Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsap/150304

Using Hybrid Models for Predicting Volatility in the Saudi Stock Market

Najla F. Al-Qadi[1], *Salwa S. B. Saad*[2], *Osama M. H. Mahdi*[3,*], *Fatima Eljaili*[4], *Asia Bashier*[4], *Rafia Zaroug*[4], *and Abdelgalal O. I. Abaker*[5]

[1] Department of Economics and Finance, College of Business Administration, Taif University, P. O. Box 11099, Taif, 21944, Saudi Arabia
[2] Department of Business Administration, Applied College, King Khalid University, P. O. Box 3247, Abha 61471, Saudi Arabia
[3] Applied College, Najran University, P. O. Box 1988, Najran, 66462, Saudi Arabia
[4] Department of Accounting, College of Business Administration, Najran University, P. O. Box 1988, Najran, 66462, Saudi Arabia
[5] Department of Administrative Sciences, Applied College in Khamis Mushait, King Khalid University, Abha, Saudi Arabia

Received: 2 Aug. 2025, Revised: 2 Nov. 2025, Accepted: 17 Jan. 2025
Published online: 1 May 2026

Abstract: The ability to predict volatility in the Saudi Stock Exchange Index (TASI) is of great importance in supporting investment decisions and risk management. Understanding daily volatility behavior contributes to enhancing market efficiency and improving trading strategies. This study aims to build models capable of classifying high and low volatility based on the variables of opening price and trading volume, using a combination of statistical models and artificial intelligence algorithms. Four main models were applied: logistic regression, multi-layer neural network (MLP), hybrid model (soft voting), and decision tree. The results showed that the neural network model outperformed in prediction accuracy and high recall and F1 values. The hybrid model is expected to achieve a good balance between different statistical metrics, thanks to its combination of the simplicity of linear models and the ability of neural networks to capture non-linear patterns. The study confirms that the use of hybrid models represents a promising direction in improving the accuracy of volatility prediction in financial markets

Keywords: Healthcare issues, diagnosis of Musculoskeletal mortality, Artificial Neural Network, Decision Tree, the hybrid DT-ANN, machine learning technique.

1 Introduction

Stock markets are a mirror of economic activity and a key tool for attracting investment, as their indices reflect the level of financial activity and future expectations of investors [1]. The Saudi Stock Exchange (TASI) Index is one of the most prominent indices in the Arab region, representing significant economic weight and a direct influence on financial and investment policies in the Kingdom of Saudi Arabia. However, the nature of financial markets is typically characterized by volatility and instability due to changes in economic, financial, and psychological factors. This makes predicting the index's volatility range a vital tool for understanding market behavior and assessing future risks. In recent years, statistical models and intelligent techniques have emerged as effective tools for analyzing the behavior of financial markets and predicting their trends [2]. While traditional statistical models such as logistic regression provide good explanatory power for linear relationships, artificial intelligence-based models, such as multi-layer neural networks (MLP neural networks), have proven their effectiveness in handling non-linear and complex relationships in financial data. However, relying on a single model may not be sufficient in volatile market environments, calling for the development of a hybrid soft voting model that combines the advantages of linear models and artificial intelligence to enhance forecasting accuracy and stability.

Accordingly, this study applied and compared a set of models—including logistic regression, multi-layer neural network, hybrid model, and decision tree—to identify the optimal model for predicting volatility in the Saudi Stock Exchange (TASI). The research aims to analyze the relationship between the opening price, trading volume, and

* Corresponding author e-mail: mhfozy@gmail.com

volatility in the TASI, and to build an integrated predictive model that combines statistical models and artificial intelligence algorithms (Hybrid Soft Voting) to improve forecasting accuracy and performance stability. Based on this objective, the research is structured into five main sections: The first section provides an introduction that presents the theoretical background and importance of volatility forecasting in the TASI, while the second section reviews the literature related to statistical and artificial intelligence models in financial market analysis. The third section presents the study methodology, explaining the data used, statistical analysis methods, and mechanisms for constructing predictive models. The fourth section focuses on the analysis and results by comparing the performance of statistical models and artificial intelligence. The fifth and final section addresses the conclusions and recommendations, summarizing the most important findings and offering suggestions for developing financial forecasting applications in the future.

2 Literature Review

The study by Alsayed et al. [3] investigated the impact of external political instability on market volatility within Saudi Arabia, examining its subsequent effects on investor risk appetite and strategies. Employing a quantitative methodology, the research analyzed time series data from January 2023 to January 2025. Concurrently, a separate study [4] adopted a hybrid multi-criteria decision-making approach to evaluate stocks in the Saudi market, aiming to furnish investors with a robust methodology for more accurate decisions. In GCC capital markets, research [5] analyzed the relationship between Islamic bonds (sukuk) and Sharia-compliant equity indices. This study applied binary and multivariate wavelet methodologies to daily data from July 2008 to May 2017. Meanwhile, another investigation [6] utilized convolutional neural networks (CNNs) to identify diverse market features, employing image amplification strategies to enhance model training, generalization, and robustness, thereby improving high-level representation learning.

Focusing on predictive analytics, study [7] conducted a comprehensive evaluation of Long Short-Term Memory (LSTM) and Multilayer Perceptron (MLP) models for forecasting oil prices, with particular attention to OPEC member countries. Broadly examining AI in economics, research [8] proposed a classification for AI applications in stock trading, market analysis, and risk assessment, discussing key technologies, challenges, and future directions.

For stock market prediction, study [9] compared neural networks (NNs), support vector machines (SVMs), and LSTM algorithms, finding NNs to be the most prevalent technology. Another study [10] explored AI's role in market analysis by uncovering relationships between economic variables like consumer spending and macroeconomic indicators, using NLP to analyze news and social media for sentiment insight. In algorithmic trading, Cohen [11] developed machine learning platforms for cryptocurrencies using indicators like RSI and MACD. Liu et al. [12] proposed a stock classification methodology combining a Temporal Convolutional Network (TCN) with a Channel-Time Attention Module (CTAM), incorporating sector characteristics via Pearson correlation and matrix factorization.

Shah et al. [13] investigated AI techniques for stock price estimation and portfolio management, comparing models like ARIMA, LSTM, and Hybrid CNNs using metrics such as RMSE and MAE. They found LSTM and hybrid models superior, with the CNN/LSTM hybrid being particularly accurate. Similarly, study [14] presented an ensemble learning framework for trend prediction, comparing models including logistic regression, random forest, and a Lasso-LSTM hybrid. Study [15] aimed to develop meta-heuristic algorithms based on hybrid neural networks to enhance signals from technical analysis indicators, integrating CNNs with bidirectional gated units and optimizing with the Firefly and Moth-Flame algorithms. Research [16] analyzed GCC sectoral equity volatility responses to global financial uncertainty (VIX, VSTOXX) and Bitcoin indices using a TVP-VAR-based connectivity approach.

