



Deep Learning Sentiment Analysis of News and Social Media in Geopolitical Crises

Neelam Tripathi^{1*} , D. Srinivasa Rao² 

¹ Department of Computer Application, Medicaps University, Indore 453331, India

² Department of CSE, Medicaps University, Indore 453331, India

Corresponding Author Email: neelamtripathi7@gmail.com

Copyright: ©2025 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/isi.300825>

ABSTRACT

Received: 11 June 2025

Revised: 14 August 2025

Accepted: 21 August 2025

Available online: 31 August 2025

Keywords:

sentiment analysis, Operation Sindoor, FinBERT, BERT, geopolitical conflict, market reaction, deep learning, Indian stock market, VADER, hybrid models, social media sentiment, news sentiment aggregation

This research analyses the impact of public sentiment sourced from news articles and social media on the Indian stock market, a period of geopolitical tension. Advanced sentiment analysis techniques were applied, including: (1) a hybrid model combining BERT with TextBlob fallback for robust inference; (2) a fine-tuned BERT model trained on domain-specific tweets and news with VADER-based weak supervision, which achieved an accuracy of 87% with strong precision and recall across sentiment classes; and (3) FinBERT, a transformer pre-trained on financial text. These models were used to analyse sentiment from news headlines and Twitter posts related to geopolitical events. Sentiment scores were aggregated over time and correlated with stock returns from key indices (Nifty 50, Sensex) and sector-specific stocks in defence, energy, IT, and banking. The study explores how sentiment dynamics align with market movements, aiming to reveal predictive and explanatory insights during high-impact geopolitical events, especially in sensitive sectors like defence, banking, and energy.

1. INTRODUCTION

The financial markets are deeply influenced by public sentiment, especially during high-impact geopolitical events. It triggered widespread public discourse across both traditional news outlets and social media platforms. In such volatile environments, investor behaviour is often driven not solely by fundamentals but also by the emotional tone and perception surrounding unfolding events.

This research aims to analyze the impact of public sentiment captured from news articles and Twitter feeds on the Indian stock market during the War. The core objective is to evaluate how sentiment, as measured through various Natural Language Processing (NLP) models, correlates with daily stock returns of key indices (e.g., Nifty 50, Sensex) and sectoral equities (e.g., defense, oil, energy). To conduct a robust sentiment analysis, three different modelling approaches were employed:

Model 1: Hybrid Sentiment Analyzer (BERT + TextBlob)

• This model leverages the bert-base-uncased transformer for initial sentiment classification and incorporates a fallback mechanism using the rule-based TextBlob when the primary model is unavailable due to resource constraints. It is especially useful in real-time or large-scale deployment scenarios.

Model 2: Fine-tuned BERT (bert-base-uncased)

• This model uses supervised fine-tuning on labelled sentiment data derived from VADER to adapt BERT to the domain-specific nuances of geopolitical crises text. It offers improved accuracy in distinguishing between Positive, Negative, and Neutral sentiments for texts related to Operation

Sindoor.

Model 3: FinBERT (Pre-trained on Financial Texts)

• FinBERT, a transformer pre-trained specifically for financial sentiment analysis, is applied directly on the news and tweets. It excels in capturing the tone of financial narratives and helps identify sentiment trends with greater contextual understanding in economic domains.

The sentiment scores from each model were aggregated on a daily basis and then aligned with historical stock return data. Through correlation matrix analysis and regression plots, the study identifies statistically significant relationships between sentiment indicators and market movements. Notably, the defense sector showed a strong correlation with Twitter sentiment, while oil and energy stocks responded more closely to traditional news sentiment.

Moreover, the study revealed that news sentiment polarity had a moderately positive correlation with Nifty 50 and Sensex returns, suggesting that public confidence or fear expressed in news media has a measurable impact on investor behaviour. Sector-wise analysis showed varying degrees of correlation, with industries such as defense, oil, and media being the most sentiment-sensitive during the Operation Sindoor timeline.

This research contributes to the growing field of sentiment-driven financial modelling by:

• Validating multiple NLP approaches for real-time sentiment classification.

• Demonstrating the effectiveness of domain-specific models like FinBERT.

• Establishing empirical links between public sentiment and equity market performance during geopolitical events.

Ultimately, these findings can inform risk assessment

strategies, investor decision-making, and policy response frameworks in periods of national conflict or emergency [1-31].

2. OBJECTIVE OF RESEARCH

The primary objective of this research is to analyze the impact of public sentiment on the Indian stock market during Operation Sindoor, a significant geopolitical event. By leveraging advanced Natural Language Processing (NLP) techniques on news and social media data, the study aims to:

- Evaluate the effectiveness of multiple sentiment analysis models, including a Hybrid BERT-TextBlob model, a fine-tuned BERT classifier, and the domain-specific FinBERT model for classifying sentiment in financial and geopolitical text.

- Investigate the correlation between daily sentiment indicators (derived from news and Twitter data) and stock market performance, including major indices (Nifty 50, Sensex) and sectoral stocks (e.g., defense, energy, media).

- Identify sectors and stocks most sensitive to sentiment changes during the event, providing insight into how market behaviour is influenced by real-time public opinion.

- Compare the strengths and limitations of general-purpose, fine-tuned, and finance-specific sentiment models in the context of high-impact national events.

3. METHODOLOGY

This study employs three sentiment analysis models: Hybrid BERT + TextBlob, fine-tuned BERT, and FinBERT to classify public sentiment from news articles and tweets during Operation Sindoor. The sentiment scores are aggregated daily and correlated with stock market returns across indices and sectors. Statistical techniques, including correlation analysis and regression plots, are used to assess the relationship between sentiment and market movements.

3.1 Data collection

This study adopts a multi-source data collection strategy combining financial market data with sentiment-rich textual data to analyze the potential impact of geopolitical tensions (specifically, Operation Sindoor) on the Indian stock market. The data was gathered from the following sources:

3.1.1 Financial market data

- Source:** Yahoo Finance API via yfinance (Python)
- Period:** 2025-04-01 to 2025-05-11 (Operation Sindoor)
- Indices:** Nifty 50 (^NSEI), Sensex (^BSESN)
- Sectors & Tickers:**
Defence: HAL.NS, BEL.NS, BDL.NS
Oil & Energy: ONGC.NS, IOC.NS, BPCL.NS
Banking: SBIN.NS, HDFCBANK.NS, ICICIBANK.NS
IT: TCS.NS, INFY.NS, WIPRO.NS

3.1.2 News article data

- Real-world dataset:** 1777 articles
- Sources simulated:** To gather relevant news data for analyzing the impact of public sentiment on the Indian stock market during geopolitical tension (April-May 2025), we employed a targeted web scraping approach leveraging

Google News RSS feeds.

- Targeted Keyword Search:** We curated diverse keywords covering political/military conflicts, economic impact, media reactions, and sector-specific stock mentions related to India-Pakistan tensions.

- Time-Frame Restriction:** Only news articles published between April 1, 2025 and May 15, 2025 were collected to focus on the most relevant geopolitical period.

- Automated Data Retrieval:** Using Python's requests and BeautifulSoup libraries, we fetched and parsed Google News RSS feeds for each keyword, extracting article titles, sources, links, and publication dates.

3.1.3 Twitter data

- Real-world dataset:** 575 tweets

- Targeted Keyword-Based Searches:** Tweets were collected using a comprehensive list of keywords related to geopolitical conflict, economic impact, sector-specific stocks, and international reactions between India and Pakistan to cover all relevant aspects. Implemented an asynchronous Python script using the twikit client to efficiently fetch tweets.

- Date Range Filtering:** Only tweets posted between April 1, 2025, and May 31, 2025, were included to focus the analysis on the defined geopolitical event timeframe.

