

MODIFYING AND PREDICTING OF CRYPTOCURRENCY PRICES USING DATA SCIENCE TECHNIQUE

Amritha A¹, Asrafunisha R², Nandhini R³, Pushpalatha S⁴

Student, Information Technology, Jeppiaar Engineering College, Tamilnadu, India

Student, Information Technology, Jeppiaar Engineering College, Tamilnadu, India

Student, Information Technology, Jeppiaar Engineering College, Tamil Nadu, India

Assistant Professor, Information Technology, Jeppiaar Engineering College, Tamilnadu, India

ABSTRACT

Nowadays, cryptocurrencies are playing a vital role in investment. societies are investing their money within the cryptocurrency share market to achieve income and it's tough to seek out precise prices. Cryptocurrencies are virtual money that's protected by cryptography. cryptography is employed for safe marketing. cryptocurrency transactions that are occurring online without a 3rd party negotiator. Cryptocurrency works using blockchain technology. They used coding for storing and transaction the cryptocurrency[1]. In this project, we used data science techniques and machine learning algorithms to predict the value of the cryptocurrency. the machine learning algorithm is employed to coach and teach the information and it's developed to search out cryptocurrency prices. the info science technique is employed for getting a better model for predicting cryptocurrency prices. We used different machine learning algorithms and compare the algorithm to work out which algorithms perform well. these algorithms are employed in the pre-processing of data and precision. And other execution metrics like precision, recall, and score are also taken for analysing the model

Keywords: *cryptocurrency, machine learning, price prediction*

1. INTRODUCTION

The project's purpose is to create a machine learning model for cryptocurrency price prediction. People investing their money in cryptocurrency and it's tough to predict the rate. Many scammers are involved in the pump and dump operation and posting fake messages regarding the cryptocurrencies prices on social media platforms and that they defrauded people. In the existing system, they only analyze the impact of social media regarding cryptocurrency pump and dump operations[2]. Those pump and dump operations started in the stock market .that are now quickly developing in the crypto market. there are various cryptocurrencies are available in the market. That are Bitcoin, Ethereum, Cardano, Binance coin, we can predict the value of the cryptocurrency by using data science techniques and machine learning algorithms. we compared different algorithms to see which algorithm gives accurate prices.

2. LITERATURE SURVEY

In the recent development, In this price prediction, they gather information from different reference papers and apply it in real-time. they will predict the sign of the daily price change with the highest possible accuracy. and they use LASSO,SVM regression [3]. In another work on bitcoin price prediction using machine learning in their work, they pointed to understanding and identify daily changes in the bitcoin market while obtaining insight into most appropriate features surrounding bitcoin prices. they tried to estimate the bitcoin price precisely taking into consideration various parameters that affect the bitcoin value [4]. In automated cryptocurrency price prediction using machine learning, this paper, focus on market prices of cryptocurrency price recorded daily over the period of 6 months with dataset consists of over nine features related to price of cryptocurrency. They could not predict the price with market price prediction[5]. By using sentiment analysis and machine learning, in this paper, the usage of machine learning tools and available in the social media data for predicting the price movement of the bitcoin using elements from twitter and market data as input features. they tried to predict prices but did not predict with accuracy[6]. bitcoin transformation of cryptocurrency into a global phenomenon, this study reports that attribute of bitcoin through a systematic literature review. the paper is based on primary data from existing literature and primary data from relevant case studies[7].

3.PROPOSED SYSTEM

Cryptocurrency processes contrarily and it is tough to predict the value. The proposed model is evolved by supervised learning algorithms it will predict the most exactitude value of cryptocurrency prices. this model will compare the various different regression algorithms and see which algorithm performs better and provides correct accurate prices by using the data science technique, machine learning algorithms. this model helps people to invest in the crypto market and authorizes business to make high precise values. and other execution metrics are also done in this model.

3.1 WORKING PRINCIPLE

The method onsets from data preprocessing it will clean the data and review the missing values. Dividing the test and training the given dataset is then employed to develop a model by dividing the dataset into 7:3 ratio. Every model has various execution and characteristics. In unseen data we can see how precise each model is by using the resampling technique. We have a new dataset the dataset can be visualize in different procedures in graphical representation by using various visualization methods we can see the average precision. in the n, ext we will similarly compare the different algorithms in the same dataset. the 6 various algorithms are compared logistic regression, linear regression, decision tree regression, random forest regression, support vector regression, and lasso regression. Regression has problemem predicting what is entailed in the numeric value. we used metrics regression to find an error score in this project. we calculated the mean squared error, root means squared error, and mean absolute error by using formulas.

