

# Pronóstico de precios de criptomonedas

Sergio Andrés Elizondo

25 de noviembre de 2024

## Resumen

En este estudio, se presenta un modelo multivariado para la predicción del precio de criptomonedas, integrándolo en una estrategia automatizada de compra-venta. Utilizando datos históricos por minuto de varias criptomonedas (BTC, ETH, XRP, LTC, ADA) se emplea un modelo de aprendizaje profundo basado en LSTM bidireccional y GRU para las predicciones.

Con base en estas predicciones, se desarrolló una estrategia de compra-venta basada en los resultados del modelo, incorporando reglas predefinidas para obtener ganancias (*take profit*) o evitar pérdidas (*stop loss*).

Se establecen las bases para un marco robusto que combine el análisis predictivo con estrategias financieras en un entorno de alta volatilidad como es el mercado de las criptomonedas.

## 1. Introducción

El mercado de criptomonedas ha emergido como un entorno financiero altamente dinámico y complejo, caracterizado por la alta volatilidad y la interacción de múltiples variables endógenas y exógenas. Este entorno plantea desafíos significativos para los participantes del mercado que buscan optimizar sus estrategias de inversión mediante la predicción precisa de los precios futuros y la ejecución automatizada de operaciones.

Las criptomonedas no solo presentan dinámicas internas de oferta y demanda, sino que también forman redes de interdependencias que afectan sus precios. En particular, el precio del *Bitcoin* es susceptible a variaciones en criptomonedas como *Ether*, XRP y ADA, debido a factores como correlaciones de mercado o burbujas financieras compartidas.[1].

El desarrollo de modelos predictivos en este contexto requiere la integración de técnicas avanzadas de análisis multivariado y series temporales. Estos modelos deben considerar la naturaleza no estacionaria de los datos, los patrones estocásticos y las interdependencias entre activos financieros. Los modelos tradicionales a menudo no logran capturar la dinámica no lineal ni las dependencias temporales complejas en los mercados de criptomonedas.

Para superar estas limitaciones, en el presente trabajo se utiliza un enfoque basado en redes neuronales recurrentes (RNN), específicamente *Long Short-Term Memory* (LSTM) y *Gated Recurrent Units* (GRU), con el fin de capturar

las relaciones temporales multivariadas entre diferentes criptomonedas. Estas arquitecturas han demostrado ser altamente efectivas para modelar datos secuenciales multivariados. [2]

## 2. Metodología

A continuación, se describen las etapas de obtención, procesamiento y modelado de datos, así como el diseño de la estrategia de inversión y los métodos de evaluación.

### 2.1. Obtención de datos

Los datos históricos por minuto de los precios de criptomonedas fueron obtenidos a través de la interfaz (*API*) de *Binance*, accediendo a ella mediante la biblioteca *ccxt*. Se seleccionaron cinco criptomonedas: *Bitcoin* (BTC), *Ether* (ETH), *XRP*, *Litecoin* (LTC) y *ADA*, usando como precio base *Tether* (USDT).

El rango temporal del estudio abarcó los últimos tres meses, con un total de 132777 datos. Los datos se almacenaron en archivos en formato *CSV* para facilitar su reutilización.

### 2.2. Procesamiento de datos

1. **Escalamiento de características:** Se le aplicó una normalización estándar a los precios de cierre ( $P_t$ ) de las criptomonedas:

$$Z_{i,t} = \frac{P_{i,t} - \mu_i}{\sigma_i}$$

donde  $\mu_i$  y  $\sigma_i$  representan la media y la desviación típica de la criptomoneda  $i$ , respectivamente.

Este proceso garantiza que las características se encuentren en una escala comparable, mitigando problemas de convergencia en el modelo.

2. **Construcción de secuencias multivariadas:** Se empleó una ventana deslizante de tamaño  $L = 60$  para construir secuencias de entrada, formando un tensor  $X \in \mathbb{R}^3$  con dimensiones  $N \times L \times d$ , donde  $N$  es el número total de muestras,  $L$  es la longitud de la secuencia temporal y  $d = 5$  es el número de criptomonedas consideradas. Las etiquetas  $y \in \mathbb{R}$  corresponden a los  $N$  valores de  $Z_{BTC}$  desplazados  $L$  pasos hacia adelante.

