International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 05 Issue: 03 | March – 2026 DOI:10.55041/ISJEM05833
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

Machine Learning-Based Time Series Analysis for Cryptocurrency Price Prediction using Synthetic Data

Dr. K. Satyam[1], Desman Likhitha[2]

[1]Associate Professor, Department of MCA, Annamacharya Institute of Technology & Sciences, Tirupati, Andhra Pradesh, India.

[2] Post Graduate, Department of MCA, Annamacharya Institute of Technology & Sciences, Tirupati, Andhra Pradesh, India.

ABSTRACT

Because of the tremendous volatility and erratic behaviour of cryptocurrency markets, it is difficult to estimate prices accurately. This research uses a synthetic time-series dataset to provide a machine learning-based method for predicting bitcoin prices. Open price, peak price, low price, close price, and trading volume are among the daily trading features in the dataset that mimic actual market conditions. The suggested approach analyses past trends and forecasts future closing prices using a variety of regression-based machine learning models. To increase the accuracy and efficiency of the model, data preprocessing methods including feature selection and normalisation are used.

Standard measures like Mean Absolute Error (MAE), Mean Squared Error (MSE), and R-squared score are used in the study to assess model performance. Even when trained on artificial data, experimental results show that machine learning algorithms may successfully identify underlying patterns in bitcoin price movements. The results show the promise of predictive modelling in financial forecasting and lay the groundwork for future studies utilising cutting-edge deep learning techniques and real-world datasets.

KEYWORDS

Cryptocurrency, Machine Learning, Price Prediction, Time Series Analysis, Synthetic Dataset, Regression Models, Financial Forecasting, Data Preprocessing

I. INTRODUCTION

Cryptocurrencies have become an important part of the world financial system in recent years. Because of their decentralised structure, quick growth, and potential for large rewards, digital currencies like Bitcoin and Ethereum have attracted a lot of attention. However, price prediction is a difficult and complex endeavour because cryptocurrency markets are extremely volatile and impacted by numerous dynamic elements.

For traders, investors, and financial analysts to make wise choices, accurate cryptocurrency price forecasting is essential. The intricate and nonlinear patterns found in financial data are frequently missed by conventional statistical techniques. In order to analyse time-series data and predict future patterns, machine learning techniques have grown in popularity.

Large datasets can reveal hidden correlations, and machine learning algorithms can adjust to evolving patterns over time. Predictive analysis in financial applications has made extensive use of algorithms including Random Forests, Decision Trees, and Linear Regression. By capturing both linear and nonlinear interactions in the data, these models outperform conventional methods in terms of accuracy.

ISSN: 2583-6129
DOI:10.55041/ISJEM05833

The goal of this project is to create an intelligent system that uses a synthetic dataset to anticipate bitcoin prices. Each record in the time-series data structure of the dataset represents daily trade information, including open, high, low, close prices, and trading volume. Predicting the upcoming closing price using historical data is the main goal.

The suggested solution shows how machine learning methods can be successfully used to solve financial forecasting issues. Despite being artificial, the dataset is appropriate for model training and evaluation because it closely mimics actual trading behaviour. This work lays the groundwork for future investigations utilising cutting-edge deep learning models and real-time cryptocurrency data.

II. PROBLEM STATEMENT

Because digital currency markets are so dynamic and volatile, predicting cryptocurrency prices is a difficult undertaking. A number of reasons, including market demand, investor behaviour, world events, and technical advancements, cause prices to change quickly. These intricate and nonlinear patterns are frequently difficult to fully capture using conventional statistical techniques.

This study's primary goal is to create an effective machine learning-based system that uses past trade data to forecast the closing price of cryptocurrencies. In order to find hidden links between features like open, high, low prices, and trade volume, the system must be able to analyse time-series data. The goal is to increase prediction accuracy and offer a trustworthy model that may help analysts and investors make wiser financial decisions.

III. METHODOLOGY

1. Data Collection

A synthetic bitcoin dataset created to mimic actual market behaviour was employed in this experiment. Open price, peak price, low price, close price, and trading volume are all included in the daily trading data. A structured time-series dataset is created, with each record representing a single trade day.

2. Data Preprocessing

To guarantee the consistency and quality of the dataset, data preparation is a crucial step. The dataset is examined for inconsistent or missing values. To improve model performance, feature scaling and normalisation techniques are used to bring all features to a similar scale. By doing this, bias is lessened and model convergence is accelerated.

3. Feature Selection

The closing price is regarded as the target variable, while pertinent characteristics like open, high, low prices, and volume are chosen as input variables. The efficiency of machine learning models is increased and superfluous complexity is reduced with the aid of feature selection.