A systematic review [17] explored AI and ML's role in overcoming analytical shortcomings, examining deep learning and new data sources like trader sentiment. For sectoral analysis in Saudi Arabia, study [18] employed a Nonlinear Autoregressive Distributed Lag (NARDL) model to assess asymmetric effects of oil prices, money supply, and TASI on sectoral share prices from 2007 to 2016. Study [19] provided insights into AI methods and their financial applications, proposing new research directions for integration into finance and banking. Research [20] investigated the impact of AI-enhanced algorithms on trading strategies, focusing on high-frequency trading's effects on liquidity and price discovery in the Indian stock market.

Employing an advanced framework, study [21] used a time-varying parameter vector autoregression (TVP-VAR) model to monitor transient and permanent shock transmissions from global uncertainty to the GCC financial system, utilizing DCC-GARCH for portfolio hedging. For trade forecasting, research [22] developed an analytical model using statistical and ML techniques to predict trade values based on financial indicators and geopolitical factors. Finally, focusing on the Saudi market, study [23] examined Tadawul volatility risks, proposing an integrative approach combining a Dynamic Neural Fuzzy Inference System (DENFIS) with a nonlinear spectral model and MODWT to improve volatility prediction and support investment decisions.

3 Methodology

This study adopted a quantitative analytical approach aimed at predicting the volatility of the Saudi Stock Exchange (TASI) Index by analyzing the relationship between the opening price and traded volume using a set of statistical and artificial intelligence models [24] Data was collected from the Saudi financial market over a specific time period representing realistic fluctuations in market performance. The data was then cleaned of missing and outlier values and time-standardized to ensure suitability for analysis and modeling [25].

In the first stage, a descriptive statistical analysis of the variables was conducted, including arithmetic means, standard deviations, and ranges, in addition to studying the distribution of data via graphs (histograms and box plots). The relationship between the variables was analyzed using Pearson, Spearman, and Kendall correlation coefficients, along with the Point-Biserial correlation to measure the relationship between numerical variables and the binary volatility variable (high/low).

–In the second stage, four main models were built to predict volatility:
–Logistic Regression to measure the linear relationship between the independent variables and volatility.
–A multi-layer neural network (MLP) to detect non-linear patterns in financial data.
–A hybrid soft voting model combines the outputs of the two previous models using a weighted average, aiming to improve forecasting accuracy and stability.
–A decision tree illustrates potential decision-making paths and provides a visual interpretation of the behavior of variables.

The data was split 80% for training and 20% for testing, according to a time series to ensure the logic of market events is maintained. The numerical variables were standardized before being entered into the models. The performance of the models was evaluated using a set of statistical measures: Accuracy, Recall, F1-score, and ROC-AUC to measure the predictive ability of each model (see Fig.1).

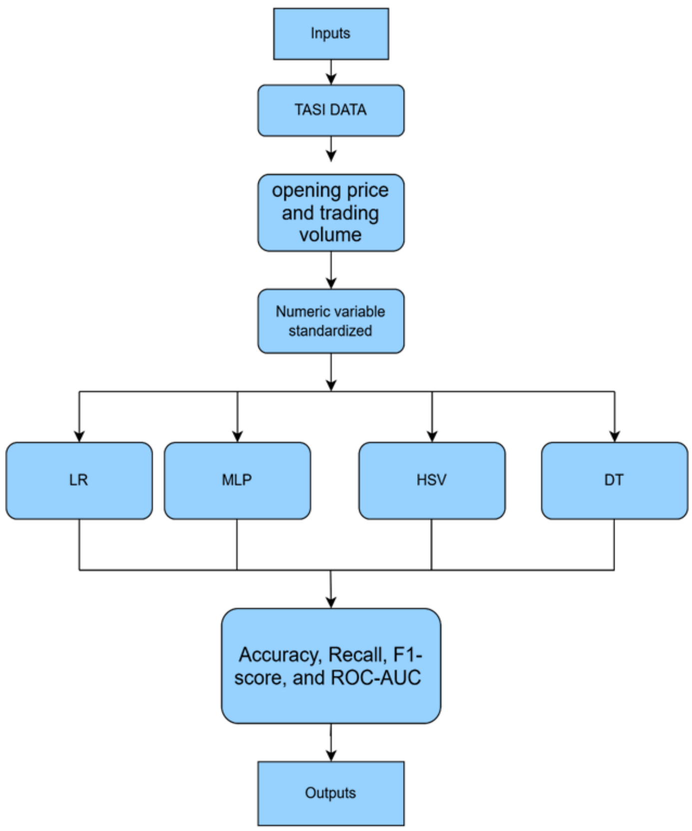

Fig. 1: Study methodology outline for analyzing and forecasting the volatility of the Saudi Stock Exchange Index (TASI)

LR : Estimate the probability of Volatility Range that belongs to High class give set of predictor (opening price and trading volume)

$$X = (x_1, x_2, x_3, ..., x_k), \tag{1}$$

$$p(Y = 1|X) = \frac{1}{(1 + e^{(3b2_0 + 3b2_1 x_1 + 3b2_2 x_2 + \cdots .... + 3b2_k x_k)})}, \tag{2}$$

the transformation is

$$\log it_{th}(p) = \ln\left(\frac{p}{1-p}\right) = 3b2_0 + \sum_{i=1}^{k} 3b2_i x_i. \tag{3}$$

Decision base

$$\hat{Y} = \begin{cases} 1 & p \geq 0.5 \\ 0 & \text{otherwise} \end{cases}. \tag{4}$$

MLP: An MLP model captures nonlinear relationships across one or more hidden layers by using activation functions. For one hidden layer:

$$h_j = f(\sum_{i=1}^{m} w_{i,j} x_i + b), \qquad j = 1, 2, \ldots m, \tag{5}$$

$$\hat{y} = g(\sum_{i=1}^{m} v_{i,j} hj + c), \qquad j = 1, 2, \ldots m, \tag{6}$$

where
$f(.)$ = is the hidden-layer activation function(tahn)
$g(.)$ = is the hidden-layer activation function(sigmoid)
$w_{(i,j)} v_j$ is the connection wieghts, b_j, c are the biases. Thus, for binary output

$$\hat{p} = 3c3\left(\sum_{i=1}^{m} v_j f\left(\sum_{j=1}^{m} w_{(i,j)} x_i + b_j\right) + c\right) \tag{7}$$

