European Economic Letters
ISSN 2323-5233
Vol 15, Issue 1 (2025)
http://eelet.org.uk

Predicting Indian Stock Prices by Using Sentiment Analysis and Natural Language Processing

Sudha Swaroop[1],
[1]G L Bajaj Institute of Technology and Management, Uttar Pradesh
Himadri Srivastava[2],
[2]BFI Linz, Austria
Richa Sinha[3]
[3]ABES Business School, Uttar Pradesh
Shishma Kushwaha[4]
[4]Noida Institute of Engineering & Technology, Uttar Pradesh

ABSTRACT
This study employs traditional empirical models to investigate the impact of sentiments on financial market volatility through the use of financial indicators. In this paper, we apply recent methods for text-based sentiment analysis of the market from relevant news articles about the financial and business markets. Two separate market sentiments—positive and negative sentiments—are created from various emotions that can be identified using standard Natural Language Processing (NLP) techniques. The research also makes an effort to apply the previously mentioned market sentiments to a refined iteration of the asymmetric GARCH model of conditional volatility for the Indian stock market (Sensex) for the period from January 1, 2011, to December 31, 2022. Empirical findings indicate that negative market sentiment predominates over positive ones and that noise trading in financially undeveloped Indian stock market.

Keywords: *Risk modelling; Stock market; Sentiment analysis; GARCH; SENSEX; NLP*

1. Introduction

Sentiment analysis can be used by traders and investors to assess market sentiment and make wiser investment choices. The study of psychological influences on investor's behaviour and decision-making in the financial markets is known as behavioural finance. Positive sentiment may represent a chance to buy, but negative sentiment may indicate the need to sell or stay away from a specific asset(Hassan et al., 2022). It is difficult and complex to predict stock prices using time series analysis and machine learning approach (Sinyak et al., 2021). Even though no model can produce stock prices with 100 percent accurate results, machine learning approach can be used to provide reliable predictions based on past data (Aggarwal, 2022). Data frequency must also be adjusted. Because so many variables can affect stock prices, it is impossible to predict the future of the stock market. As a result, even the most precise models might not always produce reliable forecasts (Jain et al., 2021). Obtain historical stock price information, taking note of the open, close, high, and low prices as well as trade volume and any other pertinent financial fcators. Data can be retrieved from financial sources, APIs, and websites (Singh, 2021). Financial institutions and portfolio management advisors can recognize and manage risk with the assistance of sentiment analysis. They can find possible market volatility or excessive assurance that can result in asset bubbles by keeping an eye on sentiment. To make the data suitable for analysis, this requires addressing disappeared values, removing outliers, and standardizing the data(Lachana & Schroeder, 2021). It seeks to control, based on news stories, social media posts, financial intelligences, and other information sources, if the general mood towards a detailed financial instrument, business, or market is positive, neutral or negative (S. Kumar et al., 2022). When using machine learning approach to make trading decisions, it is crucial to expand the investment plans and take risk management strategies into account.

To analyse and evaluate the sentiment tone of textual data relating to financial markets, assets, and economic events, sentiment analysis in finance uses natural language processing (NLP) and machine learning approach (Gonzalez-Igual et al., 2021). To systematize trading decisions, several algorithmic trading systems use sentiment analysis (Y. Chen et al., 2020). Stock trading can be done more promptly and successfully because to algorithms' quick responses to news and social media mood. Market trends and movements can be forecast using sentiment analysis. Corporations frequently publish news and earnings reports that can impact the value of their stocks. Investors and analysts can more appropriately analyse the impact of these incidences with the aid of sentiment research. Financial administrations can take anticipatory action by using sentiment analysis to identify early warning clues of financial crises, fraud, or market manipulation (M. Chen et al., 2022). Losses make people more sensitive than similar benefits. In contrast to the delight of making the same amount of money, they practice the pain of losing it more keenly.

Financial analysts can invent trends that may affect future market behavior by analyzing vast amounts of news and social media data (X. Li et al., 2020). By inspecting how sentiment might affect financial choices and market outcomes, sentiment analysis advances our knowledge of behavioural economics (Reis & Pinho, 2021). It's crucial to prove that sentiment analysis has its difficulties. Some of the presumptions of conventional financial theories, such as the Efficient Market Hypothesis (EMH), are questioned (Yadav & Chakraborty, 2022). Language shades, sarcasm, and context can make it challenging to reliably regulate sentiment from textual data. Sentiment-based exchange-traded funds (ETFs) and indices are two examples of financial instruments that are made to follow sentiment-based changes in the market. These products give investors access to sentiment trends. According to EMH, stock markets are effective because asset prices promptly and properly represent all available information. Investors are considered to be balanced in an efficient market, making choices entirely on the basis of information at their disposal and fundamental analysis, with little or no influence from emotions(Shen et al., 2021).Here are two well-known behavioural finance theories that address the feelings of investors. Vision Theory, which was created by Daniel Kahneman and Amos Tversky in 1979, is a cornerstone of behavioural finance.

