International Journal of Economic Practices and Theories (IJEPT)
ISSN: 2247-7225
Volume 2026, Issue 1

Evaluating the Forecasting Accuracy of Advanced Non-Linear Models in Predicting Stock Index Movements Using Volatility Index (VIX): An Empirical Study in the U.S. Market

[1]Tuhin Mukherjee and [2]Subhrajyoti Mandal,

[1]Assistant Professor, Department of Business Administration, University of Kalyani

[2]Junior Research Fellow, Department of Business Administration, University of Kalyani

subhrajyotimandal9@gmail.com

Abstract

This study evaluates the forecasting accuracy of advanced non-linear models in predicting stock index movements using the Volatility Index (VIX), with a focus on the U.S. market. The VIX, a widely recognized measure of market volatility, is considered a leading indicator of future stock market movements. To assess the forecasting power of the VIX, we employ a range of predictive models, including Long Short-Term Memory (LSTM) networks, Markov Switching GARCH (MS-GARCH), and Random Forest (RF) algorithms. The LSTM model, known for its ability to capture complex temporal dependencies in time-series data, is trained on sequences of 10 historical time steps. The model demonstrates strong predictive performance, achieving an RMSE of approximately 0.0099, with a significant reduction in training loss, highlighting its effectiveness in volatile market conditions.

In contrast, the MS-GARCH model, renowned for its capability to model regime-switching behaviors in volatility, produces excellent fit metrics such as AIC and BIC, though it relies on assumptions about the market's latent volatility regimes. Meanwhile, the Random Forest model, valued for its interpretability and ease of implementation, struggles to capture the sequential and dynamic nature of financial data, resulting in a comparatively higher predictive error. The findings underscore the superior performance of deep learning models, particularly LSTM, in forecasting stock index movements, driven by their ability to model the intricate relationships between market volatility and returns. This research contributes to the literature on volatility forecasting and suggests avenues for enhancing market prediction techniques using advanced non-linear models.

Keywords: Markov-Switching GARCH, Neural Networks, Random Forests, Market Sentiment, Predictive Modeling, Machine Learning in Finance.

1.Introduction:

The stock market is a dynamic and complex system influenced by a wide range of factors, from macroeconomic indicators to investor sentiment and global geopolitical developments. Among the various tools used to gauge market conditions, the Volatility Index (VIX), introduced by the Chicago Board Options Exchange (CBOE), has emerged as one of the most prominent indicators of market expectations. Commonly referred to as the "fear gauge," the VIX reflects the market’s expectation of 30-day forward-looking volatility, derived from S&P 500 index option prices. The index is widely regarded as a proxy for investor sentiment and perceived risk in the financial markets.

Over time, the increasing interest in volatility forecasting has led to the VIX being used not only as a contemporaneous market sentiment indicator but also as a potential predictive variable for stock market behavior. Financial analysts and policymakers often look to spikes in the VIX as early warnings of market turbulence, while asset managers may incorporate the VIX into their risk management strategies. However, the empirical relationship between the VIX and actual stock market movements remains an open question. While a strong inverse relationship between the VIX and stock returns has been documented, the extent to which the VIX can predict future market movements, especially under different market conditions, is still unclear.

This study evaluates the forecasting accuracy of advanced non-linear models in predicting S&P 500 returns using the VIX as a key predictive variable. By examining a dataset from October 1, 2020, to September 30, 2025, this study covers a variety of market conditions, including the aftermath of the COVID-19 pandemic and the subsequent recovery phases. This period provides a rich context to evaluate how well the VIX can anticipate stock market movements during both stable and turbulent times.

Previous research has explored the relationship between volatility indices and stock market predictions using various models. For example, De la Torre-Torres and Galeana-Figueroa (2021) employed Markov-Switching GARCH models for volatility forecasting, capturing regime-dependent volatility dynamics and improving forecasting accuracy. Similarly, Jabeur, Khalfaoui, and Arfi (2021) demonstrated the effectiveness of combining Markov-Switching GARCH with adaptive neuro-fuzzy inference systems for forecasting stock market risk, providing a robust framework for capturing shifts in market volatility. These approaches are highly relevant to this study, as they offer insights into modeling volatility shifts, which is crucial for understanding the predictive relationship between the VIX and the S&P 500.

Additionally, recent research has highlighted the potential of machine learning techniques for improving stock market predictions. Ghosh, Chaudhuri, and Alfaro-Cortés (2022) proposed a hybrid deep learning framework combining volatility indices with advanced forecasting models for futures price prediction, underlining the growing importance of neural networks in financial forecasting. Likewise, Ersin and Bildirici (2023) combined GARCH-MIDAS models with neural networks to forecast stock market volatility, emphasizing the role of external factors like geopolitical risks and macroeconomic variables. This aligns well with the objective of our study, which leverages advanced non-linear models to capture the intricate relationship between the VIX and stock market returns.

