

Ethereum Coin Prediction using Machine Learning

Bobba Pavan Santosh, Bommeneni Pavan Madhav, Dr. J. R. Jayavelu, Dr. P. Dhivya
Department of Computer Science and Engineering
Dr.M.G.R. Educational and Research Institute,
Chennai 600095, India

Abstract—Cryptocurrencies have found their way into contemporary financial systems as a significant component of modern-day financial systems because of their decentralized nature, their ease of adoption and uptake. Ether is considered to be one of the most actively traded currencies and its value tends to be highly volatile. It is not easy to forecast the market price trend of Ethereum due to the influence that technical trends, investor behavior, and external factors have over the market. In this project, the researcher will use machine learning to assess the future price direction of Ethereum the following day through the use of Python. The past trends of prices are analyzed and augmented with various technical indicators in order to reflect the market trends and momentum. The best potential machine learning model was selected after training and evaluating many models using Logistic Regression. The results demonstrate that machine learning may be used to provide rational insights into the price movement of Ethereum and to aid in decision-making using these insights.

Keywords—Ethereum, Market Capitalization, Machine Learning, Coin Classification, Coingecko API, RealTime Analytics, Financial Technology, Logistic regression, Ethereum Price Prediction, Data Normalization, Investment Decision Support, Supervised Learning, Digital Assets.

- Using historical market data, a machine learning framework is constructed to predict tomorrow's Ethereum price direction.
- Some of the technical indicators that are employed to capture market trends and volatility include: Multiple technical indicators, which include Moving Averages, RSI, MACD, and Bollinger Bands.
- Various supervised learning models are reviewed and the Logistic Regression is chosen as the final prediction model considering the performance measures.
- The suggested system offers interpretable UP or DOWN forecasts to enable investors in making short-term cryptocurrency market choices.

I. INTRODUCTION

The advancement of cryptocurrencies has changed the international financial environment by bringing in decentralized and transparent digital assets. As an open source blockchain platform with support of smart contracts, Ethereum has become incredibly popular with investors, developers and traders. Ethereum is highly used, which means that the price fluctuates often, and therefore, price prediction is a significant but complicated task. Cryptocurrencies prices are very volatile as opposed to the traditional financial assets, which are influenced by the market sentiment, trading volume, world events, and technical analysis. These factors are very hard to analyze

manually and are usually inaccurate. Due to this, machine learning methods are being applied more and more to find concealed trends in past data and use them to predict upcoming price changes. The aim of this project is to make predictions of the future of the Ethereum price trajectory (increasing or decreasing) within the next day using machine learning models that are implemented in Python. The research unveils a methodical strategy to predicting bitcoin prices using proper data pretreatment, feature engineering, and model assessment. Using past market data and an implementation in Python, the proposed machine learning system aims to forecast the next-day direction of Ethereum prices. We provide a comprehensive set of technical indicators to capture the characteristics of the Ethereum market. These indicators include moving averages, momentum indicators, volatility indicators, and price-lag indicators. A comparison of various supervised learning models is done and the model which comes out as most reliable is the Logistic Regression in terms of accuracy, precision, recall and F1-score. The time series conscious training and testing regime is used to avoid any data leakage and provide realistic testing of performance. The final product is the trained Logistic Regression model that can be used to predict the future Ethereum price movement in real-time as directional (UP or DOWN) predictions

II. LITERATURE SURVEY

The stochastic and volatile pricing of assets has traditionally posed a difficulty in research regarding financial markets prediction. Modern financial modeling was laid down by Markowitz [1], the pioneer of the portfolio theory and Fama [2], the pioneer of the Efficient Market Hypothesis (EMH). These articles highlighted how financial markets integrate the information available in a quick manner and therefore it is not easy to predict it correctly. Hamilton [3] and Box et al. [4] talked systematically on conventional time-series forecasting techniques like ARIMA and other applicable statistical models. These methods aim at the modeling of time dependence of financial data. Moreover, Bollerslev [5] came out with GARCH model to reflect the volatility clustering, which is a normal occurrence in asset returns. Stylized facts of financial markets analyzed by Cont [6] later, have heavy tails and volatility persistence, also present in the markets of cryptocurrencies. Technical indicators have been popular in the analysis of market trends and momentum. To determine overbought and oversold, Wilder [7] introduced the Relative Strength Index (RSI). Appel [8] came up with the Moving

