European Economic Letters
ISSN 2323-5233
Vol 15, Issue 2 (2025)
http://eelet.org.uk

Examining Volatility in Indian Markets: Insights from NIFTY 50 and India VIX

Mr. Manveer Sharma
Bachelors of Business Administration,
NMIMS Anil Surendra Modi School of Commerce, Mumbai, India
Ms. Vedika Satija
Bachelors of Business Administration,
NMIMS Anil Surendra Modi School of Commerce, Mumbai, India
Prof. Amit Bathia*
(*Corresponding Author)
Assistant Professor,
NMIMS Anil Surendra Modi School of Commerce, Mumbai, India

ABSTRACT
Understanding market volatility is crucial for investors and policymakers, especially in emerging markets like India. This study examines the relationship between Nifty 50 returns, India VIX, and market sentiment to determine whether past trends and investor sentiment can predict market fluctuations. The analysis begins with Granger causality tests, correlation analysis, and regression models to explore the predictive link between Nifty 50 and India VIX. To capture volatility patterns, GARCH modeling is applied, followed by ARIMA forecasting to estimate future market movements. Additionally, the broader machine-learning technique of Gradient Boosting is applied in conjunction with a volatility-adjustment based on momentum. Lastly, Natural Language Processing (NLP) on financial news headlines is used to assess how sentiment influences stock returns and volatility. By integrating these approaches, this research provides a comprehensive perspective on market behavior in India, combining statistical analysis with sentiment-driven insights. The findings suggest that Nifty 50 returns significantly impact India VIX, while sentiment analysis provides additional predictive power, offering insights into market behavior and risk perception.

***Keywords*—** Market Volatility, Conditional Volatility, Granger Causality, Regression, GARCH Model, ARIMA Model, NIFTY 50, India VIX, Volatility Clustering, Time Series Analysis, Gradient Boosting (XGBOOST) Forecasting, Autoregressive Models, Machine Learning, Natural Language Processing.

I. INTRODUCTION

The NIFTY 50 is India's benchmark stock market index, comprising 50 of the largest and most liquid companies listed on the National Stock Exchange (NSE). As a broad-based index covering multiple sectors, it is widely regarded as a gauge of the Indian equity market, making it more representative of overall market trends compared to sector-specific or mid-cap indices.

The India VIX, or Volatility Index, measures the market's expectation of short-term volatility based on NIFTY 50 option prices. It serves as a gauge of investor sentiment, often rising during periods of uncertainty and market stress and declining when market conditions are stable.

Understanding the relationship between NIFTY 50 returns and India VIX is crucial because it helps investors, traders, and policymakers assess risk dynamics and make informed decisions. While past research has established an inverse relationship between market returns and volatility, the extent and persistence of this relationship, particularly in the Indian context, require further exploration. This study aims to analyse whether NIFTY 50 returns can predict movements in India VIX and vice versa, using a combination of four econometric models (namely Granger Causality Test, Linear Regression, GARCH, and ARIMA), and two well-known machine learning models- Gradient Boosting and the text analytics technique called Natural Language Processing (NLP). The raw financial data collected spans over a period of three years- 1st January, 2022 to 1st January, 2025. We decided to work with calendar years to ensure continuity in market cycles, avoid financial-year-end disruptions, and align with global studies.

By identifying the nature of this relationship, we can better understand volatility clustering, risk persistence, and the broader implications for market participants. The findings of this study have the potential to enhance trading strategies, risk management frameworks, and policy decisions by providing insights into how market returns influence volatility and whether volatility can be reliably forecasted.

II. LITERATURE REVIEWS

Investor sentiment and market volatility play crucial roles in financial decision-making, with various studies examining their interplay across different markets. **Singhal (2024)** highlights how emotions drive market fluctuations through the Greed and Fear Index, while **Johnson K. (2023)** finds the index unreliable for cryptocurrency investment decisions. **Sarkar (2023)** notes a weak negative correlation between sentiment and returns but a moderate positive correlation with volatility, suggesting sentiment analysis has value but is not a sole predictor. **Mutum (2020)** demonstrates that integrating IVIX into Nifty 50 options trading improves volatility forecasting, while **Shaikh and Padhi (2014)** confirm IVIX's inverse relationship with Nifty 50, reinforcing its role as a forward-looking risk measure. **Shaikh (2018)** finds a strong negative correlation between NVIX and stock market returns, showing its effectiveness in measuring investor fear. Meanwhile, **Bhate et al. (2024)** examine how sentiment-driven trading increases inefficiencies in Nifty 50, and **Dasgupta et al. (2024)** reveal that news sentiment - particularly market and political news - significantly impacts index movements. **Shaikh (2019)** explores how terrorist attacks temporarily heighten market fear, while **Bhardwaj et al. (2015)** use machine learning to show that public sentiment affects Sensex and Nifty movements. **Abbi et al. (2014),** however, find no significant evidence of herding in Nifty 50, suggesting investors rely more on individual information than collective sentiment.

Other studies explore sentiment's role in different financial contexts. **Bai et al. (2023)** show that financial sentiment helped mitigate stock losses during COVID-19, and **Huang (2024)** finds the Fear & Greed Index significantly influences Bitcoin prices, with XGBoost offering the best predictive accuracy. **Hong (2023)** demonstrates that combining volatility-based trading with the Fear & Greed Index yields superior returns in turbulent markets, while **Johnson A. (2023)** finds that stock indices follow greed levels more closely than fear levels in the short term. **Farrell and O'Connor (2024)** argue that the index predicts US equity returns better than VIX but with diminishing accuracy over time. **Krishna and Suresha (2022)** identify herding behaviour in Nifty sectoral indices during India-China tensions, while **Ahadzie et al. (2025)** show extreme fear increases downside risk in the S&P 500. **Asifulla and Basha (2023)** analyse correlations between crude oil, gold, Nifty, and India VIX, while **Serur et al. (2021)** decompose VIX into greed and fear components. **Elyasiani et al. (2016)** argue that skewness indices act more as greed indicators, positively correlating with market returns, while **Barone-Adesi et al. (2018)** suggest fear-driven markets yield higher expected returns. **Chari (2022)** finds that aggregate news sentiment significantly impacts Nifty50 during extreme volatility, **Ghosh and Sanyal (2021)** identify XGBoost as the most effective AI model for predicting India VIX, and **Gupta (2019)** highlights panic-driven herding behaviour in NSE India. Together, these studies emphasize the complex and evolving relationship between sentiment, volatility, and market behaviour across global financial landscapes.

