Journal of Information Systems Engineering and Management
2025, 10(35s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Machine Learning Strategies for Accurate Cyrptocurrency Forecasting

Dr.A.Mohamed Anwar [1], Dr. R. Sabin Begum[2*], Dr.M.Vijayaraj [3], Dr. J.H.Jassma Yasmin[4], Dr. S.P. SanthoshKumar[5]
[1] Assistant Professor Senior Grade, Department of Computer Science and Engineering School of Computing ,Vel Tech Rangarajan Dr. Sagunthala R&D institute of Science and Technology, Avadi, Chennai, Tamil Nadu, India.
[2] Assistant professor Department of Computer Applications, B.S. Abdur Rahman Crescent Institute of Science and Technology,Vandalur,Chennai,Tamil Nadu,India.
[3] Assistant Professor Senior Grade, Department of Computer Science and Engineering School of Computing,Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Avadi,Chennai, Tamil Nadu, India.
[4] Assistant Professor, Department of Computer Science and Engineering,Chennai Institute of Technology ,Kunrathur,Chennai, Tamil Nadu, India.
[5] Assistant Professor Senior Grade, Department of Computer Science and Engineering School of Computing,Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Avadi, Chennai, Tamil Nadu, India.
**Corresponding Author : sabin@crescent.education*

ARTICLE INFO

Received: 30 Dec 2024

Revised: 19 Feb 2025

Accepted: 27 Feb 2025

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: Linear regression, support vector regression, SGD regression, lasso regression, XG boost regression.

1. INTRODUCTION

Cryptocurrency has become a game-changer because of its vast advantages like freedom of payment, anonymity, high security, speed, and mainly no thirdparty involvement. 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. 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.

Journal of Information Systems Engineering and Management
2025, 10(35s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

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.

2. LITERATURE SURVEY

Ahmed et al. (2023) conducted a comparative analysis of multiple ML models for cryptocurrency price prediction, concluding that deep learning models, particularly LSTMs, outperform traditional statistical methods. Similarly, Singh and Verma (2024) found that ML techniques such as Support Vector Machines (SVM), Decision Trees, and Random Forest show varying degrees of accuracy, with ensemble models offering improved performance due to their ability to generalize over diverse market conditions.

Deep learning architectures, especially Long Short-Term Memory (LSTM) and Convolutional Neural Networks (CNN), have demonstrated effectiveness in capturing temporal dependencies in cryptocurrency data. Alomari and Alksasbeh (2024) proposed a hybrid LSTM-CNN model, integrating spatial and sequential learning, which resulted in enhanced predictive accuracy. Such models leverage both short-term fluctuations and long-term trends, making them suitable for cryptocurrency price forecasting.

Macroeconomic factors, including inflation rates, interest rates, and geopolitical events, have a significant impact on cryptocurrency price movements. Gupta and Lahiani (2023) investigated the relationship between macroeconomic uncertainty and cryptocurrency volatility, highlighting the necessity of incorporating economic indicators into ML models to improve forecasting reliability.

The cryptocurrency market is influenced by external factors such as investor sentiment, news, and macroeconomic indicators. Kumar and Bansal (2024) explored sentiment analysis-based cryptocurrency price prediction, integrating social media sentiment and financial news into ML models. Their findings indicate that sentiment-driven features significantly enhance forecasting accuracy, especially during volatile market conditions.

Zhao et al. (2024) demonstrated that ensemble learning models, which combine multiple predictive techniques, perform better than single models. Their research found that techniques like bagging and boosting improve overall prediction accuracy by leveraging the strengths of individual models and minimizing prediction errors

This model analyses the current market and the current trends in social media to predictthe prices, with RNN and LSTM, the sentiment analysis is done. To map the scores collected from Twitter and the data UNIX timestamp is used. The prediction process is done using random forest regression. The performance metrics like RMSE and MSE are used for validation. The data is collected from Twitter, historical data, and as well as news

3. PROPOSED SYSTEM

The proposed system focuses on predicting the price of the cryptocurrency more precisely and accurately. It uses 7 different algorithms – linear regression, support vector regression, SGD regression, lasso regression, XG boost regression, ridge regression, and random forest regression. To find the best fit among the different algorithms for the system, various performance metrics like root mean squared, mean squared and mean absolute error are used. We have taken five different cryptocurrencies like Bitcoin, Ethereum, Litecoin, Binance coin and Maker. The key parameters like the number of days, cryptocurrency type, open, high, low and close values are taken as input from the user and the open prices for the 'n' upcoming days can be predicted as the output. The model is deployed to the user in the form of a website using flask.

