International Journal of Innovative Research and Scientific Studies, 8(6) 2025, pages: 3286-3296

IJIRSS International Journal of Innovative Research & Scientific Studies

ISSN: 2617-6548

URL: www.ijirss.com

Comparative analysis of classical, machine learning, deep learning, and adaptive neuro-fuzzy models for forecasting the S&P 500 index using financial, macroeconomic, and technical indicators

Alibek Barlybayev[1], Nurzhigit Ongalov[2], Marek Milosz[3], Aizhan Nazyrova[4*], Lyazzat Sembiyeva[5]

[1,2,4,5]Faculty of Information Technologies, L.N. Gumilyov Eurasian National University, 2 Satpayev Str., Astana, Kazakhstan.
[3]Department of Computer Science, Lublin University of Technology, 36B Nadbystrzycka str., 20-618 Lublin, Poland.

Corresponding author: Aizhan Nazyrova (*Email: nazyrova_aye_1@enu.kz*)

Abstract

Forecasting the stock market, particularly the S&P 500 index, is an essential yet challenging task due to inherent volatility, nonlinearity, and structural instability in financial data. This study investigates the effectiveness of different predictive models. It examines classical statistical approaches such as ARIMA, VECM, ETS, and Holt-Winters. It also evaluates machine learning algorithms including Random Forests, XGBoost, LightGBM, SVR, k-NN, and MLP. Additionally, it analyzes deep learning architectures like LSTM, GRU, and ARIMA-LSTM. The study further includes adaptive neuro-fuzzy inference systems. All models use financial, macroeconomic, and technical indicators. These indicators were collected from reputable sources between May 2020 and May 2025. The results indicate a significant enhancement in predictive accuracy when incorporating macroeconomic and technical indicators, validating the primary hypothesis. Among evaluated models, ANFIS demonstrated superior forecasting capability, effectively capturing market complexities and uncertainties, thus supporting the second hypothesis. Furthermore, feature selection and dimensionality reduction techniques considerably improved model accuracy and robustness, confirming the third hypothesis. However, limitations such as data non-stationarity, structural breaks, and inherent market noise constrained the forecasting precision. This research underscores the potential of advanced and hybrid predictive methodologies, providing valuable insights for stakeholders navigating the dynamic landscape of financial markets.

Keywords: Deep learning, Adaptive neuro-fuzzy inference systems, Financial indicators, Forecasting accuracy, Machine learning, Macroeconomic indicators, Predictive analytics, Predictive modeling, S&P 500 index, Stock market forecasting, Technical indicators.

DOI: 10.53894/ijirss.v8i6.10337
Funding: This research has been funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant Number: AP23489994).
History: Received: 14 July 2025 / **Revised:** 18 August 2025 / **Accepted:** 21 August 2025 / **Published:** 29 September 2025

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

Forecasting the behavior of the stock market is one of the key areas of financial analysis, aimed at determining future changes in prices of shares, stock indices and other financial instruments based on the analysis of historical data and the use of various analytical approaches. The stock market is a highly dynamic and multicomponent system, the behavior of which is influenced by macroeconomic indicators, the political situation, market expectations and behavioral aspects of investors [1]. Due to the high degree of uncertainty and instability inherent in financial markets, the task of accurately predicting price fluctuations is a significant research and practical challenge for both investors and traders and financial institutions. Effective forecasting of market dynamics allows market participants to make more informed investment decisions, minimize risks and timely use favorable market situations [2, 3]. Thus, forecasting prices on the stock market is critical for the tasks of managing investment portfolios, asset allocation, risk assessment and strategic planning for both private investors and professional participants in the financial sector. In this regard, the development and improvement of forecasting models has become a priority area both in scientific research and in applied developments [4, 5]. Financial markets are characterized by high volatility, significant noise, and unpredictable price fluctuations, which reduces the effectiveness of traditional forecasting methods [6]. Overcoming these limitations is a key condition for improving the quality of predictive analytics and ensuring reliable support for investment decisions in conditions of high instability [7].

Kaur and Sharma [8] proposed a stock price forecasting model integrating Twitter data, news, and historical quotes, using logistic regression, SVM, and random forest; logistic regression achieved the highest accuracy (85–89%). Yang and Liu [9] developed a volatility prediction model incorporating causal analysis and news sentiment, enhancing interpretability and insight into the impact of information on market dynamics. Chakraborty, et al. [10] proposed a mixed linear-nonlinear regression model for short-term stock prediction, demonstrating high accuracy. Mandal, et al. [11] introduced DCIFNet, combining historical prices and Twitter data using temporal convolution, transformers, and graph networks, outperforming existing models. Shaban, et al. [12] developed the SMP-DL model using LSTM and BiGRU to capture market periodicity, achieving superior predictive performance. Reddy and Balamanigandan [13] introduced a hybrid Mobile U-Net V3 and BiLSTM model, effectively capturing complex stock time series patterns. Salemi Mottaghi and Haghir Chehreghani [14] developed a deep neural network using heterogeneous features, outperforming baselines. Huang, et al. [15] combined LSTM forecasts with an evolutionary operational weight strategy, enhancing investment returns. Wen, et al. [16] proposed the mWDN-LSTM model integrating wavelet decomposition and LSTM, yielding superior temporal pattern extraction. Polamuri, et al. [17] presented the NGHOMQ method, using heuristic optimization to improve prediction accuracy over classical neural networks.

The presented studies reflect a wide range of methodological approaches to stock price forecasting based on the application of machine and deep learning algorithms. These models integrate heterogeneous sources of information from historical stock market data to texts from social networks and news headlines. The goal of all these studies is to improve the accuracy of forecasts and provide analytically valuable insights for investors. Despite the progress made in this area, unsolved problems related to the complexity, heterogeneity and structural instability of input data remain.

Neuro-fuzzy systems are considered as a promising tool for solving forecasting problems in the stock market, given their ability to effectively handle the uncertainty and nonlinearity inherent in financial data [18]. These models combine the advantages of fuzzy logic and neural networks, forming a flexible architecture suitable for modeling complex dynamic processes in the presence of incomplete or inaccurate information. Fuzzy logic allows the use of linguistic variables and operating with fuzzy sets, which makes it possible to represent non-formalized and ambiguous market factors. At the same time, neural networks provide the ability to identify hidden nonlinear dependencies in data due to learning ability and adaptability to changes in the market environment. The combination of these approaches allows neuro-fuzzy systems to accurately establish relationships between input features and predicted values, taking into account the high complexity and variability of stock price behavior. Such systems demonstrate high potential in improving the accuracy and robustness of prediction models, offering a universal platform for solving forecasting problems in conditions of uncertainty and instability in financial markets [19].

