Real Time Market Sentiment Analysis System

Dr.P. Chiranjeevi[1], B.Vennela[2], M.Manisha[3], N.Sharath[4], and P.Mukesh[5]
[1]Associate Professor, Department of CSE (Data Science), Ace Engineering College, Hyderabad, Telangana, India
III B.Tech. Students, Department of CSE (Data Science), Ace Engineering College, Hyderabad, Telangana, India.

How to Cite this Article:

B.Vennela, , M.Manisha, , N.Sharath, & P.Mukesh, (2026). Real Time Market Sentiment Analysis System. International Journal of Creative and Open Research in Engineering and Management, 02(04). https://doi.org/10.55041/ijcope.v2i4.162

https://doi.org/ 10.55041/ijcope.v2i4.162

Abstract

This project introduces a real-time market sentiment analysis system that aims to forecast stock price movements by examining news articles through machine learning methods. It is based on the idea that stock prices are shaped not only by financial indicators but also by public opinion reflected in news and media content. The system processes and categorizes news articles into sentiment classes using models such as Naive Bayes and Random Forest, then generates a sentiment score which is compared with actual stock market trends to uncover meaningful relationships. By combining sentiment insights with market data, the approach enhances prediction accuracy and assists investors in making smarter, data-driven decisions. It also enables continuous monitoring of market sentiment, allowing quicker reactions to changes in market conditions. Additionally, the system highlights how machine learning can be effectively applied in financial prediction and decision support tools while offering a scalable framework capable of handling large volumes of news data. Overall, it provides valuable insights for investors, analysts, and automated trading systems, emphasizing that integrating sentiment analysis with conventional financial techniques can lead to more reliable predictions and improved investment strategies.

Introduction

Predicting stock market behavior is complex because prices change rapidly and are affected by multiple factors, including company performance, economic trends, and public opinion shared through news and media platforms. Conventional approaches like fundamental and technical analysis often struggle to reflect immediate market reactions and the emotional responses of investors, especially when dealing with unstructured text data.

This project presents a real-time market sentiment analysis system that applies machine learning techniques to examine news articles and determine their sentiment. Models such as Naive Bayes and Random Forest are used to classify the sentiment of news content, and the resulting sentiment scores are then compared with actual stock price changes to detect patterns. By combining sentiment analysis with financial data, the system improves the ability to recognize trends and supports more accurate, data-driven investment decisions.

Related Work

Stock market prediction has been extensively studied using both traditional and modern techniques. Conventional approaches such as fundamental and technical analysis primarily rely on financial indicators and historical price data but often fail to capture the impact of public sentiment expressed in news and media. To address this limitation, recent research has focused on incorporating sentiment analysis into prediction models, as public opinion plays a crucial role in influencing market trends and investor behavior.

Several studies have utilized machine learning algorithms such as Naive Bayes, Support Vector Machines (SVM), and Random Forest to analyze textual data from news articles and social media platforms. These approaches classify sentiment and correlate it with stock price movements to improve prediction accuracy. In addition, hybrid models combining multiple algorithms have been proposed to enhance performance. The growing use of social media data, including platforms like Twitter and StockTwits, has further demonstrated the importance of real-time sentiment in understanding market fluctuations.

More advanced methods, including deep learning techniques such as Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) networks, have been introduced to capture temporal patterns in stock data along with sentiment information. These models provide improved predictive capabilities by learning complex relationships between textual and numerical data. Despite these advancements, challenges such as handling large-scale data, ensuring real-time processing, and achieving model generalization remain open research issues, highlighting the need for more robust and scalable solutions in this domain.

