IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 8, Issue 8, August 2019

Crypto-Currency Price Prediction using Machine Learning

Kathyayini R S[1], D G Jyothi [2]

M. Tech Student, Computer Science and Engineering, BIT College, Bengaluru, India[1]

Associate Professor, Computer Science and Engineering, BIT College, Bengaluru, India[2]

Abstract: Crypto-Currency price prediction attempt to predict the Bitcoin price accurately taking into observation various factors that affect the Bitcoin value. From this investigation, the aim is to appreciate 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 Fetch the trading exchange real time data rate in USD, from September 13th 2011 to current day's previous day is considered. For the second phase of this investigation, using the available information, the sign of the daily price change with highest possible accuracy will be predicted.

Keywords: Bitcoin, USD, Real Time Data Rate, Predict

I. INTRODUCTION

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. Investment can be done through various marketplaces known as "bitcoin exchanges". These allow people to sell/buy Bitcoin using different currencies. Bitcoin 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 Block chain. Each record in a block chain is called a block. Each block contains a pointer to a previous block of data. The data on block chain is encrypted. During transactions the user's name is not revealed, but only their wallet ID is made public.

II. LITERATURE SURVEY

In [1] the author analyzing the sentiment in Twitter and to find the relation between them. The tweets of Bitcoin collected from different news account sources are classified to positive or negative sentiments. The obtained percentage of positive and negative tweets are feed to RNN model along with historical price to predict the new price for next time frame. The accuracy for sentiment classification of tweets in two class positive and negative is found to be 81.39 % and the overall price prediction accuracy using RNN is found to be77.62%.

In [2] the author use of an Artificial Neural Network (ANN) ensemble approach called Genetic Algorithm based Selective Neural Network Ensemble (GASEN). The ensemble will be used to solve a binary classification problem. To better understand and evaluate its effectiveness, back testing was done to see how a trading strategy based on the results of the ensemble can compare against a "previous day trend following" trading strategy as well as a trading strategy that follows the single, best MLP model in the ensemble.

In [3] the author specifically, we aim at investigating the couplings among the length of training period, the choice of ARIMA parameters (*p; q; d*), and the length of time window that the prediction is carried out over it, i.e. the bitcoin price for the day after the window is predicted.

In[4] the author use 1-minute interval trading data on the Bitcoin exchange website named bitstamp from January 1, 2012 to January 8, 2018, some different regression models with scikit learn & Keras libraries had experimented. The best results showed that the Mean Squared Error (MSE) was as low as0.00002 & the R-Square (R2) was as high as 99.2%.

In [5] the author implemented of a Bayesian optimized Recurrent Neural Network (RNN)and a Long Short Term Memory (LSTM) network. The LSTM achieves the highest classification accuracy of 52% and a RMSE of 8%. The popular ARIMA model for time series forecasting is implemented as a comparison to the deep learning models. As expected, the non-linear deep learning methods outperform the ARIMA forecast which performs poorly. Finally, both deep learning models are benchmarked on both a GPU and a CPU with the training time on the GPU outperforming the CPU implementation by 67.7%.

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 8, Issue 8, August 2019

In [6] the author conduct the empirical study that compares the Bayesian neural network with other linear and non-linear benchmark models on modeling and predicting the Bitcoin process. Our empirical studies show that BNN performs well in predicting Bitcoin price time series and explaining the high volatility of the recent Bitcoin price.

In [7] the author collected data set consists of over 25 features relating to the Bitcoin price and payment network over the course of five years, recorded daily were able to predict the sign of the daily price change with an accuracy of 98.7%. For the second phase of our investigation, we focused on the Bitcoin price data alone and leveraged data at 10-minute and 10-second interval time points, as we saw an opportunity to evaluate price predictions at varying levels of granularity and noisiness are modelling and results had 50-55% accuracy in predicting the sign of future price change using 10-minute time intervals.

III. PROPOSED SYSTEM

Volatility as a measure of price fluctuations has a significant impact on trade strategies and investment decision as well as on option pricing and measure of systematic risk. Therefore, it is of great interest to machine learning community to be able to predict Bitcoin price fluctuations.