Sentiment analysis is essential to modern finance because it agreements insightful information about market sentiment, supports informed investment choices, and aids in risk management and market predicting. Furthermore, the difficulties of the algorithms utilised and the data sources used might have an impact on how reliable sentiment analysis models(Xiong et al., 2020). Sentiment analysis is projected to play a bigger role in the financial sector as technology and NLP techniques develop (Naufal Adi Nugroho & Erwin Budi Setiawan, 2021). According to this argument, people don't make decisions based on the expected value of events as suggested by conservative financial theories, but rather on the gains and losses that could be predicted relative to a benchmark(Brazdil et al., 2022). The disposition effect (a propensity to sell winning investments too soon and hang on to losing investments for too long) and the donation effect are just two examples of behavioural biases that are explained by prospect theory. Advantages and losses affect utility, which follows an S-curve rather than being linear. The Efficient Market Hypothesis (EMH), while not specifically a behavioural finance theory, is a key idea in conventional finance (Staszkiewicz & Staszkiewicz, 2022).Small gains and losses have a uniformly smaller effect, but as they grow larger, their psychological impact increases(Byun et al., 2022). Markets can overstress to news and events, leading to excessive price changes, according to behavioural finance research. They might also underreact to information, which would delay price adjustments. Understanding how investor emotions and psychosomatic biases can cause deviations from logical decision-making in financial markets is fundamentally reliant on the Prospect Theory and influences against the Efficient Market Hypothesis(Paule-Vianez et al., 2020). They discovered that under-diversification, unnecessary trading, and herding behaviour are just a few examples of the actions that individual investors frequently do that are at odds with logical decision-making. Behavioural finance offers insightful elucidations for why investors occasionally make poor decisions and how market inefficiencies might endure. They discovered that due to behavioural biases associated with loss aversion and inertia, people regularly make poor decisions in retirement savings programmes, such as not contributing enough(Lachana & Schröder, 2022). In addition to standard asset pricing models like the Capital Asset Pricing Model (CAPM), researchers have created behavioural asset pricing models that incorporate investor sentiment and anomalies like the value premium and momentum effect(Cao & Zhai, 2023).

Numerous studies, approaches, and discoveries have been made in the rich and developing field of behavioural finance that focuses on the psychological influences on financial decision-making(Gupta et al., 2020). The concepts of loss aversion, the S-curve of utility, and framing effects are among the important discoveries. They exposed that stocks that underperformed over the preceding three to five years tended to outperform during the next period, indicating that markets tend to overreact to historical data(A. Kumar & Jaiswal, 2020). A number of studies were done by Barber and Odean to observe the actions of individual investors. Due to these biases, investment performance is less than ideal (Benartzi and Thaler, 1995) ran tests to see how Prospect Theory affected people's decisions regarding their retirement funds. Researchers and financial organisations use sentiment analysis to examine enormous amounts of textual data from sources like news stories and social media to evaluate market sentiment and forecast market arrangements. In order to understand how the brain makes decisions; the interdisciplinary area of neuroeconomics incorporates ideas from economics, psychology, and neuroscience.

Studies on neuroimaging have shed light on how emotions and reasoning biases present themselves in the brain during financial decision-making(Shah et al., 2019). To study phenomena like risk aversion, framing effects, and pride, researchers employ experimental methodologies.(Y. Li & Pan, 2022) Results from experiments support and improve behavioural theories. Experiments are frequently used in behavioural finance to explore decision-making under predetermined conditions. Big data and developments in natural language processing (NLP) have made this possible (Ferreira et al., 2021). To examine real-world financial behaviour and achievement of information on investor sentiment, portfolio decisions, and financial decision-making in various circumstances, researchers perform field studies and surveys(Hossain et al., 2022). In general, behavioural finance research has grown and changed throughout time, providing important new identifications into the various ways that psychological elements, emotions, and intellectual biases affect financial decision-making. The disparity between conservative financial theories and the realities of human behaviour in financial markets has been closed in part by these studies and techniques(Ferreira et al., 2021).

2. Literature Review

A statistical method called time series analysis is used to examine and project data points gathered throughout time. When modelling and predicting the volatility of financial and economic time series data, the Autoregressive Conditional Heteroscedasticity (ARCH) and Generalised Autoregressive Conditional Heteroscedasticity (GARCH) models are frequently utilised(CHAREF & AYACHI, 2016). When modelling and predicting the variance or volatility in financial returns, such as stock prices or exchange rates, these models are especially helpful. In order to analyse and comprehend the attitudes and emotions conveyed in news stories, social media posts, and other types of textual data connected to financial markets, sentiment analysis models for the stock market have been developedClick here to enter text.Traders, investors, and financial institutions can use these models to assess market sentiment and come to wise conclusions (Gao, 2021). Stock market sentiment analysis methods are susceptible to the limits of natural language interpretation and are not infallible (J. Chen et al., 2020). Various elements, including as rumors and deceptive practices, might affect the mood of the market, which can be difficult for models to correctly interpret. As a result, in the financial sector, integrating sentiment analysis with other types of analysis and risk management is frequently a prudent course of action (Lohrbach & Schumann, 1992).

A well-liked and uncomplicated machine learning model for text categorization applications is naive bayes. Based on the Bayes theorem, it makes the "naive" assumption that, given the class label, the features—in this example, the words—are conditionally independent (Tiwari et al., 2017). Naive Bayes, despite its simplicity, can perform remarkably well for a variety of text classification problems, especially when data and computational resources are scarce. For applications like sentiment analysis, text categorization, and spam email detection, naive Bayes is frequently utilised. It's especially helpful when there aren't many training examples available, and you need a model that works well right away (Shastri, 2023).

SVMs have been effectively used in a variety of fields, including bioinformatics, spam detection, text categorization, and image classification. The problem, the dataset, and the intended trade-offs between model complexity and performance determine the machine learning technique to be used. SVMs are useful for binary classification problems, but there are other machine learning algorithms that may perform better in some situations, including gradient boosting, random forests, and deep learning (Vishwakarma et al., 2020).

