PREDICTING FUTURE CRYPTOCURRENCY PRICES USING MACHINE LEARNING

[#1]**Mr.CHADA SAMPATH REDDY,** *Assistant Professor*

[#2]**Mr.SHANIGARAPU NAVEEN KUMAR,** *Assistant Professor*

Department of Computer Science and Engineering,

SREE CHAITANYA INSTITUTE OF TECHNOLOGICAL SCIENCES, KARIMNAGAR, TS.

ABSTRACT: In this paper, 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 Bitcoin price 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.

Keywords: Lasso Algorithm, Decision Tree, Linear Regression, MSE, RMSE, MAE, RSQUARED**.**

1. INTRODUCTION

The purpose of Bitcoin Prediction is to discover the degree of accuracy that can be achieved in predicting the price of Bitcoin by assessing the accuracy of a number of different machine learning approaches and comparing their results.

Using feature selection derived from a variety of different machine learning approaches, the purpose of this research is to make price predictions for bitcoin. Before constructing prediction models that are able to simultaneously assess volatility and feature series, the objective is to first transform the data from the order book into time-based features, which are referred to as feature series.

Trading techniques, investment options, option value, and indicators of systemic risk are all heavily influenced by volatility, which is a measure of the extent to which prices fluctuate. Bitcoin is now the cryptocurrency with the largest market capitalization, and it is also an essential component of the blockchain-based financial revolution currently underway.

The ability to foresee swings in the price of bitcoin is particularly fascinating to the communities of data mining and machine learning.

Bitcoin is a form of digital currency that is increasingly being used by people all over the world for the purpose of making online investments and sending and receiving money. Due to the fact that Bitcoin is decentralized, no single person or organization can claim ownership over it.

Because they are not affiliated with any particular nation, conducting business with them is quick and easy. Bitcoin exchanges are any of a number of online platforms that make it possible to conduct financial transactions using bitcoin. These make it easier to purchase and exchange Bitcoin in a variety of other currencies.

Mt. Gox is widely recognized as the most successful Bitcoin exchange. A digital wallet, which functions in a way analogous to that of an online bank account, is where bitcoins are kept during their storage. A distributed ledger known as Blockchain is used to record and keep track of timestamps and transaction histories. A block in a blockchain acts as a unique representation of each individual entry that has been added to the chain. A reference to the data block that came before it is included in each and every data block. Encryption is applied to the information that is stored within the blockchain.

During transactions, only the wallet ID and not the user's name are made publically available information. Bitcoin (BTC) is an innovative form of decentralized digital currency that operates freely without the need for a regulating authority located in a centralized location. The processing of payments is carried out by a decentralized network of individuals connected to one another

through the internet.

Users of the Bitcoin network publicly announce new transactions, which are then verified by nodes in the network and appended to the blockchain, which is an openly available decentralized record. When compared to other cryptocurrencies, the market capitalization of Bitcoin is by far the highest.

They are generated as a reward for users who give their computational resources to a competition aimed at validating and documenting transactions on the blockchain.

The competition's goal is to increase the overall security of the blockchain. Bitcoins can be exchanged for a variety of goods and services, in addition to traditional currencies.

At the exchange office, Bitcoins are exchanged for a variety of different currencies, and the order book is used to record "buy" and "sell" orders. Offers that contain the terms "buy" or "bid" show an intention to purchase a given number of Bit coins at a specified price.

On the other hand, offers that contain the terms "sell" or "ask" represent an intention to sell a designated quantity of Bit coins at a set price. Orders are matched up in the order book based on price, which ultimately results in a genuine transaction between the buyer and the seller.

2. LITERATURE SURVEY

P. Ciaian, M. Rajcaniova, and D. Kancs, Appl. Econ., vol. 48, no. 19, pp. 1799–1815, 2016

This is the first article that examines the process of determining the value of Bit Coin. It does so by taking into account its unique qualities as a digital currency, such as its attraction to users and investors, in addition to more conventional elements, such as the market dynamics of supply and demand.