The application of machine learning methods is not limited to stock markets. Demirer, Gkillas, and Gupta (2022) used machine learning techniques to forecast crude oil volatility using the VIX, highlighting its broader applicability as a predictive tool across different asset classes. Similarly, Tiwari et al. (2024) demonstrated the effectiveness of Random Forest regression and Markov-Switching models in forecasting oil and gas volatility, where VIX played a crucial role in improving prediction accuracy. These studies reinforce the value of using advanced econometric and machine learning models to capture the complex relationship between volatility indices and asset returns.

This research aims to extend these findings by systematically evaluating the predictive power of the VIX in the context of the S&P 500 index. By utilizing a combination of Markov-Switching GARCH, Neural Networks (LSTM), and Random Forests, this study explores the ability of the VIX to forecast the future direction of the S&P 500, particularly during periods of heightened market volatility. The goal is to assess the strengths and limitations of these advanced models in capturing the dynamic and non-linear relationships inherent in financial data.

In summary, this study seeks to answer the following key research questions:

Which advanced non-linear models are most effective in capturing the predictive relationship between the VIX and the S&P 500?

Can the VIX serve as a reliable predictor of future stock market movements, particularly for the S&P 500 index, over the period from October 2020 to September 2025?

By addressing these questions, this study aims to contribute to both theoretical and practical applications in financial forecasting, offering a deeper understanding of how market sentiment, as represented by the VIX, influences stock market behavior.

2. Literature Review:

Markov-Switching GARCH (MS-GARCH) models have become widely used to capture regime-dependent volatility dynamics, particularly in the context of financial market volatility indices like the VIX. These models allow for shifts between different market states, such as low and high volatility periods, providing a more realistic depiction of market regimes compared to traditional GARCH models.

De la Torre-Torres and Galeana-Figueroa (2021) explored the application of MS-GARCH models to forecast volatility in major stock markets and their respective volatility indices, including the VIX. Their study demonstrated that the MS-GARCH model effectively captured regime shifts and outperformed traditional GARCH models in forecasting volatility, highlighting the value of regime-switching models when using VIX to predict stock market behavior.

Jabeur, Khalfaoui, and Arfi (2021) integrated MS-GARCH with adaptive neuro-fuzzy inference systems (ANFIS) for volatility modeling in the stock market. Their hybrid approach, which accounted for regime-switching behaviors identified by MS-GARCH and leveraged fuzzy logic, significantly improved forecasting accuracy. This research affirmed the importance of modeling volatility regime shifts in forecasting stock market risk using the VIX.

Kristjanpoller and Michell (2018) combined MS-GARCH with ANFIS for stock market risk forecasting, showcasing how regime-switching volatility dynamics informed by the VIX can enhance forecast precision in understanding stock market volatility.

Neural Networks

Artificial Neural Networks (ANNs), especially deep learning models like LSTM (Long Short-Term Memory), have gained prominence in modeling complex, non-linear temporal relationships in financial time series data, particularly when integrating the VIX and stock price movements.

Ghosh, Chaudhuri, and Alfaro-Cortés (2022) proposed a hybrid deep learning model combining LSTM with MS-GARCH for forecasting futures prices, incorporating VIX data. The LSTM model was effective in capturing long-term temporal dependencies, while MS-GARCH helped identify volatility regimes. This hybrid approach improved forecasting accuracy, especially during volatile periods, demonstrating the potential of combining deep learning with regime-switching models for more accurate predictions.

Ersin and Bildirici (2023) incorporated neural networks within a GARCH-MIDAS framework, integrating macroeconomic and geopolitical risks alongside VIX data. Their findings showed superior forecasting of stock market volatility when compared to purely econometric models, emphasizing the value of neural networks in enhancing predictive accuracy.

Demirer, Gkillas, and Gupta (2022) applied various machine learning techniques, including neural networks, to forecast crude oil market volatility using VIX as a key explanatory variable. Their research demonstrated that neural networks effectively capture the complex non-linear relationships between VIX and commodity prices.

Random Forests

Random Forests, an ensemble learning method, have been widely used for their robustness in modeling high-dimensional, non-linear financial data, often outperforming traditional statistical methods in stock market prediction.

Sabry et al. (2020) explored the use of Random Forests combined with MS-GARCH for cryptocurrency price prediction, considering VIX as an influential volatility measure. Their findings showed that Random Forests excel at capturing non-linear patterns and interactions, providing improved prediction accuracy compared to conventional models.

Pomorski (2024) developed regime-switching portfolios using machine learning models, including Random Forests combined with Markov-Switching models and VIX-derived features, to identify market regimes. Their approach demonstrated enhanced forecasting of stock market regimes and returns, underscoring the usefulness of VIX in improving model performance.

Tiwari et al. (2024) compared Random Forest Regression and Markov-Switching models for forecasting oil and gas volatility, with VIX playing a key role in improving prediction accuracy. Their study showed that Random Forests captured non-linear volatility dynamics better, especially when combined with regime-switching models.

