IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.39 ⁜ Vol. 11, Issue 6, June 2022

DOI: 10.17148/IJARCCE.2022.11620

BITCOIN PRICE PREDICTION USING MACHINE LEARNING TECHNIQUES

Ginjupalli Venkata Naga Sai[1], E Satish Goud[2], S Venkata Sai Kumar[3], N Naga Srinivas[4], Dr. John Basha[5]

[1-4]B. Tech (CS-DS), JAIN University, Bengaluru, Karnataka.

[5]Project Guide, JAIN University, Bengaluru, Karnataka

Abstract: Crypto-currency such as Bitcoin is more popular these days among investors. In the proposed work, it is attempted to predict the Bitcoin price accurately taking into consideration various parameters that affect the Bitcoin value. For the first phase of investigation, it is aimed to understand and identify daily trends in the Bitcoin market while gaining insight into optimal features surrounding Bitcoin price. The data set consists of various features relating to the Bitcoinprice and payment network over the course of time, recorded daily. For the second phase of investigation, using the available information, we will predict the sign of the daily price change with highest possible accuracy

INTRODUCTION:

Bitcoin

Bitcoin is a crypto currency which is used worldwide for digital payment or simply for investment purposes. Bitcoin is decentralized i.e. it is not owned by anyone. Transactions made by Bitcoins are easy as they are not tied to any country. Investment can be done through various marketplaces known as “bitcoin exchanges”. These allow people to sell/buy Bitcoins using different currencies. The largest Bitcoin exchange is Mt Gox. Bitcoins are stored in a digital wallet which is basically like a virtual bank account. The record of all the transactions, the timestamp data is stored in a place called Blockchain. Each record in a blockchain is called a block. Each block contains a pointer to a previous block of data. The data on blockchain is encrypted. During transactions the user’s name is not revealed, but only their wallet ID is made public.

Prediction:

The Bitcoin’s value varies just like a stock albeit differently. There are a number of algorithms used on stock market data for price prediction. However, the parameters affecting Bitcoin are different.Therefore it is necessary to predict the value of Bitcoin so that correct investment decisions can be made. The price of Bitcoin does not depend on the business events or intervening government unlike stock market. Thus, to predict the value we feel it is necessary to leverage machine learningtechnology to predict the price of Bitcoin.

Bitcoin (BTC) is a novel digital currency system which functions without central governing authority. Instead, payments are processed by a peer-to-peer network of users connected through the Internet. Bitcoin users announce new transactions on this network, which are verifiedby network nodes and recorded in a public distributed ledger called the blockchain. Bitcoin is the largest of its kind in terms of total market capitalization value. They are created as a reward in a competition in which users offer their computing power to verify and record transactions into the blockchain. Bitcoins can also be exchanged for other currencies, products, and services. The exchange of the Bitcoins with other currencies is done on the exchange office, where "buy" or "sell" orders are stored on the order book. "Buy" or "bid" offers represent an intention to buy certain amount of Bitcoins at some price while "sell" or "ask" offers represent an intention to sell certain amount of Bitcoins at some price. The exchange is done by matching orders by price from order book into a valid trade transaction between buyer and seller.

Volatility as a measure of price fluctuations has a significant impact on trade strategies and investment decisions as well as on option pricing and measures of systemic risk. The order book data can give us the additional information to predict future volatility by providing insights into the liquidity and trading intentions. In the proposed work, we focus on the short-term predictionof volatility from order book data from machine learning perspective, with no intention to produce another financial model of volatility or order book separately.

Bitcoin, as a pioneer in the blockchain financial renaissance plays a dominant role in a whole cryptocurrency market capitalization ecosystem. Therefore, it is of great interest of data miningand machine learning community to be able to: (i) predict Bitcoin price fluctuations and (ii) give insights to understand what drives the Bitcoin volatility and better estimate associated risks incryptocurrency domain.

