International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 04 | April - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

BIT COIN PRICE PREDICTION USING RF ALGORITHM WITH SENTIMENT FACTORS.

Keerthika. A[1] , Ajithkumar. R[2] , Tharani. K[3] , Yamuna .M[4], Indhumathi . S[5]UG

Scholars[1234], Assistant Professor[5],

Department of Computer Science and Engineering,Tirupur,

Tamil Nadu, India.

ABSTRACT

Sentiment analysis was included into a predictive model that was created for this project in order to forecast bitcoin closing values. The dataset contains daily Bitcoin data for January 2024, including sentiment scores computed for each day as well as opening, high, low, and closing values. The study assesses the predictive ability of the Random Forest regression model by using metrics like Mean Squared Error (MSE), Mean Absolute Error (MAE), and R-squared (R^2) score. The negative R^2 score indicates that, although the model may produce predictions, its overall fit to the data is still not ideal, implying that the model is unable to adequately represent the correlation between sentiment ratings and Bitcoin prices. Potential areas for development to increase the predicted accuracy of the model and gain a deeper understanding of how sentiment analysis affects Bitcoin prices include feature engineering, hyper parameter tuning, and model selection.

Keywords: Cryptocurrency, Price Prediction, Machine learning, Random Forest

1. INTRODUCTION

Digital money markets have seen touchy development and expanding intricacy as of late. With great many digital forms of money exchanged on different trades, foreseeing cryptographic money costs has turned into a difficult and rewarding undertaking. Customary monetary models frequently miss the mark in catching the dynamic and reliant nature of these computerized resources. To resolve this issue, profound learning procedures have arisen as incredible assets for digital money cost expectation. Profound learning-based digital money cost expectation plans.

1.1 CRYPTOCURRENCY

The universe of money and innovation has been changed by the development of cryptographic forms of money, a computerized upheaval that challenges customary monetary frameworks and presents new standards for exchanges, speculations, and decentralized frameworks. Cryptographic forms of money, frequently alluded to as "crypto," are advanced or virtual monetary standards that utilize cryptographic strategies to get exchanges and control the production of new units. Bitcoin, the first and most notable cryptographic money, was presented in 2009 by a mysterious substance known as Satoshi Takemoto, making way for a decentralized monetary upheaval. Digital forms of money are based on block chain innovation, a dispersed record that records all exchanges across an organization of PCs.

1.2 PRICE PREDICTION

Cost expectation has for some time been a focal concentration in monetary business sectors, filling in as a central device for financial backers, dealers, and monetary experts. Exact cost conjectures empower informed independent direction and hazard the board, enabling partners to improve their venture techniques. Lately, headways in information examination, AI, and man-made brainpower have changed the field of cost expectation, offering new systems and extraordinary degrees of precision. This development has reached out past conventional monetary business sectors to envelop a large number of resources, including digital currencies, stocks, wares, from there, the sky is the limit.

2. LITERATURE REVIEW

Shane Courbetet.al.[1] Has proposed in this paper, we analyze the co operations between cryptographic money value unpredictability and liquidity during the episode of the Coronavirus pandemic. Proof recommends that these creating computerized items play had another impact as a possible place of refuge during times of significant monetary market alarm. Results propose that cryptographic money market liquidity expanded fundamentally after the WHO recognizable proof of an overall pandemic. Huge and significant associations between digital money cost and liquidity impacts are distinguished. These outcomes add further help to the contention that significant progressions of speculation entered digital currency markets looking for a venture place of refuge during this uncommon dark swan occasion. The quickly creating Coronavirus pandemic produced a lot of disarray with respect to the seriousness and financial implications.

Thomas E. Kocheret.al. [2] has proposed in this paper We present a model for dynamic exchanging in light of support AI and apply this to five significant digital currencies available for use. Corresponding to a purchase and-hold approach, we exhibit how this model yields upgraded risk-changed returns and decreases disadvantage risk. These discoveries hold while representing genuine exchange costs. We reason that certifiable portfolio the executive's utilization of the model is feasible; yet execution can shift in view of the way things are adjusted in test tests. Is dynamic exchange practical in digital currency showcases and might it at any point yield better execution relative than a purchase and-hold approach? Utilizing an immediate support (DR) that's what model we exhibit, indeed, dynamic exchanging is both feasible and beneficial in such business sectors and can yield better gamble changed execution relative than a latent purchase and-hold approach.