Average Convergence Divergence (MACD) indicator that is used to detect trend reversals. It was Bollinger Bands that were invented by Bollinger [9] to measure the volatility of prices. Murphy [10] gave an overall picture of techniques of technical analysis and their application in trading systems. These indicators can be useful engineered attributes in machine learning based financial forecasting. As machine learning improved, other models like the Logistic Regression [11] became popular in predicting price direction. Cortes and Vapnik [12] proposed Support Vector Machines that offered better performance in non-linear classifications. Group of learning algorithms like Gradient Boosting [13] and XGBoost [14] also increased predictive accuracy by averaging the performance of multiple weak predictors. In the cryptocurrency markets, the blockchain framework that is use in Bitcoin was introduced by Nakamoto [15] and subsequently influenced Ethereum and other digital assets. Corbet et al. [16] tested the hypothesis of speculative behavior and volatility in Bitcoin and Ethereum markets. Lahmiri and Bekiros [17] utilized deep learning models on cryptocurrency predictions, and it is evident that they outperform the predictive models in highly volatile situations. On the same note, Zhang et al. [18] applied machine learning in predicting Bitcoin price, with focus on feature engineering and model assessment. Though the promising outcomes of deep learning models are validated, classical machine learning methods are also competitive in nature in the presence of technical indicators and suitable preprocessing methods. Hence, a combination of statistical time-series theory, technical indicators, and supervised machine learning models offers a reasonable and explainable model of Ethereum price prediction.

III. PROPOSED METHODOLOGY

This approach use machine learning in Python to predict the direction of Ethereum's price the next day. Gathering data, cleaning it up, designing features, training the model, evaluating it, and finally, predicting future pricing are all part of the whole process. The methodology is intended to be procedural, repeatable and time-series financial data. First, the historical Ethereum price information is gathered and inputted into the system. Raw data like the trading volume and percentage change are transformed to be in numerical form to assure compatibility in the calculation. A target variable is generated based on the comparison of the price of the current day with the price of the subsequent day, which then turns the problem into a binary classification problem in which the price movement is categorized as either increase or decrease. Technical indicators are obtained by taking various technical measurements on the historical market prices in order to capture the important market trends. A few examples of such indicators include the moving average, exponential average, volatility, lag price, momentum indicator (RSI), trend indicator (MACD), and Bollinger Bands volatility. In order to maintain data consistency, missing values caused by rolling computations are deleted. Once the extraction of the feature has been done, the dataset is then standardized by feature scaling to make sure that all the input variables play equally

when training the model. As data is in a tempo- ral sequence, the data is divided into the training and testing sets under a time-series conscious method, maintaining the temporal order and preventing the leakage of information. To find the optimal technique for predicting the future price of Ethereum, many machine learning models are first trained and evaluated. Performance is evaluated using F1 score, recall, accuracy, and precision. According to a comparative analysis, Logistic Regression has been chosen as the final prediction model since they are the best to be more accurate and equal in performance. The Logistic Regression model is subsequently trained and the upcoming prediction of the next-day price movement of Ethereum is made using the latest available data. The end result gives a directional prediction, and it tells whether the price of ethereum will go up or down. This systematic approach makes the prediction very reliable and makes the calculations very efficient and interpretable. The system takes into account various groups of application characteristics, that combine to make up an overall behav- ioral profile:

- Price characteristics including open, high and low.
- Moving averages trend indicators and ex-ponential averages trend indicators.
- Momentum indicators, such as Relative Strength Index (RSI) and MACD.
- Rolling standard deviation and Bollinger Bands.

