European Economic Letters
ISSN 2323-5233
Vol 13, Issue 1 (2023)
http://eelet.org.uk

Predicting Cryptocurrency Price Using Machine Learning

Dr. Geetanjali Sahi[1] ,Tanjot Saluja[2] and Parish Nargotra[3]
Lal Bahadur Shastri Institute of Management, New Delhi
geetanjali@lbsim.ac.in, tanjot_saluja-r21@lbsim.ac.in and parish_nargotra-r21@lbsim.ac.in

Abstract. The cryptocurrency market which is based on blockchain technology permits execution of financial transactions over a peer-to-peer network. The primary feature of the cryptocurrency market is its extreme volatility and hence requires a method of predicting its price with precision. The current study aims at predicting the price of Litecoin, Ethereum, and Bitcoin, the top three cryptocurrencies. The data for the study comprises the monthly price of these three cryptocurrencies collected from popular crypto tracking websites for the period between January 2017 and December 2021. This data is analyzed using five Machine Learning models. The results suggested that accuracy outcomes are maximum using the Support Vector Machine technique with RMSE and MAPE values closest to 0.

Keywords: Cryptocurrency, Bitcoin, Ethereum, Litecoin, Blockchain, Machine Learning, Long Short Term Memory Model, Support Vector Machine, Artificial Neural Network.

.

1 Introduction

Digital currency, often referred to as electronic cash, electronic money, or digital money is like any currency that is handled, stored, and transacted over the internet. Among these, There are three different kinds of digital currencies: cryptocurrencies, virtual currencies, and digital currencies issued by central banks. Among these three Bitcoin has garnered unprecedented attention since its creation in 2009. It may be traded using a peer-to-peer online.

Cryptocurrency networks are independent of a single administration or a central bank. The system stands out by offering a platform for transactions and/or payments that is open-source, cryptographically secure, and confidentiality-preserving. Both the number of transactions and the number of accounts owned by people and corporations are continually rising. Alongside this, an entire sector of service providers has emerged.

Cryptocurrencies have evolved into a widely accepted investment vehicle due to their use of blockchain technology.

Money is a necessity for all financial transactions and is universally recognized as a symbol of worth. Any form of capital should necessarily have an intrinsic value and be a viable means of conducting transactions at the same time. This benefit is made possible by both fiat money and cryptocurrencies, even if they have some substantial differences. While fiat money such as the rupee or dollar is legitimately issued by a government agency, cryptocurrency has a digital existence, and its worth is determined by its original blockchain. While blockchain protocols, codes, and communities control cryptocurrencies, central banks control the issuance and management of fiat currencies. While cryptocurrency is dependent on dispersed and distributed networks that enable dealings that are "trustless", the distribution of fiat money requires middlemen. The most prevalent type of money in use today is fiat money, which is the legal cash issued by governments but has no fundamental value. The only experience many people have had with money has been with coins and paper money. However, the development of cryptocurrencies and blockchain technology over the past ten years represents a fundamental upgrade to global payment and value systems.

There are many different kinds of cryptocurrencies, including, but not limited to, Bitcoin, Litecoin, Ethereum, Nem, Ripple, Iota, and Stellar. Each crypto coin's foundation in cryptography gives it importance. To get started with cryptocurrency, one only needs a small investment and some fundamental study. The recent rise in the prices of a lot of cryptocurrencies over the world, and at the same time, has made people worried that the value of "fiat" currencies would fall. As a result, more people are now interested in investing in long-valued assets like gold and cryptocurrencies rather than saving money in banks. In terms of portability and storage, one may also store bitcoins on secure, user-friendly cryptocurrency wallets for free. According to Banfield, S. [5], cryptocurrencies have had a murky history, but utilizing them is simple once a person has some experience with them. Since Bitcoin has the most recognizable brand name among all cryptocurrencies, it is responsible for the popularity of the majority of cryptocurrencies. Every cryptocurrency project seeks to address a specific issue that a community has. Due to its ability to confirm transactions more quickly than Bitcoin, Litecoin is a favorite among those seeking out quicker payment settlements lastly Ethereum not only facilitates peer-to-peer transactions and offers a framework for developing distributed apps and smart contracts.

An investigation by Allied Market Research, a global market research firm [3], predictions state that the global crypto market will increase by more than three times by the year 2030. In today's world, when everything is governed by data science, businesses must stay current with emerging technology. Artificial Intelligence is also a contributor to such progress, but it is a double-edged sword. The current study aims at forecasting prices of cryptocurrencies using multiple machine-learning models and finding out the best prediction model.

