Crypto Currency Price Prediction Using Machine Learning Techniques

Gopikrishnan. A[1], Jone Solomon. D[1], Kaviyarasan. N[1], Vignesh. T[1], Harshavardhini. S[1], Dr. S. Jothi Lakshmi[2]

[1]UG Scholar, Department of Computer Science and Engineering, Akshaya College of Engineering and Technology, Coimbatore, Tamil Nadu, India

[2]Assosiate Professor, Department of Computer Science and Engineering, Akshaya College of Engineering and Technology, Coimbatore, Tamil Nadu, India

ABSTRACT

Crypto-currency such as Bitcoin is more popular these days among investors. In the proposed work, it is studied to forecast the Bitcoin price precisely considering different parameters that influence the Bitcoin price. This study first handles, it is identified the price trend on day by day changes in the Bitcoin price while it gives knowledge about Bitcoin price trends. The dataset till current date is taken with open, high, low and close price details of Bitcoin value. Exploiting the dataset machine learning module is introduced for prediction of price values. The aim of this work is to derive the accuracy of Bitcoin prediction using different machine learning algorithm and compare their accuracy. Experiment results are compared for Random Forest and regression model.

Keywords: Machine Learning, Bitcoin, Prediction, Crypto Currency.

I. INTRODUCTION

Bitcoin is an effective crypto currency brought into the monetary market dependent on its one of a kind convention and Nakamoto's structure. Bitcoin try to achieve decentralization in currency market. Investors in Bitcoin market establish trust connections through the development of Blockchain using cryptography strategies.

Bitcoin now-a-days gaining more interest due to innovations in Blockchain and machine learning.

Highly volatile significantly affects exchange methodologies for crypto-currency market. Bitcoin is the major crypto-currency is still facing the volatility problem. Thus, there should be a significant research efforts needed for an efficient way of price prediction for crypto-currency through machine learning techniques is inevitable.

The objectives of this study to the price prediction of bitcoin by feature selection of different machine learning techniques. Volatility as a proportion of value fluctuations, it significantly affects exchange

methodologies and investment choices just as on alternative estimating and proportions of fundamental risk.

Bitcoin, as a pioneer in the blockchain money, it assumes a predominant job in an entire cryptographic money showcase capitalization. Subsequently, it is of extraordinary interest are growing these days on data mining and machine learning network to most likely predict Bitcoin value variances While senders of conventional electronic installments are generally distinguished, investors of bitcoin work in anonymous. No central authority is available for bitcoin, investors don't have to recognize themselves when sending bitcoin to another client.

Investors are identified by the address of wallet. It is mandatory for Exchanges to check identity of the clients handled by them, they are not permitted to make buy or sell trade without checking their identity. Bitcoin is more secure for investors.

The objective of the proposed study is for price prediction of bitcoin by feature selection of different machine learning techniques. Intuitively, idea is to first transform order book data into features over time, referred as feature series and then to develop prediction models to consume volatility and feature series simultaneously.

II. LITERATURE REVIEW

The McKinsey Global Institute estimates that applying machine learning techniques to better inform decision making could generate up to $100 billion in value based on optimized innovation, enhanced efficiency of clinical trials and the creation of various novel tools for physicians, insurers and consumers. Computers and Robots cannot replace doctors or nurses, however the use of life-saving technology (machine learning) can definitely transform healthcare domain. When we talk about efficiency of machine learning, more data produces effective results – and the healthcare industry is residing on a data goldmine.

Personalized treatment has great potential for growth in future, and machine learning could play a vital role in finding what kind of genetic makers and genes respond to a particular treatment or medication. Personalized medication or treatment based on individual health records paired with analytics is a hot research area as it provides better disease assessment. In future, increased usage of sensor integrated devices and mobile apps with sophisticated remote monitoring and health-measurement capabilities, there would be another data deluge that could be used for treatment efficacy. Personalized treatment facilitates health optimization and also reduces overall healthcare costs.

