

Estimación de Precios de Bitcoin mediante Regresión Lineal Múltiple y Redes Neuronales

Bitcoin Price Estimation using Multiple Linear Regression and Neural Networks

Manuel Humberto Díaz López ¹, Andrea King-Domínguez ², Luis Améstica-Rivas ³

INFORMACIÓN DEL ARTÍCULO

Fecha de recepción: 5 de Octubre de 2023.

Fecha de aceptación: 17 de Diciembre de 2023.

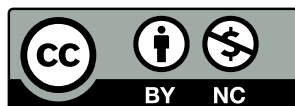
¹ Magíster en Gestión de Empresa, Universidad del Bío-Bío. Docente - investigador, Universidad del Bío-Bío - Chile.
E-mail: manuel.diaz.tkj@gmail.com
Código ORCID:
<https://orcid.org/0000-0003-5976-7118>

² Doctora en Administración y Dirección de Empresas, Universidad Politécnica de Cataluña. Docente-investigador, Universidad del Bío-Bío - Chile.
E-mail: aking@ubiobio.cl
Código ORCID:
<https://orcid.org/0000-0002-1063-4336>

³ Doctor en Administración y Dirección de Empresas, Universidad Politécnica de Cataluña. Docente - investigador, Universidad del Bío-Bío - Chile.
E-mail: lamestica@ubiobio.cl
Código ORCID:
<https://orcid.org/0000-0003-0482-0287>

CITACIÓN: Díaz López, M.H., King-Domínguez, A., & Améstica-Rivas, L. (2023). Estimación de Precios de Bitcoin mediante Regresión Lineal Múltiple y Redes Neuronales. *Podium*, 44, 119-132.
doi:10.31095/podium.2023.44.8

ENLACE DOI:
<http://dx.doi.org/10.31095/podium.2023.44.8>



Resumen

Diversos estudios se han enfocado en estimar el precio de las criptomonedas utilizando modelos de series de tiempo y variables estáticas. Este estudio se centra en la predicción del precio de Bitcoin, utilizando un modelo que combina la regresión lineal múltiple y las redes neuronales. Este enfoque permite identificar los factores que influyen en la volatilidad de Bitcoin y, mediante una selección dinámica de variables, detectar constantemente el conjunto de características más relevante para la predicción. Asimismo, se optimiza la cantidad de datos para mejorar la precisión y evitar la sobreutilización de información histórica. La combinación de estas técnicas permitió capturar patrones y tendencias subyacentes, aumentando la confiabilidad de las predicciones, con una precisión del 88%. Sin embargo, es crucial considerar la necesidad de evaluaciones continuas para adaptarse a las cambiantes condiciones del mercado. Este enfoque brinda una herramienta más precisa para tomar decisiones informadas en un mercado altamente volátil.

Palabras Clave:

Criptomonedas, Bitcoin, modelo de regresión lineal múltiple, redes neuronales, índices bursátiles, índice de volatilidad, precio del petróleo, índice de sentimiento del mercado.

Clasificación JEL: C45; C53, E37.

Abstract

Various studies have focused on estimating the price of cryptocurrencies using time series models and static variables. This study focuses on Bitcoin price prediction, using a model that combines multiple linear regression and neural networks. This approach makes it possible to identify the factors that influence Bitcoin volatility and, through a dynamic selection of variables, to constantly detect the most relevant set of characteristics for prediction. Likewise, the amount of data is optimized to improve precision and avoid overuse of historical information. The combination of these techniques captures underlying patterns and trends, increasing the reliability of predictions, with an accuracy of 88%. However, it is crucial to consider the need for continuous evaluations to adapt to changing market conditions. This approach provides a more accurate tool for making informed decisions in a highly volatile market.

Keywords:

Cryptocurrencies, Bitcoin, multiple linear regression model, neural networks, stock indices, volatility index, oil price, market sentiment index.

JEL Classification: C45; C53, E37.

Introducción

En relación con la predicción de los precios de los activos financieros, es posible distinguir dos perspectivas fundamentales. El primer enfoque postula que estos precios son inherentemente impredecibles, lo cual se apoya en la teoría de los mercados eficientes (Fama, 1970). Según esta teoría, el mercado refleja de manera completa y precisa toda la información relevante para determinar los precios de los activos. Dado que la llegada de nueva información es impredecible, los cambios en los precios de las acciones también lo serían. Este planteamiento ha llevado a varios analistas financieros y académicos a sostener que las fluctuaciones en los activos financieros siguen una trayectoria aleatoria, denominada "random walk", donde la aleatoriedad se refiere a que las variaciones de precios se generan a partir de un proceso estocástico específico. En otras palabras, si el flujo de información es constante y esta se refleja de manera inmediata en los precios de las acciones, el cambio de precio futuro reflejará únicamente las noticias del futuro y será independiente de los cambios de precio actuales. Sin embargo, dado que las noticias son, por definición, impredecibles, los cambios resultantes en los precios deben considerarse impredecibles y aleatorios. En contraste, el segundo enfoque sostiene que los precios de los activos financieros son, de hecho, predecibles, perspectiva que está vinculada al campo de las finanzas conductuales (Malkiel, 2004; Parisi y otros, 2006; Yao y otros, 2022).