Thomas E. Kocheret.al. [3] has proposed in this paper We present a model for dynamic exchanging in light of support AI and apply this to five significant digital currencies available for use. Corresponding to a purchase and-hold approach, we exhibit how this model yields upgraded risk-changed returns and decreases disadvantage risk. These discoveries hold while representing genuine exchange costs. We reason that certifiable portfolio the executive's utilization of the model is feasible, yet execution can shift in view of the way things are adjusted in test tests. Is dynamic exchange practical in digital currency showcases and might it at any point yield better execution relative than a purchase and-hold approach? Utilizing an immediate support (DR) that's what model we exhibit, indeed, dynamic exchanging is both feasible and beneficial in such business sectors and can yield better gamble changed execution relative than a latent purchase and-hold approach.

Purnima Harridanset.al. [4] Has proposed in this framework we propose, a learning framework to foresee the likeness of a given programming code to a bunch of codes that are allowed to run on a computational asset, like a supercomputer or a cloud server. This code portrayal permits us to identify harmful codes. Our framework depends on an underlying examination of the control-stream chart of the product codes and two different diagram similitude measures: Diagram Alter Distance (GED) and a particular qualities-based measurement. SiCaGCN joins components of Diagram Convolutional Brain Organizations (GCN), Case organizations, consideration instrument, and brain tensor organizations. Our exploratory outcomes incorporate an investigation of the compromises between the two likeness measurements and two varieties of our learning organizations, with and without the utilization of containers. Our fundamental discoveries are that the utilization of cases decreases mean square mistake altogether for both closeness measurements. Utilization of containers lessens the runtime to work out the GED while expanding the runtime of particular qualities calculation N the time of exascale figuring code portrayal is critical for super-registering focuses and cloud merchants.

Savva Shanaeet.al. [5] Has proposed in this framework, this paper utilizes a remarkable dataset of 120 administrative occasions from five classes to test the importance of the administrative system for digital money esteem. Time-series all-inclusive gauges and board gauges for 300 individual coins and tokens show measurably and monetarily critical effect of hostile to illegal tax avoidance and issuance guideline. Tighter guideline and more dynamic job of government decline digital money costs, confirming that possibly lower gambles and more extensive reception generally credited to the foundation of the administrative system don't make up for separate proficiency and customer utility misfortunes. The market is by and large effective in reflecting administrative data in cryptographic money costs. With the appearance of block chain innovation overall and its most well-known functional applications, digital forms of money, specifically, the issue of its guideline has been turning out to be progressively important.

3. EXISTING SYSTEM

Block chain technology has gained widespread popularity due to its applications across various domains, offering advantages such as decentralization, immutability, data integrity, and anonymity. Its most notable application is in crypto currencies, which have seen a significant surge in popularity and market capitalization in recent years, attracting investments from individual investors, institutional players, and corporate entities. However, the crypto market is characterized by greater volatility and unpredictability compared to traditional commodity markets, influenced by technical, emotional, and legal factors. Despite extensive research into crypto currency price forecasting, many existing approaches are impractical for real-time applications. To address this, we propose a novel deep-learning-based hybrid model incorporating Gated Recurrent Units (GRU) and Long Short-Term Memory (LSTM) to predict Lite coin and Zcash prices, considering the inter-dependency with their parent coin.

