ANALYSIS OF BITCOIN PRICE PREDICTION USING MACHINE LEARNING MODELS

ARSHDEEP KAUR   JASWINDER SINGH

Abstract – Bitcoin is a peer-to-peer based digital currency and one of the world's top-ranking crypto-currencies.People use Bitcoin because of the transparency and security it offers. Due to its low processing fees and worldwide payments, Bitcoin daily trends have become popular among consumers, investors, and many more. This paper conducts a study on Bitcoin and various machine learning models used for Bitcoin price prediction.

1. DIGITAL CURRENCY AND CRYPTOCURRENCY

Digital Currency also is known as electronic money is a kind of currency that is available in electronic or digital form but not in physical form. Crypto-currency is a sub element of digital currency. Examples are Bitcoin, Litecoin, Ethereum. The differences between digital and crypto-currencies are discussed below:

1) Regulation: In Digital Currency the transactions in the network are regulated by any central authority like a bank. In the case of crypto-currencies, the transactions are governed by the participants. So they are decentralized.

2) Transparency: Digital Currencies are not transparent. One cannot choose an e-wallet and see all the transactions from the beginning. All the information is kept private and confidential. Whereas crypto-currency transactions can be seen by everyone because all the transactions are stored in a public ledger called Blockchain.

3) Legal Aspect: There exists a legal framework for digital currencies in some countries,but for crypto-currencies legal status is undefined.

4) Anonymity: User identification is required in Digital Currencies. One needs to show their photograph and other unique identity documents.Crypto-currencies do not require any confidential information.The Transaction process needs a public and private key.

2. BITCOIN

Bitcoin is a crypto-currency introduced by Satoshi Nakamoto in 2008.It is a decentralized digital currency that allows participants to directly transact with each other without any intermediaries [1]. People use Bitcoin as an investment and medium of exchange where willing parties exchange Bitcoins with each other in a closed network.

As of November 19, 2020, one Bitcoin is equal to 17,946 USD. The smallest unit of Bitcoin is Satoshi, named after the creator of crypto-currency Satoshi Nakamoto.1 Satoshi is equivalent to 0.00000001 Bitcoin. The total number of Bitcoins is fixed that is 21 million units. Countries like The United States, Finland, the European Union, Canada, United Kingdom, etc. accept Bitcoin as a legal tender.

An electronic coin can be defined as a digital signature chain that uses public-key cryptography (means sign the message using the private key and verify the signature using the public key). One user transfers the coin to the next user by electronically signing the next user's public key and a hash of the former transaction and adding both of these to the end of the coin.
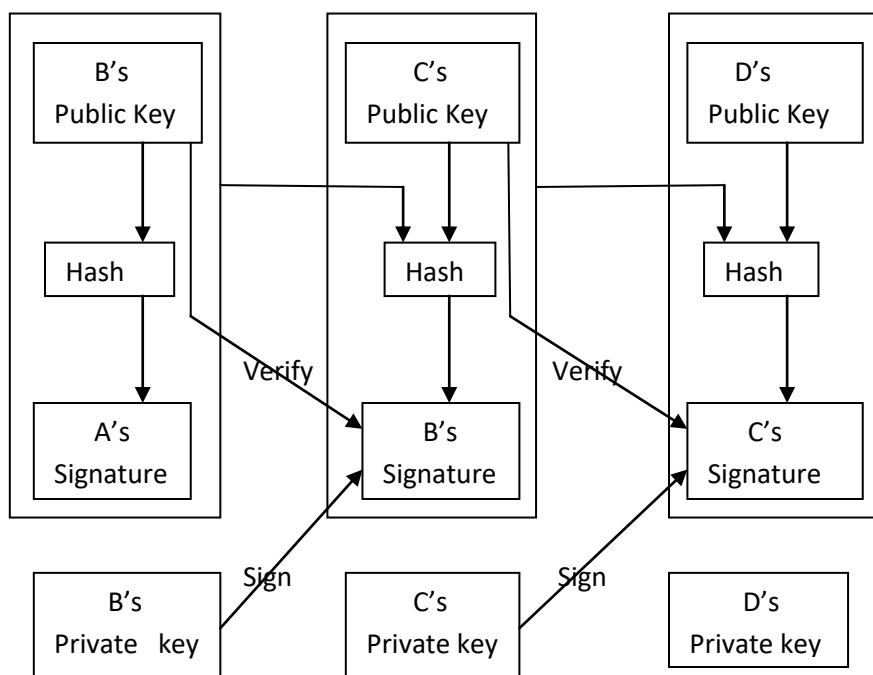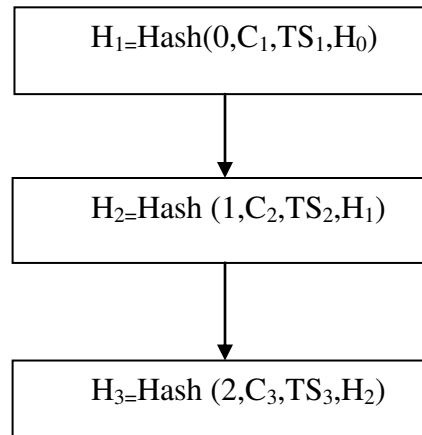

