https://doi.org/10.48047/AFJBS.6.14.2024.3835-3844

African Journal of Biological Sciences

Journal homepage: http://www.afjbs.com

Research Paper Open Access

FORECASTING ETHEREUM PRICE TRENDS USING ARTIFICIAL INTELLIGENCE TECHNIQUES: A FUTURE BROADCAST APPROACH

[1]Pradeep Patel, [2]aarti Bokade, [3]Kapilkumar Dave, [4]Ashokkumar Lakum
[1]pradip.patel.ic@gmail.com, [2]aartigyadav@gmail.com,
[3]profkcdave@gmail.com[4]aclakum@lecollege.ac.in
[1,2,3]Instrumentation and Control Engineering Department, [4]Electrical Engineering Department
[1,2,3,4]Government Engineering College Gandhinagar, [4]L. E. College, Morbi, Gujarat, India

Volume 6, Issue 14, Aug 2024

Received: 09 June 2024

Accepted: 19 July 2024

Published: 08 Aug 2024

doi: 10.48047/AFJBS.6.14.2024.3835-3844

Abstract
Machine learning algorithms have been widely used in various prediction applications, such as weather forecasting and disease prediction. This paper focuses on using different types of regression models to predict the future price of Ethereum. The data for the past year, ranging from 2021 to July 12, 2024, has been sourced from Yahoo Finance for training and testing various regression models, including Linear Regression, Decision Tree Regression, Random Forest Regression, Support Vector Regression, and Gradient Boosting Regression. Each regression model is trained using 80% of the data for training and the remaining 20% for testing. The performance of each model is evaluated using key performance measures: Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), R^2 Score, Explained Variance, and Mean Absolute Percentage Error (MAPE). Among the models, the Random Forest Regression model demonstrated superior performance across all key metrics, particularly on the test set. It achieved the lowest errors (RMSE and MAE), highest R^2, highest explained variance, and lowest MAPE. These results indicate that the trained Random Forest model is highly effective and capable of accurately predicting the future price of Ethereum.

Objective
This study aims to compare the performance of regression algorithms in predicting Ethereum prices, addressing the research question: Which regression algorithm performs best in predicting Ethereum prices, and what are the implications of the findings for investors and researchers? The price prediction serves as a crucial tool for investors, traders, and stakeholders in navigating the dynamic and often volatile nature of financial markets, enabling them to make sound investment decisions, manage risks effectively, and capitalize on opportunities for growth and profitability.

Index terms: Ethereum, cryptocurrency, Machine Learning, Regressive Integration, Time series, Regression, SARIMA, ARIMS,Autoregressive Fractional Integrated Moving Average (AFRIMA)

INTRODUCTION

Cryptocurrencies have emerged as a significant asset class, with Ethereum (ETH) being one of the leading digital assets in terms of market capitalization and adoption. Predicting cryptocurrency prices is challenging due to their high volatility and complex market dynamics.Ethereum is a decentralized platform that enables developers to build and deploy smart contracts and decentralized applications (DApps). It was proposed by Vitalik Buterin in late 2013 and development began in early 2014, with the network going live on July 30, 2015. Ethereum's native cryptocurrency is called Ether (ETH), which is used to facilitate transactions and incentivize participants in the network.

The exchange rate between Ethereum (ETH) and USD serves as a crucial economic indicator, impacting various sectors including cryptocurrency trading, cross-border transactions, and investment strategies.One of the key innovations introduced by Ethereum is its ability to execute smart contracts, which are self-executing contracts with the terms of the agreement directly written into code.

Ethereum's significance in the cryptocurrency space lies in several factors:

1. Pioneering Smart Contracts
2. Platform for innovation
3. Decentralization and security
4. Economy incentives
5. Scalability and upgrades

By leveraging historical data and predictive models, we can generate insights to assist stakeholders in mitigating risks, optimizing cryptocurrency conversions, and devising effective trading strategies.

