EdinBurg Peer Reviewed Journals and Books Publishers
Journal of Information and Technology
Vol. 5||Issue 2||pp 44-52||April||2025
Email: info@edinburgjournals.org|

A Comparative Analysis of Ensemble-Based Models for Predicting Cryptocurrency Price Movements

Valens Nkurunziza [1] & Nyesheja Muhire Enan[2]
[1,2]Faculty of Computing and Information Sciences, University of Lay Adventists of Kigali
Corresponding Email: vnziza@gmail.com; nyenani@gmail.com

Accepted: 01 April 2025 || Published: 14 April 2025

Abstract

This study, "A Comparative Analysis of Ensemble-Based Models for Predicting Cryptocurrency Price Movements," evaluates ensemble machine learning models bagging, boosting, and stacking to improve cryptocurrency price prediction accuracy. Using historical data, models like Random Forest, Gradient Boosting, and Stacking were tested, with Stacking emerging as the top performer (81.80% accuracy, 81.49% F1-score, 88.43% AUC-ROC), outperforming traditional methods like Naive Bayes and Decision Trees. The Boosting Combined model also showed strong results. The research highlights the effectiveness of ensemble techniques in handling cryptocurrency market volatility, offering valuable insights for traders and investors. It underscores the potential of advanced feature engineering and real-time testing to further enhance predictive accuracy, advancing financial decision-making and risk management in the cryptocurrency sector.

Keywords: *Stacking, Bagging, Boosting, Machine Learning, Cryptocurrency*

How to Cite: Nkurunziza, V., & Nyesheja, M. E. (2025). A Comparative Analysis of Ensemble-Based Models for Predicting Cryptocurrency Price Movements. *Journal of Information and Technology, 5*(2), 44-52.

1. Introduction

The finance industry has seen the rise of digital currencies, with accurate price predictions being crucial for traders, researchers, and regulators. However, predicting cryptocurrency prices is challenging due to their volatile and complex nature (Pintelas et al., 2020). Blockchain technology, the backbone of cryptocurrencies like Bitcoin, ensures secure, transparent, and decentralized transactions through cryptographic methods and consensus protocols like Proof of Work (PoW) and Proof of Stake (PoS). This technology not only supports digital currencies but also has applications in banking, logistics, healthcare, and more, offering secure and tamper-proof record-keeping. Cryptocurrencies, stored in digital wallets, are used for transactions, investments, and wealth preservation (Chaudhary & Sushil, 2023). Despite their potential, the market's volatility, driven by factors like market immaturity and irrational behavior, makes price prediction difficult (Omole & David, 2024).

Traditional statistical methods often fall short, prompting the use of advanced artificial intelligence (AI) and machine learning (ML) techniques. Studies have explored various models, including deep learning (CNN, LSTM, BiLSTM), ensemble methods (Random Forests, Gradient Boosting), and hybrid approaches, to improve prediction accuracy (Pintelas et al., 2020; Derbentsev et al., 2020). Recent research highlights the effectiveness of ensemble models, such as stacking, which combines multiple algorithms to enhance performance. For instance, stacking models have achieved high accuracy (81.80%) and AUC-ROC scores (88.43%), outperforming traditional methods like Naive Bayes and Decision Trees (Chaudhary & Sushil, 2023). Other techniques, such as K-Nearest Neighbors (KNN) and Support Vector Machines (SVM), also show promise but face challenges like computational inefficiency and sensitivity to data changes (MyEducator, 2024; Gulati et al., 2022). This research addresses these challenges by comparing the efficiency of ensemble learning strategies bagging, boosting, and stacking for forecasting cryptocurrency price trends. The goal is to identify the most precise and reliable approach, leveraging advanced feature engineering and real-time data to improve predictive accuracy. By doing so, the study seeks to provide actionable insights for traders and investors, enhancing decision-making and risk management in the volatile cryptocurrency market.