2 Literature Review

The cryptocurrency market is one of the ones that is growing the fastest of all the financial markets in the world. Compared to more established methods of trading such as equities, bonds, foreign exchange, and commodities, the cryptocurrencies are more unstable and capricious. An alternative payment system created with the aid of encryption techniques is referred to as a digital currency, or cryptocurrency. Cryptocurrencies can function as a virtual accounting system and a medium of commerce by leveraging encryption technology. Early studies on cryptocurrencies debated whether they were just speculative assets or different types of money, with the majority of authors siding with the latter due to their high volatility, extraordinary short-term gains, and bubble-like price behavior.

Alessandretti and ElBahrawy [2] talked about how the market capitalization of cryptocurrencies had several months of super-exponential growth in 2017, which caused their popularity to soar. They examined daily information for 1,681 coins between November 2015 and April 2018 and suggested extended short-term memory recurrent neural networks and gradient boosting decision trees accomplished better results than other models.

Mudassir and Bennbaia[20], projected short- and medium-term swings in the price of bitcoin over the periods ranging from one day to ninety days, via Artificial Neural Networks (ANN), Support Vector Machines, Long Short Term Memory (LSTM), and Subsequent Artificial Neural Networks (SANN). Their models achieved accuracy scores of up to 65% for forecasts for the following day and between 62 and 64% for forecasts for the next seventy days. The error rate for daily forecast was low at 1.4% while it increased to 2.88 to 4.10% when forecasts covered from seven to ninety days.

To anticipate the price of bitcoin, Madan, Shah, and Zhao [16] tried to use machine-learning algorithms and their dataset comprised of daily data for five years pertaining to Bitcoin prices. Both random forests and generalized linear models were used in their analysis. Using 10-minute price interval data, these findings demonstrated an accuracy of 50–55% in forecasting the direction of a future price movement. Ho and Vatambeti [13] evaluated using a few regression techniques before deep learning models are implemented. They revealed that accuracy of linear regression was very high at 99.87 in comparison to other models. Cryptocurrency future price predictions are made based on their historical data availability. Derbentsev, Datsenko et. al. [9] created a model utilizing a machine learning approach for projecting cryptocurrency prices over the short term. The time series data and standard regression tree models were used to develop the Binary Auto Regressive Tree (BART) modified model. They used the BART model to forecast prices of Bitcoin, Ethereum, and Ripple for short term (between 5 and 30 days) and posit that BART model was more accurate than the ARIMA model. For the horizons of 14, 21, and 30, the RMSE for the BART algorithm was specifically in the range of 4%, 6%, and 8%, respectively.

A comparison of the machine learning algorithms' performance across six cryptocurrencies, including Bitcoin, Ethereum, Litecoin, Nem, Ripple, and Stellar, was provided by Hitam and Ismail [12]. For prediction, Deep Learning forecasts, Artificial Neural Networks (ANNs), and Support Vector Machines (SVM) were employed. SVM surpassed other classifiers with an accuracy of 95.5%, according to the results. The outcome is additionally investigated using mean absolute percentage error (MAPE) computation. Findings show that SVM has the lowest MAPE value.

Due to the distinctive qualities that each currency possesses such as brand name among all cryptocurrencies, quicker payment settlements, and the ability to create and build smart contracts and distributed applications., we also utilized these three cryptocurrencies. The effective integration of machine learning models into trading platforms for the bitcoin market is a key component of our research study. The recent success of applying machine learning to forecast stock market prices served as the impetus for this investigation. In the present study, multiple machine learning models are applied to understand which one is the most effective at predicting bitcoin prices.

3 Research Methodology

The data for the study comprises three different cryptocurrencies based on their rankings as per Changelly [7], which is a cryptocurrency exchange as on September 2022, the first is Bitcoin, which is still the coin that is most frequently mentioned when discussing digital money (Price of Bitcoin: $23,935, Market capitalization of Bitcoin: $457 billion). In this rise of the era of cryptocurrency, Bitcoin has been the forerunner. The second term that was known in the cryptocurrency industry is

European Economic Letters
ISSN 2323-5233
Vol 13, Issue 1 (2023)
http://eelet.org.uk

Ethereum [Price of Ethereum: $1,782, Market capitalization of Ethereum: $217 billion]. Ethereum is the one of the cryptocurrency platform's for creating a peer-to-peer
network. One of the reasons that has led to the popularity and increase in the usage of Ethereum is its smart contract feature. The final cryptocurrency in the study is Litecoin which according to many market pundits is expected to have very optimistic returns. Changelly [7] predicts that litecoin could go as high as $295, with a minimum of $250.15.
For the period spanning January 2017 and December 2021, data on the monthly prices of three cryptocurrencies namely Bitcoin, Ethereum, and Litecoin—was gathered from well-known price monitoring websites Investopedia [14] and CoinMarketCap [8]. The study focuses on utilizing machine learning to predict bitcoin prices.