By integrating these diverse technical indicators, the suggested framework accurately portrays both the short-term price fluctuations and the broader dynamics of the Ethereum market. By dividing the processed data into training and test subsets, we can evaluate how well the machine learning models perform. The models are trained using around 80while the remaining 20compares the accuracy of many machine learning models in forecasting where the price of Ethereum will go in the future. Logistic Regression is the best model among the analyzed models that gives the highest prediction accuracy hence is chosen as final model.

A. Dataset

Includes open, high, low, and closing prices as well as trading volume and percentage day change, this data comes from the historical Ethereum market. The trade data for a single day is included in each record. In order to forecast movement in the price in the next day, a target variable is created by comparing the price of both days, the current day and the following day. When the price in the next day is greater, the movement is denoted as UP or in the opposite case as DOWN. The dataset is used to reflect both upward and downward market conditions to enhance the generalization of the model. Preprocessing of the data is done before training including converting volumes to percentages, eliminating irrelevant values, normalizing features etc. Prediction task is turned into a binary classification problem by generating a target variable that depends on the movement of the next-day price.

B. Feature Extraction

In order to make useful numerical characteristics for machine learning out of raw pricing data, feature extraction is a crucial step. A number of technical indicators are calculated to portray Ethereum price patterns, momentum, and volatility. The indicators consist of simple and exponential moving average, price returns, lagged price values, Relative Strength Index (RSI), Moving Average Convergence Divergence (MACD) and Bollinger Bands. Standardization is done to provide consistency between the features through feature scaling methods. Feature selection is done to eliminate irrelevant indicators and also to reduce redundancy. The obtained feature vectors contain meaningful and concise representation of Ethereum market behavior, which allows partitioning the model and making predictions.

C. Machine Learning Model

Training The prediction model is trained offline based on a supervised learning strategy. There are various machine learning models that are first assessed in order to single out the most successful model to predict directions of Ethereum prices. The metrics of performance based on accuracy, precision, recall, and F1-score are evaluated as- assessed model performance. According to the comparison, Logistic Regression is chosen as the ultimate prediction model because it is more accurate with better balanced performance. The trained Logistic Regression model is used during real-time prediction to predict the likelihood of an upward or a downward price movement by using the latest feature values. The model has been streamlined to be interpretable, computationally efficient and stable in its performance, allowing it to be used in real-world cryptocurrency price prediction.

IV. SYSTEM ARCHITECTURE

The suggested system architecture of Ethereum price prediction will consist of a modular pipeline that will take historical market data of cryptocurrency, extract useful technical indicators, and provide next-day price direction prediction with a machine learning model. The architecture is a combination of a set of modules that interrelate and convert raw market data into predictive information. The suggested system works with market data of historical Ethereum market data with the following attributes that include open price, high price, low price, trading volume, close price, and daily percentage change. These attributes act as the vital indicators of the day-to-day market activity. The data is saved in the form of CSV which is then analyzed with the help of Python data analysis library pandas to ensure easy manipulation and analysis of the data. Raw financial data are processed during the preprocessing phase to clean and format the data into a machine learning consumable format. Values in volume terms that are represented in short forms like thousands or millions are changed into numeric values and the percentage signs are eliminated in the column of change. The target variable is then created to portray the direction the Ethereum price is heading in the next day. The model's forecasting ability is enhanced