Importance of price prediction for investors, traders, and stakeholders. Price prediction holds significant importance for investors, traders, and stakeholders in the financial markets, including those involved in cryptocurrencies like Ethereum. Here are some reasons why:

- ➢ Decision making
- ➢ Risk management
- ➢ Profit maximization
- ➢ Hedging
- ➢ Strategic planning
- ➢ Market sentiment
- ➢ Regulatory compliance

II. REVIEW OF LITERATURE

The literature review highlights previous research on cryptocurrency price prediction and regression algorithms in financial forecasting. Existing studies have explored various approaches, including machine learning techniques, time series analysis, and sentiment analysis. However, few studies have comprehensively compared the performance of different regression algorithms in cryptocurrency price prediction. Here's a summary of some common methodologies and findings:

- Time series analysis
- Machine learning models
- Sentiment analysis
- Hybrid model

The paper is emphasized about the integrated model should be better than the individual model clearly shown and presented in the table along with prediction diagram. This paper will help in measuring the market more accurately for a particular store and reduce the error rate by leveraging technology. It mainly focuses on whether the current trend will continue and, if not, when it will change [2].

The paper shows ability of the ARIMA model to predict the value of Ethereum during COVID-19 pandemic. The paper results showed that using the ARIMA method to predict the value of Ethereum produced negative results, and the predicted value deviated significantly from the ideal value.

This is also reflected in the results of the accuracy test with MAPE, which gave a result of 51.94%. On side, due to the COVID-19 and the emergence of a shortage (decentralized finance) in early 2021 led to a significant increase in the value of Ethereum. It resulting in a high error rate and use the ARIMA model to predict the value of Ethereum. This paper is recommended using more advancedAFRIMA to obtain better predicted values [3].

This paper conducts comparative study of MAD, RMSE and MAPE as benchmarks among machine learning (XGBoost), statistical learning (Theta, ARIMA, Predictor) and deep learning algorithms (LSTM). Moreover, from this method of comparing different periods of comparative research, it can be shown that technological solutions of forecasting efficient results and good forecasts [4].

The power and predictive quality of the model are also emphasized in this paper. The selection of the most suitable model between these two methods is made based on Minimum Error (ME), Absolute Error (MAE) and Absolute Error (MASE). The results of paper concluded that ARIMA and ANNs are the most suitable for predicting Bitcoin prices. Finally, ANNs are believed to be a powerful model for price prediction [5].

All methods achieved an accuracy of over 50%, the best performance was achieved by ARIMA, which depends on the features and fits the data series. Alternatives were reduced due to lack of data, model assumptions about the data, and model inconsistencies. The author encouraged by the results so far and will continue to examine the dynamics of foreign exchange markets as the field grows [6].

Conclusion is that no single method guarantees accurate predictions, combining multiple methodologies and incorporating additional features can improve the robustness and reliability of Ethereum price forecasts.

III. METHODOLOGY

The methodology used for Ethereum price prediction typically involves a combination of quantitative analysis, statistical modelling, and data-driven techniques. It is as below for Ethereum price prediction [1].

Data preprocessing: It involves cleaning and preparing the collected data for analysis. This includes handling missing values, outliers, data normalization, and feature engineering.

Feature selection: This may involve conducting feature importance analysis, correlation analysis, or domain knowledge-based selection.

Model Selection: Based on the selected features and the nature of the data, a suitable prediction model or models are chosen. Common models used in Ethereum price prediction include time series models (e.g., ARIMA, SARIMA), machine learning algorithms (e.g., regression models, support vector machines, random forests, neural networks), and ensemble methods.

Model Training: The selected model is trained on historical data using a portion of the dataset. During training, the model learns the relationships between the input features and Ethereum's historical price movements.

Model Evaluation: The performance of the model is evaluated using a separate portion of the dataset reserved for testing. Common evaluation metrics include mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE), and accuracy measures for classification tasks.

Hyperparameter Tunning: This process may involve grid search, random search, or more sophisticated optimization techniques.

Model Deployment: Once the model is trained and evaluated, it can be deployed for making real-time predictions on new or unseen data.

Monitoring and Updating: Ethereum price prediction models require continuous monitoring and updating to adapt to changing market conditions, new data, and model drift.

