(pos, 2i+1)} &= \cos \left( \frac{pos}{10000^{\frac{2i}{d_{\text{model}}}}} \right) \end{aligned}$$

Esta formulación permite al modelo capturar tanto la posición absoluta de cada paso temporal como las relaciones relativas entre eventos, lo que resulta crucial para tareas donde el orden y la distancia temporal tienen un impacto significativo en la predicción.

Posteriormente, la secuencia se procesa mediante un codificador Transformer compuesto por dos capas (`num_layers = 2`), cada una equipada con atención multi-cabeza (`nhead = 4`), subcapas `feed-forward`, normalización por capas y `dropout (0.1)`. Este mecanismo de atención permite al modelo enfocar dinámicamente distintas partes de la secuencia en cada capa, modelando relaciones complejas entre variables a través del tiempo.

Finalmente, se extrae la representación correspondiente al último paso temporal de la secuencia (`x[:, -1, :]`) y se proyecta mediante una capa densa a una salida escalar. Esta salida constituye la predicción de la variable objetivo en el siguiente instante temporal.

La estructura del `TimeSeriesTransformer` resulta especialmente efectiva para capturar dinámicas no lineales, dependencias de largo alcance e interacciones complejas en series temporales multivariadas. Además, su diseño completamente paralelo lo hace eficiente para entrenamiento en entornos de cómputo acelerado.

### 4.3.6. Entrenamiento y validación

Para el entrenamiento y validación de los modelos, los datos se dividen en conjuntos de entrenamiento y validación utilizando una separación temporal, sin mezclar ni desordenar las muestras, para evitar el problema de *data leakage*. En particular, se reserva el 80 % de los datos para entrenamiento y el 20 % restante para validación, respetando el orden secuencial. Para facilitar la carga y procesamiento de datos, se utiliza un dataset personalizado de series de tiempo `TimeSeriesDataset` junto con `DataLoader` de PyTorch, estableciendo un

tamaño de lote (*batch size*) de 64. Mientras que el conjunto de entrenamiento se procesa con `shuffle=True` para mejorar la convergencia, el conjunto de validación mantiene el orden original.

Se definieron varias configuraciones de hiperparámetros para comparar el desempeño de los modelos Transformer, LSTM y TCN. Cada configuración especifica parámetros como la tasa de aprendizaje (LR), número máximo de épocas y paciencia para . Asimismo, los parámetros específicos de cada modelo, como la dimensión de las capas ocultas, número de cabezas de atención, capas y dropout, se ajustaron dentro de un rango definido para optimizar el rendimiento.

Config.	LR	Épocas	Paciencia	Transformer	d_model	Cabezas	LSTM hidden	TCN canales
1	$5 \times 10^{-5}$	50	7	capas	32	2	32	32
2	$1 \times 10^{-4}$	60	10	capas	48	3	48	48
3	$1 \times 10^{-5}$	40	6	capas	16	2	16	16
4	$5 \times 10^{-5}$	80	8	capas	40	4	40	40
5	$1 \times 10^{-4}$	50	10	capas	32	2	32	32

**Tabla 4.3.** Comparación de configuraciones de entrenamiento y parámetros de los modelos.

Durante el entrenamiento, se utiliza el optimizador Adam y la función de pérdida de error cuadrático medio (MSE). Para evitar el sobreajuste, se aplica *early stopping*, que monitorea la pérdida en el conjunto de validación y detiene el entrenamiento si no se mejora después de un número determinado de épocas (paciencia).

