

Predicción del precio del Ether (ETH) usando el modelo Ornstein-Uhlenbeck y estadística bayesiana

Ether (ETH) price prediction using the Ornstein-Uhlenbeck model and Bayesian statistics

Jorfan Vargas Ruiz

Pregrado Ciencia de Datos Universidad Externado, Bogotá - Colombia

jorfan.vargas@est.uexternado.edu.co

RESUMEN

Con el paso del tiempo las criptomonedas han tomado cada vez más relevancia en el contexto de las finanzas descentralizadas (DeFi), un ecosistema que permite intercambiar activos y servicios financieros sin un ente central. La primera criptomoneda fue bitcoin y hoy en día ya hay miles de estas, entre ellas el Ethereum. En este artículo se estimarán los parámetros del proceso de Ornstein-Uhlenbeck a partir de estadística bayesiana para predecir el precio del Ethereum.

Palabras Clave: Ethereum, criptomonedas, simulación, Ornstein-Uhlenbeck, estadística bayesiana

Entrega: 14 de 10 de 2023.

ABSTRACT

Over time, cryptocurrencies have become increasingly relevant in the context of decentralized finance (DeFi), an ecosystem that allows assets and financial services to be exchanged without a central entity. The first cryptocurrency was bitcoin and today there are already thousands of these, including Ethereum. In this article we will estimate the parameters of the Ornstein-Uhlenbeck process from Bayesian statistics to predict the price of Ethereum.

Keywords: Ethereum, cryptocurrencies, simulation, Ornstein-Uhlenbeck, Bayesian statistics

1. INTRODUCCIÓN

Un sistema financiero más justo, es lo que propone la tecnología Ethereum, hasta el día de hoy millones de personas están limitadas a abrir cuentas bancarias y otras no pueden acceder a préstamos o transferencias sin un ente central, de esta manera surge Ether (ETH) “una criptomoneda. Es dinero digital limitado que puede usar en Internet, similar al bitcoin. Si es nuevo en el mundo de las criptomonedas, así es como ETH se diferencia del dinero tradicional”. [Ethereum Foundation,], también definida como “La criptomoneda nativa utilizada por el ecosistema Ethereum, que cubre los costes de gas al ejecutar contratos inteligentes”. [Wood and Antonopoulos, 2022], ahora bien, la importancia de Ether ha venido creciendo con el tiempo desde 2015 después del lanzamiento de Bitcoin en la que Vitalik Buterin propuso la plataforma Ethereum para hacer transacciones sin permiso y con permiso. La importancia de Ether está basada en su valor.

Se utiliza dentro de la red Ethereum para

realizar una serie de funciones, incluyendo:

- utilizado para pagar las comisiones de la transacción de Ethereum (en forma de gas)
- utilizado como garantía para una amplia gama de aplicaciones de financiación abierta (MakerDAO, Compound)
- se puede prestar o pedir prestado (Dharma)
- Aceptada como pago en determinados minoristas y proveedores de servicios
- acostumbrados a ser un medio de intercambio para comprar fichas basadas en Ethereum (a través de ICOs o intercambios), criptomonedas, artículos en el juego y otras fichas no fungibles (NFT)
- ganado como recompensa por completar recompensas (Gitcoin, Bounties Network)

Además del valor de la utilidad, Ether

también tiene valor especulativo. Este es el valor que se deriva de actividades especulativas (como negociación e inversión) que actualmente representa la mayor parte del valor detrás de todos los crypto-activos. [Sassano, 2019]

Para la predicción de las criptomonedas se han usado técnicas como esquemas básicos de regresión y correlaciones, esquemas basados en el análisis de series temporales incluyen el modelo GARCH y sus derivados, el modelo ARIMA, el análisis de coherencia wavelet, también se adoptaron la regresión lineal multivariante, el modelo autorregresivo vectorial multivariante y el modelo autorregresivo vectorial ampliado, el modelo de valor en riesgo (VaR), el método de mínimos cuadrados, chainlets, el modelo de mezcla temporal, la prueba Dicky-Fuller aumentada, las pruebas de Johansen para la cointegración, el modelo de corrección de errores vectoriales, el modelo de retardo distribuido autorregresivo, entre otras como lo menciona [Khedr et al., 2021]

