Cryptocurrencies and Price Prediction: A survey

Yeray Mezquita[1*], Ana Belén Gil-González[1], Javier Prieto[1], and Juan Manuel Corchado[1]

BISITE Research Group, University of Salamanca.
Edificio Multiusos I+D+i, 37007, Salamanca, Spain
yeraymm@usal.es*

Abstract. Cryptocurrencies are fungible digital assets whose market capitalization has not stopped growing since the appearance of their first use case in 2009, Bitcoin. However, one of the biggest problems facing cryptocurrencies is the enormous fluctuation of their value in the market. To help understand different patterns in cryptocurrency ecosystems, several machine learning-based solutions have been proposed in the literature. This paper aims to study in detail the solutions proposed in the literature for the detection of patterns and anomalies in cryptocurrency ecosystems. The aim is to bring together different proposals and studies to help users of this market to understand how it works.

Keywords: Cryptocurrencies, Price, Prediction, Survey

1 Introduction

Cryptocurrencies are fungible digital assets whose market capitalization has not stopped growing since the appearance of their first use case in 2009, Bitcoin [17]. Many investors are looking to put their money in cryptocurrencies, considering them an asset that will gain value over time because of the possibilities they offer, either by the possibility of payment automation in distributed platforms with the use of smart contracts, or by the creation of a more democratic money market thanks to being considered distributed assets [14, 13–15].

However, one of the biggest problems facing cryptocurrencies is the enormous fluctuation of their value in the market. This prevents them from being used in people's daily lives as a means of transaction. Instead, they are used as means of investment, being exposed to speculation by investors, sometimes even expossed to pump and dump schemes [18, 8].

To help understand different patterns in cryptocurrency ecosystems, several machine learning-based solutions have been proposed in the literature [10, 19, 21]. Identifying these patterns can help users not to enter the game of pump and dump schemes, eliminating some risk and attracting more people to the market, which would help stabilize crypto prices [22].

This paper aims to study in detail the solutions proposed in the literature for the detection of patterns and anomalies in cryptocurrency ecosystems. The aim is to bring together different proposals and studies to help users of this market to understand how it works. To this end, the section 2 shows a theoretical context of cryptocurrencies and the technology that underlies them: the blockchain. In section 3 the research questions and search criteria are defined. In addition, the papers found are detailed and the section is developed by grouping them according to what they propose. Finally, in the 4 section, the findings of the work are discussed and future work is concluded.

2 Theoretical Context

Cryptocurrencies are fungible digital assets whose main purpose is to being used in transactions as fiat money. Blockchain technology is used to store cryptocurrency transactions and maintain user balances. This technology is a type of distributed ledger, the ledger where the transactions are stored is supported and maintained by a network of nodes that communicate with each other through a series of protocols. Thanks to this, it is possible to have a tamper-proof ledger of transactions and balances, earning the trust of their users.

Blockchain technology makes use of a public key signature mechanism, thanks to which it is possible to easily verify the source of the data generated, guaranteeing the integrity of the data and easily verifying its origin. Due to the open, decentralized, and cryptographic nature of a blockchain, it is possible to i) avoid intermediaries in transactions between non trusted parties; and ii) generate an immutable ledger kept by the network of nodes, by using consensus protocols, nodes of the network make the ledger tamper-proof [15].

There are being designed more types of consensus algorithms and their variations, each depending on the type of the blockchain network. If a blockchain is public, everyone can access it and connect to it, or if it is permissioned, only identified players can connect to it and maintain the ledger. The most widespread algorithms are [16]: i) the **Proof-of-Work** (PoW) algorithm, a network node must solve a cryptographic problem to add a new block to the blockchain. The computational cost and difficulty of solving the problem, the energy expended in finding its solution (work), and the simplicity of its verification, are sufficient reasons to deter nodes that add new blocks (miners) from engaging in illegal transactions; ii) the **Proof-of-Stake** (PoS) is a consensus algorithm, in which miners take turns adding new blocks to the blockchain. The probability of a miner getting his turn to add a block depends on the number of coins deposited for the miner as escrow (Stake). This algorithm assumes that a node is going to be honest in creating a block to avoid losing escrow; the **Practical Byzantine Fault Tolerance** (PBFT). Round is the process of adding a new block to the blockchain. In each round, a node is selected to propose a new block, and then the block is broadcasted to the network, to let the network validate it. The block is validated by each node in the network, getting one vote for each node that successfully validated it. When a block receives 2/3 of the votes from all nodes in the network, it is considered valid and is added to the blockchain.