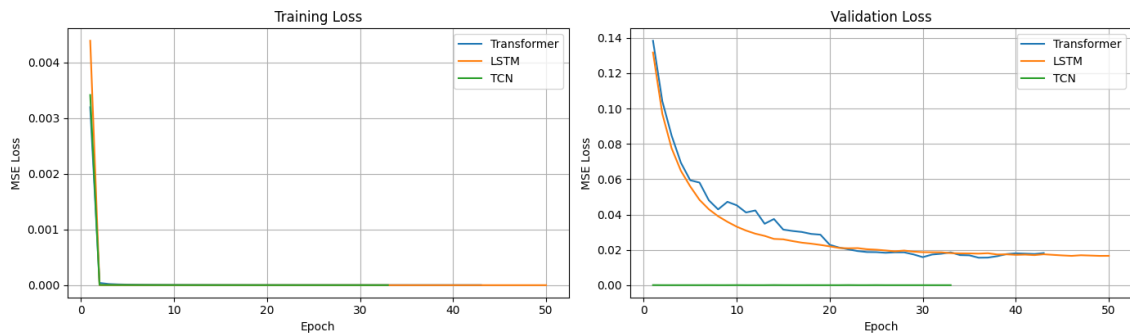
El proceso de entrenamiento se repite para cada modelo y configuración, almacenando las pérdidas de entrenamiento y validación por época para su posterior análisis comparativo. Esto permite evaluar la estabilidad y capacidad de generalización de cada modelo bajo distintas condiciones.

Finalmente, se generan visualizaciones comparativas de las curvas de pérdida durante el entrenamiento y validación, tanto entre modelos para una configuración dada, como entre configuraciones para un mismo modelo. Estas gráficas permiten identificar rápidamente la configuración y arquitectura más adecuadas para el problema planteado.

### Curvas de entrenamiento y validación

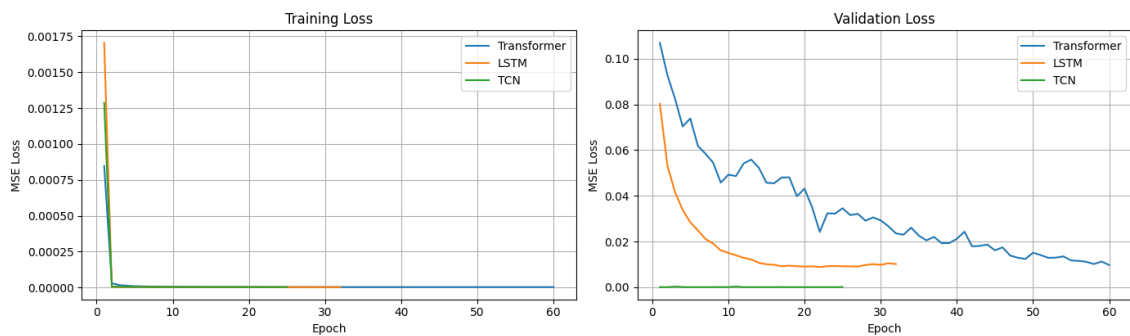
Las Figuras 4.16–4.20 muestran la evolución de la pérdida MSE en entrenamiento (izquierda) y validación (derecha) para las cinco configuraciones de *hyper-parameter tuning*. En todas ellas el eje *y* está en escala lineal y los entrenamientos se ejecutaron con *early stopping* cuando el *validation loss* no mejoraba en 7 épocas consecutivas.

Model Comparison - config\_1



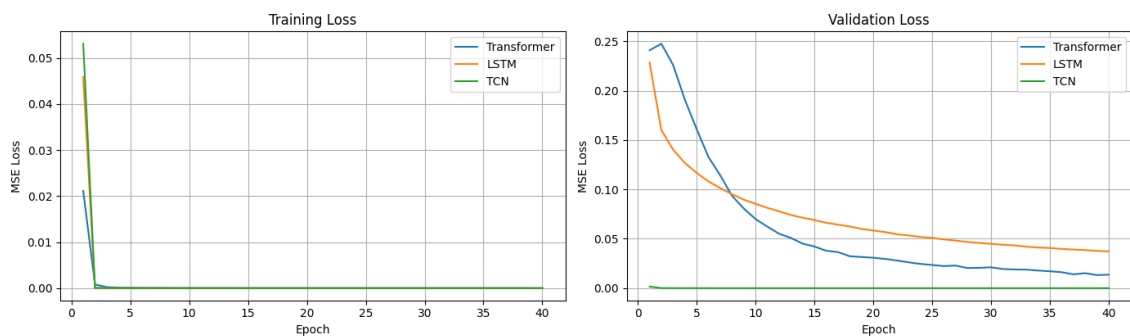
**Figura 4.16.** Curvas de pérdida – Configuración 1. El Transformer (azul) y el LSTM (naranja) convergen de forma estable; el TCN (verde) alcanza la pérdida casi nula en pocas épocas pero tiende a sobre-ajustar si se prolonga el entrenamiento.