- Robust Twitter Scraping:** Robust data collection was achieved by asynchronously querying Twitter's API with session cookies to handle rate limits and access, incrementally saving unique tweets to a CSV file while avoiding duplicates, and capturing rich metadata such as tweet IDs, user information, timestamps, retweet, and like counts to enable comprehensive and accurate downstream sentiment analysis.

3.2 Data preprocessing and visualization

To extract meaningful insights from raw textual data, a comprehensive preprocessing pipeline was implemented for both the news articles and Twitter data. The primary goal was to clean, normalize, and prepare the data for sentiment analysis and NLP-based modelling.

3.2.1 NLTK setup and text cleaning

The Natural Language Toolkit (NLTK) library was used to support text preprocessing tasks such as tokenization, lemmatization, and stopword removal.

3.2.2 Custom text preprocessing pipeline

A modular Python class, TextPreprocessor, was created to handle the following tasks:

- Lowercasing all text
- Removing URLs, hashtags, and user mentions
- Stripping non-alphabetic characters and numbers
- Stripping non-alphabetic characters and numbers
- Tokenization based on whitespace
- Rule-based lemmatization for common suffix patterns (e.g., *ing → *, ies → y)
- Stopword removal using a curated list of English stopwords

The class was applied to both the news_df and tweets_df DataFrames, creating a new cleaned_text column with preprocessed content ready for further analysis.

3.2.3 Visualization via word clouds

To identify prominent themes and keywords in both sources, word clouds were generated using the WordCloud

library in Python. The cleaned texts from all entries were concatenated, and visualizations were plotted separately for:News articles related to the war conflict, public sentiment on Twitter.

In Figure 1, word cloud visualizes the most frequent terms in news articles discussing war. Dominant themes include “india pakistan,” “stock market,” reflecting media emphasis on geopolitical analysis and market implications.

In Figure 2, the word cloud highlights recurring words from Twitter posts about the war situation. Key terms like “operation sinndoor,” “terror attack,” indicate public concern and perceived market volatility driven by escalating tensions.

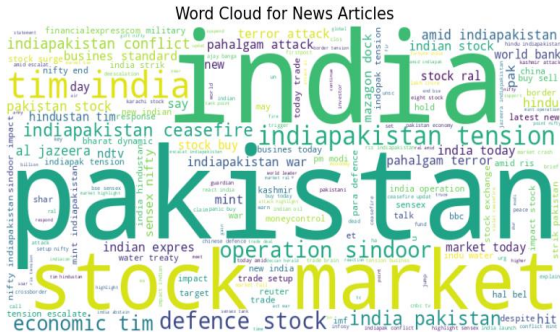


Figure 1. Word cloud of news articles

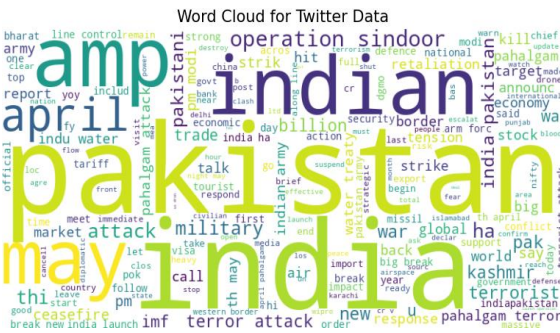


Figure 2. Word cloud of Twitter data

3.3 Sentiment analysis: Model selection, architecture and sentiment visual representation

In this study, we employed three different sentiment analysis models to classify the sentiment of news headlines and tweets related to war relations. These models vary in complexity, architecture, and domain-specificity, providing a robust comparative framework.

3.3.1 Model 1: Hybrid sentiment analyzer (BERT + TextBlob)

- Primary Model: BERT (Bidirectional Encoder Representations from Transformers)
- Model Used: bert-base-uncased
- Pretrained Source: Hugging Face Transformers library
- Justification: BERT is a state-of-the-art pre-trained transformer model well-suited for sentence-level classification tasks. It captures contextual word representations bidirectionally, improving sentiment understanding even in complex or ambiguous sentences.
- Architecture Overview:
 - Input: Tokenized text (up to 512 tokens)
 - Base Layers: 12 transformer blocks (hidden size: 768, 12 attention heads)
 - Classifier Head: A linear layer added on top of the [CLS]

token output

Output: Logits for 3 sentiment classes (positive, neutral, negative)

```
model = BertForSequenceClassification.from_pretrained('bert-base-uncased', num_labels=3)
```

Fallback Model: TextBlob (Rule-based)

- Justification: Used as a lightweight backup when BERT is unavailable (e.g., low-resource environments).
- Approach: Based on NaiveBayesAnalyzer or polarity scoring from lexicon-based sentiment analysis.
- Scoring Logic:
 - Polarity > 0.1 → Positive
 - Polarity < -0.1 → Negative
 - Else → Neutral

Visual representation of Model 1:

In Figure 3, Most news articles in the analyzed dataset were labeled as "negative," with very few classified as "positive" and almost none as "neutral." This indicates a dominant prevalence of negative sentiment in news coverage during the selected period.

In Figure 4, the vast majority of Twitter posts in the dataset exhibit negative sentiment, with almost no neutral tweets and very few positives. This pattern highlights a pronounced negative sentiment trend among social media reactions during selected period.

In Figure 5, Weekly news polarity swings from slightly negative to a brief positive peak in early May, then eases back to a mildly positive level near neutral. Overall, sentiment remains low-amplitude, indicating no sustained strong bias in news tone across the period.

In Figure 6, Twitter sentiment starts positive in early April, dips to slightly negative by late April, then recovers steadily through May to modestly positive.

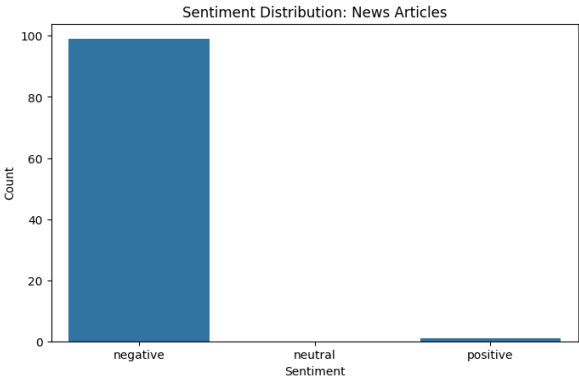


Figure 3. Distribution of sentiment in news articles (April– May 2025)

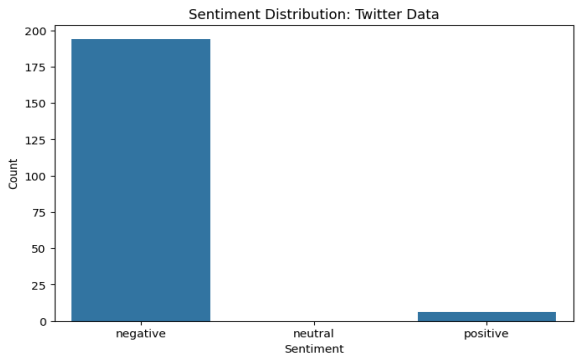


Figure 4. Distribution of sentiment in Twitter data (April– May 2025)

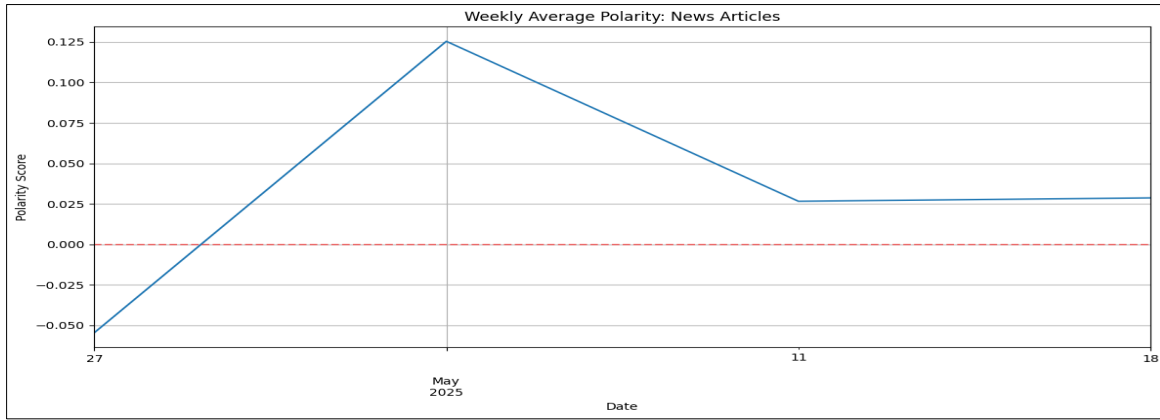


Figure 5. Trend of weekly average sentiment polarity in news (April–May 2025)

