International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2020 IJSRCSEIT | Volume 6 | Issue 2 | ISSN : 2456-3307
DOI : https://doi.org/10.32628/CSEIT206237

A Comparative Study of Bitcoin Price Prediction Using Machine Learning Algorithms

M. Akhil Sai[1], K. Sarath Chandra Sai[1], M. Manu Koushik[1], K. Gowri Raghavendra Narayan[2]
[1]UG students, Computer Science and Engineering, VVIT, Nambur, Andhra Pradesh, India
[2]Assistant Professor Computer Science and Engineering, VVIT, Nambur, Andhra Pradesh, India

ABSTRACT

ML and AI-helped exchanging have pulled in developing enthusiasm for as far back as not many years.We examine day-by-day information for different digital currencies over some stretch of time. We show that straightforward exchanging methodologies helped by innovative AI calculations outflank standard benchmarks. We have picked two Machine Learning Algorithms to play out a Comparative Study to foresee cost of a Bitcoin; we have utilized Decision tree regressor and LSTM Algorithms and watched execution of every calculation as far as anticipating the cost of Bitcoin. We saw that Decision tree regressor gives progressively effective and precise outcomes when contrasted with others.

Keywords : LSTM, Decision tree, RNN, Bitcoin

I. INTRODUCTION

ML and AI-helped exchanging have pulled in developing enthusiasm for as far back as scarcely any years. Bitcoin as of late has a great deal of consideration from the media, people in general because of its ongoing value raise, and fall. Numerous specialists have explored different components that influence the Bitcoin cost and the examples behind its vacillations, specifically, utilizing different AI strategies.Figure1.1 shows bitcoin price from January 1st 2017 to march9th 2020 on CoinMarketCap[1] which is a site, which describes the daily crypto prices of all the crypto currencies.

Figure 1.1:-Bitcoin price Movement

II. LITERATURE SURVEY

Huisu Jang et al[2], "An Empirical Study on Modelling and Prediction of Bitcoin Prices with Bayesian Neural Networks based on Block chain Information" their model is more efficient in predicting the stock prices and they provided evidence that several basic financial and economic factors have predictive power forthe market excess return. F. Andrade de Oliveira et

al[3],"The use of artificial neural networks in the analysis and prediction of stock prices" they predicted patterns from the data. M. Daniela and A. BUTOI[4],"Data mining on Romanian stock market using neural networks for price prediction" they concluded that Multi-Layer perceptron can be used for stock price prediction. D. Shah and K. Zhang[5], "Bayesian regression and Bitcoin" their survey deals with daily time series data,30-minutes,60-minutes and120 minutes time-interval data to predict the bitcoin price. McNally et al[6],"Predicting the Price of Bitcoin Using Machine Learning" they proposed two prediction models based on recurrent neural networks (RNNs) and long short-term memory (LSTM), and compared them with an autoregressive integrated moving average(ARIMA) model. K Struga et al[7], "Bitcoin Price Prediction with Neural Networks" they proposed that bitcoin price can be predicted using long short term memory(LSTM).

III. PROPOSED METHODOLOGY

We did a comparative study on predicting Bitcoin prices using Decision tree and LSTM algorithms

Decision Tree Algorithm:

Decision Tree algorithm belongs to the family of supervised learning algorithms. The decision tree algorithm tries to solve the problem, by using tree representation. Each internal node of the tree corresponds to an attribute, and each leaf node corresponds to a class label as shown in the figure 3.1[8].

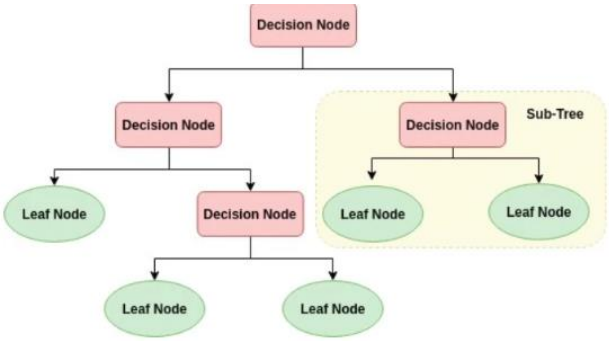

Figure 3.1:- Structure of decision tree

Pre-processing:

1.Data cleaning: Here, we performed the drop operation and removed the unwanted attributes from the dataset and also removed the incomplete columns and rows

2.Normalization: We scale the data using Min-Max Scaler to the range 0-1 so that the operations can be performed efficiently.