4. SYSTEM ARCHITECTURE

Initially pre-processing is done on the collected dataset. In the pre-processing, variousactions like feature selection, correlation, data cleaning and data encoding is performed on thedataset to make it more effective and error-free. And upon the cleaned dataset, the various regression algorithms like linear regression, support vector regression, SGD regression, lasso regression, XG boost regression, ridge regression, and random forest regression are applied toseewhich algorithm is more suitable and gives results more accurately. To evaluate this various

Journal of Information Systems Engineering and Management
2025, 10(35s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

performance metrics like Root Mean Squared Error, Mean Squared Error and Mean Absolute Error 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. It is dumped inside the web server, from the user's request, the results are predicted and given through the web application.

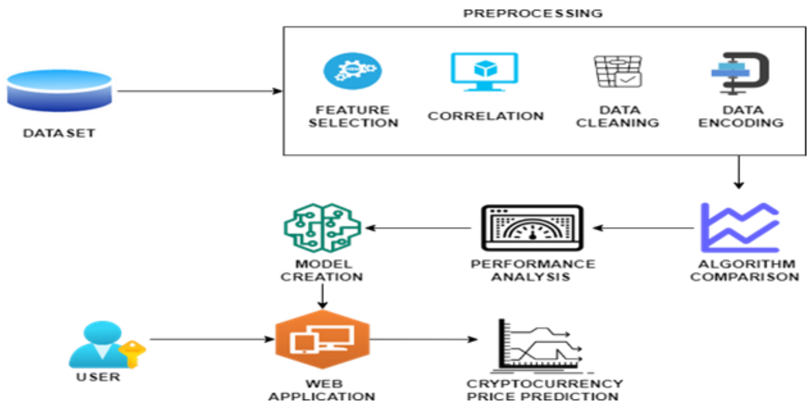

Fig. 4.1 System Architecture

The following diagram fig. 4.2 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.

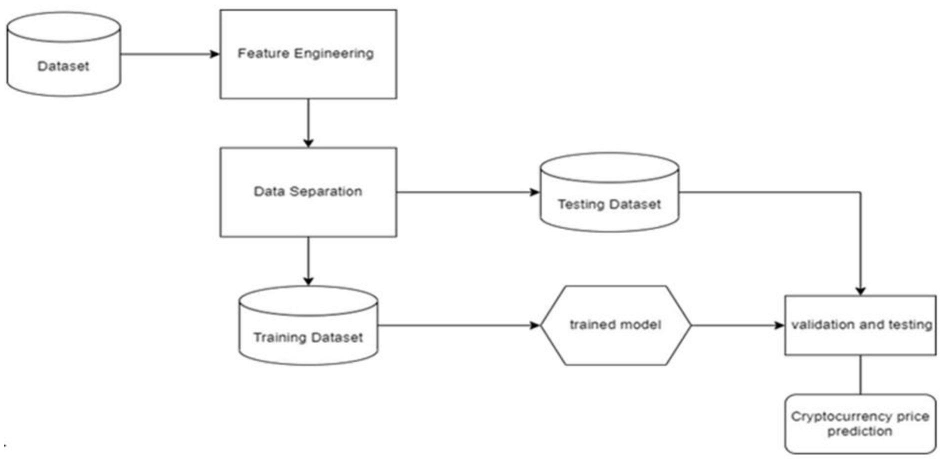

Fig 4.2 Functional architecture

5. METHODOLOGY

The modular design defines the structure of all the modules in the system. This concept breaks the system into separate components to build parallelly and to concentrate on every small detail and requirement. And finally, all the modules are combined or connected togetherto work.

The proposed system has the following list of modules

- Data Pre-Processing
- Exploratory Data Analysis

Journal of Information Systems Engineering and Management
2025, 10(35s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

- Model training
- Prediction and Deployment

5.1 Data Pre-Processing

The dataset for each cryptocurrency like Bitcoin, Ethereum, Litecoin, Binance coin and Maker were collected separately. Feature engineering is made to add a new column to representthe cryptocurrency name to identify among the others. At last, all the five different cryptocurrency dataset is integrated into a single dataset. Secondly, the null values are identified and removed from the dataset to prevent any future errors. The non-numerical parameters are changed in order to make the machine understand better. Correlation for each and every cryptocurrency is made in order to understand the parameters better. Upon analysing

the result, the parameters like date, adj close and volume are dropped. Thus, open, high, lowand close prices are taken as primary parameters for the model.