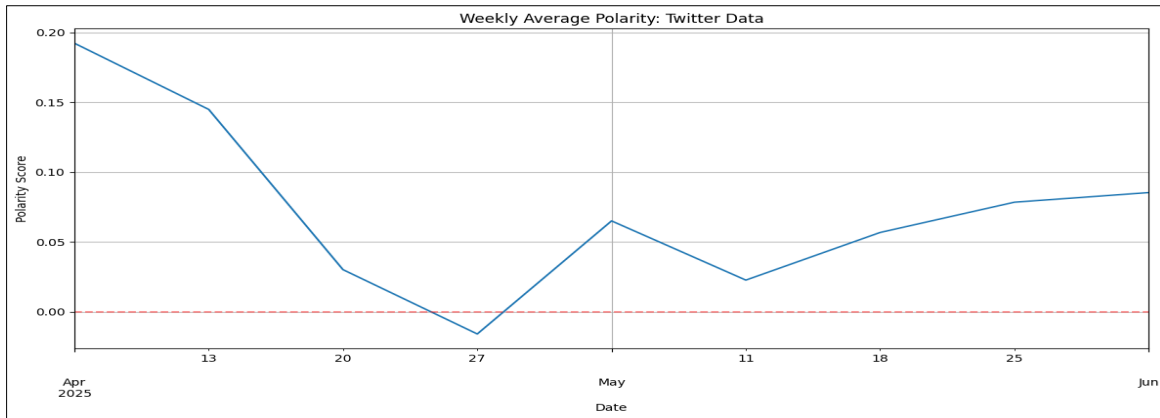


Figure 6. Trend of weekly average sentiment polarity in Twitter (April–May 2025)

3.3.2 Model 2: Fine-tuned BERT using VADER-Labelled supervision

To build a more domain-aligned model, we fine-tuned BERT on a dataset of tweets and news headlines labelled using VADER, a rule-based sentiment analyzer tailored for social media text.

- Base Model:** bert-base-uncased
12-layer Transformer encoder (110M parameters)
Hidden size: 768
12 self-attention heads
Trained on English corpus (BooksCorpus + Wikipedia)
- Fine-tuning Head:** Fine-Tuned BERT Sentiment Classifier with VADER Weak Supervision, Class-Weighted Loss, and Early Stopping

[Raw Text] → [Cleaning/Normalization] → [Tokenizer (WordPiece, max 128)] → [BERT Encoder (bert-base-uncased)] → [CLS embedding] → [Linear layer → 3 logits] → [Softmax → class probabilities] → [Argmax → label]

•**Automatic Labeling:** VADER compound polarity scores were mapped as follows:

- Score ≥ 0.05 → **Positive (2)**
- Score ≤ -0.05 → **Negative (0)**
- Else → **Neutral (1)**

•**Architecture:**

Pretrained BERT model with an added dense classification layer.

Maximum input length: 128 tokens.

Output dimension: 3 (one per sentiment class).

•**Training Configuration:**

Optimizer and scheduler handled by Hugging Face's Trainer.

Epochs: 5

Batch size: 16 (train), 64 (eval)

checkpoint every 500 steps, weight decay: 0.01, early stopping (patience=3) on eval loss.

•**Evaluation:** Performance was evaluated using accuracy, precision, recall, and F1-score, supported by a confusion matrix to analyze misclassifications.

Visual representation of Model 2:

In Figure 7, the confusion matrix shows sentiment classification with three classes: Negative, Neutral, and Positive. It highlights that the model performs best on Negative sentiments, with some misclassifications occurring mainly between Neutral and Positive.

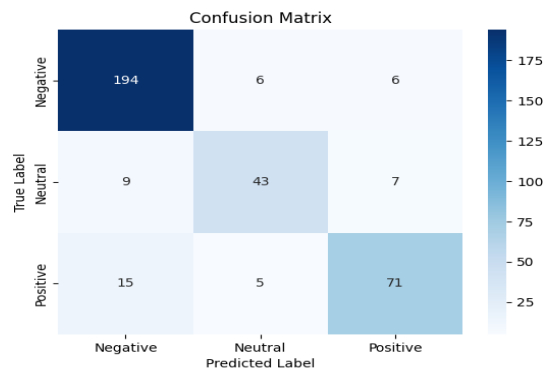


Figure 7. Confusion matrix of fine-tuned BERT using VADER sentiment classifier

In Table 1, the classification report shows an overall accuracy of 87%, with the model performing strongest on Negative sentiment. Neutral and Positive classes are slightly less accurate but still achieve solid precision and recall.

Table 1. Classification report of fine-tuned BERT using VADER-Labeled sentiment analysis model sentiment

Sentiment	Precision	Recall	F1-Score	Support
Negative	0.89	0.94	0.92	206
Neutral	0.80	0.73	0.76	59
Positive	0.85	0.78	0.81	91
Accuracy			0.87	356
Macro Avg	0.84	0.82	0.83	356
Weighted Avg	0.86	0.87	0.86	356

3.3.3 Model 3: FinBERT: Domain-specific sentiment analysis

We selected the FinBERT, a BERT model fine-tuned on financial text for sentiment classification of news articles and tweets concerning war relations. The model (yiyanghkust/finbert-tone) outputs one of three sentiment classes: Negative (0), Neutral (1), and Positive (2).

Architecture:
·Base Model: bert-base-uncased and Fine-tuning Domain: Financial documents, earnings reports, analyst statements.

- It includes 12 transformer layers, 12 attention heads, and a hidden size of 768, followed by a classification head trained for tone detection.
- Classifier Head: A softmax-based classification layer predicting: 0 → Negative, 1 → Neutral, 2 → Positive
- Data Processing and Inference Pipeline*
- Input Texts: Cleaned news headlines and tweets.
- Tokenization: Performed using AutoTokenizer compatible with FinBERT.
- Truncation & Padding: Applied to manage variable-length inputs (max 512 tokens).
- Model Inference:** For each text, FinBERT outputs logits which are passed through a softmax to get probability scores. The class with the highest probability is chosen as the sentiment label.

·**Aggregation:** Daily average sentiment scores were computed for both news and tweets to analyze trends over time.

Visual representation of Model 3:
In Figure 8, FinBERT sentiment fluctuates daily, with news oscillating between neutral and mildly positive while tweets show sharper spikes including occasional positives amid many low values. Overall, social sentiment is more volatile than news, with brief surges that may align with specific event days across April-May 2025.