A few predefined Python libraries such as pandas, NumPy, matplotlib, TensorFlow, etc. that aid in data visualization and were easier to comprehend the key features needed by the system have been integrated into this research study. A few predefined Python modules that aid in data visualization and in comprehending the key aspects needed by the system were implemented for the research paper. The research paper took into account the following characteristics; Open: The price at which a deal opened at that time, Close: The market's final price as of that time, High: It refers to the highest of the trade value recorded during that particular time, Low: It means the lowest of trade value recorded during that particular time, and Volume: The total volume recorded during that particular time.

The last step was to perform data normalization or standardization which is crucial for improving accuracy and results in order to ensure that the database constraints used to implement their dependencies are accurate and equal importance is given to each variable so that no single variable causes the performance of the model to be biassed because of their larger numbers. In the present study, since the features, i.e., open, close, volume, high and low in each of our cryptocurrency data lie in a wide range, the data is normalized to bring it to a standard scale. It is employed to reduce the database's multiple relationships' duplication. Proper data scaling is required when dealing with a Machine Learning model where we need backpropagation to be more reliable and even faster. In this study, in order to make the learning easier, lowest value is changed to zero, the greatest value to one, and all other values are altered to a decimal between 0 and 1. This is done for each feature. Five machine learning models, namely Linear Regression, Decision Tree, Long Short-Term Memory (LSTM), Support Vector Machine (SVM), and Random Forest, were used to examine the data.

4 Results

The normalized values of prices were used to create short-term projections for the chosen coins. The goal or target variable is the close price forecast value for each of the cryptocurrencies namely Bitcoin, Ethereum, and Litecoin. The same set of features was used to train the models for the purpose of analysis including Long Short-Term Memory, Support Vector Machine, Decision Tree, Random Forest, and Linear Regression. We used 30% of the testing data when testing these models. By calculating RMSE (Root Mean Square Error) and MAPE (Mean Absolute Percentage Error), performance of various models has been evaluated. Better convergence is shown by observed error values being smaller. Mean Absolute Error (MAE) and Mean Square Error (MSE) are other measures of performance. However, for measuring model performance RMSE is typically preferred over MAE and MSE because the former assists the developers in bringing down the possibility of large outliers being present as a part of predictions, and MAE is a measure which oversimplified for analyzing the performance of the overall model in this case. RMSE is more interpretable by both the model's creator and end users alike because the error is given in terms of the target as compared to MSE. MAPE is also a better measure than MSE and MAE because it does not accentuate large errors as compared to MSE and it is easy to understand for both developers and end users and can be compared across different models as compared to MAE.

4.1 i) Root Mean Square Error (RMSE)

Root Mean Square Error (RMSE) is the residuals' standard deviation (prediction errors). Residuals refer to the difference between the actual and forecasted value and the spread or dispersion of these points is measured by RMSE. RMSE is sensitive to outliers in the data. It does not penalize extreme errors while forecasting and is measured in the same units as response variables.

It is mathematically defined as-RMSE= $[\sum_{i=1}^{N} (z_{f_i} - z_{o_i})^2 / N]^{1/2}$ [23][24], Where: (zfi – Zoi)2 = differences, squared.

ii) Mean Absolute Percentage Error (MAPE)

MAPE is a measure which gives or calculates the mean or average value of the absolute percentage error of forecasts[6][11]. It is independent of measurement scale, in contrast to MSE and MAE, but is impacted by data transformation. In addition to not penalizing high forecasting errors as MAE does, MAPE also doesn't indicate the general error's direction.