4. Model Selection

Multiple machine learning algorithms are used to perform prediction, including:

- Linear Regression
- Decision Tree Regression
- Random Forest Regression

These models are selected based on their ability to handle regression tasks and capture patterns in time-series data.

5. Model Training

The dataset is divided into training and testing sets. The training set is used to train the machine learning models, allowing them to learn patterns from historical data. The testing set is used to evaluate model performance on unseen data.

ISSN: 2583-6129
DOI:10.55041/ISJEM05833

IV. EXPERIMENTAL RESULTS

The synthetic cryptocurrency dataset was used to test and execute the suggested machine learning models in order to assess how well they predicted closing values. To guarantee that the models were properly validated, the dataset was split into training and testing subsets. The dataset was subjected to several regression algorithms, such as Random Forest, Decision Tree, and Linear Regression. Standard measures including Mean Absolute Error (MAE), Mean Squared Error (MSE), and R-squared score (R2) were used to assess these models' performance.

The findings show that by identifying linear correlations in the data, linear regression offers a fundamental degree of prediction accuracy. In certain situations, Decision Tree Regression may experience overfitting, yet it performs better by detecting nonlinear patterns. Random Forest Regression outperforms all other models because of its ensemble learning strategy, which enhances generalisation and lessens overfitting.

Overall, the experimental results show that, even when trained on artificial datasets, machine learning models can successfully identify patterns from time-series cryptocurrency data and produce accurate predictions.

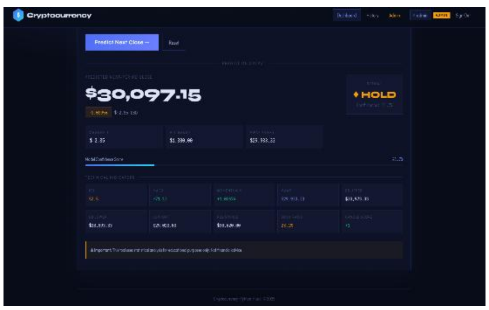

Figure: Cryptocurrency Price Prediction Dashboard

The suggested bitcoin price prediction system's graphical user interface is depicted in the above figure. Using machine learning techniques, the system's web-based dashboard allows users to forecast a cryptocurrency's next closing price.

The "Predict Next Close" button on the UI initiates the prediction process. The system shows the anticipated closing price upon execution, which in this instance is $30,097.15. The dashboard offers further information in addition to the forecast value, including price change, percentage variance, and model confidence score.

The trading signal indicator, which recommends actions like Hold, Buy, or Sell, is a crucial component of the method. With a confidence level of 21.7%, the system suggests a "HOLD" signal in the current output, indicating a modest level of prediction certainty. Additionally, the dashboard displays a number of technical indications, such as:

- Relative Strength Index (RSI)
- Moving Average Convergence Divergence (MACD)
- Momentum
- Volume Weighted Average Price (VWAP)
- Support and Resistance levels
- Bollinger Bands (Upper and Lower)

These indicators help in understanding market trends and validating the prediction results.

V. Conclusion

This study uses a synthetic time-series dataset to provide a machine learning-based method for forecasting bitcoin prices. To accurately predict future closing prices, the algorithm makes use of historical trading characteristics such open, high, low, close prices, and volume. Several regression models were used, and the findings show that ensemble approaches—Random Forest in particular—offer higher prediction accuracy than single models.

By displaying forecasts combined with technical indicators and trading signals, the created web-based dashboard improves the system's usability. The outcomes of the experiment show that machine learning methods can effectively identify trends in the fluctuations of bitcoin prices. The suggested methodology can be applied to real-world datasets even when the dataset is artificial. In general, this study shows how machine learning can be used in financial forecasting and decision support systems.

REFERENCES

[1] J. Brownlee, Machine Learning Mastery With Python: Understand Your Data, Create Accurate Models, and Work Projects End-to-End, 1st ed. Melbourne, Australia: Machine Learning Mastery, 2017.

[2] S. Russell and P. Norvig, Artificial Intelligence: A Modern Approach, 3rd ed. Upper Saddle River, NJ, USA: Pearson, 2016.

[3] T. Chen and C. Guestrin, “XGBoost: A Scalable Tree Boosting System,” in Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining (KDD), San Francisco, CA, USA, 2016, pp. 785–794.

[4] L. Breiman, “Random Forests,” Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.

[5] Y. LeCun, Y. Bengio, and G. Hinton, “Deep Learning,” Nature, vol. 521, no. 7553, pp. 436–444, May 2015.