Fig. 1. A Transaction with the private and public key

But the problem with the transactions was that of double-spending and this is solved using a timestamp server.Within timestamping the hash of the block of the items is timestamped.Hash includes the previous timestamp and thus forming a chain-like structure.

$$H_1 = Hash(0,C_1,TS_1,H_0)$$

$$\downarrow$$

$$H_2 = Hash(1,C_2,TS_2,H_1)$$

$$\downarrow$$

$$H_3 = Hash(2,C_3,TS_3,H_2)$$

Here, C is the client id, TS is the timestamp, H is the hash value.

Proof-of-Work is done for the implementation of the timestamp server. The Proof-of-Work includes finding out a value that is when hashed, the hash begins with a certain predefined complexity(no of zero bits as a prefix)

$$H_k = Hash(H_{k-1} \| T \| Nonce)$$

'H_k' denotes Hash Function, 'H_{k-1}'denotes Previous Hash, 'T' denotes set of transactions, Nonce is the random number used in cryptography.

The miners task is to search the nonce value so that they can generate certain difficulty on H_k. The miners have to change the values of nonce so that H_k has predefined complexity. At last, the longest chain is accepted because it is the one having the greatest Proof-of-Work.

3. BITCOIN NETWORK

Initially, the new transaction is broadcasted to all the clients in the network. Every node gathers the transaction into a block and starts doing Proof-of-Work. Whenever any node calculates proof-of-work then it broadcast the block to the rest clients. If the block transactions are valid and not spent twice then all nodes agree to the block. This agreement is shown by working on the next block creation using the previous block hash.

4. MODELS FOR BITCOIN PRICE PREDICTION

a) Regression model

Linear Regression (LR) reveals the relation between the dependent and independent variables.The equation for line fitting data points is as follow:

$$Y=a +b X \qquad\qquad (1)$$

'X' denotes an independent variable, 'Y' denotes a dependent variable, 'b' denotes the slope, and 'a' denotes intercept.

[4] used machine learning models such as linear regression and decision tree to forecast the Bitcoin price. The author predicted the five days price using the models. The results show that linear regression outperformed the decision tree and got the highest accuracy that is 97.59%. Bitcoin due to its peer-to-peer system and decentralization attracted a large number of users around the world. [12] The author found the link between Twitter and google searches and Bitcoin prices. The accuracy of Bitcoin price predicted using Polynomial regression is 77% with tweet volume and 66.66% with google trends.

b) Random Forest

Random Forest is one of the popular regression and classification problems. It merges multiple decision trees for better outcomes.Decision trees address a number of classification issues.The decision tree is like a tree structure where feature space is partitioned recursively.Recursion is terminated when partitioning adds no value to the forecast or until single class samples are present in each node.[6] predicted the Bitcoin daily price as well as Bitcoin 5- minute interval price.Random forest achieved an accuracy of 51% for daily price prediction and 64.8% for five-minute interval price prediction.

c) Support Vector Machine (SVM)

An SVM model is used for binary classification problems. The principle is to find out hyperplane such that there is a maximum margin between two classes of data samples. Machine learning models like SVM, Random Forest, Long Short Term Memory, Quadratic Discriminant Analysis, XGBoost were used for daily Bitcoin price prediction[2]. The Support Vector Machine performed the best with an accuracy of 65.5%.[3] analyzed the behavior of Support vector machine, Artificial Neural Network, and Ensemble algorithm(k-means clustering and recurrent neural networks). The Support Vector Machine algorithm got the best results for price predictions.