Por otra parte, y respecto al segundo enfoque mencionado anteriormente, dentro del ámbito de los mercados financieros resulta imperativo disponer de una extensa variedad de datos e información que posibilite la toma de decisiones fundamentadas y eficientes. Tanto inversionistas como académicos han empleado modelos estadísticos, econométricos y financieros para desarrollar estrategias destinadas a maximizar el rendimiento de sus inversiones y afrontar la inherente incertidumbre que caracteriza a estos contextos. Igualmente, y con el propósito de alinear los resultados obtenidos con las expectativas predefinidas, ellos han desarrollado modelos con capacidad predictiva que tienen la finalidad de atenuar la incertidumbre y el riesgo inherentes en los mercados financieros

En este contexto, Martínez y otros (2021) hace una revisión exhaustiva de los más importantes modelos predictivos y sus aplicaciones en acciones y *commodities*. Entre ellos se pueden destacar el proceso autorregresivo de primer orden, el camino aleatorio o el modelo ARIMA basado en la metodología de *Box - Jenkins*. En este estudio, para la predicción del precio de los *commodities* oro y plata, utilizan el modelo ARIMA optimizado con fuerza bruta operacional para pronosticar en el mercado financiero el precio de dos *commodities*. Otro estudio similar es Parisi-Fernández y otros (2019) que aplica el modelo ARIMA multivariable optimizado con fuerza bruta, pero esta vez aplicado al precio del petróleo y títulos accionarios acciones relacionadas

a este sector económico.

En el contexto de la predicción de precios de acciones, según lo señalado por He y Eisuke (2021), se pueden emplear dos enfoques distintos: el análisis fundamental y el análisis técnico. Adicionalmente, explican que, previo a la proliferación del aprendizaje automático, se utilizaba ampliamente un algoritmo de pronóstico basado en series temporales lineales para anticipar las variaciones en los precios de las acciones. No obstante, en los últimos años, con el avance del aprendizaje automático, se han incorporado algoritmos de vanguardia, entre los cuales destacan la utilización de memorias a largo plazo y a corto plazo (LSTM) así como las redes generativas adversas (GAN) para llevar a cabo pronósticos más precisos en el ámbito de los precios bursátiles.

En definitiva, en el ámbito de la predicción de los precios de las acciones, es viable emplear una variedad de arquitecturas de redes neuronales recurrentes. No obstante, debido a que los precios de las acciones no obedecen uniformemente a una misma tendencia, resulta imperativo no depender exclusivamente de un modelo único para anticipar las fluctuaciones en todos los tipos de precios de acciones (Dey y otros, 2021).

Más allá de los modelos aplicados a activos tradicionales, como los antes mencionados, en los últimos años se ha suscitado un interés creciente en la investigación de las predicciones de precios de las criptomonedas. Estos

activos han captado una gran atención, tanto a nivel popular como académico, por ser un novedoso instrumento financiero. Estos activos se han consolidado como instrumentos financieros ampliamente reconocidos en el mundo de las transacciones. Los inversionistas han manifestado un considerable interés en participar en operaciones relacionadas con estas monedas digitales, motivados por su capacidad potencial para generar utilidades significativas (Kang y otros, 2022).

La criptomoneda pionera fue *Bitcoin*, creada por Satoshi Nakamoto en 2008 como una moneda digital que tiene la particularidad de ser descentralizada y no requiere de la intermediación de ninguna institución financiera. Si bien inicialmente se concibió como una alternativa a las monedas nacionales, los mercados de criptomonedas han adquirido notoriedad debido a sus ciclos de auge y declive ampliamente reconocidos, así como a la potencialidad de una disminución completa en los valores del mercado. En este contexto, los primeros estudios sugieren que Bitcoin y otras criptomonedas se consideran más como activos especulativos que monedas genuinas (Fry y Ibiloye, 2023).

De acuerdo a Smales (2019), las transacciones de Bitcoin se ejecutan en un sistema de código abierto conocido como "*Blockchain*". Este sistema utiliza un protocolo altamente sofisticado para crear, registrar y autenticar las transacciones. No obstante, es importante

destacar que esta criptomoneda exhibe notables variaciones en su valor. De hecho, desde su creación en 2008, esta primera criptomoneda ha experimentado un aumento significativo en su valor. Inicialmente valorada en 1.000 dólares por unidad, en el año 2010 su cotización se situaba en un nivel ligeramente superior a los 4.000 dólares. Posteriormente, en noviembre de 2021 se registró un máximo de 68.990,6 dólares, marcando un incremento significativo en su precio. Actualmente (septiembre de 2023), su valor se encuentra en torno a los 25.200 dólares, evidenciando una considerable oscilación en su cotización a lo largo de este período.

