

Universidad de Lima
Facultad de Ingeniería
Carrera de Ingeniería Industrial



CARDANO CRYPTOCURRENCY PRICE FROM TWITTER. A PREDICTION ALGORITHM FROM MACHINE LEARNING

Tesis para optar el Título Profesional de Ingeniero Industrial

Riccardo Piccarreta Acosta

Código 20181460

Alejandra Zavala Arana

Código 20182107

Asesor

Juan Carlos Quiroz Flores

Lima – Perú

Setiembre de 2024

Título

CARDANO CRYPTOCURRENCY PRICE FROM TWITTER. A PREDICTION ALGORITHM WITH MACHINE LEARNING

Autor(es)

Riccardo Piccarreta Acosta
20181460@aloe.ulima.edu.pe
Universidad de Lima

Alejandra Zavala Arana
20182107@aloe.ulima.edu.pe
Universidad de Lima

Resumen: Criptomonedas son un mercado creciente que ha atraído la atención de varios inversionistas en los últimos años. Si bien las criptomonedas ofrecen una forma de pago segura y descentralizada, este mercado es muy volátil. Los factores que influyen en los cambios de precios incluyen el equilibrio de la oferta y la demanda, su utilidad, los indicadores de trading y la confianza del mercado. El objetivo de la presente investigación es predecir el precio de la criptomoneda Cardano mediante el uso de técnicas de aprendizaje automático, específicamente modelos de SVM, LSTM y BiLSTM. Además de tomar en cuenta índices financieros, se utilizó como fuente de datos la actividad en Twitter para medir la opinión del mercado. El estudio analiza diversos horizontes predictivos, incluyendo rangos de tiempo de 1 día, 7 días, 14 días, 21 días y 30 días. Los resultados obtenidos fueron validados con distintos indicadores de desempeño y se determinó que el modelo predice precios de Cardano con un mes de anticipación con un MAPE menor al 22% brindando información valiosa para los inversores interesados en el mercado volátil de la criptomoneda Cardano.

Palabras Clave: Cardano, criptomonedas, Machine Learning, Neural Network, Twitter

Abstract: Cryptocurrencies are a growing market that has attracted the attention of many investors in recent years. While cryptocurrencies offer a secure and decentralized form of payment, this market is highly volatile. Factors influencing price changes include the balance of supply and demand, its utility, trading indicators, and market confidence. The present research aims to predict the price of the Cardano cryptocurrency by using machine learning techniques, specifically SVM, LSTM and BiLSTM models. In addition to accounting for financial indices, Twitter activity was used as a data source to measure market sentiment. The study analyzes various predictive horizons, including time ranges of 1 day, seven days, 14 days, 21 days and 30 days. The results obtained were validated with different performance indicators, and it was determined that the model predicts Cardano prices one month ahead with a MAPE of less than 22%, providing valuable information for investors interested in the volatile Cardano cryptocurrency market.

Keywords: Cardano, Cryptocurrencies, Machine Learning, Neural Network, Twitter.

Línea de investigación IDIC – ULIMA (6) – Desarrollo empresarial

Área y Sub-áreas de Investigación: (3) Ingeniería de análisis económico

Objetivo (s) de Desarrollo Sostenible (ODS): (8) - Trabajo decente y crecimiento económico.

PLANTEAMIENTO DEL PROBLEMA

El mercado de criptomonedas, altamente volátil y sensible a factores como oferta, demanda y sentimiento del mercado, presenta un desafío para los inversores que buscan predicciones precisas de precios. Cardano (ADA), con su mecanismo de Proof-of-Stake, es de particular interés para inversores e investigadores.

Diversos estudios han explorado la predicción de precios de criptomonedas utilizando técnicas de aprendizaje automático y análisis de sentimientos. Se han utilizado librerías de Python como TextBlob Sentiment (Agarwal et al., 2021) para analizar tweets con modelos de Procesamiento de Lenguaje Natural (Regal et al., 2019) y correlacionarlos con los precios de criptomonedas. Para predecir el precio de la criptomoneda se emplearon modelos LSTM (Kilimchi, 2020).

La mayoría de los estudios se han centrado en Bitcoin y Ethereum, dejando un vacío en la investigación relacionada con otras criptomonedas como Cardano. Este estudio aborda esta brecha integrando datos de redes sociales y financieros para predecir los precios de Cardano utilizando aprendizaje automático.