4. PROPOSED SYSTEM

Our suggested solution combines machine learning and sentiment analysis to create a strong forecasting model for changes in bitcoin prices. The first step of the system's operation is data pre-processing, which gathers and cleans past bitcoin data, including opening and closing prices. To determine the sentiment scores linked to each cryptocurrency entry, sentiment analysis is also carried out on pertinent textual data sources, such as news articles or social media posts. The next step is to use feature engineering techniques to extract useful features from the data. This comprises the emotion scores produced previously, rolling window statistics, and lag features. The sentiment-driven dynamics and temporal patterns present in the cryptocurrency market are captured by these manufactured features. Next, we divide the dataset into subsets for training and testing so that we may assess the effectiveness of our machine learning model. Here, we apply a Random Forest model, which is renowned for its capacity to manage intricate interactions and reduce overfitting. The algorithm learns to forecast cryptocurrency prices based on sentiment scores and manufactured features during model training. The model's performance is evaluated using evaluation measures such Mean Squared Error (MSE), Mean Absolute Error (MAE), and R-squared (R2). These measures shed light on how well the model can predict bitcoin prices and how well it can generalize. The system then displays the outcomes of the model's predictions in relation to the testing dataset's actual bitcoin prices.

A. Data Preparation:

In order to comprehend the structure and contents of the dataset, this phase entails loading and inspecting it. Date, Cryptocurrency, Opening Price, High Price, Low Price, Closing Price, and Sentiment Score are among the columns that are included. It handles missing value management, data type conversion, and outlier detection, among other data cleaning operations. When taken as a whole, these characteristics provide a thorough grasp of Bitcoin's price movements, paving the way for additional research and forecasting.

B. Sentiment score calculation:

The dataset acquires a deeper level of understanding into the activity of the Bitcoin market by integrating sentiment analysis. Sentiment scores are obtained by textual analysis and provide important background information on how the public felt about Bitcoin at each point in time that was recorded. These ratings function as quantitative gauges of market mood and may have an impact on changes in bitcoin prices.

C. Training and testing datasets:

The creation of machine learning models requires the separation of the dataset into subsets for training and testing. Using past data to identify underlying patterns, the training set makes model parameter estimation and learning easier. On the other hand, the testing set functions as a separate validation mechanism that makes it possible to assess the performance of the model on unobserved data. Sentiment scores are added to the closing prices in both subsets, which enhances the input features of the model and improves its capacity to grasp complex interactions.

D. Model evaluation metrics:

Model performance is quantitatively assessed using evaluation metrics, which give important information about the model's resilience and accuracy of prediction. The average squared difference between the expected and actual values is quantified by Mean Squared Error (MSE), which highlights the accuracy of the model. The average absolute variation between expected and actual values is measured by the Mean Absolute Error (MAE), which offers a simple way to assess how accurate a forecast is. R-squared (R2) provides a comprehensive evaluation of the model's explanatory power by quantifying the percentage of the target variable's variance that the model explains.

E. Random Forest Results

You use the trained Random Forest model to make predictions on the testing dataset once the model has been trained and assessed. An essential stage in predicting the price of cryptocurrencies is the Random Forest model. Random Forest makes precise predictions by utilizing the combined knowledge of several decision trees through the application of ensemble learning techniques. Transparency regarding the effectiveness of the model is provided by the display of actual and anticipated closing prices for the testing subgroup.

