International Journal of Innovative Research and Scientific Studies, 9(2) 2026, pages: 1-9

IJIRSS International Journal of Innovative Research & Scientific Studies

ISSN: 2617-6548

URL: www.ijirss.com

Comparative analysis of the forecasting ability of AI and Arima models for the Vn-index

Pham Dan Khanh[1], Dang Quang Huy[2*], Le Minh Anh[3]

[1]*School of Advanced Education Program National Economics University, Hanoi, Vietnam.*
[2,3]*National Economics University, Hanoi, Vietnam.*

Corresponding author: Dang Quang Huy (*Email: huyquangdang2803@gmail.com*)

Abstract

This study compares the forecasting performance of a traditional econometric model (ARIMA) and artificial intelligence (AI)-based models, namely Multilayer Perceptron (MLP) and Extreme Gradient Boosting (XGBoost), in predicting the VN-Index during the period from 2015 to June 2025, which was characterized by heightened volatility in Vietnam's stock market. Daily VN-Index closing prices were employed and divided into an 80% training set and a 20% testing set for out-of-sample evaluation. Forecast accuracy was assessed using Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). The Diebold–Mariano test was further applied to examine the statistical significance of differences in predictive performance among the models. The results indicate that ARIMA produced the highest forecasting errors, reflecting its limitations in capturing nonlinear dynamics and market volatility. The MLP model significantly improved forecasting accuracy, while XGBoost achieved the lowest error values across all evaluation metrics, demonstrating superior performance in handling noisy and volatile financial time series. AI-based models, particularly XGBoost, outperform the traditional ARIMA model in forecasting the VN-Index during volatile periods. The findings provide useful insights for investors and financial analysts by highlighting the effectiveness of advanced machine learning models in improving short-term market forecasting and investment decision-making in emerging markets.

Keywords: ARIMA, Artificial intelligence, Financial time series, Machine learning, VN-Index forecasting, XGBoost, MLP.

DOI: 10.53894/ijirss.v9i2.11220
Funding: This research was funded by the National Economics University, Hanoi, Vietnam. The funding body had no role in the study design, data collection, data analysis, interpretation of results, or the decision to publish the manuscript.
History: Received: 26 November 2025 / **Revised:** 5 January 2026 / **Accepted:** 9 January 2026 / **Published:** 3 February 2026

Competing Interests: The authors declare that they have no competing interests.
Authors' Contributions: All authors contributed equally to the conception and design of the study. All authors have read and agreed to the published version of the manuscript.
Transparency: The authors confirm that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.
Publisher: Innovative Research Publishing

1. Introduction

The accurate forecasting of stock market index movements is critical for facilitating informed investment decisions, optimizing portfolio allocation, and enhancing financial risk management strategies. Within Vietnam's financial landscape, the VN-Index serves as a principal barometer, comprehensively reflecting the overall health and developmental trajectory of the national economy. Consequently, precise predictions of this index carry substantial implications for both investment practitioners and policy formulators. In the context of Vietnam's emerging financial market-characterized by rapid expansion and pronounced volatility-the predictive efficacy of conventional econometric models like ARIMA is increasingly being evaluated against contemporary artificial intelligence (AI) methodologies.

For decades, the Autoregressive Integrated Moving Average (ARIMA) framework has constituted the methodological benchmark in time series forecasting, renowned for its rigorous statistical foundation and proficiency in modeling linear dependencies [1]. However, the intrinsically nonlinear characteristics of financial data, compounded by noise contamination and episodic economic shocks-particularly evident during from 2018 to June 2025 period marked by global pandemic disruptions and macroeconomic instability-have exposed fundamental limitations in traditional forecasting paradigms. In contrast, AI-based approaches, including Artificial Neural Networks (ANN), Long Short-Term Memory (LSTM) architectures, Random Forest, and XGBoost algorithms, have demonstrated remarkable capability in identifying complex nonlinear patterns and capturing long-term temporal dependencies within financial time series data [2, 3].

Although machine learning models have become increasingly prevalent in financial forecasting, empirical findings regarding the superiority of AI over traditional models in the Vietnamese market remain inconclusive. Some studies argue that neural networks and tree-based models significantly outperform ARIMA during periods of high market volatility, whereas others find that ARIMA remains effective during more stable phases or when the time series exhibits a relatively clear linear structure [4, 5]. This inconsistency highlights an important research gap, necessitating a systematic, updated, and comprehensive evaluation for the period from 2018 to June 2025 a time during which Vietnam’s stock market was heavily affected by multiple economic shocks and heightened volatility.