OBJETIVOS

El objetivo de esta investigación es predecir el precio de la criptomoneda Cardano (ADA) usando técnicas de aprendizaje automático basadas en el análisis de sentimientos de Twitter y datos financieros, evaluando la precisión en diferentes horizontes temporales (Kraaijeveld & De Smedt, 2020). Se recolectarán y procesarán datos de Twitter y Yahoo Finance, se realizará un análisis de sentimientos usando NLP y se integrarán estos datos con indicadores financieros. Se desarrollarán y entrenarán modelos de aprendizaje automático, como LSTM y SVM (Ricardo Leon-Ayala et al., 2022), evaluando su precisión con métricas de error (Warres & Jurman, 2021)

JUSTIFICACIÓN

Este busca ampliar el conocimiento sobre la predicción de precios de criptomonedas mediante técnicas avanzadas de aprendizaje automático y análisis de sentimientos, proporcionando una base sólida para futuras investigaciones en el campo financiero. Técnicamente, introduce el uso de modelos como BiLSTM, LSTM y SVM, demostrando la capacidad de estas tecnologías para integrar y procesar datos de redes sociales y financieros de manera efectiva. Económicamente, la mejora en la precisión de las predicciones de precios beneficia a los inversores al reducir riesgos y aumentar oportunidades de ganancias, lo que a su vez puede incrementar la estabilidad y confianza en el mercado de criptomonedas. Socialmente, el proyecto aporta transparencia al mercado y equidad en el acceso a información crítica, ayudando tanto a pequeños inversores como a reguladores a tomar decisiones más informadas y justas. Metodológicamente, ofrece un enfoque interdisciplinario robusto y adaptable, que combina diversas fuentes de datos y técnicas avanzadas de análisis, y puede ser replicado en otros estudios financieros e informáticos. Además, contribuye al Objetivo de Desarrollo Sostenible 8 al fomentar el crecimiento económico sostenido, inclusivo y sostenible, promoviendo el empleo pleno y productivo y el trabajo decente para todos, al impulsar la innovación y eficiencia en el sector financiero.

HIPÓTESIS

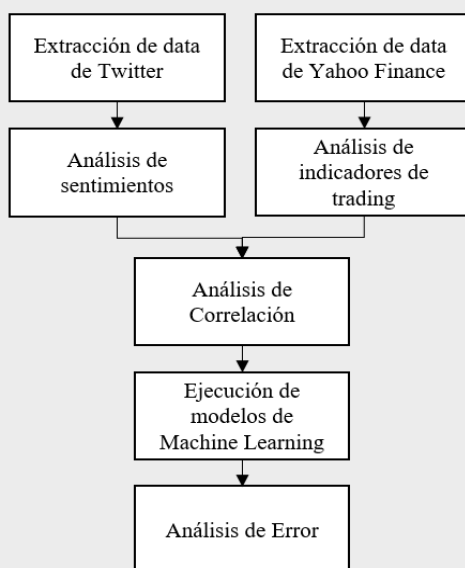
El análisis de sentimientos de Twitter, combinado con datos financieros y procesado mediante técnicas de aprendizaje automático, puede predecir con precisión el precio de la criptomoneda Cardano en diferentes horizontes temporales.

DISEÑO METODOLÓGICO

La investigación fue de carácter experimental, procesando la polaridad de los tweets y los indicadores de compra y venta de acciones como variables independientes para predecir el precio de la criptomoneda Cardano, la variable dependiente.

Con un alcance correlacional, se evaluó la relación entre estas variables y se predijo su comportamiento. El enfoque mixto integró datos cuantitativos (precio e indicadores) y cualitativos (tweets). Utilizando el procesamiento de lenguaje natural (NLP) para el análisis de sentimientos y modelos de aprendizaje supervisado como el *Support Vector Regressor* (SVR), y redes neuronales como el *Long-Short Term Memory* (LSTM) y *Bidirectional Long-Short Term Memory* (BiLSTM), se desarrollaron los algoritmos de predicción. Los datos fueron extraídos de Twitter (enero 2021 a abril 2023) y Yahoo Finance, con una muestra de 742,757 tweets en 26 meses.

La metodología de la investigación se estructuró en siete fases principales. Primero, se extrajeron tweets relacionados con Cardano y datos históricos de precios y volumen de transacciones de Yahoo Finance utilizando librerías de Python como *snsrape* y *yfinance*. Luego, se realizó un análisis de sentimientos con la librería *TextBlob* para evaluar la polaridad y subjetividad de los tweets. A continuación, se integraron estos datos de sentimiento con los indicadores financieros para crear un conjunto de datos consolidado. Posteriormente, se implementaron y entrenaron modelos de aprendizaje automático, como BiLSTM, LSTM y SVM, para predecir el precio de Cardano. Finalmente, se evaluó la precisión de los modelos utilizando métricas de error como RMSE, MAE, MAPE y MSE, y se realizó un análisis de correlación para validar las relaciones entre las variables.