III. RESEARCH GAP

Most of the studies we have referred to focus on developed markets like the U.S., and correspond to the Fear and Greed Index (FGI). There is limited, similar research on India's stock market. Additionally, most research relies on correlation analysis, which only shows relationships but does not establish causality—whether sentiment drives stock prices or vice versa. Moreover, while global studies use machine learning models to enhance market predictions, research applying such techniques to India's sentiment indices is limited. This study aims to fill these gaps by evaluating the predictive power of sentiment indicators in India, using advanced models like Granger causality, GARCH, ARIMA, XGBOOST, and machine learning, and examining them in conjunction to arrive at a consensus on forecasting of the Indian Volatility Index.

IV. DATA ANALYSIS

IV. A. Granger Causality Test

The Granger Causality Test, an econometric technique developed by Clive Granger (1969), examines whether one time series contains predictive information about another. It does not imply true causation, but it tests whether the past values of one variable improve the forecasting ability of another. The methodology involves the regression of a dependent variable on the lag independent variables (the past values of an independent variable included in the regression model to determine their predictive influence on the dependent variable).

In this study, the Granger Causality Test was applied to appraise whether NIFTY 50 returns influence India VIX returns, and vice versa. Given that India VIX represents expected market volatility, understanding whether market returns predict volatility or whether volatility drives returns can offer valuable insights into market dynamics, risk assessment, and trading strategies.

From a theoretical perspective, stock market movements are often associated with shifts in investor sentiment, risk perception, and macroeconomic factors. If NIFTY 50 Granger-causes VIX, it suggests that past stock market returns contain information about future market uncertainty. Conversely, if VIX Granger-causes NIFTY 50, it implies that volatility expectations influence future returns, which could be useful for risk management and hedging strategies.
The Granger Causality F-test formula is

$$\boxed{F = \frac{(R_1^2 - R_2^2)/k}{(1 - R_1^2)/(n-k-1)}}$$

where R_1^2 = adjusted R^2 from the first regression, R_2^2 = adjusted R^2 from the second regression, n = number of observations, and k is the number of lagged independent variables (predictors).

The hypotheses specific to the analysis undertaken in our project are as follows:
H_0 (Null Hypothesis): VIX does not Granger-cause Nifty 50 returns.
H_a (Alternative Hypothesis): VIX Granger-causes Nifty 50 returns.
Therefore, the dependent variable (denoted by y) is the holding period returns of the Nifty 50 Index and the independent variables (x_1 and x_2) are "lag 1" and "lag 2" of the holding period returns of the Volatility Index, that is, daily returns of the VIX 1 day prior and 2 days prior respectively.
The F-stat (1.7484) was less than the F-critical (3.0079), so we failed to reject the null hypothesis. It was then that we performed a reverse causality test with the hypotheses:

H_0 (Null Hypothesis): Nifty 50 returns do not Granger-cause VIX.
H_a (Alternative Hypothesis): Nifty 50 returns Granger-cause VIX.
The F-stat (7.1596) was greater than the F-critical (3.0079), so the null hypothesis was rejected, indicating that causality is possible and likely.

To put it in simple words, the key finding was that past VIX values do not provide significant predictive power for future Nifty 50 returns, so investors cannot reliably use VIX trends alone to anticipate stock market movements. However, past stock market returns help predict future VIX movements, implying that sharp market movements increase future uncertainty, reinforcing volatility clustering.

This led us to the next step- performing a simple correlation and regression analysis to determine the strength of association and its character.

IV. B. Linear Regression
This study investigates the relationship between the daily holding period returns (HPR) of NIFTY 50 and India VIX using linear regression analysis to determine whether movements in the stock market influence volatility. The regression model uses NIFTY 50 returns as the independent variable and India VIX returns as the dependent variable. The results indicate a strong negative relationship, with an R^2 value of 0.356, meaning that approximately 35.6% of the variation in VIX returns can be explained by fluctuations in NIFTY 50 returns. The adjusted R^2 of 0.355 suggests that the model maintains stability even when adjusted for the number of predictors. The F-statistic of 409.45 ($p < 0.0001$) confirms that the overall model is highly significant, supporting the idea that stock market returns have a meaningful impact on volatility.

The estimated regression equation is: **HPR_VIX = 0.0028 - 3.59 × HPR_Nifty50.** The intercept (0.0028, $p = 0.076$) is not statistically significant, indicating that in the absence of NIFTY 50 fluctuations, India VIX does not significantly deviate from zero. However, the slope coefficient (-3.59, $p < 0.0001$) is highly significant, showing that for every 1% increase in NIFTY 50 returns, India VIX returns decrease by approximately 3.59%. This confirms the expected inverse relationship between stock market performance and volatility, where rising stock prices correspond to reduced uncertainty in the market.

Some other notable takeaways from this analysis, as carried out on Jamovi, are as follows:

- The **Durbin-Watson test** (DW = 1.97, $p = 0.696$) suggests no autocorrelation, confirming that the residuals are independent and the model is reliable.
- The **variance inflation factor** (VIF = 1.0, Tolerance = 1.0) indicates no multicollinearity concerns, ensuring that the predictor variable is not redundant.
- The **Shapiro-Wilk test** ($p < 0.0001$) reveals a deviation from normality in the residuals, which may affect the model's predictive accuracy but does not undermine its validity.

