



INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X

Impact Factor: 6.078

(Volume 12, Issue 2 - V12I2-1256)

Available online at: <https://www.ijariit.com>

News-Aware Stock Market Movement Prediction for India Retail Traders

Arun Kumar K

arunkumar47954@gmail.com

Kumaraguru College of Technology, Tamil Nadu

Gokul Karthik G M

gokulkarthik1086@gmail.com

Kumaraguru College of Technology, Tamil Nadu

Shan Shad M

Shanshad.mohammediqbal@gmail.com

Kumaraguru College of Technology, Tamil Nadu

Anitha P

anitha.p.cse@kct.ac.in

Kumaraguru College of Technology, Tamil Nadu

ABSTRACT

Generally, retail investors have been experiencing various difficulties in handling financial markets due to the impact of ever-changing price movements in conjunction with ever-changing financial news. In normal circumstances of trading mechanisms, it is possible to observe historical price movements or sentiments. However, it is not possible to observe the contextual relationship between financial news and financial markets. Such cognitive complexities always affect decision-making in an unfavorable manner. In order to bridge the knowledge gap in this regard, this paper proposes the idea of developing a trading interpreter that considers financial news sentiments and financial market price data in an integrated manner. Natural language processing techniques have been used for developing a system that extracts sentiments from financial news articles. Sentiments are mapped with structured financial market price intervals. Feature engineering techniques have been used for developing financial news sentiments, price-based feature development, and interaction feature development that considers immediate reactions and lagged reactions of financial markets with respect to financial news.

Keywords: Financial Sentiment Analysis, Stock Market Prediction, Natural Language Processing, Trading Decision Support, Market Price Analysis, Web-Based Trading System, Predictive Analytics.

I. INTRODUCTION

Continuous global changes occur within financial markets due to both numeric/financial data & non-numeric/non-financial data (like financial statements to news about world events) which lead to the increased retail investor participation since they can trade online through online brokerages who are now readily available. However, retail investors will have difficulty in understanding/navigating the environment of the stock market due to the need to access many sources before determining market direction which has many different sources located all over the world. Overall, the financial markets will provide traders with two different types of analysis methods in which analysts can assess numerical price and non-price/financial data. Traders apply historical market data (price information) with predictive analytics (statistic modelling) to forecast stock price by applying traditional trading methods with successfully demonstrating an ability to identify price patterns within the numeric data used in forecasting future trends. However, because people react to news as soon as it happens, using traditional market data to forecast stock prices only address price history and do not include non-price/financial news in their models. With the continued improvements in Natural Language Processing and machine learning research new methods to extract financial-related text data from non-price/financial documents have become available for use within the financial market.

II. LITERATURE REVIEW

Stock Trend Prediction Using News Sentiment Analysis” – Kalyani et al., 2016

The significance of forecasting the stock market trends is clearly shown in the above paper. This paper has utilized different techniques of sentiment analysis in the process of retrieving the sentiment polarity and relating the sentiment polarity with the stock prices. This shows the significance of sentiment in forecasting the stock market trends.

Stock Prediction by Integrating Sentiment Scores of Financial News and Historical Stock Price Data” – Maqbool, 2023

The focus of the current study is on combining the results that are obtained from the financial news data and the historical price data. In the development of the model, the hybrid approach is used. It is found that the hybrid approach is appropriate for accuracy.

Stock Price Prediction Using FinBERT-Enhanced Sentiment with Technical and Statistical Indicators” – Ruan, 2025

The study would be a contribution to the development of the framework for sentiment analysis using FinBERT, besides being a contribution to the development of the technical indicators for the prediction of stock prices in the short term. It would also improve the accuracy of the sentiment analysis classification and the prediction of the approach.

Financial Sentiment Analysis Using FinBERT for Market Movement Prediction” – Jiang et al., 2023

The objective of this paper is to discuss the possibility of utilizing the FinBERT model in the context of sentiment extraction and the inclusion of time series models. As deduced from the results of this study, there is a significant enhancement of the accuracy of the prediction in the context of transformer-based sentiment models compared to the traditional models of NLP.

Technical Patterns and News Sentiment in Stock Markets” – Leippold, 2024

The interaction of these technical indicators and news sentiment in this research is analyzed in terms of its influence on the stock market. The effectiveness of news sentiment and technical pattern integration is proved in this research in terms of robustness.

Combining News Sentiment and Technical Analysis to Predict Trend Reversal” – Cagliero et al., 2019

This research proposes a hybrid approach that integrates news sentiment with technical indicators to identify trend reversals. The model demonstrates improved performance in capturing turning points in the market compared to standalone methods.

“Using Daily Stock Returns: The Case of Event Studies” – Brown & Warner, 1985

This basic study is an introduction to the method of Event Study, which is a method of stock price reaction to information. This is the basic foundation for the existence of abnormal returns, as they are received from external sources.