It is given by- MAPE= $\frac{1}{n}\sum_{t=1}^{n}\left|\frac{e_t}{y_t}\right| X\,100$

	LSTM		SVM		Linear Regression		Decision Tree		Random Forest	
	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE
BITCOIN	0.041	0.027	0.029	0.021	0.040	0.026	0.201	0.151	0.170	0.125
ETHEREUM	0.052	0.034	0.032	0.023	0.049	0.031	0.229	0.188	0.183	0.144
LITECOIN	0.051	0.035	0.027	0.016	0.050	0.034	0.185	0.120	0.149	0.100

The table shows that each crypto currency data performs best when processed using the Support Vector Machine algorithm (SVM), as this method has the MAPE and RMSE values that are the closest to zero i.e. for Bitcoin the RMSE and MAPE values came out to be 0.029 and 0.021 respectively, similarly for Ethereum it was 0.032 and 0.023 and lastly for Litecoin the RMSE and MAPE values were 0.027 and 0.016 and the decision tree algorithm produces the cryptocurrency price predictions with the lowest accuracy for each dataset such as for Ethereum it showed an RMSE value of 0.229 which reflects the least accuracy in comparison to other machine learning models. One of the most effective methods for resolving binary classification issues at the moment is the support vector machine (SVM) [25], according to research [24]. SVM shows brilliant results in generalization so it can produce high accuracy. It produces a binary classification model with a minimum of dataset-related assumptions. Support Vector Machine (SVM) algorithm is a popular categorization model based on the idea that there could be a hyperplane in a multidimensional space that could categorize the data. It can be adjusted to work well with unbalanced datasets, has only a few parameters that need to be tuned, is only partially immune to outliers, and is very effective when groups are fully or nearly fully separable. It can also work very well when there are more independent variables than observations. [4]

The table shows that each dataset performs best when processed using the Support Vector Machine algorithm (SVM), as this method has the MAPE and RMSE values that are the closest to zero and the decision tree algorithm produces the cryptocurrency price predictions with the lowest accuracy for each dataset.

5 Conclusion

Bitcoin, Ethereum, and Litecoin are the three cryptocurrencies that are utilized the most frequently. Data analysis is needed to forecast crypto prices because each of these cryptocurrencies has unique values that are very difficult to predict. This allows investors to trade with these currencies and make the maximum profit feasible. The transaction history data of each cryptocurrency was examined using five machine-learning approaches, and the results of this research were successful in providing information in the form of price forecasts with the highest accuracy. The accuracy outcomes are at their maximum when the data is analyzed using the SVM technique, with RMSE and MAPE values that are the closest to 0. However, the results are reflective of the current data and algorithms, more datasets, new validation methods, or algorithms Future studies may use methods different than those employed in this work to improve the accuracy of price forecasts for cryptocurrencies. Using additional variables, such as average of prices, various basic variables, including the price rate-of-change and the index of relative strength, we believe that prediction accuracy can be increased. The predictive power of the aforementioned characteristics as well as any novel elements should be further investigated in future studies.

References

1. Agarwal, A., & Kumar, K. Cryptocurrency Price Predictor (2021). http://103.47.12.35/bitstream/handle/1/9794/bt3309.pdf?sequence=1&isAllowed=y
2. Alessandretti, L., ElBahrawy, A., Aiello, L. M., &Baronchelli, A. Anticipating cryptocurrency prices using machine learning.Complexity,(2018). https://www.hindawi.com/journals/complexity/2018/8983590/