2. Literature Review

Cutting-edge artificial intelligence (AI) and deep learning (DL) methods have shown significant potential in financial forecasting, including cryptocurrency price prediction. However, the unpredictable and volatile nature of cryptocurrency markets poses challenges, often leading to overfitting and reduced model generalization (Wu et al., 2024). Ensemble learning techniques, such as bagging, boosting, and stacking, have emerged as effective alternatives, outperforming standalone models in time series forecasting (Torgo & Mariana, 2014). For instance, Basher and Perry (2022) demonstrated that random forests achieved 75-85% accuracy in predicting Bitcoin and gold prices, with technical indicators being the most influential factors. Deep learning models, particularly Long Short-Term Memory (LSTM), have also been widely explored. Ammer & Theyazn H. (2022) used LSTM to forecast prices of cryptocurrencies like Ethereum and XRP, achieving low prediction errors (MSE, RMSE). Similarly, Kunpeng et al. (2024) compared ensemble methods like LightGBM with deep learning models, highlighting their effectiveness in predicting Bitcoin, Ethereum, and other digital assets. However, these studies often focus on standalone models, neglecting hybrid or ensemble approaches that could enhance accuracy.

Research by Gudavalli and Khetan Venkata (2023) evaluated models like Random Forest (RF), Gradient Boosting (GB), LSTM, and Gated Recurrent Unit (GRU) using historical data from 2017 to 2023. While these models showed promise, the study did not explore ensemble methods, which could improve performance. Tripathy et al. (2024) compared neural network frameworks like Bi-LSTM and Facebook Prophet, finding Bi-LSTM to be the best-performing model with low MAE and RMSE values. Despite these advancements, overfitting and limited generalization remain key challenges. Several gaps persist in cryptocurrency price forecasting. First, there is limited research on ensemble learning approaches like stacking and boosting, with most studies focusing on standalone models such as RF, GB, and LSTM, which are prone to overfitting. Second, many studies rely solely on technical indicators and historical price data, ignoring alternative data sources like social media sentiment, on-chain data, and macroeconomic factors that could improve accuracy. This research aims to address these gaps by comparing ensemble learning models, exploring diverse feature selection strategies, and evaluating their performance against deep learning techniques. By doing so, it seeks to enhance the accuracy and reliability of cryptocurrency price predictions, providing actionable insights for traders and investors.

3. Methodology

3.1 Study Area and Data Sources

This study focuses on the data downloaded from kaggle.com whose data will be taken and fit into different ensemble machine learning algorithms to meet our specific objectives. The dataset has 9975 rows and 9 columns, Name of the currency, Date - date of observation, Open - Opening price on the given day, High - Highest price on the given day, Low - Lowest price on the given day, Close - Closing price on the given day, Volume - Volume of transactions on the given day and Market Cap - Market capitalization in USD respectively. The data are recorded in a CSV file. It contains information on the top cryptocurrencies such as Binance Coin, Cosmos, Crypto.com Coin, Dogecoin, Litecoin, and USD Coin from 2013 to 2021. Furthermore, the data spans 8 years, from 2023 to 2021.

3.2. Development technologies

While implementing the ensemble machine learning models in forecasting cryptocurrency prices to determine the most precise and reliable model for the task, different technologies will be used. CSV file that contains cryptocurrency information and Python, which is a versatile, widely used programming language that will facilitate the implementation of ensemble learning techniques and the visualization of outcomes. K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Ridge Classifier, Decision Tree Classifier, Multi-Layer Perceptrons (MLP) Classifier, and Gaussian Naive Bayes (NB) were used to build the staking ensemble machine learning model.

When it comes to choosing the ideal programming language for the particular objectives of the research, Python has grabbed the lead. One of the most widely used programming languages right now is Python. Along with C++, Java, and other languages, it was developed by Guido Van Rossum in 1991 and is now among the most widely used (Insights, 2016).

3 3. Model Performance Evaluation

Performance metrics like Accuracy (Bishop, (2006).) F1 Score (Hand, (2020)), Mathew's correlation coefficient (MCC) (Chicco, 2020) Area Under Curve (AUC) (Flach, 2021), and precision (Sokolova, 2020) were used to measure the best-performing ensemble model.

