IARJSET

ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 11, Issue 6, June 2024

DOI: 10.17148/IARJSET.2024.11639

BIT FORECAST

DR. HS Annapurna[1], Mrs ROOPA T[2], Unnatha D R [3], Yashwitha P R[4]

Professor and HOD, ISE Department, Sri Siddhartha Institute of Technology, Tumkur, INDIA[1]

Assistant Professor, ISE Department, Sri Siddhartha Institute of Technology, Tumkur, INDIA[2]

Student of III year, ISE Department, Sri Siddhartha Institute of Technology, Tumkur, INDIA[3]

Student of III year, ISE Department, Sri Siddhartha Institute of Technology, Tumkur, INDIA[4]

Abstract: Cryptocurrencies are a digital way of money in which all transactions are held electronically. It is a soft currency which doesn't exist in the form of hard notes physically. Here, we are emphasizing the difference of fiat currency which is decentralized that without any third-party intervention all virtual currency users can get the services. However, getting services of these cryptocurrencies impacts on international relations and trade, due to its high price volatility. There are several virtual currencies such as bitcoin, ripple, ethereum, ethereum classic, lite coin, etc. In our study, we especially focused on a popular cryptocurrency, i.e., bitcoin. From many types of virtual currencies, bitcoin has a great acceptance by different bodies such as investors, researchers, traders, and policy-makers. To the best of our knowledge, our target is to implement the efficient deep learning-based prediction models specifically long short-term memory (LSTM) and gated recurrent unit (GRU) to handle the price volatility of bitcoin and to obtain high accuracy. Our study involves comparing these two time series deep learning techniques and proved the efficacy in forecasting the price of bitcoin.

Keywords: Bitcoin , LSTM , Ethereum,GRU .

I. INTRODUCTION

Bitcoin is a decentralized digital currency that uses cryptography for security and is not controlled by any government or financial institution. It was created in 2008 by an individual or group of individuals using the pseudonym Satoshi Nakamoto (2008) with a paper titled "Bitcoin: A Peer-to-Peer (P2P) Electronic Cash System". Transactions with bitcoin are recorded on a public ledger called the blockchain, which allows anyone to view the history of a specific Bitcoin. The decentralized nature of Bitcoin allows it to operate independently of central banks and can be transferred instantly across the globe. It has gained popularity as a means of exchange and a store of value (Baur and Dimpfl 2021).

In the past 10 years, after experiencing several ups and downs, it broke through USD 68,000 per coin in November 2021, and the total current price once exceeded USD 1.2 trillion. However , as a commodity, Bitcoin has the problem of high volatility. During the seven years from April 2015 to April 2022, the standard deviation of Bitcoin's daily return rate was 3.85%, which was 2.68 times the standard deviation of gold's return rate during the same period and 3.36 times that of the S&P500. Due to the large price fluctuations, the function of Bitcoin as a store of value as a commodity and as a transaction payment function as a currency has been questioned.

While enjoying the advantages of Bitcoin's security and decentralization, how to grasp the trend of Bitcoin to minimize the risk of Bitcoin floating has become a difficult problem. Many researchers try to grasp the trend of Bitcoin through the correlation between the price of Bitcoin and the price of other commodities. But whether it is **gold (Baur and Hoang 2021; Kim et al. 2020b; Blake 2019), which is often used for comparison, stock market index (Erdas and Caglar 2018), or crude oil** price (**Selmi et al. 2018**), past studies have shown that the correlation between Bitcoin and them is weak.

Prediction of Bitcoin by AI is mainly divided into two categories. The first category is the classification research of predicting the rise or fall of Bitcoin in the future. The error standard is DA and F1. The other category is regression research on predicting Bitcoin prices, while the corresponding errors are RMSE and MAPE. Due to the sharp fluctuations in the price of Bitcoin, only grasping the rise or fall of the price of Bitcoin in the future cannot help investors avoid risks. In contrast, getting the specific bitcoin price as a reference price is more useful.

Based on the necessity of avoiding the price risk of Bitcoin as the background, this research chooses the random forest regression algorithm of machine learning and the LSTM model of neural network algorithm to predict the price of Bitcoin.

IARJSET

ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⪤ Peer-reviewed & Refereed journal ⪤ Vol. 11, Issue 6, June 2024

DOI: 10.17148/IARJSET.2024.11639

I mainly focus on the performance of random forest regression in Bitcoin price prediction when using the prediction results of LSTM as a comparison. Random forest regression is a regression form of random forest.