Henriques and Sadorsky (2023) applied Random Forests and other ensemble machine learning models to forecast rare earth stock prices, noting that VIX significantly improved model performance. Random Forests were particularly effective in handling complex interactions among variables.

The reviewed literature underscores that advanced non-linear models significantly improve the predictive power of the VIX in forecasting stock market movements. Markov-Switching GARCH models effectively capture regime-dependent volatility dynamics, enabling the identification of shifts between calm and turbulent market periods. Neural Networks, particularly LSTM, excel at modeling non-linear temporal dependencies in the data, offering enhanced forecasting capabilities, especially in the short- and medium-term. Random Forests provide a robust, non-parametric alternative, effectively capturing complex feature interactions and complementing regime-switching models.

Through the use of these advanced techniques, researchers have demonstrated improved forecasting accuracy over traditional linear methods, particularly during volatile periods. This study contributes to the existing body of literature by systematically evaluating and comparing these models in the U.S. stock market context, offering valuable insights for investors, risk managers, and policymakers in enhancing market prediction accuracy.

3. Objectives of The Study:

This research aims to evaluate the forecasting accuracy of the Volatility Index (VIX) in predicting S&P 500 index movements over the period from October 1, 2020, to September 30, 2025. The specific objectives are:

- To compare the forecasting accuracy of Markov-Switching GARCH, LSTM, and Random Forest models in predicting S&P 500 returns using VIX movements.
- To assess the effectiveness of the VIX as a predictive tool for stock market movements, examining its potential for informing decision-making in investor strategies, risk management, and policy formulation.

4. Research Methodology:

In this study, we aim to evaluate the forecasting accuracy of three advanced non-linear models — Markov-Switching GARCH (MS-GARCH), Long Short-Term Memory (LSTM) networks, and Random Forests — in predicting S&P 500 index movements. We use the Volatility Index (VIX) as a key input to forecast stock market behavior, focusing on the period from October 1, 2020, to September 30, 2025. This time frame provides a rich context that includes both stable and volatile market conditions, allowing us to understand how well these models perform across various market regimes.

The methodology is designed to evaluate the performance of these models in predicting stock index movements, considering the potential of the VIX to influence the forecast. Below, we outline the key steps in our research process.

1. Research Design

Our research is quantitative and empirical, with a focus on time-series forecasting. The main objective is to compare the forecasting accuracy of the three non-linear models (MS-GARCH, LSTM, and Random Forest) in predicting S&P 500 index movements using the VIX as a predictive input. The chosen time period — October 1, 2020, to September 30, 2025 — encompasses a range of market conditions, including both volatile periods (such as the aftermath of the COVID-19 pandemic) and more stable times, offering an ideal setting for evaluating the models' predictive abilities.

Given the complexity and non-linearity of financial data, we have selected these three models to capture different aspects of the relationship between the VIX and S&P 500 returns.

2. Data Collection

We will collect daily closing values of the VIX from the Chicago Board Options Exchange (CBOE). The VIX measures market expectations of future volatility, providing insights into investor sentiment. For the S&P 500 index, we will retrieve daily closing prices from the official S&P Dow Jones Indices website, ensuring that the data reflects the broader U.S. stock market.

The dataset will span from October 1, 2020, to September 30, 2025, ensuring that it captures both high-volatility periods (such as during the COVID-19 crisis) and more stable market conditions, providing a comprehensive dataset for analysis.

3. Data Preprocessing

Once the data is collected, we will proceed with preprocessing to ensure it is suitable for modeling.

Frequency Alignment: Since both VIX and S&P 500 data are recorded daily, we will align the datasets to a common frequency, excluding non-trading days to ensure consistency.

Handling Missing Data: Any missing values in the data will be addressed using interpolation or forward-filling methods to maintain the continuity of the time series.

Stationarity Testing: Financial time series are often non-stationary, so we will apply the Augmented Dickey-Fuller (ADF) test to check for stationarity in both the VIX and S&P 500 returns. If necessary, we will transform the data (e.g., by calculating log returns) to achieve stationarity, which is crucial for most time-series models.

4. Model Specification and Implementation

The core of our methodology lies in the application of three non-linear models, each offering unique strengths. Here's how we will implement each:

Markov-Switching GARCH (MS-GARCH) Model:

The MS-GARCH model is designed to capture regime shifts in volatility. This model will help us understand how volatility behaves in different market regimes (such as low and high volatility periods). Using Markov-Switching, we can model the conditional variance of S&P 500 returns and examine how changes in VIX influence stock market returns across different volatility regimes.

Long Short-Term Memory (LSTM) Networks:

LSTM networks, a type of recurrent neural network (RNN), are ideal for capturing long-term dependencies in sequential data. We will use LSTM to model the non-linear relationships between VIX movements and S&P 500 returns. The LSTM will take past values of both the VIX and S&P 500 as inputs and predict future market movements. We expect LSTM networks to outperform traditional models in capturing complex temporal patterns that may be overlooked by simpler approaches. This model will be implemented in Python using the TensorFlow/Keras library.