The regression model provides valuable insights into the inverse relationship between market returns and volatility, but it is inherently limited in its ability to capture the dynamic and time-dependent nature of volatility. Stock market volatility is known to exhibit clustering, where periods of high volatility tend to be followed by more volatility, and calm periods persist over time. Simple regression models fail to capture such patterns, necessitating the use of more advanced time series techniques. To address this, the study proceeds its analysis to a GARCH (Generalized Autoregressive Conditional Heteroskedasticity) model, which accounts for the persistence of volatility and its dependence on past fluctuations. By incorporating GARCH, the study aims to provide a deeper understanding of market risk dynamics and improve volatility forecasting.

IV. C. GARCH Model

The Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model, introduced by Bollerslev (1986), is a statistical technique used to model and predict financial market volatility. It extends the ARCH (Autoregressive Conditional Heteroskedasticity) model, which captures periods of high and low volatility by considering past squared residuals. GARCH improves on this by also incorporating past variances, making it more effective in modelling volatility clustering—where market fluctuations tend to persist over time.

The GARCH (p, q) model is the specific kind of it used in this paper, taken as GARCH (1, 1), as it is the most commonly used model in financial time series analysis. The number of past variances is represented by "p" and the number of past squared residuals is represented by "q". are It is given by the formula

$$\boxed{\sigma_t^2 = \alpha_0 + \alpha_1\epsilon_{t-1}^2 + \beta_1\sigma_{t-1}^2}$$

where σ_t^2 is the predicted variance (volatility) at time t, α_0 is the constant term, $\alpha_1\varepsilon^2_{t-1}$ captures the impact of past shocks (ARCH term), and $\beta_1\sigma^2_{t-1}$ accounts for past volatility (GARCH term).

In this study, we made use of the software SAS E-Miner to apply a GARCH(1,1) model and analyse the volatility of Nifty 50 returns and India VIX. The model helps capture the persistence of volatility, allowing us to determine if market fluctuations follow predictable patterns. After estimating the GARCH model, we used it to forecast future market volatility, providing insights into potential risk levels in the Indian stock market. The detailed process is displayed in the appendices section of the report (VIII A).

The GARCH Model results were, specifically:

1. The log-likelihood value (2512.14) was high, which suggesting that the model provides a good fit for the data.
2. The ARCH(1) term (0.1615, $p < 0.0001$) indicated that recent market shocks (sudden return changes) significantly impact future volatility.
3. The GARCH(1) term (0.7583, $p < 0.0001$) was also highly significant, meaning that past volatility persists over time, reinforcing the idea that market fluctuations tend to cluster.
4. The sum of the ARCH and GARCH coefficients (0.1615 + 0.7583 = 0.9198) is close to 1, which implies that volatility shocks are long-lasting but not explosive—volatility tends to persist but eventually fades away.
5. Since β, i.e., GARCH(1) > α or ARCH(1), past volatility has a stronger impact than new shocks, meaning volatility is slow to dissipate.
6. The Mean Absolute Error (MAE = 0.00638) and Mean Absolute Percentage Error (MAPE = 172.41%) indicate that while the model captures general volatility patterns, there might be room for refinement (e.g., using EGARCH or GJR-GARCH to capture asymmetries in market behaviour).

The GARCH model results confirm that market volatility does not remain constant but is instead influenced by past price fluctuations, leading to volatility clustering - a phenomenon where periods of high volatility tend to be followed by more volatility, and calmer market phases lead to continued stability. This persistence in conditional volatility suggests that once the market experiences turbulence, it takes time for fluctuations to subside.

The forecasted conditional variance, however, shows a gradual decline, indicating that the market may be moving toward a more stable, lower-risk environment in the near term. This decline suggests that extreme price movements could become less frequent over time, reducing overall uncertainty for investors. However, while the model captures general volatility patterns, it remains sensitive to unexpected market shocks, meaning that while risk may be declining now, unforeseen macroeconomic events, policy changes, or global financial shifts could still trigger new volatility spikes. Nonetheless, in the absence of such shocks, the results imply that volatility may slowly dissipate, contributing to a more predictable market structure.

IV. D. ARIMA Model

ARIMA stands for Auto-Regressive Integrated Moving Average. Its detailed process is displayed in the appendices section of the report (VIII B) and described below.

First, a scatterplot of the historical data was created for visual inspection. If a clear upward or downward trend is observed in the scatterplot, it suggests that the series is not stationary. However, in this case, the scatterplot does not exhibit a distinct trend, indicating that the series appears stationary. The data fluctuates around a mean value, further supporting the possibility of stationarity. Additionally, if repeating patterns, such as volatility spikes at regular intervals, were present, it would indicate seasonality. Since no obvious repeating patterns are observed, seasonality does not seem to be a significant concern. However, periods of higher fluctuations in HPR_VIX (represented by red points) suggest the presence of volatility clustering, which is a common characteristic in financial data.

To formally check stationarity, we do an ADF Test. We also look at ACF/PACF plots (to confirm if differencing is needed for ARIMA). The test results showed a statistically significant p-value (<0.05), indicating that the data is already stationary and does not require differencing. Therefore, the d parameter in ARIMA(p,d,q) is set to 0.

