International Research Journal of Innovations in Engineering and Technology (IRJIET)
ISSN (online): 2581-3048
Volume 9, Special Issue ICCIS-2025, pp 102-106, May-2025
https://doi.org/10.47001/IRJIET/2025.ICCIS-202516
Second International Conference on Computing and Intelligent Systems (ICCIS-2025)

Investor Sentiment-Driven Stock Price Prediction Using Optimized Deep Learning

[1]B. Rupadevi, [2]Shaik Nadheem

[1]Associate Professor, Dept. of MCA, Annamacharya Institute of Technology & Sciences, Tirupati, AP, India
[2]Post Graduate, Dept. of MCA, Annamacharya Institute of Technology & Sciences, Tirupati, AP, India
E-mails: [1]rupadevi.aitt@annamacharyagroup.org, [2]nadheembashashaik747@gmail.com

***Abstract* - Stock price prediction plays a vital role in risk evaluation, investment decision-making, and financial planning. Traditional forecasting models primarily rely on technical indicators and historical price trends; however, they often fail to account for the emotional and behavioral dynamics influencing market movements. This research introduces a hybrid model that integrates investor sentiment—extracted from online platforms such as social media and financial news—with historical market data. By merging quantitative trends with psychological insights, the proposed approach enhances prediction accuracy and adaptability. The system utilizes natural language processing techniques and a deep learning framework to analyze textual content and generate sentiment scores. Empirical results using real-world datasets demonstrate that sentiment-augmented models outperform conventional approaches in forecasting short-term trends and detecting market volatility. The paper also explores the practical implications for real-time prediction systems and addresses related ethical considerations.**

Keywords: Equity Forecasting, Market Sentiment Analysis, Trend Estimation, Long Short-Term Memory Networks, Financial Data Sequences, Predictive Analytics.

I. INTRODUCTION

Stock market forecasting remains one of the most complex and widely explored topics in the realm of financial research. Accurate predictions of stock prices carry significant importance for investors, traders, analysts, and policymakers alike. A variety of factors—including historical trends, macroeconomic indicators, corporate performance, global developments, and investor psychology—can influence stock valuations. Conventional forecasting models typically rely on technical indicators and fundamental analysis based on price history and financial statements. However, these models often neglect the psychological and behavioral aspects of market participants, which can significantly drive price fluctuations.

Investor sentiment—a critical force in financial markets—reflects the collective mood and expectations of market participants. With the increasing prominence of digital communication platforms, data from social media and financial news outlets has become a rich resource for capturing public opinion. Shifts in sentiment often precede or drive market behaviour, with sudden emotional swings sometimes leading to unexpected price movements. For instance, even a strong quarterly earnings report might not lead to a stock price increase if investors are broadly pessimistic due to wider economic concerns. Conversely, social media-driven excitement can cause stock prices to rise despite poor underlying fundamentals.

With recent advancements in artificial intelligence and machine learning, researchers have begun to integrate sentiment analysis into stock prediction models. By examining textual data from financial news, social platforms, analyst reviews, and corporate disclosures, sentiment analysis can quantify market mood. When combined with traditional price data, these sentiment-derived insights allow machine learning models to capture both historical trends and the emotional state of the market, enhancing their predictive capabilities.

An emerging area of study focuses on how emotions, beliefs, and reactions expressed through online platforms, financial articles, and discussion forums influence stock behaviour. These sentiments, often triggered by news events, rumors, or earnings results, can result in swift price movements that traditional models may fail to anticipate. Additionally, the presence of algorithmic trading systems that respond to live information further magnifies the impact of market sentiment.

As data availability and computational power continue to expand, machine learning and deep learning approaches have become vital for recognizing complex, non-linear patterns in market behavior. This paper proposes a holistic model that integrates sentiment data with historical stock records to improve the accuracy and responsiveness of price predictions.

International Research Journal of Innovations in Engineering and Technology (IRJIET)
ISSN (online): 2581-3048
Volume 9, Special Issue ICCIS-2025, pp 102-106, May-2025
https://doi.org/10.47001/IRJIET/2025.ICCIS-202516
Second International Conference on Computing and Intelligent Systems (ICCIS-2025)

II. LITERATURE SURVEY

Due to its pivotal role in guiding financial strategies, stock price prediction has long been a focus of extensive academic research. Early forecasting methods predominantly employed statistical tools to examine historical market behaviors. In recent years, however, increasing attention has been given to the influence of investor sentiment and financial news, which significantly shape market dynamics. This section reviews existing literature on stock forecasting techniques, highlighting their strengths and limitations.