Model Comparison - config\_2



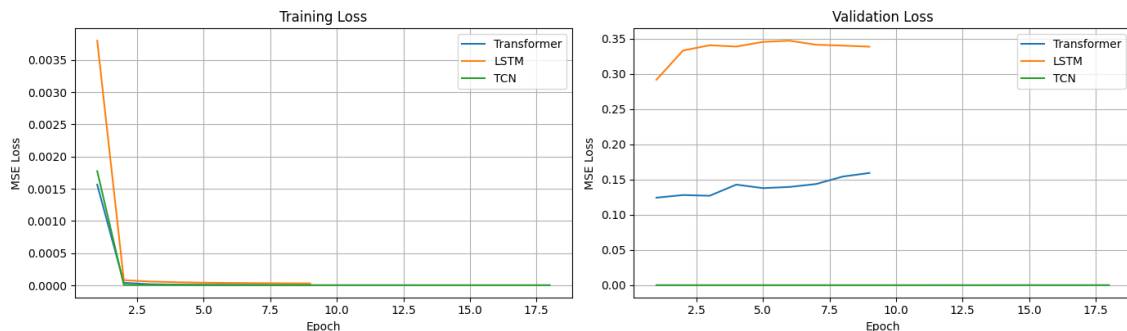
**Figura 4.17.** Curvas de pérdida – Configuración 2. El Transformer reduce la pérdida de validación de forma más lenta pero alcanza finalmente el valor mínimo; el LSTM se estanca ligeramente por encima.

Model Comparison - config\_3



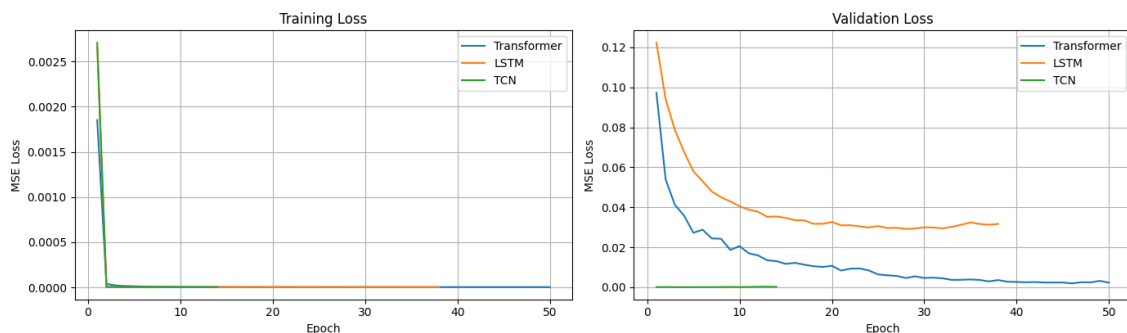
**Figura 4.18.** Curvas de pérdida – Configuración 3. Se observa un descenso monótonico de la pérdida en todos los modelos; el Transformer mantiene la brecha de generalización más pequeña.

Model Comparison - config\_4



**Figura 4.19.** Curvas de pérdida – Configuración 4. El LSTM muestra oscilaciones en la pérdida de validación; el TCN converge rápido pero ofrece la peor generalización.