This method of Event Study measures the reaction of the financial market towards certain events, such as the announcement of earnings, mergers, policy actions, news releases, etc., by comparing the actual returns with the expected returns to determine whether the event affects the stock prices significantly. This method is used the most in financial studies to determine whether the financial market is efficient or not.

This method is used to determine the speed at which information is reflected in the stock prices.

“Event Studies in Economics and Finance” – MacKinlay, 1997

The objective of this paper is to build a framework through which the event in the financial market can be assessed. It will give support to the statistical techniques used in assessing the effect of the event on the stock price in a theoretical manner. It will give an explanation on how to determine the event window and estimation period in order to assess the effect of the event on the market. It will give an explanation on the importance of determining abnormal returns and cumulative abnormal returns in order to assess the effect of the event on the stock price. It will give an explanation on different assumptions on the efficiency of the market and different statistical models used in order to assess the effect of the event on the market.

“FinBERT: Financial Sentiment Analysis with Pre-trained Language Models” – Araci, 2019

This work has proposed the usage of the language model, and this language model has been referred to as FinBERT. It is evident from this work that the accuracy level of the sentiment classification is high compared to the accuracy level achieved by the usage of the generic NLP language model. Hence, the usage of the language model can be considered highly suitable.

III.EXISTING SYSTEM

Traditional Trading Approaches

The major source of data analysis in this traditional system of trading is manual analysis of stock price charts, etc. It has been found that the source of data analysis for the investors who are using this trading platform is different. It has been observed that the investors who are using this trading platform are facing difficulty in interpreting the data in a correct manner. It has been found that the decision-making process totally depends on the experience.

Price-Based Prediction Systems

The majority of the trading models that have been developed are focused on the importance of analyzing the historical data of the market using statistical and machine learning techniques. The trading models are based on various features, such as trends in prices and volatility of the data. Although trading models are efficient in handling numerical data, there are no features such as news and sentiment in the data.

Sentiment-Based Analysis Systems

With the improvement in technology in Natural Language Processing, various models have been developed for processing financial news data as well as social media data in order to perform sentiment extraction. The models classify financial news data into positive, negative, and neutral sentiments for sentiment analysis purposes. However, most sentiment-based models are developed independently without considering the market data.

Limitations of Existing Systems

Though these developments are done in both price-based and sentiment-based techniques, some of the limitations of these techniques are as follows:

- i. The absence of integration with sentiment-based techniques and price data
- ii. The inability to carry out real-time data processing
- iii. The low level of interpretability of these techniques for retail investors
- iv. The inability to carry out analysis on the association of news events and price movements

IV.PROPOSED SYSTEM

The proposed system is a decision support system that will operate with the support of aiding retail investors in making their decisions based on sentiment analysis related to finances and live stock market data. This is because the major component of the proposed system is going to be the translation from input data to output data through different processes.

System Workflow

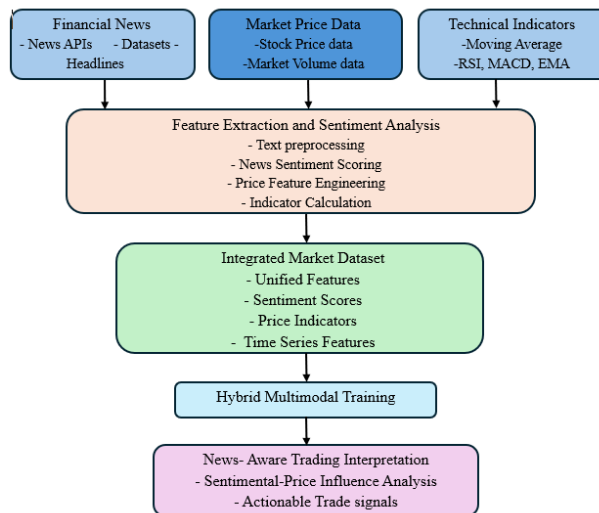


Fig-1: System Workflow

Market Data Integration

There are many types of data streams that can be utilized in the field of architecture, and some examples of these include the live market data stream and financial news data streams. The OHLC and volume data streams, as well as the financial news data streams, can be acquired through the use of APIs. These various data streams will be timestamped to ensure there is no synchronization problem between them. This will allow the analysis of both quantitative and qualitative information in the market.

Automated Analysis

In this case, the process of analyzing the data collected will be done automatically. As a result, the sentiment analysis will be carried out using natural language processing from the financial news. This data will then be translated into numerical information such as returns, volatility, trends, and others.

Market Predictions

The approach that is going to be adopted in making predictions will not only depend on the information gained through feature extraction but also the database generated at the same time. Prediction will be based on the knowledge gained about sentiments and market signals in order to make predictions about future price changes.