The mean squared error will calculate the expected target value and predicted value in the dataset.

$$MSE=1/N* \sum \text{for } I \text{ to } N(y_i-\hat{y}_i)^2$$

Root mean square is an addition of the mean square error. the square root of the error is also calculated in the RMSE. The RMSE can be calculated as follows: $RMSE=\sqrt{1/N*\sum \text{for } i \text{ to } N (Y_i-\hat{y}_i)^2}$
 $RMSE=\sqrt{MSE}$

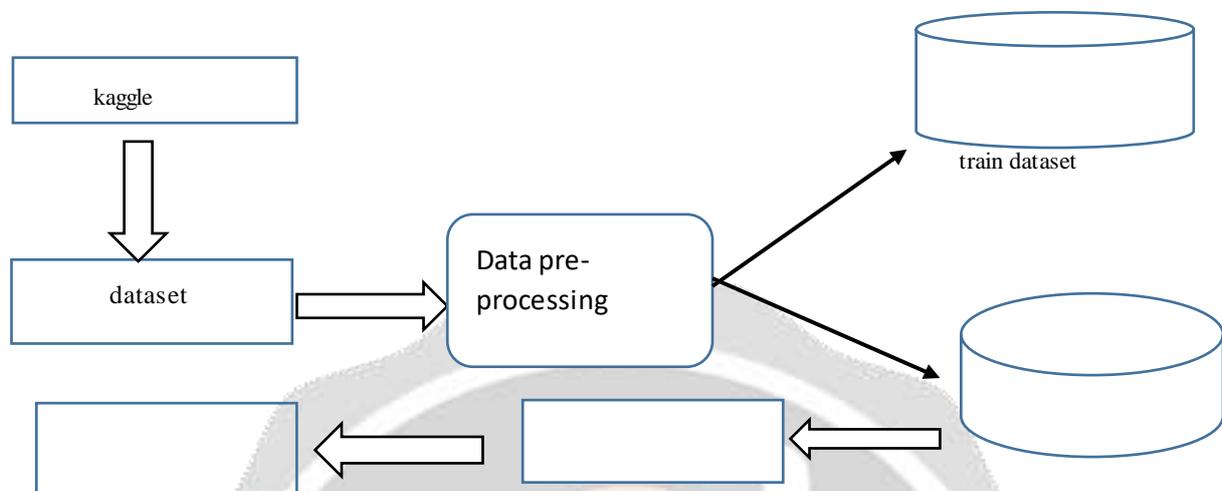
The mean square error value calculates the average value of the absolute error values. The mean absolute error can be calculated as follows:

$$MAE=1/N*\sum \text{for } I \text{ to } N \text{ abs}(y_i-\hat{y}_i)$$

Used python packages:

- Sklearn
- Numpy
- Pandas
- Matplotlib

By comparing different algorithms random forest regression gives possible accuracy because it is good in data pre-processing and discrete values.it finds out the error in the data set and is easy to tune with. Cryptocurrency price



3.1.1 workflow diagram of data pre-processing

3.2 DATASET DESCRIPTION

In this project, we collected a cryptocurrency price dataset from the Kaggle website. This contains thousands of datasets. This dataset will be trained and taught by using this model and it contains row, volume, ratio, slug, high, low, and open. We collected datasets for bitcoin, Binance coin, Ethereum, and Cardano. We collected this dataset for modifying and predicting the cryptocurrency value. the dataset is shown as a figure,

3.2.1 dataset for cryptocurrency coin

4. MODULES DESCRIPTION

The different modules will be explained as follows,

4.1 DATA PREPROCESSING

Data pre-processing is a process in machine learning to find the error in the data set. It is used to discover the missing values, values that are repeated in the dataset, and null integrity. Importing the data set in the Kaggle website with the provided dataset will find the blunders in it. it also evaluates the variable identification in the

dataset by data type and data shape. Then the correct validation dataset data will instruct the model and it will help to test the datasets and evaluate for predicting the prices.

4.2 DATA VISUALIZATION

The data visualization process is obtained in statistics and machine learning. Data visualization will show the end in the graphical expression of input and data during a graph or chart format. Data visualization used pictures to know data. Data visualization is often used to convey and indicate the information in an exceedingly plot or chart, or graph. It will import the packages and browse the info and renovate it into a graphical representation.

4.3 LINEAR REGRESSION

Linear regression is employed in machine learning and method. It is accustomed predict the values. It has dependent and independent variables. It'll find the variable supported the worth of a variable quantity. Regression toward the mean is the process of sloped line. In that line has data plots. While training the model it'll predict the worth of labeled data by the input of coaching data. Our objective is to urge an accurate value that is closest to the bulk of the focus.