A. Conventional Statistical Approaches

Initial studies on predicting stock values primarily utilized econometric and statistical techniques, relying on the identification of historical data trends. These models proved effective in recognizing short-term price fluctuations and cyclical behavior. However, they often fell short when dealing with the complexity and unpredictability of real-world markets. Traditional techniques struggled to account for abrupt disruptions caused by external factors such as political events or investor mood swings. As a result, their ability to project long-term market movements was limited, especially when sentiment-driven shifts were ignored.

B. Data-Centric Predictive Models

The emergence of data-driven methodologies, particularly machine learning, enhanced forecasting capabilities by enabling systems to learn directly from data rather than fixed algorithms. Models like decision trees and support vector machines offered greater adaptability and predictive power. Nevertheless, these techniques faced challenges in processing unstructured information—such as financial text or social media content—and were heavily reliant on manual feature selection and engineering. Their effectiveness was therefore constrained by the quality of inputs and domain-specific expertise.

C. Machine Learning Techniques

With the growth of artificial intelligence, the financial sector increasingly adopted machine learning models for stock prediction. Algorithms including ensemble methods like Random Forests and Gradient Boosting Machines demonstrated notable improvements over traditional linear models, particularly in capturing complex market patterns. Yet, models such as Support Vector Machines, while efficient for classification lacked the capability to model the sequential nature of time-series data. Furthermore, many of these systems required extensive data preprocessing and domain knowledge to fine-tune input features, making them labor-intensive to implement.

D. Deep Learning for Financial Forecasting

Recently, deep learning techniques have gained prominence due to their capacity to automatically discover intricate patterns within raw financial datasets. These models are especially adept at managing sequential data across extended timeframes, making them well-suited for time series analysis. Conventional models often faltered in preserving relevant information across long durations, but newer neural architectures addressed these challenges more effectively. These advanced models have shown superior performance in predicting stock movements compared to earlier approaches. Moreover, lightweight variants of deep learning models have been developed for real-time or large-scale applications, offering similar accuracy with reduced computational demand.

E. Role of Sentiment Analysis in Stock Forecasting

Investor sentiment is increasingly recognized as a crucial factor in analyzing financial markets. Studies have shown that opinions shared through media outlets, expert analyses, and social platforms can significantly influence stock prices. Natural Language Processing (NLP) techniques are commonly employed to interpret the emotional context and public attitudes embedded in text-based data. Incorporating sentiment information into prediction models has proven to enhance forecasting performance by capturing behavioral patterns that traditional methods often overlook. By leveraging real-time sentiment indicators from online news sources and social discussions, predictive models gain a more nuanced understanding of market dynamics, enabling more accurate projections of investor behavior and price shifts.

F. Identified Research Gap and Study Contribution

Despite extensive research into various stock forecasting techniques, a standardized approach to effectively combining advanced prediction models with sentiment data remains underdeveloped. Many existing models emphasize historical pricing trends while under representing the emotional and cognitive influences that drive market decisions. While sentiment analysis has been examined independently in financial research, its integration into unified forecasting systems is still limited. This study aims to address this deficiency by presenting a holistic model that merges traditional financial indicators with real-time sentiment insights. By fusing quantitative market data with qualitative behavioral cues, the proposed approach offers a more robust and accurate method for predicting stock market behavior.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 9, Special Issue ICCIS-2025, pp 102-106, May-2025

https://doi.org/10.47001/IRJIET/2025.ICCIS-202516

Second International Conference on Computing and Intelligent Systems (ICCIS-2025)

III. METHODOLOGY

A. Data Acquisition

This research utilizes a composite dataset that merges sentiment-based textual data with financial time series information. The preprocessing phase involved handling incomplete entries, unifying date formats, and refining textual content through regular expressions and natural language processing methods. The financial dataset, sourced from publicly available repositories, included daily metrics such as opening and closing prices, intraday highs and lows, trading volume, and adjusted close prices for selected companies. In parallel, sentiment scores were derived from textual analysis of investor commentary and financial news articles. To ensure consistency and usability, categorical variables were properly encoded, while missing data points were addressed using forward-filling and interpolation strategies. Additionally, all date-related attributes were standardized into a unified date time format to align records accurately across datasets during the feature integration process.