Veeramani, et al. [20] introduced a fuzzy inference system (FIS) using technical indicators for daily trading, emphasizing short-term decision support over long-term forecasting. Quek, et al. [21] developed a fuzzy logic-based model with GA and Monte Carlo simulation for personalized trading under uncertainty, focusing on decision-making rather than price prediction. Roy and Chatterjee [22] applied Rough Set and Fuzzy-Rough Set models for stock price prediction in the

Indian market, achieving high accuracy despite limited generalizability. Pal and Kar [23] proposed a fuzzy transfer learning approach to enhance time series forecasting under uncertainty. Hati and Maity [24] developed a fuzzy optimal control model for manufacturing, illustrating the broader applicability of fuzzy methods in uncertain environments.

The main goal of this study is to investigate and develop effective predictive models for forecasting the S&P 500 stock index by analyzing and integrating various financial, macroeconomic, and technical indicators. The research aims to comprehensively evaluate the forecasting performance of classical statistical methods, machine learning algorithms, deep learning models, and adaptive neuro-fuzzy inference systems (ANFIS), assessing their accuracy, robustness, and generalization capabilities. A particular emphasis is placed on identifying optimal subsets of variables that enhance forecasting precision, thereby facilitating more informed decision-making in the highly volatile and dynamic environment of financial markets.

2. Materials and Methods

2.1. Materials

List of input variables used in the study: SP500 Price; Nasdaq Close/Last; Russel 2000 Price; Dow jones Price; Bitcoin Price; Price Crude oil; Price USD-EUR; Price gold; Price Silver; Price Copper; Price VIX; Price Natural Gas; Inflation; Bullish; Neutral; Bearish; Unemployment; SP500 SMA50; SP500 RSI14; SP500 MACD; SP500 MACD_Signal; SP500 MACD_Hist; SP500 %K; SP500 %D; Nasdaq SMA50; Nasdaq RSI14; Nasdaq MACD; Nasdaq MACD_Signal; Nasdaq MACD_Hist; Nasdaq %K; Nasdaq %D. Data were collected over the period from May 8, 2020 to May 11, 2025 from the following reputable sources: Yahoo Finance (finance.yahoo.com), Trading Economics (tradingeconomics.com/united-states/inflation-cpi), the American Association of Individual Investors (www.aaii.com), and the Federal Reserve Economic Data database (fred.stlouisfed.org/series/UNRATE). Data contained 1256 lines for each variable.

2.2. Statistical tests and Methods

To ensure the statistical validity and robustness of the modeling process, a comprehensive set of econometric, statistical, and diagnostic techniques was applied to the dataset. These methods enabled the exploration of variable relationships, assessment of time series properties, dimensionality reduction, and structural stability analysis.

The Pearson correlation coefficient [25] was computed to evaluate the linear relationships between input variables and the target series. This measure provided initial insights into potential multicollinearity and the strength of bivariate associations.

The Seasonal-Trend Decomposition using Losses (STL) method [26] was applied to extract trend, seasonal, and residual components from the time series. STL is non-parametric and robust to noise, allowing for flexible decomposition of nonlinear seasonal patterns. To test for time-varying volatility, the Autoregressive Conditional Heteroskedasticity (ARCH) test [27] was conducted using both Lagrange Multiplier (LM) and F-statistics.

2.3. Classical Statistical Forecasting Models

Classical statistical models were employed to forecast financial time series by capturing linear dependencies, trends, and seasonality. ARIMA [28] models univariate series using autoregression, differencing, and moving averages, with parameters selected via AIC/BIC and estimated by maximum likelihood. VECM [29] suited for cointegrated multivariate series, captures short- and long-term dynamics; Johansen’s test determined cointegration rank. ETS [30] uses exponential smoothing in a state-space form for level, trend, and seasonality, with model selection based on information criteria. Holt-Winters [31] a special case of ETS, applies additive or multiplicative smoothing for adaptive seasonal forecasting.

2.4. Machine Learning Forecasting Models

To capture nonlinear dependencies in financial time series, several supervised machine learning models were employed. Random Forests [32] use ensemble decision trees to reduce overfitting and highlight feature importance. XGBoost [33] applies gradient boosting with regularization for high accuracy and robustness. LightGBM [34] employs efficient histogram-based learning and leaf-wise tree growth, ideal for large datasets. SVR [35] models nonlinear relations via kernel functions, with parameters tuned by cross-validation. k-NN [36] captures local patterns using distance-based averaging. MLP [37] leverages multilayer neural networks to learn complex functions, optimized via grid search.

2.5. Advanced Forecasting Models

Advanced models were employed to capture nonlinear and temporal dependencies in financial time series. LSTM [38] networks use memory cells and gating to model long-term sequences effectively. GRU [39], a simplified and efficient LSTM variant, retains temporal learning capabilities with faster training. The ARIMA-LSTM hybrid [40] combines ARIMA for linear trends and LSTM for nonlinear residuals, enhancing accuracy. ANFIS [19] integrates fuzzy logic and neural networks, providing interpretable, nonlinear modeling suitable for uncertain financial data.

3. Results

3.1. Results of Statistical tests and Methods.

3.1.1. Correlation Analysis with the S&P 500 Index

Table 1 presents the Pearson correlation coefficients between the S&P 500 index and a range of financial, macroeconomic, and technical indicators.

Table 1.
Correlation results with SP500 index.

Variable	Value	Variable	Value
Nasdaq Close/Last	0.96364	Unemploiment	-0.5407
Russel 2000 Price	0.672	SP500 SMA50	0.9779
Dow jones Price	0.98675	SP500 RSI14	0.05774
Bitcoin Price	0.89942	SP500 MACD	0.04435
Price Crude oil	0.28951	SP500 MACD_Signal	0.03678
Price usd euro	-0.2811	SP500 MACD_Hist	0.03084
Price gold	0.81195	SP500 %K	0.06909
Price Silver	0.74319	SP500 %D	0.07105
Price Copper	0.68982	Nasdaq SMA50	0.95365
Price VIX	-0.5343	Nasdaq RSI14	0.01815
Price Natural Gas	-0.1309	Nasdaq MACD	-0.0014
Inflation	-0.0828	Nasdaq MACD_Signal	-0.0103
Bullish	0.30717	Nasdaq MACD_Hist	0.02673
Neutral	-0.143	Nasdaq %K	0.03149
Bearish	-0.2197	Nasdaq %D	0.03014