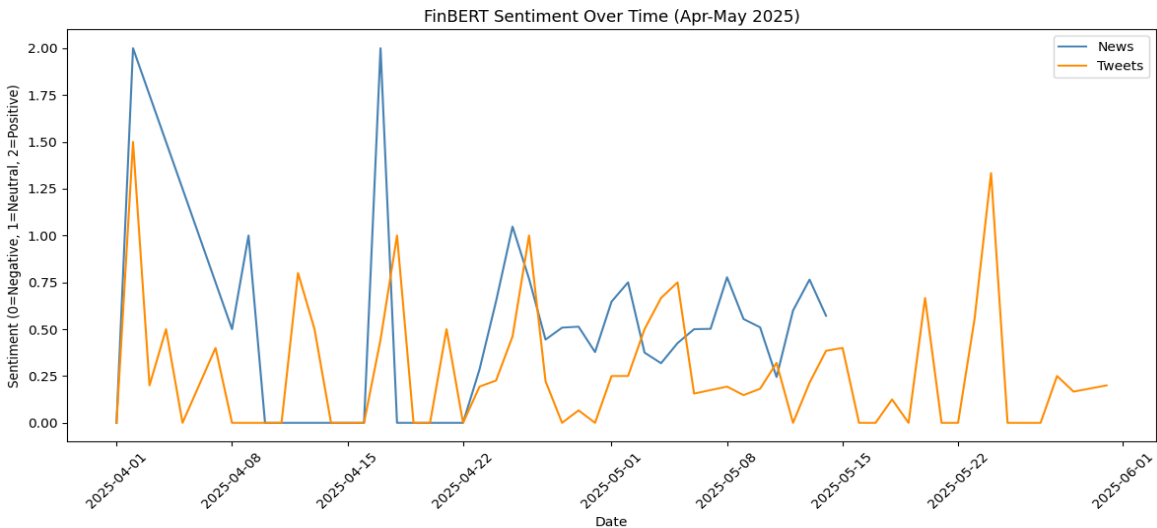


Figure 8. FinBERT sentiment trends in news and tweets (April-May 2025)

3.4 Sentiment-stock market correlation analysis

Objective: This section explores the relationship between sentiment from news and tweets surrounding Operation Sindoor, both before and during the event and the daily returns of selected stocks across key sectors (Defense, Oil, Banking, IT) as well as major indices (Nifty 50 and Sensex). By aggregating sentiment scores and analyzing their correlation with market returns, the objective is to determine whether sentiment acts as a meaningful leading or coincident indicator of stock performance.

3.4.1 Sentiment aggregation

- News and Twitter datasets were processed to capture sentiment signals using a combination of lexicon-based and categorical methods. Each text entry (headline or tweet) was evaluated using:
- We computed:

- Polarity (TextBlob): A continuous sentiment intensity score ranging from −1 (negative) to +1 (positive).
- VADER Compound Score: A normalized sentiment measure (−1 to +1) derived from the VADER sentiment model, which accounts for linguistic nuances such as negations, intensifiers, and emoticons.
- Sentiment Score (Categorical): Calculated as the difference between the proportion of positive and negative labels on a given day:

$$Sentiment\ Score = \frac{\#Positive - \#Negative}{\#Total} \tag{1}$$

For each trading day, sentiment values were aggregated as daily averages of polarity, VADER compound scores, and sentiment score separately for news and tweets. This resulted in six sentiment indicators:

- News: news_polarity, news_vader, news_sentiment
- Twitter: tweets_polarity, tweets_vader, tweets_sentiment

3.4.2 Financial data

- Daily percentage returns were calculated for selected stocks and indices as:

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} \times 100 \quad (2)$$

where,

R_t is the return at time t ,

P_t is the price at time t ,

P_{t-1} is the price at time $t-1$.

- Stocks were selected from sectors potentially influenced by geopolitics (Defense, Oil, Banking, IT) and broad market indices (Nifty 50 and Sensex).

'Defense': ['HAL.NS', 'BEL.NS', 'BDL.NS'],

'Oil': ['ONGC.NS', 'IOC.NS', 'BPCL.NS'],

'Banking': ['HDFCBANK.NS', 'ICICIBANK.NS',

'SBIN.NS'] 'IT': ['TCS.NS', 'INFY.NS', 'WIPRO.NS']

3.4.3 Correlation analysis (static Pearson + heatmaps)

The sentiment and return data were merged by date and a Pearson correlation matrix was computed to assess relationships. Figure 9 represents correlation analysis.

$$\text{Correlation Matrix} = \text{corr}(X) \quad (3)$$

where,

- X represents the combined dataset (e.g., sentiment scores and stock returns),

- $\text{corr}()$ computes the pairwise Pearson correlation coefficients between columns.

A heatmap visualization was used to identify the most significant sentiment–return relationships:

- Tweet sentiment correlates most with returns (notably Nifty and large banks), while news polarity also shows broad positive links.

- Lexicon-based VADER scores are weaker and turn negative for IT stocks, so model-derived polarity is the more reliable signal for those sectors.

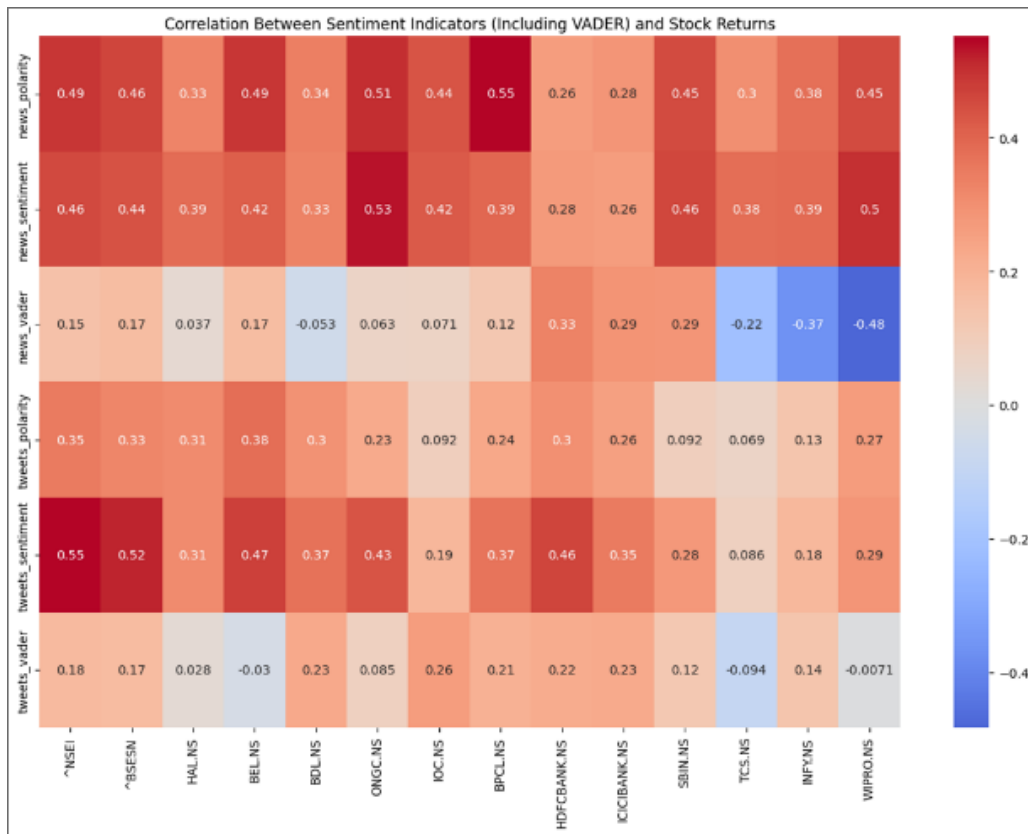


Figure 9. Correlation between sentiment indicators (news and tweets) including VADER and stock returns

3.4.4 Correlation and pairwise relationships

- Method: Daily sentiment indicators (news_polarity, news_vader, tweets_vader) were aligned with same-day log returns for benchmark and sectoral equities; Pearson correlations were reported, and linear fits were visualized with scatter plots and regression lines to examine pairwise associations.

- Illustrative pairs: Figures 10, 11, and 12 present the following analyses:

Nifty 50 Returns vs News VADER Compound shows the benchmark's sensitivity to continuous news tone (news_vader → ^NSEI).

HAL (Defense) Returns vs Twitter VADER Compound highlights social sentiment's relationship with defense sector performance (tweets_vader → HAL.NS).

ONGC (Oil) Returns vs News Polarity examines model-based news polarity against energy sector returns (news_polarity → ONGC.NS).