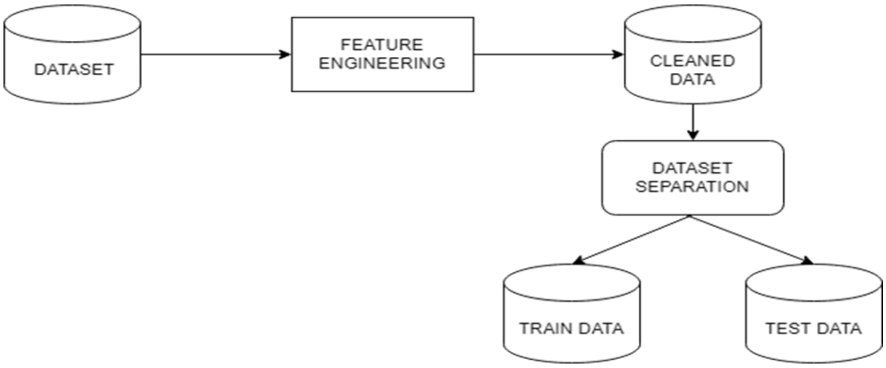

Fig. 5.1 Data Pre-processing

5.2 Exploratory Data Analysis

The ED Analysis was made to translate the information into a visual context to make iteasier 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 andnature. 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 oftime. And the past patterns were analysed 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.

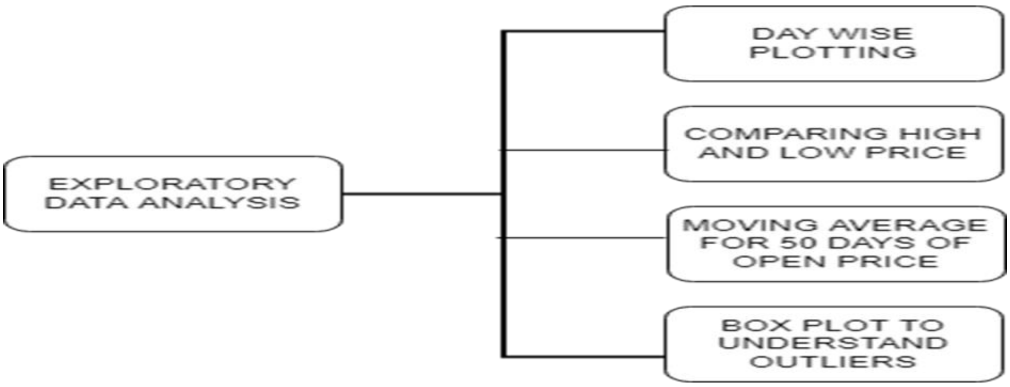

Fig.5.2 Exploratory Data Analysis

5.3 Model training

Initially, the target variable is created by shifting the respective parameters like open, high, low and close values.

Journal of Information Systems Engineering and Management
2025, 10(35s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

And upon completing it, again the null values are checked and removed. Data scaling is made to reduce the range which enormously helps in processing morequickly and effectively. It further increases the accuracy of the trained model. The dataset is split in the ratio of 8:2 for training and testing respectively. And finally, the 7 different algorithms (i.e.) linear regression, support vector regression, SGD regression, lasso regression,XG boost regression, ridge regression, and random forest regression are applied to the training model. Their performance is tested by using different performance metrics like mean squared error, mean absolute error, root mean squared error and R squared. And the best algorithm which is more suitable with less error and produces more accuracy is chosen among the other algorithms**.**