4. Results and discussion

4.1. Introduction

The focus was on assessing the performance of various machine learning models using metrics like accuracy, F1-score, precision, recall, and Matthews Correlation Coefficient (MCC). The

EdinBurg Peer Reviewed Journals and Books Publishers
Journal of Information and Technology
Vol. 5||Issue 2||pp 44-52||April||2025
Email: info@edinburgjournals.org|

outcomes are displayed using both graphical and tabular formats by the established methodology. Testing was conducted on an HP laptop featuring 500GB of storage, 16GB RAM, and an Intel Core i7-8250U processor running at 2.00GHz, with the system supporting both Windows 11 (64-bit) and Ubuntu 24.10 (64-bit). The analysis involved using Support Vector Machine (SVM), Linear Discriminant Analysis (LDA), and Naïve Bayes (NB) models to forecast student sentiment based on the available data.

4.2. Comparison of individual Classifiers and ensemble models

Traditional models like **MLP, Decision Tree**, and **SVM** showed moderate performance, with MLP achieving the third-highest accuracy of **81.55%** and a strong AUC-ROC of **87.41%**. However, these models were outperformed by the ensemble methods, indicating their limitations in handling complex trends within the dataset. Among all models, the Naive Bayes classifier performed the weakest, achieving the lowest accuracy of 63.31%, precision of 57.47%, F1-score of 72.50%, MCC of 37.81%, and AUC-ROC of 75.40%. This suggests that Naive Bayes is not well-suited for this dataset, further emphasizing the superiority of ensemble methods for achieving high predictive accuracy and robustness.

Table 1: Summary of Models' Performance Metrics

Model	Accuracy (%)	Precision (%)	F1-Score (%)	MCC (%)	AUC-ROC (%)
knn	79.00	80.39	78.07	58.04	86.48
naive_bayes	63.31	57.47	72.50	37.81	75.40
mlp	81.55	82.54	80.91	63.13	87.41
decision_tree	79.15	79.50	78.60	58.29	82.91
svm	71.53	65.66	75.39	46.09	85.46
bagging_combined	81.05	83.35	80.00	62.24	NaN
boosting_combined	81.75	82.27	81.26	63.51	88.39
stacking	81.80	81.70	81.49	63.60	88.43