d) Recurrent Neural Network

In Simple feed-forward Networks or multilayer perceptrons, the data simply flows from the input layer to the hidden layer and then to the output layer.[10] researched the influence of the network's most frequent edges on prices of Bitcoin by using a single hidden layer feed-forward network and got an accuracy of about 60%. The main limitation was that the crucial information was ignored. Hence the Recurrent Neural Networks come into the picture. In RNN signals can flow both forward and backward in a repetitive manner. The network consists of two inputs, one of historic data and the other for present data. [5] predicted Bitcoin price using RNN and LR model. The RNN model got a low Root Mean Square Error(RMSE) than the LR model because of its capability to identify long term dependencies. In [6]

comparison between linear regression, RNN, and Random forest is performed. The RNN with LSTM increased the efficiency for price prediction of Bitcoin because data is highly fluctuating. The model got 0.0043 mean absolute error(MAE) which is less than that of linear regression and random forest.[11] predicted Bitcoin price by analyzing the Twitter sentiments.Sentiment Analyzer classified the collected tweets into negative and positive tweets. The tweets were fed to the RNN model and the overall accuracy achieved is 77.62%

e) Long Short Term Memory (LSTM)

The limitation of RNN is that the network suffers from problems of vanishing gradient and exploding gradient. The low values of the gradient cause a vanishing problem whereas in the exploding gradient the model mainly focuses on the weight of the network. Thus LSTM came into view. The LSTM increases memory for solving both the problems. The model contains three gates, namely: Output gate, Forget gate, and Input gate. In the case of sequence data, LSTM performs better in extracting long-term dependencies and in representing both future and past information.[7] the author compared the conventional LSTM model with LSTM to AR (2) model. The proposed model outperformed the conventional model as LSTM with AR (2) model evaluation index values are much smaller than that of LSTM. McNally compared the accuracy of Automated Regressive Integrated Moving Average (ARIMA), Recurrent Neural Network (RNN), and Long Short Term Memory (LSTM) models and got the highest 52% accuracy by using the LSTM network [8].

f) Autoregressive Integrated Moving Average (ARIMA)

ARIMA is used for forecasting and examining the time series.

$$AR (p) + MA (q) = ARIMA (p, d, q) \qquad (2)$$

p, d, q are the three orders of ARIMA. 'p' is Auto Regression order, 'd' is differencing order and 'q' is moving average order. [9] used ARIMA model for short term predictions. ARIMA(4,1,4) predicted the Bitcoin price with greater accuracy. The mean absolute error generated was 0.87 for first-day prediction and 5.98 for seventh-day prediction. Hence ARIMA performed better in the short term.

5. OBSERVATION

The increased competition in the mining of bitcoins led to a challenge in future price prediction. The main aim of forecasting the prices of crypto-currency is to maximize the profit of the people. With the increase in prices, the interest of people in Bitcoin
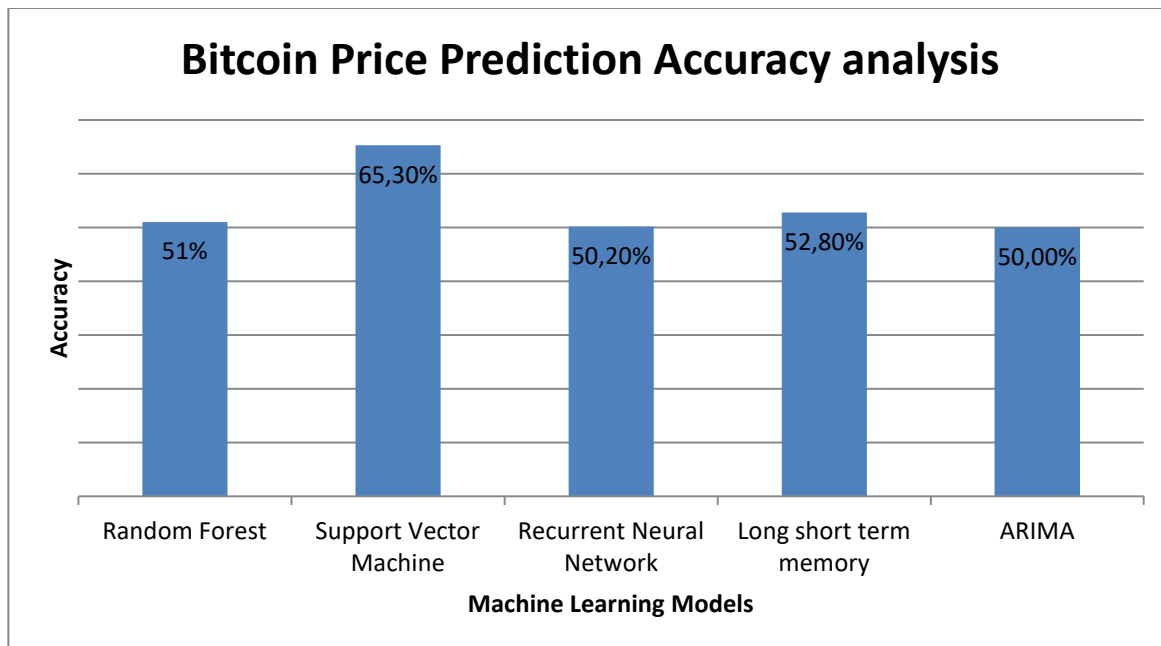
Fig.2. Bitcoin Price Prediction Accuracy Analysis

