International Journal of Advance and Innovative Research
Volume 12, Issue 2 (XVII): April - June 2025
ISSN 2394 - 7780

IMPACT OF MAJOR NEWS EVENTS ON INVESTOR SENTIMENT AND HERD BEHAVIOR: A STUDY ON COVID-19 PANDEMIC

[1]Swarda Ankush Parab and [2]Sunita Jena
[1]Student and Assistant Professor, Department of Big Data Analytics, Jai Hind College, Autonomous, Mumbai - 400020

ABSTRACT
To test the impact of the significant news events during the COVID-19 epidemic to the Indian stock market, this study makes the use of NIFTY-50 index for the analysis of the investors' sentiment and the extent of the herd mentality. As a result of analysing news data, sentiment scores have been obtained to measure the reactions of investors with the help of various sentiment analysis tools, including VADER, BERT, and FinBERT. The study then used the Cumulative Abnormal Returns (CAR) results to demonstrate the positive and significant relationship between sentiment and market returns using the random forest, gradient boosting and extreme gradient boosting regressors. Besides, cross sectional measures of CSAD and CSSD established that herding behaviour was prominent during periods of high volatility. These results indicate that particularly during crises investor activity significantly influences market dynamics and encourages group action. Thus, this research helps to expand knowledge of the psychological factors influencing financial markets and to provide recommendations for practical improvements in decision-making and improving the stability of financial markets in conditions of uncertainty.

Keywords *- Investor Sentiment, Herd Behaviour, COVID-19, Sentiment Analysis, Stock Market*

I. INTRODUCTION

Investor sentiment refers to the overall attitude or perception of investors about particular stocks or market conditions. Such sentiments play an important role in influencing market movements and often result in prices deviating from their intrinsic values. By taking data points from informational sources such as financial data, social media activity, and news articles, sentiment analysis has become an effective tool in predicting market behaviour because of advancements made in machine learning and natural language processing. Thus, today investors as well as analysts have a better understanding of what comprises public opinion and subsequently can forecast market actions more effectively due to technological innovations.

Investor sentiment is the general feeling that investors have about specific stocks or market situation. Opinions of this kind are essential in determining prices in the market and are frequently responsible for the disparity between the actual values and the prices. Because of the development of the machine learning and natural language processing, sentiment analysis provides the data taken from informational sources including financial data, social activity, and articles to predict market trends. Thus, today the investors and analysts have improved understanding what constitutes public opinion and accordingly can predict the market actions because of technological advancement.

While herd behaviour is that situation where investors relinquish their own discretion and decision-making capabilities and begin to emulate others. The action of one investor ends up influencing the next then the next till it reaches a large group of investors who have switched their actions according to trends this is like the ripple effect. The two main reasons are FOMO or you think others have more information than you do. However, such marketplace inefficiencies result from price deviation from their fundamental values through the involvement of more than one group-market bubbles or crashes may follow. Daniel Kahneman's dual-system theory offers an account for this phenomenon because he argues that investors tend to make fast, unconscious-System 1- or slow, conscious-System 2- decisions most of the time. Herd behaviour is particularly so because emotional reactions are high in events of significant news communications-health pandemics, policy changes or economic crises-and this indeed makes the markets volatile also.

This paper seeks to establish how large news events influence the investors' feelings and hence create the herding behaviour. It can be appreciated that events in the news are random while the responses of the markets are predictable when the right analytical tools are applied to make the estimates hence the need for a study of this nature. Investors and analysts will be equipped with adequate tools once academicians undertake research into the direction and herding activity. However, at the same time, this research demonstrates how the use of sentiment analysis makes it possible to determine that herding is dominant, and therefore, investors should not make unsuitable decisions influenced by prejudice emotions and policymakers should be capable of taking necessary actions that would contribute to the establishment of stable market condition.

In today's interconnected global financial market, it is important to investigate how big news items bear on investor feelings and trends. This research aims at increasing the current knowledge on how investor psychology reacts to external stimuli, as well as at providing practical implications to maintain stability in the markets, refining decision-making, and advancing risk management for analysts, investors, and regulators.