RNNs have drawbacks, including the inability to capture very long-range relationships and the propensity for gradients to evaporate or explode during training. Modern state-of-the-art models for many NLP applications include more complex models like Transformers, particularly pre-trained models like BERT and GPT, which are better at capturing contextual information(C. C. Chen et al., 2019). RNNs are still a useful tool for some sequential data processing jobs, particularly when dealing with shorter sequences like tweets(Agarwal et al., 2022). CNNs can be helpful for jobs that largely entail identifying and categorising local patterns within text, even though they are typically not the first choice for text classification. This is especially true when working with fixed-length sequences or limited computational resources (Wang et al., 2016). Since there are so many unforeseen factors that might affect financial markets, it is difficult to accurately estimate stock values. Instead, then trying to generate exact forecasts, machine learning models work best for providing insights and spotting trends (Rayegan et al., 2022). When making investing decisions based on predictive models, it's also critical to take risk management measures into account and seek the advice of financial professionals.

Convolutional Neural Networks (CNNs) are generally used in computer vision tasks because they are particularly good at identifying regional patterns and feature hierarchies (B. Ma et al., 2023). Even if CNNs aren't the most popular option for text categorization, they can still be used efficiently when you wish to analyse text with a focus on local patterns, such finding n-grams or brief phrases (Abbahaddou & Chiadmi, 2022). Because it is impossible to foresee the future of the stock market with any degree of accuracy, models should be used to help make decisions rather than as absolute truths. Due to the numerous factors impacting market movements, predicting stock market values using machine learning and time series analysis is a difficult and complex undertaking. Nevertheless, it is feasible to create prediction models that can offer perceptions and projections(Rajeswari et al., 2022). Natural language processing (NLP) and machine learning techniques are used in investor sentiment analysis on social media to analyse and quantify the sentiment or feelings expressed by investors and traders on social media platforms, forums, and financial news websites (Q. Ma, 2020).

Knowing how investors feel about a market might help investors make more informed investing decisions. Social media investor sentiment analysis can offer insightful information and serve as an additional tool for investing decisions(Pandji et al., 2019). But it's important to recognise that investor attitude is only one of many variables affecting the financial markets. Investment strategies that are more informed can be developed by combining sentiment analysis with fundamental analysis and technical analysis(Trierweiler Ribeiro et al., 2021). For a variety of time series prediction and classification problems, a hybrid prediction model that combines Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN) components can be effective (Baliyan et al., 2021).

Utilising the benefits of both CNN, which is excellent at collecting local patterns and LSTM, which is excellent at capturing temporal dependencies, is the rationale behind such a hybrid model (Adebiyi et al., 2014). It can be used for many different tasks, such as time series forecasting and classification jobs where local patterns and temporal dependencies are both important. To get the best results, model design, hyperparameters, and feature experimentation and optimization are crucial. In financial research, technical indicators are frequently employed to forecast stock prices and market trends (Alroobaea, 2022; Khan et al., 2022). These indicators offer perceptions into market behaviour and prospective price fluctuations because they are based on historical price and volume data. It's vital to remember that no one indicator can guarantee precise forecasts, and their efficacy can change depending on the state of the market and the duration of the forecast(Gao et al., 2022).

To decide on their trading methods more intelligently, traders frequently combine various technical indicators. Technical analysis is only one component of financial analysis; for a more complete understanding of the market, it is frequently supplemented with fundamental analysis and market sentiment analysis(Mahadevaswamy & Swathi, 2022). The Moving Average Convergence Divergence (MACD) technical indicator is a popular and flexible tool in technical analysis and stock trading(Wiranata & Djunaidy, 2021). It is used to recognise changes in the strength, direction, momentum, and duration of a trend in the price of a stock or other financial asset(J. Chen et al., 2020). It was created by Gerald Appel in the late 1970s. While more modern methods like transformers (e.g., BERT, GPT) have emerged as the state-of-the-art for natural language comprehension tasks, including sentiment analysis, CNNs can still be useful(Basiri et al., 2021). Transformers have revolutionised the field of NLP and are renowned for their efficiency in capturing contextual information. It may be good to investigate both RNN and transformer-based models to see which performs better for sentiment analysis, depending on your particular goal and dataset(Chaudhuri, 2022).

Long Short-Term Memory (LSTM) neural networks have gained popularity as a method for stock price prediction in the world of algorithmic trading and financial analysis. Recurrent neural networks (RNNs) of the LSTM type are well suited for time series prediction because they can capture sequential dependencies(Baliyan et al., 2021). Although predicting stock prices is intrinsically difficult due to the myriad factors driving market movements, LSTM-based models can offer insights and forecasts. When utilising predictive models to make investing decisions, it is crucial to limit risk, diversify, and seek advice from financial professionals (Alroobaea, 2022).

For a more complete understanding of the stock market, it is essential to combine technical analysis with fundamental analysis and market sentiment analysis. It's critical to compare the effectiveness of various sentiment classifier models using the same dataset and standardised evaluation metrics (Khan et al., 2022). Remember to consider the unique goals of your stock market sentiment study as well as the potential effects of imbalanced data, overfitting, and other factors. given addition, given the dynamic financial markets, continual monitoring and revaluation of model performance over time are essential (Lohrbach & Schumann, 1992). It is essential to assess and contrast the performance of various sentiment classifier models for the stock market to ascertain how well they anticipate market trends or stock prices. Depending on the dataset, model configurations, and requirements of the sentiment classification task, the assessment metrics chosen and the performance itself may change (Trierweiler Ribeiro et al., 2021).