Another feature of blockchain technology that some cryptocurrencies like Ethereum are taking most profit from, is the possibility of deploy smart contracts. Smart contracts are a sequence of code stored within a blockchain that can be executed in a distributed manner alongside the network, reaching a consensus among the nodes on the result obtained from its execution. These programs facilitate, verify, and enforce an agreement on a set of predefined conditions [7]. Smart contracts are self-enforcing and self-verifying contractual agreements that automate the lifecycle of a contract to improve compliance, mitigate risk and increase efficiency on any platform where entities with different interests have to interact with each other [16].

Although blockchain technology offers a great variety of possibilities, there also poses some challenges. E.g. each consensus algorithm has its risks and vulnerabilities, like PoW that spends a huge amount of energy to solve cryptographic problems to produce new blocks, and it is very limited in terms of scalability [1]. In the case of the PoS algorithm, its nothing at stake theory causes forks to occur on the blockchain more frequently than with other consensus algorithms [11]. Meanwhile, the main risk of PBFT is that it is a permissioned protocol and not a truly decentralized one, making it not suitable for public networks [24]. Finally, the use of smart contracts is not covered by the law, making it difficult to resolve disputes that are not well covered in the contract.

To address the scalability problems that PoW based blockchain platforms, are arising PoS based blockchain platforms, and some of its variations like the delegated Proof-of-Stake (dPoS) algorithm, in which participants vote, based on the stake they hold, which nodes can add new data to the blockchain. Some cryptocurrencies, like Tron, Cardano, EOS, and possibly Ethereum, adopt this approach to provide solutions to the scalability problems of the PoW based solutions and make it possible for Decentralized Applications (DApps) to become mainstream.

3 Literature Review

The main objective of this paper is to study the possibility of predicting cryptocurrency market fluctuations and how they can be done. To achieve the objectives, a study has been carried out in the Science direct database to find papers that help to find answers to the following hypotheses:

- **RQ1**: What has been used in the literature to forecast prices in the crypto market?
- **RQ2**: Are the methods used reliable?

The keywords used were: i) *("distributed ledger technology" OR blockchain* OR cryptocurrenc*)*, identified as Cryptocurrency; ii) *price*; iii) and *prediction*. Using the keywords, the following search string has been formulated: *Cryptocurrency AND Price AND Prediction*. Using this search string, 1004 results were obtained, of which the most relevant ones were used in this work, obtaining 13 papers in which prediction in cryptocurrency markets is studied.

In previous studies such as the one conducted in [20], the influence that news and social networks have on emerging technology, such as the case of distributed Ledger Technologies (DLT), more specifically blockchain technologies, and somewhat unknown to the general public, has been tested. In that study, only English-language news referring to such technologies from the period from 2010 to 2018 have been used. It is revealed a pattern of DLT diffusion in transition from its infancy, through a growing awareness of its potential, and favor and opposition to its applications. Additional sentiment analysis reveals that the likelihood of unfavorable sentiment toward Bitcoin remained higher than favorable sentiment.

Following the steps of the previous study, in [10] we seek to use news, more specifically published Tweets, and use them as a means to predict cryptocurrency prices. Authors obtained daily and hourly Twitter sentiment polarity for 9 major cryptocurrencies: Bitcoin (BTC), Ethereum (ETH), Bitcoin Cash (BCH), Litecoin (LTC), Cardano (ADA), Stellar (XLM), EOS, TRON, Ripple (XRP). Causality tests indicate Twitter sentiment help predict the returns of BTC, BCH, LTC. Using a bullishness ratio, predictive power is found for the returns of EOS TRON. Tweet volume only has predictive power for the price returns of LTC and XRP. Finally, it is estimated that 14% of Tweets were identified as posted by Twitter bot accounts.