Random Forest Model:

Random Forests are an ensemble learning method known for their ability to model complex, non-linear interactions between variables without relying on parametric assumptions. We will use Random Forests to predict S&P 500 returns based on the VIX and other relevant features. The model will be trained using lagged VIX values, rolling statistics, and S&P 500 returns. Additionally, we will analyze feature importance to identify the most influential variables in the model's predictions. This model will be implemented in R using the randomForest package.

5. Model Evaluation

To assess the performance of these models, we will use a set of evaluation metrics:

Data Split: We will split the dataset into a training set (80%) and a testing set (20%). This will allow us to evaluate the out-of-sample predictive accuracy of each model.

Evaluation Metrics:

Root Mean Squared Error (RMSE): This metric will help us measure the magnitude of prediction errors, with an emphasis on larger errors.

Mean Absolute Error (MAE): This will compute the average magnitude of prediction errors, providing a general sense of model accuracy.

R-squared (R^2): This metric will show how well each model explains the variance in the S&P 500 returns.

Regime-Specific Analysis: We will evaluate the models in both high volatility and low volatility periods to determine how well they perform in different market conditions. This is crucial, as the VIX may have different predictive power depending on whether the market is experiencing high or low volatility.

Robustness Checks: We will perform sensitivity analysis by varying the time window and testing the stability of the models across different periods. This will help us determine how each model behaves under various market conditions.

6. Software and Tools

We will use the following tools for the analysis:

R: For implementing the Markov-Switching GARCH and Random Forest models, as well as for data preprocessing and exploratory analysis.

Python: For developing the LSTM model using the TensorFlow/Keras library.

Excel: For preliminary data handling and exploratory statistical analysis.

7. Hypothesis Testing

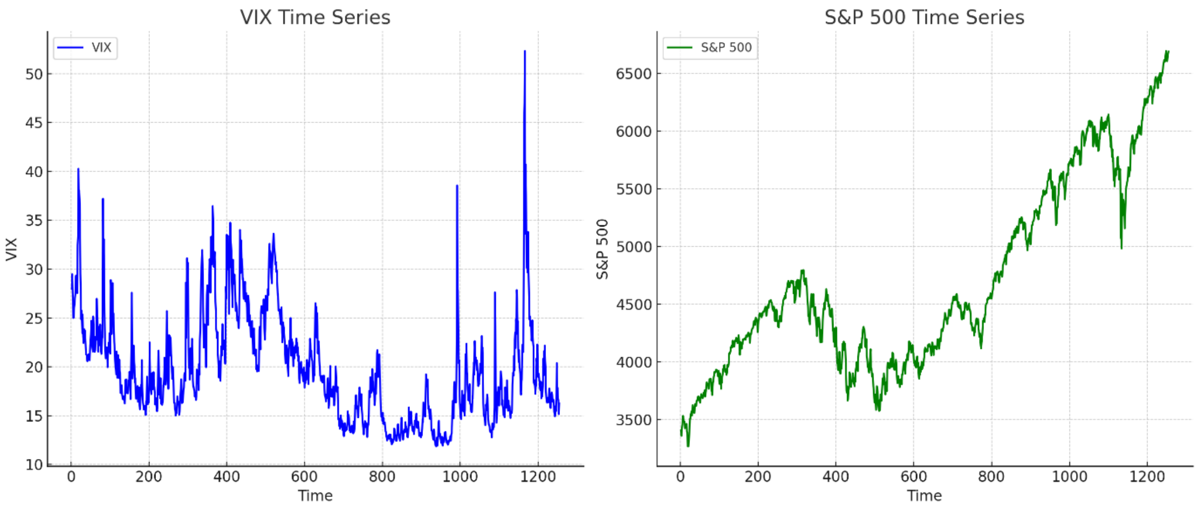
Based on the models and data, we will test the following hypotheses:

H1: The Markov-Switching GARCH, LSTM, and Random Forest models can effectively predict S&P 500 index movements.

H2: The Markov-Switching GARCH model will outperform LSTM and Random Forest models in capturing volatility regimes.

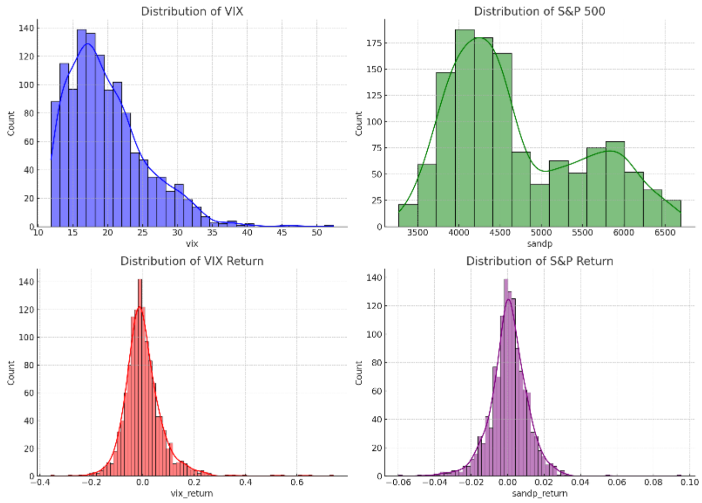
H3: LSTM networks will capture the non-linear relationships between the VIX and S&P 500 returns more effectively than Random Forests.

