Bitcoin Financial Forecasting

Dhyanendra Jain¹, Ashu Jain², Amit Kumar Pandey³, Jogender Kumar⁴
Dr. Akhilesh Das Gupta Institute of Technology & Management, New Delhi

¹dhyaneendra.jain@gmail.com, ²ajainashu@gmail.com, ³amitpandey33@gmail.com, ⁴jogender@rediffmail.com

Abstract. Bitcoin is one of the crypto currencies and is most unpredictable currencies. In the world of crypto currencies, the value of a coin can unpredictably upgrade or degrade. In this study, the model has been trained to predict the value of Bitcoin in USD at any given time stamp. For this prediction, three algorithms of machine learning - Linear Regression (LR), Support Vector Regression (SVR) and Neural Network Regression (NNR) have been used. The model is trained using the collected dataset. After the collection of data set, we first applied SVR algorithm, then we used LR and then NNR to calculate the error compared to the actual value. Root mean squared error (RMSE) is used as the predictive measure. Out of the three algorithms, LR was found out to be more accurate for predicting the value of bitcoin.

1. Introduction

Bitcoin is an inclining crypto currency on the planet with appropriation developing reliably after some time. Because of the open idea of bitcoin it likewise represents another worldview instead of customary financial markets. It works on a decentralized, shared and trustless framework in which all transactions are presented on an open record called the Blockchain. This kind of straightforwardness is incredible in other financial markets.

Foreseeing the time series is the use of a model to predict the upcoming values which are dependent on the current values. In the crypto currency world, costs are exceptionally unstable. This implies the estimation of a coin can go up or down actually rapidly, with regularly no clarification concerning why.

Prediction of develop financial markets, for example, the securities exchange has been looked into finally. Bit coin displays an intriguing corresponding to this as it is a period arrangement expectation issue in a market still in its transient stage. Subsequently, there is high instability in the market and this gives an open door as far as expectation. Customary time arrangement expectation techniques depend on linear suppositions and require information that can be separated into pattern, occasional and clamor to be compelling. This kind of procedure is increasingly appropriate for an errand, for example, anticipating deals where regular impacts are available. Because of the absence of regularity in the Bit-coin market and its high instability, these techniques are not successful for this undertaking. Given the multifaceted nature of the assignment, Machine learning (ML) makes for a fascinating innovative arrangement dependent on its exhibition in comparable zones. Assignments, for example, natural language processing which is additionally successive in nature and has demonstrated promising outcomes. This kind of assignment utilizes information of a consecutive sort and thus is like a value expectation task. All through this undertaking, Linear Regression (LR), Neural Network Regression (NNR) and Support Vector Regression (SVR) have been utilized to show up at the ideal aftereffect of anticipating the bitcoin.
Through this paper we are aiming to train our model to predict the value of the bitcoin at any given instant of time and see the error value as compared to the actual value at that instant. For this purpose, root mean squared value of the error has been calculated.

Organization of the paper is: Section 2 includes the literature survey done for the predictive analysis using machine learning algorithms. Section 3 gives the overview of the models used. Section 4 discusses about the dataset used. Section 5 includes the predictive analysis. Section 6 includes conclusion and directions of future work.