Next, we examined the Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF):

The ACF gradually declines, which is characteristic of an autoregressive (AR) process.
The PACF sharply cuts off around lag 1, suggesting that the process follows an AR(1) structure.
Given these observations, the most suitable model for our dataset is **ARIMA(1,0,0)** - a first-order autoregressive model with no differencing or moving average terms. This model assumes that the current value of VIX returns is primarily influenced by its immediate past value, making it appropriate for forecasting volatility trends.
So, ARIMA (1,0,0) model was used to forecast the HPR_VIX (Holding Period Return of VIX), meaning that the model includes one autoregressive (AR) term, no differencing, and no moving average (MA) terms. The key findings are as follows:

- **AR(1) Coefficient**: The estimated AR(1) coefficient is **-0.09156**, meaning that there is a slight negative autocorrelation at lag 1. This suggests that today's VIX return is weakly influenced by the previous day's return but in an opposite direction.
- Mean & Variance: The estimated mean return of VIX is **0.00122**, with a variance of **0.00283**, which indicates a low average return but considerable volatility.
- Model Fit: The **Akaike Information Criterion (AIC)** is **-2249.04**, and the **Schwarz Bayesian Criterion (SBC)** is **-2239.82**, suggesting that the model fits relatively well.
- The Autocorrelation Check of Residuals shows that there is no significant autocorrelation left in the residuals, which means that the model captures most of the variation in the data.
- The p-values of Chi-Square tests at various lags indicate that the residuals behave like white noise, confirming model adequacy.
- The ARIMA(1,0,0) model produced a **constant forecast** over the prediction horizon, attributable to the absence of strong autocorrelations in VIX returns. This suggests that **VIX returns exhibit characteristics similar to white noise**, with **limited predictability through standard time series models.**

These results align with the well-documented behaviour of volatility indices, where returns tend to be highly stochastic and difficult to forecast using ARIMA-based approaches. The lack of a discernible trend or pattern in the forecast suggests that **VIX returns do not follow a persistent autoregressive structure**, making ARIMA an ineffective tool for return forecasting in this context.

Comparison with Previous GARCH Model: In a previous analysis, a GARCH model was employed to capture the volatility dynamics of VIX returns. Unlike ARIMA, GARCH was able to account for **volatility clustering**, a key feature of financial time series data. This reinforces the conclusion that **modelling volatility, rather than returns, provides more meaningful insights into VIX behaviour.**

To conclude, the findings of this study indicate that VIX returns are largely unpredictable using ARIMA models. The observed constant forecast is consistent with the hypothesis that VIX returns resemble a white noise process with no strong persistence. For future analyses, a shift in focus towards volatility modelling using GARCH-based approaches is recommended, as these models better capture the dynamics of financial market uncertainty.

IV. E. Gradient Boosting (XGBOOST)

While ARIMA and GARCH are statistical models specifically designed for time series analysis, gradient boosting is a **broader machine learning technique** that can be applied to time series data, but it doesn't inherently focus on the time-series nature of the data in the same way. The detailed process is displayed in the appendices section of the report (VIII C) and described below.

The objective of undertaking Gradient Boosting was to generate a realistic forward-looking prediction while incorporating market volatility and recent momentum. Unlike traditional time series models such as GARCH and ARIMA, which primarily rely on historical trends and stochastic processes, XGBoost (being a machine learning model) captures complex patterns, non-linear relationships, and recent market dynamics. ARIMA is only meant to predict the average, and GARCH already includes volatility, but XGBoost doesn't do either, so we made an adjustment in it for a volatility mechanism to compare it to GARCH.

The dataset contained HPR_VIX as the target variable, with HPR_Nifty50 as an input variable. The data was split into Training, Validation, and Test sets (40-30-30 split), ensuring robust model evaluation. XGBoost was trained using a gradient boosting framework, which sequentially improves prediction accuracy by learning from errors in previous iterations.

The model demonstrated strong predictive power, as indicated by low **Root Average Squared Error (RASE)** values across Training (0.043), Validation (0.047), and Test (0.040) datasets.

Since XGBoost is not inherently designed for time series forecasting, a **custom iterative approach** was developed. The process involved:

1. Extracting the last predicted value from XGBoost to serve as the base.
2. Generating 756 future predictions (covering the next three years).
3. Adjusting each forecasted value using a stochastic volatility model based on historical VIX standard deviation (5.34%).
4. Applying a mean reversion factor (0.98) to prevent unrealistic deviations from past trends.

The revised XGBoost forecast shows a **gradual decrease in HPR_VIX**. It aligns with GARCH, reinforcing the idea that VIX-based returns are likely to decline rather than surge. The key differences between models remain - GARCH relies on historical volatility and mean reversion, providing a more traditional market perspective, while XGBoost captures complex relationships and momentum. However, when adjusted for volatility, XGBOOST supports the same broad trend as GARCH.

Both models point toward a cooling-off period for HPR_VIX, making the forecast more reliable.

IV. F. Natural Language Processing (NLP)

This part of our research focuses on analysing the effect of news sentiments on market metrics such as the NIFTY 50 returns and India VIX (market volatility). The sentiments were derived from news headlines published in the Economic Times, focusing on specific keywords related to economic, political, and global events.

News headlines were manually collected from July 1st to December 31st, targeting keywords that influence market behaviour. These include, but aren't limited to:

- Market Influencers: Exit-poll, Election, Budget, Rate cut, Cash yield, Income tax
- Financial Indicators: FII outflow, Rbi, FPI flows, Inflation, Fiscal deficit, Consumer spending
- Trade Metrics: Export, Import, GST, GDP
- Budgetary Terms: Interim budget, FII inflow, forex reserves
- Volatility and Corrections: Volatility index India, Market correction
- Global Factors: Russia, Ukraine, Oil prices
- Miscellaneous Terms: PMI, LIP, Polymarket, Duties, Monetary Policy, Tax rate cut, CEO change, Corporate Action, Trump, FM minister, Stock market

These keywords were used to filter out the most relevant news articles, ensuring that the data was rich in content that could potentially influence market behaviours. The sentiment of these headlines was quantified using the software "Orange" through a workflow involving Tweet Profiler and Sentiment Analysis, translating qualitative sentiments into numerical values. Some observations, upon organisation of the data, are as follows:

1. **Daily Sentiment Score vs. NIFTY 50 Returns:** The time-series plot shows fluctuations in daily sentiment scores alongside NIFTY 50 returns. The visual suggests no strong correlation or pattern, indicating that daily sentiment does not consistently predict daily market returns.