Table 1: stock_yfinance_data.csv

	Date	Open	High	Low	Close	Adj Close	Volume	Stock Nam
2	30-09-2021	260.333343	263.043334	258.333343	258.493347	258.493347	53868000	TSLA
3	01-10-2021	259.466674	260.260009	254.529998	258.406677	258.406677	51094200	TSLA
4	04-10-2021	265.5	268.989990	258.706665	260.510009	260.510009	91449900	TSLA
5	05-10-2021	261.600006	265.769989	258.066680	260.196655	260.196655	55297800	TSLA
6	06-10-2021	258.733337	262.220001	257.739990	260.916656	260.916656	43898400	TSLA
7	07-10-2021	261.820007	268.333343	261.126678	264.536682	264.536682	57587400	TSLA
8	08-10-2021	265.40332	265.459991	260.303344	261.829986	261.829986	50215800	TSLA
9	11-10-2021	262.549987	267.079986	261.833343	263.980010	263.980010	42600900	TSLA
10	12-10-2021	266.976654	270.773345	265.523345	268.573333	268.573333	66060000	TSLA
11	13-10-2021	270.156677	271.803344	268.593322	270.359985	270.359985	42360300	TSLA
12	14-10-2021	271.829986	273.416656	271.116668	272.773345	272.773345	36741600	TSLA
13	15-10-2021	274.579986	281.070007	274.116668	281.010009	281.010009	56773800	TSLA
14	18-10-2021	283.929992	291.753326	283.823333	290.036682	290.036682	72621600	TSLA
15	19-10-2021	292.510009	292.649993	287.503326	288.089996	288.089996	52143300	TSLA
16	20-10-2021	288.450012	289.829986	285.793334	288.600006	288.600006	42096300	TSLA
17	21-10-2021	285.333343	300	285.166656	298	298	94444500	TSLA
18	22-10-2021	298.5	303.333343	296.986663	303.226654	303.226654	68642400	TSLA
19	25-10-2021	316.843322	348.339996	314.733337	341.619995	341.619995	188556300	TSLA
20	26-10-2021	341.563323	364.980010	333.813323	339.476654	339.476654	187245000	TSLA
21	27-10-2021	346.553344	356.959991	343.593322	345.953338	345.953338	115579500	TSLA
22	28-10-2021	356.103332	360.333343	351.399993	359.013336	359.013336	81639600	TSLA
23	29-10-2021	360.619995	371.736663	357.736663	371.333343	371.333343	89755200	TSLA
24	01-11-2021	381.666656	403.25	372.886657	402.863342	402.863342	168146100	TSLA

Table 2: Stock_tweets.csv

	Date	Tweet	Stock Name	Company Name
2	2022-09-29	Mainstrean	TSLA	Tesla, Inc.
3	2022-09-29	Tesla delive	TSLA	Tesla, Inc.
4	2022-09-29	3/ Even if I	TSLA	Tesla, Inc.
5	2022-09-29	@RealDanC	TSLA	Tesla, Inc.
6	2022-09-29	@RealDanC	TSLA	Tesla, Inc.
7	2022-09-29	@RealDanC	TSLA	Tesla, Inc.
8	2022-09-29	For years @	TSLA	Tesla, Inc.
9	2022-09-29	$NIO just b	TSLA	Tesla, Inc.
10	2022-09-29	50 likes for some $SPY $TSLA charts to study!	TSLA	Tesla, Inc.

- Date: The date and time when the tweet was posted.
- Tweet: The content of the tweet that references the TSLA stock.
- Stock Symbol: The ticker symbol related to the stock.
- Company Name: The full official name of the business (Tesla, Inc.).

B. Sentiment Analysis

Text data was first tokenized using the Natural Language Toolkit (NLTK), followed by the elimination of stopwords, punctuation, and the application of lemmatization. Sentiment polarity scores were derived using pre-trained sentiment analysis tools such as VADER. These scores were normalized and then integrated into the main dataset as numerical features. By reflecting investor sentiment and outlook, these metrics offered supplementary insights beyond standard technical indicators.