Sin embargo, Ether es vulnerable a ataques, [Tikhomirov, 2018] menciona que aunque el precio del Ether lo determina el mercado, en septiembre de 2016, un atacante explotó una debilidad de la fijación del precio del gas aprovechando que ciertas operaciones tenían un precio inferior, es por esto que la predicción del precio de la criptomoneda está expuesta a los distintos factores externos que lo pueden afectar además del mercado, esto representa un problema a la hora de usar indicadores técnicos, pues no identifican estos comportamientos dado a eventos independientes y particularidades del mercado que aumentan la volatilidad de la criptomoneda, para abordar este inconveniente se hará el desarrollo de una metodología que mejore la eficiencia en la estimación de los parámetros del modelo Ornstein-Uhlenbeck(1) para predecir el precio del Ether (ETH), “El proceso Ornstein - Uhlenbeck es un proceso estocástico que fue propuesto por el físico Leonard Solomon Ornstein y el físico George Eugene Uhlenbeck”... [Mejia, 2017]....” se requiere algún tipo de tendencia para describir adecuadamente la situación modelizada”. [Maller et al., 2009]

$$dS_t = \lambda \cdot (\mu - S_t)dt + \sigma \cdot dW_t \quad (1)$$

2. Contextualización Teórica

La predicción de las criptomonedas es un campo de investigación relativamente nuevo, pero ha experimentado un crecimiento considerable en los últimos años. Esto se debe a la creciente popularidad de las criptomonedas y al aumento de

la demanda de herramientas que permitan a los inversores tomar decisiones informadas sobre sus inversiones.

Los métodos estadísticos son una herramienta importante para la predicción de las criptomonedas. Estos métodos permiten a los investigadores identificar patrones en los datos históricos del precio de las criptomonedas y utilizarlos para predecir su precio futuro.

Los métodos estadísticos más utilizados para la predicción de criptomonedas son el análisis de series temporales, el análisis de regresión y el aprendizaje automático. El análisis de series temporales utiliza datos históricos del precio de una criptomoneda para predecir su precio futuro. El análisis de regresión utiliza datos históricos del precio de una criptomoneda y de otros factores que influyen en su precio para predecir su precio futuro. El aprendizaje automático utiliza algoritmos de aprendizaje automático para identificar patrones en los datos históricos del precio de una criptomoneda y predecir su precio futuro.

3. Análisis Exploratorio

Los datos usados fueron extraídos por medio de Yahoo Finance por medio de el siguiente código en python:

```
import yfinance as yf
ethsymbol = "ETH-USD"
startdate = "2023-08-01"
enddate = "2023-10-05"
ethdata = yf.download(ethsymbol,
                      start=startdate, end=enddate,
                      interval="1h")
```

	Open	High	Low	Close	Adj Close	Volume
2023-08-01 05:00:00+00:00	1824.8751220703125	1829.3817138671875	1824.6591796875	1829.3817138671875	1829.3817138671875	0
2023-08-01 06:00:00+00:00	1829.566162109375	1835.2398681640625	1827.620361328125	1831.893310546875	1831.893310546875	80370688
2023-08-01 07:00:00+00:00	1833.091163089375	1834.305908203125	1829.5287890625	1834.305908203125	1834.305908203125	69719552
2023-08-01 08:00:00+00:00	1833.94189453125	1835.922607421875	1832.641845703125	1835.359130859375	1835.359130859375	24462848
2023-08-01 09:00:00+00:00	1836.501220703125	1836.501220703125	1832.1307373046875	1832.475830078125	1832.475830078125	34542592

Tabla. 1. Primeros 5 registros del Ethereum

La base de datos generada obtiene los registros del ETH desde 2023-08-01 hasta 2023-10-05 con una frecuencia de una hora, esta base de datos incluye los precios de apertura, cierre, cierre ajustado, precio más alto registrado en el intervalo y el más bajo y el volumen, donde:

Precio de Apertura: Es el precio en el que inicia el ETH. Se representa en la tabla como *Open*.

Precio Más Alto Registrado en el Intervalo: El precio más alto alcanzado durante un período de tiempo específico. Se representa en la tabla como *High*.

Precio Más Bajo Registrado en el Intervalo: El precio más bajo alcanzado durante un período de tiempo específico. Se representa en la tabla como *Low*.

Precio de Cierre: Es el precio en el que cierra el ETH y se representa en la tabla como *Close*.

Cierre Ajustado: Es el precio de cierre modificado después de contabilizar cualquier acción corporativa. Se representa en la tabla como *AdjClose*.

Volumen: La cantidad total de activos negociados en el período de tiempo determinado. Se representa en la tabla como *Volume*.

3.1 Análisis Descriptivo

	Open	High	Low	Close	Adj Close	Volume
count	1560.000000	1560.000000	1560.000000	1560.000000	1560.000000	1.560000e+03
mean	1689.993080	1693.052965	1686.860711	1689.880609	1689.880609	6.178017e+07
std	91.848748	91.502386	92.183591	91.784077	91.784077	1.716701e+08
min	1538.777588	1545.406250	1533.426758	1539.252930	1539.252930	0.000000e+00
25%	1629.933289	1632.357086	1627.104919	1629.926636	1629.926636	0.000000e+00
50%	1650.570496	1653.825500	1648.034119	1650.520325	1650.520325	0.000000e+00
75%	1788.628235	1790.413513	1783.222168	1787.096558	1787.096558	5.430618e+07
max	1873.468140	1877.511475	1867.970581	1873.029907	1873.029907	2.530970e+09

Tabla. 2. Estadística descriptiva de las variables de la base de datos.