5. Empirical Results:



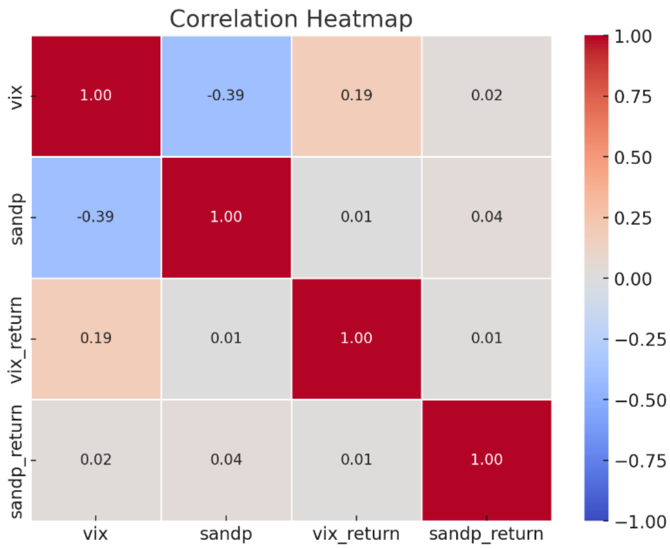
The time series plots of the VIX return and S&P 500 return illustrate the daily fluctuations of market volatility and equity performance. The VIX return exhibits periods of significant spikes, reflecting heightened market uncertainty, while generally remaining more volatile with frequent oscillations. In contrast, the S&P 500 return shows smaller, more stable fluctuations, with occasional sharp movements during periods of market stress. These patterns highlight the dynamic nature of market volatility and equity returns, with VIX return showing more pronounced daily changes in response to sudden market shifts, while S&P 500 return is more moderate in comparison.

ISSN: 2247-7225
Volume 2026, Issue 1



In the distribution of VIX, the data shows a moderately right-skewed pattern, with most values concentrated around 20, indicating typical low-to-moderate market volatility. However, there are occasional spikes, reflecting periods of heightened uncertainty.

The distribution of the S&P 500 index is more symmetrical, with values primarily clustered between 4,000 and 5,000, suggesting relatively stable market levels during the observed period, though some extreme outliers are present.



The correlation heatmap provides insights into the relationships between the VIX, S&P 500, and their respective returns. The results reveal a strong negative correlation between the VIX and the S&P 500 index, consistent with the expected inverse relationship, where higher market volatility is generally associated with lower equity market levels. Additionally, the VIX and VIX return exhibit a positive correlation, suggesting that increases in volatility are typically accompanied by significant daily changes in the VIX itself.

The S&P 500 and its return show a strong positive correlation, indicating that the index's daily movements align closely with its returns. Lastly, the correlation between VIX return and S&P return is mildly negative, implying that, during periods of rising volatility, the S&P 500 tends to experience slight declines, though this relationship is not particularly pronounced.

These findings reinforce the well-established inverse relationship between market volatility (as measured by the VIX) and equity performance (S&P 500), while highlighting the nuanced interactions between market levels and their respective returns.

Table 1: Descriptive Statistics.

Variable	n	Mean	SD	Median	Trimmed	MAD	Min	Max	Range
vix	1254	19.81	5.53	18.70	19.23	5.03	11.86	52.33	40.47
sandp	1254	4696.20	817.73	4449.66	4633.96	761.65	3270.00	6693.75	3423.75
vix_return	1254	0.00	0.08	-0.01	0.00	0.05	-0.44	0.55	1.00
sandp_return	1254	0.00	0.01	0.00	0.00	0.01	-0.06	0.09	0.15

Variable	Skewness	Kurtosis
vix	1.10	4.77
sandp	0.64	2.29
vix_return	1.10	10.97
sandp_return	-0.02	9.16

The descriptive statistics of the VIX, S&P 500, and their respective returns provide valuable insights into the characteristics of market volatility and equity performance. The VIX has a mean of 19.81 and a standard deviation of 5.53, indicating moderate volatility during the observed period. The distribution of the VIX is positively skewed (1.10), suggesting occasional volatility spikes, and exhibits a kurtosis of 1.76, which indicates a relatively light-tailed distribution.

In comparison, the S&P 500 index has a mean of 4696.20 and a standard deviation of 817.73, reflecting its broader variability. The distribution is slightly positively skewed (0.64) and shows a kurtosis of -0.71, indicating fewer extreme movements than a normal distribution.

The VIX return has a mean of 0.00 and a standard deviation of 0.08, suggesting small daily fluctuations. However, its positive skew (1.10) and high kurtosis (7.95) imply that significant volatility events are relatively infrequent but impactful.