To acquire time series data recorded daily for 9 certain time period at different time instances, select the parameter from available features that are fed to the model compared with the models and final prediction

- Bitcoin Dataset
- Extraction Dataset
- Data Preparation

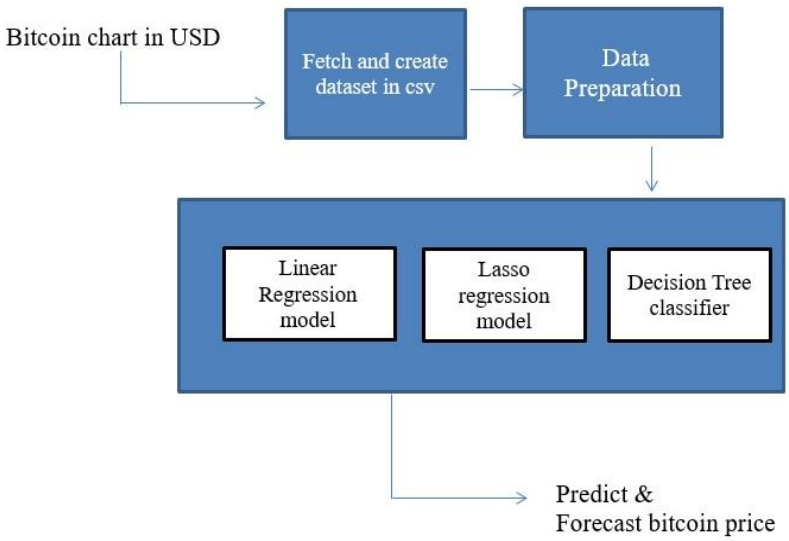


Fig. 1 Architectural diagram

A. Linear Regression –

Linear regression is a type of regression analysis where the number of independent variables is one and there is a linear relationship between the independent(x) and dependent(y) variable.

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

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

B. Lasso Regression –

Lasso (least absolute shrinkage and selection operator; also Lasso or LASSO) is a regression analysis method that performs both variable selection and regularization in order to enhance the prediction accuracy and interpretability of the statistical model it produces.

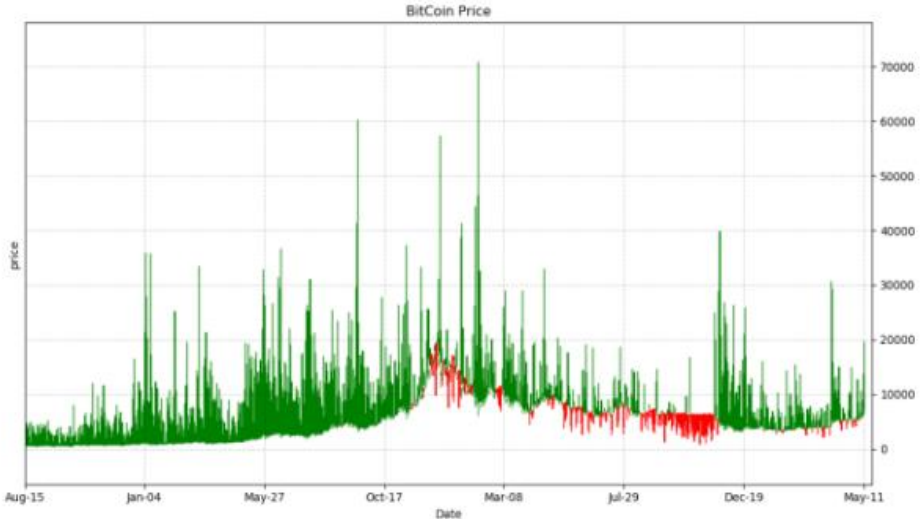
C. Decision Tree –

Decision trees are constructed via an algorithmic approach that identifies ways to split a data set based on different conditions. It is one of the most widely used and practical methods for supervised learning. Decision Trees are a non-parametric supervised learning method used for both classification and regression tasks.

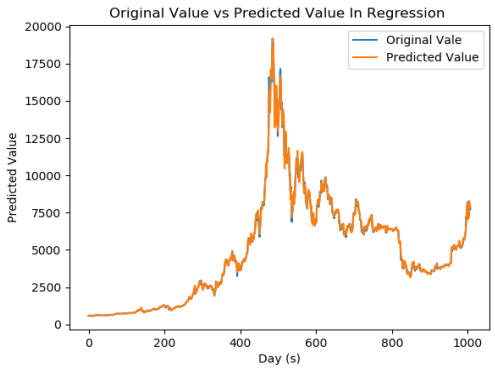
IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 8, Issue 8, August 2019