- Interpretation: Positive slopes and significant Pearson coefficients indicate that more positive sentiment is associated with higher same-day returns; weaker or negative slopes suggest limited or inverse association, guiding which sentiment source/measure is most informative by sector.

3.4.5 Extended correlation and causality analysis

To complement the static Pearson correlation analysis, additional statistical techniques were applied to capture both nonlinear and time-lagged relationships between sentiment indicators and stock returns:

•Granger Causality Tests: Granger causality was used to determine whether sentiment time series (e.g., news_polarity, tweets_polarity) provide predictive information for future stock returns. Tests were conducted with lags up to 5 trading days. This analysis highlights whether changes in sentiment precede movements in stock prices, beyond simple contemporaneous correlations.

•Spearman Rank Correlation: Since financial data often exhibit nonlinear or monotonic relationships, Spearman’s rho

was calculated between sentiment indicators and returns. This non-parametric measure captures rank-based associations that Pearson correlation may miss.

•Rolling Window Correlation: To assess the stability of sentiment–return relationships over time, rolling Pearson correlations were computed using 7-day and 14-day windows. These plots reveal periods when sentiment and stock returns are more strongly aligned (positive or negative) and when correlations weaken or reverse.

Figure 13, 14 showing Rolling Window Correlation plot. Together, these extended analyses provide a more dynamic view of sentiment–market interactions by detecting predictive causality, nonlinear associations, and temporal variation in correlations.