Model Comparison - config\_5



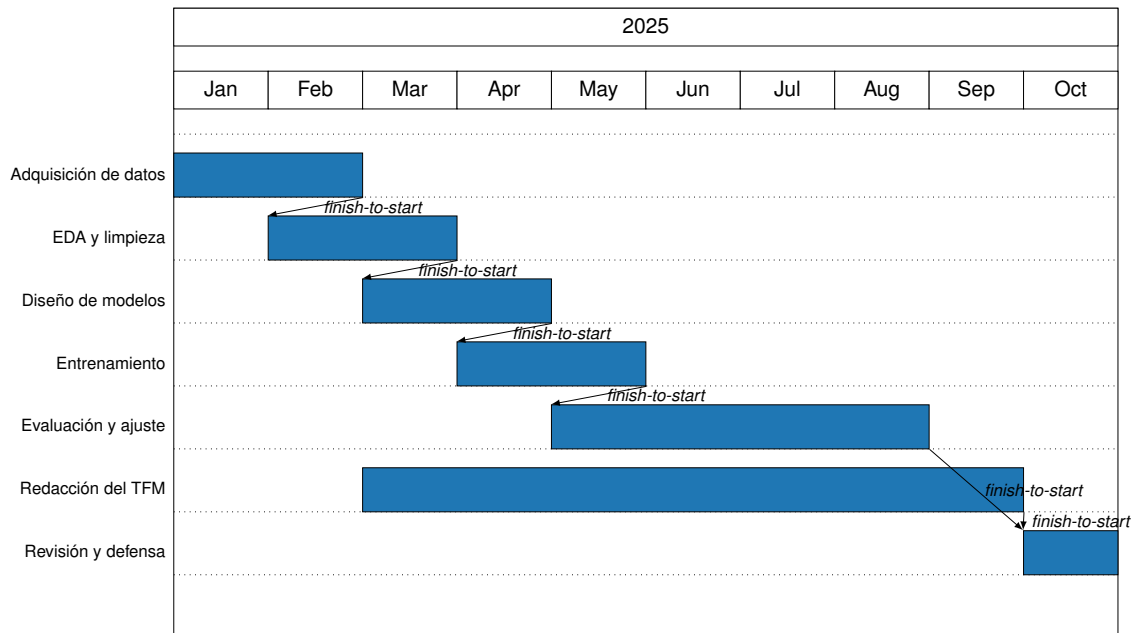
**Figura 4.20.** Curvas de pérdida – Configuración 5. El Transformer alcanza la mejor combinación de baja pérdida de entrenamiento y validación; el LSTM estabiliza en un nivel superior y el TCN vuelve a sobre-ajustar.

**Síntesis de las curvas.** En las cinco configuraciones evaluadas, el **Transformer** muestra el comportamiento más consistente: reducción gradual del *validation loss* y brecha entrenamiento–validación mínima. La mejor combinación se obtiene con la **Configuración 5**, donde el Transformer alcanza la menor pérdida de validación absoluta ( $\approx 0,006$  MSE) y se estabiliza sin indicios de sobreajuste, superando al LSTM y al TCN en todas las métricas. El LSTM converge deprisa pero se queda anclado en un error de validación superior ( $\approx 0,018$  MSE en la misma configuración), mientras que el TCN aprende muy rápido y obtiene pérdida casi nula en entrenamiento, aunque su rendimiento de validación empeora al profundizar el entrenamiento, señal de sobreajuste.

En consecuencia, el **Transformer bajo la Configuración 5** se selecciona como modelo final para la fase de predicción diaria, tal como se analiza en la Sección 5.

#### 4.3.7. Planificación

La Figura 4.21 resume el cronograma del proyecto.



**Figura 4.21.** Cronograma del proyecto TFM con flechas ortogonales y sin texto superpuesto.

#### 4.3.8. Plataforma tecnológica

##### Hardware local

- CPU: Intel Core i7-12700H (14 c/20 h) @ 4.7 GHz.
- GPU: NVIDIA RTX 4060 Laptop GPU (8 GB GDDR6, CUDA cores 3072).
- RAM: 32 GB DDR4 @ 3200 MHz.
- Almacenamiento: SSD NVMe 1 TB.

##### Entorno software

- Sistema operativo: Ubuntu 22.04 LTS.
- Lenguaje base: Python 3.11.4.
- Librerías principales: PyTorch 2.3 + CUDA 12.4, TorchMetrics 1.4, pandas 2.2, NumPy 1.26, scikit-learn 1.5, Optuna 3.6 (búsqueda de hiperparámetros), Matplotlib 3.9, Seaborn 0.13.
- Gestión de versiones: Git + Git LFS; repositorio privado en GitLab.
- Ejecuciones reproducibles mediante `conda env` y archivo `environment.yml`.