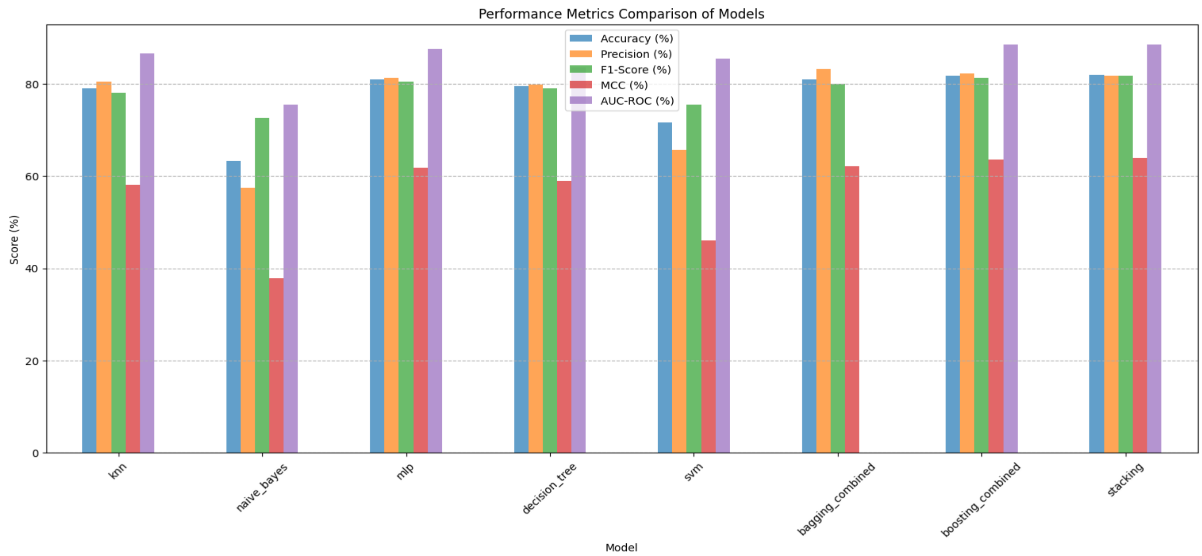


Figure 1: Performance Metric Comparison of Models

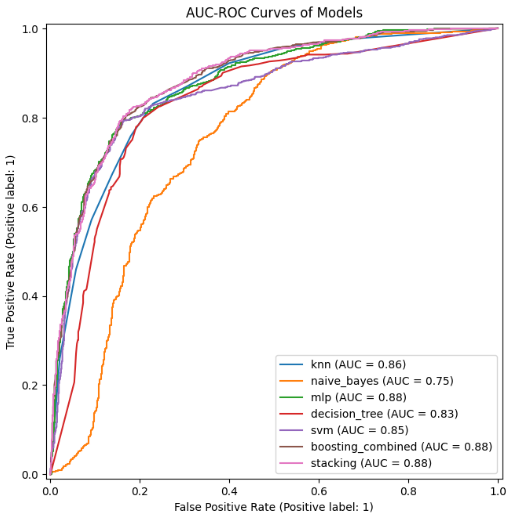


Figure 2: Area Under Curve for all Models

Achieving the highest accuracy of 81.80%, F1-score of 81.49%, Matthews Correlation Coefficient (MCC) of 63.60%, and Area Under the Curve (AUC-ROC) of 88.43%, the Stacking model was identified as the most effective performer. This demonstrates the effectiveness of combining multiple models to leverage their strengths, resulting in superior predictive performance. The stacking approach outperformed all other models, highlighting its robustness and reliability for complex classification tasks. Its ability to balance precision and recall, along with its strong class separation capability, makes it the preferred choice for this dataset.

The **Boosting Combined model** also performed exceptionally well, getting 81.75% accuracy**,** an F1-score of **81.26%**, and an AUC-ROC of **88.39%.** While it slightly trailed the stacking model, it outperformed the bagging combined model, emphasizing the strength of boosting techniques in improving model accuracy and generalization. The **Bagging Combined model** (Random Forest + Extra Trees) achieved the highest precision of **83.35%,** indicating its ability to minimize false positives. However, its accuracy of **81.05%** and F1-score of **80.00%** were slightly lower than those of the boosting and stacking models, suggesting that while it performs well in certain aspects, it is not as balanced as the other ensemble methods. The performance metrics for the stacking model, which outer outperformed other models, revealed the following.

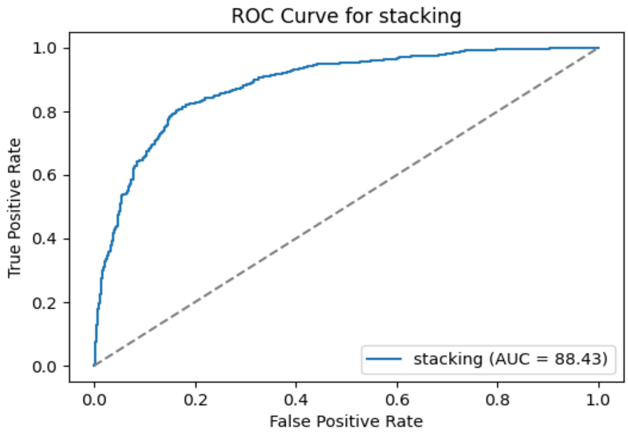
EdinBurg Peer Reviewed Journals and Books Publishers
Journal of Information and Technology
Vol. 5||Issue 2||pp 44-52||April||2025
Email: info@edinburgjournals.org|

Table 2: Classification Report for Stacking Ensemble Model

Classification Report for stacking:

	precision	recall	f1-score	suppo
Class 0	0.82	0.82	0.82	10
Class 1	0.82	0.81	0.81	9
accuracy			0.82	19
macro avg	0.82	0.82	0.82	19
weighted avg	0.82	0.82	0.82	19