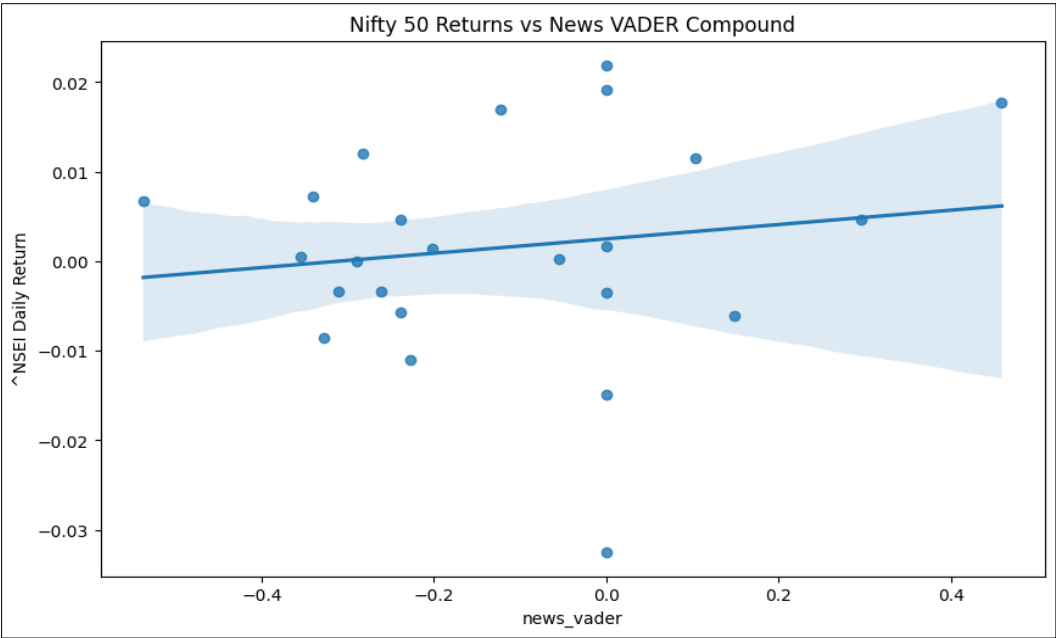


Figure 10. Impact of news VADAR on Nifty 50 returns

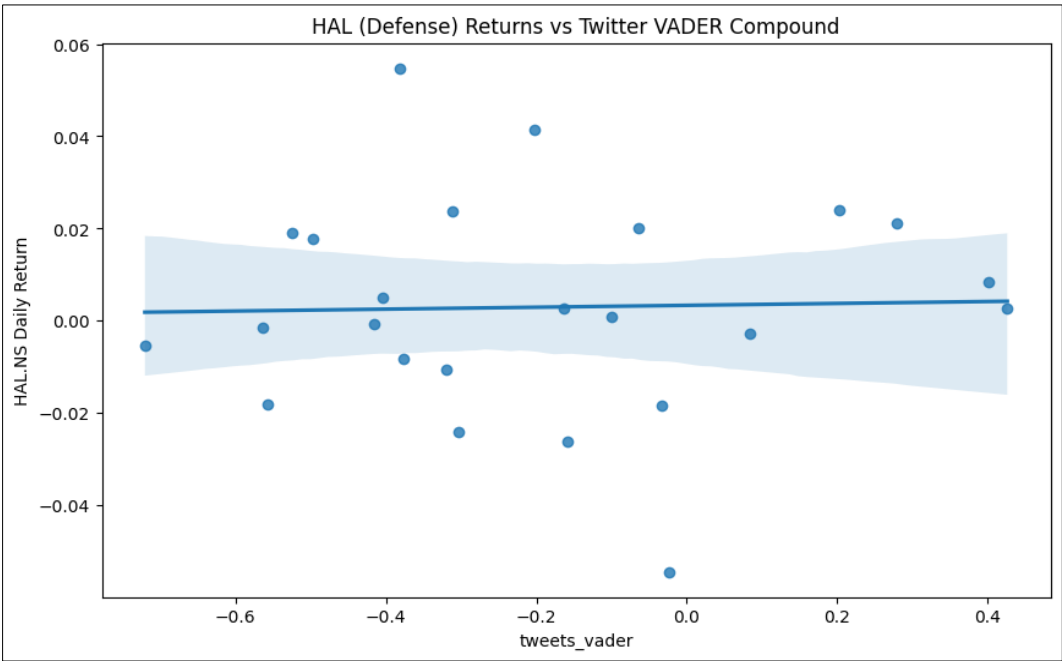


Figure 11. Sentiment vs. performance: HAL’s stock returns and Twitter buzz

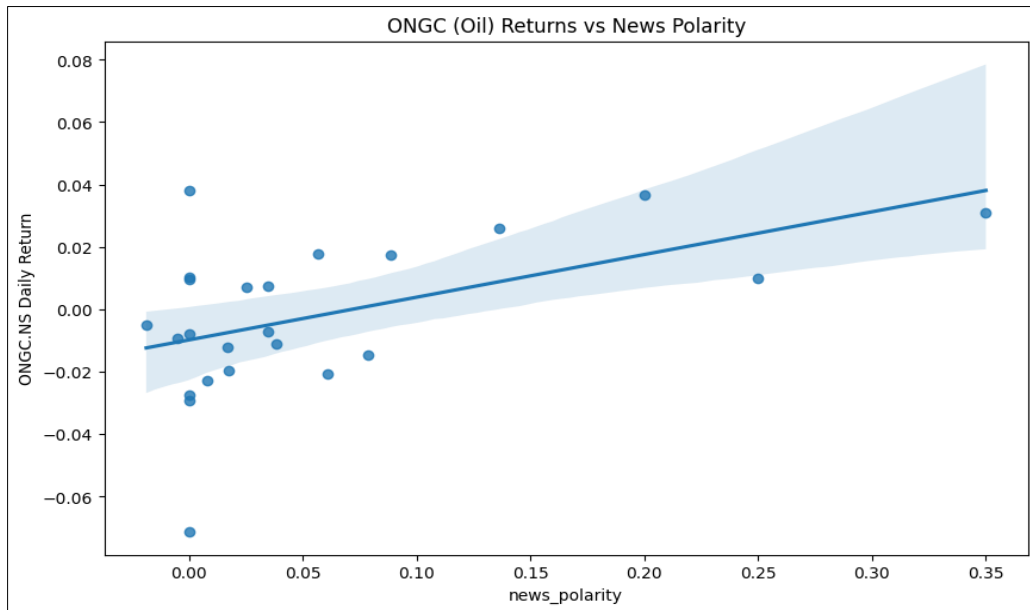


Figure 12. News polarity VS ONGC daily returns

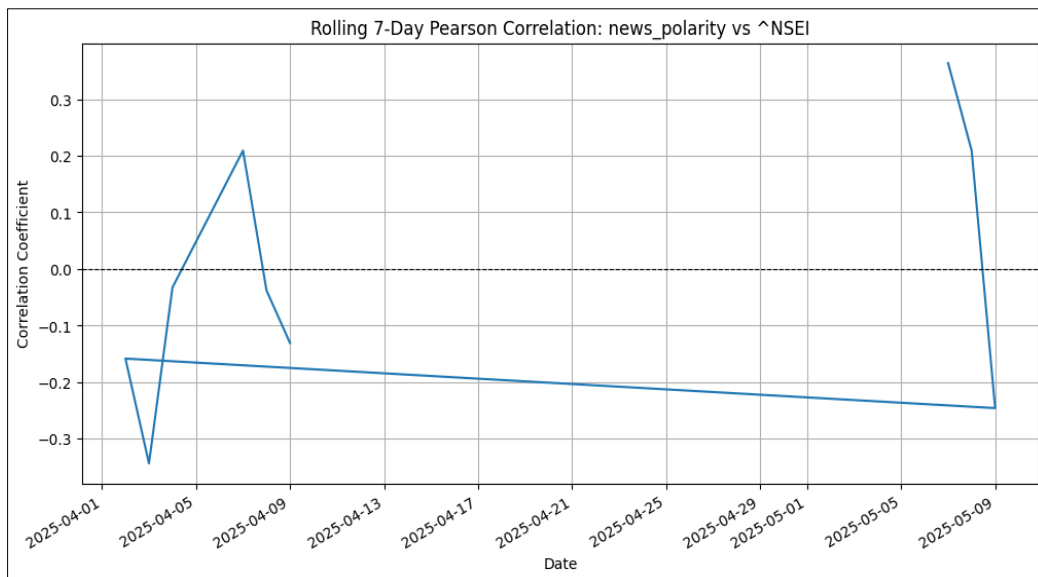


Figure 13. Rolling 7-day correlation between news polarity and Nifty 50 returns

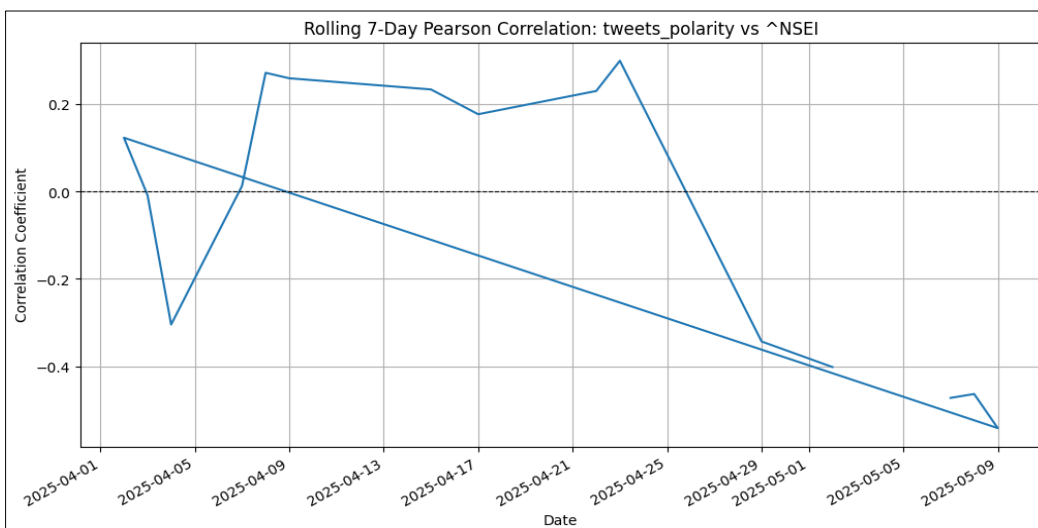


Figure 14. Rolling 7-day correlation between tweets polarity and Nifty 50 returns

3.5 Results and discussion

To interpret the impact of public and media sentiment on the Indian stock market, we computed correlation coefficients between sentiment indicators and daily stock returns across various indices and sectors. These insights were generated using a custom Python function that analyzes the correlation matrix produced from our aligned sentiment and market data.

3.5.1 Correlation analysis key insight

- Tweet-based sentiment shows the strongest and broadest positive correlations with returns across indices and sectors, notably for Nifty ($\wedge$ NSEI ≈ 0.55) and HDFCBANK (≈ 0.46), suggesting social sentiment is a better contemporaneous signal than news.

- News polarity/sentiment also correlates positively with many stocks (e.g., ONGC ≈ 0.51 via polarity), but VADER-based continuous scores underperform and even flip negative for IT names (TCS ≈ -0.22 , INFY ≈ -0.37 , WIPRO ≈ -0.48), indicating lexicon signals may misread tech-domain tone; model-based polarity is preferable for these sectors.

3.5.2 Correlation and pairwise relationships key insight

(1) Impact of News Vader compound on Nifty 50 Returns. The Figure 10 analysis insights are:

- The scatter with regression line indicates a weak positive association between daily news VADER compound and Nifty 50 returns: more positive news tone tends to coincide with slightly higher same-day returns.

- Wide dispersion and broad confidence bands suggest low explanatory power; outliers and limited samples can bias visual perception, so this relation should be corroborated with Pearson/Spearman coefficients and significance tests, not visual inspection alone.

(2) Impact of Twitter sentiments on HAL's Stock Returns. Figure 11 analysis insights are

- The plot shows a near-flat positive regression between Twitter VADER sentiment and HAL daily returns, indicating

only a very weak relationship.

- Wide confidence bands and scattered points imply high uncertainty; Twitter tone alone offers limited explanatory power for HAL's same-day moves.

(3) Impact of News Polarity on ONGC Daily Returns. Figure 12 analysis insights are:

In Figure 12, a scatter plot illustrates the relationship between news sentiment and ONGC (Oil and Natural Gas Corporation) daily stock returns.

- ONGC returns rise with more positive news polarity; the upward regression slope indicates a meaningful positive association.

- Confidence bands narrow around low-to-moderate polarity values, suggesting better fit in the common range; extreme polarity days are few and widen uncertainty.

3.5.3 Extended correlation and causality analysis key insight

(1) Rolling 7-Day Correlation between News Polarity and Nifty 50 Returns showing in Figure 13.

- Short positive bursts appear around early April and mid-May, but the relationship quickly reverts to mildly negative for most of the window, reflecting fragile, event-driven alignment.

- The rapid sign flips and shallow magnitudes imply low persistence; fixed-window correlations should be supplemented with robustness checks (e.g., different windows, lagged tests) before drawing predictive conclusions.

(2) Rolling 7-Day Correlation between Tweet Polarity and Nifty 50 Returns showing in Figure 14.

- The 7-day rolling correlation between tweet polarity and Nifty 50 flips from mildly positive in mid-April (~ 0.25) to distinctly negative by early May (~ -0.5), indicating regime change.

- This instability suggests short, event-driven windows where social sentiment aligns with returns, followed by periods where optimism coincides with drawdowns, so fixed-window correlations should be interpreted cautiously.