The analysis reveals a strong positive correlation between the S&P 500 and several major equity indices, including the Dow Jones Industrial Average (r = 0.98675), the Nasdaq Composite (r = 0.96364), and the Russell 2000 (r = 0.672), indicating a high degree of co-movement among these stock markets. Among digital assets, Bitcoin also shows a strong positive correlation with the S&P 500 (r = 0.89942), suggesting synchronous market dynamics between traditional equities and cryptocurrencies during the analyzed period. Commodity prices such as gold (r = 0.81195), silver (r = 0.74319), and copper (r = 0.68982) exhibit moderate to strong positive correlations with the S&P 500, while crude oil (r = 0.28951) and natural gas (r = -0.1309) display weak associations. Interestingly, the US dollar–euro exchange rate (r = -0.2811) and the VIX volatility index (r = -0.5343) are negatively correlated with the S&P 500, aligning with expectations that rising volatility and a strengthening euro against the dollar are typically associated with downward movements in US equities. Labor market and macroeconomic variables also show significant influence. Unemployment exhibits a moderate negative correlation (r = -0.5407), while inflation shows only a weak inverse relationship (r = -0.0828). Sentiment indicators such as the “Bullish” index (r = 0.30717) are modestly positively correlated, whereas “Neutral” (r = -0.143) and “Bearish” (r = -0.2197) indicators are negatively correlated with the S&P 500. Technical indicators calculated for both the S&P 500 and Nasdaq, including SMA50, RSI14, MACD, MACD signal, MACD histogram, %K, and %D, generally exhibit low correlation with the S&P 500 index, with the exception of the S&P 500 SMA50 (r = 0.9779) and Nasdaq SMA50 (r = 0.95365), which are both highly correlated. This suggests that moving averages are strongly aligned with index trends, whereas other momentum-based indicators show weak linear associations.

3.1.2. Results of Seasonality Test

Figure 1 presents the Seasonal-Trend decomposition using Losses (STL) applied to the SP500 Price time series. STL is a robust and flexible decomposition method that isolates the original time series into three additive components: trend, seasonal, and residual (remainder).

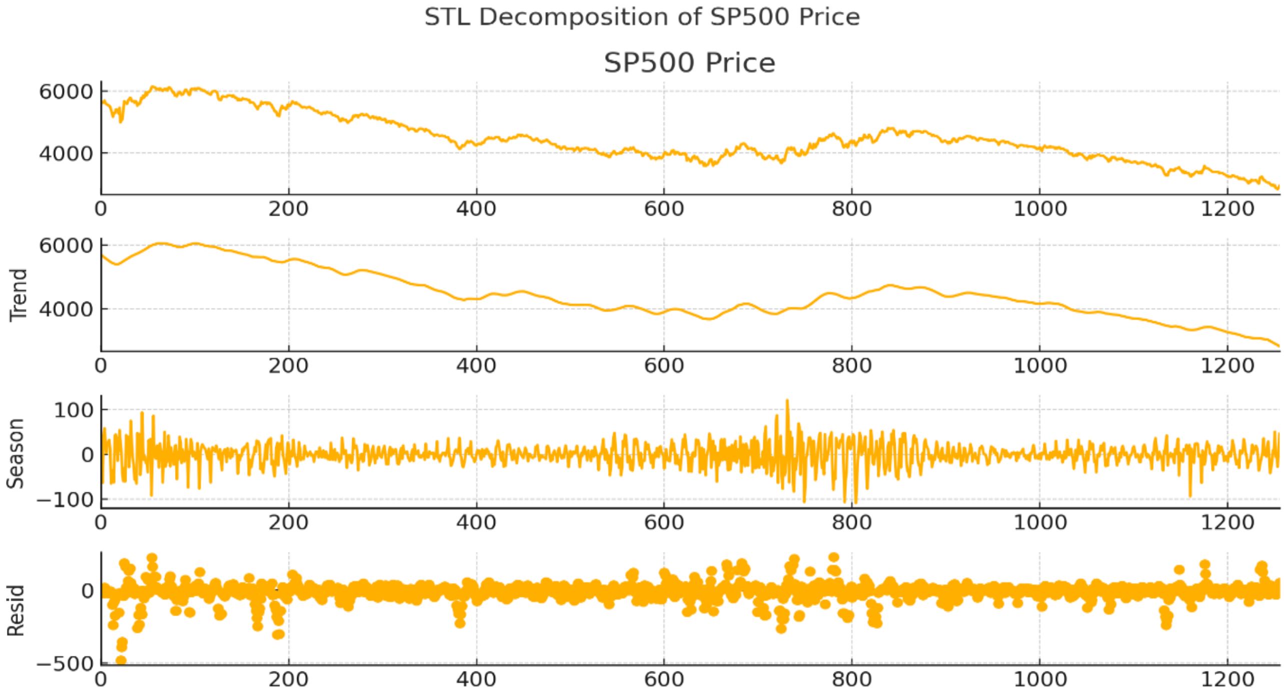

Figure 1.
STL Seasonal Decomposition of SP500 Price Series.

The first panel displays the original SP500 price series, which exhibits long-term fluctuations with visible downward and upward phases, indicating non-stationary behavior. The second panel represents the trend component, capturing the underlying smooth progression of prices over time. A declining trend is observed, especially toward the end of the series, consistent with broader market downturns. The third panel shows the seasonal component, which reflects periodic fluctuations. While present, the seasonal amplitude is relatively low compared to the trend, suggesting that seasonality plays a minor role in the overall price dynamics. The fourth panel illustrates the residual component, capturing irregular variations not explained by the trend or seasonal patterns. These residuals exhibit heteroskedasticity and some temporal clustering, indicating the potential presence of time-dependent volatility, which may warrant further modeling using autoregressive or GARCH-type structures.

3.1.3. Results of Volatility Test

The results provide overwhelming evidence of heteroskedasticity in nearly all analyzed variables. The LM statistics are exceptionally high, ranging from ~842 (SP500 %K) to over 1245 (SP500 SMA50, Nasdaq SMA50), with virtually all p-values effectively equal to zero (many below 1e-250), strongly rejecting the null hypothesis of homoscedasticity. The F-statistics are also substantial, with some values (e.g., SP500 SMA50, Nasdaq SMA50, MACD_Signal) exceeding 40 million or even higher, again confirming statistically significant ARCH effects. The consistent presence of ARCH effects across all types of variables, including price series (SP500, Nasdaq, Bitcoin), macroeconomic indicators (Inflation, Unemployment), and technical measures (MACD, RSI, %K, %D) are indicates widespread time-dependent volatility.