Against this backdrop, the present study focuses on comparing the forecasting performance of the VN-Index between the ARIMA model and two widely used machine learning models: MLP and XGBoost. The dataset comprises daily closing prices from 2018 to June 2025. Forecasting accuracy is assessed using standard error metrics, including Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). Additionally, the Diebold–Mariano statistical test is employed to determine whether differences in forecasting accuracy among the models are statistically significant. This study aims to provide robust empirical evidence on whether AI-based models offer meaningful predictive advantages over traditional approaches under real-world conditions in the Vietnamese market.

2. Theoretical Framework

2.1. Theoretical Background on Forecasting the VN-Index

Forecasting stock market indices, including the VN-Index, is grounded in the theoretical foundations of financial time series analysis and assumptions regarding market behavior. One of the most influential frameworks is the Efficient Market Hypothesis (EMH), proposed by Fama [6] which posits that stock prices fully incorporate all available information. Under this hypothesis, predicting future price movements based solely on historical data becomes inherently ineffective. However, empirical studies suggest that emerging markets such as Vietnam often exhibit only weak or semi-strong forms of efficiency, implying the presence of exploitable patterns, cycles, or market anomalies within price data [7, 8].

In the specific context of the VN-Index, prior research has documented several characteristic features typical of financial time series: non-stationarity, heteroskedasticity, autocorrelation, and nonlinear dependencies [9, 10]. These properties present considerable challenges for traditional linear models such as ARIMA, which assume relative stability and linear structure in the underlying data-generating process.

Conversely, machine learning models-including neural networks and boosting algorithms-have been recognized for their capacity to capture nonlinear relationships, complex variable interactions, and dynamic patterns inherent in financial markets [11, 12]. Nonetheless, their forecasting superiority is not universal. Their performance may degrade during periods marked by macroeconomic shocks, policy changes, or structural market disruptions-factors that are particularly prevalent in the Vietnamese stock market.

Given these considerations, comparing the forecasting performance of traditional models (e.g., ARIMA) with machine learning approaches (e.g., MLP, XGBoost) is essential for identifying the most suitable method for the period from 2018 to June 2025a timeframe characterized by heightened volatility and substantial changes in market structure.

2.2. The ARIMA (Autoregressive Integrated Moving Average) Model in Financial Forecasting

The ARIMA (Autoregressive Integrated Moving Average) model, introduced by Box [13] stands as one of the most prevalent traditional models for the analysis and forecasting of financial time series. ARIMA operates based on three core components: (i) the autoregressive (AR, order p) component, which describes the dependence of the current value on its own past values; (ii) the integration (I, order d) component, which involves differencing the data to remove trends and handle non-stationarity-a common characteristic of stock market data; and (iii) the moving average (MA, order q) component, which models the relationship between the current value and past random error terms.

The primary advantages of ARIMA lie in its simplicity, high interpretability, and solid theoretical foundation. As it only requires a univariate time series, the model is suitable for analysts seeking to generate rapid and reliable forecasts without the need for assembling numerous additional explanatory variables. Contemporary research, such as that by

Hyndman and Athanasopoulos [1] also affirms that ARIMA remains a crucial benchmark model in quantitative forecasting due to its stability and ease of tuning.

However, ARIMA possesses significant limitations when applied to financial data, particularly for stock indices like the VN-Index. Constructed upon linear assumptions, ARIMA struggles to capture nonlinear relationships, threshold effects, or complex market behaviors that are prevalent in emerging market contexts. Furthermore, the model is slow to react to unusual shocks, sharp volatility, or market phases influenced by macroeconomic shifts, investor sentiment, and geopolitical events. This can lead to a considerable decline in forecasting accuracy during periods of high market turbulence-a defining characteristic from 2018 to June 2025 period.

Consequently, ARIMA is often employed as a benchmark model when comparing against more advanced methodologies, particularly machine learning models capable of capturing nonlinearity and complex interactions within the data.

2.3. The Multilayer Perceptron (MLP) Model in Financial Forecasting

The Multilayer Perceptron (MLP) is a feedforward artificial neural network consisting of one or more hidden layers and is regarded as one of the foundational architectures of deep learning. MLPs learn to approximate the relationship between inputs and outputs through weight optimization using backpropagation algorithms. Due to their multilayer nonlinear structure, MLPs possess the ability to approximate any continuous nonlinear function, as established by the universal approximation theorem.