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁜ Impact Factor 7.39 ⁜ Vol. 11, Issue 6, June 2022
DOI: 10.17148/IJARCCE.2022.11620

The main contributions are:

(1) We formulate the problem of predicting short-term Bitcoin price fluctuations from order book as learning predictive models over both volatility series and order book features;

(2) We propose interpretable temporal mixture models to capture the dynamical effect of order book on the volatility evolution;

(3) The comprehensive experimental evaluation demonstrates the superior performance of the mixture model in comparison to numerous time series models and ensemble methods

(4) By analyzing the components of the mixture model, we detect regimes when order book dynamics affect future short-term volatility;

(5) In addition, we adopt rolling and incremental learning and evaluation schemes to study the robustness of models with respect to the look-back horizon of historical data.

LITERATURE SURVEY

Generalized Autoregressive Conditional Heteroskedasticity (GARCH) volatility analysis is performed to explore the time series of Bitcoin price. Various studies on statistical or economical properties and characterizations of Bitcoin prices refer to its capabilities as a financial asset; these research focus on statistical properties, inefficiency of Bitcoin according to efficient market hypothesis, hedging capability, speculative bubbles in Bitcoin, the relationship between Bitcoin and search information, such as Google Trends and Wikipedia, and wavelet analysis of Bitcoin.Bitcoin price prediction done by P. Ciaian, M. Rajcaniova, and D. Kancs evaluates Bitcoin price formation based on a linear model by considering related information that is categorized into several factors of market forces, attractiveness for investors, and global macro-financial factors. They assume that the first and second factors mentioned above significantly influence Bitcoin prices but with variation over time. S. McNally et.al predicts the Bitcoin pricing process using machine learning techniques, such as recurrent neural networks (RNNs) and long short-term memory (LSTM), and compare results with those obtained using autoregressive integrated moving average (ARIMA) models. A machine trained only with Bitcoin price index and transformed prices exhibits poor predictive performance.Existing techniques compares the accuracy of predicting Bitcoin price through binomial logistic regression, support vector machine, and random forest.

EXISTING SYSTEM

Existing technique have created three time series data sets for 30, 60 and 120 minutes followed by performing GLM/Random Forest on the datasets which produces three linear models. These three models are linearly combined to predict the price of Bitcoin.

Another existing work analyzed what has been done to predict the U.S. stock market. The conclusion of work is the mean square error of the prediction network was as large as the standard deviation of the excess return.

Drawbacks

This method providing evidence that several basic financial and economic factors have predictive power for the market excess return.

Another existing work, predict trend of the stock. The trend can be considered as a pattern. They perform both short term predictions (day or week predictions) and also long-term predictions (months). They found that the latter produced better results with 79% accuracy.

Drawbacks

Instead of directly forecasting the future price of the stock, it analyzed trend only

Another work proposed performance evaluation criteria of the network. Based on the predicted output the performance evaluation algorithm decides to either buy, sell or hold the stock.

Drawbacks

There is no strong method for price prediction.

PROPOSED SYSTEM

Acquire time-series data recorded daily for five certain time period at different time instances, it must be normalized and smoothened. The next step is to select parameters that will be fed to the predictive network. From an array of available features, some are mentioned below:

After feature selection, the sample inputs will be fed to the model.

The accuracy can be compared with different models after the final prediction.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.39 ⁜ Vol. 11, Issue 6, June 2022

DOI: 10.17148/IJARCCE.2022.11620

METHODOLOGY OF PROPOSED SYSTEM

We have collected the dataset for the document with following details from quandl.com and we applied machine learning algorithm such as decision tree and regression for prediction and price forecast. We have collected the dataset for the document with following details from quandl.comand we applied machine learning algorithm such as decision tree and regression for prediction and price forecast.