NOTAS (AGRADECIMIENTOS)

Queremos agradecer a nuestros padres por habernos dado la oportunidad de estudiar en la Universidad de Lima y permitirnos expandir nuestros conocimientos y visión sobre el mundo.

REFERENCIAS

Agarwal, B., Harjule, P., Chouhan, L., Saraswat, U., Airan, H., & Argawal, P. (2021). Prediction of dogecoin price using deep learning and social media trends. *EAI Endorsed Transactions on Industrial Networks and Intelligent Systems*, 8(29). doi:10.4108/EAI.29-9-2021.171188

Altan A, Karasu S, Bekiros S (2019) Digital currency forecasting with chaotic meta-heuristic bio-inspired signal processing techniques. *Chaos Soliton Fract* 126:325–336

Chicco, D., Warrens, M., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *Computer Science*. doi:10.7717/peerj-cs.623

Cohen G (2020) Forecasting bitcoin trends using algorithmic learning systems. *Entropy* 22(8):838

Das, H., Kumar, J., Chandra, S., & Dey, N. (2020), *Applied Intelligent Decision Making in Machine Learning*. Florida: CRC

Delfabbro, P., King, D. L., & Williams, J. (2021). The psychology of cryptocurrency trading: Risk and Protective factors. *Journal of Behavioral Addictions*, 10(2), 201-207. doi:10.1556/2006.2021.00037

Dupire, M., Fibien, J., & M'Zali, B. (2022). Non-Governmental Organization (NGO) Tweets: Do Shareholders Care? *Business & Society*, 61(2), 419-456. doi:10.1177/0007650320985204

Gestión. (2023, 05, 20). Luna: el caso de la criptomoneda que colapsó en cuestión de días. Recuperado de <https://gestion.pe/tendencias/luna-el-caso-de-la-criptomoneda-que-colapso-en-cuestion-de-dias-bitcoin-nnda-nnlt-noticia/>

Hamraoui, I., & Boubaker, A. (2022). Impact of Twitter sentiment on stock price returns. *Social Network Analysis and Mining*, 12(1), 1-15. doi:10.1007/s13278-021-00856-7

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-term Memory. *Neural Computation*, 9(8), 1735-1780. doi:10.1162/neco.1997.9.8.1735

Johnson, R. A., & Wichern, D. W. (2022). *Applied Multivariate Statistical Analysis* (6th ed.). Pearson.

Jolly, J. (4 de febrero de 2021). Price of dogecoin rises by 50% following Elon Musk tweet. *The Guardian*. <https://www.theguardian.com/business/2021/feb/04/price-of-dogecoin-rises-by-50-following-elon-musk-tweet>

Kelley, N. J., Davis, W. E., Dang, J., Liu, L., Wildschut, T., & Sedikides, C. (2022). Nostalgia confers psychological wellbeing by increasing authenticity. *Journal of Experimental Social Psychology*, (102) doi: 10.1016/j.jesp.2022.104379

Kilimci, Z. (2020). Sentiment analysis based direction prediction in bitcoin using deep learning algorithms and word embedding models. *International Journal of Intelligent Systems and Applications in Engineering*, 8(2), 60-65. doi:10.18201/ijisae.2020261585

Kraaijeveld, O., & De Smedt, J. (2020). The predictive power of public Twitter sentiment for forecasting cryptocurrency prices. *Journal of International Financial Markets, Institutions and Money*, 65. doi:10.1016/j.intfin.2020.101188

Lahmiri, S., Bekiros, S., & Bezzina, F. (2022). Complexity analysis and forecasting of variations in cryptocurrency trading volume with support vector regression tuned by Bayesian optimization under different kernels: An empirical comparison from a large dataset. *Expert Systems with Applications*, 209, 118349. doi:10.1016/j.eswa.2021.118349

Leon, R., Vicente, N., Quispe, N., Mesa, L., Saldaña, B., Delgado, A., & Huamaní, E. (2022). Predictive Analytics to Determine the Bitcoin Price Rise using Machine Learning Techniques. *International Journal of Engineering Trends and Technology*, 70(3), 275-283. doi:10.14445/22315381/IJETT-V70I3P231

Li, Y., Jiang, S., Li, X., & Wang, S. (2022). Hybrid data decomposition-based deep learning for Bitcoin prediction and algorithm trading. *Financial Innovation*, 8(1), 1-24. doi:10.1186/s40854-022-00336-7

Maxwell, Scott & Delaney, Harold & Kelley, Ken. (2017). *Designing Experiments and Analyzing Data: A Model Comparison Perspective*. doi: 10.4324/9781315642956.