3. Allied Market Rresearch, (2018) https://www.prnewswire.co.uk/news-releases/cryptocurrency-market-to-reach-4-94 billion-globally-by-2030-at-12-8-cagr-allied-market-research-836503660.html
4. Auti, A., Patil, D., Zagade, O., Bhosale, P., &Ahire, P. BITCOIN PRICE PREDICTION USING SVM (2022).
5. Banfield, S., Cryptocurrencies : Decrypting the past, regulating the future (2018) https://www.withersworldwide.com/en-gb/insight/cryptocurrencies-decrypting-the-past-regulating-the-future
6. Cao, L., & Tay, F. Support vector machine with adaptive parameters in financial time series forecasting. IEEE Transactions on Neural Networks, 14(6), 1506–1518. (2003).
7. Changelly, 2022, https://changelly.com/
8. CoinMarketCap URL (for accessing the data) - https://coinmarketcap.com/
9. Derbentsev, V., Datsenko, N., Stepanenko, O., &Bezkorovainyi, V. Forecasting cryptocurrency prices time series using machine learning approach. In SHS Web of Conferences (Vol. 65, p. 02001). EDP Sciences (2019).
10. Fang, F., Chung, W., Ventre, C., Basios, M., Kanthan, L., Li, L., & Wu, F.. Ascertaining price formation in cryptocurrency markets with machine learning. The European Journal of Finance, 1-23(2021) https://www.tandfonline.com/doi/abs/10.1080/1351847X.2021.1908390
11. Hamzaçebi, C. Improving artificial neural networks' performance in seasonal time series forecasting. Information Sciences, 178(23), 4550-4559(2008).
12. Hitam, N. A., & Ismail, A. R. Comparative performance of machine learning algorithms for cryptocurrency forecasting. Ind.J.Electr.Eng. Comput. Sci, 11(3), 1121-1128.(2018). https://ijeecs.iaescore.com/index.php/IJEECS/article/view/13469
13. Ho, A., & Ravichandran, S. KBitcoin Price Prediction Using Machine Learning and Artificial Neural Network Model. Indian Journal of Science and Technology, 14(27), 2300-2308 (2021). https://sciresol.s3.us-east-2.amazonaws.com/IJST/Articles/2021/Issue-27/IJST-2021-878.pdf.
14. Investopedia URL (for accessing the data) - https://www.investopedia.com/
15. Jagannath, N., Barbulescu, T., Sallam, K. M., Elgendi, I., Okon, A. A., McGrath, B., ... &Munasinghe, K. A self-adaptive deep learning-based algorithm for predictive analysis of bitcoin price. IEEE Access, 9, 34054-34066 (2021). https://ieeexplore.ieee.org/abstract/document/9359745/
16. Madan, I., Saluja, S., & Zhao, A. Automated bitcoin trading via machine learning algorithms. URL: http://cs229. stanford. edu/proj2014/Isaac% 20Madan, 20 . (2015). http://cs229.stanford.edu/proj2014/Isaac%20Madan,%20Shaurya%20Saluja,%20Aojia%20Zhao,Automated%20Bitcoin%20Trading%20via%20Machine%20Learning%20Algorithms.pdf
17. Marne, S., Correia, D., Churi, S., & Gomes, J. Predicting Price of Cryptocurrency-A deep learning approach. NTASU-9 (3(2020).). https://www.academia.edu/download/66254763/predicting_price_of_cryptocurrency_a_IJERTCONV9IS03083.pdf
18. McNally, S., Roche, J., & Caton, S. Predicting the price of bitcoin using machine learning. In 2018 26th euromicro international conference on parallel, distributed and network-based processing (PDP) (pp. 339-343). IEEE (2018). https://ieeexplore.ieee.org/abstract/document/8374483/
19. Mittal, M., & Geetha, G. Predicting Bitcoin Price using Machine Learning. In 2022 International Conference on Computer Communication and Informatics (ICCCI) (pp. 1-7). IEEE. (2022). https://ieeexplore.ieee.org/abstract/document/9740772/
20. Mudassir, M., Bennbaia, S., Unal, D., &Hammoudeh, M. Time-series forecasting of Bitcoin prices using high-dimensional features: a machine learning approach. Neural computing and applications, 1-15(2020).. https://link.springer.com/article/10.1007/s00521-020-05129-6
21. Sebastião, H., &Godinho, P. Forecasting and trading cryptocurrencies with machine learning under changing market conditions. Financial Innovation, 7(1), 1-30(2021).. https://jfin-swufe.springeropen.com/articles/10.1186/s40854-020-00217-x
22. Thomas, C., Watson, Z., Kim, M., Baidya, A., Lamsal, M., Chowdhury, M. H., ... & Poudel, K. N. Cryptocurrency Analysis using Machine Learning and Deep Learning. In 2021 IEEE Signal Processing in Medicine and Biology Symposium (SPMB) (pp. 01-03). IEEE (2021).. https://ieeexplore.ieee.org/abstract/document/9672268/

23. Velankar, S., Valecha, S., & Maji, S. Bitcoin price prediction using machine learning. In 2018 20th International Conference on Advanced Communication Technology (ICACT) (pp. 144-147). IEEE (2018). https://ieeexplore.ieee.org/abstract/document/8323676/
24. Widodo, A., Yang, B.S.: Support vector machine in machine condition monitoring and fault diagnosis. Mechanical Systems and Signal Processing 21 (2007).
25. .Xu W, Cooper A. Digital spillover: measuring the true impact of the digital economy. Huawei and Oxford Economics(2017).
https://www.huawei.com/minisite/gci/en/digital-spillover/files/ gci-digital-spillover.pdf
26. Zhang, G. P. Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing, 50, 159-175.(2003).