Different from the black box technology of neural networks, random forest regression as machine learning can deliver the importance of each explanatory variable in predicting Bitcoin through the results of its weak-learners

The prediction effect of random forest in predicting stock price direction . However, unlike random forest classifier, whose research goal is to classify ups or downs, there are not many papers that use random forest regression to study the cryptocurrency market in the existing literature. In the literature using random forest regression, the explanatory variables used by **Parvez** (**2022**) focus on the highly correlated OHLC (Open, High, Low, Close) and transaction volume of Bitcoin itself as explanatory variables. On this basis, I think it is of great research value to add explanatory variables in other fields.

II. RELATED WORK

Aggarwal et al. (**2019**) studied whether gold price can predict Bitcoin price through three deep learning algorithms of CNN, LSTM, and GRU. The conclusion is that the predicted price of the model which only uses gold price deviates from the true Bitcoin price, and the prediction accuracy of the LSTM model is the best of three. **Liu et al.** (**2021**) expanded the range of explanatory variables, based on the cryptocurrency market and macro market index (stock market index, crude oil price, exchange rate, etc.) and search index, a total of 40 explanatory variables for Bitcoin price prediction.

SDAE algorism shows better prediction performance than BPNN, PCA-SVR, and SVR.
Regarding the prediction research of Bitcoin price, the methods are divided into time series and machine learning. Multiple studies have concluded that the prediction accuracy of ARIMA is not as good as that of machine learning (**McNally et al. 2018**; **Shin et al. 2021**; **Chen et al. 2020a**; **Akyildirim et al. 2021**)

STM, as a controlled study of random forest regression in this study, has been studied as a target model many times in the past literature (**Shin et al. 2021**; **Jagannath et al. 2021**; **Rizwan et al. 2019**). **Phaladisailoed and Numnonda** (**2018**) used four deep learning algorithms (Theil–Sen regression, Huber regression, LSTM, and GRU) to predict the price of Bitcoin. The 52.78% accuracy of the LSTM algorithm is the highest. Based on the same explanatory variables, **Tandon et al.** (**2019**) found that adding 10-fold cross-validation to the LSTM training process can increase the accuracy of LSTM by 14.7%. However, the selection of explanatory variables in Phaladisailoed's and Tandon's studies is limited to OHLC, volume from top exchange and market cap. In the research done by **Aggarwal et al.** (**2019**), in addition to the price of Bitcoin itself, gold price was added to explanatory variables. The experimental results show that the RMSE of the LSTM algorithm is 47.91, which is better than CNN and GRU P.

In **McNally et al.**'s (**2018**), **García-Medina and Duc Huynh**'s (**2021**), and **Chen et al.**'s (**2020a**) studies, it is mentioned that adding Dropout layers between each layer of LSTM can reduce the effect of overlearning. But there are differences in the choice of dropout coefficients (0.1, 0.3, 0.5) among the three works of literature above.

Regarding the selection of explanatory variables, in addition to the macroeconomic variables used in many works of literature, **Jagannath et al.**'s (**2021**) research focuses on the core variables of the Bitcoin blockchain, including users, miners, and exchanges. Technical indicators have proven useful for predicting Bitcoin prices (**Jaquart et al. 2021**; **Mudassir et al. 2020**). The LSTM based on the self-adaptive technique also gets good prediction performance, but the article lacks a comparative experiment with the model added macroeconomic variables.

III. SYSTEM DESIGN

The data research sample collected data from a total of 7 natural years from 31 March 2015 to 1 April 2022. However, due to the particularity of Bitcoin having two price bubbles at the end of 2017 and 2021, and the longest span of a single sample in past studies is no longer than 4 years.

Based on the above two reasons, to improve the price prediction accuracy of the model, the total sample is divided into Period 1 (from 31 March 2015 to 30 September 2018) and Period 2 (1 October 2018 with 1 October 2018). Conduct independent research on two sub-samples, train models for their respective periods and predict respectively. Machine learning is the process of training initial samples through training samples and then substituting them into test samples for evaluation.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁑ Peer-reviewed & Refereed journal ⁑ Vol. 11, Issue 6, June 2024

DOI: 10.17148/IARJSET.2024.11639

Usually, training samples occupy 75% to 90% of the samples. The specific division of training and testing samples in this study is shown in **Table 3** and **Figure 6**. The last 10% of the training data is set as validation data.