Confidence Scoring

However, for the prediction to be of any significance, it is imperative that there be some sort of certainty ratings associated with them, whereby the degree of certainty with regard to making the particular prediction using the signals is rated. There should also be an evaluation of the impact of all the signals combined.

Information Reduction

In the final stage, the focus is on reducing information overload and highlighting the key findings for the user. Instead of displaying all of the information and index numbers to the user, what is actually presented to the user is an organized output which can be achieved through their computer screens.

V.METHODOLOGY

This approach is primarily focused on the connection that exists between the sentiment expressed in financial news and the structured price data present in the market. There are many ways to achieve this objective, including data pre-processing, feature extraction, sentiment-price correlation, and prediction.

Data Preprocessing

Pre-processing happens on the financial data that have been collected from several sources before they are processed. Pre-processing of sentiments collected from news requires processes such as tokenization, removing stop words, and normalization of the data. For the market data, sorting happens according to different time frames, while any missing data are filled forward. Considering the similarities that sentiments and prices should exhibit in terms of temporal data processing, sentiments and price data will be aligned according to their respective timestamps.

Sentiment Feature Extraction Sentiment Feature Extraction

T Sentiment Analysis is carried out by analyzing the news from the financial market, where the sentiment in the market is calculated. The sentiment can be classified either as negative, neutral, or positive according to the analysis of the news.

Cumulative Sentiment score over a particular period can be calculated as:

$$S_t = \sum_{i=1}^N N s_i$$

where:

S_t = aggregated sentiment at time t

s_i = sentiment score of individual news item

N = number of news articles in the interval

Market Feature Engineering

The data on the market prices is then transformed into numerical data for the analysis of the trends within short periods of time. The logarithmic returns can be calculated using the following formula:

$$r_t = \ln(P_t - 1P_t)$$

where, P_t is the price at time t .

Other parameters that can be calculated using the data set include:

- i. Volatilities,
- ii. Moving averages,
- iii. Momentum. effectively lead users.

Sentiment–Price Interaction Modeling

Interaction variables for measuring the effect of news on market mechanisms are derived using sentiment and price variables. The following are interaction variables:

- i. sentiment and return interaction,
- ii. sentiment lag variable, and
- iii. sentiment and volatility interaction.

The sentiment lag variable can be measured by:

$St-k$

where k is the lag of market reaction.

Prediction Framework

These variables cover all possible factors, and the idea is to make market forecasts using them. The resulting forecast produced by the model shows the probability of the positive/negative trend on the market, as follows:

$$yt=f(St,rt,Xt)$$

yt= predicted market direction

St = sentiment features

rt= return features

Xt= additional technical indicators

This model provides the ability to build a linear and non-linear connection between the variables of sentiment and market.

Output Interpretation

Results obtained from this system will be presented in an organized manner using the following format:

Market Trend (Up/Down)

Probability

Confidence Interval

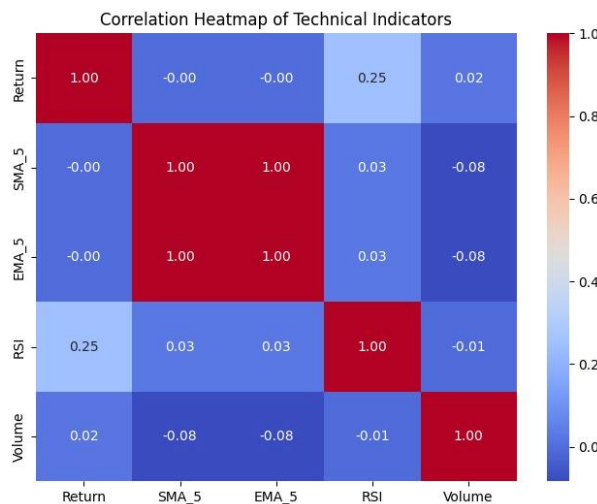


Fig-2: Correlation Heatmap

An organized presentation of the outputs ensures that traders can easily make decisions regardless of their knowledge of technicalities.

VI.RESULTS

The performance of this trading interpreter shall be assessed based on how he or she interprets the financial data in order to forecast using the online trading system. From all the things that have been discussed so far, it is clear that the evaluation of this trader shall revolve around the effectiveness of conducting price relationship sentiment modeling using the trading interpreter.

Result of the Implementation Process

Where the implementation process proves successful, it will be possible to acquire the ability of performing the analysis of information and market data using Flask. Some of the activities include live analysis of the market, data mining, and analysis of sentiments. The following results have been obtained from the use of the application.