Figure 3: Area Under Curve for Staking Ensemble model



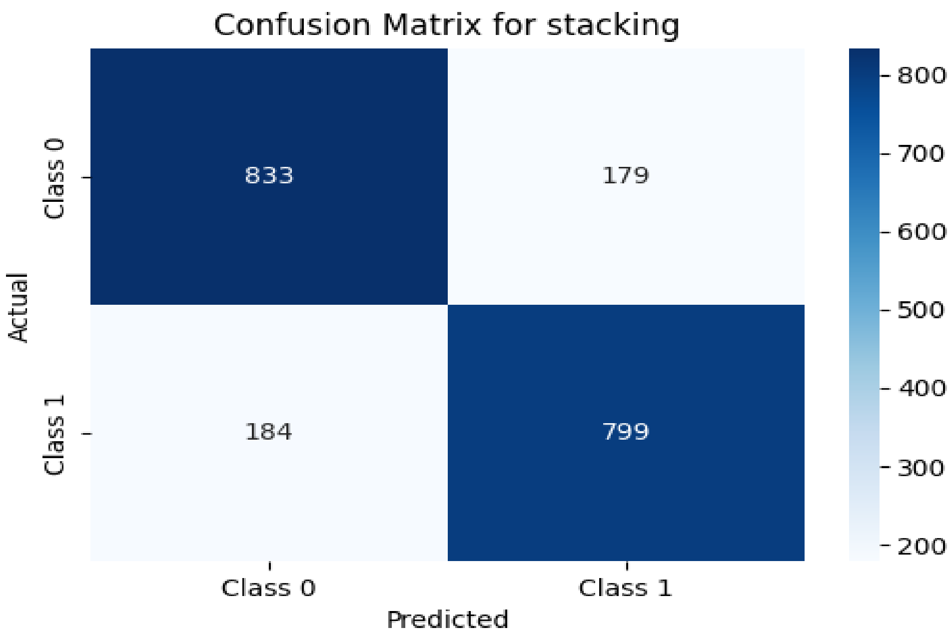


Figure 4: Confusion Matrix for Stacking Ensemble model

5. Conclusion

The study evaluated the performance of various machine learning algorithms and ensemble methods by comparing their performance in forecasting Bitcoin price fluctuations. The results revealed that ensemble techniques outperformed traditional classifiers such as Naive Bayes, Decision Trees, and Support Vector Machines (SVM), with the Stacking model standing out as the most effective. By leveraging the strengths of multiple base models, the Stacking approach achieved superior performance, with an accuracy of 81.80%, F1-score of 81.49%, Matthews Correlation Coefficient (MCC) of 63.60%, and Area Under the Curve (AUC-ROC) of 88.43%. The Boosting Combined model (which incorporates AdaBoost and Gradient Boosting) also showed impressive results, with an accuracy of 81.75% and AUC-ROC of 88.39%, highlighting the power of boosting techniques in improving model generalization.

However, conventional models such as Naive Bayes fared poorly, achieving the lowest accuracy 63.31%) and AUC-ROC 75.40%), demonstrating their inability to handle the intricacy of bitcoin data. Even if models like Random Forest and Multi-Layer Perceptron (MLP) performed moderately, the ensemble approaches still outperformed them. These results highlight the value of ensemble approaches, which successfully balance precision, recall, and class separation skills, for precise and trustworthy predictions of bitcoin price movement. The research demonstrates how ensemble learning may be used to overcome the difficulties associated with predicting bitcoin prices, including their high volatility and intricate patterns. Ensemble methods like stacking and boosting combine the advantages of several models to offer a strong foundation for raising prediction accuracy and dependability. By providing insightful information for traders, investors, and academics looking to create cutting-edge tools for cryptocurrency market analysis, the research complements the expansion of corpus of studies in financial artificial intelligence.

6. Recommendation

Advanced ensemble models, such as stacking and boosting, are suggested to enhance Bitcoin price predictions since they successfully capture intricate patterns. Models (like LSTM, Random Forest) and a meta-learner are combined in stacking, and predictions are iteratively improved using boosting (like AdaBoost, and Gradient Boosting). Accuracy will be further increased by addressing class imbalance with methods like SMOTE and improving data quality through feature engineering (e.g., technical indicators, and social media sentiment). To capitalize on their advantages, future studies should investigate hybrid strategies that combine classical models (like XGBoost) and deep learning (like LSTM). To ensure practical applicability, models will be validated using current Bitcoin data (e.g., 2022–2024) and projections will be compared to actual trading strategies (e.g., buy-and-hold). Furthermore, enhancing interpretability with the use of programs like SHAP (SHapley Additive exPlanations) can give traders useful information about the main factors influencing price changes. The precision, dependability, and usefulness of Bitcoin price projections will all be improved by these tactics.