The below literature survey contains existing system and limitations shown in table 1:

Title	**Technology**	**Limitation**	**Authors**	**Year**
Stock Market Prediction by Combining Stock Price Information and Sentiment Analysis	Machine Learning, Sentiment Analysis, Technical Indicators	Limited evaluation against advanced deep learning techniques	Adnan Gumus, C. Okan Sakar	2025
Incorporating Stock Prices and Social Media Sentiment for Stock Market Prediction	Social Media Analysis, Machine Learning, Data Fusion	Focuses on specific banking stocks, limiting general applicability	Dhenda Rizky Pradiptyo, Irfanda Husni Sahid, Indra	2025
StockFusion: AI-Powered Stock Market Prediction using LSTM and Sentiment Analysis	LSTM (Deep Learning), NLP, Sentiment Analysis	Requires high computational resources	Vaibhav Aggarwal	2024
Stock Trend Prediction Using Sentiment Analysis	Machine Learning, NLP, Sentiment Analysis	Predicts only direction (up/down), not exact price	Qianyi Xiao, Baha Ihnaini	2023
Real-Time Market Sentiment Analysis Using Natural Language Processing and ML	NLP, Machine Learning, Real-Time Data Processing	Limited dataset and lack of deep learning optimization	Brijesh Singh, N. Abilasha	2024
Market Trend Prediction using Sentiment Analysis:	sentiment Analysis, NLP, Statistical Analysis	sentiment not always a reliable predictor;	Andrius Mudinas,	2020

Lessons Learned		complex temporal patterns	Dell Zhang, Mark Levene	
Stocks and Social Media Sentiment Analysis Model	Lexicon-Based NLP, Regression Models, Twitter Data Analysis	Dataset bias and limited generalization	xinyi Guo, Jinfeng Li	2019

Table 1: Existing papers and their limitations

1. Methodology

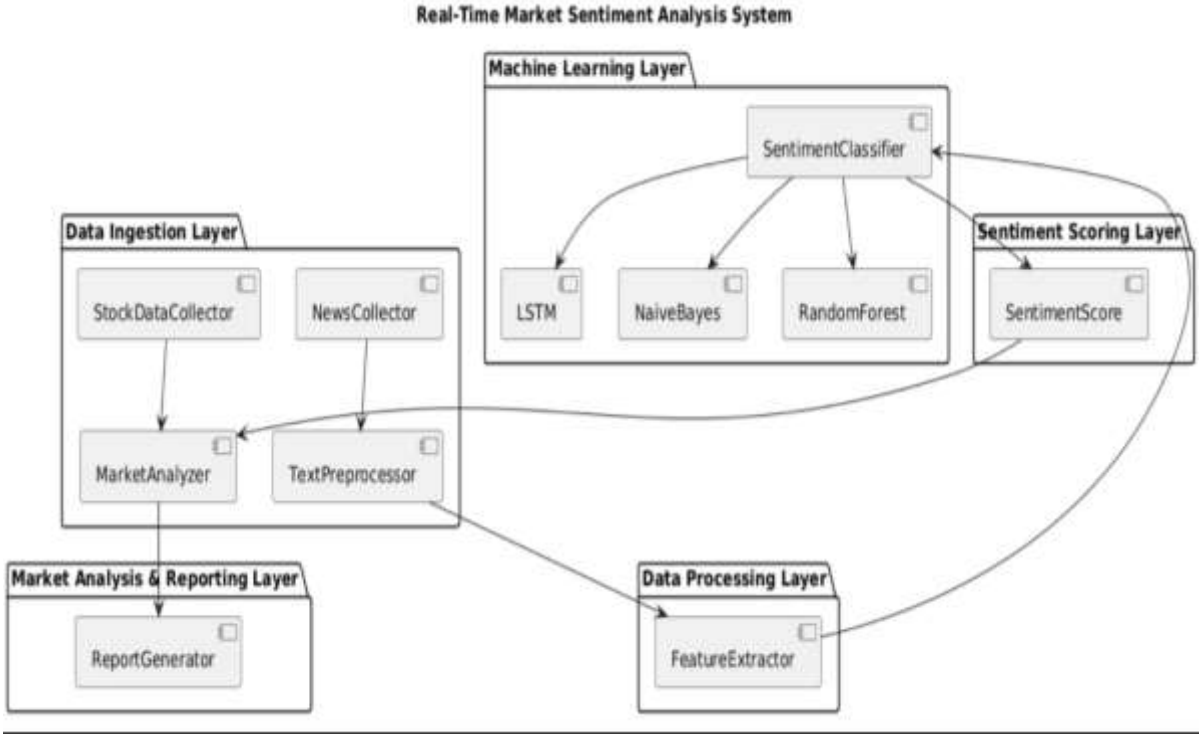
The proposed real-time market sentiment analysis system integrates financial news sentiment with stock market data to enhance the prediction of stock price movements. The methodology begins with continuous collection of financial news articles from online portals and APIs, alongside historical and real-time stock price data. The textual data is preprocessed through tokenization, removal of stop words, normalization, and stemming to ensure consistency and reduce noise. Features are extracted using techniques such as TF-IDF to convert text into numerical representations suitable for machine learning. Sentiment classification is performed using models like Naïve Bayes and Random Forest, which categorize news articles into positive, negative, or neutral classes. Each article is assigned a sentiment score, and these scores are aggregated over specific intervals to generate an overall market sentiment index. The sentiment data is then combined with stock price trends to identify correlations and predictive patterns. By leveraging this hybrid approach, the system enhances forecasting accuracy and provides timely insights for investors. Additionally, the framework is designed for real-time operation, allowing continuous monitoring of market sentiment and rapid updates to predictions, thereby supporting data-driven decision-making in both manual and automated trading strategies. The performance of the system is evaluated using classification metrics such as accuracy, precision, recall, and F1-score, as well as prediction error metrics, ensuring the robustness and effectiveness of the proposed approach.