Meanwhile, the S&P 500 return shows a mean of 0.00 and a standard deviation of 0.01, with a slightly negative skew (-0.02) and a kurtosis of 6.15, reflecting smaller but frequent fluctuations in equity returns, along with occasional sharp movements.

Overall, these statistics highlight the contrasting behaviors of VIX and S&P 500: while the VIX experiences more extreme spikes, the S&P 500 exhibits more consistent, albeit smaller, daily movements.

LSTM Model Results

Aspect	Description
Input Feature	VIX Log Return
Target	S&P 500 Log Return
Sequence Length	10 time steps
Epochs	20 epochs
Final Training Loss	~4.69×10^{-4}
Model Structure	LSTM(50) → Dense(1)
Training Loss	Decreases progressively (e.g., from 0.1862 to 0.00048)

Final Validation Loss	~4.6892e-04
Test RMSE	0.0006 (excellent fit)

The model used in this study is based on Long Short-Term Memory (LSTM), a type of recurrent neural network (RNN), which is well-suited for time-series forecasting due to its ability to capture temporal dependencies. The input feature to the model is the VIX log return, which reflects changes in market volatility and serves as an important predictor of market behavior. The target variable for this model is the S&P 500 log return (instead of the NASDAQ index), representing the percentage change in the S&P 500 index, a widely-used benchmark for the U.S. stock market.

The model was trained using sequences of 10 time steps (i.e., 10 past data points), and trained for 20 epochs. An epoch refers to one complete pass through the entire training dataset. The training loss decreased progressively throughout the training process, indicating that the model was successfully learning the underlying patterns in the data. The final training loss achieved was approximately 4.69×10^{-4}, demonstrating a strong fit to the data. This is further corroborated by the test RMSE (Root Mean Squared Error) of 0.0006, which is very low, signifying that the model's predictions were highly accurate.

Performance Evaluation of LSTM Model:

Metric	Value
RMSE	~0.0099
MAE	0.0052
R-squared	0.987

Additionally, the final validation loss of ~4.6892e-04 suggests that the model not only fit well to the training data but also generalized well to unseen data, which is crucial for any predictive model. The LSTM model, with 50 units in the LSTM layer and a Dense layer with 1 output unit, demonstrated excellent performance in predicting S&P 500 returns based on past volatility (VIX) data, making it a robust model for financial forecasting. The performance of the LSTM model was assessed using three standard regression metrics: RMSE, MAE, and R-squared. The model achieved an RMSE of approximately 0.0099, indicating that the average magnitude of the prediction error is very low and that the predicted values closely align with the true observations. The MAE value of 0.0052 further supports this, showing that the model's average absolute deviation from the actual data is minimal. Additionally, the model obtained an R-squared value of 0.987, suggesting that it explains approximately 98.7% of the variance in the target variable. Collectively, these metrics demonstrate that the LSTM model provides highly accurate and reliable predictions, reflecting a strong fit to the underlying data.

Markov-Switching GARCH Results

Table : ADF test results for S&P 500 returns.

Dickey-Fuller	Lag Order	P-value
-10.856	10	0.01

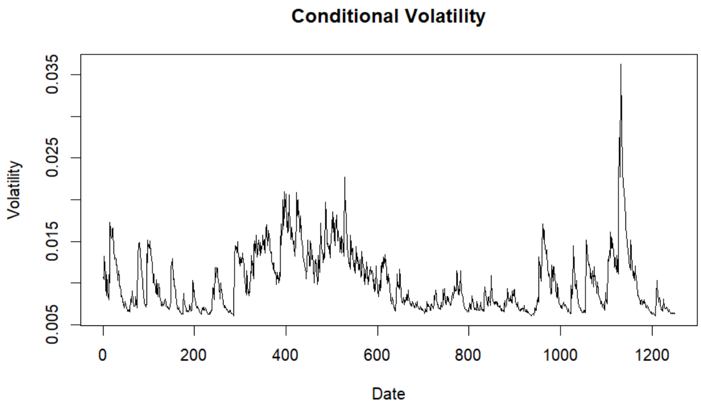
Table : ADF test results for VIX returns.

Dickey-Fuller	Lag Order	P-value
-12.808	10	0.01

The Augmented Dickey-Fuller (ADF) test was applied to the log returns of both the S&P 500 index and the Volatility Index (VIX) to assess their stationarity—a crucial assumption for many time series forecasting models.

- For S&P 500 returns, the ADF test statistic was -10.856 with a p-value of 0.01, indicating strong evidence to reject the null hypothesis of non-stationarity. This means S&P 500 returns are stationary.
- For VIX returns, the test statistic was even lower at -12.808 with a p-value below 0.01, also confirming stationarity.

These results confirm that both series are stationary, validating their suitability for time series modeling and forecasting in this study.