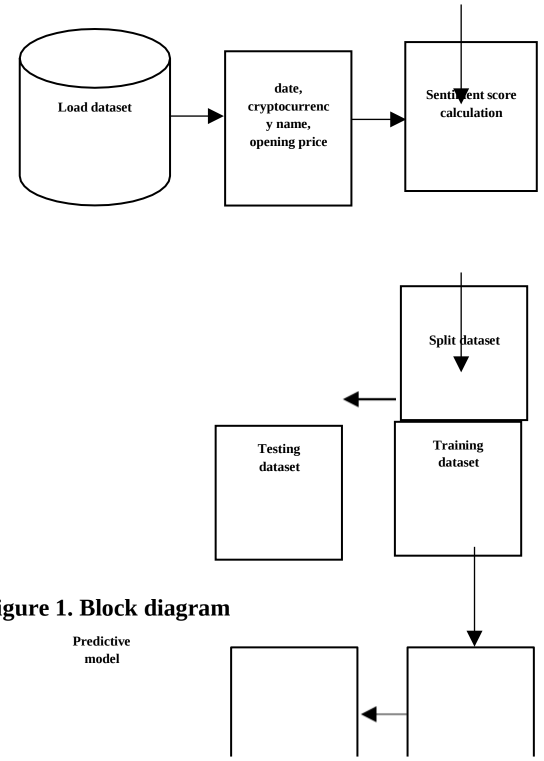

Figure 1. Block diagram

5. RESULT ANALYSIS

The performance and efficacy of the bitcoin price prediction model are revealed through the examination of the data from the provided code. Performance measures such as an MSE of 1.536973e+06, an MAE of 1102.357433, and an R2 of -0.157484 were attained by the Random Forest model. The model's predictions show variety across the testing dataset, even with the negative R2 value showing a performance below that of a simple horizontal line. While some forecasts nearly match the real cryptocurrency prices, others show notable discrepancies, indicating that the model's generalization and accuracy are limited. Additionally, as the model's performance does not regularly correspond with sentiment trends, the effect of sentiment ratings on prediction accuracy appears to be unclear. In order to better represent market dynamics, this suggests possible areas for feature engineering and model optimization improvement. Notwithstanding these drawbacks, the model offers a starting point for additional investigation and improvement, emphasizing the necessity of ongoing study and testing to improve forecast accuracy in the bitcoin market.

Metric	Value
Mean Squared Error	1.536973e+06
Mean Absolute Error	1102.357433
R-squared	-0.157484

Table 1. Comparison table

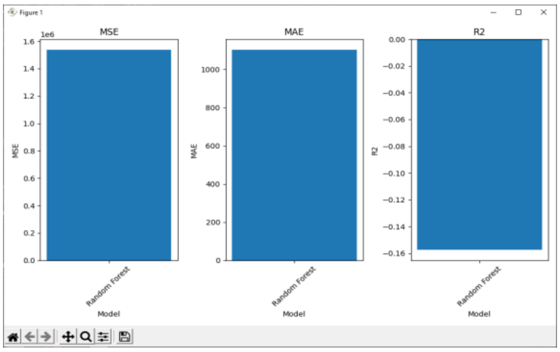

Figure 2. Comparison graph

6. CONCLUSION

To sum up, our suggested solution combines machine learning and sentiment analysis methods to create a strong prediction model for changes in bitcoin prices. By combining sentiment ratings with past bitcoin data, we are able to capture intricate market dynamics and improve prediction accuracy. The model's capacity to extract significant features and adjust to shifting market conditions offers stakeholders insightful information that helps them make well-informed decisions when investing in cryptocurrencies. We guarantee the dependability and efficacy of our prediction model by thorough assessment and testing, providing a data-driven strategy for navigating the unstable cryptocurrency market.

7. FUTURE WORK

Future research on improving the suggested approach for predicting Bitcoin prices could concentrate on a number of areas to increase its resilience and accuracy. First off, a more thorough picture of market dynamics may be obtained by including data from sources other than historical price data, such as macroeconomic indicators, social media trends, and market sentiment analysis. Second, investigating more complex machine learning models or ensemble approaches—like deep learning models or hybrid approaches—beyond Random Forest may present chances to identify more complex patterns in the data.

8. REFERNCES

[1] During the COVID-19 epidemic, cryptocurrency liquidity and volatility interrelationships were examined by S. Corbet, Y. Hou, Y. Hu, C. Larkin, B. Lucey, and L. Oxley. Early access to Finance Res. Lett., May 2021, Art. no. 102137.

[2] "Cryptocurrency trading using machine learning," by T. E. Koker and D. Koutmos J. Risk Financial Manage., Aug. 2020; vol. 13, no. 8, p. 178.

[3] BATS: A blockchain and AI-powered drone-assisted telesurgery system towards 6G, R. Gupta, A. Shukla, and S. Tanwar. Early access to IEEE Trans. Netw. Sci. Eng., Dec. 8, 2020

[4] "Code characterization with graph convolutions and capsule networks," by P. Haridas, G. Chennupati, N. Santhi, P. Romero, and S. Eidenbenz IEEE Access, 8, (2020), 136307–136315.

[5] Shanaev, Sharma, Ghimire, and Shuraeva, S. (2015) "Taming the blockchain beast? The cryptocurrency market's regulatory ramifications are discussed in Res. Int. Bus. Finance, vol. 51, Jan. 2020, Art. no. 101080.