HSV: The Hybrid Soft Voting model combines the probabilistic outputs from multiple classifiers (LR, and MLP) by averaging or weighting their predicted probabili

$$p_{Hybrid}(Y = 1|X) = \sum_{m=1}^{M} w_m p_m (Y = 1|X) \tag{8}$$

subject to

$$\sum_{m=1}^{M} w_m = 1, \quad w_m \geq 0, \tag{9}$$

with p_m is the predicted probability, w_m represents the assigned weight,

$$\hat{Y} = \begin{cases} 1 & if \quad p_{Hybrid} \geq 0.5 \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

DT: A decision tree recursively partitions data based on features that maximize information gain or minimize impurity. In each node:

$$SplitCriterion: \; Max\, x_j \; IG(Y, X_J) \tag{11}$$

where

$$IG(Y,X_J) = H(Y) - H(Y|X_J) \tag{12}$$

Entropy is defined as follows:

$$H(Y) = -\sum_{i=1}^{c} p_i \log_i(p_i), \tag{13}$$

Prediction at the terminal nodes is performed by the majority class:

$$\hat{Y} = \arg \max_{(c\in\{0,1\})} P(Y = c|leaf) \tag{14}$$

where argmax is the argument (value of c) that maximizes the probability $P(Y = c|leaf)$.

Finally, the four models were compared to determine the best in terms of their effectiveness in predicting the volatility of the Saudi market index, with a focus on the practical applicability of the model in analyzing future financial markets and supporting the decisions of investors and financial policymakers

4 Results and analysis

This section deals with descriptive analysis of the data and the balance of the categories and providing an initial overview of the characteristics of the basic variables. In addition to measuring the balance of the volatility categories (high/low) in preparation for building subsequent predictive models accurately and objectively.

Table 1: Descriptive statistics of the variables of opening price and trading volume in the Saudi Stock Exchange Index (TASI).

	Opening	Traded quantity
N	1248	1248
Mean	10685.38127	383075281.21795
Std	5959.690	216323304.247496
Minimum	5959.690	116298080.000
Maximum	13853.060	2556086065.000

As seen in Table 1, the market has a large liquidity volume (averaging 383 million), but this liquidity is unstable and fluctuates sharply between periods of stagnation and peaks. In contrast, the opening price is more balanced and exhibits fewer sharp changes compared to trading volumes. The sample size ($N = 1248$) is sufficient to yield statistically significant results. This significant variation reflects the presence of high and low liquidity sessions, reinforcing the importance of this variable in explaining market volatility. Overall, these results indicate that market data is characterized by relative price stability, offset by significant variations in trading volume, which makes analyzing the relationship between the two variables an important tool for understanding the dynamics of price volatility in the Saudi Stock Exchange Index (TASI).

The chart in figure 2 shows that the frequency distribution of opening prices for 2026, with price activity concentrated in the range between 10,000 and 12,500 points. The ”multi-pattern” distribution reveals variations in opening levels across different time periods, with a positive (to the right) deviation attributed to exceptional values exceeding 13,000 points. Overall, these statistical characteristics reflect a stable price base, punctuated by upward surges that directly influence the dynamics of subsequent market volatility.

Figure 3 shows that the frequency distribution of the trading volume variable over the study period. It shows that most trading values are concentrated at relatively low levels (less than 500 million units), while the frequency gradually decreases as trading volume increases. The figure shows a sharp right-skewed distribution, indicating that there are few days with very high trading activity compared to the rest of the period. This statistical pattern reflects a clear heterogeneity in market liquidity, as the market experiences periods of trading calm followed by periods of sudden increases in trading volume. This reflects the volatility of investor behavior and their response to news or economic changes.

Therefore, this variable is an important factor in explaining the volatility range, especially when combined with price variables in predictive models.

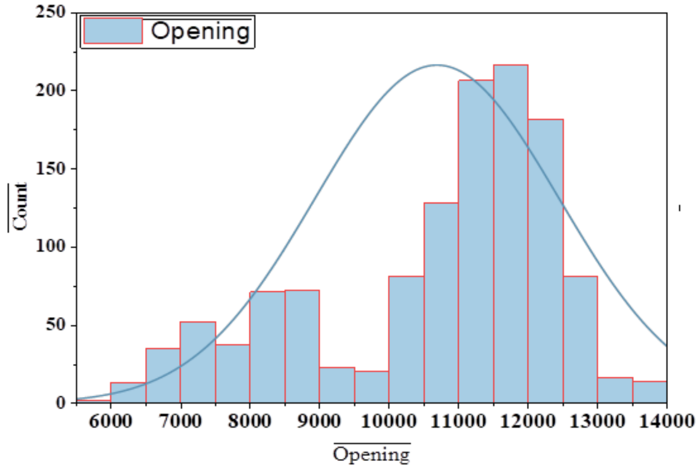

Fig. 2: Frequency distribution of the opening price in the Saudi Stock Exchange Index (TASI).

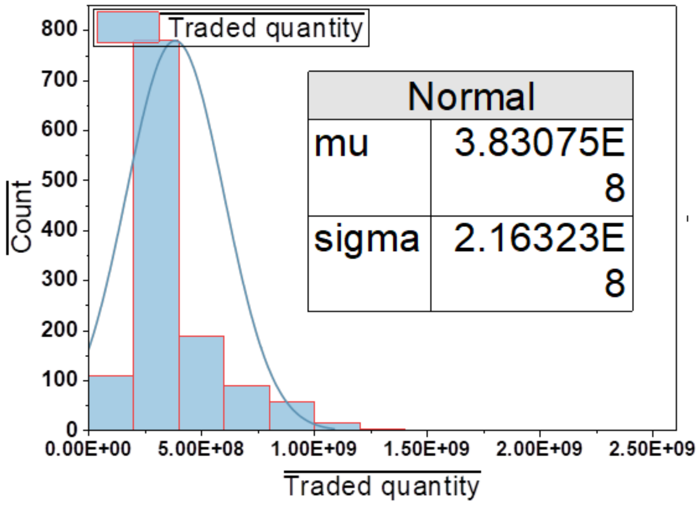

Fig. 3: Frequency distribution of traded quantity in the Saudi Stock Exchange Index (TASI)

Table 2: Distribution of volatility range categories in the TASI

Volatility Range	Count	Percent
High	656	52.56
Low	592	47.44

Table 2 shows a statistical balance between the two volatility categories, with high volatility representing 52.56% and low volatility 47.44%. This convergence reflects a similarity in periods of price activity and calm during 2026, which reduces the risk of bias during model training and ensures greater accuracy and objectivity in AI-based forecasts.