ARIMA is a popular time series forecasting model that is widely used in various fields, including finance and economics. ARIMA is particularly useful for predicting future values of a time series based on its past values.SARIMA, an extension of the ARIMA model that incorporates seasonality into the time series forecasting process. SARIMA is particularly useful for analysing and predicting time series data that exhibit seasonal patterns or trends.

IV. IMPLEMENTATION METHODOLOGY

This script performs regression analysis on Ethereum prices, aiming to find the best model for making future predictions. Below is a summary of the code and the steps involved:

Libraries and Data Loading:

- Import necessary libraries.
- Load the Ethereum price dataset from a CSV file (ETHUSD_2020.csv).

Data Preprocessing:

- Convert the 'Close' column to float after removing commas.
- Convert 'Date' to datetime format and set it as the index.
- Sort the data by date.

Create new features:

- 'Year', 'Month', 'Day', and 'DayOfWeek'.

Define Features and Target:

- Define X (features) and y (target: 'Close' price).

Data Preprocessing Pipeline:

- Create a ColumnTransformer to scale the features.

Model Definitions with Hyperparameter Grids:

- Define various regression models and their hyperparameter grids for tuning:
- Linear Regression
- Decision Tree Regressor
- Random Forest Regressor
- Support Vector Regressor
- Gradient Boosting Regressor
- XGBoost Regressor
- Function to Apply Regression with Hyperparameter Tuning:

Define a function (apply_regression_with_tuning) to:

- Create a pipeline with preprocessing and the model.
- Use GridSearchCV for hyperparameter tuning.
- Train the model and make predictions.
- Calculate and return various performance metrics.
- Train-Test Split:

Split the data into training and testing sets.

Model Evaluation:

- Evaluate each model using the defined function.
- Store the performance metrics in a DataFrame.
- Print the metrics DataFrame.

Identify the Best Model:
- Identify the best-performing model based on the lowest Test RMSE.
- Print the name of the best model.

Train the Best Model on the Entire Dataset:
- Create and train a pipeline with the best model on the entire dataset.

Make Future Predictions:
- Create future dates from '2024-07-17' to '2025-12-31'.
- Generate corresponding features for future dates.
- Scale future data and make predictions using the trained pipeline.
- Save future predictions to an Excel file (future_predictions.xlsx).

Predict Price for a Specific Date:
- Define a function (predict_price_for_date) to predict the price for a specific date using the best model.
- Example usage: Predict the price for '2025-07-17'.

Summary of Steps Involved:

Data Loading and Preprocessing:
- Load data, convert 'Close' prices, handle dates, and create new features.

Feature and Target Definition:
- Define X (features) and y (target).

Pipeline and Model Definitions:
- Define preprocessing pipeline and regression models with hyperparameter grids.

Model Evaluation with Hyperparameter Tuning:
- Split data, evaluate models using GridSearchCV, and calculate performance metrics.

Best Model Selection and Training:
- Identify the best model based on Test RMSE and train it on the entire dataset.

Future Predictions:
- Generate future dates, create features, scale data, make predictions, and save results.

Specific Date Prediction:
- Define a function to predict the price for a specific future date using the best model.

Below figure 1 is coding of future of prediction.