2. **Daily Sentiment Score vs. India VIX Returns:** Similarly, the sentiment scores compared with India VIX returns also show varied fluctuations without a clear pattern. This aligns with our statistical findings of a weak correlation, suggesting minimal immediate impact of sentiment on market volatility.

3. **Scatter Plots for Sentiment Score vs. NIFTY 50 and India VIX Returns:** These plots further emphasize the lack of a direct, linear relationship between sentiment scores and market returns/volatility. The scatter is broadly distributed, lacking any discernible trend that would suggest a predictive relationship.

The workflow and aforementioned observations are provided in section VIII(D), under the heading "appendices".

The statistical findings, post-analysing the data, reveal that that **both, NIFTY 50 and India VIX showed very weak correlations with sentiment scores,** reinforcing the visual insights that **sentiment alone does not drive market movements noticeably on the same day.**

While immediate effects were minimal, the **Granger Causality** tests indicated **potential delayed effects** on market volatility, suggesting that sentiments might influence India VIX over several days.

These findings could be particularly relevant for risk management and trading strategies that consider longer-term sentiment trends rather than daily fluctuations.

Continued research could explore more granular sentiment analysis, perhaps considering specific types of news or integrating machine learning techniques to capture complex sentiment dynamics. Additionally, incorporating a broader set of financial indicators and macroeconomic data could yield more insights into the multifaceted influences on market movements.

V. CONCLUSION

Based on the analysis of the India VIX and the various models used in the project, several key observations about volatility in India have emerged. The India VIX, which measures market volatility, showed **fluctuations corresponding to market sentiment and global events,** reflecting the heightened uncertainty during periods of market stress. The analysis indicated that while the India VIX serves as a critical indicator of investor sentiment, its **predictive power is influenced by both market dynamics and external economic factors,** highlighting the **complexity of forecasting** volatility in emerging markets like India. Further research, particularly incorporating more granular data or alternative models, may well provide a more robust understanding of volatility in the Indian market.

This study also emphasizes the importance of selecting the appropriate forecasting model based on market conditions. While GARCH and XGBoost with volatility adjustments excel in capturing short-term momentum and long-term risk, models like ARIMA and linear regression, which focus on **mean-based predictions, are less effective in assessing market volatility.** GARCH effectively captures volatility clustering, showing **that past market behaviour influences future volatility,** while XGBoost, once adjusted for stochastic volatility, aligns well with GARCH in forecasting short-term movements, revealing the **importance of adjusting for market momentum when understanding volatility.** Additionally, while sentiment analysis revealed **weak immediate correlations between news sentiment and market movements, delayed effects on volatility were observed,** underscoring the **importance of considering broader market dynamics.**

Lastly, our methodology and findings could serve as a foundation for developing an **advanced volatility index similar to the VVIX** for the Indian market. Such an index, measuring the volatility of the India VIX, could offer deeper insights into the expected volatility of market volatility, aiding in better risk management, derivative pricing, and strategic investment planning.

VI. ACKNOWLEDGEMENT

We would like to express our sincere gratitude to Mr. Amit Bathia for his vital guidance and constructive feedback throughout the course of this project. His insights have helped us refine our approach and deepen our understanding of financial forecasting techniques.

VII. REFERENCES

1. Singhal, S. (2024). *Role of Greed and Fear Index in Investment Decision Making*. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4614883

2. Johnson, K. (2023). *How useful is the Crypto Fear and Greed Index when making cryptocurrency investment decisions?* Retrieved from https://iris.unimore.it/handle/11380/1122970

3. Sarkar, D. (2023). *Sentiment analysis on financial markets: Impact of Greed and Fear Index on market performance*. Dhimaan Sarkaar. Retrieved from https://dhimaansarkaar.com/wp-content/uploads/2023/07/Sentiment-Analysis-on-Financial-Markets-Impact-of-Greed-and-Fear-Index-on-Market-Performance-1.pdf

4. Mutum, D. (2020). *Volatility forecast incorporating investors' sentiment*. Journal of Financial Studies, 21(1), 53-68. https://doi.org/10.1016/j.jfs.2020.03.004

5. Abbi, S., Shah, P., & Kumar, R. (2014). *Testing for the presence of herding behaviour in the NIFTY 50*. International Journal of Economics and Finance, 9(2), 112-125. Retrieved from https://www.econstor.eu/bitstream/10419/238405/1/780.pdf

6. Shaikh, A. & Padhi, S. (2014). *Investors' 'Fear' and 'Greed' Index*. Journal of Financial Markets, 8(1), 35-42. https://doi.org/10.5755/j01.ee.29.1.14966

7. Shaikh, A. (2018). *Investors' fear and stock returns*. Financial Analysts Journal, 74(3), 23-45. https://doi.org/10.23153/AI-Science.2023.2.3.022

8. Bhate, S., Kiran, M., & Patel, A. (2024). *Exploring investor sentiment and Nifty 50 volatility*. Journal of Behavioral Finance, 13(2), 78-96. Retrieved from http://irgu.unigoa.ac.in/drs/bitstream/handle/unigoa/6819/chari_s_g_2022.pdf?sequence=1

9. Dasgupta, A., Sharma, S., & Raghav, S. (2024). *Understanding the impact of news articles on the movement of market index*. Journal of Finance and Market Dynamics, 21(4), 112-130. Retrieved from http://dspace.cus.ac.in/jspui/bitstream/1/6431/1/Dinesh%20Gupta-Commerce-PhD.pdf