References

1. Bishop, C. M. ((2006).). *Pattern Recognition and Machine Learning.* Springer.
2. Chicco, D. &. (2020). *The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation.* BMC Genomics, 21(1), 6.
3. Flach, P. A.-O. (2021). *Classifier performance evaluation with ROC curves: A survey. ACM Computing Surveys, 54(8), 1-35.*
4. Hand, D. J. ((2020)). *An interpretable transformation of the F-measure. arXiv preprint arXiv:2008.00103.*
5. https://library.fiveable.me/key-terms/technology-and-engineering-in-medicine/classification-accuracy, F. (. (March 15, 2025,). *Classification accuracy*. Fiveable. (n.d.). Classification accuracy.
6. Sokolova, M. &. (2020). *Systematic analysis of performance measures for classification tasks. Information Processing & Management, 57(2), 102138.*
7. Pintelas, E., Ioannis , E., Stavros , S., Theodore , K., & Panagiotis , P. (2020, May). Investigating the Problem of Cryptocurrency Price Prediction: A Deep Learning Approach. *https://pmc.ncbi.nlm.nih.gov/*. Retrieved December 03, 2024, from https://pmc.ncbi.nlm.nih.gov/articles/PMC7256561/
8. Gulati, K., Saravana, K. S., Raja Sarath, k., & Ketan, S. (2022, February). Comparative analysis of machine learning-based classification models using sentiment classification of tweets related to COVID-19 pandemic. *ResearchGate*. doi:10.1016/j.matpr.2021.04.364
9. Gudavalli, H. N., & Khetan Venkata, R. (2023, June). Predicting Cryptocurrency Prices with Machine Learning Algorithms: A Comparative Analysis. *https://www.diva-portal.org/*. Retrieved September 17, 2023, from https://www.diva-portal.org/smash/get/diva2:1778251/FULLTEXT03
10. Kunpeng, Y., Mohammad , Z., Petr , H., & Ahmed , B. (2024, March). Cryptocurrency price forecasting – A comparative analysis of ensemble learning and deep learning methods. *ScienceDirect*. Retrieved September 11, 2024, from https://www.sciencedirect.com/science/article/pii/S1057521923005719#ab0005
11. Torgo, L., & Mariana, O. (2014, January). Ensembles for Time Series Forecasting. *Dalhousie University*. Retrieved September 18, 2024, from https://www.researchgate.net/publication/272833422_Ensembles_for_Time_Series_Forecasting
12. Chaudhary, D., & Sushil, K. (2023). Cryptocurrency Price Prediction Using Supervised Machine Learning Algorithms. *ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal*. doi:https://doi.org/10.14201/adcaij.31490

13. Derbentsev, V., Datsenko, N., Babenko, V., Pushko, O., & Pursky, O. (2020, October 06-09). Forecasting Cryptocurrency Prices Using Ensembles-Based Machine Learning Approach. *https://ieeexplore.ieee.org/*. doi:10.1109/PICST51311.2020.9468090
14. MyEducator. (2024). *Advantages and disadvantages of KNN*. Retrieved December 14, 2024, from https://app.myeducator.com: https://app.myeducator.com/reader/web/1421a/11/q07a0/
15. Omole, O., & David, E. (2024, August 5). Deep learning for Bitcoin price direction prediction: models and trading strategies empirically compared. *https://jfin-swufe.springeropen.com/*. Retrieved December 02, 2024, from https://jfin-swufe.springeropen.com/articles/10.1186/s40854-024-00643-1
16. Ammer, M. A., & Theyazn H., H. (2022, July 5). Deep Learning Algorithm to Predict Cryptocurrency Fluctuation Prices: Increasing Investment Awareness. *https://www.mdpi.com/*. doi:https://doi.org/10.3390/electronics11152349
17. Wu, J., Xinyi, Z., Fangyixuan, H., Haochen, Z., & Rohtiash, C. (2024, May 19). Review of deep learning models for crypto price prediction: implementation and evaluation. *https://arxiv.org/*. Retrieved September 20, 2024, from https://arxiv.org/html/2405.11431v1#:~:text=Deep%20Learning%20models%20are%20prominent,due%20to%20high%20market%20volatility.