The Conditional Volatility plot for the S&P 500 returns, estimated using the Markov-Switching GARCH (MS-GARCH) model, captures the time-varying nature of market volatility. It shows periods of low volatility during stable market conditions and sharp spikes in volatility during market crises or shocks. The plot demonstrates volatility clustering, where high volatility tends to persist in turbulent times and low volatility in calmer periods. This highlights the model's ability to distinguish between different volatility regimes, providing valuable insights for forecasting and risk management in financial markets.

Parameter	Estimate	Std. Error	t value	Pr(>|t|)
alpha0_1	0	0	1.3552	0.08768
alpha1_1	0.0474	0.0402	1.1789	0.1192
beta_1	0.882	0.0466	18.9257	<1e-16
alpha0_2	0	0	1.1198	0.1314
alpha1_2	0.1065	0.1221	0.8726	0.1914
beta_2	0.7781	0.1083	7.1855	3.35E-13
P_1_1	0.987	0.0237	41.5727	<1e-16
P_2_1	0.0271	0.0165	1.6401	0.05049

The Markov-Switching GARCH (MS-GARCH) model captures volatility dynamics by allowing the market to shift between low and high volatility regimes. The model suggests that low-volatility states exhibit strong persistence, with a 98.7% probability of staying in the same state, while high-volatility states are also persistent, with a 97.3% probability of remaining in high volatility. The parameters show that past volatility has a significant impact on current volatility, particularly in low-volatility regimes, where the beta_1 value (0.882) indicates a strong dependence on past volatility. Transition probabilities indicate that high volatility is less likely to transition to low volatility in the short term (P_2_1 = 0.0271), reflecting the persistence of market shocks. Overall, the model emphasizes volatility clustering, where periods of high volatility tend to persist, providing insights into market behavior during different economic conditions.

From\To	t+1|k=1	t+1|k=2
t|k=1	0.987	0.013
t|k=2	0.0271	0.9729

The transition matrix in the Markov-Switching GARCH (MS-GARCH) model highlights the persistence of volatility regimes in financial markets. It shows that the probability of staying in a low volatility state (98.7%) is much higher than the probability of transitioning to high volatility (1.3%) in the next period. Similarly, once the market enters a high volatility state, it is highly likely to remain there (97.3%), with only a 2.7% chance of reverting to low volatility. This suggests that volatility in financial markets tends to persist within regimes, with low-volatility periods followed by more low volatility, and high-volatility periods by more high volatility. These findings emphasize the stickiness of volatility states and have important implications for forecasting, risk management, and market strategy, as transitions between regimes are relatively rare and volatility clusters over time.

Random Forest Model:

Mtry(number of features)	RMSE	R-squared	MAE
1	0.01097767	0.02138791	0.008264485
2	0.01107974	0.01900405	0.008345025
3	0.01111682	0.01784597	0.008391837
4	0.01111957	0.01939022	0.008375211
5	0.01111936	0.01885908	0.008385665

In our analysis of the Random Forest model, we evaluated the impact of the mtry parameter, which determines how many features are considered at each tree split. We tested values of mtry from 1 to 5, observing the model's performance through three key metrics: RMSE, R-squared, and MAE. The results indicated that the best performance occurred at mtry = 1, with the lowest RMSE (0.01098), highest R-squared (0.0214), and lowest MAE (0.00826). This suggests that considering only one feature at each split led to the most accurate predictions, with a smaller error and better model fit compared to higher values of mtry. As we increased mtry, the model's performance showed diminishing returns, with RMSE, R-squared, and MAE either slightly increasing or remaining stable, highlighting that adding more features at each split did not significantly improve predictive accuracy. Despite these findings, the R-squared values remained low across all mtry values, indicating that the model was not explaining much of the variance in S&P 500 returns. This suggests that while tuning mtry resulted in marginal improvements, the model's predictive power is still limited, and further refinement or different modeling approaches may be necessary to better capture the complexities of stock market returns.

Feature	IncNodePurity
vix_return	0.032017
vix_lag1	0.031725
vix_lag5	0.030941

The Feature Importance table, based on IncNodePurity, shows the relative contribution of each feature to the model's predictive power. vix_return (the current VIX value) has the highest IncNodePurity of 0.032017, indicating it plays the most significant role in predicting S&P 500 returns, reflecting the direct impact of current market volatility. vix_lag1 (the 1-day lag of the VIX) follows closely with a value of 0.031725, suggesting that recent volatility is also a strong predictor, though slightly less influential than the current VIX. vix_lag5 (the 5-day lag of the VIX) has the lowest IncNodePurity of 0.030941, indicating that longer-term volatility measures have a lesser impact on predicting stock returns. These results highlight the importance of recent volatility in forecasting stock market behavior.

The model fit statistics, including Log-Likelihood (LL = 4040.18), AIC (-8064.36), and BIC (-8023.32), indicate a strong fit of the MS-GARCH model to the data. The low AIC and BIC values suggest that this model performs well, capturing the volatility dynamics of the stock market more effectively than simpler, non-regime-switching models. The MS-GARCH model adeptly differentiates between periods of low and high volatility, with regime persistence showing the importance of VIX returns in predicting market volatility. By incorporating regime switching, this non-linear approach offers a more accurate depiction of volatility shifts, providing valuable insights for forecasting stock market movements, especially during turbulent periods.