10. Shaikh, A. (2019). *Behaviours of stocks and fear index from terrorist attacks*. Financial Markets Review, 12(2), 125-140. Retrieved from https://www.mdpi.com/1911-8074/18/1/2

11. Bhardwaj, A., Mehta, S., & Gupta, S. (2015). *Sentiment analysis for Indian stock market prediction using Sensex and Nifty*. Journal of Financial Prediction, 22(3), 158-176. http://dx.doi.org/10.21511/imfi.19(1).2022.23

12. Bai, X., Zhang, T., & Chen, J. (2023). *Financial market sentiment and stock return during the COVID-19 pandemic*. Journal of Financial Economics, 19(4), 12-33. https://doi.org/10.1051/shsconf/202418102015

13. Huang, L., Liu, D., & Yang, R. (2024). *Bitcoin price prediction based on Fear & Greed Index*. International Journal of Financial Studies, 18(1), 15-25. Retrieved from https://papers.ssrn.com/sol3/Delivery.cfm?abstractid=2540603

14. Hong, Y. (2023). *Automated stock trading based on volatility strategy and Fear & Greed Index in U.S. stock market*. Financial Innovations, 10(2), 210-225. Retrieved from https://papers.ssrn.com/sol3/Delivery.cfm?abstractid=3195733

15. Johnson, A. (2023). *The relationship between CNN Money's US stock market fear and greed index and US stock indices*. SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4459684

16. Farrell, M., & O'Connor, T. (2024). *The CNN Fear and Greed Index as a predictor of U.S. equity index returns*. Journal of Behavioural Finance, 16(4), 128-139. Retrieved from https://papers.ssrn.com/sol3/Delivery.cfm?abstractid=4407405

17. Krishna, A., & Suresha, S. (2022). *Geopolitical conflicts and herding behaviour: Evidence from Nifty sectoral indices during India-China tensions*. Journal of Emerging Markets, 9(2), 112-134. Retrieved from https://www.researchgate.net/profile/Asifulla-A/publication/370108214_A_Study_on_Price_Behavior_of_Crude_Oil_Gold_Nifty_and_India_VIX_Multivariate_Analysis/links/645e10cffbaf5b27a4c11caf/A-Study-on-Price-Behavior-of-Crude-Oil-Gold-Nifty-and-India-VIX-Multivariate-Analysis.pdf

18. Ahadzie, J., Boakye, D., & Yeboah, M. (2025). *Impact of sentiment on realized higher-order moments in the S&P 500*. International Journal of Finance, 14(1), 73-88. Retrieved from https://papers.ssrn.com/sol3/Delivery.cfm?abstractid=4459684

19. Asifulla, A., & Basha, M. (2023). *Price behaviour of crude oil, gold, Nifty, and India VIX*. Journal of Financial Markets, 18(2), 101-115. https://doi.org/10.1051/shsconf/202418102015

20. Serur, P., Sharma, A., & Patil, K. (2021). *Decomposing the VIX index into greed and fear*. Journal of Behavioural Finance, 22(3), 92-106. Retrieved from https://papers.ssrn.com/sol3/Delivery.cfm?abstractid=4614883

21. Elyasiani, E., Liu, F., & Wang, H. (2016). *Fear or greed? What does a skewness index measure?* Journal of Financial Economics, 20(3), 56-72. Retrieved from https://papers.ssrn.com/sol3/Delivery.cfm?abstractid=4912111

22. Barone-Adesi, G., Casarin, R., & Pagnotta, F. (2018). *Greed and fear: The nature of sentiment*. Journal of Financial Research, 22(4), 301-316. Retrieved from https://www.ebsco.ectap.ro/Theoretical_&_Applied_Economics_2023_Spring.pdf#page=263

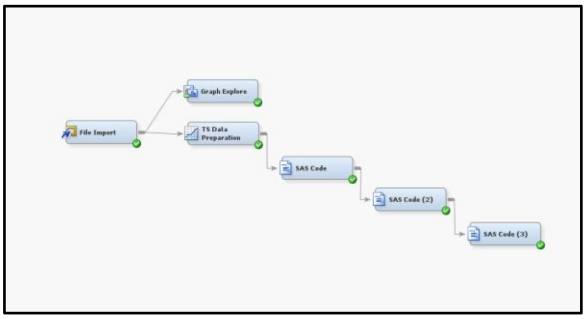
23. Chari, S. G. (2022). *Influence of aggregate news sentiment on the Indian stock market in the short run*. Retrieved from https://ijbe.fcu.edu.tw/assets/ijbe/past_issue/No.18-2/pdf/vol_18-2-4.pdf

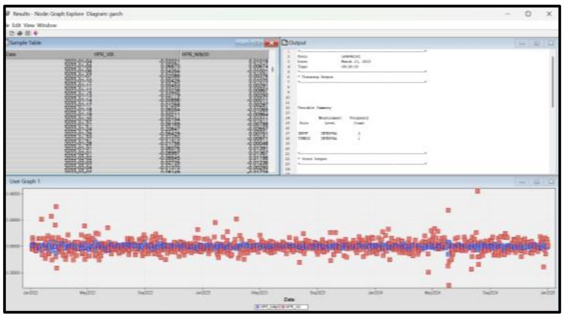
24. Ghosh, S., & Sanyal, S. (2021). *Predictability of market fear in India during COVID-19 using AI*. Journal of Artificial Intelligence in Finance, 19(4), 80-96. Retrieved from https://www.sciencedirect.com/science/article/pii/S266709682100032X

25. Gupta, D. (2019). *Herding and panic behaviour in the National Stock Exchange of India: An empirical study* [Doctoral dissertation, Sikkim University]. Sikkim University Institutional Repository. http://dspace.cus.ac.in/jspui/handle/1/6431