We derive testable hypotheses from the conceptual groundwork laid by Barro's model (1979), which serves as the conceptual underpinning for our work.

We determine that market dynamics and Bit Coin's appeal to users and investors have a significant impact on its price by utilizing time-series analysis techniques on daily data spanning the period from 2009 to 2015; this allows us to establish that the price of Bit Coin is significantly impacted by these factors. However, it is essential to keep in mind that the magnitude of this influence varies throughout the course of time.

Our calculations run counter to earlier findings that show macroeconomic variables would ultimately determine the price of bitcoin over the long term.

S. McNally, Ph.D. dissertation, School Comput., Nat. College Ireland, Dublin, Ireland, 2016.

The purpose of this paper is to evaluate how well the direction of Bitcoin's price expressed in US dollars can be predicted and how accurate those predictions are. For the purpose of acquiring pricing data, the Bitcoin pricing Index is applied. The challenge is tackled by developing a Bayesian optimized recurrent neural network (RNN) and a Long Short Term Memory (LSTM) network. The results of these two networks are mixed in terms of how successfully they tackle the challenge.

The LSTM model obtains the highest performance, with a classification accuracy of 52% and an RMSE (Root Mean Square Error) of 8%. Both of these figures are better than the other models. The ARIMA model, which is widely recognized, is created as an alternative to deep learning approaches for the purpose of time series forecasting.

As was to be expected, the non-linear deep learning algorithms perform better than the ARIMA forecast, which has a poor track record. In the end, the performance of the two deep learning models is compared on both the central processing unit (CPU) and the graphics processing unit (GPU). It can be noticed that the training time on the GPU takes 67.7% more time than the implementation on the CPU.

Madan, S. Saluja, and A. Zhao, Dept. Comput. Sci., Stanford Univ., Stanford, CA, USA, Tech. Rep., 2015

The purpose of this research is to utilize machine learning methods in order to make price predictions for bitcoin. In the beginning, our objective was to comprehend and precisely recognize the daily patterns that occur on the

bitcoin market, all while researching the optimal conditions that are associated with the price of bitcoin.

The data collection contains over 25 elements that were taken from the Bitcoin pricing and payment network. These elements were recorded on a daily basis for a period of five years, making up the entirety of the time period covered by the study. When we used these data, we had an accuracy percentage of 98.7% when attempting to forecast the path that the daily price movement would take. During the second stage of our analysis, we concentrated solely on the data pertaining to the price of Bitcoin. We made use of data that was acquired at intervals of both ten minutes and ten seconds. Because of this, we were able to assess price estimates with varying degrees of granularity and degree of fluctuation.

The challenge of predicting future prices is posed in the form of a binomial classification job, and our objective is to speculate on the path that future price shifts will take. In order to overcome this issue, we make use of a specialized method that integrates random forests with generalized linear models. According to the findings of the research, using time intervals of ten minutes allowed them to accurately forecast future price movement with an accuracy rate of between 50 and 55 percent.

P. Katsiampa, Econ. Lett., vol. 158, pp. 3–6, Sep. 2017

The purpose of this study was to determine which conditional heteroskedasticity model is most suited for conducting an analysis of Bitcoin price data.

It has been discovered that the AR-CGARCH model is the most effective one; this finding places an emphasis on the requirement of incorporating both a short-term and a long-term component of conditional variance.

The end of the paper contains a list of the references that were used.

3. METHODOLOGY

System Architecture

The dataset for bitcoin utilized in this project is obtained from quandl. Initially, the dataset undergoes both testing and training processes. The dataset is classified using three independent methods: the Lasso algorithm, linear regression, and decision tree.

The accuracy of each algorithm is assessed by evaluating the predictions made by the three algorithms. Finally, findings are contrasted to identify the precise cost.