3. Data

The analysis is based on the closing price of India's main market index, the Sensex, which is the best proxy for assessing the performance of the country's financial system. For empirical analysis, daily data from January 1, 2011, to December 31, 2022, has been collected. Sensex return is computed using the following standard formula:

Return = In (Pt /P t-1) x 100 …………..(1)

where Pt represents the Sensex index on day t and Pt-1 denotes the Sensex value from the day before.

In this empirical study, market sentiment is the other relevant variable. The task of creating an appropriate market index for intangible feelings and emotions is difficult from an empirical standpoint. Using market fundamentals, some

European Economic Letters
ISSN 2323-5233
Vol 15, Issue 1 (2025)
http://eelet.org.uk

studies have created sentiment indices. Proxies for investor sentiment in the market are constructed using macroeconomic variables such as interest rates, daily exchange rates, stock trading volume, etc. as in (Mahadevaswamy & Swathi, 2022) and (C. C. Chen et al., 2019) . On the other hand, empirical finance has recently become more interested in a few alternative approaches to measuring sentiment. These are behavioural indices that have been developed using dynamic data from news sources and other web portals. Market sentiment has been measured using a new machine learning mechanism-based method. Natural Language Processing (NLP) methods and the NRC word-emotion association lexicon are used in our study to create sentiment and emotional indices (Lohrbach & Schumann, 1992). For sentiment analysis and emotion categorization, the most well-known websites covering business and economic policy in India have been taken into account. The Hindu Business Line, Moneycontrol.com, Livemint Business News, and Reuters India's Business and Economic section are a few of them. The steps that follow are used to construct the sentiment indices from various emotions.

In the beginning, we collected headline text and summaries from pertinent sections of all the previously mentioned sources. Additionally, we filtered the data using "Natural Language Toolkit" in order to classify the emotions in the text. Lower case text for all texts to facilitate the process.In order to process the words in the string further, tokenization is used to split them up. Using the "Natural Language Toolkit," stop words are eliminated from the text. This is done to get rid of words that are frequently used in English but don't contribute any meaningful insights. In the following step, lemmatization is performed to reduce a variety of similar words to a common root form in order to improve our understanding of frequency. After preprocessing, the data is changed to a frequency.

The mentioned feelings can be divided into two main groups: positive sentiment and negative sentiment. Anger, fear, sadness, and disgust are categorised as elements of negative sentiment, while the remaining four emotions stand for positive sentiment. Principal component analysis (PCA) is used to create scores for the two sentiments by assigning appropriate weights to each of the derived factor loadings for the respective groups of emotions. The statistically more suitable method for creating the sentiment scores is PCA. In the end, the relative dominance of two types of market sentiments is measured by calculating the relative share of each sentiment type. The following formula is used to determine the relative share of each market sentiment:

POSIT = Sp/ (Sp + Sn) ………………………………..(2)

Whereas POSIT denotes the proportion of positive sentiment on a given day, Sp denotes the day's positive sentiment score and Sn denotes the day's negative sentiment score. This suggests that we can designate the following as a share of negative market sentiment (NEGAT):

NEGAT = 1-POSIT ………………………………………..(3)

4. Volatility Measurement & Result Analysis

Our goal is to evaluate how two opposing market sentiments affect the conditional volatility of the Indian stock market. To examine the influence of market sentiments, three distinct generalised autoregressive conditional heteroscedasticity (GARCH) models are employed. The GJR-GARCH model of Glosten, Jagannathan, and Runkle, as well as the standard GARCH model and its two variations, EGARCH, or exponential GARCH model, have been used. First, stationarity is tested for each of the three variables in the model. To verify that the data were non-stationarity, we performed two different unit root tests. Table 1 displays the results of the two tests, Augmented Dickey-Fuller (1979) and Kwiatkowski, Phillips, Schmidt, and Shin (1992).

Table 1: Unit Root Test

Tests	Model specification	Return	Share of Positive Sentiment	Share of Negative Sentiment
ADF unit root test	Intercept	-42.00216*	-35.023*	-35.00*
	Trend and Intercept	-22.53*	-16.31*	-17.96*
KPSS unit root test	Intercept	0.523	0.200	0.100
	Trend and Intercept	0.069	0.01†30	0.110

Source : Author's Compilation

European Economic Letters
ISSN 2323-5233
Vol 15, Issue 1 (2025)
http://eelet.org.uk

It is evident from Table 1 that both variables are stationary in level and need no further differencing for conversion to a stationary time series.

Since the work of Engle (1982) and Bollerslev (1986), GARCH models have been a highly popular conditional volatility modelling. A basic GARCH model is unable to capture the asymmetric impact of volatility, even though it can effectively estimate a conditional variance in a model. Differentiating between the effects of positive and negative shocks led to the conceptualization of the leverage effect, which in turn caused the GARCH model to be extended. The asymmetric effects of two distinct shock types on conditional volatility in financial data were captured by Glosten, Jagannathe, and Runkle's 1993 augmented version of the GARCH model. Two equations make up a standard GARCH(1,1) model: the conditional volatility equation (5) and the mean equation (4).