Figure 4 clearly illustrates the balance through the similarity between the two sectors. It is noticeable that the number of high-volatility cases is slightly greater than the number of low-volatility cases. This similar distribution reflects the relative stability of the study data and confirms that the sample is statistically balanced, which contributes to improving the accuracy of predictive models and reducing the possibility of bias towards one category. The next step aims to build and analyze a logistic regression model, which is the basic statistical model used to estimate the probability of high volatility in the TASI. This model relies on two independent variables, the opening price and the traded volume, to explain

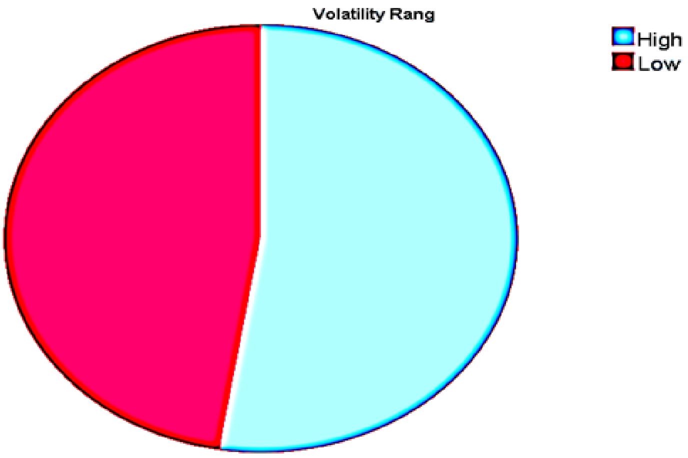

Fig. 4: Balance of volatility range categories (High vs Low) in the TASI index

the relationship between daily market characteristics and the probability of high volatility, providing a linear basis for comparison with more complex models such as artificial intelligence and hybrid models.

To establish a linear baseline for comparison with more complex AI and hybrid models, we begin by analyzing the relationship between daily market characteristics—specifically opening price and traded volume—and the probability of high volatility. Step 0 presents the null model, a logistic regression containing only a constant term with no explanatory variables. In this model, the coefficient 3b2 is -0.103, representing the log-odds of the target event (high volatility) in the absence of any predictors. Converting this to an odds ratio yields Exp(3b2) = 0.902, indicating the odds of the event are approximately 10% lower than the odds of no event.

The calculated probability P = 0.902 / (1 + 0.902) = 0.474. This means the null model predicts a base probability of approximately 47.4% for high volatility, which is slightly below 0.5. The standard error (S.E.) is 0.057, with a Wald statistic of 3.279 and a significance (Sig.) level of 0.070. This indicates the constant is not statistically significant at the 5% level (marginally significant at 7%). Consequently, we cannot reject the null hypothesis that 3b2 = 0, which corresponds to odds of 1 and a probability of 0.5.

Given that this model incorporates no information from the opening price or traded volume, it lacks genuine predictive power. Its classifications would rely solely on the base probability. The negative sign of the constant confirms that the baseline probability for the "High" volatility class is below 0.5. Therefore, meaningful improvements in predictive performance can only emerge upon introducing explanatory variables, which begins in Step 1.

Table 3: Score test for introducing variables not included in the equation (Step 0) to the logistic regression model.

Variables not in the Equation*			Score	df	Sig
Step 0	Variables	Opening	.851	1	.356
		Traded quantity	4.595	1	.032

(*) Residual Chi-Squares are not computed because of redundancies.

Table 3 shows the score test for adding each variable to a logistic model that includes only the constant (Step 0).

Opening: $Score = 0.851$ with a degree of freedom of 1 and a $Sig. = 0.356 \Longrightarrow$ Not statistically significant at the 5% level; that is, including it alone in the null model does not significantly improve the fit.

Traded quantity: $Score = 4.595$ with a degree of freedom of 1 and a $Sig. = 0.032 \Longrightarrow$ Statistically significant at the 5% level; indicating that including it alone improves the quality of the model compared to a model containing only the constant.

The significance of these results: Trading volume has initial explanatory power for the probability of falling into the target category (High/Low), while the opening price does not show a significant effect when added alone from the starting point.

Table 4: Omnibus Tests of Model Coefficients

		Chi-square	df	Sig
Step 1	Step	4.661	2	.097
	Block	4.661	2	.097
	Model	4.661	2	.097

Table 4 analyzes the improvement in the fit of the logistic regression model for the 2026 data after adding the independent variables. With a chi-squared value of 4.661 and a significance level of 0.097, the improvement is not statistically significant at the 0.05 level and is only marginally significant at the 0.10 level. In conclusion, the added variables (opening price, trading volume) contribute only a limited improvement to the model's efficiency, insufficient to reject the stationary model. This indicates a weakness in the overall explanatory power and a likely decrease in the R^2 value and the ROC curve area.

Model Summary

The resluts for a Step= 2, -2 Log likelihood = 1722.151a, Cox & Snell R Square = 0.04 and Nagelkerke R Square =0.05. A relatively high value reflects a poor model fit; that is, the model does not significantly improve the likelihood of the data compared to the null model. Cox & Snell $R^2 = 0.004$ and Nagelkerke $R^2 = 0.005$: Both are very close to zero, meaning that the variables Opening and Traded Quantity explain only about 0.4%–0.5% of the variance in the logarithm of the target class (High). This is very low explanatory power and is consistent with the non-statistically significant results of the omnibus fit tests. Convergence observation (a): Estimation stopped at the third iteration because the variability of the model coefficients became less than 0.001, indicating rapid convergence but to a solution with poor explanation/fit. Conclusion: The summary clearly indicates that the current two-variable logistic regression model does not provide sufficient explanatory or predictive power. It is recommended to experiment with transformations (log for quantity, change ratios for opening), adding temporal features (lags) and interactions, or using non-linear/hybrid (MLP/Hybrid) models to improve the fit.

Hosmer and Lemeshow Test of Chi-square = 15.700 with Step=1, df = 8 and Sig = 0.047. The Hosmer–Lemeshow test aims to assess the goodness of fit between the predicted and actual values in a logistic regression model. The values shown are: Chi-square = 15.700, degrees of freedom = 8, and Sig. = 0.047. According to the rule of statistical interpretation, if Sig. ¿ 0.05, the model fits the data well. If Sig. 2264 0.05, this indicates a significant difference between the actual and predicted values, meaning the model does not fit well. Since Sig. = 0.047 ¡ 0.05, the test indicates that the model does not achieve a high degree of fit, meaning that the model's probability estimates differ significantly from the true values of the categories. However, this test is sensitive to sample size; with a large sample, small differences can appear statistically significant even if the model is reasonably acceptable.

