CRYPTOCURRENCY MARKET PRICE PREDICTION USING MACHINE LEARNING

Arul Dalton, G., Anand, T and P. Jino

Department of Computer Science and Engineering
Madha Engineering College, Kundrathur, Chennai-600069.

ABSTRACT

The cryptocurrency trend is gearing up among investors these days as they want to escape the boring norms of fiat money. The features like secured transactions, no intermediaries, and high speed have paved the way for the growth of cryptocurrencies across the globe. There may be a sudden rise or fall in the value and it is difficult to predict them. This might be a huge challenge for the investors as they can face a huge loss. In the proposed work, the price of the various cryptocurrencies like Bitcoin, Ethereum, Litecoin, Binance Coin, and Maker is forecasted by considering the different parameters that influence the price. The dataset is collected till the current date with the open, high, low, and close prices of the cryptocurrencies. For the price prediction, different machine learning algorithms like linear regression, support vector regression, SGD regression, lasso regression, XG boost regression, ridge regression, and random forest regression are used and compared their performance. These regression algorithms are chosen because of their predictive analysis and by comparing them we can find the best fit for the data. The prediction is improved by using the regression algorithms as they are great for forecasting because of their exploratory nature between the data points which will, in turn, predict both long-term and short-term values with fewer errors. This helps in predicting the price of the cryptocurrency more precisely and accurately by which the investors and beginners can be easily able to choose and invest in a way more profitable.

Keywords: Cryptocurrency, secured transactions, regression algorithms and linear regression.

INTRODUCTION

Cryptocurrency has become a game-changer because of its vast advantages like freedom of payment, anonymity, high security, speed, and mainly no third-party involvement. The snag in the current system of using the currency is that every single transaction is being recorded and it

is highly vulnerable (Mirtaheri *et al.,* 2021). But in the case of cryptocurrency, as it uses blockchain technology, there is no single factor that makes it vulnerable in any way (Rathan *et al*., 2019). This attracts the investors to invest as it also helps in gaining an enormous amount of profit. These reasons have significantly gained its growth over the years and more people have started to invest in this. So, beginners and investors might need a foretelling system to predict the market price which will help everyone in spending their money more optimally. This system is proposed to predict the price of the cryptocurrency more precisely and minutely (Wimalagunaratne and Poravi 2018). The input of the system is the cryptocurrency name, the number of days to be predicted, open, close, high, and low prices while the output is the open price prediction for the upcoming days. Initially, exploratory data analysis is made separately for each cryptocurrency to understand its nature and characteristics. Then feature engineering, scaling is made on the collected dataset to increase the efficiency and accuracy (Inamdar 2019). Then the model is trained using the different regression algorithms that we have taken into account. Separate modules are built for each parameter like open, close, high, and low prices in order to loop over to predict the target open price for 'n' number of days. The model is validated using the test dataset and the most suitable algorithm with high accuracy is identified by comparing with the different algorithms using the performance metrics. The user interface is built with the help of flask to make the system more accessible and user-friendly. This will, in turn, produce a flawless system where one can predict the price of the cryptocurrency without any obstacles or barriers (Wu *et al*., 2019).

PROBLEM DESCRIPTION

Investing in the right cryptocurrency at the right time is very difficult. And if you are a beginner in the crypto world, it will be a huge challenge to invest wisely. Nowadays, there is a wide range of unique cryptocurrencies each has its own nature and characteristics over time (Roy *et al*., 2018). The price differs from time to time. So, we need a system that can predict the price of cryptocurrency more accurately and precisely. By using different regression algorithms, we will find the best fit for the proposed system. So, this will in turn produce a foretelling system that helps in predicting the price of the cryptocurrency which results in more profit (Xiaomeng 2020).

OBJECTIVES

- To predict the price of the cryptocurrency and help investors to make better investments.
- To analyze and study the hidden patterns and relationship between the various parameters like high, low, open and close values to predict the price of the cryptocurrency more accurately.
- To find the best among the different algorithms which give greater accuracy in predicting the market price of the cryptocurrency.