2. Related Work

Günay and Ensarı [1] broke down the notable ML calculations which are for the most part utilized in the past investigations to plan another model to foresee customer churn. Kanawaday and Sane[ 2] investigated the utilization of AutoRegressive Integrated Moving Average (ARIMA) anticipating on the time arrangement information gathered from different sensors from a Slitting Machine, to foresee the potential disappointments and quality deformities, hence improving the general assembling process. Charleonnan et al [3] presented ML strategies for anticipating the chronic kidney disease utilizing clinical information. Four ML techniques are investigated which includes K-closest neighbors (KNN), support vector machine (SVM), logistic regression (LR), and choice tree classifiers. These models are developed by using the dataset of chronic kidney disease and then all the models are compared to select the best one. Rejawat et al [4] proposed a hybrid ML procedure to classify individuals into patient and healthy classes, separately. The patient class is the people having diabetes and the healthy class is the people having no diabetes. Proposed model was assessed using various implementation measures. Pushpa et al [5] utilized the idea of ML to anticipate the consequence of class students. In view of the presentation of the students in past semester, and the scores of internal examinations of the current semester, the final outcome, regardless of whether the understudy passes or fails the current semester is registered before the final assessment really happens. Hassan et al [6] showed up with the point of aggregating a major piece of the information about solar power anticipating, concentrating on the latest headways and future patterns. Karasu et al [7] performed predictive analysis of bitcoin value with LR and SVM from ML techniques by utilizing time arrangement comprising of every day Bitcoin shutting costs during 2012-2018. Velankar et al [8] endeavored to anticipate the Bitcoin cost precisely thinking about different parameters that influence the Bitcoin value. For the principal period of our examination, they expected to comprehend and distinguish day by day inclines in the Bitcoin showcase while picking up understanding into ideal highlights encompassing Bitcoin cost. Sakiz and Kutlugun [9] in the wake of clarifying the fundamental ideas driving distributed architecture and blockchain innovation behind crypto cash, artificial intelligence calculations were exploited, and dependent on most recent three years values of bitcoin determining was performed for Bitcoin which has a colossal piece of the overall industry in since nine years.

3. Models Used

3.1. Linear Regression

Regression is a strategy for exhibiting a target esteem dependent on independent pointers. This strategy is generally used for assessing and finding conditions and legitimate outcomes association between factors. Regression methods generally differentiate dependent on the amount of independent components and the sort of association between the independent and dependent variables.

Direct linear regression is a kind of regression assessment where the amount of independent components is one and there is a linear association between the independent(x) and dependent(y) variable.

3.2. Neural Network Regression
Albeit neural systems are comprehensively known for use in profound learning and showing complex issues, for instance, picture acknowledgment, they are viably changed in accordance with regression issues. Any class of quantifiable models can be named a neural system in case they use adaptable loads and can assessed non-linear components of their data sources. Along these lines neural system regression is fit to issues where an inexorably standard regression model can't fit an answer.

Neural system regression is a directed learning procedure, and subsequently requires a named dataset, which consolidates a name area. Since a regression model predicts a numerical worth, the name fragment must be a numerical data type.

You can prepare the model by giving the model and the labeled dataset as a contribution to Train Model or Tune Model hyper parameters. The prepared model would then be able to be utilized to foresee values for the new information models.

Neural network wording is enlivened by the natural activities of specific cells called neurons. A neuron is a cell that has a few information sources that can be enacted by some outside procedure. Contingent upon the measure of activation, the neuron creates its own action and sends this along its yields. What's more, explicit info or yield ways might be "fortified" or weighted higher than different ways. The speculation is that since the human mind is only a network of neurons, we can imitate the cerebrum by displaying a neuron and interfacing them through a weighted diagram.

3.3. Support Vector Regression
SVRs are unmistakable class of calculations, portrayed by utilization of pieces, nonappearance of neighborhood minima, meager condition of the arrangement and limit control acquired by following up on the edge, or on number of support vectors, and so forth.

4. Dataset
We have used the data bitstamp USD 1-min data from January 2012 to April 2019. The dataset is a 266775 row and 8 column dataset. Rows represent different data points. Columns represent parameters of the collected data. The columns are classified as:-

- Timestamp- Represents time at which the values are recorded.
- Open- denotes the opening price in the stock market.
- High- Represents the highest value incurred within opening and closing price.
- Low- denotes the lowest value incurred within opening and closing price.
- Close- denotes the closing price in the stock market.
- Volume of Bitcoin Traded- Represents the volume of bitcoin traded.
- Volume of Bitcoin Traded in price- Represents the actual price of the traded volume.
- Weighted price- Represents the average price corresponding to a timestamp.