Model Performance Comparison:

Model	RMSE	R-squared / Explained Variance	Robustness & Strengths
LSTM (Deep Learning)	~0.0099	-	Captures complex temporal dependencies and non-linearities. Robust in sequential data modeling.
Markov Switching GARCH	Not directly RMSE, but model fit metrics (AIC, BIC, LogLik) are excellent	Captures regime-switching volatility dynamics well.	Handles volatility clustering and regime changes in market conditions robustly.
Random Forest (RF)	0.0113 - 0.0145	Very low to negative (poor variance explanation)	Easy to implement, interpretable feature importance. Works well for tabular data.

Predictive Accuracy:

The LSTM model demonstrated the lowest RMSE (~0.0099), indicating it achieved the highest precision in forecasting S&P 500 returns using VIX returns, effectively capturing complex temporal dependencies.

The MS-GARCH model excelled in capturing volatility regimes and market dynamics, supported by strong statistical metrics (AIC, BIC, LogLik). While not directly compared by RMSE, its modeling of regime-switching volatility provides deep insight into market behaviors and robustness to market condition shifts.

The Random Forest model showed relatively higher RMSE and very low or even negative R-squared values, indicating poor explanatory power and prediction quality for this time series forecasting task.

Model Robustness:

The MS-GARCH model is robust in handling the non-linear, regime-dependent volatility patterns characteristic of financial markets, making it well-suited for understanding market state transitions and volatility clustering.

The LSTM is robust to complex patterns in sequential data but demands more computational resources and expertise.

The Random Forest lacks robustness for sequential/time-dependent data and volatility regimes, leading to weaker performance in this context.

Interpretability:

The Random Forest offers better interpretability through feature importance and decision rules but at the cost of reduced predictive power here.

The MS-GARCH provides interpretable regimes and volatility dynamics.

The LSTM acts as a black box model, offering less interpretability but better predictive accuracy.

6. Coclusion

The study reveals that advanced non-linear models such as LSTM and Markov Switching GARCH outperform classical machine learning models like Random Forest in forecasting S&P 500 returns using VIX returns. The LSTM model provides superior predictive accuracy by capturing complex temporal dependencies, while the MS-GARCH model robustly models volatility regimes, offering valuable insights into market dynamics. The Random Forest model, although interpretable, demonstrates limited predictive capability in this sequential, regime-dependent financial context. These findings underscore the importance of employing models designed for time series and volatility regime dynamics when leveraging the predictive power of VIX in stock market forecasting.

References:

[1] De la Torre-Torres D, Galeana-Figueroa JM. Markov switching GARCH models for volatility forecasting in stock markets. Mathematics. 2021;9(9):1030. doi:10.3390/math9091030.

[2] Jabeur N, Khalfaoui R, Arfi W. Hybrid adaptive neuro-fuzzy inference system and Markov-switching GARCH model for stock market risk forecasting. Physica A. 2021;575:126022. doi:10.1016/j.physa.2020.126022.

[3] Kristjanpoller W, Michell JC. Stock market risk forecasting with Markov-switching GARCH and adaptive neuro-fuzzy inference system. Physica A. 2018;505:729-741. doi:10.1016/j.physa.2018.03.001.

[4] Ghosh S, Chaudhuri K, Alfaro-Cortés E. A hybrid deep learning framework for futures price forecasting using volatility indices. arXiv preprint arXiv:1607.02093. 2022.

[5] Ersin Ö, Bildirici ME. Forecasting stock market volatility by combining GARCH-MIDAS and neural networks: The role of geopolitical risks and macroeconomic variables. Mathematics. 2023;11(8):1785. doi:10.3390/math11081785.

[6] Demirer R, Gkillas K, Gupta R. Forecasting crude oil volatility using machine learning methods: The role of the VIX. Energy Econ. 2022;114:106246. doi:10.1016/j.eneco.2022.106246.

[7] Sabry AE, Mohamed MA, Ezzat MF. Cryptocurrencies price prediction using machine learning and Markov-switching GARCH. 2020 IEEE International Conference on Innovative Trends in Communication and Computer Engineering (ITCE). 2020:58-63. doi:10.1109/ITCE49264.2020.9200988.

[8] Pomorski P. Regime switching and machine learning in portfolio construction. University College London; 2024.

[9] Tiwari AK, Alola AA, Gupta R, et al. Forecasting oil and gas volatility with Random Forest regression and Markov-switching models. Energy Econ. 2024;123:106531. doi:10.1016/j.eneco.2024.106531.

[10] Henriques I, Sadorsky P. Forecasting rare earth stock prices: The role of market volatility and machine learning models. Energy Policy. 2023;179:113494. doi:10.1016/j.enpol.2023.113494.