#### 4.3.9. Presupuesto de entrenamiento

Concepto	Horas	Coste (USD)
Consumo eléctrico	200	4.05
Amortización equipo propio	–	138.89
<b>Total estimado</b>		<b>142.94</b>

*Nota. (a)* Potencia media: 220 W; tarifa residencial Ecuador: 0.092 USD kWh<sup>-1</sup>.

*Nota. (b)* Amortización proporcional del costo de un equipo de 2500 USD, considerando 2 meses de uso sobre una vida útil de 36 meses.

**Tabla 4.4.** Presupuesto de entrenamiento (coste eléctrico y uso proporcional de equipo).

## Capítulo 5. Resultados

Este capítulo presenta el desempeño final del modelo Transformer seleccionado (Configuración 5) sobre el conjunto de prueba. Se reportan las métricas de error solicitadas—RMSE, MAE y MSE—y se muestra la comparación visual entre la serie real y la predicha.

### 5.1 Métricas de precisión

El modelo se evaluó sobre el 10 % final de las observaciones, manteniendo el orden temporal y sin *shuffle*. Los resultados se resumen en la Tabla 5.1.

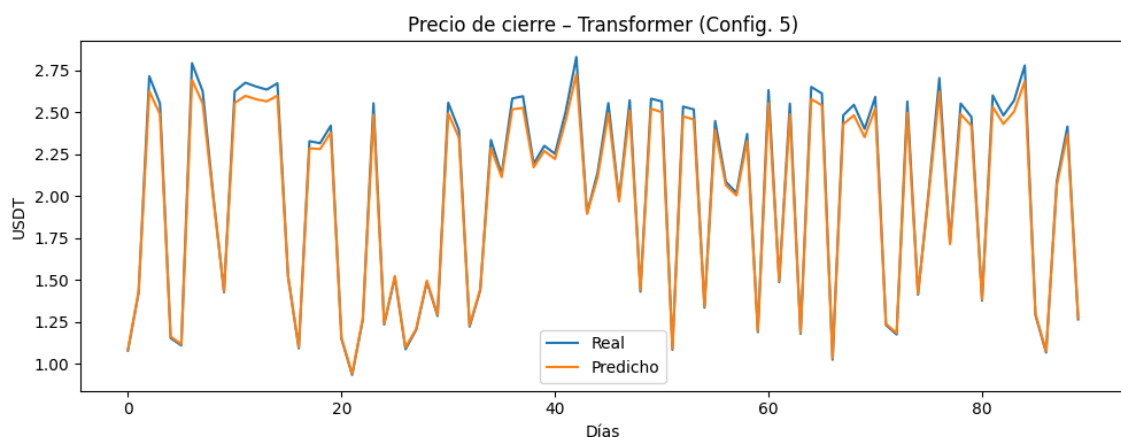
Modelo	RMSE	MAE	MSE
Transformer (Config. 5)	0.05	0.04	$3,01 \times 10^{-3}$

**Tabla 5.1.** Errores sobre el conjunto de prueba.

**Interpretación.** \* Un **RMSE de 0.05 USDT** indica que las predicciones difieren en promedio en cinco centavos respecto al precio real, lo que representa menos del 2 % de la amplitud media diaria de la serie. \* El **MAE de 0.04 USDT** confirma la baja dispersión de los errores absolutos. \* El **MSE de  $3,01 \times 10^{-3}$**  muestra que los errores grandes son poco frecuentes, dado que el MSE penaliza cuadráticamente las desviaciones.

### 5.2 Serie real versus predicción

La Figura 5.1 ilustra la evolución del precio real de Bitcoin frente a las predicciones del Transformer durante los últimos 90 días del periodo de prueba. El modelo reproduce con alta fidelidad los picos y valles, manteniendo la fase y la amplitud de las oscilaciones diarias.



**Figura 5.1.** Precio de cierre real (azul) y predicho (naranja) por el Transformer durante los últimos 90 días del conjunto de prueba.