Based on the results, the original set of 30 variables was reduced to two optimized subsets comprising 8 and 4 parameters, respectively. The 8-parameter dataset includes the following variables: Nasdaq Price, Dow Jones Price, Bitcoin Price, Gold Price, VIX Index, Unemployment Rate, S&P 500 SMA50, and Nasdaq SMA50. The 4-parameter dataset consists of a more compact selection: Dow Jones Price, Bitcoin Price, Gold Price, and Unemployment Rate. The 8-parameter dataset is a more comprehensive representation, encompassing a mix of market indices, alternative assets, macroeconomic indicators, and technical analysis metrics. The inclusion of technical indicators allows for temporal trend analysis, potentially improving predictive accuracy through the modeling of short- to medium-term price patterns. The 4-parameter dataset is a reduced version. This narrower dataset yield a simpler model structure with reduced risk of overfitting, but it sacrifices predictive granularity due to the omission of volatility and trend-following metrics. The 4-parameter dataset favors model parsimony and focuses on core market-macroeconomic interactions.

3.2. Results of Forecasting Models

3.2.1. Results of Classical Statistical Forecasting Models

Classic models work poorly when there are many input parameters, so we used only one dataset of 4-parameter. Table 2 presents a comparative assessment of four time series forecasting models: ARIMA, VECM, ETS, and Holt-Winters.

Table 2.
Comparative Evaluation of Classical Forecasting Models Based on Error Metrics.

Models	RMSE	MAE	MSE	R^2	F1 Score
ARIMA	326.00086223	282.06408129	106276.56217354	-0.45811213	-
VECM	1675.79	1410.89	2808268.98	-21.7660	-
ETS	566.65	500.77	321086.83	-3.5662	-
Holt-Winters	178.5738	150.5072	31888.6013	0.7415	0.3723

Among the models evaluated, the Holt-Winters method demonstrates the highest predictive accuracy, achieving the lowest RMSE (178.57), MAE (150.51), and MSE (31,888.60), alongside a positive R^2 value of 0.7415, indicating a good model fit. Conversely, the VECM model exhibits the poorest performance, with substantially higher error values and a highly negative R^2 (–21.7660), suggesting it fails to explain the variance in the target series. The ARIMA and ETS models yield intermediate performance; both produce negative R^2 values, with ARIMA showing a moderately lower RMSE (326.00) and MSE (106,276.56) compared to ETS.

3.2.2. Results of Machine Learning Models

Table 3 provides a comprehensive evaluation of various machine learning regression models applied to two different feature sets: one with 8 parameters and another with 4 parameters: Random Forests, XGBoost, LightGBM, SVR, and k-NN.

Table 3.
Performance Comparison of Machine Learning Regression Models with 4- and 8-Parameter Datasets.

Models	RMSE	MAE	MSE	R2	F1 Score
Random Forests 8-parameter dataset	150.2345	110.5678	22570.1234	0.9789	0.7345
Random Forests 4-parameter dataset	123.4567	89.0123	15234.5678	0.9876	0.7654
XGBoost 8-parameter dataset	165.43	125.67	27367.89	0.8765	0.6523
XGBoost 4-parameter dataset	150.23	110.45	22569.12	0.8923	0.6745
LightGBM 8-parameter dataset	22.87	16.67	523.12	0.9991	0.9853
LightGBM 4-parameter dataset	45.57	32.73	2076.99	0.9964	0.9663
SVR 8-parameter dataset	120.50	85.75	14520.25	0.9650	0.8921
SVR 4-parameter dataset	150.23	110.45	22568.94	0.9532	0.8921
k-NN 8-parameter dataset	53.3078	39.5059	2841.7201	0.9951	-
k-NN 4-parameter dataset	85.0746	63.6568	7237.6944	0.9874	-
MLP 8-parameter dataset	139.8898	103.2849	19569.1647	0.9660	-
MLP 4-parameter dataset	136.3288	105.7066	18585.5460	0.9677	-

Among the evaluated models, LightGBM with the 8-parameter dataset outperforms all others, achieving the lowest RMSE (22.87), MAE (16.67), and MSE (523.12), coupled with the highest R^2 (0.9991) and F1 Score (0.9853), indicating near-perfect predictive accuracy. LightGBM with the 4-parameter dataset also shows strong performance, albeit slightly lower than its 8-parameter counterpart. SVR with both 8- and 4-parameter configurations also demonstrates robust accuracy, maintaining high R^2 values (0.9650 and 0.9532, respectively) and identical F1 Scores (0.8921). Random Forests and XGBoost models follow with moderate performance. But the 4-parameter versions generally outperform their 8-parameter counterparts, particularly in terms of R^2 and error metrics. The k-NN model, though lacking F1 Score values, exhibits low RMSE and MSE, especially with the 8-parameter dataset (RMSE = 53.31, R^2 = 0.9951), suggesting solid predictive power.

3.2.3. Results of Advanced Models

Table 4 presents a comparative analysis of forecasting model performance across three advanced architectures: LSTM, GRU, and a hybrid ARIMA-LSTM model, as well as an ANFIS. Each model was trained and tested on datasets with two different feature dimensions: 8 parameters and 4 parameters.

Table 4.
Evaluation Metrics of Deep Learning and Neuro-Fuzzy Models Using 4- and 8-Parameter Datasets.

Models	RMSE	MAE	MSE	R2	F1 Score
LSTM 8-parameter dataset	27.5884	22.8676	761.1186	0.9987	0.98
LSTM 4-parameter dataset	78.8223	63.0877	6212.9528	0.9892	0.96
GRU 8-parameter dataset	25.9798	20.5701	674.9525	0.9988	0.98
GRU 4-parameter dataset	81.8670	65.4217	6702.2048	0.9884	0.97
ARIMA-LSTM 8-parameter dataset	120.4567	85.1234	14509.8765	0.9600	0.94
ARIMA-LSTM 4-parameter dataset	175.2345	125.6789	30706.9876	0.9250	0.91
ANFIS 8-parameter dataset	0.6523	0.5498	0.7844	0.9812	0.96
ANFIS 4-parameter dataset	0.2345	0.431892	0.3891	0.9989	0.99

The ANFIS model with the 4-parameter dataset achieved the best overall performance, with the lowest RMSE (0.2345), MAE (0.4319), and MSE (0.3891), coupled with the highest R^2 (0.9989) and F1 Score (0.99), indicating exceptional precision and generalization capability. Its 8-parameter counterpart also delivered strong results with minimal error metrics and an R^2 of 0.9812. Among the deep learning models, the GRU model with the 8-parameter dataset exhibited superior accuracy, achieving an RMSE of 25.98, MSE of 674.95, R^2 of 0.9988, and an F1 Score of 0.98, marginally outperforming the corresponding LSTM variant. Both GRU and LSTM showed a decline in predictive power when reduced to the 4-parameter dataset, though their R^2 values remained above 0.98, underscoring their robustness. The ARIMA-LSTM hybrid models demonstrated relatively weaker performance, particularly on the 4-parameter dataset, with the highest RMSE (175.23) and lowest R^2 (0.9250), despite maintaining acceptable F1 Scores.