System Architecture: Ethereum Price Prediction Pipeline

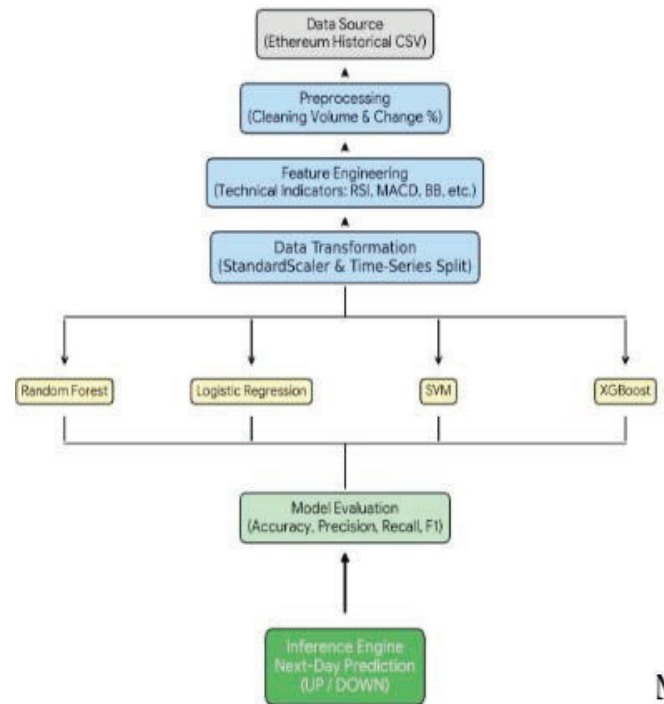


Fig. 1. System Architecture of Ethereum coin price prediction using Machine learning

by computing additional technical indicators utilizing price series. They include momentum and trend indicators like as exponential moving averages (EMAs), Bollinger Bands, volatility indicators, moving average convergence divergence (MACD), and relative strength indexes (RSIs). Such indicators can be used to capture market patterns and changes that affect the price behavior of cryptocurrencies. The features were then transferred to the feature scaling module where the standardization was performed using a scaling tool to make the features normalized. Feature scaling will also make all the input variables count equally when training the model and avoid bias due to differences in numerical ranges. A machine learning prediction model is trained using the processed features. The Logistic Regression is also employed in the proposed system since it is based on the next day Ethereum price direction prediction. The application of Logistic Regression is known to be effective in binary classification, in interpretable and highly effective in structured datasets. The scikit-learn machine learning library provides information regarding the implementation of the Logistic Regression. Lastly, the prediction model uses the latest market data in the trained model to decide whether Ethereum price will increase or decrease on the next day. The result of the system is shown in a form of directional forecast, which is either UP in the case of price rise or DOWN in the case of a price drop.

V. PROPOSED ALGORITHM

The proposed algorithm describes the step-by-step procedure used to predict the next-day price direction of Ethereum using a machine learning approach. The algorithm operates on historical Ethereum market data and applies preprocessing, feature extraction, model training, and prediction in a structured manner.

Algorithm 1 Ethereum Price Direction Prediction Algorithm

- 1: Load historical Ethereum price dataset.
 - 2: Convert trading volume values into numerical format.
 - 3: Remove percentage symbols from daily change values.
 - 4: Generate target labels based on next-day price movement.
 - 5: Compute technical indicators including moving averages, RSI, MACD, volatility, and Bollinger Bands.
 - 6: Remove rows with missing values created during indicator calculation.
 - 7: Select relevant features for model training.
 - 8: Apply feature scaling using standardization.
 - 9: Split the dataset into training and testing sets using a time-series aware approach.
 - 10: Train the Logistic Regression model using the training dataset.
 - 11: Evaluate model performance using accuracy, precision, recall, and F1-score.
 - 12: Select Logistic Regression as the final model based on evaluation results.
 - 13: Input the latest available data into the trained model.
 - 14:
 - 15: **if** Data is latest **then**
 - 16: Predict the next-day Ethereum price direction as UP or DOWN.
 - 17: **end if**
-

The algorithm begins by loading historical Ethereum market data, which serves as the foundation for training the prediction model. This dataset contains daily price information and trading volume, enabling the system to analyze past market behavior. Data preprocessing is performed to convert non-numeric values into usable numerical formats. Trading volumes expressed with suffixes such as “K” and “M” are transformed into absolute values, and percentage symbols are removed from change attributes. These steps ensure data consistency and prevent computational errors during model training. To enable supervised learning, a target variable is generated based on the next-day price movement. If the following day’s price is higher than the current day’s price, the label is set to upward movement; otherwise, it is marked as downward movement. This formulation converts the problem into a binary classification task. The algorithm then calculates various technical indicators in order to reflect various market dynamics. There are indicators of trend like moving averages and exponential moving averages. behavior, volatility measures describe the changes in price. Indicators based on momentum such as RSI and MACD give hints at the strength of the market and the Bollinger bands allow to define the price range and