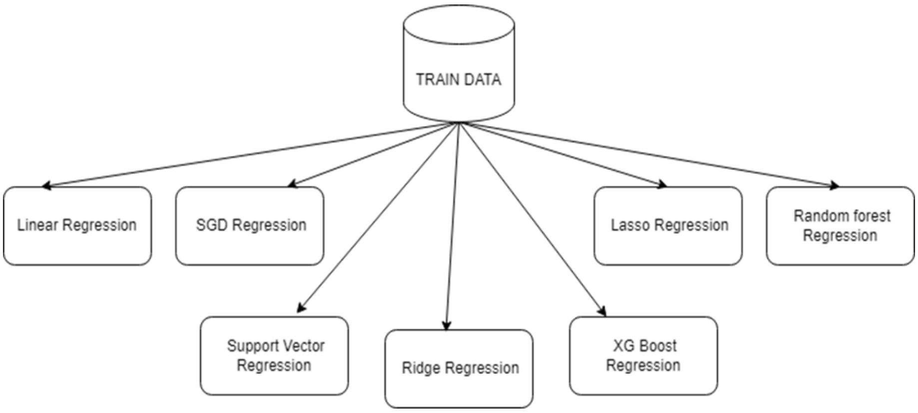

Fig 5.3.Model Training

5.4 Prediction and Deployment

When the user enters the input parameters like number of days, cryptocurrency type, open, high, low and close prices through the website. It is made as a request to the flask serverwhere the trained model has been dumped. And similarly, the predicted price for the 'n' numberof days is displayed as an output to the user through the response from the server which is shown in fig 5.4.

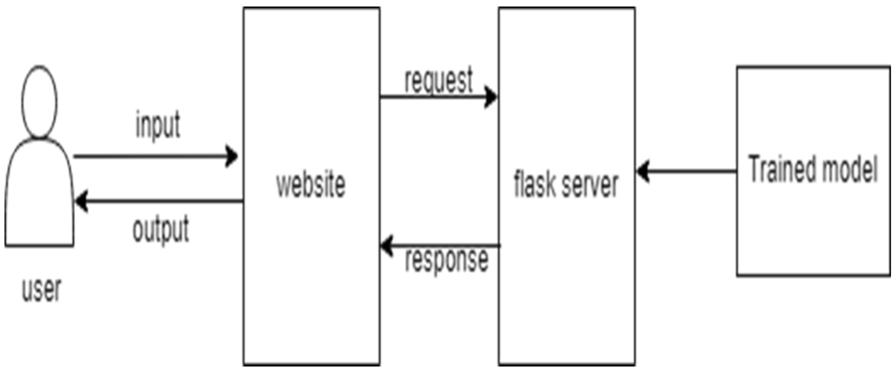

Fig 5.4 Prediction and Deployment

6. RESULTS AND DISCUSSION

The various steps are given in below section to the successful development and deployment of our cryptocurrency market price prediction system.

6. Data Preprocessing

A. REMOVAL OF NULL VALUES

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 theperformance of the algorithm.

Journal of Information Systems Engineering and Management
2025, 10(35s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

Table 6.1 Identification of null values

	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

The following table 6.2 depicts the aftermath of the null value removal from the datasetbefore applying it to the machine learning algorithms.

Table 6.2 Removal of null values

	Number of Mising values	Percentages of Missing Values
Unnamed: 0	0	0.0
Date	0	0.0
Open	0	0.0
High	0	0.0
Low	0	0.0
Close	0	0.0
Adj Close	0	0.0
Volume	0	0.0
Cripto	0	0.0

B. 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. Table 6.3 represents the correlation of different attributes.

Table 6.3 Correlation of attributes

	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

Upon analysing the result, the parameters like date, adj close and volume are dropped. Thus,open, high, low and close prices are taken as primary parameters for the model which is shown in table 6.4.

Journal of Information Systems Engineering and Management
2025, 10(35s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Table 6.4 Final parameters