In financial time series forecasting, MLPs are frequently employed to identify nonlinear patterns and complex interactions within the data-features that linear models such as ARIMA are unable to capture [11]. When applied to forecasting the VN-Index, MLPs can exploit characteristics such as price momentum, multi-frequency market movements, and nonlinear fluctuations that are typical of emerging markets.

The key advantages of MLPs include:

- Strong capability in modeling nonlinear relationships
- Flexibility in adjusting the number of hidden layers and neurons
- Robustness in environments with noisy data

However, MLPs also exhibit notable limitations. The model is highly sensitive to hyperparameter selection and is prone to overfitting, especially when the training dataset is insufficient or improperly normalized [14, 15]. In addition, MLPs function as “black-box” models, providing limited interpretability-an important consideration in finance, where transparency and explainability are often required.

Despite these limitations, given the highly nonlinear dynamics of stock market time series such as the VN-Index, MLP remains a strong candidate for evaluating forecasting performance in comparison with traditional models like ARIMA.

2.4. The Xgboost (Extreme Gradient Boosting) Model in Financial Forecasting

XGBoost (Extreme Gradient Boosting), introduced by Chen and Guestrin [16] is a machine-learning algorithm belonging to the family of boosted decision tree models. Leveraging gradient boosting mechanisms, XGBoost constructs a sequence of decision trees in which each subsequent tree is optimized to correct the residual errors of the preceding one. The model is distinguished by its capacity to effectively handle nonlinear structures, high levels of noise, and complex interactions among variables.

In financial forecasting, XGBoost has gained prominence due to several advantages:

- Stronger nonlinear modeling capability compared with traditional models such as ARIMA
- Built-in L1–L2 regularization that mitigates overfitting
- High computational efficiency through system-level distributed optimization
- Excellent performance on structured (tabular) data, which typifies financial index forecasting tasks

Empirical studies have demonstrated that XGBoost often outperforms other methods in stock market prediction, particularly in emerging markets where nonlinear dynamics are more pronounced [12]. Nevertheless, XGBoost is not without limitations. The model is sensitive to hyperparameter configurations and can become overly complex if not properly tuned.

Given the characteristics of the VN-Index-marked by high volatility, nonlinear patterns, and sensitivity to macroeconomic conditions-XGBoost is expected to be a highly competitive model relative to both ARIMA and MLP in forecasting performance.

2.5. Overview of Empirical Research Findings

2.5.1. International Studies

Globally, the debate regarding the comparative efficacy of AI-based versus traditional models in financial forecasting has generated a substantial body of empirical research. The general trend over the past decade has leaned towards the superiority of machine learning and deep learning models, particularly in handling the inherent nonlinearity and noise present in market data.

A seminal study by Siami-Namini and Siami Namin [2] systematically compared the performance of LSTM and ARIMA across diverse financial time series. The results consistently demonstrated that LSTM reduced forecast errors by 85% to 90% compared to ARIMA, reinforcing the argument that recurrent network architectures with long-term memory are powerful tools for capturing complex data dependencies. Similarly, in the cryptocurrency domain, Kara, et al. [17] and

Latif, et al. [18] demonstrated that a hybrid deep learning model combining CNN and LSTM (CNN-LSTM) achieved superior accuracy over ARIMA in both short- and medium-term forecasting of Bitcoin prices, even during periods of extreme volatility.

However, the superiority of AI is not absolute. A study by Ahmed, et al. [19] which compared ARIMA and ANN for forecasting the S&P 500 index, found that while ANN generally performed better, ARIMA remained competitive or even superior during periods of stable market conditions with clear linear trends. This implies that the intrinsic characteristics of the data in specific periods are a critical determinant of model success. Furthermore, tree-based ensemble models like XGBoost have also demonstrated robust capabilities. In a large-scale comparative study, Patel, et al. [20] concluded that Random Forest and Support Vector Machines (SVM) yielded significantly better forecasting results than ARIMA across multiple global stock indices, attributable to their capacity to handle nonlinear interactions among numerous input variables.

2.5.2. Empirical Studies in the Vietnamese Market

The Vietnamese stock market, as an emerging market characterized by high volatility and strong susceptibility to speculative capital flows and macroeconomic events, provides an ideal testing ground for comparing forecasting models. Domestic studies conducted in recent years have reflected a diverse picture, further highlighting the context- and methodology-dependent nature of forecasting performance.

The study by Trung [5] is considered among the first systematic applications of Long Short-Term Memory (LSTM) networks for forecasting the VN-Index. Using daily data from 2010 to 2019, the authors concluded that LSTM produced significantly lower RMSE and MAPE values than the ARIMA model, particularly in predicting key reversal points. These findings support the view that AI models can offer clear advantages in the Vietnamese market context.