SYSTEM DESIGN

The system design plays an important role in providing the solutions to the issues raised in the model. It mainly deals with design and planning. The design phase is the most important as it is the primary factor affecting the quality of the system. This chapter depicts the flow of the operations involved in the system (Oggier *et al*., 2018). The proposed system is explained using the system architecture, functional architecture and modules using various charts and diagrams. It clearly explains the entire program flow and its functions.

SYSTEM ARCHITECTURE

Initially pre-processing is done on the collected dataset. In the pre-processing, various actions like feature selection, correlation, data cleaning and data encoding is performed on the dataset to make it more effective and error-free. And upon the cleaned dataset, the various regression algorithms are applied to see which algorithm is more suitable and gives results more accurately (Liang *et al*., 2018). To evaluate this various performance metrics are used to test the performance of the algorithm. Lastly, the machine learning model is built with the most suitable algorithm among the others to complete the system (Pavlyshenko 2019). It is dumped inside the web server, from the user's request, the results are predicted and given through the web application.
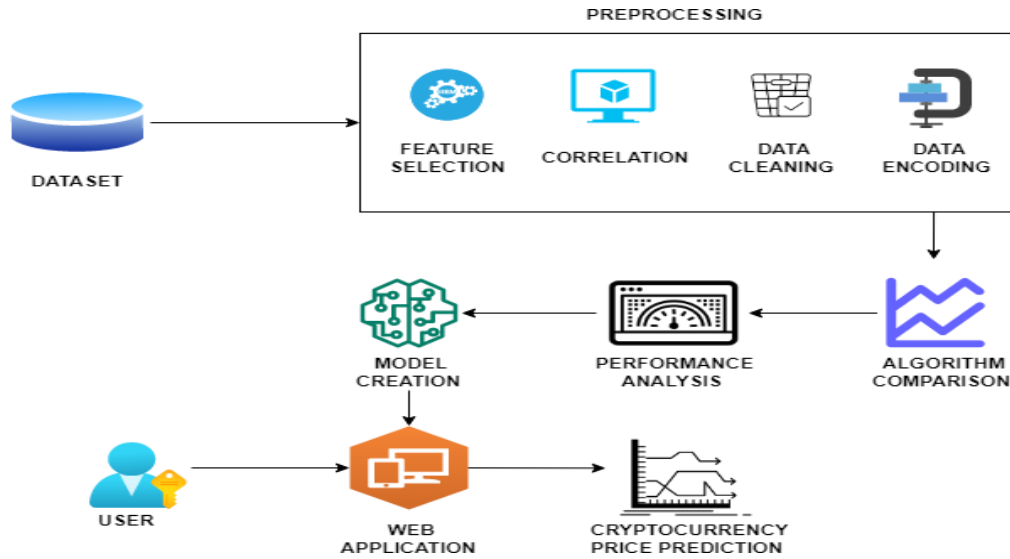
SYSTEM ARCHICTECTURE FUNCTIONAL ARCHITECTURE