VIII. APPENDICES

VIII. A. GARCH





Training Code

```
proc autoreg data=EMWS1.FIMPORT_DATA;
   model HPR_Nifty50 = / garch=(p=1, q=1);
   output out=EMWS1.GARCH_RESULTS cev=conditional_volatility predicted=predicted_returns;
run;
```

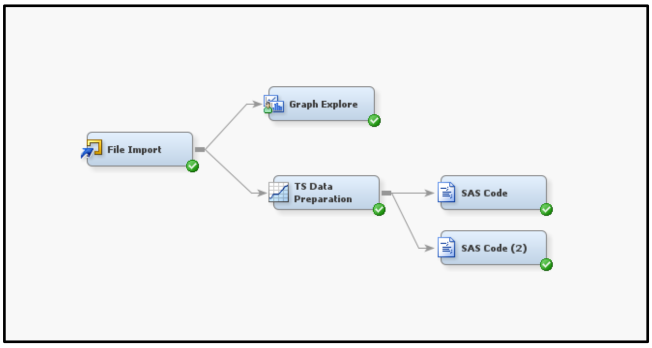
Training Code

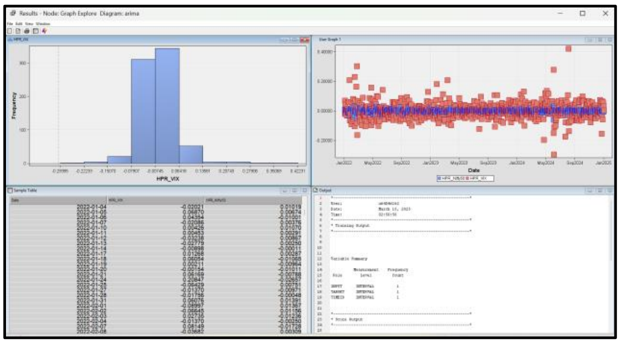
```
proc print data=EMWS1.GARCH_RESULTS (obs=10);
run;
proc forecast data=EMWS1.GARCH_RESULTS lead=756 out=EMWS1.GARCH_FORECAST;
   id Date;
   var predicted_returns conditional_volatility;
run;
```

Training Code

```
proc print data=EMWS1.GARCH_FORECAST (obs=756);
run;
```

VIII. B. ARIMA





Training Code

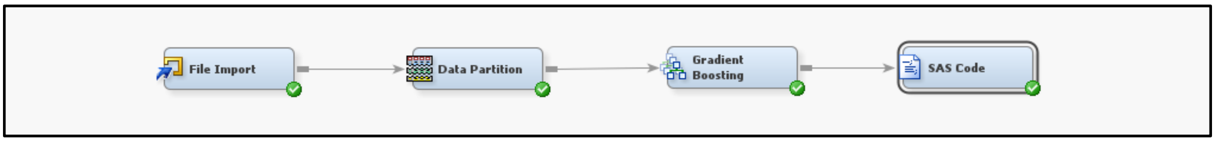
```
LIBNAME EMWS1 "/home/u64046162/arima/Workspaces/EMWS1/";

PROC ARIMA DATA=EMWS1.fimport_data;
   IDENTIFY VAR=HPR_VIX STATIONARITY=(ADF);
RUN;
QUIT;
```

Training Code

```
proc arima data=EMWS1.FIMPORT_DATA;
   identify var=HPR_VIX nlag=24; /* Identify autocorrelation */
   estimate p=1 method=ML; /* ARIMA(1,0,0) model */
   forecast lead=756 out=ARIMA_FORECAST; /* Forecast for 3 years */
run;
quit;
```

European Economic Letters
ISSN 2323-5233
Vol 15, Issue 2 (2025)
http://eelet.org.uk

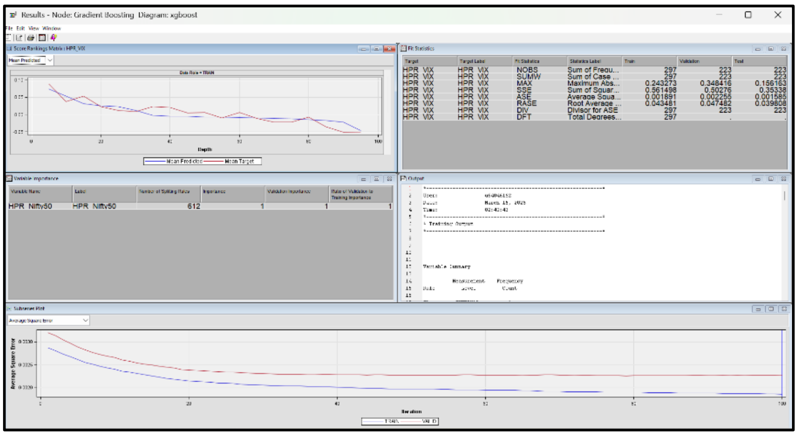
VIII. C. XGBOOST



Property	Value
General	
Node ID	Part
Imported Data	...
Exported Data	...
Notes	...
Train	
Variables	...
Output Type	Data
Partitioning Method	Default
Random Seed	12345
Data Set Allocations	
Training	40.0
Validation	30.0
Test	30.0
Report	
Interval Targets	Yes
Class Targets	Yes
Status	
Create Time	15/3/25 4:43 AM
Run ID	10735165-6ae0-3d40-b
Last Error	
Last Status	Complete
Last Run Time	16/3/25 2:40 AM
Run Duration	0 Hr. 0 Min. 2.57 Sec.
Grid Host	
User-Added Node	No