After performing the pre-processing of the dataset , next we must select a feature from the data set which will used to predict the price. Here, we are taking the feature with name 'last' which is the price of the bitcoin on that specific time and date. Then, we applied the fit-transform method to the selected feature and converted it into the array. So that it will be suitable to give it was the input to the decision tree regressor algorithm.

Then, we selected 67 % data samples as training set and remaining percent of data samples as testing set. Next, By taking train set (trainX) we formed trainY . Next , we give both the sets as the input for the decision tree regressor algorithm and predict the price using the training set . Same will be followed for the test set and the performance of our model will be checked by giving the test set as the input to the model.

LSTM Algorithm:

Long short-term memory is an artificial recurrent neural network (RNN) architecture used in the field of deep learning. LSTM systems are appropriate to arranging , handling and make expectations dependent on time-arrangement information, since there can be slacks of obscure term between significant occasions in a period arrangement. LSTMs were created to manage the detonating and disappearing angle issues that can be experienced when preparing conventional RNNs. Below figure3.2[9] shows the architecture of LSTM.

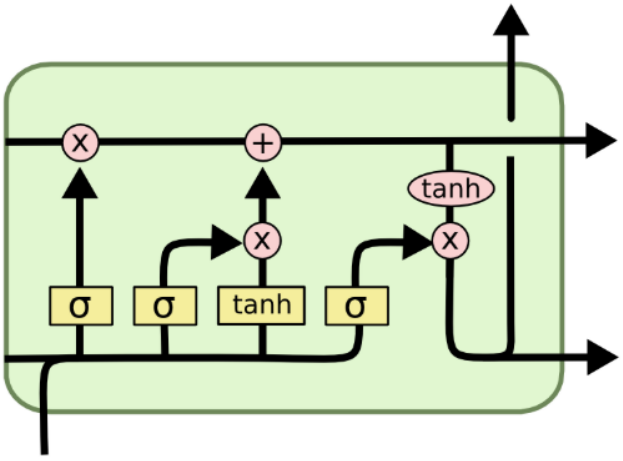

Figure 3.2:-LSTM Architectural diagram

MODULES

1. Dataset Collection
2. Data Processing
3. Predicting Polarity using RNN

1.Dataset Collection:

We first started with Bit coin market data that was publicly available on Kaggle. The time span comprises of 1,574,274 minutes. For each timestamp, the information remembered data for the initial worth, the end esteem, the most noteworthy worth, the least worth, the volume exchanged, and the weighted cost.

2.Data Processing:

1. Data cleaning
2. Normalization

3. **Predicting Polarity using RNN**:

The undeniable ultimate objective of making a digital currency based neural system is to anticipate value vacillations progressively. In view of this objective, we were anxious to begin with a profoundly transiently settled dataset. In the event that we could get data on a moment by moment or step by step timescale, we could improve employment of foreseeing costs and remaining in front of the market. Moreover, there would be a huge number of information focuses and that would be a dataset size that neural systems exceed expectations at. In any case, as we suggested above, we understood that there likewise issues with profoundly settled information. When taking a gander at our moment dataset, we had an instinct that there would be no change on a moment timescale, or if there was change, that it was little and loud.

After performing these three steps we take 70 % of our dataset as training set and remaining percent as test data. Training set is used to train the LSTM model and its performance can be analysed using the test data.

IV. RESULTS AND DISCUSSION

RMSE means root mean square error. The RMSE is calculated using the formula shown in figure 4.1[10]

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (Predicted_i - Actual_i)^2}{N}}$$

Figure 4.1:-RMSE

Decision tree:

When we calculated the RMSE value for the decision tree regressor model test data we got the test score as 3.57 RMSE and the figure 4.2 shows the graph for model.