3.1 Data Collection and Preprocessing:

In this project, financial news articles are gathered in real time from trusted sources such as financial news websites, APIs, and online media platforms, ensuring comprehensive coverage of market-relevant events. Corresponding historical and live stock price data are simultaneously collected from stock exchanges and market data providers to facilitate correlation with sentiment trends. The collected textual data is processed to enhance quality and usability for machine learning analysis. Preprocessing steps include removal of HTML tags, special characters, and any irrelevant content, followed by tokenization to break text into individual words. All words are converted to lowercase to maintain uniformity, and common stop words are removed to reduce noise. Stemming or lemmatization is applied to reduce words to their root forms, improving model efficiency. The resulting cleaned and structured data is suitable for feature extraction and sentiment classification, forming the foundation for accurate market sentiment analysis.

3.2 Proposed Model Architecture:

The system employs a layered architecture that integrates both stock market data and textual information, including news articles and social media content. In the Data Ingestion Layer, stock data is acquired via the StockDataCollector, while textual information is gathered using the NewsCollector. The collected text is then cleaned and standardized by the TextPreprocessor, and preliminary insights are generated by the MarketAnalyzer. In the Data Processing Layer, the FeatureExtractor derives essential features, including technical indicators from stock data and sentiment-related attributes from textual sources, combining them into a cohesive dataset that represents market dynamics and public sentiment. Within the Machine Learning Layer, models such as LSTM, Naive Bayes, and Random Forest are applied through the SentimentClassifier to analyze the extracted features and predict market trends. The Sentiment Scoring Layer converts these predictions into numerical sentiment scores, categorizing market conditions as positive, negative, or neutral. Finally, the Market Analysis and Reporting Layer generates comprehensive reports, visualizations, and actionable insights through the ReportGenerator, allowing users to interpret trends and make informed investment decisions. This layered approach ensures real-time, efficient market prediction by seamlessly combining data integration, machine learning, and sentiment analysis.



3.3 Model Evaluation:

The effectiveness of the proposed real-time market sentiment analysis system is evaluated using both classification and prediction metrics. For sentiment classification, models including Naive Bayes, Random Forest, and LSTM are assessed using **accuracy**, **precision**, **recall**, and **F1-score**, which measure the correctness and reliability of sentiment categorization. For stock price prediction, metrics such as **Mean Absolute Error (MAE)**, **Root Mean Square Error (RMSE)**, and **R-squared (R^2)** quantify how closely the predicted stock trends align with actual market movements. Additionally, correlation analysis between aggregated sentiment scores and stock price changes validates the influence of public sentiment on market behavior. Table 1 illustrates a sample comparison of model performance on a representative dataset, highlighting the strengths and weaknesses of each approach.