Y_t □ □ □□ *$Y_{t□1}$ □ □ *□$_{t□1}$ □□$_t$ ………………………………………………(4)

$ht = c*\varepsilon 2(t-1) + \delta *h(t-1) + \Upsilon *\varepsilon 2\ (t-1)\ * D(t-1) + \phi 1\ POSIT + \phi 2\ NEGAT$……………………..(5)

The conditional volatility equation of the GJR-GARCH(1,1) model follows a deterministic path, in contrast to mean equation (5), because conditional volatility, represented by ht, is dependent on both its own past and past error variance ($\varepsilon 2$). A dummy variable (Dt-1) is added to the model to evaluate the asymmetric impact of error. It assumes a value of zero otherwise and a value of unity if past error, $\varepsilon\varepsilon t-1$, is negative. The coefficient γ represents the unequal influence of error on conditional volatility. A positive ϕ indicates a greater impact of negative shock on conditional volatility, while a negative ϕ indicates a predominance of positive shock. Researchers are becoming more interested in the recent trend of capturing market sentiment through news articles, online information, etc., even though the traditional assumption of this GJR GARCH model to consider positive error ($\varepsilon\varepsilon t-1 > 0$) as a proxy for good news and its counterpart as bad news in the market has been empirically proven to be very successful in literature. This paper aims to improve the current GJR GARCH model by incorporating two previously suggested variables: the share of negative market sentiment (NEGAT) and the share of positive market sentiment (POSIT). The purpose of doing this is to more thoroughly evaluate the ways in which these two opposing sentiments influence the dynamics of conditional volatility.

Table 2: Estimated coefficients of Mean Equation (4)

Parameters	Values (SE)
M	0.0120(0.011)*
α	0.805(0.331)*
β	-0.930(0.226)*

Source : Author's compiltaion

Table 3: Estimated coefficients of augmented GJR GARCH model

Parameters	Values (SE)
C	0.123(0.011)*
ε	0.675(0.003)*
Υ	-0.239(0.456)*
$\phi 1$	-0.390(0.481)*
$\phi 2$	4.664 (1.438)*
Log likelihood = 3425.90 Wald Chi = 2080.34 No. of Obs. = 2990	

Source : Author's compilation

The GJR GARCH model's outcomes are shown in Table 3.The conditional volatility equation is used to determine the extent of influence on the overall risk profile of the market by incorporating coefficients of positive and negative sentiment based on the emotional quotients of news information. The relative influence of positive sentiment on conditional volatility is represented by the coefficient $\varphi 1$, which has a negative sign, while the relative importance of negative market sentiment on market volatility is measured by parameter $\varphi2$, which has a positive sign. It follows that while negative sentiment increases market volatility, positive sentiment appears to mitigate it. But since the estimated value of $\varphi2$ is 4.664 compared to $\varphi1$=-0.390, the impact of negative sentiment is found to be significantly larger than that of positive sentiment. It shows that the volatility of the Indian stock market is primarily influenced by negative sentiments rather than positive ones. Our experimental results clearly show that noise traders are in the lead. Positive sentiment's contractionary impact on market volatility signals the departure of noise traders, but when such traders are seen to be dominant in the market, the effect of positive sentiment is amplified, outweighing the departure of the noise traders. Our finding is in consonance with empirical works on Indian market by (Ferreira et al., 2021) and by (Simões Vieira, 2011). Nevertheless, utilising real stream data that is accessible on a daily basis is a more appropriate approach to assess market sentiments than the traditional macroeconomic and financial variables used in both of these studies to measure market sentiments.

5. Conclusion

The augmented asymmetric GARCH model with two opposing investor sentiments—positive and negative sentiments, for example has been examined in this paper. This analysis suggested news-based sentiment analysis for the Indian stock market in order to categories the positive and negative values of errors as corresponding sentiments in the market. Empirical results support the dominance of noise traders in the Indian stock market and indicate that the influence of negative sentiment is stronger than that of positive sentiment. The study's conclusions are helpful for financial policy because it highlights the existence and prominence of noise traders in the Indian stock market, which forces many sane investors to reconsider their current investment approaches. The immaturity of an emerging financial market such as India is evident when irrational exuberance shapes market volatility and market fundamentals are not the key determining factors. Appropriate regulatory actions and practices are needed to contain this situation.

Future research for this study will compare various Indian stock market sectors, such as energy, telecommunication, and metals, among others. These sectors can be evaluated, and their individual effects on the stock market as a whole can be examined to gain a detailed grasp of the dynamics of the market.