IV. EXPERIMENT AND RESULT

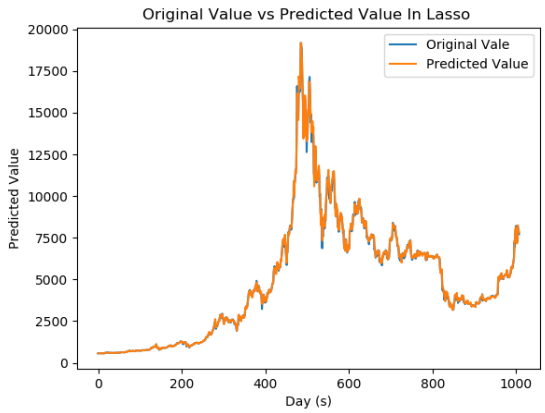
The test set for this evaluation experiment calculated value. Python software platform is use to perform the experiment. The PC for experiment is equipped with an Intel Core 2.4GHz Personal laptop and 3GB memory. The proposed scheme is tested using ordinarily machine learning. From the simulation of the experiment results, we can draw to the conclusion that this method is robust to many kinds Bitcoin data set images.



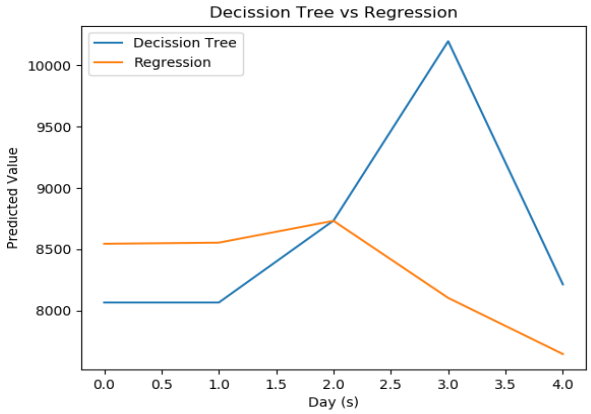
(a)



(b)



(c)



(d)

Fig 2. (a) Graph Generated by data which having collected from Quandl website
(b) Represents the original value and predicted value in Regression
(c) Represents the original value and predicted value in Lasso (d) Represents the decision tree vs linear regression

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 8, Issue 8, August 2019

Table I Experiment & Result

(a)

	Lasso	Regression
MSE	17836.11	17049.54
MAE	68.77	66.04
R-Squared	0.99	0.99
RMSE	133.55	130.57
Accuracy	99.86	99.87

(b)

	Decision Tree	Regression
Accuracy	97.53	97.76

Table 1 shows the Predicted values of Bitcoin data set using Lasso and Regression algorithm MSE MAE R-Squared RMSE & Accuracy are compared each other.
Table 2 shows forecasted values of Bitcoin data set using Decision tree and Regression algorithm finding the accuracy and compared each other.

V. CONCLUSION

As Bitcoin is very fluctuating in nature, segregated witness and distributed immutable ledger as fetching the real-time data and put into the regression models. According to the model analysis, the Lasso regression model provides the accuracy is 98.6% and linear regression model accuracy is 98.7% for the prediction. The forecasting using the decision tree accuracy is 97.5% and by the linear regression will be the 97.7%. Therefore, to the best results of all models, datasets should be always updated.

REFERENCES

[1]. Dibakar Raj Pant, Prasanga Neupane, Anuj Poudel, Anup Kumar Pokhrel, Bishnu Kumar "Recurrent Neural Network Based Bitcoin Price Prediction by Twitter Sentiment Analysis" Lama 2018IEEE.
[2]. Edwin Sin, Lipo Wang "Bitcoin Price Prediction Using Ensembles of Neural Networks" 2017IEEE.
[3]. Amin Azari "Bitcoin Price Prediction an ARIMA Approach" 2018IEEE.
[4]. Thearasak Phaladi, Thanisa Numnonda "Machine Learning Models Comparison for Bitcoin Price Prediction" 2019IEEE.
[5]. Sean Mc Nally, Jason Roche, Simen Caton "Predicting the Price of Bitcoin Using Machine Learning" 2018IEEE.
[6]. Huisu Jang, Jaewook Lee "An Empirical Study onJ Modeling and Prediction of Bitcoin Prices with Bayesian Neural Networks Based on Blockchain Information" 2017IEEE.
[7]. Isaac Madan, Saluja, Aojia Zhao "Automated Bitcoin Trading via Machine Learning Algorithms" 2017 IEEE