Se obtiene las medidas de dispersión y tendencia central de cada variable de la base de datos, se puede evidenciar que los precios en el intervalo de tiempo establecido tienen una variabilidad de 91.84 en promedio aproximadamente.

#	Column	Non-Null Count	Dtype
0	Open	1560 non-null	float64
1	High	1560 non-null	float64
2	Low	1560 non-null	float64
3	Close	1560 non-null	float64
4	Adj Close	1560 non-null	float64
5	Volume	1560 non-null	int64

Tabla. 3. Verificación de datos faltantes y tipos de datos presentes en la base de datos.

No hay datos faltantes en la base de datos, y estos son de tipo entero y flotante.

3.1.1 Procesamiento de Datos

A partir de la información recolectada en la base de datos se generaron dos variables más para analizar y visualizar de mejor forma el comportamiento del ETH en el período de tiempo determinado, se generó la nueva variable SMA24 que indica la media móvil de los precios de cierre del Ethereum con 24 horas de referencia y de la misma manera se generó la nueva variable Volatility24 que hace referencia a la volatilidad del precio de cierre con 24 horas de referencia.

Fecha	SMA_24	Volatility_24
2023-10-05 00:00:00+00:00	1643.133691	5.340981
2023-10-05 01:00:00+00:00	1643.482697	5.368766
2023-10-05 02:00:00+00:00	1643.618890	5.355127
2023-10-05 03:00:00+00:00	1643.532336	5.349175
2023-10-05 04:00:00+00:00	1643.438531	5.308202

Tabla. 4. Variables media móvil y volatilidad generadas a partir de los datos

3.2 Presentación de Gráficas

EN esta se presentan solo las gráficas pertinentes para la solución del problema. Deben haber al menos cuatro gráficas.

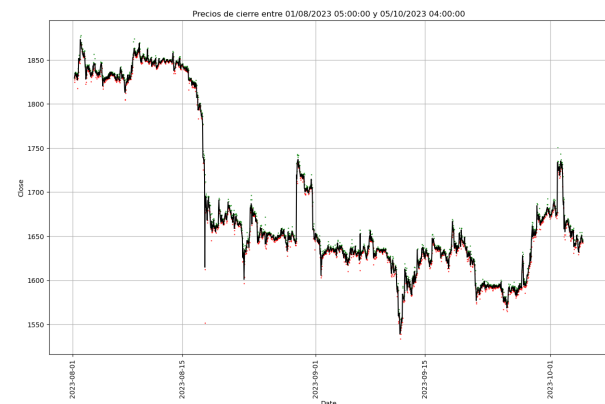


Fig. 1. Precios de cierre del ETH entre la fecha de inicio y la fecha final, en cada fecha está representado por un punto verde el precio más alto y por un punto rojo el precio más bajo.

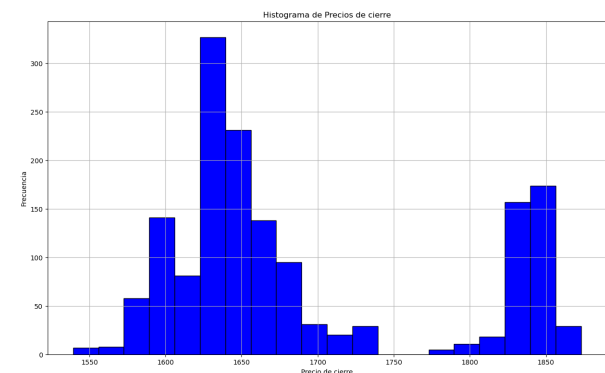


Fig. 2. Distribución de los precios de cierre.

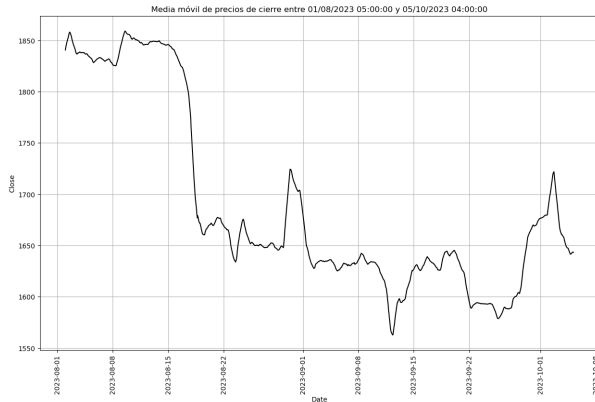


Fig. 3. Media móvil de 24 horas respecto a la fecha de inicio y la fecha final.