Sobre la base de lo anterior, este trabajo se centra en la predicción del precio de *Bitcoin*, utilizando un modelo que combina la regresión lineal múltiple y las redes neuronales. Mediante la utilización de registros históricos y algoritmos avanzados que facultan el análisis de patrones previos y la proyección del comportamiento futuro de la criptomoneda *Bitcoin*, se persigue proporcionar una herramienta que facilite la comprensión y la explotación de las oportunidades que las criptomonedas ofrecen en el contexto financiero contemporáneo.

Revisión de literatura

Como explica Almeida y Gonçalves (2023), las criptomonedas -incluida *Bitcoin*- a pesar de su función original como medios de intercambio, han evolucionado predominantemente hacia el ámbito de activos de inversión. En

efecto, han emergido como una destacada categoría de activos en los mercados financieros globales, caracterizándose por su rápido desarrollo y penetración en diversas regiones del mundo. Este fenómeno ha dado lugar a la consolidación de las criptomonedas como uno de los sectores financieros de mayor crecimiento a nivel global. La irrupción de este novedoso mercado, acompañada por el surgimiento de plataformas de inversión, ha llevado a que las personas inviertan en ellos, con la promesa de obtener ganancias sustanciales de forma relativamente sencilla.

El creciente uso de las criptomonedas como activo de inversión y su alta volatilidad en sus cotizaciones ha generado la necesidad de desarrollar modelos predictivos. La fijación de precios en el mercado de criptomonedas está primordialmente vinculada al comportamiento de los *traders* en lugar de depender en gran medida de los fundamentos económicos, tales como la producción bruta, la tasa de desempleo, los tipos de cambio y las políticas monetarias. Consecuentemente, se ha suscitado un creciente interés en comprender los factores determinantes de los rendimientos de las criptomonedas, las interacciones entre distintos criptoactivos, así como sus derivados, y los *commodities*, además de su dependencia, movimientos e interacciones con los activos financieros convencionales (Ivanovski y Hailemariam, 2023).

Así, distintos estudios se han centrado en el análisis de la relación entre los precios de *commodities*, específicamente

el precio del petróleo, y las criptomonedas (Mohali y Palm, 2021; Lahiani y otros, 2021). En relación al impacto de los precios accionarios, se han utilizado distintos índices bursátiles, como el Dow Jones, S&P 500, NASDAQ, Nikkei, Shanghai, entre otros (Chuang y otros, 2009; Ivanovski y Hailemariam, 2023; Lahiani y otros, 2021; Sun y otros, 2020).

Relacionado a lo anterior está el índice VIX, comúnmente referido como el "Índice del Sentimiento de Mercado". Corresponde a un indicador que proporciona una estimación de la volatilidad esperada en el futuro cuyo cálculo se basa en la medición de la volatilidad implícita en los contratos de opciones vinculados al índice S&P 500, con un horizonte temporal de 30 días. Entonces, mientras el S&P 500 es un buen indicador de cómo les está yendo a los mercados financieros, el VIX pretende proporcionar una medida instantánea sobre cuánto cree el mercado que fluctuará el S&P 500 en los próximos 30 días. Al ser incorporado ambas variables en un modelo predictivo, se consideran tanto los rendimientos numéricos como los riesgos en los mercados financieros (Kjærland, y otros, 2018).

Otra variable analizada como factor que impacta en el precio de las criptomonedas son las divisas. Esto se fundamenta en que la mayoría de las cotizaciones y modificaciones en los operadores financieros se limitan a realizar intercambios con un conjunto selecto de divisas, siendo el euro y el

dólar los más utilizadas. Este fenómeno sugiere que estas monedas dependen en gran medida de las principales divisas para su comercialización y están vinculadas a las economías de los países que las respaldan (Mohali y Palm, 2021).

Al igual que la relación entre el volumen de transacciones y el rendimiento ha sido objeto de extenso análisis en el contexto de acciones, materias primas, futuros de tasas de interés y divisas, también ha sido analizada en el mercado de las criptomonedas. Balcilar y otros (2017) explican que el análisis de esta relación es relevante, puesto que proporciona una comprensión más profunda de la transmisión de información al mercado y su posterior incorporación en los precios de los activos. Asimismo, este enfoque contribuye de manera significativa al fortalecimiento de la capacidad de pronóstico, tanto del rendimiento como de la volatilidad de los activos financieros, lo que es especialmente relevante en momentos de tensiones económicas, lo que permitiría mejorar la comprensión de los movimientos alcistas y bajistas observados en los mercados financieros.