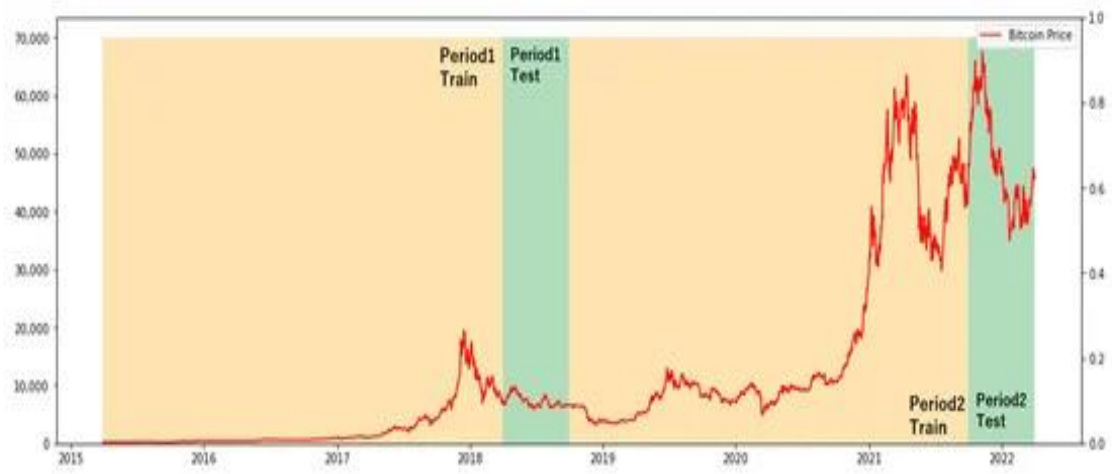

Fig1. Prediction Analysis

Among all the explanatory samples, only ETH has the problem of missing sample data because it came out (7 August 2015) later than April 2015, so the training samples used for ETH in the Period 1 model all start from 7 August 2015, not 31 March 2015.

Bitcoin is available for trading 24 h a day and 365 days a year, while the variables such as stock market indices, exchange rates, and commodity price indices are not traded during weekends and holidays, so there is missing data. There are two ways to deal with samples with these missing data, one is to delete the data with missing data before training, and the second method is to fill in the missing data.

IV. CONCLUSION

In this paper, to predict the price of Bitcoin on the next day, (a) Bitcoin price variables, (b) the specific technical features of Bitcoin, (c) other cryptocurrencies, (d) commodities, (e) market index, (f) foreign exchange, (g) public attention, and (h) dummy variables of the week, a total of eight categories (47 variables) were used as explanatory variables. Random forest regression has the better price prediction accuracy than LSTM. In previous research, LSTM was widely used and recognized as an algorithm with high accuracy when predicting Bitcoin prices. This paper uses the random forest regression machine learning algorithm, which has not been widely used by other researchers in the previous literature and obtains a result with higher prediction accuracy than LSTM. Although random forest regression has the disadvantage of being unable to predict the results that did not appear in the training samples.

For example, when the price of Bitcoin broke the record high, random forest regression could not provide a higher price result than the previous historical high. But with the increase in Bitcoin transaction history, I think random forest regression will perform better when Bitcoin price stabilizes

REFERENCES

[1] Aggarwal, Apoorva, Isha Gupta, Novesh Garg, and Anurag Goel. 2019. Deep Learning Approach to Determine the Impact of Social Economic Factors on Bitcoin Price Prediction. Paper presented at 2019 Twelfth International Conference on Contemporary Computing (IC3), Noida, India, August 8–10. [**Google Scholar**]

[2] Akyildirim, Erdinc, Oguzhan Cepni, Shaen Corbet, and Gazi Salah Uddin. 2021. Forecasting mid-price movement of Bitcoin futures using machine learning. *Annals of Operations Research*, 1–32. [**Google Scholar**] [**CrossRef**] [**PubMed**]

[3] Awoke, Temesgen, Minakhi Rout, Lipika Mohanty, and Suresh Chandra Satapathy. 2021. Bitcoin Price Prediction and Analysis Using Deep Learning Models. In *Communication Software and Networks*. Singapore: Springer, pp. 631–40. [**Google Scholar**]

[4] Basak, Suryoday, Saibal Kar, Snehanshu Saha, Luckyson Khaidem, and Sudeepa Roy Dey. 2019. Predicting the direction of stock market prices using tree-based classifiers.