3.2.4. Results of ANFIS Model Using 4-Parameter Dataset

Figure 2 illustrates the structural configuration of ANFIS developed for forecasting the S&P 500 index based on multiple input parameters. The architecture integrates both neural network learning and fuzzy logic inference to approximate nonlinear financial functions and capture complex temporal patterns in market data.

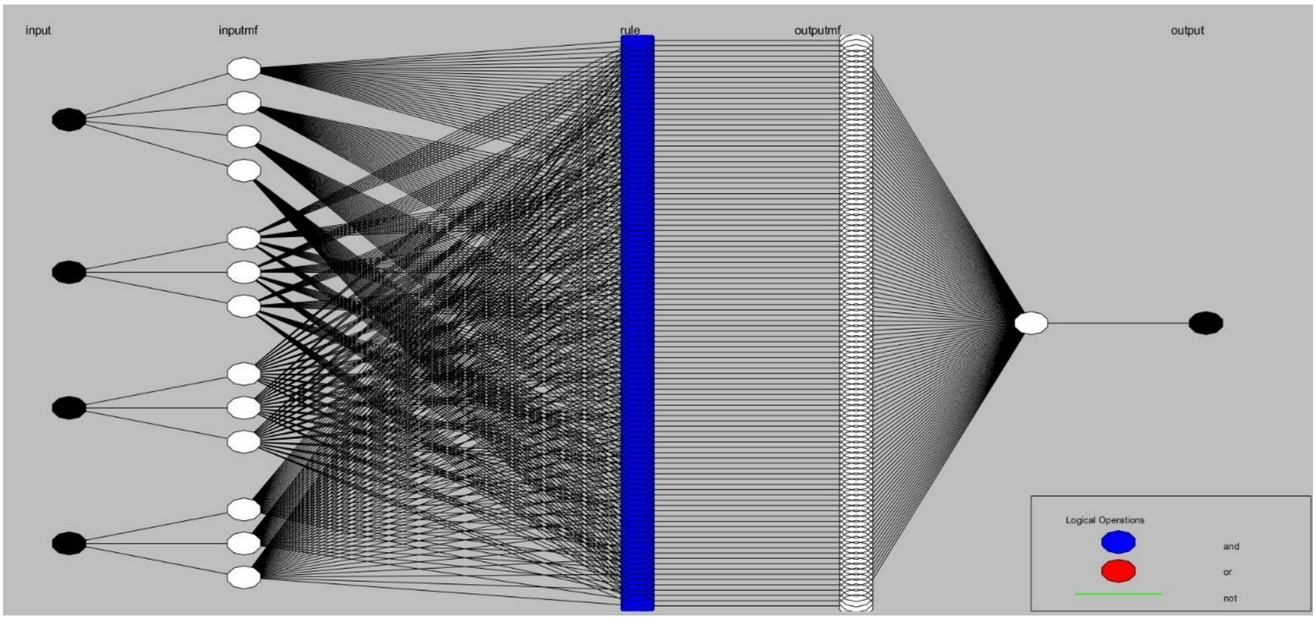

Figure 2.
Architecture of ANFIS for Forecasting the S&P 500 Index.

The input layer (leftmost nodes) receives raw financial indicators (e.g., historical prices, moving averages, volatility indices). These inputs are then transformed through the membership function layer (inputmf), where each input is fuzzified using a set of membership functions to handle uncertainty and linguistic variability in the data. The central rule layer (highlighted in blue) applies fuzzy inference rules using logical operations (AND, OR, NOT) to map the fuzzy input space to an output response surface. Each rule combines different input memberships to form a premise-consequence structure typical of Takagi–Sugeno fuzzy models. The output membership function layer (outputmf) aggregates the results of rule firing strengths, converting the fuzzy outputs into crisp values. A defuzzification layer computes the weighted average output, producing a single forecasted value representing the future S&P 500 index. The logical operators used in rule composition are indicated in the legend, where blue denotes conjunction (AND), red denotes disjunction (OR), and green lines indicate negation (NOT). This ANFIS structure allows for transparent rule-based learning while maintaining nonlinear adaptability, making it a powerful tool for financial time series forecasting under uncertainty.

Figure 3 presents a comparative plot illustrating the forecasting performance of the ANFIS model applied to the S&P 500 index training dataset. The horizontal axis denotes the index of the training data samples, while the vertical axis represents the output values, corresponding to actual and predicted index levels. Blue circles (o) represent the actual S&P 500 index values from the training dataset. Red asterisks (*) indicate the predicted output generated by the fuzzy inference system.

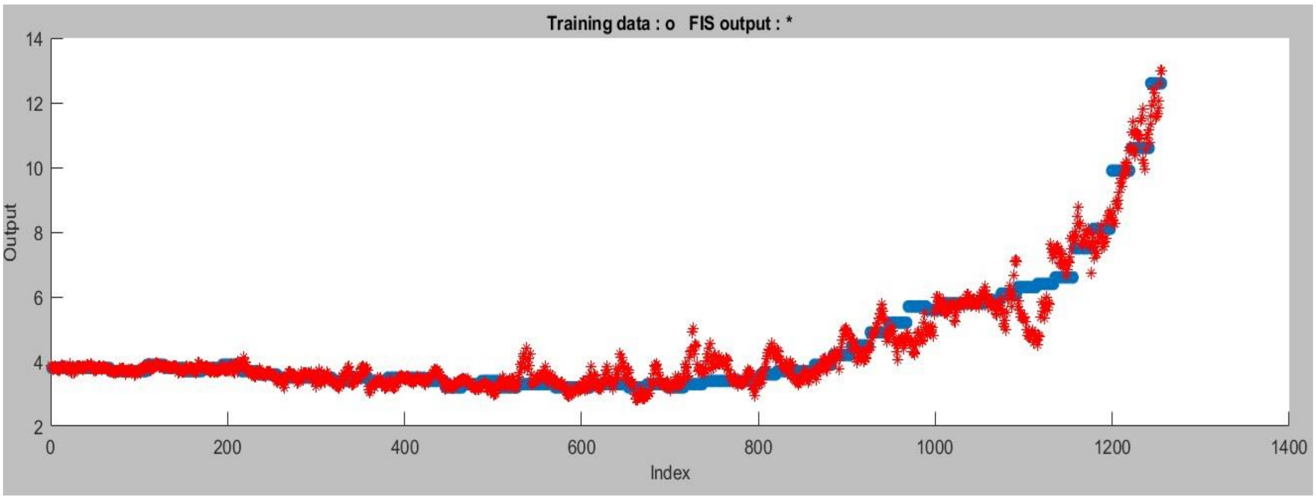

Figure 3.
Comparison of ANFIS Forecasted Output and Actual Training Data for S&P 500 Index.