is also increasing. The prices predicted may differ from the actual prices due to the high volatility of the Bitcoin. Fig 2. Shows the accuracy of various machine learning models. From the figure, it is clear that the accuracy levels achieved are less.

6. CONCLUSION

Bitcoin is a unique digital currency that provides possibilities for research. This paper is the analysis of the various machine learning models used in forecasting Bitcoin prices. The prices of Bitcoin are highly volatile that needs to be maintained. The observation shows the various machine learning models along with their accuracy levels.This analysis could be extended by researching more of the advanced models for price prediction.

REFERENCES

[1] S. Nakamoto, Bitcoin: A peer to peer electronic cash system,2008

[2] Zheshi Chen, Chunhong Li, Wenjun Sun, "Bitcoin price prediction using machine learning: An approach to sample dimension engineering", Journal of Computational and Applied Mathematics 365 (2020) 112395

[3] D.C.A. Mallqui and R.A.S. Fernandes, "Predicting the direction, maximum, minimum and closing prices of daily Bitcoin exchange rate using machine learning techniques", Applied Soft Computing Journal(2018),https://doi.org/10.1016/j.asoc.2018.11.038

[4] K. Rathan, S .V. Sai, and T. S. Manikanta, "Crypto-Currency price prediction using Decision Tree and Regression techniques",3rd International Conference on Trends in Electronics and Informatics(ICOEI),2019, pp. 190-194

[5] H. Kavitha, U.K. Sinha, and S.S. Jain, "Performance Evaluation Of Machine learning Algorithms for Bitcoin Price Prediction", Fourth International Conference on Inventive Systems and Control (ICISC),2020, pp. 110-114

[6] S. Tandon, S. Tripathi, P. Saraswat, and C. Dabas, "Bitcoin Price Forecasting using LSTM and 10 Fold Cross-Validation", International Conference on Signal Processing and Communication(ICSC),2019, pp. 323-328

[7]C. Wu, C. Lu, C. Ma, and R. Lu, "A New Forecasting Framework for Bitcoin price with LSTM", IEEE International Conference on Data Mining Workshops(ICDMW),2018, pp. 168-175

[8] S. McNally, J. Roche, and S. Caton," Predicting the price of Bitcoin Using Machine Learning",26th Euromicro International Conference on Parallel, Distributed and Network-based Processing(PDP),2018, pp. 339-343

[9]I.M. Wirawan, T.Widiyaningtyas, and M.M.Hasan, "Short Term Prediction on Bitcoin Price using ARIMA method", International Seminar on Application for Technology of information and communication,2019, pp.260-265

[10] Marcell Tamas Kurbucz," Predicting the price of Bitcoin by the most frequent edges of its transaction network", Economics Letters, 184 (2019),108655

[11] D.R. Pant, P. Neupane, A. Poudel, A.K. Pokhrel, and B.K. Lama, "Recurrent Neural Network Based Bitcoin Price Prediction by Twitter Sentiment Analysis", IEEE 3rd International Conference on Computing, Communication and Security(ICCCS),2018, pp. 128-132

[12] A. Mittal, V. Dhiman, A. Singh and C. Prakash, "Short –Term Bitcoin Price Fluctuation Prediction Using Social Media and Web Search Data", Twelfth International Conference on Contemporary Computing (IC3), 2019, pp. 1-6