II. RELATED WORK

A. Sentiment-Related Studies

The study by Verma et al. (2017) categorize news events into economic, political, and social views using Support Vector Machines and other machine learning methodologies with an intention to establishing the impact of news on Indian stock markets. [1] Their research argues for news as an important factor of the market in consideration, by employing Granger causality: instead, providing evidence of causality of news types to changes in the markets. Similarly to above Hanna et al. (2020) who investigated the impact of the media concerning the mood of the investors in the bull and bear markets. It was proposed that news sentiment would be in the Financial Times and would lead to stock returns particularly during bullish periods when trading activity was linked to attitude [2]. In 2019, Papakyriakou and his team examined the impacts of G7 terrorist attacks on international stock exchanges; they also found that with attacks and subsequent attacks, the stock return sentiment decreases with higher market losses. This scheme is worsened by social and news media [3].

Ren et al. (2020) suggest a BERT-BiLSTM hybrid model for improved sentiment analysis in the energy industry, which results in better trend forecasts in the energy market [4]. Similarly, Day and Lee (2019) use deep learning for financial sentiment research, demonstrating that the quality of news sources affects prediction accuracy and supporting deep learning models over lexicon-based methods for longer-term, more precise market predictions [5].

B. Psychological Biases and Emotional Influences

Aigbovo & Ilaboya (2019) talk about how behavioural biases affect Nigerian investors and point to aspects of prospect theory, such as loss aversion, as major causes of less-than-ideal investment choices [6]. They contend that emotional elements, such as fear and greed, lead to heuristicdriven behaviours that frequently result in unfavourable consequences, and they recommend that financial education and assistance could lessen these biases. Othman (2021) delves deeper into psychological factors by illustrating biases such anchoring and overconfidence in investor decision-making using Daniel Kahneman's dual-system theory. The paper proposes methods such as journaling and financial advising to avoid biases and improve rational decisions regarding investments by incorporating AI for logical decision assistance [7].

C. Herding Behaviour in Financial Markets

By differentiating between information-based and reputation-based herding, Bikhchandani and Sharma (2000) offer fundamental insights into herding. They advocate for legislative actions to increase transparency and reduce market volatility, emphasizing the significance of distinguishing between real herding and illusory converge [8]. Chiang and Zheng (2010) look at herding tendencies around the world and find that there is a lot of herding in Asian and developed markets (but not in the US), and that herding gets worse globally during financial crises due to contagion from US markets, which raises systemic risk [9]. Herding is common during market turbulence and increased trade volumes, particularly in China, according to Lao & Singh's (2011) comparison of the Chinese and Indian markets. Because of increased herding during crises, like the 2008 financial crisis, they demand stronger rules and regulations in China [10].

Dang and Lin (2016) examine herding's reliance on heterogeneous information with an emphasis on emerging markets, especially for amateur investors who must contend with high information costs. [11] They believe that these investors frequently imitate their profitable colleagues, a behaviour that necessitates a more distinct factual distinction between genuine and fraudulent herding. In an examination of media-driven herding in the cryptocurrency market during COVID-19, Youssef & Waked (2020) point out how different news sources, influenced by reporting style and media bias, affect investor choices [12]. They uncover media-induced herding trends by using deep learning to improve sentiment accuracy. Ren and Wu (2018) expand on this strategy by measuring herding in blue-chip stocks using forum sentiment analysis [13]. They discover that herding is exacerbated by negative mood and that macroeconomic considerations play a significant role in influencing such conduct.

ISSN 2394 - 7780

III. DATA AND METHODOLOGY

A. Data Collection and Preprocessing 1. News Data Collection

The initial step was to use Google Advanced Search and the WebScraper Chrome extension to scrape news articles from multiple internet sources. The following strategy was used:

Strategy for Searching: News articles from January 1, 2020, to December 31, 2020 were gathered using particular keywords associated with market indices (like the NIFTY50) and economic events (like the economic crisis). Using Google's date filter, the search was honed to collect URLs inside this particular time window.

WebScraping: All recognized URLs were extracted and saved in a CSV file using the WebScraper Chrome extension. The foundation for additional data extraction was this CSV file.

2. Data Extraction

BeautifulSoup Extraction: Each scraped URL was accessed using the BeautifulSoup library in Python to extract relevant fields. The resulting dataset was stored in a structured CSV format, containing the columns: text, url, Published Date, Title, Text, and Summary.

3. Data Preprocessing

The dataset underwent a series of preprocessing steps to clean and standardize the data. Stopwords were removed, NaN values were handled appropriately, and additional text processing techniques such as stemming or lemmatization were applied where necessary. The final dataset, which had a shape of (1437, 6) after preprocessing, formed the input for sentiment analysis.