References

- Abbahaddou, K., & Chiadmi, M. S. (2022). Comparative Analysis of the Performance of Expert Systems and Machine Learning Models in the Context of the Islamic Stock Market. *WSEAS Transactions on Environment and Development*, *18*. https://doi.org/10.37394/232015.2022.18.93
- Adebiyi, A. A., Adewumi, A. O., & Ayo, C. K. (2014). Comparison of ARIMA and artificial neural networks models for stock price prediction. *Journal of Applied Mathematics*, *2014*. https://doi.org/10.1155/2014/614342
- Agarwal, S., Rajput, P., & Jena, A. K. (2022). A Hybrid Evolutionary model for Stock Price Prediction Using Grey Wolf Optimizer. *Proceedings - 2022 OITS International Conference on Information Technology, OCIT 2022*. https://doi.org/10.1109/OCIT56763.2022.00062
- Alroobaea, R. (2022). Sentiment Analysis on Amazon Product Reviews using the Recurrent Neural Network (RNN). *International Journal of Advanced Computer Science and Applications*, *13*(4). https://doi.org/10.14569/IJACSA.2022.0130437
- Aggarwal, D. (2022). Defining and measuring market sentiments: a review of the literature. In *Qualitative Research in Financial Markets* (Vol. 14, Issue 2). https://doi.org/10.1108/QRFM-03-2018-0033
- Baliyan, A., Batra, A., & Singh, S. P. (2021). Multilingual sentiment analysis using RNN-LSTM and neural machine translation. *Proceedings of the 2021 8th International Conference on Computing for Sustainable Global Development, INDIACom 2021*. https://doi.org/10.1109/INDIACom51348.2021.00126
- Basiri, M. E., Nemati, S., Abdar, M., Cambria, E., & Acharya, U. R. (2021). ABCDM: An Attention-based Bidirectional CNN-RNN Deep Model for sentiment analysis. *Future Generation Computer Systems*, *115*. https://doi.org/10.1016/j.future.2020.08.005
- Barbaglia, L., Consoli, S., & Manzan, S. (2020). Forecasting with Economic News. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3698121
- Bourezk, H., Raji, A., Acha, N., & Barka, H. (2020). Analyzing Moroccan Stock Market using Machine Learning and Sentiment Analysis. *2020 1st International Conference on Innovative Research in Applied Science, Engineering and Technology, IRASET 2020*. https://doi.org/10.1109/IRASET48871.2020.9092304
- Brazdil, P., Muhammad, S. H., Oliveira, F., Cordeiro, J., Silva, F., Silvano, P., & Leal, A. (2022). Semi-Automatic Approaches for Exploiting Shifter Patterns in Domain-Specific Sentiment Analysis. *Mathematics*, *10*(18). https://doi.org/10.3390/math10183232
- Byun, J., Kim, K., & Park, K. H. (2022). Review of Behavioral Finance Studies on Korean Capital Markets. In *Asian Review of Financial Research* (Vol. 35, Issue 4). https://doi.org/10.37197/ARFR.2022.35.4.3

- Cao, Y., & Zhai, J. (2023). Bridging the gap–the impact of ChatGPT on financial research. *Journal of Chinese Economic and Business Studies*, *21*(2). https://doi.org/10.1080/14765284.2023.2212434
- Chen, M., Guo, Z., Abbass, K., & Huang, W. (2022). Analysis of the impact of investor sentiment on stock price using the latent dirichlet allocation topic model. *Frontiers in Environmental Science*, *10*. https://doi.org/10.3389/fenvs.2022.1068398
- Chen, W., Cai, Y., Lai, K., & Xie, H. (2016). A topic-based sentiment analysis model to predict stock market price movement using Weibo mood. *Web Intelligence*, *14*(4). https://doi.org/10.3233/WEB-160345
- Chen, Y., Zhao, H., Li, Z., & Lu, J. (2020). A dynamic analysis of the relationship between investor sentiment and stock market realized volatility: Evidence from China. *PLoS ONE*, *15*(12 December). https://doi.org/10.1371/journal.pone.0243080
- CHAREF, F., & AYACHI, F. (2016). A Comparison between Neural Networks and GARCH Models in Exchange Rate Forecasting. *International Journal of Academic Research in Accounting, Finance and Management Sciences*, *6*(1). https://doi.org/10.6007/ijarafms/v6-i1/1996
- Chaudhuri, A. (2022). Sentiment Analysis of COVID-19 Reviews Using Hierarchical Version of d-RNN. *Computacion y Sistemas*, *26*(2). https://doi.org/10.13053/CyS-26-2-4143
- Chen, C. C., Chang, J. H., Lin, F. C., Hung, J. C., Lin, C. S., & Wang, Y. H. (2019). Comparison of Forcasting Ability between Backpropagation Network and ARIMA in the Prediction of Bitcoin Price. *Proceedings - 2019 International Symposium on Intelligent Signal Processing and Communication Systems, ISPACS 2019*. https://doi.org/10.1109/ISPACS48206.2019.8986297
- Chen, J., Yan, S., & Wong, K. C. (2020). Verbal aggression detection on Twitter comments: convolutional neural network for short-text sentiment analysis. *Neural Computing and Applications*, *32*(15). https://doi.org/10.1007/s00521-018-3442-0
- Fedorova, E., Chertsov, P., & Kuzmina, A. (2023). COVID-19: the impact of the pandemic fear on IPO underpricing. *Journal of Financial Reporting and Accounting*, *21*(4). https://doi.org/10.1108/JFRA-05-2022-0201
- Ferreira, F. G. D. C., Gandomi, A. H., & Cardoso, R. T. N. (2021). Artificial Intelligence Applied to Stock Market Trading: A Review. In *IEEE Access* (Vol. 9). https://doi.org/10.1109/ACCESS.2021.3058133
- Gao, Z. (2021). Stock Price Prediction with ARIMA and Deep Learning Models. *2021 IEEE 6th International Conference on Big Data Analytics, ICBDA 2021*. https://doi.org/10.1109/ICBDA51983.2021.9403037
- Gao, Z., Li, Z., Luo, J., & Li, X. (2022). Short Text Aspect-Based Sentiment Analysis Based on CNN + BiGRU. *Applied Sciences (Switzerland)*, *12*(5). https://doi.org/10.3390/app12052707
- Gonzalez-Igual, M., Corzo Santamaria, T., & Rua Vieites, A. (2021). Impact of education, age and gender on investor's sentiment: A survey of practitioners9. *Heliyon*, *7*(3). https://doi.org/10.1016/j.heliyon.2021.e06495
- Gupta, A., Dengre, V., Kheruwala, H. A., & Shah, M. (2020). Comprehensive review of text-mining applications in finance. In *Financial Innovation* (Vol. 6, Issue 1). https://doi.org/10.1186/s40854-020-00205-1
- Hassan, M. K., Hudaefi, F. A., & Caraka, R. E. (2022). Mining netizen's opinion on cryptocurrency: sentiment analysis of Twitter data. *Studies in Economics and Finance*, *39*(3). https://doi.org/10.1108/SEF-06-2021-0237
- Hossain, E., Sharif, O., & Hoque, M. M. (2022). MemoSen: A Multimodal Dataset for Sentiment Analysis of Memes. *2022 Language Resources and Evaluation Conference, LREC 2022*.
- Huma, Z. e, & Ayub, R. (2018). An Emerging Need of Introducing Behavioral Finance in Pakistan: Evidence from Investor's Sentiments and Actions. *JISR Management and Social Sciences & Economics*, *16*(1). https://doi.org/10.31384/jisrmsse/2018.16.1.2
- Jain, P. K., Pamula, R., & Srivastava, G. (2021). A systematic literature review on machine learning applications for consumer sentiment analysis using online reviews. In *Computer Science Review* (Vol. 41). https://doi.org/10.1016/j.cosrev.2021.100413
- Katsafados, A. G., Nikoloutsopoulos, S., & Leledakis, G. N. (2023). Twitter sentiment and stock market: a COVID-19 analysis. *Journal of Economic Studies*. https://doi.org/10.1108/JES-09-2022-0486
- Kumar, A., & Jaiswal, A. (2020). Systematic literature review of sentiment analysis on Twitter using soft computing techniques. *Concurrency and Computation: Practice and Experience*, *32*(1). https://doi.org/10.1002/cpe.5107
- Kumar, S., Rao, S., Goyal, K., & Goyal, N. (2022). Journal of Behavioral and Experimental Finance: A bibliometric overview. In *Journal of Behavioral and Experimental Finance* (Vol. 34). https://doi.org/10.1016/j.jbef.2022.100652
- Khan, L., Amjad, A., Afaq, K. M., & Chang, H. T. (2022). Deep Sentiment Analysis Using CNN-LSTM Architecture of English and Roman Urdu Text Shared in Social Media. *Applied Sciences (Switzerland)*, *12*(5). https://doi.org/10.3390/app12052694