Context
This data set contains minute wise data on bitcoin values

Content
Time Period: January 2012 – April 2019 (7 years 4 months)
Granularity: Minute Wise

5. Predictive analysis
For the predictive analysis of the bitcoin using the ML algorithms, RMSE value is used for evaluating the performance

\[
RMSE = \sqrt{(p - a)^2}
\]
Where p is the predicted value and a is the original value.
The values of mean squared root errors on test obtained on the three machine learning algorithms: Support Vector Regression, Linear Regression, Neural Network Regression (Table 1). Graphs are also obtained by plotting training data points with green and test data points with red. (Figure 1-3)

| S. No. | Algorithm Used                  | RMSE value |
|--------|---------------------------------|------------|
| 1.     | Support Vector Regression       | 1.279      |
| 2.     | Linear Regression               | 1.176      |
| 3.     | Neural Network Regression       | 3.71       |

5.1. Support Vector Regression
In figure 1, it can be seen that the green and red points are clearly distinct at certain data points. This indicates significant errors in the predictions made by the model.

![Figure 1. Mean squared root errors on Support Vector Regression](image)

5.2. Linear Regression
In figure 2, it can be seen that the green and red points are not much distinct. This indicates quite low error in the predictions made by the model. However, there is significant error at some data points.
5.3. Neural Network Regression

From figure 3, the distinction between green and red points is between the observations of 1st and 2nd model, indicating low error compared to 1st model but higher than 2nd model.

By the analysis of Table 1 and figure 1-3, it can be said that linear regression predicted the value of the bitcoin most accurately among the three algorithms.

6. Conclusion

In this paper, machine learning models – LR, NNR and SVR have been used to predict the bitcoin value. To compare the algorithms, root mean squared error value is used as the accuracy measure. Results show that LR performed best among the three algorithms.

In this paper, only three algorithms have been used. Other various ML algorithms can be explored in predicting the value of the bitcoin. Also, here only one dataset is used for the training and the testing purpose as well. Multiple datasets can be used to get more accurate results.

References

[1] M. Günay and T. Ensarı, "Predictive churn analysis with machine learning methods," 2018 26th Signal Processing and Communications Applications Conference (SIU), Izmir, 2018, pp. 1-4.

[2] A. Kanawaday and A. Sane, "Machine learning for predictive maintenance of industrial machines using IoT sensor data," 2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS), Beijing, 2017, pp. 87-90.

[3] A. Charleonnan, T. Fufaung, T. Niyomwong, W. Chokchueypattanakit, S. Suwannawach and N. Ninchawee, "Predictive analytics for chronic kidney disease using machine learning techniques," 2016 Management and Innovation Technology International Conference (M1Ticon), Bang-San, 2016, pp. MIT-80-MIT-83.

[4] P. S. Rajawat, D. K. Gupta, S. S. Rathore and A. Singh, "Predictive Analysis of Medical Data using a Hybrid Machine Learning Technique," 2018 First International Conference on
Secure Cyber Computing and Communication (ICSCCC), Jalandhar, India, 2018, pp. 228-233.

[5] S. K. Pushpa, T. N. Manjunath, T. V. Mrunal, A. Singh and C. Suhas, "Class result prediction using machine learning," 2017 International Conference On Smart Technologies For Smart Nation (SmartTechCon), Bangalore, 2017, pp. 1208-1212.

[6] M. Z. Hassan, M. E. K. Ali, A. B. M. S. Ali and J. Kumar, "Forecasting Day-Ahead Solar Radiation Using Machine Learning Approach," 2017 4th Asia-Pacific World Congress on Computer Science and Engineering (APWC on CSE), Nadi, 2017, pp. 252-258.

[7] S. Karasu, A. Altan, Z. Saraç and R. Hacioğlu, "Prediction of Bitcoin prices with machine learning methods using time series data," 2018 26th Signal Processing and Communications Applications Conference (SIU), Izmir, 2018, pp. 1-4.

[8] S. Velankar, S. Valecha and S. Maji, "Bitcoin price prediction using machine learning," 2018 20th International Conference on Advanced Communication Technology (ICACT), Chuncheon-si Gangwon-do, Korea (South), 2018, pp. 144-147.

[9] B. Sakiz and E. Kutlugün, "Bitcoin price forecast via blockchain technology and artificial intelligence algorithms," 2018 26th Signal Processing and Communications Applications Conference (SIU), Izmir, 2018, pp. 1-4.