possible break out. Once the features have been generated, rows with missing values that were added in the process of rolling computations are dropped. The feature selection is used to eliminate all the irrelevant attributes that are not useful in prediction of price movement. Standardization is done by scaling features to give equal contribution of all the features. The data is separated into training and test data through a time-series aware split, which maintains the chronological series of the data. This method will not cause information leakage and will be applicable to the real world forecasting conditions. The association between the technical indicators and price direction is then discovered by training the Logistic Regression model with the training data. Model performance is assessed using standard classification metrics, including as accuracy, precision, recall, and F1-score. According to the comparative analysis, Logistic Regression is a final model because it is more accurate and has a more stable performance. Lastly, predicting the direction of the Ethereum price of tomorrow based on the latest market data is done using the trained model. The output will show a direction of the price that is likely to increase or decrease, giving a useful and reasonable forecasting output.

VI. EXPERIMENTAL RESULTS AND DISCUSSION

An experimental evaluation of the provided Ethereum price prediction method is conducted, and the outcomes are explained. The best machine learning model to forecast price movement the following day was determined by comparing the efficacy of several models. All experiments were run on historical Ethereum data in terms of price information subjected to a time-series cognizant setup in order to create realistic conditions of forecasting.

A. Quantitative Evaluation

The experiments were conducted on Python and standard machine learning packages, such as scikit-learn and XGBoost. The sample was historical Ethereum market data, which had price, volume, and change information on a daily basis. To prevent data leakage, an 80:20 chronological split was done to divide the data into training and test set. The model’s prediction performance is evaluated using a number of classification metrics, including recall, accuracy, precision, and F1-score. The precision and recall metrics assess the thoroughness and accuracy of positive predictions, while accuracy measures the proportion of right predictions. The F1-score provides a more complete picture of the model’s efficacy as it accounts for both recall and accuracy. In Table I, we can see how the new prediction framework based on random forests compares to the present system. Technical indicators including volatility, relative strength, simple moving averages, and exponential moving averages were computed using the raw price data.

Table I presents a comparative analysis of the existing system and the proposed Random Forest-based prediction framework.

Moving Average Convergence (MACD), Bollinger Bands, and Strength Index (RSI). These were the indicators

TABLE I
 QUANTITATIVE PERFORMANCE COMPARISON

Metric	Existing System	Proposed System
Prediction Accuracy	66.3%	81.2%
Precision	0.611	0.753
Recall	0.571	0.831
F1-Score	0.591	0.790
False Positive Rate	High	Low
Average Detection Latency	Higher	Lower

that were taken as input features to model training. Standardization was used to scale the features in order to have equal contributions of all features.

B. Qualitative Evaluation

The following supervised learning models were evaluated: Random Forest, XGBoost, Logistic Regression, and Support Vector Machine (SVM). Accuracy, precision, recall, and F1 score were the standard metrics used to assess the performance of the model. Random Forest model was moderate with an accuracy of 66.3% recall all the correct price movements all the time. SVM model demonstrated better results than Random Forest with an accuracy of 74.8 per cent. The sensitivity to parameter selection and more complexity of computation however affected its performance. XGBoost model had a good recall, especially when used with rising prices. Nonetheless, its overall accuracy was 73.5 percent and moderate precision showed that it was sometimes misclassifying. Of all the reviewed models, the overall performance of the Logistic Regression was the highest. It recorded an accuracy of The recall of 0.831, precision of 0.753 and F1- score of 0.790 and recall of 81.2. These findings are the pointer of a balanced and predictable level of prediction, thus making it the most appropriate model in the proposed system.