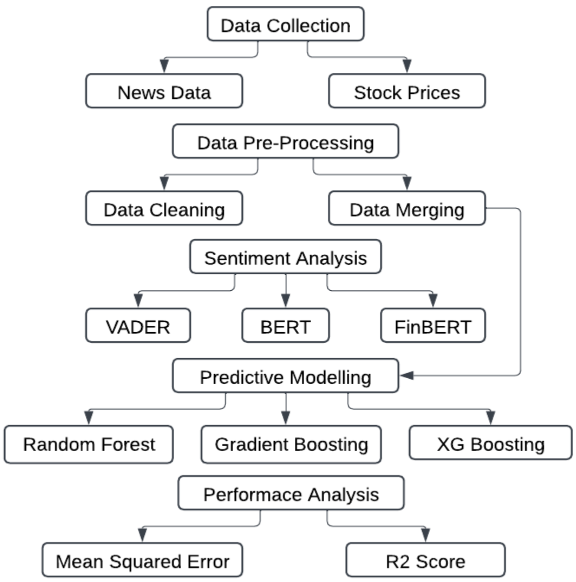

Fig 1: Workflow for sentiment analysis

B. Sentiment Analysis 1. Sentiment Analysis Models

Three pre-trained sentiment analysis models were employed to analyse investor sentiment:

VADER (Valence Aware Dictionary and sEntiment Reasoner): A rule-based model used for general sentiment analysis.

BERT (Bidirectional Encoder Representations from Transformers): A state-of-the-art NLP model that excels in sentiment analysis tasks.

FinBERT: A financial model emerged from BERT and specially used for the analysis of the sentiment of the financial news.

2. Sentiment Scoring

The sentiment scores which have been estimated using each model was then summed up and incorporated with the stock market data. In this study the stock price data of Nifty 50 index stocks were taken including; Open prices, High, Low, Close and the Adjusted close price data. Daily returns were calculated as the first difference of the closing stock prices for a specific day.

C. Feature Engineering 1. Lagged Features and Event Annotations
New features were included in order to better enhance the model's prediction ability:

Lagged Sentiment Features: As the impact of news on stock prices may initially be delayed, lagged sentiment scores were created.

Event Dates: In order to determine whether or not significant news events have a major effect on market sentiment, some of the most noteworthy events were listed and highlighted.

Moving Averages: To identify trends, moving averages (5 day and 10 day) of sentiment scores were computed.

D. Predictive Modelling 1. Machine Learning Models
Three machine learning models were employed to predict stock returns using the engineered features:

Random Forest Regressor: An iterative learning approach to the prediction of stock return based on sentiment and market characteristics.

Gradient Boosting Regressor: An improved boosting algorithm to capture the intricate features in the data by making refinements on the model's estimations. The accuracy of the models was assessed by means of Mean Square Error (MSE) and R2 score.

Extreme Gradient Boosting: A fast and easily parallelized version of gradient boosting, XGBoost enhances the prediction of outcome by iteratively learning from the mistakes of previous models.

E. Herd Behaviour Analysis 1. Calculation of CSAD and CSSD

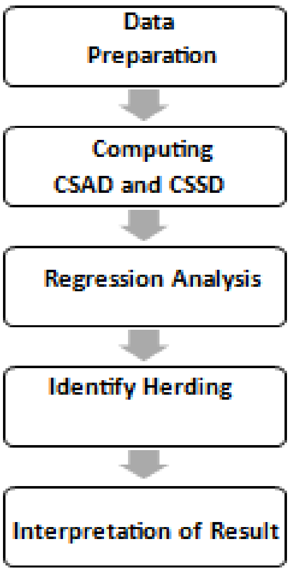

Fig 2: Flow chart For Herding Analysis

To investigate herd behaviour, the Cross-Sectional Absolute Deviation (CSAD) and Cross-Sectional Standard Deviation (CSSD) methodologies were applied: **Market Benchmark:** The market value benchmark was NIFTY-50 index.

Stock Selection: A sample of ten stocks was chosen at random across different sectors for the analysis of individual stock performance against the benchmark index.

CSAD and CSSD Calculation: The CSAD and CSSD values were calculated for the selected stocks used in the analysis. These measures were useful in identifying the variances of specific stock from the overall market index during situations of market volatility or risk.