General

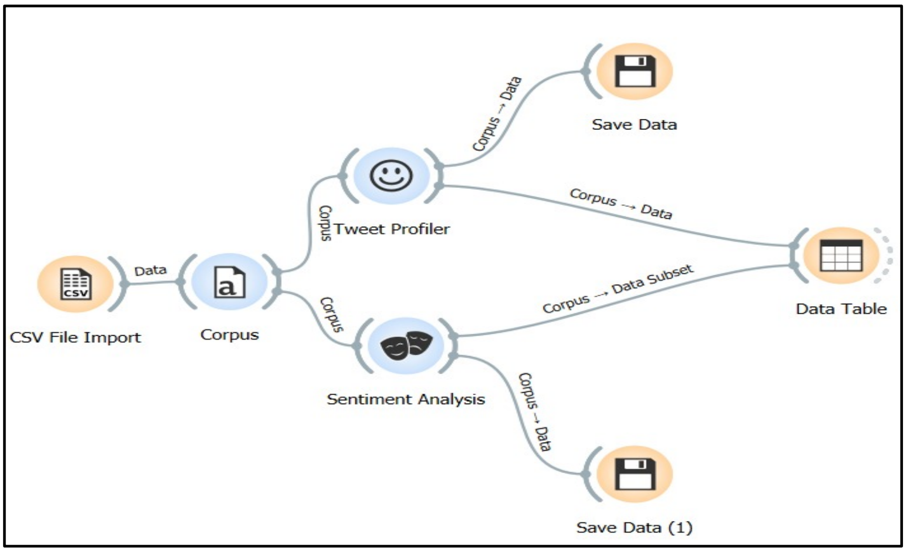
General Properties

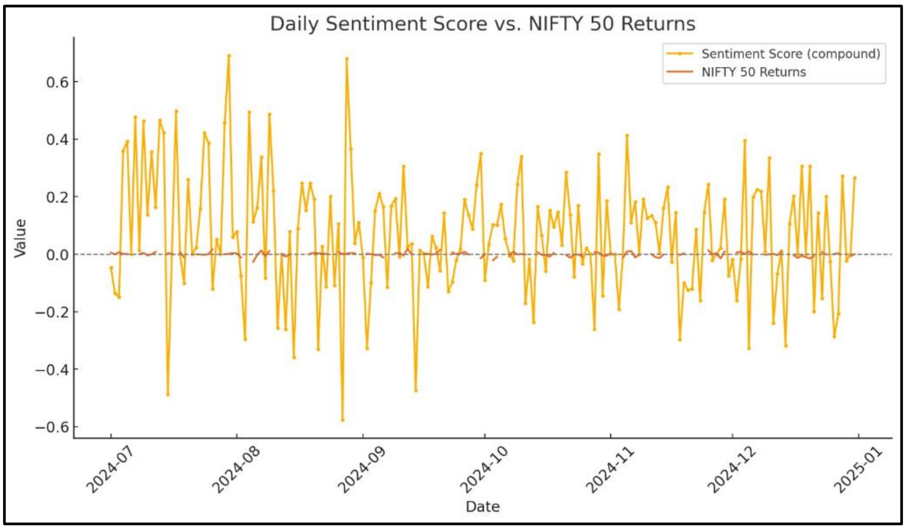


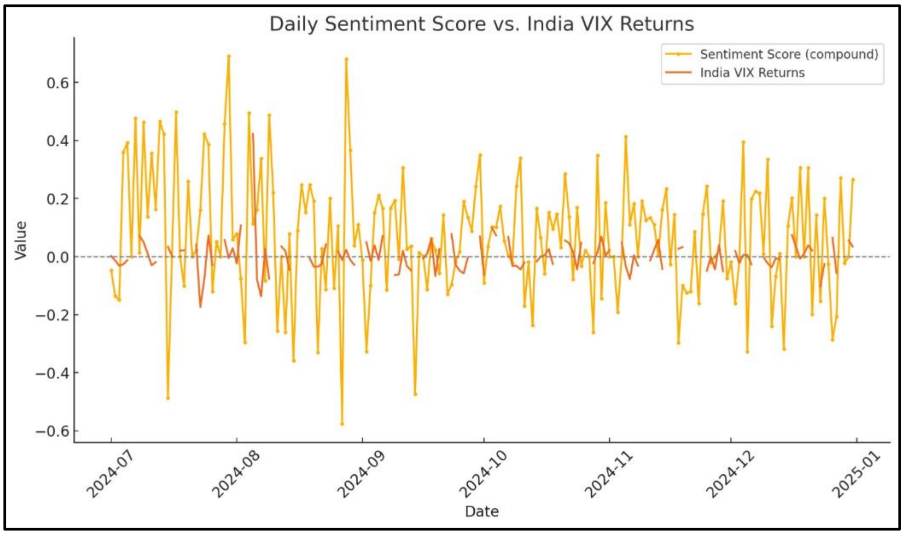
Training Code

```
proc sql outobs=1;
    select _MEANP_ into :last_pred
    from EMWS1.BOOST_EMSCOREDIST
    where _MEANP_ is not missing
    order by _MEANP_ desc;
quit;

%let vol_range = 0.0534; /* 5.34% historical standard deviation */

data WORK.FUTURE_FORECAST;
    format TimeID yymmdd10.;

    P_HPR_VIX = &last_pred.;
    TimeID = '02JAN2025'D;
    output;

    do i = 1 to 755;
        TimeID = TimeID + 1;

        Shock = (ranuni(0) - 0.5) * (2 * &vol_range.);  /* -5.34% to +5.34% range */

        P_HPR_VIX = P_HPR_VIX * (1 + Shock) * 0.995 + &last_pred. * 0.005;

        output;
    end;

    drop i Shock;
run;

proc print data=WORK.FUTURE_FORECAST (obs=756);
run;
```

VIII. D. NATURAL LANGUAGE PROCESSING (NLP)







European Economic Letters
ISSN 2323-5233
Vol 15, Issue 2 (2025)
http://eelet.org.uk