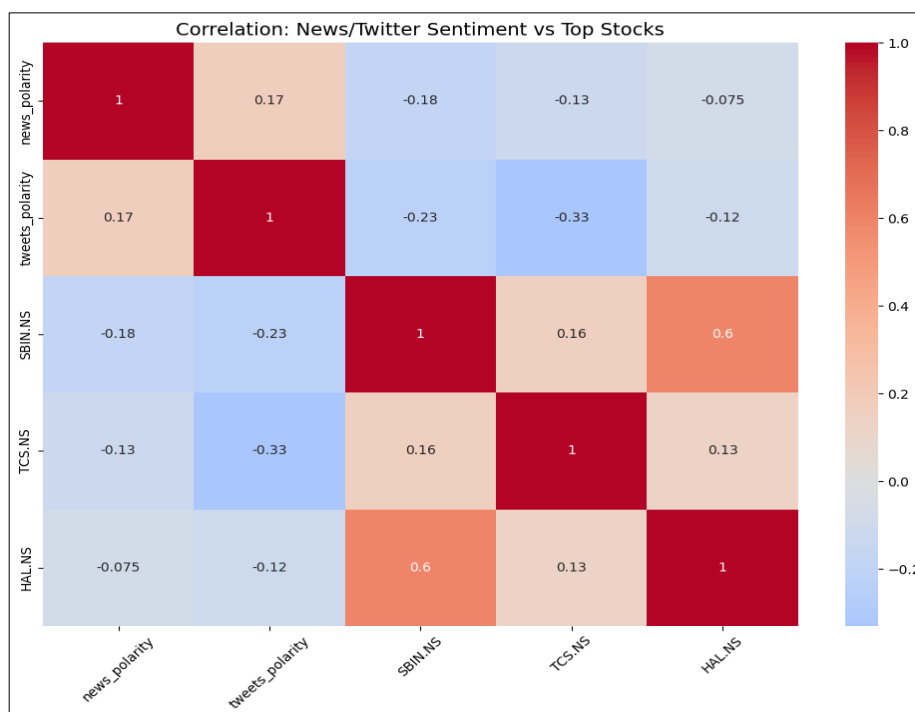


Figure 15. Sentiment VS top stock correlation

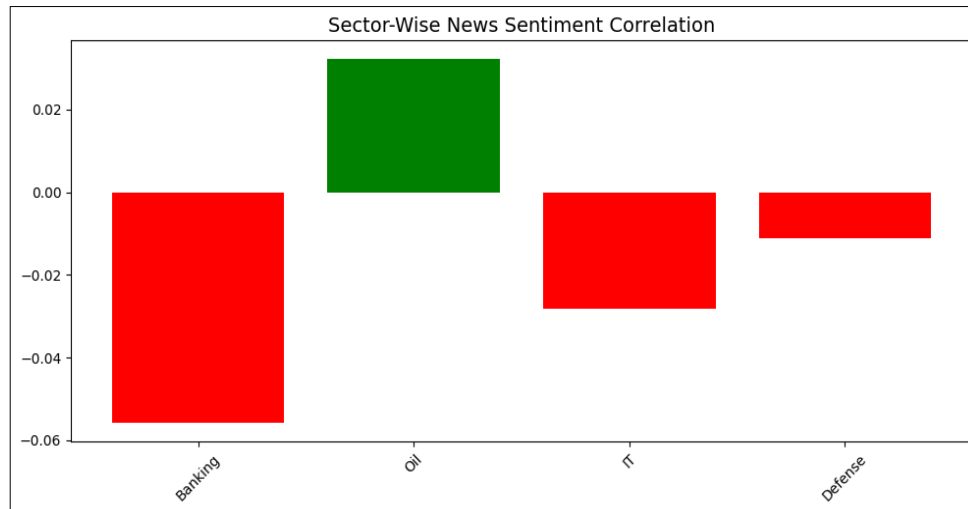


Figure 16. Sector-wise correlation between news sentiment and returns

3.5.4 General market sentiment impact

In Figures 15 and 16, the Correlation Matrix on News and Twitter Sentiment vs Returns of Top Indian Stocks insight.

(1) News Sentiment vs Market Returns

•Index-level impact:

Nifty 50 → correlation 0.03

Sensex → correlation 0.02

•Sector-wise impact:

Banking sector → -0.06 (negative correlation)

Oil sector → +0.03 (positive correlation)

IT sector → -0.03 (negative correlation)

Defense sector → -0.01 (negative correlation)

•Stock-level impact:

Most positively correlated with news sentiment

Most negatively correlated with news sentiment → SBIN.NS, TCS.NS, HAL.NS

(2) Twitter Sentiment vs Market Returns

Average correlation with stocks → -0.02

Compared to news sentiment average → +0.08

returns (e.g., Nifty ≈ 0.55 , HDFC Bank ≈ 0.46) compared to news sentiment, making social media a more sensitive and timely market signal.

News polarity also demonstrated meaningful associations with certain stocks (e.g., ONGC ≈ 0.51), but lexicon-based methods like VADER underperformed, particularly in the IT sector (TCS ≈ -0.22 , Infosys ≈ -0.37 , Wipro ≈ -0.48), highlighting the superiority of transformer-based approaches for domain-specific sentiment. Sector-wise, defense and oil stocks were most responsive to sentiment, while banking stocks showed minimal influence, suggesting they are more tied to macroeconomic fundamentals than public mood.

Rolling correlation analysis further indicates that sentiment–return linkages are short-lived, event-driven, and unstable, with frequent sign flips and weak persistence over time. Overall, the results emphasize that while sentiment, especially from social media, provides valuable short-term signals, its predictive power is fragile and sector-dependent. These insights support integrating sentiment analysis with traditional financial indicators to enhance robustness in stock prediction during geopolitical crises.

4. FUTURE SCOPE

This research can be extended in several meaningful directions:

•Integrating macroeconomic indicators – Combine sentiment with financial fundamentals (e.g., interest rates, global indices) to improve robustness and reduce noise.

•Cross-event validation – Test sentiment–market dynamics across multiple geopolitical events to assess consistency and generalizability.

•Multimodal sentiment analysis – Incorporate images, videos, and multilingual text from social media alongside news for richer sentiment signals.

5. CONCLUSION

This research analyzed the impact of public and media sentiment on the Indian stock market during Operation Sindoor using advanced sentiment analysis models, including a fine-tuned BERT model that achieved 87% accuracy with strong precision and recall. The findings reveal that Twitter sentiment shows stronger and broader correlations with stock

6. LIMITATION

•**Short-lived and unstable correlations** – Sentiment–return relationships were highly event-driven, flipping signs frequently, which limits long-term predictive reliability.

•**Sector-specific variability** – Sentiment impact differed widely across sectors: strong in defense and oil, weak or inconsistent in IT and banking, reducing generalizability.

•**Data and model bias** – Social media sentiment may not represent all investor groups, while transformer models, though effective, still misclassify and require large datasets and computational resources.

REFERENCES

- [1] Koval, R., Andrews, N., Yan, X. (2024). Financial forecasting from textual and tabular time series. In Findings of the Association for Computational Linguistics: EMNLP 2024, Miami, Florida, USA, pp. 8289-8300. <https://doi.org/10.18653/v1/2024.findings->