Finalmente, un foco de interés es el análisis del impacto que tienen los sentimientos en el precio estas monedas. Esto se fundamentaría en un estudio del Instituto Americano de Investigación Económica (AIER), que asevera que las noticias y los sentimientos que influyen a nivel mundial pueden provocar grandes fluctuaciones en el precio de *Bitcoin* (Ye y otros, 2022). Con este propósito Shen y

otros (2019) analizaron como impacta el número de *tweets* en Twitter, un enfoque novedoso, considerando que estudios previos habían utilizado *Google Trends*. Anamika y Subramaniam (2023) utilizaron, para medir el sentimiento de los inversores, una medida directa basada en encuestas que captura el sentimiento de los inversores sobre Bitcoin, obtenida de la base datos de *Sentix database*.

Metodología

La metodología empleada en esta investigación se basa en un enfoque dinámico que combina técnicas de regresión lineal múltiple (RLM) y redes neuronales. El objetivo principal es aprovechar las fortalezas de ambos modelos y mejorar la capacidad de predicción y adaptarse de manera óptima a la volatilidad inherente a las criptomonedas, en específico, el *Bitcoin*. Las variables seleccionadas por el modelo de RLM se utilizaron como entrada para el modelo de redes neuronales.

Mientras que la RLM es eficaz para identificar relaciones lineales entre las variables, las redes neuronales tienen la capacidad de descubrir patrones complejos y no lineales en los datos. Esta colaboración entre ambas técnicas permite enfrentar la complejidad y no linealidad inherente en los conjuntos de datos relacionados con las criptomonedas, en última instancia, mejorando la capacidad predictiva del modelo dinámico.

Para ello, se seleccionaron datos históricos relevantes sobre las

criptomonedas en estudio, como precios, volúmenes de transacciones, datos del mercado, noticias y otros factores importantes para el análisis. Cabe mencionar que se obtuvieron registros diarios de cada una de estas variables a través de los sitios web www.investing.com y <https://trends.google.es/home>, durante el período comprendido entre el 01 de abril y el 30 de julio de 2023. Estos datos se utilizaron para entrenar y evaluar los modelos.

Como primera etapa, se realizó un análisis exploratorio de los datos recopilados con el objetivo de identificar posibles correlaciones y patrones entre las variables. Esto es, determinar la correlación entre variables independientes y la variable dependiente (precio del Bitcoin) para así comprender la relación entre ellas. Luego, se determinó el valor del factor de inflación de las varianzas (VIF), con el objeto de detectar la presencia de multicolinealidad entre las variables. Como resultado de ambos procesos se seleccionaron las variables a incluir en el modelo.

Mediante el uso de un modelo de RLM, se llevó a cabo un análisis estadístico para determinar las variables que tienen mayor influencia en la predicción del precio del *Bitcoin*. El modelo de RLM fue implementado en Python, usando bibliotecas como *Scikit-learn*. Se ajustó el modelo usando los datos seleccionados y se evaluó su desempeño mediante métricas adecuadas, como el coeficiente de determinación (R^2) y el error cuadrático medio (MSE). Se realizaron pruebas en diferentes

intervalos de tiempo para evaluar y comparar resultados con el propósito de identificar si los datos utilizados son adecuados para la elección de las variables que serán utilizadas en las redes neuronales.

Como tercera etapa, se desarrolló un modelo de redes neuronales en Python, usando las variables que se seleccionaron del modelo anterior, junto a la cantidad de datos seleccionados para poder predecir el valor de los siguientes 20 días. Para esto se utilizaron bibliotecas como TensorFlow, Keras Numpy, pandas, tensorflow LSTM, sklearn. Se definió la arquitectura de la red, incluyendo el número de capas LSTM (Long Short-Term Memory), neuronas y funciones de activación. El modelo se entrenó usando los datos seleccionados y se evaluó su rendimiento mediante distintas métricas para determinar la precisión y el rendimiento del modelo dinámico: error absoluto medio (MAE) y el error porcentual absoluto medio (MAPE).

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (1)$$

$$MAPE = \frac{1}{100} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (2)$$

Donde:

n: número de muestras en el conjunto de prueba

y_i : valor real de la variable objetivo en la muestra i

\hat{y}_i : predicción del modelo para la muestra i

El modelo consta con dos capas principales:

1. Capa LSTM: Tipo de capa de red neuronal recurrente que es especialmente útil para trabajar con secuencias de datos, como series de tiempo. La capa se define utilizando LSTM (128, activation="relu", input_shape=(X_train.shape[1],1)). Significa que la capa LSTM tiene 128 unidades (neuronas) LSTM con una función de activación ReLU. La entrada a esta capa debe tener la forma (X_train.shape[1],1), donde X_train.shape[1] representa el número de características (variables predictoras) y 1 representa la dimensión del tiempo. En este caso, la capa LSTM está configurada para procesar secuencias unidimensionales de longitud X_train.shape[1].