This additional layer allowed the model to detect emotional and behavioral patterns that often precede market fluctuations, especially in response to major announcements like earnings reports or geopolitical events.

C. Feature Engineering and Selection

To boost the model's predictive performance, a variety of preprocessing strategies were applied. Feature normalization and standardization were used to ensure consistency across variables. Dimensionality reduction was performed through principal component techniques, retaining critical information while streamlining the input space. Statistical methods were employed for feature selection to identify the most relevant predictors. The dataset was further enriched with calculated technical indicators, such as moving averages, volatility measures, and momentum-based metrics. When combined with sentiment-derived features from text analysis, this comprehensive dataset captured both market trends and investor behavior, reducing the risk of overfitting and improving model interpretability.

D. Model Development and Assessment

Multiple predictive models were constructed and analyzed to identify the most effective approach for forecasting stock prices. These models were built within a deep learning framework, with particular attention given to preserving the sequential integrity of time-series data. To avoid look-ahead bias, time-series cross-validation was implemented, ensuring that the training and test data were properly ordered. The models were evaluated using common regression metrics including the R^2 score, mean absolute error (MAE), and root mean squared error (RMSE). These metrics provided a thorough assessment of model accuracy, prediction

International Research Journal of Innovations in Engineering and Technology (IRJIET)
ISSN (online): 2581-3048
Volume 9, Special Issue ICCIS-2025, pp 102-106, May-2025
https://doi.org/10.47001/IRJIET/2025.ICCIS-202516
Second International Conference on Computing and Intelligent Systems (ICCIS-2025)

error, and explanatory strength in capturing future price movements.

```
Model Performance Comparison:
           LSTM    XGBoost     Hybrid
RMSE  18.190353  42.596815  26.980990
MAE    8.017213  21.172971  13.356727
R²     0.963595   0.800369   0.919908
MAPE        inf        inf        inf

Best performing model based on RMSE: LSTM
```

Figure 1: Model Performance

IV. IMPLEMENTATION AND OUTCOMES

A. System Architecture Summary

The proposed stock prediction platform integrates both investor sentiment and historical price data to enhance forecasting precision. Its architecture follows a structured pipeline that includes data acquisition, preprocessing, model development, validation, and deployment through a web interface. The frontend, built with standard web technologies, offers an intuitive and responsive user experience, while the backend manages data processing and machine learning operations. The system begins by sourcing financial data from credible platforms and extracting sentiment from online content such as news articles and social media posts. This data is then cleaned and formatted to ensure reliability and relevance. Once processed, it is input into a predictive engine that combines machine learning techniques with time-series analysis. Through the web interface, users can receive live stock predictions along with contextual insights.

B. Online Application Interface

The web-based interface allows users to input stock symbols and obtain predictions enriched with sentiment insights. With a minimalist and accessible layout, the application enables users to enter a desired ticker symbol into an input form. Upon submission, the backend gathers up-to-date historical stock data and corresponding sentiment metrics, processes them through the trained model, and displays the forecasted value along with key indicators. The output includes visual charts showing stock trends, highlighting both historical movement and projected performance. Sentiment metrics are incorporated to help users grasp public perception and market mood surrounding the selected stock, thereby supporting more informed decision-making.

C. Predictive Model Strategy

Selecting an effective model was critical for delivering accurate stock forecasts. Traditional analytical methods were initially explored but proved insufficient for capturing the temporal dynamics typical of financial markets. To address this limitation, the study explored advanced approaches capable of recognizing sequential patterns in price movement and market sentiment. The forecasting system incorporated sentiment derived from financial media and social platforms, in addition to historical stock prices. Sentiment data was quantified through established scoring algorithms and fused with structured financial inputs, resulting in a more holistic analysis. This hybrid approach significantly outperformed traditional models, especially during volatile market periods. Model performance was evaluated using metrics such as MAE, RMSE, and R^2, confirming the model's improved precision and robustness.

D. Forecasting Results and Visual Representation

Once trained, the model generated future stock price forecasts by leveraging both historical trends and sentiment indicators. Compared to standard forecasting models, the integrated approach demonstrated superior responsiveness to market shifts. Accuracy was validated using statistical metrics including MAE, RMSE, and the R^2 score. Visualizations further illustrated the model's effectiveness: one chart displayed a side-by-side comparison of actual versus predicted stock prices, showcasing alignment and deviations. Another visual linked sentiment fluctuations to price movements, shedding light on the emotional undercurrents driving investor behavior. These insights equip analysts and traders with more reliable tools for interpreting and anticipating market trends.