The main idea of principal component analysis (PCA) is to reduce the dimensionality of a data set consisting of many variables correlated with each other, either heavily or lightly, while retaining the variation present in the dataset, up to the maximum extent. The same is done by transforming the variables to a new set of variables, which are known as the principal components (or simply, the PCs) and are orthogonal, ordered such that the retention of variation present in the original variables decreases as we move down in the order. So, in this way, the 1st principal component retains maximum variation that was present in the original components. The principal components are the eigenvectors of a covariance matrix, and hence they are orthogonal.

Importantly, the dataset on which PCA technique is to be used must be scaled. The results are also sensitive to the relative scaling. As a layman, it is a method of summarizing data. Imagine some wine bottles on a dining table. Each wine is described by its attributes like colour, strength, age, etc. But redundancy will arise because many of them will measure related properties. So what PCA will do in this case is summarize each wine in the stock with less characteristics.

Cryptocurrency Price Prediction using Time Series Forecasting (ARIMA)

Cryptocurrency is a tangible or digital currency protected with the help of Cryptography, making it almost impossible to counterfeit or double. Many cryptocurrency networks are categorized primarily based on blockchain technology. The present socio-economic situation also creates an environment for people to hold less cash and remain marginalized by the market trends. The objective of the project is to build a profitable Machine Learning prediction model. We begin by collecting the data from Yahoo Finance website using inbuilt python libraries. Our objective was to perform price prediction of various Cryptocurrencies using Machine Learning, and we have implemented the Autoregressive Integrated Moving Average (ARIMA) Model.

Crypto-Currency price prediction using Decision Tree and Regression techniques

Crypto-currency such as Bitcoin is more popular these days among investors. In the proposed work, it is studied to forecast the Bitcoin price precisely considering different parameters that influence the Bitcoin price. This study first handles, it is identified the price trend on day by day changes in the Bitcoin price while it gives knowledge about Bitcoin price trends. The dataset till current date is taken with open, high, low and close price details of Bitcoin value. Exploiting the dataset machine learning module is introduced for prediction of price values.

Cryptocurrency Price Prediction by Using Hybrid Machine Learning and Beetle Antennae Search Approach

Cryptocurrencies are defined as digital mediums of exchange, that use strong cryptography for securing the transactions and verifying the ownership of the coins. Blockchain operates in the background to guarantee the security, transparency and traceability of the transactions. Consequently, cryptocurrencies became more and more popular and established their considerable presence in financial sector. However, one of the major drawbacks in the cryptocurrency market is the unreliability and unpredictability of their values, that poses a major risk for any kind of investment.

III. PROPOSED SYSTEM

In this project , the proposed system is based on Random forest Regression and liner regression with Machine learning.

Advantages

- Random Forest is based on the bagging algorithm and uses Ensemble Learning technique. It creates as many trees on the subset of the data and combines the output of all the trees. In this way it reduces overfitting problem in decision trees and also reduces the variance and therefore improves the accuracy.
- Random Forest can be used to solve both classification as well as regression problems.
- Random Forest works well with both categorical and continuous variables.
- Random Forest can automatically handle missing values.
- No feature scaling required: No feature scaling (standardization and normalization) required in case of Random Forest as it uses rule based approach instead of distance calculation.
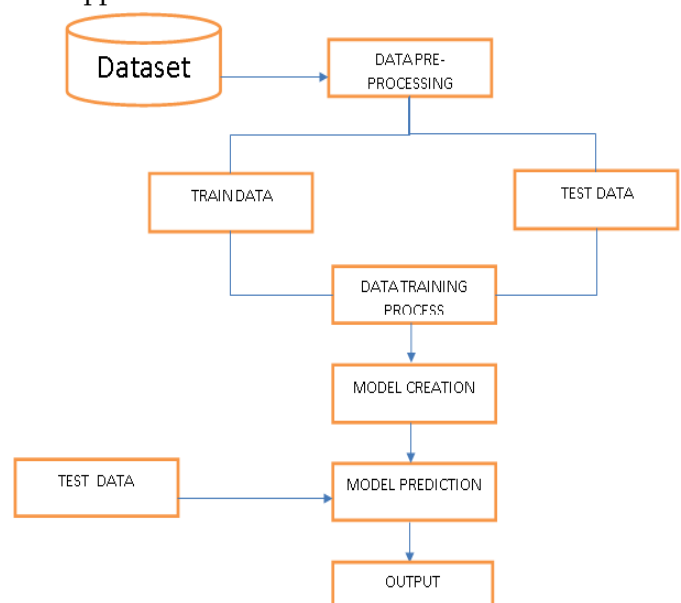