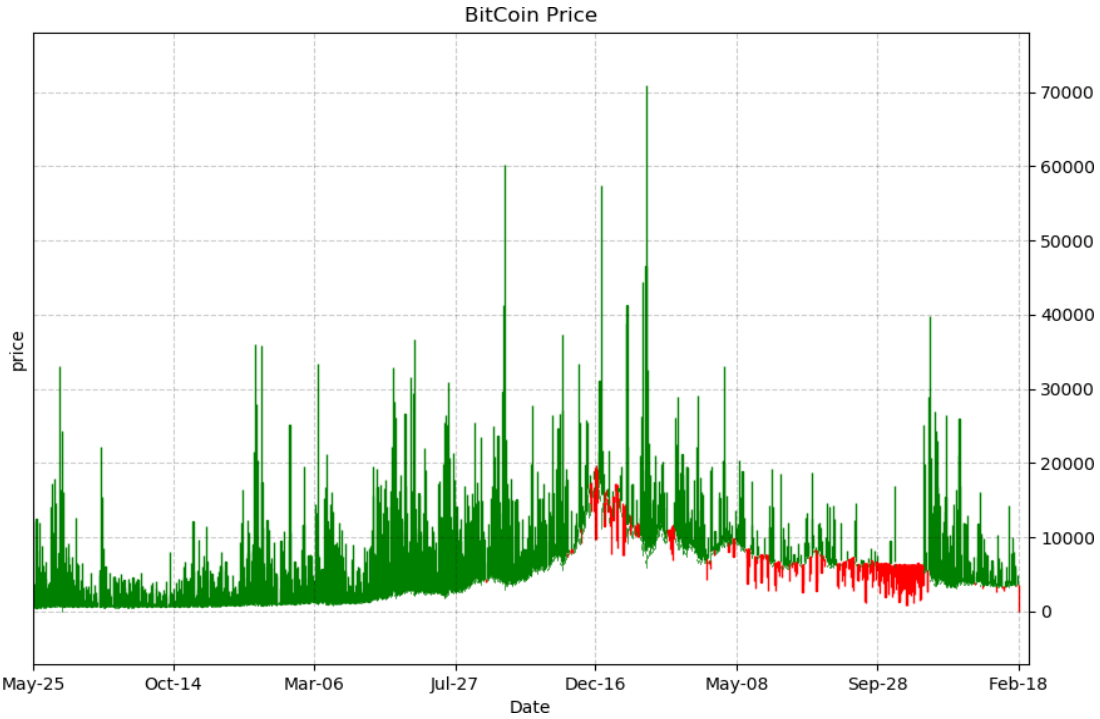
PRICE PREDICTION AND PRICE FORECAST

In this section, we will explain our research steps in detail, with the focus to describe priceprediction and price forecast of bitcoin using regression and decision tree techniques.

DATASET DETAILS

As Bitcoin is a kind of stock traded in stock market, dataset will be available in plenty with all timeintervals. Live data from 2011 to till date is collected from quandl.com, which provided us the most comprehensive bitcoin price in datewise data. Dataset is extracted to CSV file.As Bitcoin is a kind of stock traded in stock market, dataset will be available in plenty with all time intervals. Livedata from 2011 to till date is collected from quandl.com, which provided us the most comprehensive bitcoin price in datewise data. Dataset is extracted to CSV file.Though there are many authorize websites are available for collecting bitcoin dataset for study, CoinMarketCap is one of the other authorized websites, which provides the transactions that bitcoin traded for the 24 hours of a day. These data are fed from various exchanges handling crypto currency.

Quandl.com has dataset related to finance, economic data from five hundred publishers. Data published in Quandl.com can be exploited for different development platforms and analysis tools.In this proposed work, we have collected the Quandl.com data with name mentioned as “BITSTAMPUSD”.The data collected with following features and stored as data.csv “Time_stamp, Open, High, Low, Close, Volume_btc, Volume_currency, Weighted_price”.