In [23], it has been tried to assess the properties of the cryptocurrency market and the associated phenomena, by studying the characteristics of the complexity of exchange rates on the cryptocurrency market and comparing it to traditional and mature markets, such as stocks, bonds, commodities or currencies. With the help of statistical physics methods like the multifractal cross-correlation analysis and the dependent detrended cross-correlation coefficient, the non-linear correlations and multiscale characteristics of the top 100 cryptocurrencies of the cryptocurrency market are analyzed. Thanks to the study carried out, it has stated that due to the Covid-19 pandemic it has been witnessed a "phase transition" of the cryptocurrencies from being a hedge opportunity for the investors fleeing the traditional markets to become a part of the global market that is substantially coupled to the traditional financial instruments like the currencies, stocks, and commodities.

Cryptocurrency price prediction has become a trending research topic globally. Many machine learning and deep learning algorithms such as Gated Recurrent Unit (GRU), Neural Networks (NN), and Long short-term memory (LSTM) have been used by the researchers to predict and analyze the factors affecting the cryptocurrency prices [19]. In the referenced paper, a LSTM and GRU-based hybrid cryptocurrency prediction scheme is proposed, which focuses on only two cryptocurrencies, namely Litecoin and Monero. The results depict that the proposed scheme accurately predicts the prices with high accuracy, revealing that the scheme can be applied in various cryptocurrencies price predictions.

Regarding the ETH cryptocurrency, [21] has proposed the use of two machine learning methods, namely linear regression (LR) and support vector machine (SVM), by using a time series consisting of daily ETH closing prices. When using the proposed model of this work, the SVM method has higher accuracy (96.06%) than the LR method (85.46%).

Meanwhile, in [3], it is made an attempt to apply machine learning techniques on the index and constituents of cryptocurrency with a goal to predict and forecast prices thereof. In particular, the purpose of the referenced work is to predict and forecast the closing price of the cryptocurrency index 30 and nine constituents of cryptocurrencies using machine learning algorithms and models so that it becomes easier for people to trade these currencies. Results show a 92.4% accuracy using an ensemble learning method, which is considered as the best among all the models used in the referenced paper. Using such prediction and forecasting methods, people can easily understand the trend and it would be even easier for them to trade in a difficult and challenging financial instrument like cryptocurrency.

The study of [4], employed a variety of mathematical tools for the analysis of six significant cryptocurrencies: BTC, ETH, XMR, LTC, XRP, ETC. It used fractional and fractal mathematical tools as an instrument that can help market agents and investors to more clearly assess the cryptocurrencies price dynamics and, guiding investment decisions more assertively while mitigating risks. The BTC was the only cryptocurrency that presented more consistent long-memory behavior. The LTC exhibited the lowest predictable horizon compared to the other cryptocurrencies, pointing to chaotic behavior. The ETH and the XMR behaved mainly as an anti-persistent process showing a short memory effect. In this study, it can be concluded that, with exception of BTC, the other five cryptocurrencies analyzed are mean reverting, showing lower predictability.

When it comes to making predictions, the literature gives us a warning on the subject of regulations in the different countries [2]. In the referenced paper, it is examined whether cryptocurrency traders perceive the market regulation beneficially. Using an event study methodology for daily data covering the period 2015-2019, it is assessed how regulatory news and events have affected returns in the cryptocurrency market. The results suggest that events that increase the probability of a regulation adoption are associated with a negative abnormal return for concerning cryptocurrencies. It is also stated that investors reacted less negatively for most illiquid cryptocurrencies, as well as for cryptocurrencies that had more information asymmetry. From this work, it can be concluded that the performance in the pre-event period is positive and significant; however, it appears to be not significantly different from zero in the post-event period.

In [6], it has been examined the returns and volatility spillover with the cryptocurrency market over the period 1 September 2017 to 7 June 2018, on cryptocurrency-linked stocks (CLS) engagement with blockchain technology in Australia. In the referenced work, it is used the spillover methodology of [5], finding significant unidirectional return spillover and weak volatility spillover from the cryptocurrency market to CLS, after controlling return dynamics of the Australian dollar, gold, and commodity. The results indicate that investors incorporate the price dynamics of cryptocurrencies into their trading decisions for CLS.

From [9], can be found that several corporate blockchain developments have generated significant spillovers. So, after analyzing whether corporate blockchain patent developments influence BTC volatility, it can be concluded that its usage by corporations reaffirms investor confidence in BTC. Those results also provide evidence of the maturity of the cryptocurrency markets.