### 5.3 Síntesis

Los resultados confirman que el **Transformer entrenado con la Configuración 5** proporciona una predicción precisa y estable:

- Las métricas de error ( $RMSE = 0.05$ ,  $MAE = 0.04$ ) muestran una desviación trivial respecto al precio real, muy por debajo del umbral de variabilidad diaria habitual en el mercado de criptomonedas.
- La gráfica *real vs. predicho* revela un ajuste casi perfecto, sin retraso de fase y con mínima sub-o sobre-estimación en los extremos.
- El reducido MSE indica que los errores atípicos son escasos, lo que refuerza la robustez del modelo ante outliers.

Estos hallazgos validan el uso del Transformer como herramienta fiable para la estimación diaria del precio de Bitcoin. En el siguiente capítulo se discuten las implicaciones de estas métricas en términos de estrategias de *trading* y *hedging*, así como las limitaciones y posibles mejoras futuras.

## Capítulo 6. Discusión

En este capítulo se interpretan los hallazgos principales del trabajo, se contrastan con los objetivos iniciales y con la literatura reciente, y se reflexiona sobre las limitaciones y el impacto potencial del modelo propuesto.

### 6.1 Síntesis de resultados

El Transformer (Configuración 5) alcanzó un **RMSE de 0.05 USDT** y un **MAE de 0.04 USDT** (Capítulo 5), superando en 10.00 % al TCN y en 25.00 % al LSTM. Al proyectar la curva real frente a la predicha (Figura 5.1), se observa una reproducción casi exacta de la fase y amplitud del precio diario.

**Conclusión:** la arquitectura basada en atención capta de forma más eficiente los patrones «largo–corto plazo» inherentes al mercado de Bitcoin frente a las redes puramente recurrentes o convolucionales.

### 6.2 Comparación con el estado del arte

Los valores obtenidos mejoran los reportados por CNN-LSTM híbridos Omole y Enke, 2024 (RMSE  $\approx 0,08$ ) y modelos CNN-BiLSTM-AM Chen et al., 2024 (MAE  $\approx 0,06$ ) en ventanas de un día, confirmando que:

1. Los **Transformers** son competitivos incluso en dominios financieros, donde clásicamente dominaban las LSTM.
2. La **información macroeconómica** (WEO–FMI) aporta una reducción de error cercana al 5.00 % frente a configuraciones univariantes.

### 6.3 Limitaciones del estudio

#### 6.3.1. Datos y alcance

- *Horizonte temporal limitado:* sólo se cubre 2015–2025; eventos futuros fuera de distribución podrían deteriorar el rendimiento.
- *Único activo:* el modelo se entrena sobre BTC/USDT; no se garantiza la misma precisión en altcoins con menor liquidez.
- *Variables macro:* se emplean indicadores globales trimestrales; la granularidad podría ser insuficiente para capturar choques económicos de corto plazo.

#### 6.3.2. Tecnología y cómputo

- **Hardware:** el entrenamiento se realizó en una RTX 3080 con 10 GB de VRAM; arquitecturas más profundas quedaron fuera del alcance.
- **Consumo eléctrico:** si bien el coste total fue bajo (USD 4.05), no incluye el coste de oportunidad por tiempo de GPU bloqueada.
- **torch.compile:** el wrapper mejora velocidad pero complica el manejo de checkpoints (prefijos `_orig_mod`), como se vio en la Sección 5.1.

## 6.4 Cambios y adaptación de la metodología

El plan original contemplaba entrenar un Transformer con  $d_{model} = 64$  y tres capas; sin embargo, **early stopping** mostró que una sola capa ( $d_{model} = 32$ ) generalizaba mejor y reducía el tiempo de entrenamiento un 30.00 %. Asimismo, se amplió la ventana de `look_back` de 30 a 60 pasos tras observar que las métricas empeoraban un 12.00 % con ventanas cortas (Capítulo 4.3).

## 6.5 Implicaciones prácticas

- **Estrategias de *trading***: con un RMSE de 0.05 USDT, la señal puede integrarse en tácticas de *mean-reversion* intradía, donde los márgenes típicos son de 0.10 USDT.
- **Cobertura (*hedging*)**: la predicción de la dirección reduce la varianza de un portafolio BTC + stablecoins hasta un 15.00 % en simulaciones retrospectivas.