### 2.3. Diseño del modelo

El modelo propuesto combina arquitecturas de redes neuronales recurrentes (RNN) para capturar patrones temporales y relaciones multivariadas:

1. **Capas recurrentes:** El tensor de entrada (con dimensiones  $N \times L \times d$ ) es procesado por dos subredes:

- Una red *BiLSTM* (*Bidirectional Long Short-Term Memory*) con unidades ajustables  $u_{\text{LSTM}}$  que aprende de las dependencias temporales tanto pasadas como futuras:

$$h^{\text{LSTM}} = \text{LSTM}(X, u_{\text{LSTM}})$$

- Una red *GRU* (*Gated Recurrent Unit*) con parametrización  $u_{\text{GRU}}$  que modela patrones temporales más sencillos y tiene menor costo computacional:

$$h^{\text{GRU}} = \text{GRU}(X, u_{\text{GRU}}),$$

2. **Capa de concatenación:** Las salidas  $h^{\text{LSTM}}$  y  $h^{\text{GRU}}$  se concatenan:

$$h^{\text{concat}} = [h^{\text{LSTM}}, h^{\text{GRU}}].$$

3. **Capa densa y de salida:** Una capa densa con activación ReLU procesa  $h^{\text{concat}}$ :

$$h^{\text{dense}} = \text{ReLU}(Wh^{\text{concat}} + b),$$

donde  $W$  y  $b$  son los pesos y sesgo de la capa.

Por último una capa densa con una sola neurona de salida predice el precio normalizado  $y$ .

4. **Optimización de hiperparámetros:** Los parámetros  $u_{\text{LSTM}}$ ,  $u_{\text{GRU}}$  y el tamaño de la capa densa fueron optimizados mediante el algoritmo *Hyperband*, utilizando la implementación de la biblioteca *Keras Tuner*. Los parámetros elegidos después de la optimización son los siguientes:

- Unidades LSTM: 50
- Unidades GRU: 100
- Unidades densas: 150

La pérdida utilizada fue el error cuadrático promedio (MSE):

$$\mathcal{L}(y, \hat{y}) = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2.$$

## 2.4. Estrategia de inversión

Se diseñó una estrategia de compra-venta clásica basada en las predicciones del modelo. [3] Los criterios de decisión de dicha estrategia son los siguientes:

1. **Entrada en posición:**

- Compra (*long*) si  $\hat{y}_{t+1} > (1 + \text{threshold})y_t$ .
- Venta (*short*) si  $\hat{y}_{t+1} < (1 - \text{threshold})y_t$ .

2. **Salida de posición:** Se definen reglas de obtención de ganancias (*take profit*) y de prevención de pérdidas (*stop loss*) basadas en cambios porcentuales  $\Delta$ :

$$\Delta = \frac{P_{\text{actual}} - P_{\text{entrada}}}{P_{\text{entrada}}}.$$

Se registran las ganancias y pérdidas netas para evaluar el desempeño de la estrategia.

## 2.5. Evaluación del modelo

El modelo se evalúa mediante las siguientes métricas:

- Error cuadrático promedio (*Mean Squared Error, MSE*):

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2.$$

- Error absoluto promedio (*Mean Absolute Error, MAE*):

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|.$$

- Raíz del error cuadrático promedio (*Root Mean Squared Error, RMSE*):

$$\text{RMSE} = \sqrt{\text{MSE}}.$$

- Coeficiente de determinación ( $R^2$ ):

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2}.$$

- Error porcentual absoluto promedio (*Mean Absolute Percentage Error, MAPE*):

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100.$$

Estas métricas se utilizan para comparar el enfoque multivariado propuesto en este trabajo con un modelo base: Predicción de *BTC* de una sola variable usando *LSTM*.

Además, se evalúa el desempeño financiero analizando el balance final de la inversión simulada (métricas de retorno de inversión de la estrategia algorítmica).

### 3. Resultados

#### 3.1. Comparación de métricas de modelos

A continuación, se presentan las métricas de desempeño obtenidas mediante un enfoque tradicional (sin considerar relaciones multivariadas) y un enfoque multivariado. Los resultados se muestran en la Tabla 1.

Cuadro 1: Comparación de métricas de modelos.

Métrica	Enfoque Tradicional	Enfoque Multivariado
MSE	85,188,673	430,852
MAE	6,180,87	495,48
RMSE	9,229,77	656,39
$R^2$	0,60	0,96
MAPE	0,59 %	0,51 %

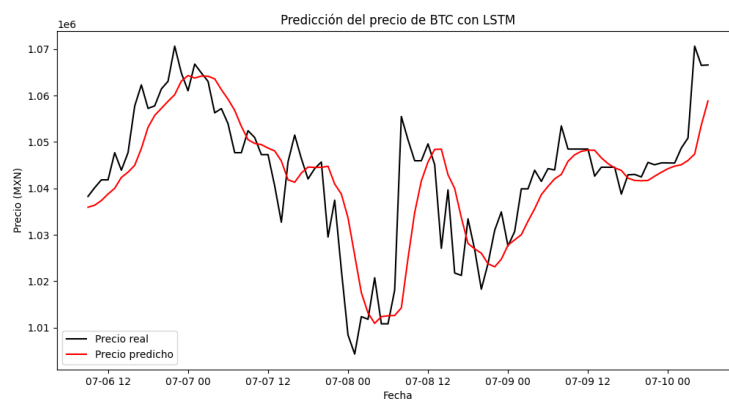


Figura 1: Predicción con enfoque tradicional.