To overcome the limitations of individual models, some researchers have explored hybrid approaches. Quang [21] in his thesis, proposed an ARIMA-ANN hybrid model in which ARIMA handles the linear component while ANN models the nonlinear residuals. Empirical results showed that this hybrid approach outperformed both ARIMA and ANN individually, providing more accurate and stable forecasts of the VN-Index. This finding aligns with the research by Trung [5] who extended the hybrid approach by combining ARIMA with LSTM and observed a substantial improvement in MAPE values.

However, there remain perspectives emphasizing the continued relevance of traditional models. Several qualitative studies and brokerage reports indicate that during sideways or trending market phases, which are less affected by shock news, analysts can still rely on ARIMA to generate sufficiently reliable short-term forecasts with low computational costs. This inconsistency highlights a critical research gap: whether the advantages of complex AI models are sufficiently large to justify their implementation and operational costs under all market conditions in Vietnam, or if their superiority is only pronounced during periods of crisis and heightened volatility.

2.5.3. Research Gaps and Contributions of This Study

The literature review reveals several research gaps that need to be addressed:

Timeliness: Most domestic studies utilize data only up to 2019 or 2020, consequently omitting the critically important period from 2020 to 2024. This period was marked by unprecedented volatility stemming from the COVID-19 pandemic, global inflationary pressures, and significant monetary policy shifts.

Scope of Comparison: Existing research typically compares only one or two AI models against ARIMA. This study expands the comparative framework to include both deep learning models (LSTM) and tree-based ensemble models (Random Forest, XGBoost), thereby providing a more comprehensive perspective on the relative strengths of different AI paradigms.

Statistical Rigor: Many studies limit their evaluation to comparing error metrics without employing formal statistical tests to verify whether the observed performance differences are statistically significant. The application of the Diebold-Mariano test in this study is designed to address this specific methodological limitation.

By systematically targeting these identified gaps, the present research aims to provide a more updated, comprehensive, and statistically robust evaluation of forecasting models for the Vietnamese stock market, seeking to meaningfully contribute to the extant body of knowledge.

3. Data and Research Methodology

3.1. Research Data

The VN-Index data were collected based on the closing price of the last trading day of each month, covering the period from January 1, 2015, to June 30, 2025. The data source was obtained from the FiinPro database.

3.2. Data Preprocessing

After collection, the dataset was divided into two periods: (i) Period 1, from January 1, 2015, to April 30, 2023, and (ii) Period 2, from May 1, 2023, to June 30, 2025. Period 1 accounts for approximately 80% of the total study timeframe and was used as the training set for the AI models. Period 2, representing the remaining 20%, was used as the testing set to evaluate model performance through backtesting.

To ensure data consistency and comparability, the Min-Max normalization method was applied during preprocessing. This technique scales all variable values to a uniform range between 0 and 1, thereby enhancing convergence and improving the operational efficiency of machine learning models. All data processing, normalization, and model development procedures were conducted using the Python programming language.

3.3. Forecasting Models

This study employs three machine learning and artificial intelligence models to forecast the VN-Index, comprising:

(i) ARIMA (Autoregressive Integrated Moving Average): Introduced by Box [13]. ARIMA represents a classical tool in time series analysis. The model integrates three components: Autoregression (AR - p), Integration (I - d) to address non-stationarity, and Moving Average (MA - q). Its strengths lie in a well-defined theoretical framework, high interpretability, and relatively simple data requirements [1]. However, a fundamental limitation of ARIMA is its assumption of linear relationships within the data, rendering it less effective in capturing nonlinear structures and relatively inflexible in responding to unexpected shocks or sharp financial market fluctuations.

(ii) MLP (Multilayer Perceptron): The MLP model is a type of artificial neural network featuring multiple hidden layers, enabling it to learn complex features from data. MLP performs well with nonlinear data and possesses the capacity for accurate forecasting when properly trained.

(iii) XGBoost Regressor: XGBoost is a powerful boosting algorithm, optimized for efficient performance with large datasets and capable of minimizing generalization error. It utilizes decision trees to generate predictions and iteratively improves results through a weighted learning process.

(i) Root Mean Squared Error (RMSE)

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}$$

where:

- y_t represents the actual observed value at time t,
- $\hat{y}_t$ denotes the predicted value at time t, and
- n is the total number of observations in the evaluation dataset.

$$\text{MSE} = \frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2$$

where:

- y_t is the actual observed value at time t,
- $\hat{y}_t$ is the predicted value at time t, and
- n represents the total number of observations.