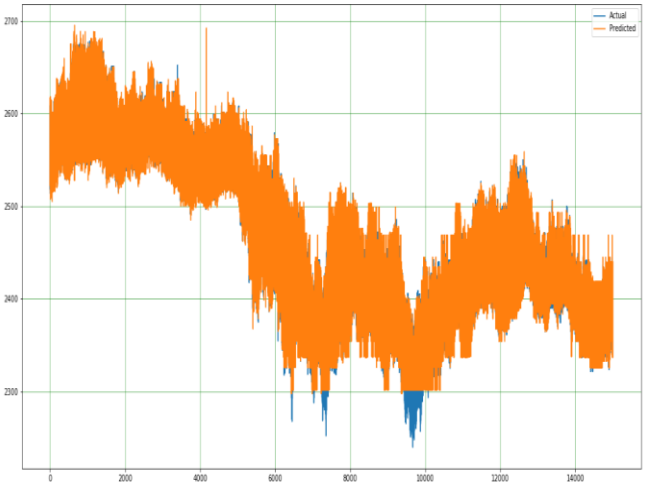

Figure 4.2:- decision tree regressor model

LSTM:

When we calculated the RMSE value for the long short term memory model test data, we got the test score as 5.96 RMSE and the figure 4.3 shows the graph For the model.

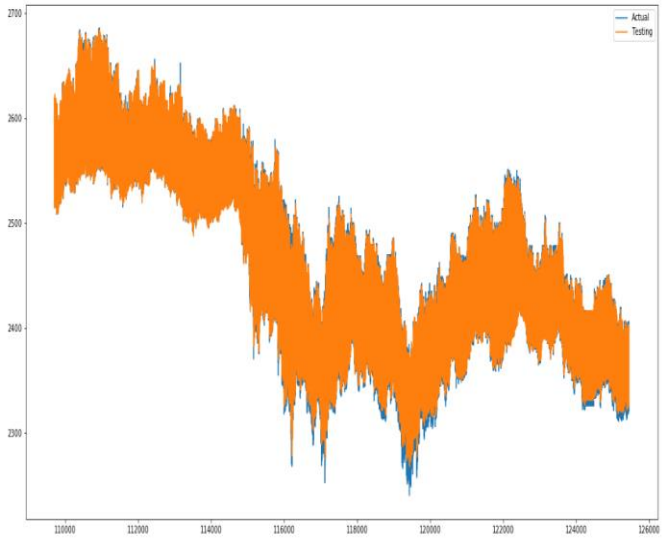

Figure 4.3:-LSTM model

V. CONCLUSION

As it is a regression problem, we calculated the RMSE values of our two models. The Model which will have the less RMSE is the best model. From our results we got less RMSE value for the decision tree regressor. So, we conclude that decision tree regressor is best compared to LSTM when the data taken is having the minimum difference.

VI. REFERENCES

[1]. They track the capitalization of different crypto currencies https://coinmarketcap.com/

[2]. Huisu Jang et al, "An Empirical Study on Modeling and Prediction of Bitcoin Prices with Bayesian Neural Networks based on Block chain Information" IEEE vol.6,pp.5427-5437,2017.

[3]. F. Andrade de Oliveira et al, "The use of artificial neural networks in the analysis and prediction of stock prices" IEEE International Conference on Systems, Man, and Cybernetics, 2011.

[4]. Daniela and A. BUTOI, "Datamining on Romanian stock market using neural networks for price prediction" Informatica Economica, Academy of Economics Studies - Bucharest, Romania, vol. 17(3), pages 125-136, 2013.

[5]. D. Shah and K. Zhang, "Bayesian regression and Bitcoin" IEEE 52nd Annual Allerton Conference on Communication, Control, and Computing (Allerton), 2014.

[6]. McNally et al," Predicting the Price of Bitcoin Using Machine Learning" IEEE 26th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP), 2018.

[7]. K Struga et al[7], "Bitcoin Price Prediction with Neural Networks" RTA-CSIT, 2018.

[8]. Avinash Nalvani, Decision tree classification https://www.datacamp.com/community/tutorials/decision-tree-classification-python

[9]. Architectural diagram of LSTM from Researchgatehttps://www.researchgate.net/figure/Basic-LSTM-Unit-Transfer-Function-Diagram-from-10_fig8_308804546.

[10]. Root mean square error formula https://www.includehelp.com/ml-ai/root-mean-square%20error-rmse.aspx.

Cite this article as :

M. Akhil Sai, K. Sarath Chandra Sai, M. Manu Koushik, K. Gowri Raghavendra Narayan, "A Comparative Study of Bitcoin Price Prediction Using Machine Learning Algorithms ", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 6 Issue 2, pp. 160-163, March-April 2020. Available at doi : https://doi.org/10.32628/CSEIT206237

Journal URL : http://ijsrcseit.com/CSEIT206237