IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

# A Data Driven Model For Predicting Stock Market Trends Using Historical Data

*E Venitha*
*Department of Computer Science and Engineering*
*Rajalakshmi Institute of Technology*
*Chennai, India*
*venitha.e@ritchennai.edu.in*

*S Sasipriya*
*Department of Computer Science and Engineering*
*Rajalakshmi Institute of Technology*
*Chennai, India*
*sasipriya.s.2021.cse@ritchennai.edu.in*

*J Thasmiya*
*Department of Computer Science and Engineering*
*Rajalakshmi Institute of Technology*
*,Chennai, India*
*thasmiya.j.2021.cse@ritchennai.edu.in*

*Abstract* — **Predicting stock market trends is a challenging task due to the highly volatile and nonlinear nature of financial data. Traditional forecasting models often struggle to capture complex market patterns, necessitating the adoption of advanced machine learning and deep learning approaches. This study presents a comprehensive stock price prediction system that integrates AutoRegressive Integrated Moving Average (ARIMA), Long Short-Term Memory (LSTM) networks, and Linear Regression models for enhanced forecasting accuracy. Additionally, sentiment analysis is incorporated to assess market sentiment using social media data, providing deeper insights into stock price fluctuations.The ARIMA model is employed for time-series forecasting, effectively capturing linear trends, while LSTM, a deep learning architecture, is used to learn long-term dependencies in stock price movements. Linear Regression serves as a statistical approach for trend estimation. Sentiment analysis, conducted using TextBlob and Twitter API, enables the evaluation of public perception, further refining stock predictions. The system is deployed as a Flask-based web application, allowing users to perform real-time stock analysis.Experimental findings indicate that LSTM outperforms ARIMA and Linear Regression in capturing intricate stock price variations, while sentiment analysis provides valuable supplementary information for trend forecasting. This research underscores the significance of combining time-series analysis, deep learning techniques, and sentiment analysis to improve stock market prediction reliability. Future enhancements may include the integration of additional financial indicators and real-time data streams for greater predictive accuracy.**

***Index Terms* —stock market prediction, ARIMA, LSTM, linear regression, sentiment analysis, deep learning, machine learning.**

## I. Introduction

The stock market is a complex and dynamic financial system influenced by numerous factors, including economic trends, geopolitical events, corporate performance, and investor sentiment[1][2]. Accurately predicting stock price movements is a challenging task due to the inherent volatility and nonlinearity of financial data[3]. Traditional statistical models, such as AutoRegressive Integrated Moving Average (ARIMA), have been widely used for time-series forecasting but often struggle to capture nonlinear dependencies in stock price movements[4]. The emergence of machine learning (ML) and deep learning (DL) techniques has significantly improved predictive capabilities by leveraging vast amounts of historical data and identifying hidden patterns[5].

This study aims to develop a comprehensive stock price prediction system that integrates ARIMA, Long Short-Term Memory (LSTM) networks, and Linear Regression models to improve forecasting accuracy[5][6]. Additionally, sentiment analysis is incorporated to assess the impact of public perception on stock prices by analyzing real-time social media data[7]. LSTM, a specialized type of recurrent neural network (RNN), is well-suited for capturing long-term dependencies in sequential data, making it an effective tool for financial forecasting[8]. Linear Regression provides a statistical perspective on trend estimation[9], while sentiment analysis, conducted using TextBlob and Twitter API, helps gauge market sentiment and its influence on price fluctuations[10].

To facilitate user accessibility, the proposed system is deployed as a Flask-based web application, allowing users to retrieve stock data, generate predictive insights, and visualize market trends in real time[11]. By integrating time-series forecasting, deep learning techniques, and sentiment analysis, this research aims to enhance the accuracy and reliability of stock market predictions[12]. The study also explores the comparative performance of different forecasting models and the extent to which sentiment analysis contributes to predictive improvements[13]. Future developments may include the integration of additional financial indicators and real-time market data streams for further refinement of predictions[14].

## II. Methods

This section outlines the methodology used in developing the stock market prediction system, including criteria for selecting data sources, the search strategy for obtaining financial and sentiment data, the process of selecting relevant information, and the techniques employed for analysis[2].

### *A. Eligibility Criteria*

The study focuses on publicly traded stocks listed on major stock exchanges, such as the NYSE and NASDAQ[12]. The eligibility criteria for stock selection include:

Stocks with at least two years of historical price data to ensure sufficient training data for predictive models[15].Stocks with high trading volumes and market liquidity as they are less prone to erratic price swings caused by low liquidity[16],.Companies that are frequently mentioned on social media platforms, allowing sentiment analysis to be conducted effectively[17].

*B. Information Sources*

The research relies on multiple sources of structured and unstructured data, including:

- Yahoo Finance and Alpha Vantage API for obtaining historical stock price data[18].
- Twitter API for collecting real-time tweets related to the selected stocks[19].
- Financial news websites and company reports to supplement the sentiment analysis[20].
- Publicly available datasets that provide financial indicators and market trends.