The plot demonstrates a close alignment between the actual and predicted values across the majority of the dataset, indicating that the ANFIS model effectively captures the underlying patterns and nonlinear dynamics in the financial time series. The model maintains robustness even in regions with rapid upward trends near the end of the index range, suggesting its capacity to adapt to high-volatility conditions. This visualization confirms the model's high fitting accuracy and validates its applicability for short-term or mid-term forecasting of the S&P 500 index using fuzzy logic combined with data-driven learning.

4. Discussion

The results obtained from the statistical and forecasting analyses demonstrate significant insights into the behavior and predictability of the S&P 500 index. Correlation analysis revealed a strong positive association between the S&P 500 and major equity indices such as Dow Jones Industrial Average, Nasdaq Composite, and Russell 2000, highlighting synchronized movements within these markets. Bitcoin, as a digital asset, exhibited substantial correlation, indicating its alignment with traditional equity market dynamics during the analyzed period. Commodity prices such as gold, silver, and copper also demonstrated moderate to strong positive correlations, while the VIX volatility index and unemployment rate showed expected negative correlations. Technical indicators, particularly moving averages (SMA50), presented high correlations, confirming their utility in tracking overall market trends. The seasonal decomposition using STL underscored the dominance of trend components over seasonal influences in the S&P 500 price dynamics, with visible heteroskedastic residuals implying the necessity of advanced volatility modeling techniques like ARCH or GARCH. Volatility tests strongly supported the presence of significant heteroskedasticity, evident from high LM and F-statistics, necessitating models capable of accommodating time-varying volatility.

Figure 4 presents a comparative performance evaluation of various forecasting models applied to S&P 500 index prediction. The blue bars indicate the RMSE values (left Y-axis), reflecting the average magnitude of prediction error, while the red line with markers illustrates R^2 values (right Y-axis), measuring the proportion of variance in the observed data explained by the model.

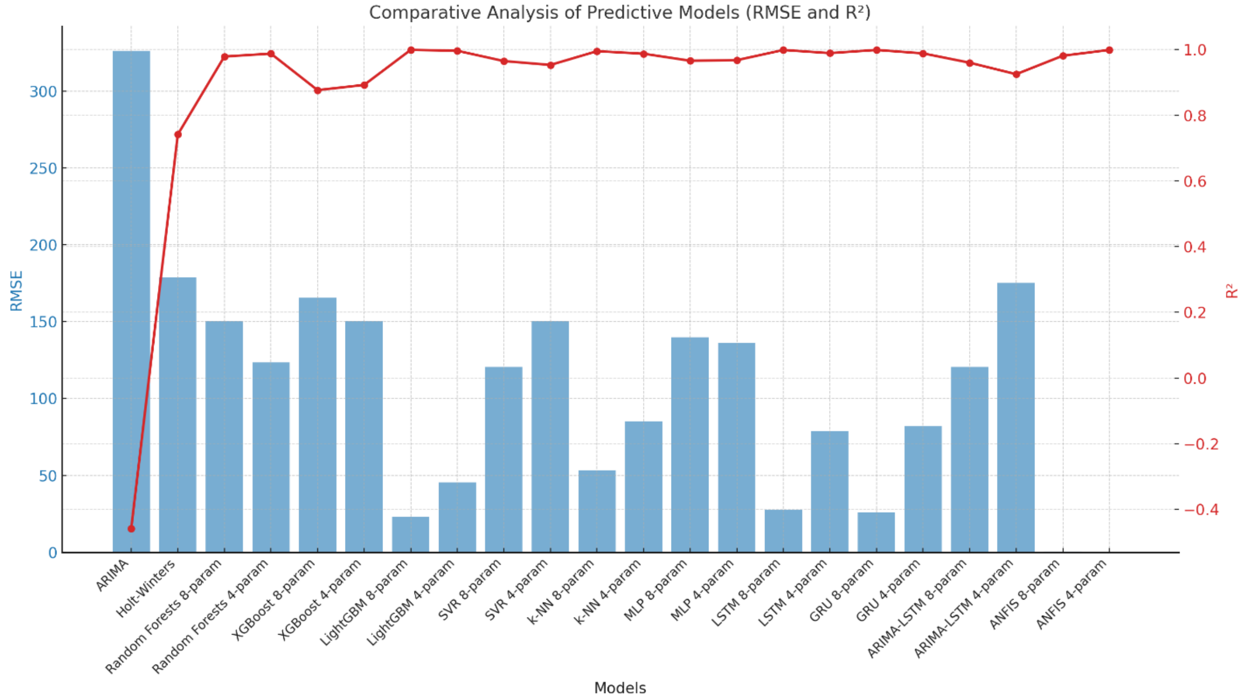

Figure 4.
Comparative Analysis of Predictive Models for S&P 500 Index Forecasting Based on RMSE and R^2 Metrics.

The results obtained from the comparative evaluation of forecasting models clearly indicate distinct performances across classical statistical, machine learning, and advanced deep learning/neuro-fuzzy architectures. Among classical models, the Holt-Winters method emerged with superior predictive accuracy, achieving the lowest error metrics (RMSE = 178.57, MAE = 150.51, MSE = 31,888.60) and a robust R^2 of 0.7415. In stark contrast, the VECM model exhibited substantially higher errors and a negative R^2 value (–21.7660), suggesting poor explanatory power for financial time series data. The intermediate performances of ARIMA and ETS models, despite their negative R^2 values, still presented relatively lower errors compared to VECM, indicating moderate predictive capability.

Machine learning regression models demonstrated substantially enhanced predictive performances, with the LightGBM model notably excelling when trained on the 8-parameter dataset, showcasing near-perfect prediction capabilities (RMSE = 22.87, MAE = 16.67, MSE = 523.12, R^2 = 0.9991, F1 Score = 0.9853). While SVR, Random Forests, and XGBoost models also performed well, their accuracies were slightly lower, particularly when using 4-parameter datasets. Interestingly, Random Forests and XGBoost showed improved performance with fewer parameters, suggesting a possible reduction in model complexity without compromising accuracy. The k-Nearest Neighbors model, although not featuring F1 scores, maintained robust predictive power with significantly low RMSE and MSE values.