2. Regression Analysis for Herd Detection
To do this, a regression of the market return on the CSAD/CSSD scores was conducted in an attempt to determine the extent of the association between the two. Nonlinear patterns are expected to suggest the presence of investor herding behaviour, as established by the objective of the study. Specifically, the regression model explored if one of the key herding indicators, namely CSAD or CSSD, reduced when the absolute market return increased.

International Journal of Advance and Innovative Research
Volume 12, Issue 2 (XVII): April - June 2025
ISSN 2394 - 7780

IV. RESULTS

A. Sentiment Analysis Results 1. Sentiment Analysis using VADER, BERT, FinBERT

The analysis of sentiment was done using VADER, BERT and FinBERT models. These models estimate the sentiment to turn in news articles that tracked changes in investor emotions following significant news events during COVID-19. FinBERT which is a finance-oriented sentiment analysis model provided a high level of correlation with the market movements of the sentiment scores and highlighted the importance of the application of specialized models in the financial domain.

2. Stock Data Analysis (NIFTY-50)

Besides, the sentiment assessment was performed and compared with stock market data of the NIFTY-50 index. It revealed the change in price trends and returns during the Covid-19 period by focusing on the daily movement of price and other quantitative indicators including moving averages. It allowed a direct comparison of sentiment scores with market behaviour and established a framework for subsequent predictive analysis.

3. Predictive Modelling –

Random Forest, Gradient Boosting and Extreme Gradient Boosting Regressors

A quantitative assessment of the positive relationship between market investor sentiments and stock market return was performed through the use of predictive analytics. In this paper, this research aimed at predicting complex, nonlinear relationships between the scores of sentiment and market returns using machine learning methods including Random Forest, Gradient Boosting and Extreme Gradient Boosting Regressors. These models are suitable for pioneering financial market forecasting due to their ability to capture numerous interactions in the relationship between stock returns and investors, due to the consideration of the patterns of large quantities of data. The main findings from the predictive models are presented in table 1.

Table 1: Model Performance

Model \ Metric	R2 Score	Mean Square Error (MSE)
Random Forest Regressor with Cross Validation	0.8668	0.000511
Random Forest Regressor with Hyperparameter Tuning	0.8668	0.000511
Gradient Boosting with Cross Validation	0.8961	0.000398
Gradient Boosting with Hyperparameter Tuning	0.8950	0.000403
Ensemble Model (combination of RF and GB)	0.8821	0.000452
Extreme Gradient Boosting	0.9017	0.00037
Extreme Gradient Boosting Hyperparameter Tuning	0.9155	0.00032

These findings show that the models were able to explain the degree of association between the sentiment indicators and the stock returns evidenced by high R^2 scores and low MSE values. Out of the models, specifically, the Extreme Gradient Boosting Regressor showed high accuracy, which implies that sentiment has a strong influence in predicting the stock market during such episodes as the COVID-19 crisis.

CUMULATIVE ABNORMAL RETURNS (CAR) ANALYSIS

To examine market reaction to specific major news events during the pandemic, the analysis of Cumulative Abnormal Returns (CAR) was done. CAR sums the excess of stock return over the expected amount around event windows and can explain how the market views and responds to significant events. The following study looked at CAR around important event dates in order to find out whether event specific information altered investor sentiment with a resultant effect on stock market returns.

Table 2: CAR Analysis result

Event Date	Value
March 24, 2020 (Nationwide lockdown announced)	-0.1464
April 15, 2020 (Lockdown extended)	-0.2194
May 1, 2020	0.0249

(Lockdown extension)	
July 1, 2020 (Unlock phase 2.0)	-0.6657

These values show relatively significant market response to the major events evidenced by increased investors' risk mitigation and negativity during the event releases. That is why cases like on March 24 and April 15, which show negative CAR values but higher investor concerns due to the pandemic, indicate the impact of sentiment.

B. Herding Behaviour Analysis 1. Cross-Sectional Absolute Deviation (CSAD) and Cross-Sectional Standard Deviation (CSSD) Results

In the CSAD and CSSD studies, herding behaviour among the investors was detected. It is a phenomenon whereby investors all act in unison instead of independently by keeping in tow with the market.