The following diagram is the functional architecture of the system. It tells us about the complete flow of the system with the utmost detail. Feature engineering is done on the collected dataset according to the model development. The dataset is separated into a training dataset and a testing dataset in the ratio of 8:2 respectively. Using the training dataset, the final model is built by applying the different machine learning algorithms. And it is validated and tested using the testing dataset. And finally, with the trained model, the price of the selected cryptocurrency is predicted and displayed to the user.
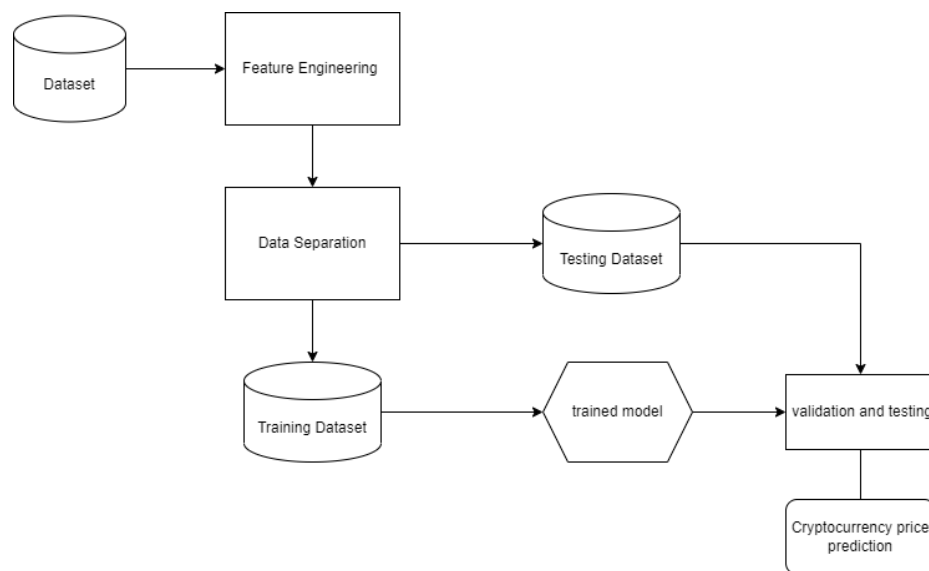
Exploratory Data Analysis: The ED Analysis was made to translate the information into a visual context to make it easier to understand the data to pull more insights. Numerous patterns and outliers are identified in this module. Day-wise plotting is made to understand the movements of the prices over the past years. The high and low prices are compared in order to understand their movements and nature. Box plot is used to understand the outliers and the moving average for 50 days is calculated for the open price to understand the movement and pattern for a particular period of time. And the past patterns were analyzed using the candlestick charts to understand the movement of the prices as it helps in showing the four parameters (open, high, low and close) for the time period, the user has specified.

MODULE IMPLEMENTATION

Removal of null values: In this module, the null values are removed from the dataset to prevent any future errors. In data wrangling, removing null values is one of the important steps, as these may reduce the performance of the algorithm (Fig 1).

Data correlation: Correlation for each and every cryptocurrency is made in order to understand how one parameter is related to another and this will help in forecasting the target variable for the model (Fig 2 &3).

Day-wise close price plotting: In this, the close prices for each and every individual day are plotted in order to understand the movement of the price pattern and movement (Fig 4).

	Number of Mising values	Percentages of Missing Values
Unnamed: 0	0	0.000000
Date	0	0.000000
Open	3	0.031496
High	3	0.031496
Low	3	0.031496
Close	3	0.031496
Adj Close	3	0.031496
Volume	3	0.031496
Cripto	0	0.000000

Fig 1: Identification of null values

	Unnamed: 0	Open	High	Low	Close	Adj Close	Volume
Unnamed: 0	1.000000	0.571611	0.571401	0.572203	0.571806	0.571806	0.583141
Open	0.571611	1.000000	0.999639	0.999298	0.999106	0.999106	0.707236
High	0.571401	0.999639	1.000000	0.999253	0.999603	0.999603	0.709782
Low	0.572203	0.999298	0.999253	1.000000	0.999535	0.999535	0.701759
Close	0.571806	0.999106	0.999603	0.999535	1.000000	1.000000	0.706476
Adj Close	0.571806	0.999106	0.999603	0.999535	1.000000	1.000000	0.706476
Volume	0.583141	0.707236	0.709782	0.701759	0.706476	0.706476	1.000000