	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

6.2 Exploratory Data Analysis

A. 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. 6.1 represents the plot of day-wise close price**.**

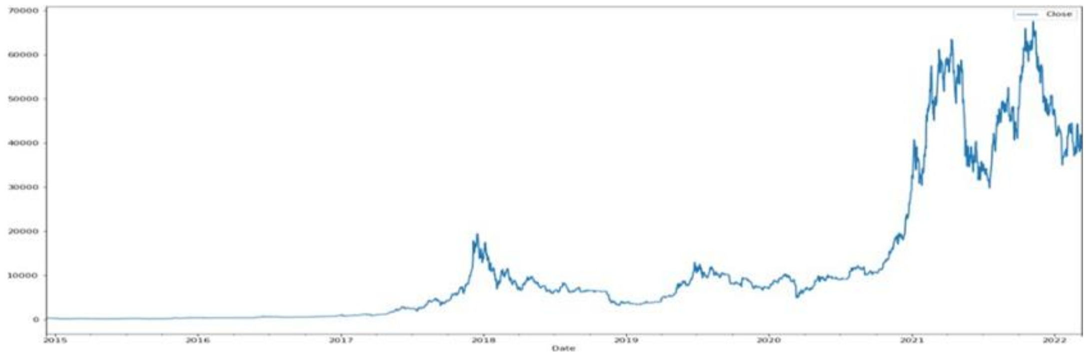

Fig 6.1 Day wise close price plotting

B. Comparing High And Low Price

Fig. 6.2 represents the comparison of high and low prices of the cryptocurrency.This method is important in trading and investing, as it helps in cost estimation and alsosegregates cost with minimal data points.

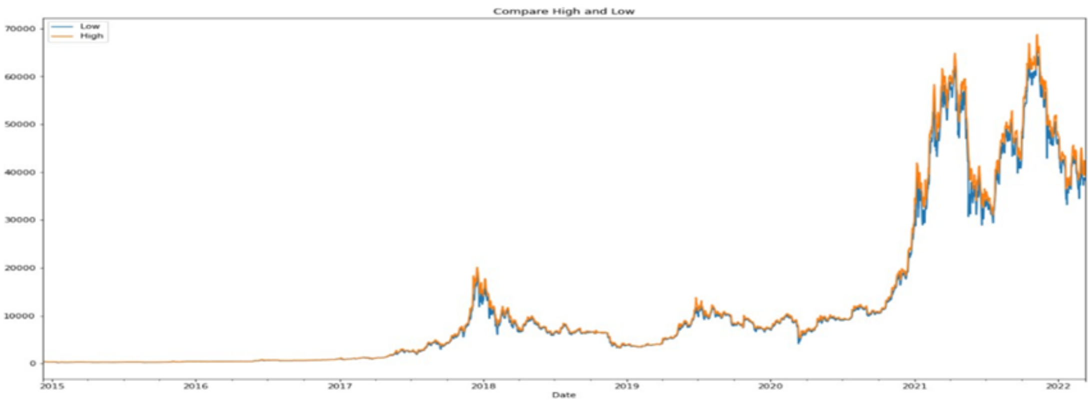

Fig. 6.2 Comparison of high and low price

6.3 Model Training

Data scaling is made to reduce the range which enormously helps in processing more quickly and effectively. It further increases the accuracy of the trained model. The 7 different

algorithms (i.e.) XG boost regression, linear regression, SGD regression, support vector regression, lasso regression, XG boost regression, ridge regression, and random forest regression are applied to the training model.

Journal of Information Systems Engineering and Management
2025, 10(35s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

And the best algorithm which is more suitable with less error and produces more accuracy is chosen among the other algorithms.