- i. Live prices in the market
- ii. Sentiment scores
- iii. Probabilistic trends of the market
- iv. Decision making for trade

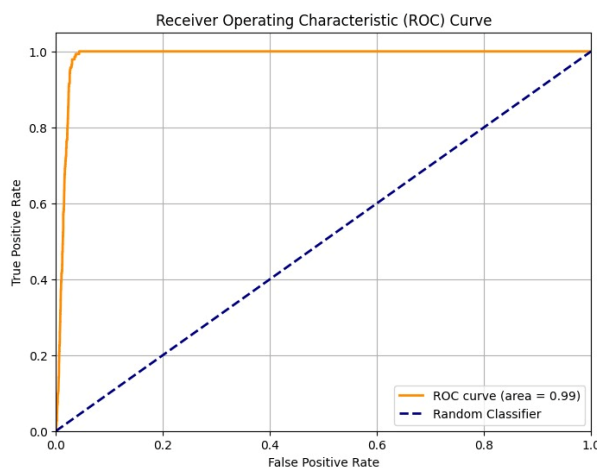


Fig-3: ROC Curve

Performance Analysis of Real-Time Data Processing System

Based on the analysis of the system described above, it becomes evident that the analyzed system looks extremely interesting thanks to its superb performance capabilities. In fact, it should be stated that they seem to be really great since there is no delay associated with processing real-time data since it is new because it comes from API data sources. At the same time, another point that is worth paying attention to is the following – processing real-time data takes less time compared to any other data.



Fig-4: User - Interface

Limitations

Even after the success of the model, some of the other limitations include:

It is not enough to conduct back-tests using only historical information.

There has not been any analysis of the model's performance.

Unrelated news may affect the model.

Some of the issues that should be addressed to enhance the performance of the model include those mentioned above.

VII. CONCLUSION

With regards to our contribution to research, we advise the use of the trading interpreter to gauge the influence of news on the psychological state of the trader and market data to get quick results. Therefore, when designing our trading strategy, we intend to design a new one which integrates quantitative and qualitative approaches to trading and analyzes the effects of informational events at various stages.

Additionally, it is worth noting that there is another important characteristic of our trading strategy – its effective execution within the Flask framework. Hence, as a result of our trading strategy, we can analyze real-time information and the performance of the market based on probabilities to generate interpretable results. Our trading strategy can also be interesting for retail traders because of its ease of application.

Overall, considering the stated characteristics of our trading strategy, we can state that our trading strategy is appropriate for designing an instrument that can aid retail traders in their decision-making process.

VIII. FUTURE WORK

Despite the illustration of the procedure above on the analysis process of sentiment analysis coupled with stock prices, there are numerous modifications that can be done to guarantee that the process becomes perfect.

The development of a machine equipped with machine learning algorithms will be the first step towards achieving this goal. The subsequent step will involve testing back to determine how accurate the algorithm is.

Apart from this, there are certain other aspects that can be taken into consideration as well, such as the analysis of sentiments via social media, macroeconomics, and global economics. Apart from adding other traits, it will also ensure efficiency.

A second approach for the same can be the Explainable Artificial Intelligence approach, through which the entire process can be understood better.

As far as the overall architecture of the system is concerned, certain modifications may need to be made.

REFERENCES

- [1] Kalyani, J., et al. (2016). "Stock Trend Prediction Using News Sentiment Analysis." (arXiv PDF) Shows how financial news sentiment can be used to predict stock trend.
- [2] Maqbool, J. (2023). "Stock Prediction by Integrating Sentiment Scores of Financial News and Historical Stock Price Data." (Procedia Computer Science, ScienceDirect) Directly supports your idea of mixing sentiment + price features.

- [3] Ruan, L. (2025). "Stock Price Prediction Using FinBERT-Enhanced Sentiment with Technical and Statistical Indicators." (MDPI Mathematics) Supports the exact approach: FinBERT + indicators → short-term prediction.
- [4] Financial sentiment analysis using FinBERT... (2025). (arXiv HTML) Shows combining FinBERT sentiment extraction and time series model improves performance
- [5] Leippold, M. (2024). "Technical patterns and news sentiment in stock markets." (ScienceDirect) Supports that technical patterns/indicators + sentiment provide predictive information.
- [6] Brown & Warner (1985). "Using Daily Stock Returns: The Case of Event Studies." (Journal of Financial Economics) This supports your logic: events/news cause abnormal returns and reaction around days.
- [7] MacKinlay, A.C. (1997). "Event Studies in Economics and Finance." (Review paper, PDF) Best reference to justify event impact approach academically.
- [8] Nifty 50 Price Forecasting with NLP Technique (2025). (IJISAE PDF) Direct NIFTY-based reference for your Indian market project.
- [9] Dogu Araci (2019) FinBERT: Financial Sentiment Analysis with Pre-trained Language Models (arXiv)
- [10] Jiang et al. (2023) Financial sentiment analysis using FinBERT... predicting market movement (arXiv PDF)
- [11] Hutto & Gilbert (2014) VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text (ICWSM)
- [12] Cagliero et al. (2019) Combining News Sentiment and Technical Analysis to predict trend reversal (SenticNet PDF)