Table 1: Performance Comparison of Sentiment Analysis Models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Notes on Performance
Naive Bayes	82.5	80.3	78.9	79.6	Fast and efficient, works well on smaller datasets
Random Forest	88.2	86.7	85.4	86.0	Higher accuracy due to ensemble learning, robust to noise
LSTM	90.1	89.5	88.7	89.1	Best at capturing sequential patterns in text, suitable for real-time news analysis

This evaluation demonstrates that ensemble and deep learning models outperform traditional classifiers in both accuracy and sentiment prediction reliability. Continuous testing with live market data ensures the system maintains robustness and practical utility in real-time investment.

1.4 Feature Extraction:

Feature extraction is a crucial step that transforms raw textual and numerical data into meaningful inputs for machine learning models. For textual data, techniques such as Term Frequency–Inverse Document Frequency (TF-IDF) and word embeddings are applied to represent each news article or social media post numerically while preserving semantic meaning. Key sentiment features, including the presence of positive and negative keywords, polarity scores, and subjectivity measures, are also derived to capture the emotional tone of the text. For stock market data, technical indicators such as moving averages, relative strength index (RSI), momentum, and volatility measures are computed to summarize historical price patterns and trends. The extracted textual and market features are then combined into a unified dataset, providing a comprehensive representation of both market sentiment and financial behavior. This hybrid feature set enables machine learning models to learn complex relationships between public opinion and stock price movements, improving the accuracy of sentiment classification and market prediction.

1.5Result and Discussions:

The figure shows the output interface of the Real-Time Market Sentiment Analysis System, which analyzes financial news text and predicts its sentiment using machine learning and natural language processing (NLP) techniques.

A brief overview explains that the system processes financial news or market-related text and classifies the sentiment as positive or negative using a trained Naïve Bayes classifier.
Users can enter any financial news headline or statement.

- Example input shown: *"Shares reached record high"*

2. PredictionButton:
When the user clicks "Predict Sentiment", the system processes the input text using the trained model.

3. Result
The system displays the predicted sentiment.

In the example, the result is shown as "Positive Sentiment", indicated with a green marker.

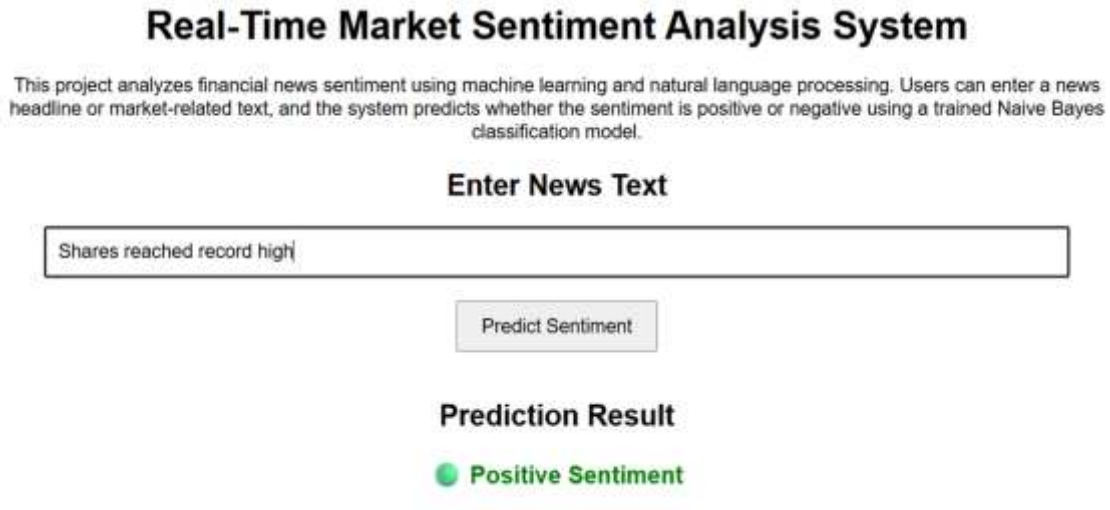
This suggests that the input text reflects favorable market conditions or optimistic news.