2. Capas densas: Después de la capa LSTM, se utilizan dos capas densas (capas completamente conectadas). Estas capas se definen como Dense (64, activation="relu", kernel_regularizer=regularizers.l2(0,001)) y Dense(1). La primera capa densa tiene 64 unidades con una función de activación ReLU y una regularización L2 de 0,001. La regularización L2 ayuda a prevenir el sobreajuste al penalizar los pesos más grandes en la función de pérdida. La segunda capa densa tiene una sola unidad y no tiene función de activación, lo que indica que es una capa de salida para la predicción.

Este tipo de arquitectura es comúnmente utilizado en problemas de series de tiempo, pronósticos y otras tareas que involucran datos secuenciales.

Como variable dependiente se utilizó el valor de mercado de Bitcoin, y las variables independientes se presentan en la Tabla 1.

En definitiva, la implementación del modelo dinámico propuesto que integra la RLM con las redes neuronales se realizó de la siguiente manera:

a) Una vez cargados los datos y realizada la normalización, se aplicó una

RLM. Esto implicó ajustar un modelo de regresión lineal a los datos, donde se buscó establecer una relación lineal entre las variables independientes (características) y la variable dependiente (etiqueta). La selección de las variables más significativas para el modelo se llevó a cabo mediante un análisis de correlación de Spearman y pruebas estadísticas.

b) En la prueba t de Student, se compararon las medias de dos grupos y se

Tabla 3.

Resultado obtenido de la RLM

Variable	Definición	Fundamentos teóricos
Oil	Precio del petróleo crudo.	(Mohali y Palm, 2021) (Lahiani y otros, 2021)
Nikkei	Índice bursátil de la Bolsa de Valores de Tokio en Japón.	(Lahiani y otros, 2021)
Shanghai	Índice bursátil de la Bolsa de Valores de Shanghái en China.	(Sun y otros, 2020)
DAX	Índice DAX 30, el índice bursátil de Alemania.	(Lahiani y otros, 2021)
Dow Jones	Índices bursátiles de Estados Unidos (Bolsa de Nueva York y <i>NASDAQ</i>).	(Chuang y otros, 2009) (Ivanovski y Hailemariam, 2023)
SP	Índice <i>Standard & Poor's</i> 500.	(Chuang y otros, 2009) (Lahiani y otros, 2021) (Ivanovski y Hailemariam, 2023)
NASDAQ	Índice <i>NASDAQ Composite</i>	(Lahiani y otros, 2021) (Ivanovski y Hailemariam, 2023)
VIX	Índice de volatilidad VIX (o índice del miedo). Mide las expectativas de volatilidad en el mercado de valores de Estados Unidos.	(Lahiani y otros, 2021) (Kjærland y otros, 2018)
Dólar	Tipo de cambio del dólar estadounidense con respecto a otra moneda.	(Mohali y Palm, 2021) (Sun y otros, 2020)
Volcripto	Volumen de inversión de una criptomoneda. Mide la variabilidad de los precios de las criptomonedas en un período de tiempo.	(Balcilar y otros, 2017) (Chuang y otros, 2009)
Sentimiento	Índice de sentimiento del mercado. Mide la opinión en redes sociales o el sentimiento general de los inversores y <i>traders</i> con respecto a las condiciones del mercado.	(Anamika y Subramaniam, 2023) (Georgoula y otros, 2015)

Fuente: Elaboración propia.

determinó si existen diferencias significativas entre ellos. Si la diferencia es significativa, puede indicar que la variable es relevante para la predicción.

c) Una vez obtenidos los coeficientes de la RLM, se utilizaron como entrada para el modelo de redes neuronales. Estas características son una representación ponderada de las variables originales y se utilizaron para capturar la información relevante obtenida a través de la regresión lineal.

d) En la implementación del modelo de redes neuronales, se definió la arquitectura y se utilizaron las características derivadas de la RLM como entrada. Se construyó un modelo LSTM (*Long Short-Term Memory*) usando capas densas. Las capas LSTM ayudan a capturar la información secuencial y las dependencias a largo plazo en los datos, mientras que las capas densas ayudan a mejorar la capacidad de generalización

del modelo.

Resultados

Como se explicó anteriormente, en una primera etapa se analizó la correlación existente entre las variables independientes y dependientes. Como muestra la matriz de correlación (ver Tabla 2), las variables que tienen la mayor correlación (en valor absoluto) con el precio del *Bitcoin* son: Nikkei, Shanghai, NASDAQ, Oil, DAX, SyP, VIX, y Sentimiento. De esta manera se determinó que, en un primer análisis, podrían ser aplicadas en la RLM. Sin embargo, la determinación del factor VIF, detectó la presencia de multicolinealidad entre algunas variables, llevando a que el índice Shanghai no fuese considerado en el análisis subsiguiente.