(ii) Mean Squared Error (MSE)

(iii) Mean Absolute Error (MAE)

$$\text{MAE} = \frac{1}{n}\sum_{t=1}^{n}|y_t - \hat{y}_t|$$

where:

- y_t represents the actual observed value at time t,
- $\hat{y}_t$ denotes the predicted value at time t, and
- n is the total number of observations in the evaluation dataset.

(iv) Mean Absolute Percentage Error (MAPE)

$$\mathrm{MAPE} = \frac{100\%}{n} \sum_{t=1}^{n} \left| \frac{y_t - \hat{y}_t}{y_t} \right|$$

where:

- y_t represents the actual observed value at time t,
- $\hat{y}_t$ denotes the predicted value at time t, and
- n is the total number of observations.

Lower values of MSE, RMSE, MAE, and MAPE indicate smaller forecast errors and correspondingly higher predictive accuracy of the model in capturing the underlying dynamics of the original time series.

3.4. Forecasting Results of the ARIMA Model

In this study, the ARIMA model was employed as a benchmark to assess potential improvements from applying machine learning models. After testing for stationarity and performing differencing when necessary, the model parameters (p, d, q) were identified based on the ACF/PACF plots and the AIC/BIC criteria, following the Box–Jenkins procedure. The optimal ARIMA model was then trained on historical data and used to forecast the VN-Index closing price at step t+1. According to Quang [21] the ARIMA(2,2,1) configuration was found to be the most suitable for the VN-Index dataset.

The empirical results indicate that the ARIMA model produced relatively high forecast errors: MSE = 16,928.43; RMSE = 130.10; MAE = 121.11; and MAPE = 9.64%. These substantial errors highlight the limitations of ARIMA in forecasting highly volatile and nonlinear financial time series such as the VN-Index. This observation is consistent with recent studies comparing ARIMA with machine learning models, which suggest that ARIMA often underperforms when data are noisy and exhibit strong nonlinear patterns.

3.5. Forecasting Results of the MLP Model

The MLP (Multilayer Perceptron) model-a fundamental feed-forward neural network architecture-has been extensively examined in financial research for benchmarking against ARIMA. In a study focused on stock price forecasting [22] demonstrated that MLP achieved superior predictive accuracy using only basic features derived from historical price series. In our experiment, the MLP model yielded results with MSE = 16,928.43, RMSE = 66.07, MAE = 50.81, and MAPE = 4.2%, which are substantially lower than those produced by the ARIMA model. This finding aligns with conclusions from prior literature: When a time series exhibits nonlinear characteristics or contains noise, neural networks such as MLP demonstrate an enhanced capacity to capture complex underlying relationships, consequently improving forecasting performance over ARIMA.

3.6. Forecasting Results of the XGBoost Model

The XGBoost model was constructed using lag features similar to the MLP model and subsequently trained with carefully tuned hyperparameters (e.g., learning_rate, max_depth, n_estimators). Leveraging its ability to capture nonlinear relationships, handle noisy data, and efficiently exploit temporal features, XGBoost has proven to be one of the most robust machine learning models for time series forecasting.

Recent studies have compared ARIMA and XGBoost to evaluate forecasting performance on highly volatile time series. In this study, the XGBoost model achieved MSE = 3,342.43, RMSE = 57.81, MAE = 43.12, and MAPE = 3.5%, representing the lowest errors among the three models. These results indicate that XGBoost is better suited for VN-Index data, which are characterized by volatility, noise, and nonlinearity. This finding aligns with the research of Kontopoulou, et al. [23] which highlighted that XGBoost is more “sensitive” to fluctuations and noise in the data compared to ARIMA, which tends to produce overly “smooth” or linear forecasts.

3.7. Comparison of Forecasting Models

As reported in Table 1 the XGBoost model achieves the lowest forecasting errors across all evaluation metrics, followed by the MLP model, while ARIMA records the highest error values

Table 1.
Forecasting performance comparison of ARIMA, MLP, and XGBoost Models.

Model	MSE	RMSE	MAE	MAPE
ARIMA	16,928.43	130.10	121.11	9.64%
MLP	4,365.28	66.07	50.81	4.22%
XGBoost	3,342.43	57.82	43.12	3.5%

Note: MSE = Mean Squared Error; RMSE = Root Mean Square Error; MAE = Mean Absolute Error; MAPE = Mean Absolute Percentage Error. Lower values indicate higher forecasting accuracy

Table 2.
Diebold–Mariano Test Results for Pairwise Model Comparisons.