McCoy, M., & Rahimi, S. (2020). Prediction of Highly Volatile Cryptocurrency Prices Using Social Media. *International Journal of Computational Intelligence and Applications*, 19(4), 1-28. doi:10.1142/s146902682050025x

Mojtaba, N., Pooyan, N., Hamed, J., Shahab, S., & Amir, M. (2020). Predicting Stock Market Trends Using Machine Learning and Deep Learning Algorithms Via Continuous and Binary Data; a Comparative Analysis. *IEEE Access*, 8, 150199-150212. doi:10.1109/ACCESS.2020.3015966

Montgomery, D. C. (2017). *Design and Analysis of Experiments*. Wiley.

Mudassir, M., Bennbaia, S., Unal, D., & Hammoudeh, M. (2020). Time-series forecasting of Bitcoin prices using high-dimensional features: a machine learning approach. *Neural Computing and Applications*, 1, 1-15. doi:10.1007/s00521-020-05129-6

Nyakurukwa, K., & Seetharam, Y. (2022). Can a 280-character message explain stock returns? Evidence from South Africa. *Managerial Finance*, 48(4), 663-686. doi:10.1108/MF-12-2021-0598

Pradhyumna, R., Nishit, B., Puneet, G., & Lakshita, A. (2021). Crypto currency portfolio allocation using Machine Learning. *International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)* (págs. 1522-1527). Delhi-NCR: Institute of Electrical and Electronics. doi:10.1109/ICAC3N53548.2021.9725704

Regal, A., Morzán, J., Fabbri, C., Herrera, G., Yaulli, G., Palomino, A., & Gil, C. (2019). Cryptocurrency price projection based on tweets using LSTM. *Ingeniare*, 27(4), 696-706. Obtenido de <https://doi.org/10.4067/S0718-33052019000400696>

Rouhani, S., & Abedin, A. (2020). Crypto-currencies narrated on tweets: a sentiment analysis approach. *International Journal of Ethics and Systems*, 36(1), 58-72. doi:10.1108/IJOES-12-2018-0185

Valle, D., Fernandez, V., López, A., & Sandoval, R. (2022). Does Twitter Affect Stock Market Decisions? Financial Sentiment Analysis During Pandemics: A Comparative Study of the H1N1 and the COVID19 Periods. *Cognitive Computation*, 14(1), 372-387. doi:10.1007/s12559-021-09819-8

Wirawan IM, Widiyaningtyas T, Hasan MM (2019) Short term prediction on bitcoin price using ARIMA method. In: 2019

International Seminar on Application for Technology of Information and Communication (iSemantic). IEEE

Yamamoto, H., Sakaji, H., Matsushima, H., Yamashita, Y., Osawa, K., Izumi, K., & Shimada, T. (2020). Forecasting Crypto-Asset Price Using Influencer Tweets. *Advances in Intelligent Systems and Computing*. 926, págs. 940-951. Matsue: University of Tokyo. doi:10.1007/978-3-030-15032-7

ANEXO. Datos del artículo publicado

- **Nombre del artículo:** Cardano Cryptocurrency Price from Twitter. A Prediction Algorithm from Machine Learning
- **Autores:** Riccardo Piccarreta Acosta, Alejandra Zavala Arana
- **Co autor(es):** Yvan Jesús García López, Juan Carlos Quiroz Flores

Publicación en revista

- **Nombre de la revista:** International Journal of Electronics and Communication Engineering
- **Volumen:** 10
- **Número:** 12
- **Año:** 2023
- **Pp:** 33-44
- **Enlace web donde se encuentra publicado el artículo (identificador DOI, ISBN, ISSN o equivalentes):** <https://doi.org/10.14445/23488549/IJECE-V10I12P104>

9% Similitud general

El total combinado de todas las coincidencias, incluidas las fuentes superpuestas, para ca...




Filtrado desde el informe

- ▶ Bibliografía
- ▶ Coincidencias menores (menos de 15 palabras)

Exclusiones

- ▶ N.º de coincidencias excluidas

Fuentes principales

- 8%  Fuentes de Internet
- 5%  Publicaciones
- 7%  Trabajos entregados (trabajos del estudiante)

Marcas de integridad

N.º de alertas de integridad para revisión

No se han detectado manipulaciones de texto sospechosas.

Los algoritmos de nuestro sistema analizan un documento en profundidad para buscar inconsistencias que permitirían distinguirlo de una entrega normal. Si advertimos algo extraño, lo marcamos como una alerta para que pueda revisarlo.

Una marca de alerta no es necesariamente un indicador de problemas. Sin embargo, recomendamos que preste atención y la revise.