Figure 1. Block diagram

➢ **Data collection**

There are five components of the data collected, the following being

- Cost: The average daily digital price
- Open: The starting price for the digital currency of the day
- Close: Regular currency digital closing price
- High: The highest per-day digital price
- Low: the lowest daily digital price
- Volume: The digital sum of money sold every day

➢ **Data Pre-Processing**

Pre-processing refers to the transformations applied to our data before providing the data to the algorithm. Data Preprocessing technique is used to convert the raw data into an understandable data set. In other words, whenever the information is gathered from various sources it is collected in raw format that isn't possible for the analysis.

➢ **Training data and Test data**

- For choosing a model we split our dataset into train and test
- Here data's are split into 3:1 ratio that means
- Training data having 70 percent and testing data having 30 percent
- In this split process preforming based on train_test_split model
- After splitting we get xtrainxtest and ytrainytest

➢ **Model Creation**

- Contextualise machine learning in your organisation.
- Explore the data and choose the type of algorithm.
- Prepare and clean the dataset.
- Split the prepared dataset and perform cross validation.
- Perform machine learning optimisation.
- Deploy the model.

➢ **Model Prediction**

Predictive modeling is a statistical technique using machine learning and data mining to predict and forecast likely future outcomes with the aid of historical and existing data. It works by analyzing

current and historical data and projecting what it learns on a model generated to forecast likely outcomes.

IV. CONCLUSION

The Bitcoin is a booming crypto-currency market, and various researches have been studied in fields of economics and price prediction. In our proposed work, Bitcoin dataset is considered from 2011 to till date price and applied machine learning models such as Random forest and regression models. Also the price forecast for five days is done using Random forest and regression . The proposed learning method suggest the best algorithm to choose and adopt for crypto currency prediction problem. The experimental study results show that linear regression outperforms the other by high accuracy on price prediction.

V. REFERENCES

[1]. Chan, S., Chu, J., Nadarajah, S. and Osterrieder, J., 2017. A statistical analysis o fcryptocurrencies. Journal o f Risk and Financial Management, 10(2), p. 12.

[2]. Sovetov, Y., 2018. Factors influencing cryptocurrency prices: Evidence from bitcoin, ethereum, dash, litecoin, and monero. Journal o f Economics and Financial Analysis, 2(2), pp.1-27.

[3]. Miglani, A. and Kumar, N., 2019. Deep learning models for traffic flow prediction in autonomous vehicles: A review, solutions, and challenges. Vehicular Communications, 20, p.100184.

[4]. Makkar, A. and Kumar, N., 2020. An efficient deep learning-based scheme for web spam detection in the IoT environment. Future Generation Computer Systems, 108, pp.467-487.

[5]. Teker, D., Teker, S. and Ozyesil, M., 2019. Determinants of Cryptocurrency Price Movements.

[6]. Peng, Y., Albuquerque, P.H.M., de Sa, J.M.C., Padula, A.J.A. and Montenegro, M.R., 2018. The

best o f two worlds: Forecasting high frequency volatility for cryptocurrencies and traditional currencies with Support Vector Regression. Expert Systems with Applications, 97, pp.177-192.

[7]. Song, H., Sui, S., Han, Q., Zhang, H. and Yang, Z., 2020. Autoregressive integrated moving average model-based secure data aggregation for wireless sensor networks. International Journal o f Distributed Sensor Networks, 16(3), p.1550147720912958.

[8]. Garg, S., 2018, December. Autoregressive integrated moving average model based prediction o f bitcoin close price. In 2018 International Conference on Smart Systems and Inventive Technology (ICCSIT) (pp. 473-478). IEEE.

[9]. Roy, S., Nanjiba, S. and Chakrabarty, A., 2018, December. Bitcoin price forecasting using time series analysis. In 2018 21st International Conference o f Computer and Information Technology (ICCIT) (pp. 1-5). IEEE.