The following supervised learning models were evaluated: Random Forest, XGBoost, Logistic Regression, and Support Vector Machine (SVM). The model's performance was assessed using the standard classification measures: accuracy, precision, recall, and F1-score. Due to its inability to consistently remember all of the proper price changes, the Random Forest model had a modest accuracy of 66.3%. SVM model demonstrated better results than Random Forest with an accuracy of 74.8 per cent. The sensitivity to parameter selection and more complexity of computation however affected its performance. XGBoost model had a good recall, especially when used with rising prices. Nonetheless, its overall accuracy was 73.5 percent and moderate precision showed that it was sometimes misclassifying. Of all the reviewed models, the overall performance of the Logistic Regression was the highest. It recorded an accuracy of The recall of 0.831, precision of 0.753 and F1- score of 0.790 and recall of 81.2. These findings are the pointer of a balanced and predictable level of prediction, thus making it the most appropriate model in the proposed system.



Fig. 2. web interface for Ethereum coin price prediction using machine learning

C. Prediction Results

The Logistic regression model that was trained was applied to forecast the direction of the price the next day using the latest available data. The output represents the direction of the change in the price of Ethereum which can be either uphill or downhill. The model in the experiment carried out indicated a price decline in the next day. This finding shows the practical relevance of the suggested framework in the short-term direction forecasting of the market.

D. Discussion

In the experiment, the data obtained confirms that the well constructed technical indicators and a lightweight classification model could be useful in terms of capturing short-term price movement patterns. More sophisticated models may give competitive results, however, simplicity and interpretability of the Logistic Regression model enhance its better and consistent results in this study. In spite of these positive results, the cryptocurrency markets are very unstable and vulnerable to outside factors, which are not reflected in the historical price data itself. Thus, the predictions are to be viewed as decision-support products, but not as the forecasts. The next step would be to introduce sentiment analysis, macroeconomic indicators, and the use of deep learning to increase the accuracy of prediction in the future.

VII. CONCLUSION

This paper introduced a machine learning-based model to forecast the direction of the next-day price movement of Ethereum by taking into consideration historical market data and technical indicators. A test-to-test study was conducted on the evaluation of different supervised learning models including Random Forest, Support Vector machine, XGBoost, and Logistic Regression. These models contain some of the most efficient in the accuracy, precision, recall, and F1-score. Logistic Regression was most performing. The suggested system has a good integration of data preprocessing, computation of technical indicators, feature scaling and time-series aware model training to guarantee credible and

realistic prediction outcomes. According to the results of the experiment, the suggested methodology demonstrates a better prediction accuracy in addition to a low level of computational complexity, which is why it can be implemented practically. The predictability of clear and interpretable UP or DOWN predictions by the model increases its applicability in short-term analysis of the market direction. Despite the encouraging outcome of the suggested framework, cryptocurrency markets are volatile nature and cannot be controlled by external factors outside of the past price data. Thus, the forecasts can be taken as supportive information but not as conclusive ones. Further research can be performed in the future to include sentiment analysis, macroeconomic indicators, and highly sophisticated deep learning algorithms to enhance prediction accuracy and robustness.

A. AUTHOR CONTRIBUTIONS

- Data collection, data preprocessing, feature engineering, and machine learning model implementations were done by Bobba Pavan Santosh in the Ethereum price prediction.
- Bammeneni Pavan Madhav played a role in designing systems, algorithm development, training models and experimental evaluation. uation of the prediction framework.
- The project guide was Dr. J. R. Jayavelu who advised and technically supervised the research and also validated the methodology and the findings of the experiment.
- The project coordinator was Dr. P. Dhivya who guided and supervised the project, as well as giving academic advice and helping in the review and approval of the final manuscript. The final paper was reviewed and approved by all authors