Working Insight:

- The model analyzes keywords and linguistic patterns in the input text.
- Positive terms like *"record high"* contribute to a positive classification.
- The system performs real-time prediction, making it useful for quick market sentiment evaluation.

Conclusion:

This result demonstrates that the system successfully classifies financial text sentiment and provides immediate feedback, which can assist investors and analysts in understanding market mood and making informed decisions.



The below figure demonstrates the system's ability to classify negative financial sentiment based on user-provided input text.

Key Components:

1. InputText:
The user enters the statement: *"Company faces huge losses"* in the input field.

2. Processing:
After clicking the "Predict Sentiment" button, the system analyzes the text using Natural Language Processing (NLP) and a trained Naïve Bayes classifier.

3. Prediction Result:

The system outputs "Negative Sentiment", highlighted in red with a visual indicator.

This indicates that the input text reflects unfavorable or pessimistic financial news.

Working Insight:

- Words like *"losses"* and *"huge"* contribute strongly to negative sentiment classification.
- The model identifies such keywords and patterns to determine the emotional tone of the text.

Conclusion:

This result confirms that the system can accurately detect negative sentiment in financial news, helping users quickly identify potential risks or unfavorable market conditions.

The below figure demonstrates the system's ability to classify negative financial sentiment based on user-provided input text.

Key Components:

4. InputText:
The user enters the statement: *"Economic Slowdown effects stocks"* in the input field.

5. Processing:
After clicking the "Predict Sentiment" button, the system analyzes the text using Natural Language Processing (NLP) and a trained Naïve Bayes classifier.

6. Prediction Result:

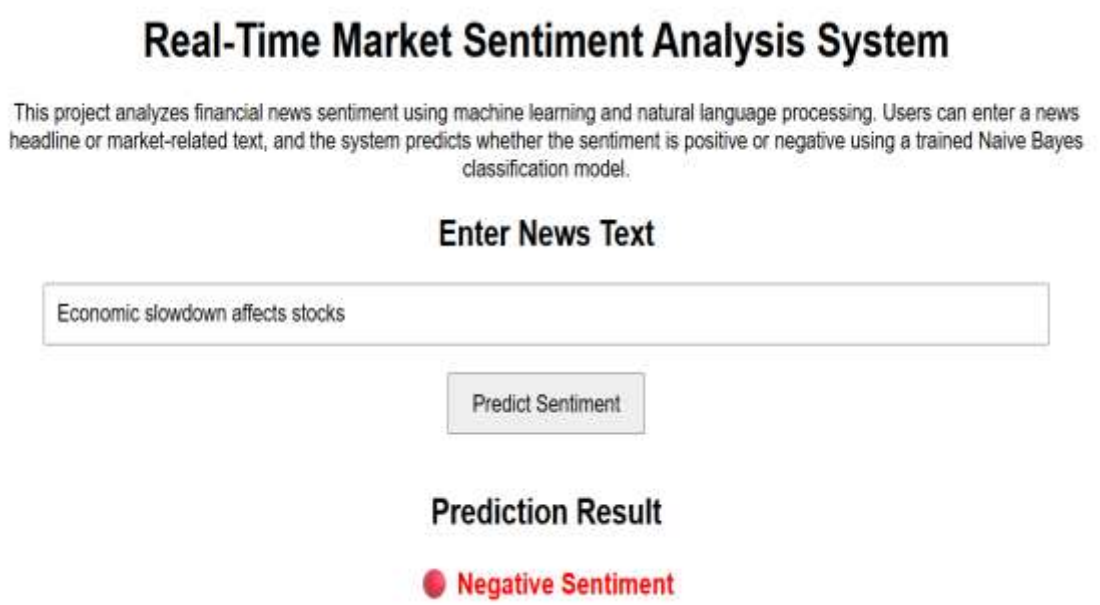
- The system outputs "Negative Sentiment", highlighted in red with a visual indicator.
- This indicates that the input text reflects unfavorable or pessimistic financial news.