Fig 2: Correlation of attributes

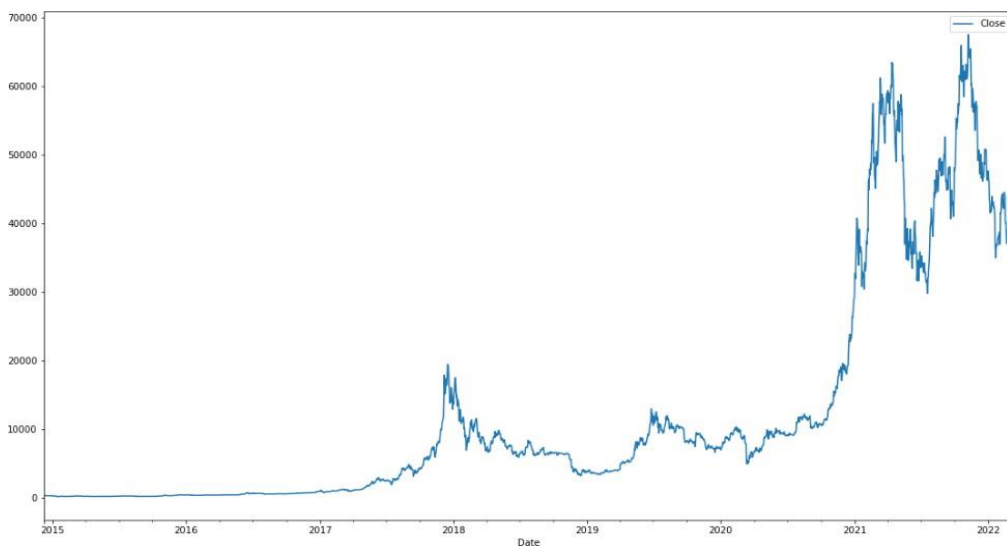|	Open	High	Low	Close
Open	1.000000	0.999639	0.999298	0.999106
High	0.999639	1.000000	0.999253	0.999603
Low	0.999298	0.999253	1.000000	0.999535
Close	0.999106	0.999603	0.999535	1.000000

Fig 3: Final parameters

Fig 4: Day wise plotting

CONCLUSION

The market of the cryptocurrency will go double its value as estimated by many analysts around the world. The main challenge is the instability of the price as it changes rapidly. In this situation, we need a system to help the investors and beginners to guide them.   So, this system would be a great foretelling system to predict the price more clearly and accurately to invest more ideally. The dependency between the parameters has been studied carefully to predict the exact price.   From analyzing with the help of several performance metrics, it is concluded that Linear Regression is the most suitable among the others. Followed by it, Random Forest Regression produces similarly fewer errors. The trained model is saved and deployed using the Flask framework in which the price of the cryptocurrency is predicted successfully

REFERENCES

Inamdar A, A. Bhagtani, S. Bhatt and P. M. Shetty, "Predicting Cryptocurrency Value using Sentiment Analysis," 2019 International Conference on Intelligent Computing and Control Systems (ICCS), 201: 932-934.

Liang J, L. Li, W. Chen and D. Zeng, "Towards an Understanding of Cryptocurrency: A Comparative Analysis of Cryptocurrency, Foreign Exchange, and Stock," 2019 IEEE International Conference on Intelligence and Security Informatics (ISI), 2019, 137-139.

Mirtaheri M, S. Abu-El-Haija, F. Morstatter, G. V. Steeg and A. Galstyan, "Identifying and Analyzing Cryptocurrency Manipulations in Social Media," in IEEE Transactions on Computational Social Systems, 2021, 8 (3): 607-617.

Oggier F, S. Phetsouvanh and A. Datta, "BiVA: Bitcoin Network Visualization & Analysis," 2018 IEEE International Conference on Data Mining Workshops (ICDMW), 2018, 1469-1474.

Pavlyshenko B M, "Bitcoin Price Predictive Modeling Using Expert Correction," 2019 XI International Scientific and Practical Conference on Electronics and Information Technologies (ELIT), 2019, 163-167.

Rathan K, S. V. Sai and T. S. Manikanta, "Crypto-Currency price prediction using Decision Tree and Regression techniques," 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), 2019, 190-194.

Roy S, S. Nanjiba and A. Chakrabarty, "Bitcoin Price Forecasting Using Time Series Analysis,"

2018 21st International Conference of Computer and Information Technology (ICCIT), 2018, 1-5.

Wimalagunaratne M and G. Poravi, "A Predictive Model for the Global Cryptocurrency Market: A Holistic Approach to Predicting Cryptocurrency Prices," 2018, 8th International Conference on Intelligent Systems, Modelling and Simulation (ISMS), 2018, 78-83.

Wu Y, A. Luo and D. Xu, "Forensic Analysis of Bitcoin Transactions," 2019 IEEE International Conference on Intelligence and Security Informatics (ISI), 2019, 167-169.

Xiaomeng J, Z. Fan, L. Shenwen, Y. Jinglin and H. Ketai, "Data Analysis of Bitcoin Blockchain Network Nodes," 2020 15th IEEE Conference on Industrial Electronics and Applications (ICIEA), 2020. 1891-1895.