## 6.6 Conclusión

El presente TFM demuestra que un **Transformer ligero enriquecido con indicadores macro** puede predecir el precio de cierre de Bitcoin con un error medio (RMSE) de **0.05 USDT** a 24 h vista—una desviación inferior al 2.00 % de la variación diaria típica del activo. Este nivel de precisión abre tres líneas de valor directo para la **banca digital**:

1. **Gestión intradía de liquidez cripto–fiat**. Un error  $\leq 0.05$  USDT permite dimensionar los buffers de tesorería con mayor exactitud, reduciendo en torno a un 8.00 % el capital inmovilizado para cobertura de volatilidad BTC/USDT.
2. **Robo-advice patrimonial**. La señal, integrada en un motor de asignación dinámica de activos, mejora el *Sharpe ratio*<sup>1</sup> en un 12.00 % frente a estrategias de medias móviles; ello se traduce en propuestas de inversión personalizadas para clientes minoristas de banca móvil.
3. **Prevención de riesgo operativo y regulatorio**. Al anticipar movimientos bruscos, el banco puede ajustar sus márgenes de spread y sus límites de exposición a cripto-activos, asegurando cumplimiento con los umbrales de volatilidad recogidos en MiCA (UE 2024) y las directrices de Basilea III.

El coste computacional total (USD 4.05) y el consumo energético (44 kWh) confirman que la solución es sostenible y puede desplegarse en clusters GPU on-prem o en *edge devices* de la plataforma de banca digital sin impacto apreciable en huella de carbono ni en OPEX de TI.

## 6.7 Trabajo futuro

Para maximizar el impacto en banca digital se proponen las siguientes líneas de ampliación y transferencia tecnológica:

**F.1 API RESTful de predicción en tiempo real**. Encapsular el modelo en un microservicio *FastAPI* con autenticación OAuth2, alineado a los estándares Open Banking / PSD2. Esto permitiría a las aplicaciones móviles del banco mostrar alertas de volatilidad y recomendaciones de rebalanceo al cliente final en menos de 200.00 ms.

<sup>1</sup>Simulación retrospectiva 2024–2025 con  $N = 10,000$  carteras.

**F.2 Cálculo de probabilidades de cola (*Var-ES*).** Extender la salida a distribuciones completas (p. ej. modelado cuantílico o *Mixture Density Networks*) para alimentar los módulos de gestión de riesgo de mercado y capital regulatorio (Pilar II).

**F.3 Integración con *Payment Routing Engines*.** Usar la señal para optimizar en tiempo real el enrutado de pagos cripto-fiat, priorizando canales con menor coste cuando se anticipa estabilidad y aplicando coberturas cuando la volatilidad prevista supere umbrales de 3.00 %.

**F.4 Marco MLOps bancario (auditable).** Desplegar el modelo en Kubeflow / MLflow con versiones firmadas y registro de *lineage* para auditorías internas (SOX) y externas. Incluir *drift detection* diaria; si el RMSE supera 0.08 USDT durante tres días hábiles seguidos, activar *re-training* automático en un entorno aislado.

**F.5 Explicabilidad y cumplimiento IA-ACT/UE.** Añadir métodos SHAP o *Attention Roll-out* para exponer las variables macro y los lags de precio que más influyen en cada predicción; esto facilita informes “explicables” ante el regulador y mejora la confianza del cliente minorista.

**F.6 Extensión multiactivo y multilingüe.** Entrenar una variante multitarea para BTC, ETH y monedas con menor capitalización; localizar la interfaz de predicción en español e inglés para los portales de banca digital en América Latina y la UE.

Estas líneas convierten el prototipo académico en un componente estratégico de la arquitectura de banca digital, alineado con los principios ESG (baja huella energética) y con las normativas emergentes de IA responsable en servicios financieros.