4.4 DECISION TREE REGRESSION

Decision tree regression is put forth in machine learning. it is employed for predicting the cryptocurrency values during a tree node structure. its ridges into minor and minor subsets at an analogous time decision tree are formulated. the worthwhile upshot may be a tree with decision and leaf lump a decision lump has likewise more branches. The leaf node depicts the numerical value.

4.5 LOGISTIC REGRESSION

Logistic regression is utilized for the method. The most objective of the logistic regression is to point to and predict the expectations of the variable. And it evaluates a dataset whether there are one or more independent variables for inferring the result. Logistic regression contains binary variables the ultimate result is 0 or 1.

4.6 SUPPORT VECTOR REGRESSION

Support vector regression is obtained in supervised learning algorithms. it finds the unique values in the dataset. Support vector regression is similar to support vector machine it has little disparity. It reduces the errors in the dataset. Support vector regression will predict the high probable value. This regression mainly uses linear data to predict the value.

4.7 LASSO REGRESSION

Lasso regression is one of the regression procedures. The method of lasso regression first imports the library package so reads the given data set. The info set is regularized then the information set is split and trained the data set and eventually applying the lasso regression algorithms

4.8 RANDOM FOREST REGRESSION

Random forest regression is employed in the machine learning algorithm. it refers to a supervised learning technique. Random forest regression contains a decision tree and is divided into subsets into smaller and smaller subsets and it will improve the precision of the dataset. The random forest takes precision from each tree and based on high accuracy it gives the final output.

4.9 DEPLOYMENT

The final output of the price prediction is in the flask deployment and it is written in python language. flask is evaluated as a micro web framework because it does not accomplish particular tools or library packages. It is formulated on top of python and uses all characteristics.

5. CONCLUSION

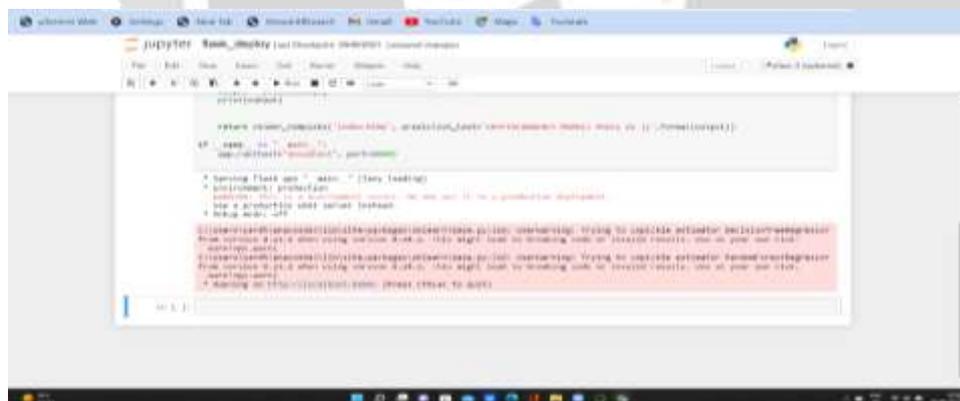
We concluded that it's feasible to predict the value of cryptocurrency value with the best precision. we correlated different algorithms and discovered which algorithms give the most effective precision and picked up the datasets from the Kaggle website. This model pre-processing, visualization, and eventually building the model. By comparing different algorithms random forest regression gives possible accuracy because it is good in data pre-processing and discrete values.it finds out the error in the data set and is easy to tune with. This model can help to seek out cryptocurrency price prediction.

5.2 FUTURE ENHANCEMENT

- Our future development is to develop an application that can guide you on where to invest and which coin will give you compensation.
- Cryptocurrency price prediction to pertain with an AI model.

5.3 RESULT

The below figure shows how the cryptocurrency price is predicted. The cryptocurrency price is predicted by given datasets