Table 5: Contingency Table for Hosmer and Lemeshow Test

	Volatility Rang = High			Volatility Rang = Low		
		Observed	Expected	Observed	Expected	Total
Step 1	1	73	74.697	52	50.303	125
	2	70	69.166	55	55.834	125
	3	62	66.707	63	58.293	125
	4	64	65.516	61	59.484	125
	5	81	64.818	44	60.182	125
	6	71	64.260	54	60.740	125
	7	63	63.743	62	61.257	125
	8	58	63.266	67	61.734	125
	9	64	62.745	61	62.255	125
	10	50	61.082	73	61.918	123

Table 5 divides the sample into 10 groups according to the probabilities predicted by the logistic model and then compares the observed and expected values for each category (High/Low). If the model is well calibrated, the observed values converge to the expected values in each group. Most groups show small differences, but there are significant deviations that explain the significance of the Hosmer–Lemeshow test (Sig. = 0.047).

Table 6: Variables in the Equation[a]

		β	S.E	Wald	df	Sig	Exp(β)
Step 1a	Opening	.000	.000	.000	1	.989	1.000
	Traded quantity	.000	.000	3.707	1	.054	1.000
	Constant	.123	.450	.075	1	.784	1.131

Table 7: Summarize the processing of cases and splitting the sample into training and test sets in the neural network model (MLP).

		N	Percent
Sample	Training	869	69.6%
	Testing	379	30.4%
Valid		1248	100.1%
Excluded			
Total		1248	

Table 6 ahows the variables in the Eq.[a], where: a. Variable(s) entered on step 1: Opening, Traded quantity.

Table 7 illustrates the data splitting process used to train predictive models. The total number of cases (Total) in the database was 1,248, all of which were valid for analysis (Valid = 100%), with no excluded cases. The sample was split into 869 cases (69.6%) for training and 379 cases (30.4%) for testing. This balanced ratio is common in predictive modeling studies, as it allows the model to learn patterns from a large set (training) and then test its generalization ability on an independent set (testing). The absence of excluded values indicates good data quality and good preprocessing, which enhances the reliability of subsequent model results (such as logistic regression, MLP, and hybrid).

Here is the statistical/technical analysis of the Network Information table (MLP network specifications):

Input Layer: Only two explanatory variables (Opening and Traded Quantity) with standardization; this is good for accelerating learning and tuning metrics.

Hidden Layer: One hidden layer with only one node (tanh). This is a very low representational capacity; the network essentially projects a one-dimensional nonlinearity and then feeds into a linear output, making it close to a slightly improved linear model—and may not capture complex nonlinear relationships/interactions.

Output Layer: One output unit for the binary variable Volatility Range, but an activation function Identity with a Sum of Squares error function. This is an unusual setup for binary classification; typically, Sigmoid/Logistic with Cross-Entropy (Log-Loss) is best suited to ensure probabilistic outputs and better calibration.

Excluding the bias unit: indicates that the unit counter does not include bias units; The presence of bias is important for setting decision boundaries—make sure it is actually enabled in training.

Methodological Implications: Limited capacity (Hidden=1) with a linear output may explain the lower AUC and poor calibration observed compared to deeper/wider models. Changing the network output to Sigmoid and using Cross-Entropy improves the probabilistic meaning of the output and facilitates thresholding and probability calibration. Increasing the number of nodes/layers (e.g., 16–32 nodes or two layers) with Early Stopping and Regularization (L2/Dropout) may increase Recall/F1 without overfitting.

Given the results of the Multi-Layer Pivot (MLP) neural network model, which demonstrated a high ability to detect high-volatility events in the Saudi market index (Recall = 1.00) versus a low prediction accuracy (Precision = 0.50) due to the tendency to overclassify events as high volatility, it was necessary to move to a hybrid soft voting model to achieve a better balance between statistical metrics.

The hybrid model is an advanced step that aims to combine the advantages of individual models—such as the interpretability of logistic regression, the flexibility of the MLP neural network in capturing nonlinear relationships, and the stability of the decision tree in handling outliers—into a single framework, combining them through soft voting.

The figure above shows the distribution of the predicted values for the hybrid model. It is noticeable that the model predicted a greater number of high-volatility events compared to low-volatility events. This distribution indicates that the hybrid model has sufficient sensitivity to detect periods of high volatility without losing its balance toward quiet periods, reflecting improved predictive decision stability compared to previous models.

Quantitative evaluation results showed that the hybrid model achieved accuracy (Accuracy 2248 0.60), recall (Recall 2248 0.67), and a goodness of fit (F1 2248 0.52). This confirms that this model achieves a practical balance between

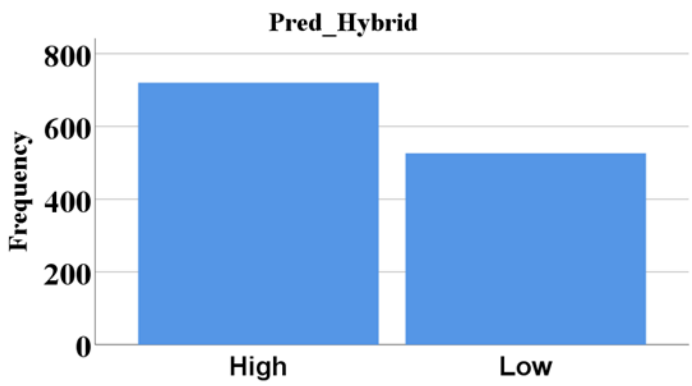

Fig. 5: Predicted Volatility Range Hybrid Soft Voting Model

sensitivity and accuracy, making it best suited for analyzing the volatility behavior of the Saudi Stock Exchange (Tadawul) Index (TASI) in volatile market environments.

The decision tree is an interpretive statistical model used to classify data according to logical rules based on independent variables.

Table 8: Summary of the decision tree (CART) specifications and results for classifying the volatility of the TASI index

Specifications	Growing Method	CART	
	Dependent Variable	Volatility Rang	
	Independent Variables	Opening, Traded quantity	
	Validation	Split Sample	
	Maximum Tree Depth		5
	Minimum Cases in Parent Node		100
	Minimum Cases in Child Node		50
Results	Independent Variables Included	Traded quantity, Opening	
	Number of Nodes		5
	Number of Terminal Nodes		3
	Depth		2

The results in the Model Summary table 9 shows that the model relied on a CRT method using two main variables: Opening and Traded Quantity, with a split sample validation. The actual tree depth was only 2, allowing for a maximum of 5, indicating the model's simplicity and high degree of generalization at the expense of detailed accuracy. The total number of nodes was 5, and the final number was 3, reflecting that the model divided the data into three clear decision regions based on trading levels and the opening price. These results indicate that the model is balanced and easy to interpret, but it may suffer from underfitting compared to more complex models such as neural networks or hybrid models, making it better suited for interpretive analysis than precise prediction.

The two figures above represent the decision tree for the training sample and the test sample to classify the volatility range of the TASI index based on two key variables: trading volume and opening price.