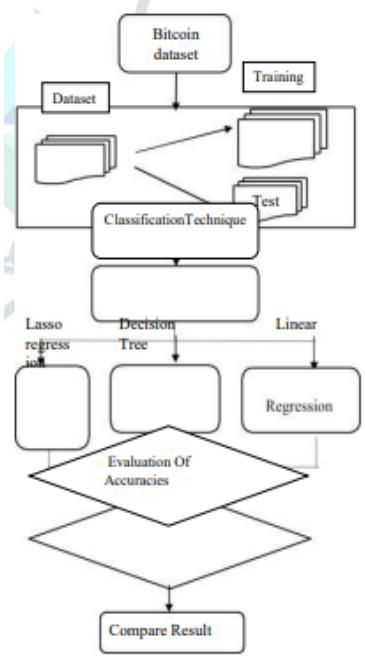

Fig 1. Architecture of the system

Machine Learning Classifiers

Linear regression is a form of regression analysis that involves a single independent variable and assumes a linear relationship between the independent variable (x) and the dependent variable (y).

$$minimize \frac{1}{n}\sum_{i=1}^{n}(pred_i - y_i)^2$$

$$J = \frac{1}{n}\sum_{i=1}^{n}(pred_i - y_i)^2$$

The Lasso, also known as the least absolute shrinkage and selection operator (LASSO), is a regression analysis method that combines variable selection and regularization to enhance the predictability and interpretability of the resulting statistical model.

A decision tree is constructed using an algorithmic approach that determines the division of a dataset

into sections depending on multiple criteria. This method is widely used and highly efficient for supervised learning.

Decision trees, a non-parametric supervised learning technique, are utilized for classification and regression tasks.

4. RESULTS AND DISCUSSION

A value was computed by the test set of the evaluation experiment. The Python software platform is utilized to carry out the experiment. The test system is equipped with 3GB of RAM and a 2.4GHz Intel Core CPU for personal laptops. Machine learning is commonly employed to evaluate the proposed methodology. From the simulation of the experiment's findings, it can be inferred that this method is robust against diverse Bitcoin data set images.