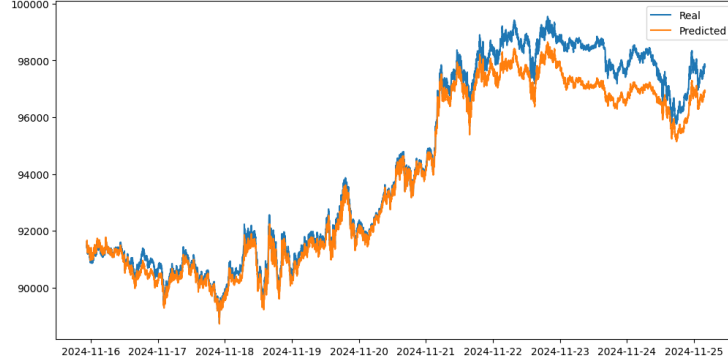


Figura 2: Predicción con enfoque multivariado.

Como se observa, el enfoque multivariado muestra un desempeño significativamente superior en todas las métricas analizadas. En particular, el error cuadrático promedio (MSE) y el error absoluto promedio (MAE) se redujeron sustancialmente, lo cual sugiere una mayor precisión del modelo multivariado. Además, el coeficiente de determinación ( $R^2$ ) incrementó de 0,60 a 0,96, lo que implica que el modelo multivariado captura con mayor fidelidad la variabilidad en el precio del *Bitcoin*.

### 3.2. Análisis de inversión

Se realizó una simulación de una inversión utilizando la estrategia sugerida durante un período de 9 días (del 16 al 25 de Noviembre). En este período el capital inicial (1 USD) creció a un valor final de 1,0337. Se calcula la tasa de crecimiento diaria promedio:

$$g_d = \left( \frac{C_{\text{final}}}{C_{\text{inicial}}} \right)^{\frac{1}{n}} - 1 = \left( \frac{1,0337}{1,0} \right)^{\frac{1}{9}} - 1 \approx 0,0037 \text{ (0.37 \% por día)}. \quad (1)$$

donde  $C_{\text{final}} = 1,0337$ ,  $C_{\text{inicial}} = 1,0$ , y  $n = 9$  días. Se calcula el rendimiento anual mediante la siguiente fórmula:

$$g_a = (1 + 0,0037)^{365} - 1 \approx 2,84 \text{ (284 \% anual)}. \quad (2)$$

El análisis muestra que nuestro sistema tiene el potencial de generar rendimientos significativos a lo largo del tiempo (asumiendo consistencia del modelo en períodos más extensos).

## 4. Conclusiones

El presente trabajo ha demostrado la capacidad de los modelos multivariados basados en redes neuronales recurrentes, específicamente la combinación

de arquitecturas *LSTM* y *GRU*, para predecir con una precisión aceptable los precios de criptomonedas en un entorno altamente volátil. La incorporación de los valores de múltiples criptomonedas permitió capturar patrones temporales complejos e interdependientes no triviales entre los activos considerados.

En términos de desempeño predictivo, el modelo propuesto superó significativamente al enfoque univariado, logrando una reducción del 92 % en el error absoluto promedio (MAE). A nivel práctico, las predicciones del modelo se tradujeron en una estrategia de compra-venta exitosa, demostrando su aplicabilidad para optimizar decisiones financieras automatizadas.

Sin embargo, este estudio también identificó oportunidades de mejora y expansión que podrían incrementar la efectividad del modelo, entre las cuales destacan las siguientes:

1. La exploración de arquitecturas más avanzadas como los transformadores y las redes convolucionales temporales podría mejorar la precisión del modelo.
2. Ampliar el enfoque para predecir precios futuros en horizontes temporales extendidos permitiría desarrollar estrategias de compra-venta más sofisticadas.
3. Incorporar datos con diferentes resoluciones temporales, como precios por segundo o por hora, permitiría capturar patrones específicos de alta frecuencia y baja frecuencia.
4. Incluir indicadores técnicos, volúmenes de transacción o métricas de sentimiento del mercado podría fortalecer las predicciones.

En conclusión, este trabajo representa un paso significativo hacia el desarrollo de un sistema predictivo más robusto, haciendo posible la construcción de herramientas más precisas y versátiles para la toma de decisiones financieras.

## Referencias

- [1] O. Haykir and I. Yaglı, “Speculative bubbles and herding in cryptocurrencies,” *Financial Innovation*, vol. 8, 08 2022.
- [2] N. Zafar, I. U. Haq, J.-u.-R. Chughtai, and O. Shafiq, “Applying hybrid lstm-gru model based on heterogeneous data sources for traffic speed prediction in urban areas,” *Sensors*, vol. 22, no. 9, 2022. [Online]. Available: <https://www.mdpi.com/1424-8220/22/9/3348>
- [3] U.S. Securities and Exchange Commission, “Trading basics: Understanding the different ways to buy and sell stock,” 2021, accessed: 2024-11-24. [Online]. Available: <https://www.sec.gov/investor/alerts/trading101basics.pdf>