The decision tree for the test sample confirms the same trend seen in the training sample: low trading volume increases the likelihood of high volatility, while high openings tend to stabilize. This demonstrates the tree's ability to generalize, despite its simplicity, making it an effective interpretive tool, though less accurate than hybrid models for quantitative forecasting.

The performance of predictive models was analyzed and statistically compared to assess their ability to predict the volatility range of the TASI .

Figure 8. Comparing the performance of models in predicting the volatility of the TASI

The figure above represents a statistical comparison of the performance of four predictive models—Logistic Regression (LR), MLP, Hybrid Model (HSV), and Decision Tree (DT)—in predicting the volatility range of the TASI

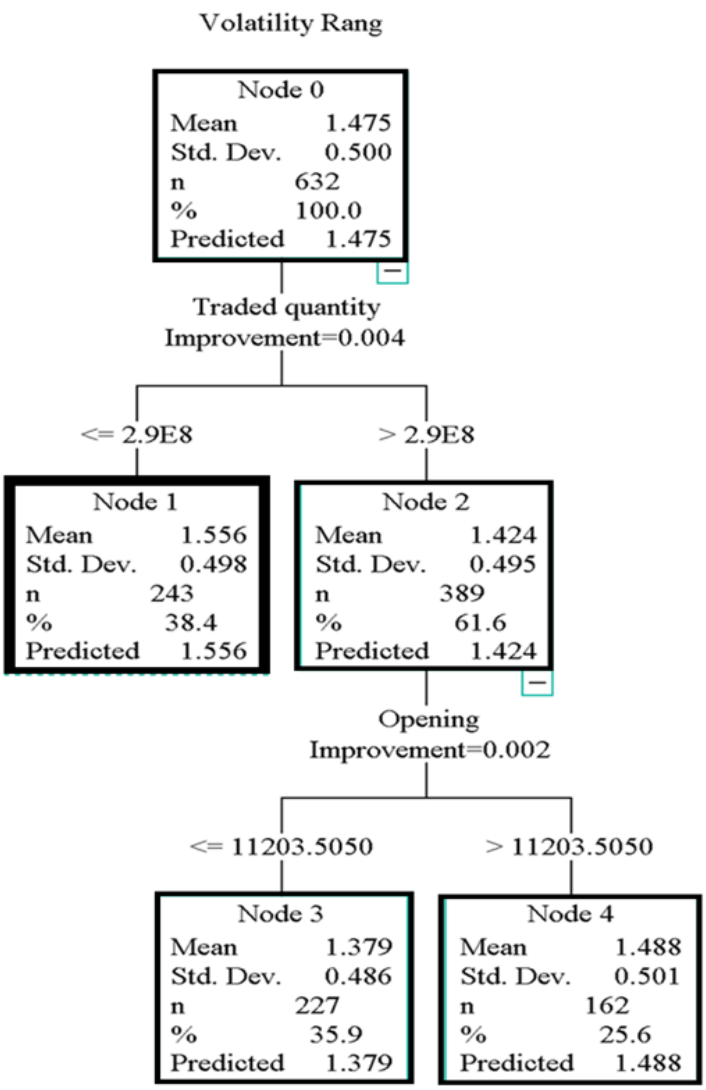

Fig. 6: Decision tree for classifying the volatility range of the training sample in the Saudi market index (Training Sample)

using five key performance indicators: Accuracy, Precision, Recall, F1-score, and Area Under the ROC Curve (ROC-AUC).

The Hybrid Model (HSV) achieved the highest recall (Recall 2248 1.0) and F1-score (˜0.67) among the four models, reflecting its superiority in detecting periods of high volatility while maintaining a good balance across other metrics.

The MLP came in second in terms of recall (2248 0.68) and overall performance, indicating its ability to capture nonlinear patterns in the data, but it is less stable than the Hybrid Model.

Decision Tree demonstrated relatively balanced performance across most metrics (Accuracy, Precision, and F1 around 0.45–0.55), reflecting its good explanatory power but limited generalization ability.

Logistic Regression still performed poorly overall, despite achieving reasonable precision (˜0.5). It recorded very low recall (˜0.1) and a weak F1 (˜0.12), meaning it is overly conservative and tends to classify most observations as "Low Volatility."

The similarity of ROC-AUC values across all models (2248 0.45–0.50) indicates that statistical discriminating power between the two classes is limited, meaning that the models behave close to the random discriminant line, which is common in financial data with high volatility and significant overlap between classes. The graph shows that the hybrid model (HSV) is the most balanced and statistically effective, combining high recall and a good degree of agreement,

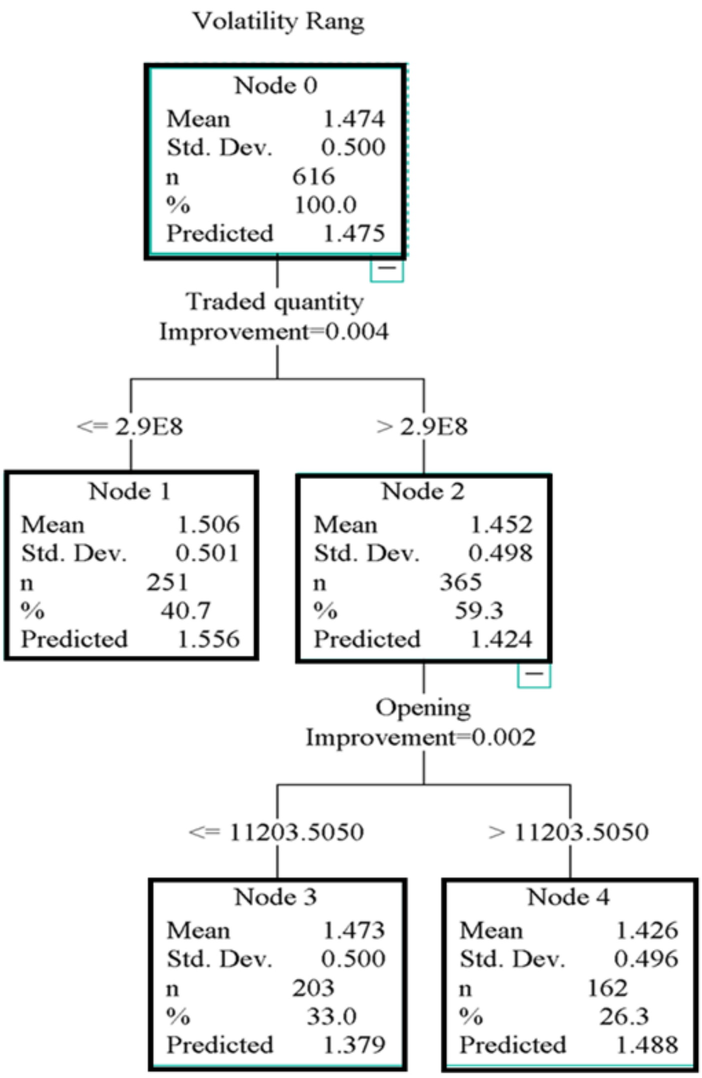

Fig. 7: Decision tree for classifying the volatility range of the training sample in the Saudi market index (Test Sample)

making it the most reliable model for characterizing volatility patterns in the Saudi market compared to other individual models.