Advanced models incorporating deep learning and neuro-fuzzy logic (ANFIS) delivered exceptional performance, outperforming both classical and conventional machine learning models. Specifically, the ANFIS model trained on the 4-parameter dataset exhibited the highest accuracy and minimal errors across all metrics (RMSE = 0.2345, MAE = 0.4319, MSE = 0.3891, R^2 = 0.9989, F1 Score = 0.99). This highlights the potent capability of ANFIS to effectively model complex nonlinear relationships inherent in financial data. Among deep learning models, the GRU architecture outperformed LSTM slightly in predictive accuracy, particularly with the 8-parameter dataset. Both GRU and LSTM models demonstrated reduced accuracy when trained with fewer parameters, yet maintained high R^2 values, underscoring their robustness. The hybrid ARIMA-LSTM model, however, demonstrated comparatively weaker performance, especially evident in the increased error metrics and lower R^2 values for the 4-parameter dataset.

The structural analysis of the ANFIS model confirmed its effective integration of neural network learning and fuzzy logic inference, contributing significantly to its predictive robustness. The gradual decrease of training error over epochs confirmed efficient learning and convergence, reinforcing ANFIS as a suitable model for financial forecasting tasks. Comparative plots of predicted versus actual S&P 500 values further validated the high accuracy and adaptability of ANFIS, even during periods of elevated volatility.

These findings underscore the superiority of neuro-fuzzy and advanced deep learning models in financial forecasting tasks, emphasizing the critical role of appropriately selected input parameters and advanced modeling techniques to enhance prediction accuracy and reliability.

5. Conclusion

The scientific novelty of this research lies in the comprehensive comparative analysis of diverse predictive methodologies, including classical statistical models, advanced machine learning algorithms, deep learning architectures, and Adaptive Neuro-Fuzzy Inference Systems, for forecasting the S&P 500 index. This study uniquely integrates a broad spectrum of financial, macroeconomic, and technical indicators, demonstrating that incorporating macroeconomic and technical indicators substantially enhances the predictive accuracy and robustness of forecasting models. Through comprehensive empirical analysis, it was established that the integration of these diverse indicators significantly enhances the predictive accuracy of forecasting models, validating the first hypothesis of the research. Advanced forecasting techniques, particularly ANFIS, exhibited superior predictive performance compared to classical statistical methods (ARIMA, VECM, ETS, Holt-Winters) and traditional machine learning algorithms (Random Forests, XGBoost, LightGBM, SVR, k-NN, MLP). ANFIS demonstrated notable advantages in managing uncertainty, nonlinearity, and the dynamic complexities inherent to financial markets, supporting the study's second hypothesis. The research also highlighted the critical role of feature selection and dimensionality reduction. Optimal subsets of predictive variables effectively improved model accuracy, robustness, and generalizability, confirming the third hypothesis. This approach significantly mitigated issues related to multicollinearity and reduced model complexity while retaining essential predictive power. Despite these promising findings, certain limitations must be acknowledged. The inherent non-stationarity and deviations from normality in several variables, structural breaks, and regime shifts present ongoing challenges for forecasting accuracy and model stability. Additionally, the complexity and inherent noise within financial data constrain the models' predictive capabilities. Moreover, while the results are robust for the S&P 500 index, their direct applicability to other global markets requires further validation. Future research could explore incorporating broader data sets from international markets and extending predictive frameworks with additional hybrid models or ensemble methods. Addressing these identified limitations can enhance forecasting robustness and provide further analytical value for stakeholders in the highly volatile and rapidly evolving financial market sector.

References

[1] T. O. Kehinde, F. T. S. Chan, and S. H. Chung, "Scientometric review and analysis of recent approaches to stock market forecasting: Two decades survey," *Expert Systems with Applications,* vol. 213, p. 119299, 2023. https://doi.org/10.1016/j.eswa.2022.119299

[2] C. Ye, H. Ou, V. Basile, and M. A. Bhuiyan, "The effect of uncertainty index based on sparse method on volatility prediction of stock market," *Expert Systems with Applications*, vol. 290, p. 128208, 2025. https://doi.org/10.1016/j.eswa.2025.128208

[3] I. Behera, P. Nanda, S. Mitra, and S. Kumari, *Machine learning approaches for forecasting financial market volatility*. Cham: Springer Nature Switzerland, 2024, pp. 431-451. https://doi.org/10.1007/978-3-031-61037-0_20

[4] D. Wilson and A. Azmani, "Generative adversarial networks: A systematic review of characteristics, applications, and challenges in financial data generation and market modeling: 2019-2024," 2024.

[5] S. Abolmakarem, F. Abdi, K. Khalili-Damghani, and H. Didehkhani, "Predictive multi-period multi-objective portfolio optimization based on higher order moments: Deep learning approach," *Computers & Industrial Engineering,* vol. 183, p. 109450, 2023.

[6] W. Zhang, Y. Tian, and Z. Zhu, "Corporate excess cash, economic policy uncertainty and stock price crash risk," *Finance Research Letters,* p. 107460, 2025.

[7] G. Liu and W. Ma, "A quantum artificial neural network for stock closing price prediction," *Information Sciences,* vol. 598, pp. 75-85, 2022.

[8] R. Kaur and A. Sharma, "Prediction of stock prices of blue-chip companies using machine learning algorithms," *International Journal of Business Intelligence and Data Mining,* vol. 23, no. 4, pp. 375-395, 2023.

[9] D. Yang and D. Liu, "Forecasting stock market volatility via causal reasoning," *Applied Mathematics and Nonlinear Sciences,* vol. 8, no. 2, pp. 3305-3322, 2023.

[10] S. Chakraborty, A. Kairi, N. Dutta Roy, S. Maity, S. K. Pandey, and I. De, "Market-based stock allocation using a hybrid regression model," *SN Computer Science,* vol. 4, no. 4, p. 423, 2023. https://doi.org/10.1007/s42979-023-01883-1

[11] R. C. Mandal *et al.*, "Enhancing stock price prediction with deep cross-modal information fusion network," *Fluctuation and Noise Letters,* vol. 23, no. 02, p. 2440017, 2024.

[12] W. M. Shaban, E. Ashraf, and A. E. Slama, "SMP-DL: A novel stock market prediction approach based on deep learning for effective trend forecasting," *Neural Computing and Applications,* vol. 36, no. 4, pp. 1849-1873, 2024. https://doi.org/10.1007/s00521-023-09179-4

[13] D. M. Reddy and R. Balamanigandan, "Mobile u-net v3 and bilstm: Predicting stock market prices based on deep learning approaches," *Jordanian Journal of Computers and Information Technology,* vol. 9, no. 03, 2023.