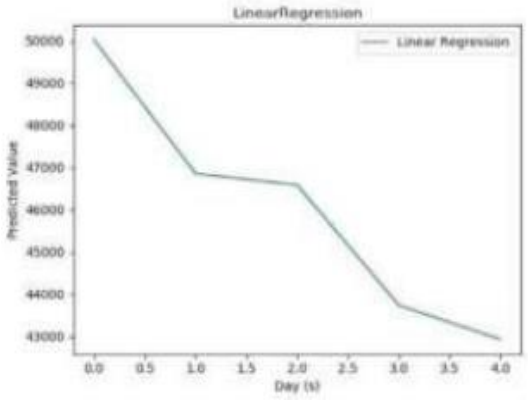

Fig 2. Cost forecasting using linear regression

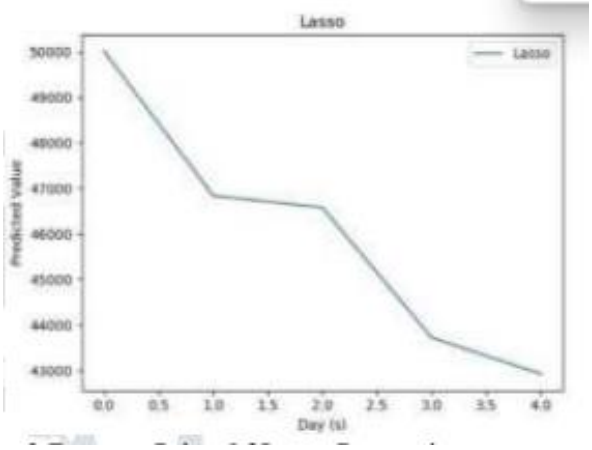

Fig 3. Lasso regression for price forecasting

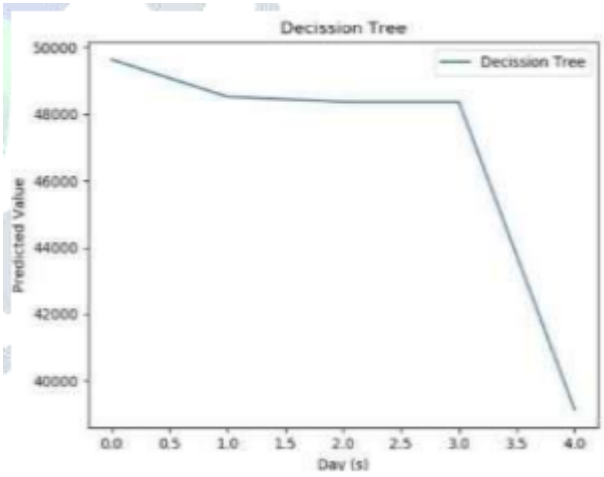

Fig 4. Price prediction using a decision tree algorithm.

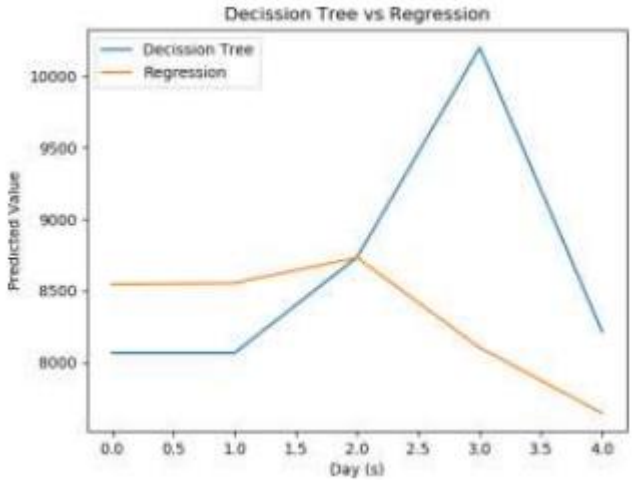

Fig 5. Comparison between Decision Tree and Linear Regression Model

Table 1

	Lasso	Regression
MSE	126154.7505	105644.6069
MAE	137.1256	130.8279
R-Squared	0.9991	0.9993
RMSE	355.1827	325.0301
Accuracy	99.9176	99.9310

Table 2

	Decision Tree	Linear Regression
Accuracy	98.3895	99.0558

Table 1 compares the accuracy, mean squared error (MSE), mean absolute error (MAE), R-squared, and root mean squared error (RMSE) of

the projected values of the Bitcoin data set created using the Lasso and Regression algorithms.

Table 2 presents a comparison of the predicted accuracy of the Bitcoin data set using the Decision Tree and Regression approaches.

5. CONCLUSION

Due to the high level of instability in Bitcoin, segregated witnesses and distributed immutable ledgers are employed to gather up-to-date information and integrate it into regression models. Based on the model analysis, the Lasso regression model and linear regression model have prediction accuracies of 98.6% and 98.7% respectively. The decision trees have a prediction accuracy of 97.5%, but linear regression achieves a slightly higher accuracy of 97.7%. To optimize the performance of all models, it is imperative to regularly update datasets.

REFERENCES

1. D. Shah and K. Zhang, “Bayesian regression and Bitcoin,” in 52nd Annual Allerton Conference on Communication, Control, and Computing (Allerton), 2015, pp. 409-415.

2. Huisu Jang and Jaewook Lee, “An Empirical Study on Modelling and Prediction of Bitcoin Prices with Bayesian Neural Networks based on Blockchain Information,” in IEEE Early Access Articles, 2017, vol. 99, pp. 1-1.

3. F. Andrade de Oliveira, L. Enrique ZÃ¡rate and M. de Azevedo Reis; C. Neri Nobre, “The use of artificial neural networks in the analysis and prediction of stock prices,” in IEEE International Conference on Systems, Man, and Cybernetics, 2011, pp. 2151-2155.

4. M. Daniela and A. BUTOI, “Data mining on Romanian stock market using neural networks for price prediction”. informatica Economica, 17,2013

5. Thearasak Phaladi, Thanisa Numnonda “Machine Learning Models Comparison for Bitcoin Price Prediction” 2019IEEE.