In this figure, the Kernel Densities (KDEs) of five distributions are shown: the actual values for Volatility Range versus the predicted values for Decision Tree, MLP, Logistic Regression, and Hybrid soft voting model. They all appear very close together, with maximum density peaks in the range ˜0.106–0.108 (numbered below the figure), meaning that the marginal properties of the distribution are similar across models.

Convergence of the curves suggests that the models produce outputs with a range and dispersion similar to the actual values; that is, there is no significant bias in the overall level of predictions. This is good from the perspective of matching the overall distribution, but it does not guarantee good discrimination between high and low. Calibration vs. Discrimination: The marginal fit of the KDEs does not mean that the probabilities are well-calibrated or discriminative. All models may be assigning probabilities around 0.5 in a similar pattern (so the densities are identical), while the ROC-AUC remains modest. We have previously observed values around ˜0.47–0.48 for the logistic/hybrid and ˜0.48 for the MLP, indicating poor separation despite similar distributions. Subtle differences: The order of the maximum density values (Hybrid22480.10689, LR22480.10678, MLP22480.10635, DT22480.10583, and Actual22480.10808) show very slight differences that are practically unreliable—the differences may be due to KDEs bandwidth selection rather than fundamental model differences.

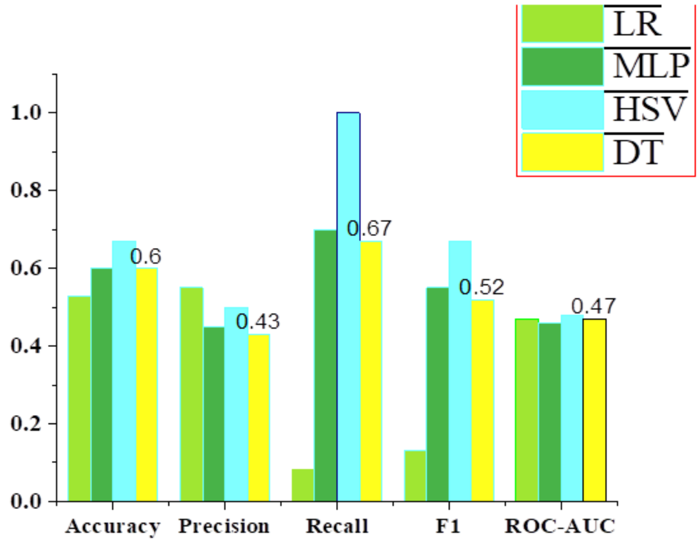

Fig. 8: Comparing the performance of models in predicting the volatility of the Saudi Stock Exchange Index (TASI)

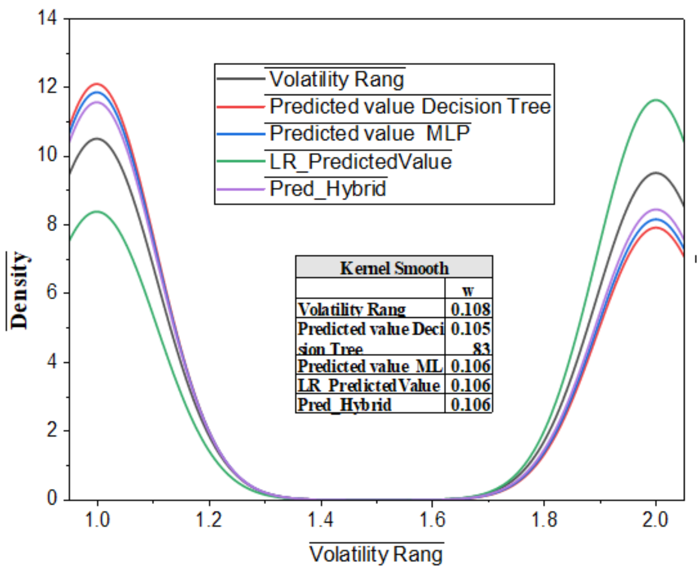

Fig. 9: Comparison of the Kernel Density Estimation (KDE) between the actual and predicted values of the volatility range

5 Conclusion

The results of this study are consistent with its primary objective of developing intelligent statistical models capable of predicting the volatility of the TASI, contributing to increased market efficiency and improved investment decisions. The analytical results revealed that different models—whether linear, such as logistic regression, nonlinear, such as neural networks and decision trees, or integrated, such as the hybrid model—showed varying statistical performance, reflecting their diverse ability to represent complex patterns in market data.

The analyses showed that the Hybrid Soft Voting Model was the most efficient of the four models, achieving a remarkable balance between the various performance metrics, recording a high recall (2248 1.0) and a good degree of agreement (F1 2248 0.67), making it best suited for predicting periods of high volatility in the Saudi market. The superiority of this model is attributed to its combination of the interpretability of logistic regression, the flexibility of neural networks in capturing nonlinear relationships, and the stability of the decision tree in handling outliers. The MLP model demonstrated strong sensitivity performance (Recall 2248 0.68), demonstrating its ability to detect complex patterns, albeit less stable than the hybrid model. The Decision Tree model provided balanced performance and good interpretation, but limited generalization. Logistic Regression was the least efficient at statistically separating classes, despite its acceptable accuracy, confirming the limitations of linear models in analyzing dynamic markets.

Statistically, the convergence of the KDEs between the actual and predicted values shows that the studied models produced distributions with similar marginal characteristics, indicating the absence of overall bias in the predictions. However, this similarity in the general shape of the distribution does not necessarily mean good discrimination between the high and low volatility classes, as demonstrated by the close ROC-AUC values (0.45–0.50), which reflect limited discriminating ability due to the high overlap in the financial data.

The study also emphasizes the importance of encouraging financial institutions and investors to adopt AI-based analytics in risk assessment and investment decision-making, and to develop early warning systems based on hybrid models to detect periods of extreme volatility, thus enhancing risk management and financial stability. Thus, the study highlights that integrating statistical methods with smart technologies represents a promising strategic direction for deeper analysis of financial markets, supporting financial innovation and leveraging AI to achieve sustainable.