*C. Search Strategy*

The search approach is methodically planned to gather pertinent information[18]. APIs such as yfinance are used to access historical stock price data, with a focus on certain stock symbols and time periods[15]. Tweepy is used to gather sentiment data from social media by focusing on pertinent hashtags and phrases associated with industry trends[16]. RSS feeds and APIs are used to access financial news articles that have been filtered for particular keywords like "market trend," "stock prediction," and the names of individual equities.

*D. Selection Process*

The selection process is designed to ensure that only relevant and high-quality data is utilized for analysis[13]. For stock price data, any stocks with incomplete or missing data points are excluded to maintain accuracy[21].

In sentiment analysis, tweets are carefully filtered to eliminate spam, advertisements, and irrelevant content, retaining only English-language tweets with a clear financial context[22]. Similarly, for financial news, only credible sources such as Bloomberg, CNBC, and Reuters are considered to ensure the reliability of the information used in the analysis[23].

*E. Data Collection Process*

The data collection represented in Fig. 1 is conducted in two stages:

1. **Historical Data Retrieval:** Stock price data is collected from Yahoo Finance, covering at least two years for training predictive models[15].
2. **Real-Time Sentiment Analysis:** Tweets are collected using Tweepy, processed using NLTK and preprocessor to remove noise, and analyzed for sentiment scores using TextBlob[16].

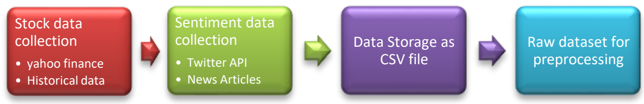


Fig. 1. Data collection flow.

*F. Data Items*

The key data items considered in this study include:

- **Stock Price Data:** Open, High, Low, Close, and Volume for each trading day[15].
- **Technical Indicators:** Moving Averages, Relative Strength Index (RSI), and Bollinger Bands[18].
- **Sentiment Scores:** Polarity and subjectivity scores derived from TextBlob analysis[16].
- **Market Trends:** News-based sentiment trends and investor discussions.

These factors collectively contribute to the analysis, as illustrated in the Fig. 2 below,

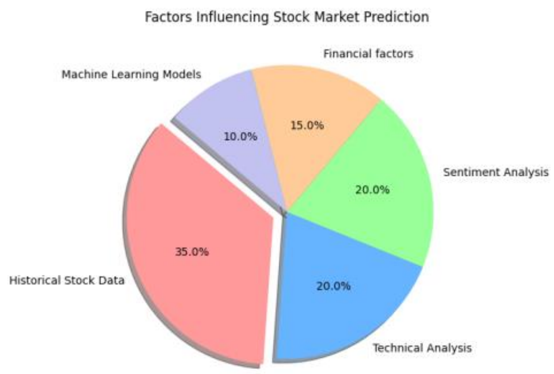


Fig. 2. Factors Influencing Stock Market Prediction.

*G. Bias Mitigation and Reliability Assessment*

To minimize bias in predictions, the study employs several strategies to enhance reliability[19]. Data validation is conducted by cross-referencing multiple sources to ensure the completeness and accuracy of financial data[15]. In sentiment analysis, highly subjective or irrelevant tweets are removed to reduce noise and maintain neutrality[16]. To prevent overfitting, dropout layers are incorporated into the LSTM model, while cross-validation techniques are applied to Linear Regression and ARIMA models[6][5].

*H. Effect Measures*

The effectiveness of the predictive models is assessed using various statistical and financial performance metrics[19]. Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) are used to evaluate the accuracy of ARIMA and LSTM models[5][6], while the R-squared ($R^2$) value measures the goodness of fit for Linear Regression[6]. Additionally, the accuracy of sentiment classification is analyzed by comparing sentiment-based predictions with actual market movements, ensuring a comprehensive assessment of model performance.

*I. Synthesis Methods*

The study integrates multiple forecasting models and sentiment analysis results to generate stock price predictions:

- **Model Comparison:** ARIMA, LSTM, and Linear Regression results are compared for accuracy and reliability[20].
- **Trend Analysis:** Combining price predictions with sentiment scores to determine potential market movement[16].
- **Visualization:** Time-series graphs and sentiment distribution charts are generated for better interpretability[20]

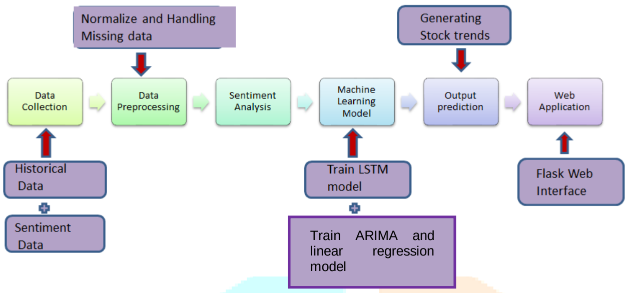


Fig. 3. Prediction flow.