In [8] it is gathered the most comprehensive dataset on cryptocurrency pump and dump schemes. From that study it is found that i) pumped coins experience a modest price increase, but it is higher for less popular coins; ii) brazen pumps are more “successful” than those obscuring their corrupt intent; iii) and high concentration among pump channels creates an opportunity for regulators. In the work [18], it has been found that i) fraud groups use Telegram and other social media to organize and conduct pumps; ii) market signals are better than administrative social signals as predictive factors; iii) and changes in market movements could influence robustness in performance over time.

To conclude our study, in [12], it is stated that Directed Acyclic Graphs (DAGs) are an increasingly valuable substitution for blockchain technology, possessing several key strategic advantages over blockchains, namely their lack of transaction fees and increased speeds. Three cryptocurrencies already employ DAGs: IOTA, NANO, and Obyte. All three assets are built around a different data structure (The Tangle, a Block Lattice, and a Main-Chain) and are resistant to quantum computing, meaning they already possess an important long-term advantage over currencies such as BTC, ETH, and LTC. In the referenced paper, it is examined the volatility reaction of these groups of currencies to market shock events (in the form of regulatory announcements), stating that DAG-based assets have become increasingly responsive to market shocks as they mature, evidencing signals of substantial market maturity in recent years.

4 Conclusion

This paper aims to provide, in the most detailed way possible, information on the current state of the cryptocurrency market and the patterns that can be extracted from its activity. For this purpose, a study has been made of the literature that focuses on the possible prediction of prices in this market. A total of 13 papers have been studied that focus their work in this area, in one way or another: some seek to find patterns in the activity of social networks and news; others seek the best machine-learning algorithm to predict its price; or simply analyze the behavior of the market in the face of external factors such as its regulation by different governments.

In this study, responding to RQ1, works have been found that make use of different methods of machine learning or statistical analysis to find the market closing prices. We have also found works that analyze different factors such as i) the sentiment of Tweets, ii) of news appearing in the media, iii) how investors react to announcements of regulatory measures affecting the market, iiii) and the technology behind each crypto.

Answering RQ2, it can be said that methods have been seen with a very good percentage of predicting the market closing price, although other works have also been studied that deny that one can ever predict the price of all crypto. Moreover, if external and social factors, such as regulatory measures and community sentiment in social networks, are not taken into account, it is not possible to accurately identify prediction patterns. In addition, there are pump and dump schemes that make prediction tasks difficult. Therefore, it can be concluded that, separately, the works are not effective, but as a whole, important conclusions can be drawn for users who intend to operate in this market.

This work is far from perfect, in the future we intend to extend the study to more databases and articles. We will also emphasize the possibilities offered by each of the algorithms used and what attributes they use to create their models.

Acknowledgements

The research of Yeray Mezquita is supported by the pre-doctoral fellowship from University of Salamanca and co-funded by Banco Santander. This research was also partially Supported by the project “Computación cuántica, virtualización de red, edge computing y registro distribuido para la inteligencia artificial del futuro”, Reference: CCTT3/20/SA/0001, financed by Institute for Business Competitiveness of Castilla y León, and the European Regional Development Fund (FEDER).

References

1. Beikverdi, A., Song, J.: Trend of centralization in bitcoin’s distributed network. In: 2015 IEEE/ACIS 16th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD). pp. 1–6. IEEE (2015)
2. Chokor, A., Alfieri, E.: Long and short-term impacts of regulation in the cryptocurrency market. The Quarterly Review of Economics and Finance (2021)
3. Chowdhury, R., Rahman, M.A., Rahman, M.S., Mahdy, M.: An approach to predict and forecast the price of constituents and index of cryptocurrency using machine learning. Physica A: Statistical Mechanics and its Applications 551, 124569 (2020)
4. David, S., Inacio Jr, C., Nunes, R., Machado, J.: Fractional and fractal processes applied to cryptocurrencies price series. Journal of Advanced Research (2021)
5. Diebold, F.X., Yilmaz, K.: Better to give than to receive: Predictive directional measurement of volatility spillovers. International Journal of forecasting 28(1), 57–66 (2012)
6. Frankovic, J., Liu, B., Suardi, S.: On spillover effects between cryptocurrency-linked stocks and the cryptocurrency market: Evidence from australia. Global Finance Journal p. 100642 (2021)
7. Gazafroudi, A.S., Mezquita, Y., Shafie-khah, M., Prieto, J., Corchado, J.M.: Islanded microgrid management based on blockchain communication. In: Blockchain-based Smart Grids, pp. 181–193. Elsevier (2020)
8. Hamrick, J., Rouhi, F., Mukherjee, A., Feder, A., Gandal, N., Moore, T., Vasek, M.: An examination of the cryptocurrency pump-and-dump ecosystem. Information Processing & Management 58(4), 102506 (2021)
9. Hu, Y., Hou, Y.G., Oxley, L., Corbet, S.: Does blockchain patent-development influence bitcoin risk? Journal of International Financial Markets, Institutions and Money 70, 101263 (2021)