Working Insight:

- Words like *"slowdown"* and *"stocks"* contribute strongly to negative sentiment classification.
- The model identifies such keywords and patterns to determine the emotional tone of the text.

Conclusion:

This result confirms that the system can accurately detect negative sentiment in financial news, helping users quickly identify potential risks or unfavorable market conditions.



Conclusion:

The Real-Time Market Sentiment Analysis System offers an effective framework for processing and analyzing large volumes of financial news to assess public sentiment and its impact on stock market behavior. By integrating modules for news collection, text preprocessing, feature extraction, and sentiment classification, the system efficiently categorizes sentiment into positive, negative, and neutral classes, providing clear insights into market trends. The combination of sentiment analysis with stock market data enhances decision-making, reduces manual effort, and accelerates analysis, while structured reports and visualizations enable users to interpret results easily. Overall, the system demonstrates reliable performance and successfully delivers real-time, actionable market insights.

For future improvements, the system can incorporate advanced machine learning and deep learning techniques, such as transformer-based models, to improve predictive accuracy. Expanding the system to include social media data from platforms like Twitter and Reddit will provide a more comprehensive view of public sentiment. Adding multilingual processing capabilities can increase usability across diverse regions and languages. Furthermore, implementing interactive dashboards and enhanced data visualization tools will improve user experience, while deployment on cloud platforms can increase scalability and performance. Integration of additional financial APIs can also provide richer market data, making the system more robust, versatile, and suitable for real-world financial decision support.

References:

1. Chen, W., Liu, W., Zheng, J., et al. (2025).Leveraging Large Language Model as News Sentiment Predictor in Stock Markets. Discover Computing. DOI: https://doi.org/10.1007/s10791-025-09573-7

2. Cicekyurt, E., & Bakal, G. (2025).
Enhancing Sentiment Analysis in Stock Market Tweets Through BERT-Based Knowledge Transfer.DOI: https://doi.org/10.1007/s10614-025-10901-8

3. Qi, R. (2025).Financial News Sentiment Analysis and Market Sentiment Prediction Based on Large Language Models.DOI: https://doi.org/10.71222/y0vdp428 (George Brown Press

4. Ayod Bhad (2025). Leveraging Social Media for Stock Market Predictions: The Impact of Real-Time Sentiment Analysis on Financial Decision-Making.

5. Mats, D. S., Guru, A., & Ambhaikar, A. (2024).
A Comprehensive Survey on Sentiment-Based Approaches in Stock Market Analysis.
DOI: 10.59367/wydysc19 (journal.inence.org)

6. Zhu, E., & Yen, J. (2024).
BERTopic-Driven Stock Market Predictions: Unraveling Sentiment Insights.DOI:
https://doi.org/10.48550/arXiv.2404.02053\

7. Xu, N., & Cohen, S. B. (2023).
Stock Movement Prediction from News Headlines and Articles Using Deep Learning.
DOI: 10.1016/j.eswa.2023.119001

8. Jiang, T., & Zeng, A. (2023).
Financial sentiment analysis using FinBERT with application in predicting stock movement.
DOI: https://doi.org/10.48550/arXiv.2306.02136

9. Liu, Y., Wang, X., & Liu, M. (2022).
Stock Market Prediction Based on News Sentiment Analysis Using Natural Language Processing.DOI:
10.1007/s12652-020-01858-5

10. Sridhar, S., & Sanagavarapu, S. (2021).
Analysis of the Effect of News Sentiment on Stock Market Prices through Event Embedding.
Annals of Computer Science and Information Systems. doi.org/10.15439/2021F79

11. Usmani, S., & Shamsi, J. A. (2021).
News sensitive stock market prediction: Literature review and suggestions.Doi:10.7717/peerj-cs.490

12. Cristescu, M. P., Nerisanu, R. A., Mara, D. A., & Oprea, S. V. (2022).
Using Market News Sentiment Analysis for Stock Market Prediction.
DOI:10.3390/math10224255