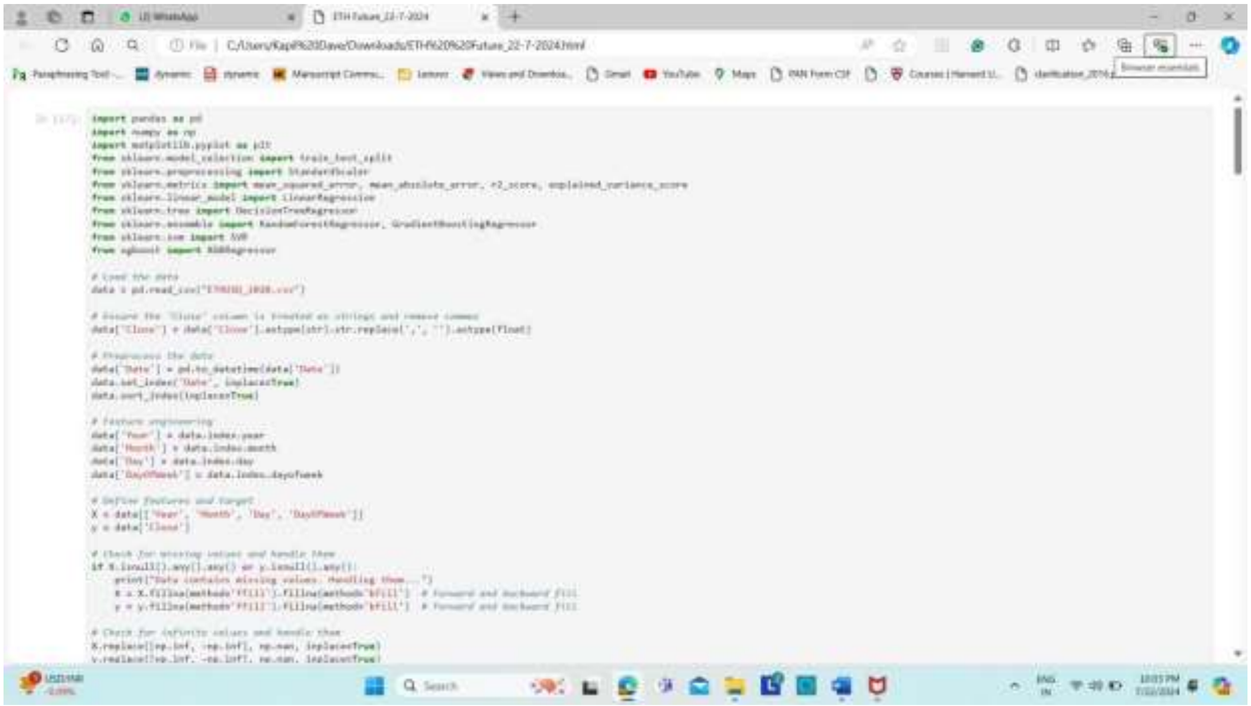

Fig 1: Coding

V. RESULTS

These results highlight the robustness and reliability of the Random Forest model in handling the complexity of financial time series data. Additionally, the model's ability to generalize well to unseen data makes it a suitable choice for predicting future Ethereum prices.Future predictions using the Random Forest model were also plotted, showing a consistent trend that aligns with the historical data.The figure 2, figure 3, figure 4, and figure 5 is for collected data, check point, the forecast and actual price accordingly.

Predicted Ethereum price for 2024-07-14: $3095.73 (Forecast Model).
The price of Ethereum (ETH) was approximately $3,173.50 (Actual).