10. Kraaijeveld, O., De Smedt, J.: The predictive power of public twitter sentiment for forecasting cryptocurrency prices. Journal of International Financial Markets, Institutions and Money 65, 101188 (2020)
11. Martinez, J.: Understanding proof of stake: The nothing at stake theory (June 2018), https://medium.com/coinmonks/understanding-proof-of-stake-the-nothing-at-stake-theory-1f0d71bc027, [Accessed; 09/10/2019]
12. Meegan, A., Corbet, S., Larkin, C., Lucey, B.: Does cryptocurrency pricing response to regulatory intervention depend on underlying blockchain architecture? Journal of International Financial Markets, Institutions and Money 70, 101280 (2021)
13. Mezquita, Y., Casado, R., Gonzalez-Briones, A., Prieto, J., Corchado, J.M., AETiC, A.: Blockchain technology in iot systems: review of the challenges. Annals of Emerging Technologies in Computing (AETiC), Print ISSN pp. 2516–0281 (2019)
14. Mezquita, Y., Casado-Vara, R., GonzÁlez Briones, A., Prieto, J., Corchado, J.M.: Blockchain-based architecture for the control of logistics activities: Pharmaceutical utilities case study. Logic Journal of the IGPL (2020)
15. Mezquita, Y., González-Briones, A., Casado-Vara, R., Chamoso, P., Prieto, J., Corchado, J.M.: Blockchain-based architecture: A mas proposal for efficient agri-food supply chains. In: International Symposium on Ambient Intelligence. pp. 89–96. Springer (2019)
16. Mezquita, Y., Valdeolmillos, D., González-Briones, A., Prieto, J., Corchado, J.M.: Legal aspects and emerging risks in the use of smart contracts based on blockchain. In: International Conference on Knowledge Management in Organizations. pp. 525–535. Springer (2019)
17. Nakamoto, S., et al.: Bitcoin: A peer-to-peer electronic cash system (2008)
18. Nghiem, H., Muric, G., Morstatter, F., Ferrara, E.: Detecting cryptocurrency pump-and-dump frauds using market and social signals. Expert Systems with Applications p. 115284 (2021)
19. Patel, M.M., Tanwar, S., Gupta, R., Kumar, N.: A deep learning-based cryptocurrency price prediction scheme for financial institutions. Journal of Information Security and Applications 55, 102583 (2020)
20. Perdana, A., Robb, A., Balachandran, V., Rohde, F.: Distributed ledger technology: Its evolutionary path and the road ahead. Information & Management 58(3), 103316 (2021)
21. Poongodi, M., Sharma, A., Vijayakumar, V., Bhardwaj, V., Sharma, A.P., Iqbal, R., Kumar, R.: Prediction of the price of ethereum blockchain cryptocurrency in an industrial finance system. Computers & Electrical Engineering 81, 106527 (2020)
22. Valdeolmillos, D., Mezquita, Y., González-Briones, A., Prieto, J., Corchado, J.M.: Blockchain technology: a review of the current challenges of cryptocurrency. In: International Congress on Blockchain and Applications. pp. 153–160. Springer (2019)
23. Wątorek, M., Drożdż, S., Kwapień, J., Minati, L., Oświęcimka, P., Stanuszek, M.: Multiscale characteristics of the emerging global cryptocurrency market. Physics Reports (2020)
24. Witherspoon, Z.: A hitchhiker's guide to consensus algorithms (November 2017), https://hackernoon.com/a-hitchhikers-guide-to-consensus-algorithms-d81aae3eb0e3, [Accessed; 09/10/2019]