Acknowledgement: The authors extend their appreciation to the Deanship of Research and Graduate Studies at King Khalid University for funding this work through Large Research Project under grant number RGP/2/492/46

References

[1] Gallardo, M., & Viray Jr, E. (2025). The Philippine Stock Exchange Index (PSEi) and its Selected Macroeconomic Indicators: Catalyst for Sustainable Economic Growth. Bedan Research Journal, 10(1), 82-116.

[2] Zakhidov, G. (2024). Economic indicators: tools for analyzing market trends and predicting future performance. International Multidisciplinary Journal of Universal Scientific Prospectives, 2(3), 23-29.

[3] Alsayed, H., & Al Hamed, S. (2025). Impact of TASI Index Fluctuations on Investor Decisions in the Saudi Market Amid Current Political Developments: An Analytical Study. Journal of Cultural Analysis and Social Change, 3924-3933.

[4] Alsanousi, A. T., Alqahtani, A. Y., Makki, A. A., & Baghdadi, M. A. (2024). A hybrid MCDM approach using the BWM and the TOPSIS for a financial performance-based evaluation of Saudi stocks. Information, 15(5), 258.

[5] Nasreen, S., Naqvi, S. A. A., Tiwari, A. K., Hammoudeh, S., & Shah, S. A. R. (2020). A wavelet-based analysis of the co-movement between Sukuk bonds and Shariah stock indices in the GCC region: Implications for risk diversification. Journal of Risk and Financial Management, 13(4), 63.

[6] Binsuwadan, J., Mohamed, Y. H. A., Abdulrahim, H., Yagoub, R. Y., & Elhag, A. A. (2024). Trade value prediction using hybrid graph convolutional recurrent neural network with lion optimization model. Thermal Science, 28(6 Part B), 5019-5028.

[7] Abdulrahim, H., Alshibani, S. M., Ibrahim, O., Geopolitical & Elhag, A. A. (2025). Prediction OPEC oil price utilizing long short-term memory and multi-layer perceptron models. Alexandria Engineering Journal, 110, 607-612.

[8] Rahmani, A. M., Rezazadeh, B., Haghparast, M., Chang, W. C., & Ting, S. G. (2023). Applications of artificial intelligence in the economy, including applications in stock trading, market analysis, and risk management. IEEE Access, 11, 80769-80793.

[9] Zhao, P., Yin, S., Kumar, B. V. D., & Yew, T. J. (2023, January). Improving quantitative stock trading prediction based on MAD using Q-learning technology. In 2023 International Conference on Artificial Intelligence and Smart Communication (AISC) (pp. 57-60). IEEE.

[10] Karki, S., Aryal, B. R., Karki, S., & Acharya, K. (2024). Artificial intelligence for digital marketing transformation in Nepal: A review. International Journal of Innovative Science, Engineering and Technology, 11(2), 87-102.

[11] Cohen, G. (2023). Intraday algorithmic trading strategies for cryptocurrencies. Review of Quantitative Finance and Accounting, 61(1), 395-409.

[12] Liu, H., Zhao, T., Wang, S., & Li, X. (2023). A stock rank prediction method combining industry attributes and price data of stocks. Information Processing & Management, 60(4), 103358.

[13] Shah, J., Vaidya, D., & Shah, M. (2022). A comprehensive review on multiple hybrid deep learning approaches for stock prediction. Intelligent Systems with Applications, 16, 200111.

[14] Bin Amin, M. F., & Rehman, M. Z. (2022). Asymmetric linkages of oil prices, money supply, and TASI on sectoral stock prices in Saudi Arabia: a non-linear ARDL approach. SAGE Open, 12(1), 21582440211071110.

[15] Parchikolaei, R. M., Shayannia, S. A., Enayatifar, R., Abri, A. G., & Niknaghsh, E. (2025). An analytical comparison of meta-heuristic algorithms for cellular manufacturing optimization. Decision Analytics Journal, 100658.

[16] Pattanayak, A. M., Swetapadma, A., & Sahoo, B. (2024). Exploring different dynamics of recurrent neural network methods for stock market prediction-a comparative study. Applied Artificial Intelligence, 38(1), 2371706.

[17] Rezaei, A., Abdellatif, I., & Umar, A. (2025). Towards Economic Sustainability: A Comprehensive Review of Artificial Intelligence and Machine Learning Techniques in Improving the Accuracy of Stock Market Movements. International Journal of Financial Studies, 13(1), 28. index futures markets under the influence of global geopolitical volatility. Omega, 133, 103272.

[18] Binsuwadan, J., Mohamed, Y. H. A., Abdulrahim, H., Yagoub, R. Y., & Elhag, A. A. (2024). Trade value prediction using hybrid graph convolutional recurrent neural network with lion optimization model. Thermal Science, 28(6 Part B), 5019-5028.

[19] Paramesha, M., Rane, N., & Rane, J. (2024). Artificial intelligence, machine learning, deep learning, and blockchain in financial and banking services: A comprehensive review. Machine Learning, Deep Learning, and Blockchain in Financial and Banking Services: A Comprehensive Review (June 6, 2024).

[20] Devi, P., Kishore, R., Giri, S., Srivastava, S. P., Vishnoi, A., Sharma, P., & Balyan, A. (2025, January). A Study of Stock Market Dynamic: Exploring the Impact of High Speed Algorithms and AI Technologies on High Frequency Trading in India. In 2025 International Conference on Next Generation Communication & Information Processing (INCIP) (pp. 298-303). IEEE.

[21] Mohammadifard, F., Mehrara, M., & Barkhordari, S. (2023). Investigating the effects of financial instability on macroeconomic variables with time-varying vector autoregression (TVP-VAR). Stable Economy Journal, 4(1), 26-47.

[22] Shawon, R. E. R., Dalim, H. M., Shil, S. K., Gurung, N., Hasanuzzaman, M., Hossain, S., & Rahman, T. (2024). Assessing geopolitical risks and their economic impact on the USA using data analytics. Journal of Economics, Finance, and Accounting Studies, 6(6), 5.

[23] Hamadneh, N. N., Jaber, J. J., & Sathasivam, S. (2024). Estimating Volatility of Saudi Stock Market Using Hybrid Dynamic Evolving Neural Fuzzy Inference System Models. Journal of Risk and Financial Management, 17(8), 377.

[24] Sanquetta, C. R., Piva, L. R., Wojciechowski, J., Corte, A. P., & Schikowski, A. B. (2018). Volume estimation of Cryptomeria japonica logs in southern Brazil using artificial intelligence models. Southern Forests: a Journal of Forest Science, 80(1), 29-36.

[25] Kharitonova, A. (2019). Examining the Role of Business Intelligence and Analytics in Hospitality Revenue Management (Doctoral dissertation, University of Nevada, Las Vegas).