- Lohrbach, T., & Schumann, M. (1992). Artificial Neural Networks and ARIMA-Models within the Field of Stock Market Prediction - A Comparison. In *Adaptive Intelligent Systems*. https://doi.org/10.1016/b978-0-444-89838-8.50004-7
- Lachana, I., & Schröder, D. (2022). Investor sentiment, social media and stock returns: Wisdom of crowds or power of words? *Social Science Research Network (SSRN).*
- Lachana, I., & Schroeder, D. (2021). Investor Sentiment, Media and Stock Returns: The Advancement of Social Media. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3842039
- Lakhchini, W. (2022). Artificial Intelligence & Machine Learning in Finance: A literature review. *Finance, Auditing, Management and Economics*, *3*(1).
- Li, X., Wu, P., & Wang, W. (2020). Incorporating stock prices and news sentiments for stock market prediction: A case of Hong Kong. *Information Processing and Management*, *57*(5). https://doi.org/10.1016/j.ipm.2020.102212
- Li, Y., & Pan, Y. (2022). A novel ensemble deep learning model for stock prediction based on stock prices and news. *International Journal of Data Science and Analytics*, *13*(2). https://doi.org/10.1007/s41060-021-00279-9
- Luo, Y. T., & Zhang, M. (2022). Comparison and Analysis of Financial Text Information Value of Listed Companies Based on WoE. *Procedia Computer Science*, *214*(C). https://doi.org/10.1016/j.procs.2022.11.259
- Ma, B., Yang, Y., Zhang, J., & Zhang, K. (2023). A Comparison of Stock Price Prediction with ANN and ARIMA. *BCP Business & Management*, *38*. https://doi.org/10.54691/bcpbm.v38i.3719
- Ma, Q. (2020). Comparison of ARIMA, ANN and LSTM for Stock Price Prediction. *E3S Web of Conferences*, *218*. https://doi.org/10.1051/e3sconf/202021801026
- Mahadevaswamy, U. B., & Swathi, P. (2022). Sentiment Analysis using Bidirectional LSTM Network. *Procedia Computer Science*, *218*. https://doi.org/10.1016/j.procs.2022.12.400
- Naas, S. E., Bendob, A., & Bensania, A. (2021). Towards construction of an investor sentiment index in the Arab stock markets. *An-Najah University Journal for Research - B (Humanities)*, *35*(4). https://doi.org/10.35552/0247-035-004-005
- Naufal Adi Nugroho, & Erwin Budi Setiawan. (2021). Implementation Word2Vec for Feature Expansion in Twitter Sentiment Analysis. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, *5*(5). https://doi.org/10.29207/resti.v5i5.3325
- Paule-Vianez, J., Gómez-Martínez, R., & Prado-Román, C. (2020). A bibliometric analysis of behavioural finance with mapping analysis tools. *European Research on Management and Business Economics*, *26*(2). https://doi.org/10.1016/j.iedeen.2020.01.001
- P.Kavya, S. S. R. T. S., & C.G.Nivetheni, Dr. M. M. Dr. P. V. (2021). Stock Market Analysis Using Mapreduce and Pyspark. *International Research Journal of Modernization in Engineering Technology and Science* , *3*(7).
- Rajeswari, K., Vispute, S. R., Gangatirkar, S., Dixit, K., Chaudhari, S., & Bhole, S. (2022). Comparative Analysis of various Machine Learning Algorithms for Stock Price Prediction. *2022 International Conference on Smart Generation Computing, Communication and Networking, SMART GENCON 2022*. https://doi.org/10.1109/SMARTGENCON56628.2022.10084008
- Rayegan, A., Shiri, A., & Bahrak, B. (2022). A comparative study of machine learning techniques for stock price prediction. *Proceedings - 2022 8th International Iranian Conference on Signal Processing and Intelligent Systems, ICSPIS 2022*. https://doi.org/10.1109/ICSPIS56952.2022.10044075
- Reis, P. M. N., & Pinho, C. (2021). A dynamic factor model applied to investor sentiment in the European context. *Investment Management and Financial Innovations*, *18*(1). https://doi.org/10.21511/imfi.18(1).2021.25
- Shastri, V. (2023). Comparing Statistical, Deep Learning, and Additive Models for Forecasting in the Indian Stock Market. In *Artificial Intelligence for Capital Markets*. https://doi.org/10.1201/9781003327745-9
- Shah, D., Isah, H., & Zulkernine, F. (2019). Predicting the Effects of News Sentiments on the Stock Market. *Proceedings - 2018 IEEE International Conference on Big Data, Big Data 2018*. https://doi.org/10.1109/BigData.2018.8621884
- Shen, X., Wang, G., & Wang, Y. (2021). The Influence of Research Reports on Stock Returns: The Mediating Effect of Machine-Learning-Based Investor Sentiment. *Discrete Dynamics in Nature and Society*, *2021*. https://doi.org/10.1155/2021/5049179
- Simões Vieira, E. (2011). Investor sentiment and the market reaction to dividend news: European evidence. *Managerial Finance*, *37*(12). https://doi.org/10.1108/03074351111175100
- Singh, B. (2021). A bibliometric analysis of behavioral finance and behavioral accounting. *American Business Review*, *24*(2). https://doi.org/10.37625/ABR.24.2.198-230