Data Visualization of Bitcoin price

PRICE PREDICTION

Price prediction on the considered dataset is done using two different machine learning algorithms such as Decision Tree and linear regression. The predicted value is compared for thepredicted accuracy and error values.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.39 ⁜ Vol. 11, Issue 6, June 2022

DOI: 10.17148/IJARCCE.2022.11620

A. Decision Tree

Decision tree is a type of supervised learning algorithm that is mostly used in classification problems. It works for both categorical and continuous input and output variables. In this technique, we split sample into two or more homogeneous sets (or sub-populations) based on most significant splitter / differentiator in input variables. In decision tree internal node represents a test on the attribute, branch depicts the outcome and leaf represents decision made after computing attribute.

Decision Tree works in following manner

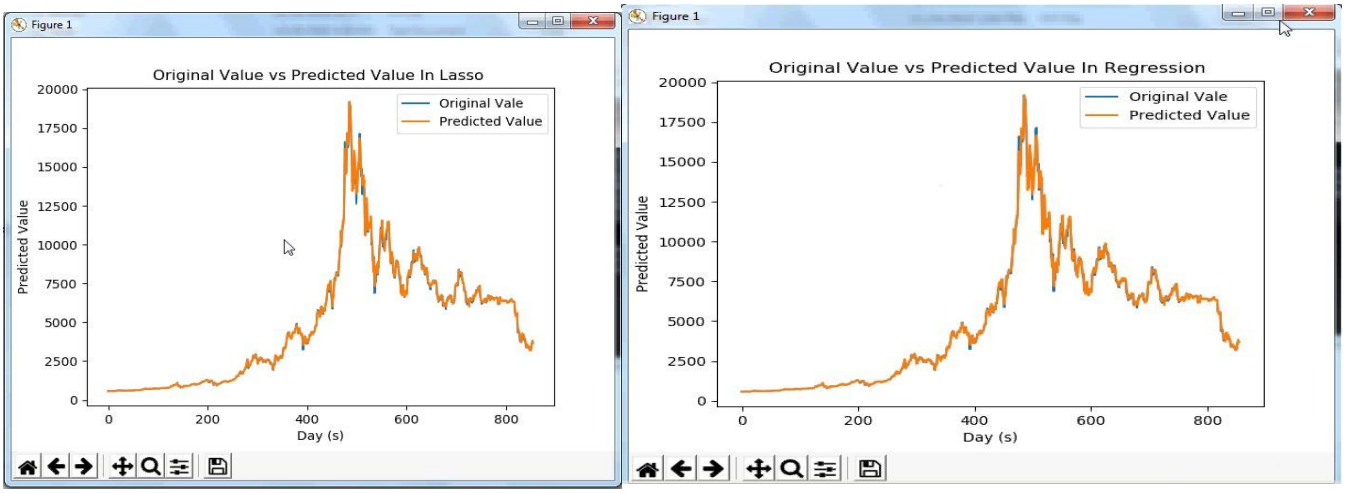
Place the best attribute of the dataset at the root of the tree.

Split the training set into subsets. Subsets should be made in such a way that each subset contains data with the same value for an attribute.

Repeat step 1 and step 2 on each subset until you find leaf nodes in all the branches of the tree. In decision trees, for predicting a class label for a record we start from the root of the tree. Then compare the values of the root attribute with record's attribute. On the basis of comparison,follow the branch corresponding to that value and jump to the next node.

B. Linear regression

In simple, linear regression, predict scores on one variable from the scores on a second variable. The variable that predicted is called the criterion variable and is referred to as Y. The variable base for predictions on is called the predictor variable and is referred to as X. When there is only one predictor variable, the prediction method is called simple regression. In simple linear regression, the topic of this section, the predictions of Y when plotted as a function of X form a straight line.



PRICE FORECAST

Price forecast is done for 5 days using machine learning techniques such as Decision tree and regression. The result is compared with the score value to identify the accuracy value and plotted.

df['label'] = df[forecast_col].shift(-forecast_out)X = np.array(df.drop(['label'],1))

X = preprocessing.scale(X) X_lately = X[-forecast_out:]X = X[:-forecast_out]

The X_lately variable contains the most recent features, which we're going to predict against. Asyou should see so far, defining a classifier, training, and testing was all extremely simple.