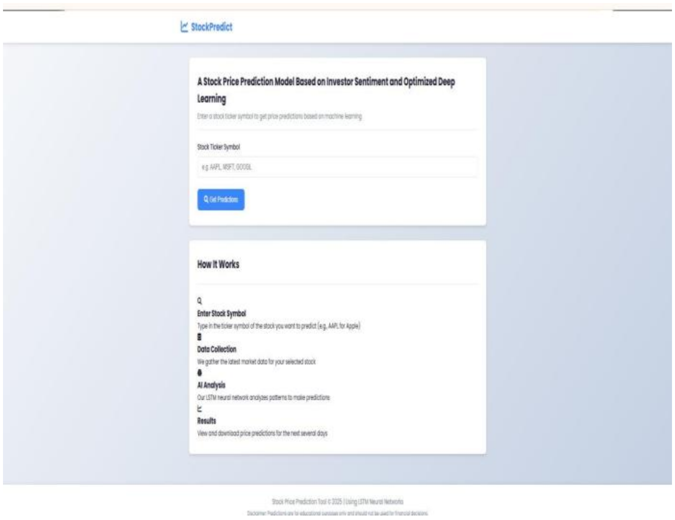

Figure 2: Stock Prediction Page

International Research Journal of Innovations in Engineering and Technology (IRJIET)
ISSN (online): 2581-3048
Volume 9, Special Issue ICCIS-2025, pp 102-106, May-2025
https://doi.org/10.47001/IRJIET/2025.ICCIS-202516
Second International Conference on Computing and Intelligent Systems (ICCIS-2025)

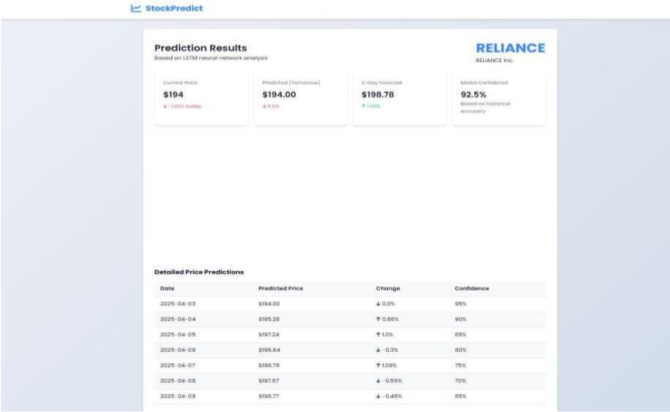

Figure 3: Result page of Stock Price Prediction

V. CONCLUSION

This research presents a comprehensive approach to stock price prediction by integrating investor sentiment extracted from financial media and social platforms with traditional historical market data. The proposed model delivers a well-rounded analysis of market dynamics, effectively capturing both numerical patterns and emotional signals. Incorporating sentiment analysis enhances the model's ability to detect shifts in public mood—often early indicators of market volatility. The findings demonstrate that blending behavioral data with time-series analysis significantly boosts predictive accuracy, particularly in fast-changing or uncertain market conditions.

The system's accessible and interactive web interface adds practical value, enabling real-time user engagement and clear visualization of predictions. This makes it a valuable resource for traders, analysts, and financial institutions alike. Looking ahead, enhancements could include the integration of live sentiment data streams, incorporation of broader macroeconomic indicators, and implementation of adaptive learning methods to maintain performance over time. Expanding the platform to cover a wider range of assets and enabling scalable deployment would further strengthen its role as a robust decision-support tool in the financial sector.

Citation of this Article:

B. Rupadevi, & Shaik Nadheem. (2025). Investor Sentiment-Driven Stock Price Prediction Using Optimized Deep Learning. In proceeding of Second International Conference on Computing and Intelligent Systems (ICCIS-2025), published in *IRJIET*, Volume 9, Special Issue ICCIS-2025, pp 102-106. Article DOI https://doi.org/10.47001/IRJIET/2025.ICCIS-202516

****