- Sinyak, N., Tajinder, S., Madhu Kumari, J., & Kozlovskiy, V. (2021). Predicting real estate market trends and value using pre-processing and sentiment text mining analysis. *Real Estate: Economics, Management*, *1*. https://doi.org/10.22337/2073-8412-2021-1-35-43
- Staszkiewicz, P., & Staszkiewicz, R. (2022). Security exchange commission forms K-10 filings – Positive and negative word occurrence dataset 1995–2008. *Data in Brief*, *42*. https://doi.org/10.1016/j.dib.2022.108110
- Trierweiler Ribeiro, G., Alves Portela Santos, A., Cocco Mariani, V., & dos Santos Coelho, L. (2021). Novel hybrid model based on echo state neural network applied to the prediction of stock price return volatility. *Expert Systems with Applications*, *184*. https://doi.org/10.1016/j.eswa.2021.115490
- Vieira, E. F. S. (2016). Investor sentiment and share returns: evidence on family firms. *Academia Revista Latinoamericana de Administración*, *29*(1). https://doi.org/10.1108/arla-08-2015-0234
- Vishwakarma, G. K., Paul, C., Elsawah, A. M., Tang, L., Wang, G., Liu, J., Blitz, A., Kawabata, N., سيادت سعيد, Abdul-Wahab, S. A., Al-Alawi, S. M., Janik, A., Gautam, A., Singh, V., Taylor, S. J., Letham, B., Li, D., Moghaddam, M. R., Monjezi, M., … Statista. (2020). Forecasting steel prices using ARIMAX model: A case study of Turkey. *Resources Policy*, *35*(8).
- Wang, S., Tang, Z., & Chai, B. (2016). Exchange rate prediction model analysis based on improved artificial neural network algorithm. *Proceedings of the International Conference on Communication and Electronics Systems, ICCES 2016*. https://doi.org/10.1109/CESYS.2016.7889912
- Wiranata, R. B., & Djunaidy, A. (2021). Optimasi Hyper-Parameter Berbasis Algoritma Genetika Pada Ensemble Learning Untuk Prediksi Saham Yang Mempertimbangkan Indikator Teknikal & Sentimen Berita. *JATISI: Jurnal Teknik Informatika Dan Sistem Informasi*, *8*(3).
- Wu, C. C., Yan, Y., Yuan, T., Huang, C. C., & Tsai, Y. J. (2022). A Study of Network Negative News Based on Behavioral Finance Analysis of Abnormal Fluctuation of Stock Price. *Discrete Dynamics in Nature and Society*, *2022*. https://doi.org/10.1155/2022/7952532
- Xiong, G., Wang, L., & Wu, D. (2020). Influence of investor sentiment on the return rate of transnational investment behaviour. *Revista Argentina de Clinica Psicologica*, *29*(1). https://doi.org/10.24205/03276716.2020.65
- Yadav, A., & Chakraborty, A. (2022). Relationship between investor sentiment and stock returns: a bibliometric analysis using SCOPUS database. *Global Business and Economics Review*, *28*(1). https://doi.org/10.1504/gber.2023.127554