Table 3: Regression result for CSAD and CSSD

	CSAD Analysis	**CSSD Analysis**
Intercept Coefficient	0.0139	0.0174
Coefficient for Absolute Market Return	0.1222 (p-value: 0.003)	0.1616 (p-value: 0.002)
Coefficient for Squared Market Return	-1.0614 (p-value: 0.018)	-1.3777 (p-value: 0.017)
R^2 Score	0.048	0.051

Thus, negative Squared Market Return coefficients in both models further indicate that the larger the market returns, the lower the dispersion of returns. This study provides evidence for the presence of herding behaviour, that is; the investors seem to act as one in the course of the period of high volatility in the market.

Using Regression Analysis to Identify Herding: The regression model was used to analyse the relationship between CSAD/CSSD values and market returns. Herding behaviour was confirmed by the large negative coefficient for Squared Market Return, which indicated that as the figure of market returns increased (up or down) the variability of returns decreased.

Volume-Return Relationship and Correlation Analysis: Contrary to conjecture, this paper did not establish any strong evidence of herding, based solely on trading volumes, when trading volume was correlated with stock returns. Yet, the results of the CSAD and CSSD provided some evidences to herding and indicated that the investor choices were closer to market returns rather than trade frequency.

V. CONCLUSION

In the present research while testing the hypothesis, the NIFTY-50 index was taken as the benchmark to study how important event triggered variation in investors' perception and herd mentality in the Indian stock market during the COVID-19 outbreak. Market and news data indicators were combined with sentiment scores derived with help of stateof-art sentiment analysis models such as VADER, BERT and FinBERT. There was a significant relationship between sentiment and stock market returns which was established using Random Forest and Gradient Boosting Regressors for predictive modelling. CAR research also showed that investor attitude came into play during uncertain times as changes in market behaviour were found to be due to significant news events.

Furthermore, cross sectional measures (CSAD and CSSD) were employed to analyse the herd behaviour. The results showed that the investors acted in a herd, the evidence being when return dispersion declined during periods of high turbulence. In conclusion, this work establishes an understanding of how news-driven emotion influences market returns and how often people tend to herd in emergency situations. These findings extend the existing literature on how external factors affect the decision-making process of investors and offer useful information about the psychological processes that underlie trading in finance.

ACKNOWLEDGEMENT

I would like to express my gratitude to my college faculty for their invaluable guidance and support during my research work. I am immensely thankful of my parents' and sister's unwavering support and faith, which served as my inspiration. Rahul deserves special mention for his insightful feedback and recommendations. This paper would not have been possible without all of you.

REFERENCES

1. A. J. Hanna, J. D. (2020). "News media and investor sentiment during bull and bear markets,". The Eauropean Journal of Finance.
2. I. Verma, L. D. (2017). "Detecting, Quantifying and Accessing impact of News events on Indian Stock Indices,". Association for Computing Machinery.
3. J., A. O. (2019). "Does Behavioural Biases Influences Individual Investment,". Management Science Review, vol. Vol 10(1).
4. Lee, M. -Y.-.. (2016). "Deep Learning for Financial Sentiment Analysis on Finance News Providers," . International Conference on Advances in Social Networks Analysis and Mining (ASONAM), San Francisco, CA.
5. Lin, H. V. (2016). "Herd Mentality in the Stock Market: On the Role of Idiosyncractric Participants with Heterogeneous Informtion". SSRN.
6. Othman, N. N. (2024). "Emotional Economics: The Role of Psychological Biases in Personal Investment Outcomes,". SSRN.
7. P. Papakyriakou, A. S. (2019). "Impact of terrorist attacks in G7 countries on international stock markets and the role of investor sentiment,". Elsevier.
8. Pagolu, V., K.N., R., G., P., & B., M. (2016). "Sentiment Analysis of Twitter Data for Predicting Stock Market Movements," International conference on Signal Processing, Communication, Power and Embedded System.
9. R. Cai, B. Q. (2020). "Sentiment Analysis About Investors and Consumers in Energy Market Based on BERTBiLSTM,". IEEE Access.
10. Sharma, S. B. (2001). "Herd Behavior in Financial Markets,". IMF Staff Papers, vol. Vol. 47.
11. Singh, P. L. (2011). "Herding Behaviour in the Chinese and Indian Stock Markets,". SSRN.
12. Waked, M. Y. (2022). "Herding behavior in the crypocurrency market during COVID-19 pandemic: The role of media coverage. Elsevier.
13. Wu, R. R. (2018). An Innovative Sentiment Analysis to Measure Herd Behavior," IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS: SYSTEM.
14. Zheng, T. C. (2010). "An empirical analysis of herd behavior in global stock markets,". Elsevier.