Model Comparison	DM Statistic	p-value	Significance
XGBoost vs ARIMA	5.95	0	***
MLP vs ARIMA	4.89	0	***
XGBoost vs MLP	0.64	0.52	Not Significant

***Note:** ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

The dynamic comparison between actual and forecasted VN-Index values is illustrated in Figure 1, highlighting the superior tracking ability of the XGBoost model during volatile periods.

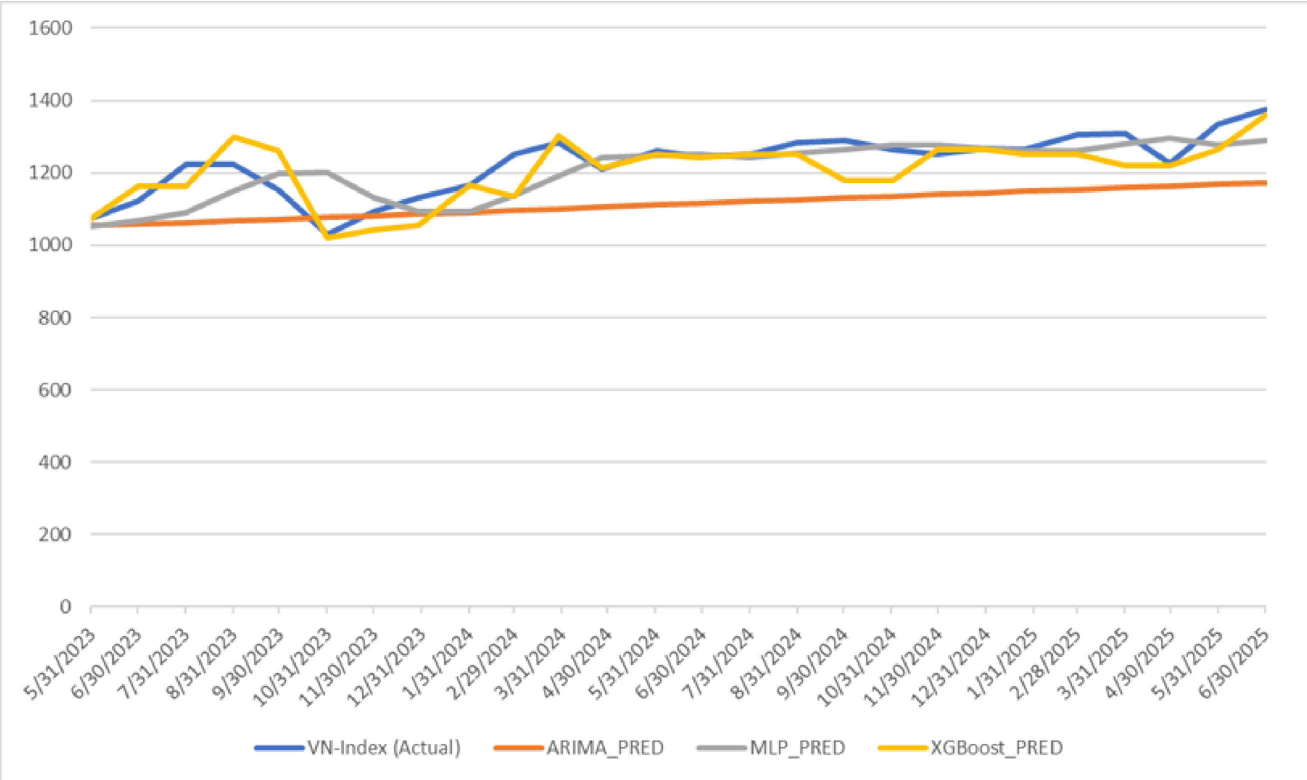

Figure 1.
Actual and Forecasted VN-Index Values Using ARIMA, MLP, and XGBoost Models.

In addition, Figure 2 provides a visual comparison of forecasting errors, clearly illustrating the superior performance of XGBoost across all evaluation metrics.