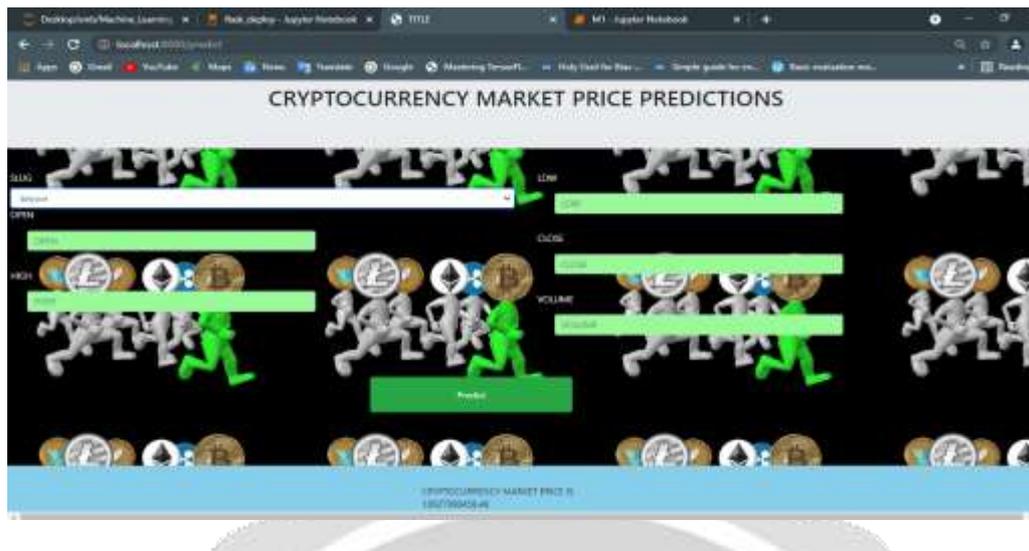
```

from flask import Flask, jsonify, request
from prediction import predict_price

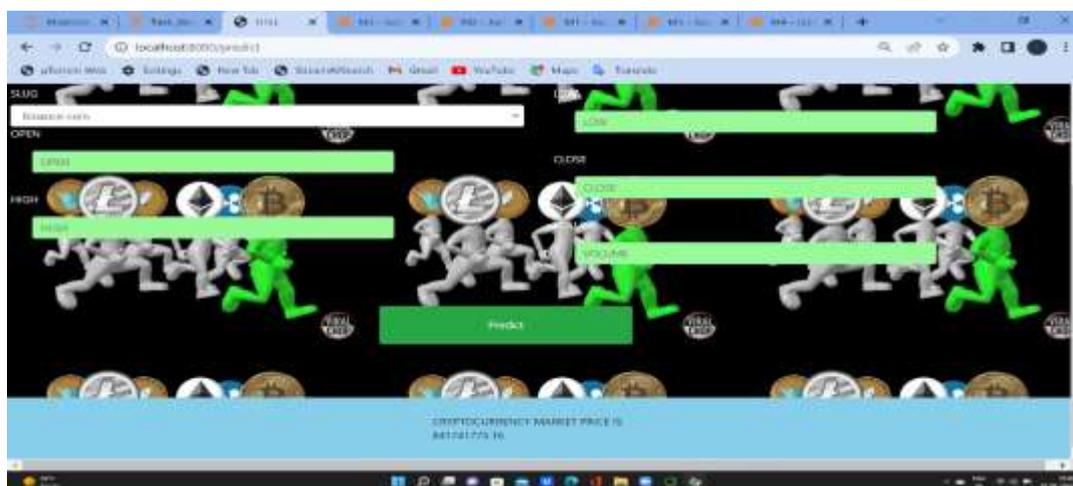
app = Flask(__name__)

@app.route('/predict', methods=['POST'])
def predict():
    data = request.get_json()
    price = predict_price(data)
    return jsonify({'price': price})

if __name__ == '__main__':
    app.run(host='0.0.0.0', port=5000)
  
```



5.3.1 Bitcoin cryptocurrency market price



5.3.2 Binance coin cryptocurrency market price

6. REFERENCES

- [1] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," Bitcoin.org, Tech. Rep., 2008.
- [2] U.S Securities and Exchange Commission. (2017). Customer Advisory: Beware Virtual Currency Pump-and-Dump Schemes. [Online]. Available:
- [3] C. Hutto and E. Gilbert, "Vader: A parsimonious rule-based model for sentiment analysis of social media text," in Proc. 8th Int. Conf. Weblogs Social Media (ICWSM), 2014, pp. 216–225. [Online]. Available: <http://comp.social.gatech.edu/papers/icwsm14.vader.hutto.pdf>
- [4] A Research On Bitcoin Price Prediction Using Machine Learning Algorithms Lekkala Sreekanth Reddy, Dr.P. Sriramya
- [5] PREDICTION OF CRYPTOCURRENCY PRICES USING MACHINE LEARNING TECHNIQUES by Alireza Ashayer April 2019

[6] Price Movement Prediction of Cryptocurrencies Using Sentiment Analysis and Machine Learning Franco Valencia *, Alfonso Gómez-Espinosa * and Benjamín Valdés-Aguirre

[7] AUTOMATED CRYPTOCURRENCIES PRICES PREDICTION USING MACHINE LEARNING Ruchi Mittal, Shefali Arora, and M.P.S Bhatia