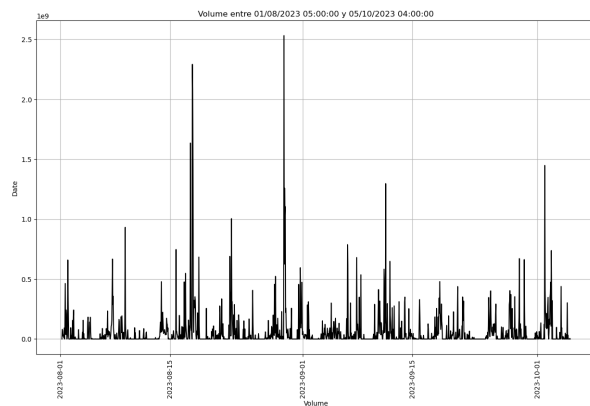


Fig. 4. Cantidad de activos negociados en el periodo de una hora respecto a la fecha de inicio y la fecha final.

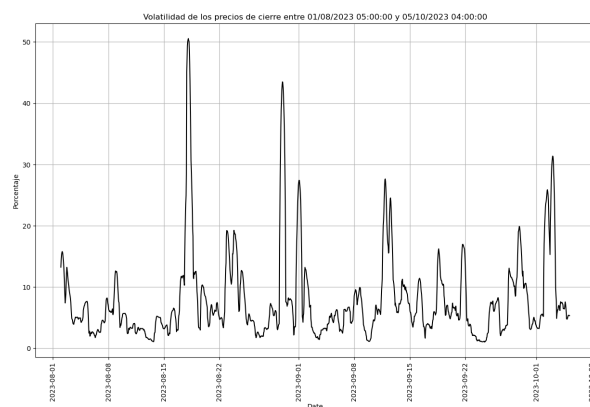


Fig. 5. Volatilidad de los precios de cierre con ventana móvil de 24 horas desde la fecha de inicio y la fecha final.

3.3 Presentación de las Priors

Proceso Ornstein-Uhlenbeck

$$dS_t = \lambda \cdot (\mu - S_t)dt + \sigma \cdot dW_t \quad (2)$$

Presentación de la priors

El objetivo de esta investigación es estimar los hiperparámetros a partir de estadística bayesiana del proceso Ornstein-Uhlenbeck, para esto se debe definir los parámetros y sus distribuciones prior:

$$\lambda \sim \mathcal{N}(\mu_0, \tau_0)$$

λ : Indica la velocidad con la que el precio del Ethereum vuelve a un estado de normalidad luego de eventos en los que el precio cambia considerablemente como se ve en el gráfico 1 a mediados de 18-08-2023, λ toma valores positivos y negativos cercanos a 0 por lo que se usará una distribución normal con parámetros μ_0 Y τ_0 .

$$\mu \sim \mathcal{N}(\mu_0, \tau_0)$$

μ : Indica el promedio de los precios del ETH, este hiperparámetro toma valores continuos y sin restricciones, y con información para μ_0 Y τ_0 , pues son valores que se pueden suponer con información previa del ETH.

$$\sigma \sim \text{Inv} - \text{Gamma}(\alpha, \beta)$$

σ : La desviación estándar solo puede tomar valores positivos como podemos ver en el gráfico 5.

BIBLIOGRAFÍA

- [Ethereum Foundation,] Ethereum Foundation. ¿Qué es el ether (ETH)?
- [Khedr et al., 2021] Khedr, A. M., Arif, I., P V, P. R., El-Bannany, M., Alhashmi, S. M., and Sreedharan, M. (2021). Cryptocurrency price prediction using traditional statistical and machine-learning techniques: A survey. *Intelligent Systems in Accounting, Finance and Management*, 28(1):3–34.
- [Maller et al., 2009] Maller, R. A., Müller, G., and Szymayer, A. (2009). Ornstein–uhlenbeck processes and extensions. pages 421–437.
- [Mejia, 2017] Mejia, C. (2017). The ornstein-uhlenbeck process. an introduction for commodity modelling with some extensions. an example with the london cocoa net spot instantaneous convenience yield (el proceso ornstein-uhlenbeck. una introducción para el modelaje en commodities con algunas extensiones. un ejemplo con la

tasa neta, spot e instantanea de conveniencia asociada al cocoa futures de londres). *SSRN Electronic Journal*.

[Sassano, 2019] Sassano, A. (2019). Why ether is valuable - ethhub - medium. *EthHub*.

[Tikhomirov, 2018] Tikhomirov, S. (2018). *Ethereum: State of knowledge and research perspectives*. Springer International Publishing.

[Wood and Antonopoulos, 2022] Wood, G. and Antonopoulos, A. M. (2022). *Mastering Ethereum: Creando Contratos Inteligentes y DApps*. Independently Published.