Aplicando las variables seleccionadas a la RLM, se obtuvo un coeficiente de determinación R^2 ajustado de 0,876, es

Tabla 2.

Matriz de correlación entre variables dependiente e independientes

	<i>Bitcoin</i>	<i>Nikkei</i>	<i>Shanghai</i>	<i>Oil</i>	<i>DAX</i>	<i>DJIA</i>	<i>SyP</i>	<i>NASDAQ</i>	<i>VIX</i>	<i>Dolar</i>	<i>Volc.</i>	<i>Sent.</i>
<i>Bitcoin</i>	1											
<i>Nikkei</i>	0,55	1,00										
<i>Shanghai</i>	0,61	0,13	1,00									
<i>Oil</i>	-0,23	-0,42	0,22	1,00								
<i>DAX</i>	0,69	0,83	0,50	-0,13	1,00							
<i>DJ</i>	-0,06	0,01	0,13	0,53	0,23	1,00						
<i>SyP</i>	0,59	0,79	0,30	-0,08	0,83	0,51	1,00					
<i>NASDAQ</i>	0,72	0,91	0,27	-0,36	0,85	0,13	0,92	1,00				
<i>VIX</i>	-0,43	0,11	-0,41	-0,24	-0,15	-0,42	-0,20	-0,05	1,00			
<i>Dolar</i>	-0,05	-0,01	-0,11	0,11	-0,01	-0,01	0,01	0,01	0,12	1,00		
<i>Volc.</i>	-0,12	-0,33	-0,04	0,02	-0,32	-0,30	-0,39	-0,31	0,05	0,10	1,00	
<i>Sent.</i>	0,35	-0,22	0,22	-0,06	-0,19	-0,33	-0,22	-0,08	-0,28	0,02	0,45	1

Fuente: Elaboración propia.

decir, 87,6%. Los indicadores estadísticos relacionados a cada una de las variables se muestran en la Tabla 3. Cada uno de los coeficientes de las variables son los que se aplicarán más tarde en las predicciones diarias del precio del *Bitcoin*.

Una vez elegidas las variables a utilizar en el modelo en las redes neuronales, éste determinó la serie de tiempo adecuado para la predicción. Para ello se normalizaron los datos de las características (variables independientes) y el precio de *Bitcoin* (variable dependiente) para que estén en un rango entre 0 y 1. Luego, se definieron los parámetros del modelo LSTM, como el número de épocas, tamaño del lote y el incremento en el que se realizará el análisis. Para ello, se realizó un análisis de incrementos de 20 registros.

Así, fue posible ejecutar el modelo con los datos de entrenamiento (40%) y se validó con los datos de prueba (60%), y evaluar el rendimiento del modelo calculando la pérdida en los datos de prueba. Finalmente, se desnormalizaron las predicciones para obtener los valores

reales del precio de Bitcoin.

Como resultado de lo anterior, el modelo permite encontrar el mejor conjunto de datos para predecir el precio del *Bitcoin*. De esta manera, al variar el número de registros utilizados, se pudo constatar que el intervalo de tiempo más efectivo para estimar el precio futuro es el comprendido en los últimos 60 días, ya que se obtienen como resultado un MAE de 596,1 y un MAPE igual a 2,16% de error comparando con el real. En el caso de 40 y 80 días los resultados serían un MAE de 1075,91 y 1838,01, y un MAPE de 3,82% y 8,23, respectivamente (ver Figura 1).

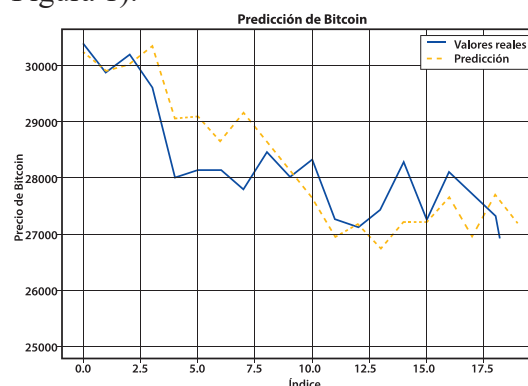


Figura 1. Estimaciones del precio del Bitcoin, considerando un intervalo de 60 días.
Fuente: Elaboración propia.

Tabla 3.

Resultado obtenido de la RLM

	Coef.	Error típico	Est. t	Prob.	Inferior 95%	Superior 95%	Inferior 95,0%	Superior 95,0%
Intercepción	117002,0	19833,2	5,90	0,00	77672,0	156332,1	77672,0	156332,1
Nikkei	-1,02	0,23	-4,46	0,00	-1,48	-0,57	-1,48	-0,57
Oil	65,49	43,03	1,52	0,13	-19,84	150,81	-19,84	150,81
DAX	4,02	0,63	6,42	0,00	2,78	5,26	2,78	5,26
SyP	-33,37	4,40	-7,59	0,00	-42,09	-24,65	-42,09	-24,65
NASDAQ	9,17	0,97	9,48	0,00	7,25	11,09	7,25	11,09
VIX	-1,03	0,12	-8,57	0,00	-1,27	-0,80	-1,27	-0,80
Sentimiento	48,78	14,50	3,36	0,00	20,03	77,53	20,03	77,53

Fuente: Elaboración propia.