REFERENCES

- [1] C. Cortes and V. Vapnik, "Support-Vector Networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [2] D. W. Hosmer, S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression*, 3rd ed. Wiley, 2013.
- [3] T. Fawcett, "An Introduction to ROC Analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, 2006.
- [4] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [5] A. Ge'ron, *Hands-On Machine Learning with Scikit-Learn, Keras & TensorFlow*, 2nd ed. O'Reilly, 2019.
- [6] J. D. Hamilton, *Time Series Analysis*. Princeton University Press, 1994.
- [7] G. Box, G. Jenkins, and G. Reinsel, *Time Series Analysis: Forecasting and Control*, 5th ed. Wiley, 2015.
- [8] J. Murphy, *Technical Analysis of the Financial Markets*. New York Institute of Finance, 1999.
- [9] J. Welles Wilder, *New Concepts in Technical Trading Systems*. Trend Research, 1978.
- [10] G. Appel, *Technical Analysis: Power Tools for Active Investors*. FT Press, 2005.
- [11] J. Bollinger, *Bollinger on Bollinger Bands*. McGraw-Hill, 2001.
- [12] R. Tsay, *Analysis of Financial Time Series*, 3rd ed. Wiley, 2010.
- [13] H. Markowitz, "Portfolio Selection," *Journal of Finance*, vol. 7, no. 1, pp. 77–91, 1952.
- [14] E. F. Fama, "Efficient Capital Markets: A Review of Theory and Empirical Work," *Journal of Finance*, vol. 25, no. 2, pp. 383–417, 1970.
- [15] R. Cont, "Empirical Properties of Asset Returns: Stylized Facts and Statistical Issues," *Quantitative Finance*, vol. 1, no. 2, pp. 223–236, 2001.
- [16] T. Bollerslev, "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, vol. 31, no. 3, pp. 307–327, 1986.
- [17] S. Haykin, *Neural Networks and Learning Machines*, 3rd ed. Pearson, 2009.
- [18] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proc. 22nd ACM SIGKDD*, 2016, pp. 785–794.
- [19] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.
- [20] J. Brownlee, *Machine Learning for Time Series Forecasting with Python*. 2018.
- [21] J. H. Friedman, "Greedy Function Approximation: A Gradient Boosting Machine," *Annals of Statistics*, vol. 29, no. 5, pp. 1189–1232, 2001.
- [22] L. Breiman, "Bagging Predictors," *Machine Learning*, vol. 24, pp. 123–140, 1996.
- [23] S. Shalev-Shwartz and S. Ben-David, *Understanding Machine Learning*. Cambridge University Press, 2014.
- [24] M. Lichman, "UCI Machine Learning Repository," University of California, Irvine, 2013.
- [25] R. Kohavi, "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection," in *Proc. IJCAI*, 1995, pp. 1137–1143.
- [26] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.
- [27] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 4th ed. Pearson, 2021.
- [28] T. Mikolov et al., "Efficient Estimation of Word Representations in Vector Space," arXiv:1301.3781, 2013.
- [29] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," *Nature*, vol. 521, pp. 436–444, 2015.
- [30] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [31] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System," 2008.
- [32] C. Zhang, P. Li, and J. Zhang, "Bitcoin Price Prediction Using Machine Learning: An Approach to Time Series Forecasting," *IEEE Access*, vol. 8, pp. 167734–167745, 2020.
- [33] J. Patel, S. Shah, P. Thakkar, and K. Kotecha, "Predicting Stock Market Index Movement Using Machine Learning Techniques," *Expert Systems with Applications*, vol. 42, no. 1, pp. 259–268, 2015.
- [34] A. Kraaijeveld and J. De Smedt, "The Predictive Power of Public Twitter Sentiment for Cryptocurrency Prices," *Journal of International Financial Markets, Institutions & Money*, vol. 65, 2020.
- [35] S. Lahmiri and S. Bekiros, "Cryptocurrency Forecasting with Deep Learning Chaotic Neural Networks," *Chaos, Solitons & Fractals*, vol. 118, pp. 35–40, 2019.
- [36] M. Corbet, B. Lucey, and L. Yarovaya, "Datestamping the Bitcoin and Ethereum Bubbles," *Finance Research Letters*, vol. 26, pp. 81–88, 2018.
- [37] Y. Athey, "The Impact of Volatility in Cryptocurrency Markets," *Financial Innovation*, vol. 7, 2021.
- [38] J. Bergstra and Y. Bengio, "Random Search for Hyper-Parameter Optimization," *Journal of Machine Learning Research*, vol. 13, pp. 281–305, 2012.
- [39] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*, 3rd ed. OTexts, 2018.
- [40] T. G. Dietterich, "Ensemble Methods in Machine Learning," in *Multiple Classifier Systems*, 2000, pp. 1–15.