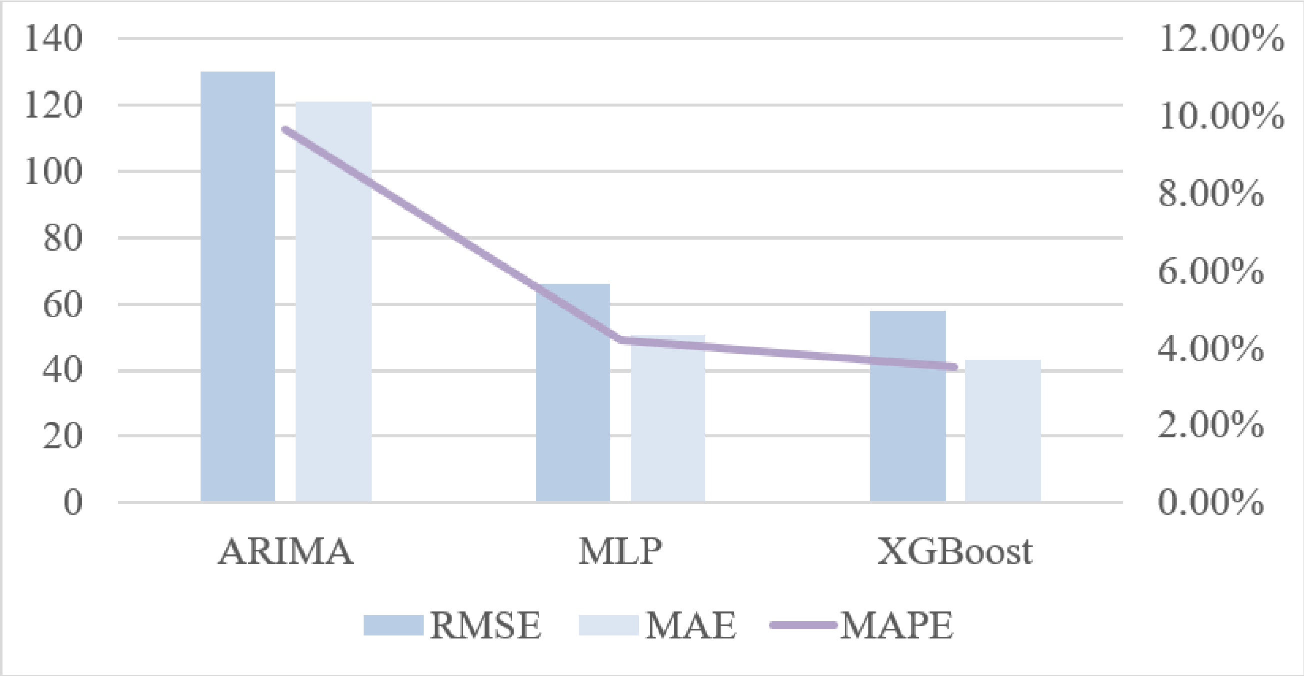

Figure 2.
Comparison of Forecasting Errors (RMSE, MAE, and MAPE) Across Models.

Empirical results and recent scholarly evidence demonstrate a clear distinction between traditional statistical models and machine learning/boosting approaches. The ARIMA model-despite its simplicity and interpretability-yielded the highest forecasting errors and proved inadequate for predicting highly volatile and nonlinear financial time series. The MLP model significantly reduced these errors, demonstrating the advantage of multi-layer neural networks in capturing nonlinear patterns. However, the most effective model in this study was XGBoost, achieving the lowest error rates and showcasing the power of gradient boosting algorithms for forecasting volatile time series.

This conclusion is supported not only by our experimental results but also by numerous recent international studies. These findings confirm that stock markets-particularly highly volatile indices-exhibit nonlinear structures with significant noise, making machine learning/boosting models and neural networks generally more suitable than traditional linear models.

Therefore, we conclude that for short-term forecasting of the VN-Index, XGBoost represents the superior approach; MLP serves as a viable alternative for simpler modeling requirements; while ARIMA should be primarily employed as a benchmark or in contexts with stable, linear data assumptions.

The VN-Index exhibits constant fluctuations influenced by multiple factors, making stock price forecasting a significant challenge for investors, financial analysts, and market professionals. In this study, we developed three distinct forecasting models using the Python programming language. Analytical results demonstrate that the XGBoost method achieved superior performance compared to the other two approaches in predicting VN-Index closing prices. Notably, despite the MLP model generating a “ConvergenceWarning”-indicating incomplete convergence after 2000 iterations-it still delivered stable and significantly more effective results than ARIMA. This research recommends that investors, portfolio managers, and market analysts should prioritize XGBoost methodologies for index price forecasting, as accurate prediction capabilities can substantially impact profitability and support informed investment portfolio decisions for all stakeholders.

References

[1] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and practice*. Melbourne, Australia: OTexts, 2018.