A partir de lo anterior, se observa que el modelo más adecuado para esa cantidad de variables es aquella con contempla 60 registros. Igualmente, las líneas de tendencia y la del valor real muestran un resultado con tendencia similar al alza. Adicionalmente, se desprende del gráfico que, para los últimos cinco días, el modelo no predice con certeza debido a la volatilidad de la criptomoneda.

Conclusiones

A partir de los hallazgos encontrados, se puede inferir la creciente importancia de encontrar un modelo más efectivo para predecir las fluctuaciones de precios en los activos financieros, en particular en el caso de las criptomonedas, cuyo uso ha crecido exponencialmente en la última década, complejizando aún más la dinámica de los mercados financieros mundiales.

Se han observado distintas metodologías, pero sin duda al intentar innovar utilizando complementariamente RLM, redes neuronales, la ponderación de resultados de la matriz de peso, la selección dinámica de variables y datos, ha permitido capturar patrones y tendencias subyacentes en el mercado criptográfico, brindando resultados de correlación de datos con una precisión cercana al 88%. Por otro lado, se destaca el hecho que cuando se desea proyectar una ventana de 10 o 20 días, es sugerible seguir la tendencia de los resultados.

Estos resultados apoyarían el enfoque de que los precios de estos activos

financieros son parcialmente predecibles, contrariamente a lo propuesto por la teoría de los mercados eficientes y *random walk*.

Es importante tener en cuenta que el rendimiento futuro de las criptomonedas puede verse afectado por las regulaciones específicas de cada país, aún incipiente en muchas economías. Por tanto, la evolución de las políticas y normativas en torno a las criptomonedas puede influir en su adopción y aceptación en los mercados lo que, a su vez, podría impactar en sus precios y volatilidad.

El modelo propuesto difiere significativamente de los enfoques previamente examinados, que se basaban principalmente en el análisis de series de tiempo y variables estáticas. Es relevante destacar que muchos de estos modelos no están diseñados para adaptarse a nuevas realidades o condiciones ambientales cambiantes. Sin embargo, es importante tener en cuenta que los resultados obtenidos en los artículos analizados han sido exitosos.

En este sentido, el modelo propuesto en este trabajo representa un avance al incorporar la selección dinámica de variables y la selección dinámica de los datos, lo que permite adaptar constantemente el conjunto de características relevantes para la predicción. Esta mejora ha demostrado su capacidad al aumentar la precisión del modelo y brindar predicciones más confiables en un mercado tan cambiante y complejo como el de las criptomonedas.

Dado el carácter experimental del modelo propuesto, lo cual insta a buscar mejoras que permitan levantar las limitaciones explicitadas a lo largo del estudio, como a su vez indagar en nuevas líneas de investigación, especialmente para el activo financiero criptomonedas. Se propone como futura investigación la aplicación del modelo propuesto a distintas criptomonedas y activos financieros.

Contribución de autores

M.H.D.L.: Idea, revisión de literatura y estado del arte, diseño de la investigación, metodología, análisis de datos, y redacción del artículo.

A.K.D.: Idea, revisión de literatura y estado del arte, diseño de la investigación, metodología, análisis de datos, y redacción del artículo.

L.A.R.: Validación del modelo, revisión y análisis de resultados y redacción del artículo.

Referencias

Almeida, J., y Gonçalves, T. C. (2023). A Decade of Cryptocurrency Investment Literature: A Cluster-Based Systematic Analysis. *International Journal of Financial Studies*, 11(2), 71. <https://doi.org/10.3390/ijfs11020071>

Anamika, M. C., y Subramaniam, S. (2023). Does sentiment impact cryptocurrency? *Journal of Behavioral Finance*, 24(2), 202–218. <https://doi.org/10.1080/15427560.2021.1950723>

Balcilar, M., Bouri, E., Gupta, R., y Roubaud, D. (2017). Can Volume Predict Bitcoin Returns and Volatility? A Quantiles-Based Approach.