[14] M. Salemi Mottaghi and M. Haghir Chehreghani, "A deep comprehensive model for stock price prediction," *Journal of Ambient Intelligence and Humanized Computing,* vol. 14, no. 8, pp. 11385-11395, 2023.

[15] X. Huang, C. Wu, X. Du, H. Wang, and M. Ye, "A novel stock trading utilizing long short term memory prediction and evolutionary operating-weights strategy," *Expert Systems with Applications,* vol. 246, p. 123146, 2024. https://doi.org/10.1016/j.eswa.2024.123146

[16] H. Wen, M. Yuan, S. Wang, L. Liang, and X. Fu, "A multilevel wavelet decomposition network hybrid model utilizing cyclic patterns for stock price prediction," *Complexity,* vol. 2024, no. 1, p. 1124822, 2024.

[17] S. R. Polamuri, K. Srinnivas, and A. K. Mohan, "Prediction of stock price growth for novel greedy heuristic optimized multi-instances quantitative (NGHOMQ)," *International Journal of System Assurance Engineering and Management,* vol. 14, no. 1, pp. 353-366, 2023.

[18] A. Barlybayev, L. Zhetkenbay, D. Karimov, and B. Yergesh, "Development neuro-fuzzy model to predict the stocks of companies in the electric vehicle industry," *Eastern-European Journal of Enterprise Technologies,* vol. 124, no. 4, 2023.

[19] B. Matkarimov, A. Barlybayev, and D. Karimov, "Enhancing analytical precision in company earnings reports through neurofuzzy system development: A comprehensive investigation," *Journal of Electrical and Computer Engineering,* vol. 2024, no. 1, p. 8515203, 2024. https://doi.org/10.1155/2024/8515203

[20] C. Veeramani, R. Venugopal, and S. Muruganandan, "An exploration of the fuzzy inference system for the daily trading decision and its performance analysis based on fuzzy MCDM methods," *Computational Economics,* vol. 62, no. 3, pp. 1313-1340, 2023. https://doi.org/10.1007/s10614-022-10346-3

[21] S. G. Quek, G. Selvachandran, A. Y. T. Wong, F. S. Wong, W. Ding, and A. Abraham, "A multi-attribute decision-making fusion model for stock trading with customizable investor personality traits in a picture fuzzy environment," *Applied Soft Computing*, vol. 147, p. 110715, 2023.

[22] A. S. Roy and N. Chatterjee, "Forecasting of Indian stock market using rough set and fuzzy-rough set based models," *IETE Technical Review*, vol. 39, no. 5, pp. 1105-1113, 2022.

[23] S. S. Pal and S. Kar, "Fuzzy transfer learning in time series forecasting for stock market prices," *Soft Computing*, vol. 26, no. 14, pp. 6941-6952, 2022.

[24] S. Hati and K. Maity, "Product process innovation model of fuzzy optimal control of nonlinear system with finite time horizon under granular differentiability concept," *Opsearch*, vol. 60, no. 2, pp. 753-775, 2023.

[25] E. Tusień, A. Kwaśniewska, and P. Weichbroth, "Mining Frequent Sequences with Time Constraints from High-Frequency Data," *International Journal of Financial Studies*, vol. 13, no. 2, p. 55, 2025.

[26] V. Khullar, "Exploring techniques and approaches for time series analysis and its attacks," in *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 2024, pp. 1-6.

[27] S. Ravichandran and M. Afjal, "Investigating the impact of investor attention on AI-based stocks: A comprehensive analysis using quantile regression, GARCH, and ARIMA models," *PLoS One*, vol. 20, no. 5, p. e0324450, 2025.

[28] M. Afeef, A. Ihsan, and H. Zada, "Forecasting stock prices through univariate ARIMA modeling," *NUML International Journal of Business & Management*, vol. 13, no. 2, pp. 130-143, 2018.

[29] A. Suharsono, A. Aziza, and W. Pramesti, "Comparison of vector autoregressive (VAR) and vector error correction models (VECM) for index of ASEAN stock price," *AIP Conference Proceedings*, vol. 1913, no. 1, p. 020032, 2017.

[30] Z. Sun, "Comparison of trend forecast using ARIMA and ETS models for S&P500 close price," in *Proceedings of the 2020 4th International Conference on E-Business and Internet*, 2020, pp. 57-60.

[31] M. Mohan, P. Raja, P. Velmurugan, and A. Kulothungan, "Holt-winters algorithm to predict the stock value using recurrent neural network," *Methods*, vol. 8, no. 10, 2023.

[32] P. Sadorsky, "A random forests approach to predicting clean energy stock prices," *Journal of Risk and Financial Management*, vol. 14, no. 2, p. 48, 2021.

[33] P. H. Vuong, T. T. Dat, T. K. Mai, and P. H. Uyen, "Stock-price forecasting based on XGBoost and LSTM," *Computer Systems Science & Engineering*, vol. 40, no. 1, 2022.

[34] A. D. Hartanto, Y. N. Kholik, and Y. Pristyanto, "Stock price time series data forecasting using the light gradient boosting machine (LightGBM) model," *International Journal on Informatics Visualization*, vol. 7, no. 4, pp. 2270-2279, 2023.

[35] Y. Guo, S. Han, C. Shen, Y. Li, X. Yin, and Y. Bai, "An adaptive SVR for high-frequency stock price forecasting," *IEEE Access*, vol. 6, pp. 11397-11404, 2018.

[36] T. Ban, R. Zhang, S. Pang, A. Sarrafzadeh, and D. Inoue, "Referential k NN regression for financial time series forecasting," presented at the International Conference on Neural Information Processing, 2013.

[37] J. Fan and Y. Shen, "StockMixer: A simple yet strong MLP-based architecture for stock price forecasting," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 38, no. 8, pp. 8389-8397, 2024.

[38] J. Du, Q. Liu, K. Chen, and J. Wang, "Forecasting stock prices in two ways based on LSTM neural network," presented at the 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), 2019.

[39] Y. Gao, R. Wang, and E. Zhou, "Stock prediction based on optimized LSTM and GRU models," *Scientific Programming*, vol. 2021, no. 1, p. 4055281, 2021.

[40] H. K. Choi, "Stock price correlation coefficient prediction with ARIMA-LSTM hybrid model," *arXiv preprint arXiv:1808.01560*, 2018. https://doi.org/10.48550/arXiv.1808.01560