[6] Gomes, Gomes, and Correia, G., "CryingJackpot: Network flows and performance counters against crypto jacking," in Proceedings of the IEEE 19th International Symposium on Network and Computer Applications (NCA), Nov. 2020, pp. 1–10.

[7] "Bitcoin price prediction using ensembles of neural networks," by E. Sin and L. Wang, was published in the July 2017 issue of Proc. of the 13th International Conference on Natural Computing, Fuzzy System Knowledge Discovery (ICNC-FSKD), pp. 666–671.

[8] Toward describing blockchain-based cryptocurrencies for extremely accurate predictions: M. Saad, J. Choi, D. Nyang, J. Kim, and A. Mohaisen, IEEE Syst. J., vol. 14, no. 1, pp. 321–332, Mar. 2020.

[9] "Stochastic neural networks for cryptocurrency price prediction," by P. Jay, V. Kalariya, P. Parmar, S. Tanwar, N. Kumar, and M. Alazab IEEE Access, 8, (2020), 82804–82818,

[10] In Proc. 12th Int. Conf. Contemp. Comput. (IC3), Aug. 2019, pp. 1–5, A. Aggarwal, I. Gupta, N. Garg, and A. Goel, "A novel approach for abo blood group prediction using fingerprint through optimized convolutional neural network"

[11] "An optimal least square support vector machine-based earnings prediction of blockchain financial products," by M. Sivaram, E. L. Lydia, I. V. Pustokhina, D. A. Pustokhin, M. Elhoseny, G. P. Joshi, and K. Shankar IEEE Access, 8, (2020), 120321–120330.

[12] A data selection methodology to train a linear regression model to predict the price of bitcoin was presented by M. Ali and S. Shatabda in Proc. 2nd Int. Conf. Adv. Inf. Commun. Technol. (ICAICT), Nov. 2020, pp. 330–335.

[13] Cryptocurrency price prediction using neural networks and deep learning, S. Biswas, M. Pawar, S. Badole, N. Galande, and S. Rathod, Proc. 7th Int. Conf. Adv. Compute. Commun. Syst. (ICACCS), Mar. 2021, pp. 408–413

[14] "A dropout weight-constrained recurrent neural network model for forecasting the price of major cryptocurrencies and CCi30 index," by I. E. Livieris, S. Stavro Yiannis, E. Pintelas, T. Kotsilieris, and P. Pintelas Developing System, early access, January 2021, pp. 1–16.

[15] "LSTM based sentiment analysis for cryptocurrency prediction," Huang, W. Zhang, Y. Huang, X. Tang, M. Zhang, J. Surbiryala, V. Iosifidis, Z. Liu, and J. Zhang, 2021, arXiv:2103.14804. [Online]. [http://arxiv.org/abs/2103.14804] is accessible. Energy Econ., vol. 87, Mar. 2020, Art. no. 104703;

[16] D. I. Okorie and B. Lin, "Crude oil price and cryptocurrencies: Evidence of volatility connectedness and hedging strategy."

[17] "The Nexus between black and digital gold: Evidence from US markets," by T. L. D. Huynh, R. Ahmed, M. A. Nasir, M. Shahbaz, and N. Q. A. Huynh Ann. Oper. Res., July 2021, vol. 22, pp. 1–26.

[18] "Asymmetric correlation and hedging effectiveness of gold & cryptocurrencies: From pre-industrial to the 4th industrial revolution," by N. Thampanya, M. A. Nasir, and T. L. D. Huynh Science and Technology Forecasting, vol. 159, Oct. 2020, art. no. 120195

[19] Foglia, M., and Dai, P.-F., "Ubiquitous uncertainties: Spillovers across cryptocurrency uncertainty indices and economic policy uncertainty," Early access, J. Asian Bus. Econ. Stud., pp. 1–15, July 2021.

[20] "Do FEARS drive bitcoin?" by T. Burggraf, T. L. D. Huynh, M. Rudolf, and M. Wang Rev. Behav. Finance, July 2021, vol. 13, no. 3, pp. 229–258.