[2] S. Siami-Namini and A. Siami Namin, "Forecasting economics and financial time series: ARIMA vs," *LSTM (arXiv:1803.06386) [Preprint]. arXiv,* 2018. https://doi.org/10.48550/arXiv.1803.06386

[3] N. T. Co, H. H. Son, and N. T. Hoang, *Comparison between ARIMA and LSTM-RNN for stock index prediction," in Intelligent Human Systems Integration 2020*. Cham, Switzerland: Springer, 2020.

[4] T. Q. Bao and B. T. T. My, "Forecasting stock index based on hybrid artificial neural network models," *Journal of Economics, Law and Management,* vol. 3, no. 1, pp. 52-57, 2019. https://doi.org/10.32508/stdjelm.v3i1.540

[5] V. T. Trung, "ARIMA performance in stock index forecasting," 2023.

[6] E. F. Fama, "Efficient capital markets: A review of theory and empirical work," *The Journal of Finance,* vol. 25, no. 2, pp. 383-417, 1970. https://doi.org/10.2307/2325486

[7] K. P. Lim and R. Brooks, "The evolution of stock market efficiency over time: A survey of the empirical literature," *Journal of Economic Surveys,* vol. 25, no. 1, pp. 69-108, 2011.

[8] T. H. Hoang, "Market efficiency in emerging markets: Evidence from Vietnam," *Asian Economic and Financial Review,* vol. 10, no. 6, pp. 623–635, 2020.

[9] T. M. Nguyen and Q. P. Tran, "Nonlinear characteristics of stock index: An empirical analysis," *Journal of Asian Finance, Economics and Business*, vol. 8, no. 4, pp. 103–112, 2021.

[10] A. N. Q. Huynh, G. T. M. Vu, H. Q. Bui, and H. V. Hoang, "Corporate investment efficiency in response to national innovation: The moderating role of state ownership," *Journal of Economics and Development*, vol. 27, no. 3, pp. 194–209, 2025.

[11] G. P. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model," *Neurocomputing*, vol. 50, pp. 159-175, 2003. https://doi.org/10.1016/S0925-2312(01)00702-0

[12] T. Fischer and C. Krauss, "Deep learning with long short-term memory networks for financial market predictions," *European Journal of Operational Research*, vol. 270, no. 2, pp. 654-669, 2018. https://doi.org/10.1016/j.ejor.2017.11.054

[13] G. Box, *Box and Jenkins: Time series analysis, forecasting and control. In A Very British Affair: Six Britons and the Development of Time Series Analysis During the 20th Century (pp. 161-215)*. London: Palgrave Macmillan UK, 2013.

[14] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Networks*, vol. 2, no. 5, pp. 359-366, 1989.

[15] C. M. Bishop, *Neural networks for pattern recognition*. Oxford, U.K.: Oxford Univ. Press, 1995.

[16] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785–794). ACM*, 2016.

[17] Y. Kara, M. A. Boyacioglu, and Ö. K. Baykan, "Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange," *Expert Systems with Applications*, vol. 38, no. 5, pp. 5311–5319, 2011. https://doi.org/10.1016/j.eswa.2010.10.027

[18] N. Latif, J. D. Selvam, M. Kapse, V. Sharma, and V. Mahajan, "Comparative performance of LSTM and ARIMA for the short-term prediction of bitcoin prices," *Australasian Accounting, Business and Finance Journal*, vol. 17, no. 1, pp. 256-276, 2023. https://doi.org/10.14453/aabfj.v17i1.15

[19] I. Ahmed *et al.*, "Does service quality affect students' performance? Evidence from institutes of higher learning," *African Journal of Business Management*, vol. 4, no. 12, p. 2527, 2010.

[20] J. Patel, S. Shah, P. Thakkar, and K. Kotecha, "Predicting stock market index using fusion of machine learning techniques," *Expert Systems with Applications*, vol. 42, no. 4, pp. 2162-2172, 2015. https://doi.org/10.1016/j.eswa.2014.10.031

[21] L. T. Quang, "Comparing AI and ARIMA models for forecasting stock indices," 2024.

[22] Y. Sun, "A stock price prediction model based on MLP," *Advances in Economics, Management and Political Sciences*, vol. 165, pp. 208–214, 2025.

[23] V. I. Kontopoulou, A. D. Panagopoulos, I. Kakkos, and G. K. Matsopoulos, "A review of ARIMA vs. machine learning approaches for time series forecasting in data driven networks," *Future Internet*, vol. 15, no. 8, p. 255, 2023. https://doi.org/10.3390/fi15080255