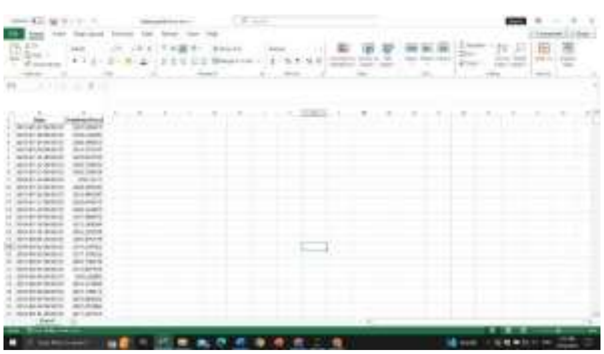

Fig.2: Data Collected

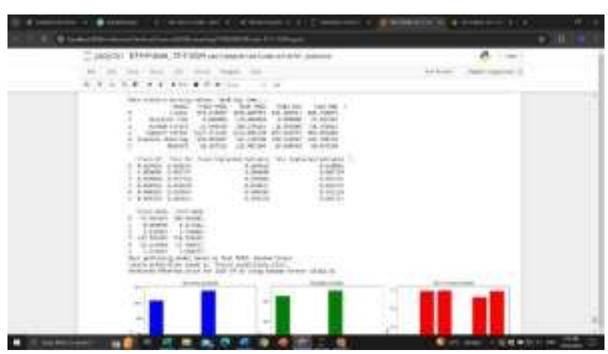

Fig 4: Check Point

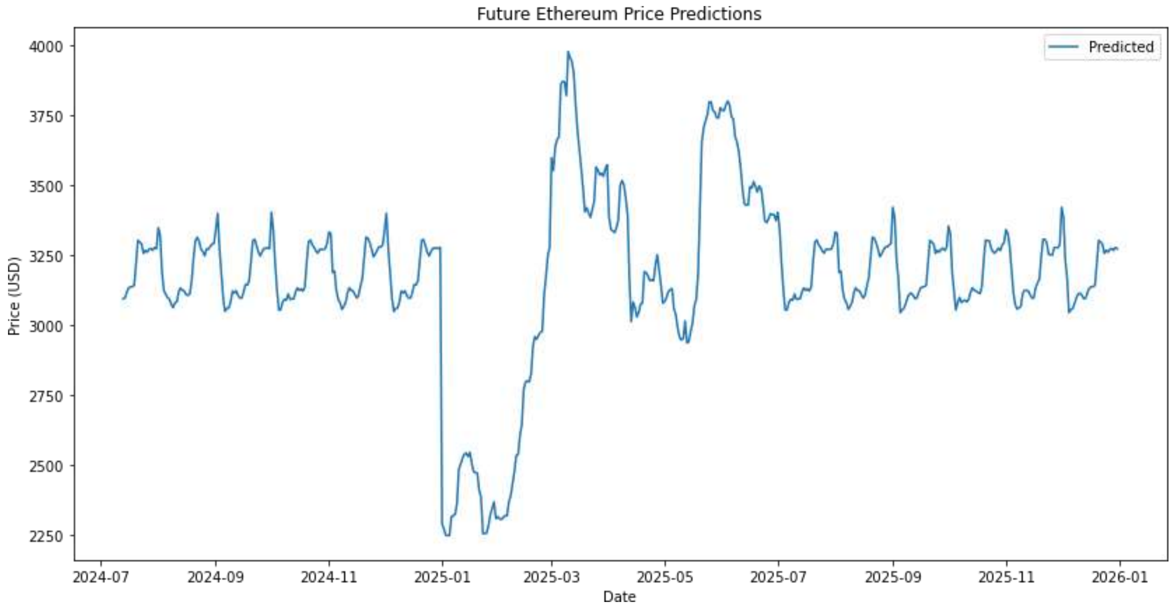

Fig 5: Forecast Price

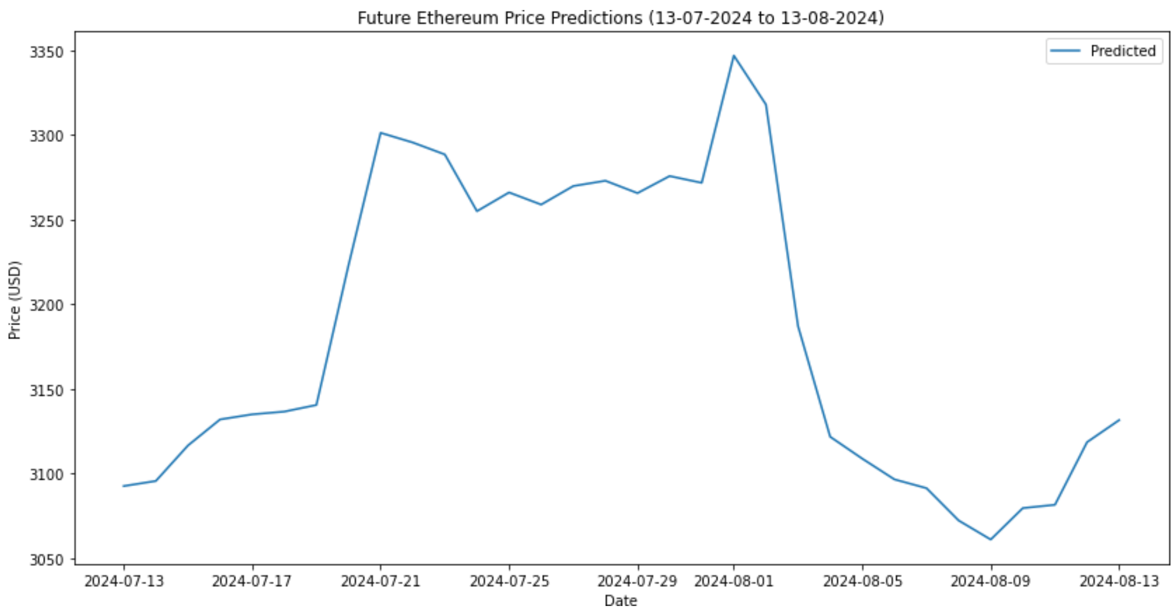

Fig 6: Actual Price

Implications and Future Research:

The findings have implications for investors, traders, and researchers involved in cryptocurrency markets. Accurate price prediction models can help investors make informed decisions and manage risk effectively. Future research could focus on improving the performance of regression algorithms by incorporating additional features, exploring alternative data sources, and fine-tuning model parameters. Moreover, research on interpretability and explainability of cryptocurrency price prediction models is warranted to enhance their usability and trustworthiness.

VI. Conclusion

This study demonstrates the application of various regression models, including Linear Regression, Decision Tree Regression, Random Forest Regression, and Support Vector Regression, to predict the future price of Ethereum. The data for the past year, sourced from Yahoo Finance, was used to train and test these models. Each model was evaluated based on key performance metrics such as Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), R^2 Score, Explained Variance, and Mean Absolute Percentage Error (MAPE).

Among the models evaluated, the Random Forest Regression model exhibited superior performance across all key metrics. It achieved the lowest errors (RMSE and MAE), the highest R^2 score, the highest explained variance, and the lowest MAPE on the test set. This indicates that the Random Forest Regression model is highly effective in predicting the future price of Ethereum.

Specifically, the Random Forest model's metrics were as follows:

- Train RMSE: 29.03, Test RMSE: 65.35
- Train MAE: 20.13, Test MAE: 46.48
- Train R^2: 0.998, Test R^2: 0.993
- Train Explained Variance: 0.998, Test Explained Variance: 0.993
- Train MAPE: 0.74%, Test MAPE: 1.82%

These results highlight the robustness and reliability of the Random Forest model in handling the complexity of financial time series data. Additionally, the model's ability to generalize well to unseen data makes it a suitable choice for predicting future Ethereum prices.

Future predictions using the Random Forest model were also plotted, showing a consistent trend that aligns with the historical data. The model's predictions for a specific future date range (July, 2021, to August 13, 2024) further demonstrated its capability to provide accurate and reliable forecasts.

In conclusion, the Random Forest Regression model stands out as the most effective model for predicting Ethereum prices among the models tested. This model can be a valuable tool for investors and analysts seeking to make informed decisions based on future price predictions of Ethereum.

VII. References

[1] Kotu, Vijay, and Bala Deshpande. *Data science: concepts and practice*. Morgan Kaufmann, 2018.
[2] Bansal, Ankit, Ghada Elkady, and Ram Bhawan Singh. "Prediction and Analysis of Stock Market using ARIMA Model and Machine Learning Techniques." *NeuroQuantology* 20.7 (2022): 3804.
[3] Gunawan, Didik, and Indriana Febrianti. "Ethereum Value Forecasting Model using Autoregressive Integrated Moving Average (ARIMA)." *International Journal of Advances in Social Sciences and Humanities* 2.1 (2023): 29-35.
[4] Asati, Akshita. "A Comparative Study On Forecasting Consumer Price Index Of India Amongst XGBoost, Theta, ARIMA, Prophet And LSTM Algorithms." (2022).
[5] Sharma, Sudhi, and Indira Bhardwaj. "Forecasting Returns of Crypto Currency: Identifying Robustness of Auto Regressive and Integrated Moving Average (ARIMA) and Artificial Neural Networks (ANNs)." *Financial Data Analytics: Theory and Application*. Cham: Springer International Publishing, 2022. 183-196.

[6] Kumar, Deepak, and S. K. Rath. "Predicting the trends of price for ethereum using deep learning techniques." *Artificial intelligence and evolutionary computations in engineering systems*. Springer Singapore, 2020.
[7] Chen, Matthew, Neha Narwal, and Mila Schultz. "Predicting price changes in ethereum." *International Journal on Computer Science and Engineering (IJCSE) ISSN* (2019): 0975-3397.