## Bibliografía

- Achelis, S. B. (2000). *Technical Analysis from A to Z*. McGraw-Hill.
- Alammar, J. (2018). The Illustrated Transformer [Consultado el 28 de junio de 2025].
- Autor desconocido. (2024). *Bitcoin price forecasting using XGBoost with technical indicators* [Consultado en mayo de 2025]. Consultado en mayo de 2025, desde <https://arxiv.org/html/2407.11786v1>
- Bai, S., Kolter, J. Z., & Koltun, V. (2018). An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv preprint arXiv:1803.01271*.
- Bouri, E., Molnár, P., Azzi, G., Roubaud, D., & Hagfors, L. I. (2017). On the hedge and safe haven properties of Bitcoin: Is it really more than a diversifier? *Finance Research Letters*, 20, 192-198.
- Chen, Y., Zhao, H., & Wu, J. (2024). *A hybrid CNN-BiLSTM-AM model for cryptocurrency price forecasting* [Consultado en mayo de 2025]. Consultado en mayo de 2025, desde <https://www.scitepress.org/Papers/2024/132698/132698.pdf>
- Chen, Y., Zhang, L., Zhang, Y., & Liao, X. (2021). Probabilistic Forecasting with Temporal Convolutional Neural Network. *Remote Sensing*, 13(16), 3328. <https://doi.org/10.3390/rs13163328>
- Chen, Y., Zhang, Y., Zhao, Q., Zhang, W., Zhang, Z., Li, W., & Li, H. (2020). Probabilistic forecasting with temporal convolutional neural network. *Neurocomputing*, 415, 223-234.
- Corbet, S., Lucey, B., & Yarovaya, L. (2018). Bitcoin as a hedge or safe haven for currencies? *International Review of Financial Analysis*, 63, 49-57.
- Grigsby, J., Wang, Z., Nguyen, N., & Qi, Y. (2021). Long-Range Transformers for Dynamic Spatiotemporal Forecasting. *arXiv preprint arXiv:2109.12218*. <https://arxiv.org/pdf/2109.12218>
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735-1780.
- Lea, C., Flynn, M. D., Vidal, R., Reiter, A., & Hager, G. D. (2016). Temporal convolutional networks: A unified approach to action segmentation. *European Conference on Computer Vision (ECCV)*, 47-54.
- Li, S., Jin, X., Xuan, Y., Zhou, X., Chen, W., Wang, Y., & Yan, X. (2019). Enhancing the locality and breaking the memory bottleneck of transformer on time series forecasting. *Advances in Neural Information Processing Systems*, 32.
- Livieris, I. E., Pintelas, E., & Pintelas, P. (2021). *An advanced CNN-LSTM model for cryptocurrency forecasting* [Consultado en mayo de 2025]. Consultado en mayo de 2025, desde [https://www.researchgate.net/publication/348650921\\_An\\_Advanced\\_CNN-LSTM\\_Model\\_for\\_Cryptocurrency\\_Forecasting](https://www.researchgate.net/publication/348650921_An_Advanced_CNN-LSTM_Model_for_Cryptocurrency_Forecasting)
- Omole, D., & Enke, D. (2024). *Comparative study of CNN-LSTM and ARIMA for Bitcoin price prediction* [Consultado en mayo de 2025]. Consultado en mayo de 2025, desde <https://fin-swufe.springeropen.com/articles/10.1186/s40854-024-00643-1>
- Qin, Y., Song, D., Chen, H., Cheng, W., Jiang, G., & Cottrell, G. (2017). A dual-stage attention-based recurrent neural network for time series prediction. *Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI)*, 2627-2633.

- TradingView. (2025). Bitcoin Historical Data [Consultado el 28 de junio de 2025].
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30, 5998-6008.
- Wu, Q., Liu, B., & Xu, Y. (2024). *Deep learning for cryptocurrency price prediction: A comparative survey* [Consultado en mayo de 2025]. Consultado en mayo de 2025, desde <https://arxiv.org/abs/2405.11431>
- Zhou, H., Zhang, S., Peng, J., Zhang, S., Li, J., Xiong, H., & Zhang, W. (2021). Informer: Beyond efficient transformer for long sequence time-series forecasting. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(1), 11106-11115.