- emnlp.486
- [2] Wang, L., Shen, J., Shen, C., Ma, Y. (2025). Public sentiment-based stock selection strategy using the BERT model and LightGBM algorithm. *Journal of Computational Methods in Sciences and Engineering*, 14727978251355788. <https://doi.org/10.1177/14727978251355788>
 - [3] Agrawal, M., Mukherjee, A. (2025). Predicting Stock market trends using machine learning and sentiment analysis. In *SoutheastCon*, Concord, NC, USA, pp. 1001-1006, <https://doi.org/10.1109/SoutheastCon56624.2025.10971605>
 - [4] Zakir, U., Daykin, E., Diagne, A., Faile, J. (2025). Advanced deep learning techniques for analyzing earnings call transcripts: Methodologies and applications. *arXiv preprint arXiv: 2503.01886*. <https://doi.org/10.48550/arXiv.2503.01886>
 - [5] Kirtac, K., Germano, G. (2024). Enhanced financial sentiment analysis and trading strategy development using large language models. In *Proceedings of the 14th Workshop on Computational Approaches to Subjectivity, Sentiment, & Social Media Analysis*, Bangkok, Thailand, pp. 1-10. <https://doi.org/10.18653/v1/2024.wassa-1.1>
 - [6] Liu, Y., Wang, J., Long, L., Li, X., Ma, R., Wu, Y., Chen, X. (2025). A multi-level sentiment analysis framework for financial texts. *arXiv preprint arXiv: 2504.02429*. <https://doi.org/10.48550/arXiv.2504.02429>
 - [7] Jiang, T., Zeng, A. (2023). Financial sentiment analysis using FinBERT with application in predicting stock movement. *arXiv preprint arXiv: 2306.02136*. <https://doi.org/10.48550/arXiv.2306.02136>
 - [8] Choe, J., Noh, K., Kim, N., Ahn, S., Jung, W. (2023). Exploring the impact of corpus diversity on financial pretrained language models. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, Singapore, pp. 2101-2112. <https://doi.org/10.18653/v1/2023.findings-emnlp.138>
 - [9] Raju, A., Arun, K., Sudha, S.K., Aji, S. (2024). SMT-PREDICT: An efficient framework for stock market trend prediction using historical and sentimental data. In *2024 3rd International Conference for Advancement in Technology (ICONAT)*, GOA, India, pp. 1-5, <https://doi.org/10.1109/ICONAT61936.2024.10774856>
 - [10] Iacovides, G., Konstantinidis, T., Xu, M., Mandic, D. (2024). Finllama: Llm-based financial sentiment analysis for algorithmic trading. In *Proceedings of the 5th ACM International Conference on AI in Finance*, New York, USA, pp. 134-141. <https://doi.org/10.1145/3677052.3698696>
 - [11] Ukhalkar, P., Zirmite, R., Hingane, S. (2023). Sentiment analysis models for bank nifty index: An overview of predicting stock market sentiment in India. In *2023 7th International Conference on Computing, Communication, Control and Automation (ICCUBEA)*, Pune, India, pp. 1-9, <https://doi.org/10.1109/ICCUBEA58933.2023.10392130>
 - [12] Chiong, R., Fan, Z., Hu, Z., Dhakal, S. (2022). A novel ensemble learning approach for stock market prediction based on sentiment analysis and the sliding window method. *IEEE Transactions on Computational Social Systems*, 10(5): 2613-2623. <https://doi.org/10.1109/TCSS.2022.3182375>
 - [13] Liapis, C.M., Karanikola, A., Kotsiantis, S. (2023). Investigating deep stock market forecasting with sentiment analysis. *Entropy*, 25(2): 219. <https://doi.org/10.3390/e25020219>
 - [14] Chandra, J., Mondal, A.C. (2025). Studies of sentiment analysis for stock market prediction using machine learning: A survey towards new research direction. *Sch J Eng Tech*, 13(1): 56-65. <https://doi.org/10.36347/sjet.2025.v13i01.007>
 - [15] Zhang, B., Yang, H., Liu, X.Y. (2023). Instruct-FinGPT: Financial sentiment analysis by instruction tuning of General-Purpose large language models. *FinLLM at IJCAI 2023*. <http://doi.org/10.2139/ssrn.4489831>
 - [16] Echambadi, V. (2025). Financial market sentiment analysis using LLM and RAG. Available at SSRN: <https://ssrn.com/abstract=5145647>.
 - [17] Fakhurroja, H., Chiqamara, T.M., Hamami, F., Pramesti, D. 2024. Sentiment analysis of local water company customer using naive bayes algorithm. In *2024 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT)*, BALI, Indonesia, pp. 168-173, <https://doi.org/10.1109/IAICT62357.2024.10617472>
 - [18] Khonde, S.R., Virmodkar, S.S., Nemade, S.B., Dudhedia, M.A., Kanawade, B., Gawande, S.H. (2024). Sentiment analysis and stock data prediction using financial news headlines approach. *Revue d'Intelligence Artificielle*, 38(3): 999-1008. <https://doi.org/10.18280/ria.380325>
 - [19] Biswas, S., Ghosh, S., Roy, S., Bose, R., Soni, S. (2023). A study of stock market prediction through sentiment analysis. *Mapana Journal of Sciences*, 22(1), 89-120. <https://doi.org/10.12723/mjs.64.6>
 - [20] Maqbool, J., Aggarwal, P., Kaur, R., Mittal, A., Ganaie, I.A. (2023). Stock prediction by integrating sentiment scores of financial news and MLP-regressor: A machine learning approach. *Procedia Computer Science*, 218: 1067-1078. <https://doi.org/10.1016/j.procs.2023.01.086>.
 - [21] Parvatha, L.S., Tarun, D.N.V., Yeswanth, M., Kiran, J.S. (2023). Stock market prediction using sentiment analysis and incremental clustering approaches. In *2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Coimbatore, India, pp. 888-893. <https://doi.org/10.1109/ICACCS57279.2023.10112768>
 - [22] Ajmain, M.R., Khatun, M.F., Bandan, S.S., Rejuan, A.R., Ria, N.J., Noori, S.R.H. (2022). Enhancing sentiment analysis using machine learning predictive models to analyze social media reviews on junk food. In *2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, Kharagpur, India, pp. 1-7. <https://doi.org/10.1109/ICCCNT54827.2022.9984355>
 - [23] Sawale, G.J., Rawat, M.K. (2022). Stock market prediction using sentiment analysis and machine learning approach. In *2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT)*, Tirunelveli, India, pp. 1-6. <https://doi.org/10.1109/ICSSIT53264.2022.9716326>
 - [24] Mehta, P., Pandya, S., Kotecha, K. (2021). Harvesting social media sentiment analysis to enhance stock market prediction using deep learning. *PeerJ Computer Science*, 7: e476. <https://doi.org/10.7717/peerj-cs.476>
 - [25] Chen, X., Xie, H., Li, Z., Zhang, H., Tao, X., Wang, F.

- L. (2025). Sentiment analysis for stock market research: A bibliometric study. *Natural Language Processing Journal*, 10: 100125. <https://doi.org/10.1016/j.nlp.2025.100125>
- [26] Bharathi, S., Geetha, A. (2017). Sentiment analysis for effective stock market prediction. *International Journal of Intelligent Engineering and Systems*, 10(3): 146-154. <https://doi.org/10.22266/ijies2017.0630.16>
- [27] Bhardwaj, A., Narayan, Y., Dutta, M. (2015). Sentiment analysis for Indian stock market prediction using Sensex and nifty. *Procedia Computer Science*, 70: 85-91. <https://doi.org/10.1016/j.procs.2015.10.043>
- [28] Fernandes, D.S., Fernandes, M.G., Borges, G.A., Soares, F.A. (2019). Decision-making simulator for buying and selling stock market shares based on twitter indicators and technical analysis. In 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC), Bari, Italy, pp. 2626-2632. <https://doi.org/10.1109/SMC.2019.8913879>
- [29] Fernandes, D.S., Fernandes, M.G., Borges, G.A., Soares, F.A. (2019). Decision-making simulator for buying and selling stock market shares based on twitter indicators and technical analysis. In 2016 International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES), Paralakhemundi, India, pp. 1345-1350. <https://doi.org/10.1109/SCOPES.2016.7955659>
- [30] Bing, L., Chan, K.C., Ou, C. (2014). Public sentiment analysis in Twitter data for prediction of a company's stock price movements. In 2014 IEEE 11th International Conference on e-Business Engineering, Guangzhou, China, pp. 232-239. <https://doi.org/10.1109/ICEBE.2014.47>
- [31] Mohan, S., Mullapudi, S., Sammeta, S., Vijayvergia, P., Anastasiu, D.C. (2019). Stock price prediction using news sentiment analysis. In 2019 IEEE Fifth International Conference on Big Data Computing Service and Applications (BigDataService), Newark, CA, USA, pp. 205-208. <https://doi.org/10.1109/BigDataService.2019.00035>