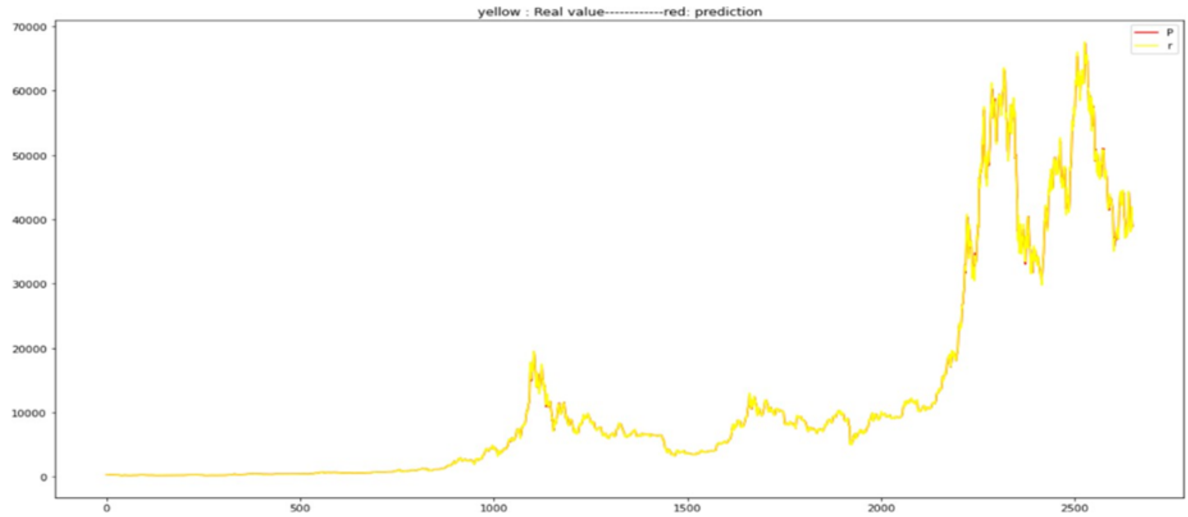

Fig. 6.3 Trained model result

6.4 Prediction And Deployment

Fig. 6.4 is the user interface where we have to enter the cryptocurrency type, the numberof days, open, high, low and close price of the current date. On clicking the predict button, thepredicted price for 'n' number of days is displayed to the user.

Fig. 6.4 User Interface

In fig. 6.5, the different parameters have been entered into the form for submission to predict the price. Here the request and response are through the flask server where the trained modelis dumped.

Journal of Information Systems Engineering and Management
2025, 10(35s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

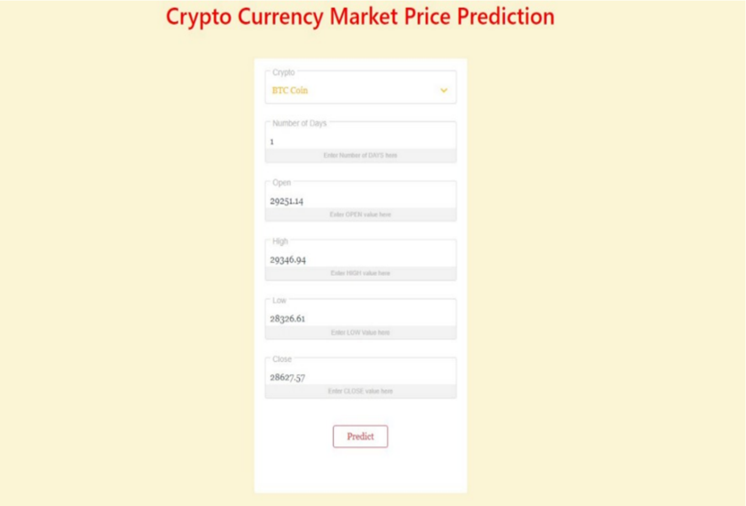

Fig. 6.5 Updation of input values

Fig. 6.6 Represents the aftermath of the submission. Here the predicted price for the 'n'number of days is displayed.

Fig. 6.6 Price prediction

7. CONCLUSION

The market of the cryptocurrency will go double its value as estimated by many analystsaround 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 analysing 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

[1] Ahmed, S., Rahman, M. M., & Islam, M. T. (2023). Machine learning-based cryptocurrency price prediction: A comparative analysis of different models. Algorithms, 16(12), 542. https://doi.org/10.3390/a16120542.
[2] Alomari, R., & Alksasbeh, A. (2024). A hybrid deep learning model for cryptocurrency forecasting using LSTM and CNN. Proceedings of the Artificial Intelligence Conference (ACE). https://www.ewadirect.com/proceedings/ace/article/view/9912
[3] Dr.R.Sabin Begum, Dr.A.Mohamed Anwar, Dr. S. Anu Priya,(2025) A Health-Based Deep Learning System for

Rapid and Precise Detection of Acute Lymphoblastic Leukemia.Journal of Neonatal Surgery,14(4s), 87-94.
[4] Gupta, R., & Lahiani, A. (2023). Macroeconomic uncertainty and cryptocurrency predictability. International Review of Financial Analysis, 88, 102311. https://doi.org/10.1016/j.irfa.2023.102311.
[5] Kumar, P., & Bansal, A. (2024). Sentiment analysis-based cryptocurrency price prediction using machine learning techniques. European Journal of Management and Business Economics, 33(1), 99–117. https://doi.org/10.1108/ejmbe-08-2023-0244
[6] Singh, D., & Verma, A. (2024). Comparative study of machine learning models for Bitcoin price forecasting. Expert Systems with Applications, 236, 120615. https://doi.org/10.1016/j.eswa.2023.120615.
[7] Zhao, Y., Li, M., & Zhang, Q. (2024). Time-series forecasting of cryptocurrencies using ensemble machine learning models. MDPI Applied Sciences, 14(3), 196. https://doi.org/10.3390/app14030196.