Economic Modelling, 64, 74–81. <https://doi.org/10.1016/j.econmod.2017.03.019>

Chuang, C. C., Kuan, C. M., y Lin, H. Y. (2009). Causality in quantiles and dynamic stock return-volume relations. *Journal of Banking and Finance*, 33(7), 1351–1360. <https://doi.org/10.1016/j.jbankfin.2009.02.013>

Dey, P., Hossain, E., Hossain, M. I., Chowdhury, M. A., Alam, M. S., Hossain, M. S., y Andersson, K. (2021). Comparative analysis of recurrent neural networks in stock price prediction for different frequency domains. *Algorithms*, 14(8), 1–20. <https://doi.org/10.3390/a14080251>

Fama, E. F. (1970). Efficient Market Hypothesis: A Review of Theory and Empirical Work. *The Journal of Finance*, 25(2), 383–417. <https://doi.org/10.2307/2325486>

Fry, J., y Ibiloye, O. (2023). Towards a taxonomy for crypto assets. *Cogent Economics and Finance*, 11(1). <https://doi.org/10.1080/23322039.2023.2207266>

Georgoula, I., Pournarakis, D., Bilanakos, C., Sotiropoulos, D. N., y Giaglis, G. M. (2015). Using Time-Series and Sentiment Analysis to Detect the Determinants of Bitcoin Prices. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2607167>

He, B., y Eisuke, K. (2021). The Application of Sequential Generative Adversarial Networks for Stock Price Prediction. *The Review of Socionetwork Strategies*, 15, 455–470. <https://doi.org/10.1007/s12626-021-00097-2>

Ivanovski, K., y Hailemariam, A. (2023). Forecasting the stock-cryptocurrency relationship: Evidence from a dynamic GAS model. *International Review of Economics and Finance*, 86(February 2022), 97–111. <https://doi.org/10.1016/j.iref.2023.03.008>

Kjærland, F., Khazal, A., Krogstad, E., Nordstrøm, F., y Oust, A. (2018). An Analysis of Bitcoin's Price Dynamics. *Journal of Risk and Financial Management*, 11(4), 63. <https://doi.org/>

10.3390/jrfm11040063

Kang, T. S., Joo, M. I., Kim, B. S., y Lee, T. G. (2022, February). Blockchain-based Lightweight Transaction Process Modeling and Development. In *2022 24th International Conference on Advanced Communication Technology (ICACT)* (pp. 113-118). IEEE.

Lahiani, A., Jeribi, A., y Boukef Jlassi, N. (2021). Nonlinear tail dependence in cryptocurrency and stock market returns: The role of Bitcoin futures. *Research in International Business and Finance*, 56. <https://doi.org/10.1016/j.ribaf.2020.101351>

Malkiel, B. G. (2004). Models of stock market predictability. *Journal of Financial Research*, 27(4), 449–459. <https://doi.org/10.1111/j.1475-6803.2004.00102.x>

Martínez, J., Améstica-Rivas, L., Parisi, A., y Gurrola, C. (2021). Predictor alternativo del precio de los activos financieros. El caso de la plata y el oro. *Oikos Polis, Revista Latinoamericana*, 30–54. http://www.scielo.org.bo/pdf/rlces/v6n2/v6n2_a04.pdf

Mohali, A., y Palm, F. (2021). Efecto de indicadores económicos adelantados sobre la predicción de precios de criptomonedas. *Visión Gerencial*, 1(21), 59–68. <https://doi.org/10.53766/vigeren/2021.21.01.03>

Parisi-Fernández, A., Améstica-Rivas, L., y Chileno-Trujillo, Ó. (2019). Predicción de variaciones en el precio del petróleo con el modelo de optimización ARIMA, innovando con fuerza bruta operacional. *Tec Empresarial*, 13(1), 53–70. <https://doi.org/10.18845/te.v13i1.4302>

Parisi, A., Parisi, F., y Díaz, D. (2006). Modelos de algoritmos genéticos y redes neuronales en la predicción de índices bursátiles asiáticos. *Cuadernos de Economía - Latin American Journal of Economics*, 43(128), 251–284. <https://www.jstor.org/stable/41951551>

Shen, D., Urquhart, A., y Wang, P. (2019). Does twitter predict Bitcoin? *Economics Letters*, 174, 118–122. <https://doi.org/10.1016/j.econlet.2018.11.007>

Smales, L. A. (2019). Bitcoin as a safe haven: Is it even worth considering? *Finance Research Letters*, 30, 385–393. <https://doi.org/10.1016/j.frl.2018.11.002>

Sun, X., Liu, M., y Sima, Z. (2020). A novel cryptocurrency price trend forecasting model based on LightGBM. *Finance Research Letters*, 32. <https://doi.org/10.1016/j.frl.2018.12.032>

Yao, Y., Li, X., y Li, Q. (2022). A Comparison on LSTM Deep Learning Method and Random Walk Model Used on Financial and Medical Applications: An Example in COVID-19 Development Prediction. *Computational Intelligence and Neuroscience*, 2022. <https://doi.org/10.1155/2022/4383245>

Ye, Z., Wu, Y., Chen, H., Pan, Y., y Jiang, Q. (2022). A Stacking Ensemble Deep Learning Model for Bitcoin Price Prediction Using Twitter Comments on Bitcoin. *Mathematics*, 10(8). <https://doi.org/10.3390/math10081307>