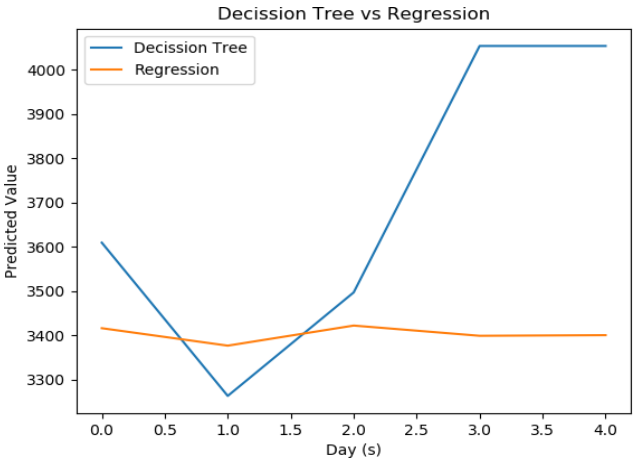
The forecast_set is an array of forecasts, showing that not only could you just seek out a singleprediction, but you can seek out many at once.

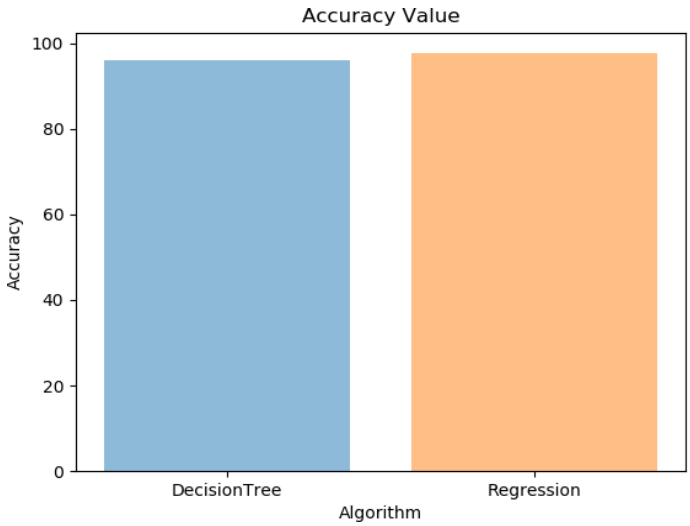
IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.39 ⁜ Vol. 11, Issue 6, June 2022

DOI: 10.17148/IJARCCE.2022.11620





CONCLUSION

Price Forecast for five days

Bitcoin is a successful cryptocurrency, and it has been extensively studied in fields of economics and computer science. In this study, we analyze the time series of Bitcoin price with a Decision Tree and Linear regression models. Also the price forecast for five days is done using Lasso and Linear regression models. After establishing the learning framework and completing the normalization, we intend to use the two methods mentioned above and choose the best method to solve the crypto currency prediction problem. The experimental results show that linear regression outperforms the other by high accuracy on price prediction.

REFERENCES:

[1] https://medium.com/@rohansawant7978/forecasting-of-bitcoin-price-using-machine-learning-deep-learning techniques-93bf662f46ab
[2] Bitcoin (BTC) price prediction 2021-2040 | StormGainhttps://sci-hub.ru/
[3] Bitcoin Price Prediction (swapspace.co)
[4] Bitcoin (BTC) Price Prediction 2022, 2023, 2024 (tradingbeasts.com)
[5] https://www.mdpi.com/2076-3417/10/15/5324https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8022579/
[6] https://coinjournal.net/news/bitcoin-price-prediction-can-bitcoin-breach-resistance/
[7] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8022579/
[8] https://www.researchgate.net/publication/333162007_Bitcoin_Price_Prediction_Using_Machine_Learning