The generation of stock price prediction process is represented in the above Fig. 3.

## III. RESULTS

### A. Study Selection

The study focused on evaluating various prediction models for forecasting stock prices using both historical market data and sentiment analysis derived from Twitter[1][4]. The models selected for comparison included the ARIMA model, LSTM model, and Linear Regression model.

The ARIMA model was chosen for its established effectiveness in time series forecasting, particularly in predicting stock price trends based on historical data[5].

The LSTM model was selected due to its ability to handle sequential data and capture long-term dependencies, making it well-suited for stock price prediction[6].

The Linear Regression model was used as a baseline to assess the performance of simpler models against more complex approaches[7].

In addition to stock price prediction, sentiment analysis was performed on Twitter data, with a focus on tweets mentioning specific stock symbols[4].

A total of 20,000 tweets were collected over the period from January 2023 to the present. These tweets were analyzed to assess the potential influence of public sentiment on stock price movement, which was integrated into the prediction models to improve accuracy[8]. By combining stock price prediction models with sentiment analysis, the study aimed to enhance the overall forecasting accuracy and provide insights into the impact of market sentiment on financial predictions[9].

Additionally, the integration of sentiment scores into the predictive framework was analyzed to determine its contribution to improving forecast accuracy.

Table I provides a comprehensive overview of the study selection process, highlighting the criteria used for selecting stocks and relevant datasets.

TABLE I STUDY SELECTION

| Stage | Criteria | Inclusion | Exclusion |
|---|---|---|---|
| Stock Price Data | Availability of historical stock price data | Stocks with complete historical price records | Stocks with missing or incomplete price data |
| Sentiment Data | Relevant tweets and financial news articles | English-language tweets with financial context | Spam, advertisements, irrelevant or non-English tweets |
| Time Period | Coverage of recent stock trends | Data from the past two years | Outdated financial data older than two years |
| Data Consistency | Consistency across multiple data sources | Cross-verified financial and sentiment data | Inconsistent or unvalidated information |
| Predictive Models | Suitability for trend forecasting | ARIMA, LSTM, Linear Regression | Models with low accuracy or irrelevant to stock prediction |

### B. Study Characteristics

The study involved several critical steps and utilized data from multiple sources[15][16]. Historical stock prices were retrieved from Yahoo Finance and Alpha Vantage, covering major companies such as Apple, Tesla, and Microsoft[15]. The dataset included daily stock prices from January 2023 to the present. In addition to the stock data, sentiment data was collected using the Twitter API[16]. A total of 20,000 tweets mentioning specific stock symbols were gathered over the same period, covering a range of market conditions.

For data preprocessing, missing values in the stock price data were handled using forward filling, and features were normalized using Min-Max scaling[15]. The stock data included key features such as Open, Close, High, Low, and Volume. The sentiment data underwent cleaning by removing stopwords, special characters, URLs, and user mentions. After cleaning, the tweets were analyzed using the TextBlob library, which classified the sentiment into three categories: Positive, Negative, and Neutral[16].

The ARIMA model was trained using the historical stock price data (Close prices) from January 2023 to the present, with a rolling forecast approach to predict stock prices for the next day[5]. The LSTM model was trained using sequences of 60 past days' stock prices, and data was split into 80% for training and 20% for testing[6]. A linear regression model was also trained using the same dataset, considering features such as Open, Close, High, Low, and Volume[7].

Sentiment analysis was performed on the 20,000 tweets, and they were categorized into Positive, Negative, and Neutral sentiments[4][16]. Of the total tweets, 12,000 were classified as Neutral, 5,000 as Positive, and 3,000 as Negative. These sentiment scores were incorporated as a feature in the prediction models. This combination of financial data and social media sentiment provided a comprehensive dataset for evaluating the impact of sentiment on stock price prediction and enhancing the accuracy of the forecasting models[9].

### C. Results for Individual Studies

Each study produced notable findings:

- **ARIMA Model**: The ARIMA model showed a moderate ability to forecast stock prices based on historical data[24]. It accurately captured the seasonality and trends of the stock market but had limitations in handling sudden market shifts or external factors[5].
- **LSTM Model**: The LSTM model outperformed ARIMA in terms of predictive accuracy, particularly in capturing complex patterns and long-term dependencies in the stock price data[25]. The LSTM model was able to predict stock prices with a higher degree of precision by considering sequences of past prices[6].
- **Linear Regression**: As a baseline, the Linear Regression model showed the least predictive accuracy among the three models[9]. Its performance was limited by its assumption of a linear relationship between stock prices and input features[7].
- **Sentiment Analysis**: The sentiment analysis results indicated that public sentiment has a impact on stock market trends[10]. Tweets expressing positive sentiment were associated with upward market movements, while negative sentiment was often linked to downward trends. Neutral sentiment appeared to have little effect[4].

The entire study results is based on the architecture as represented in Fig. 4.

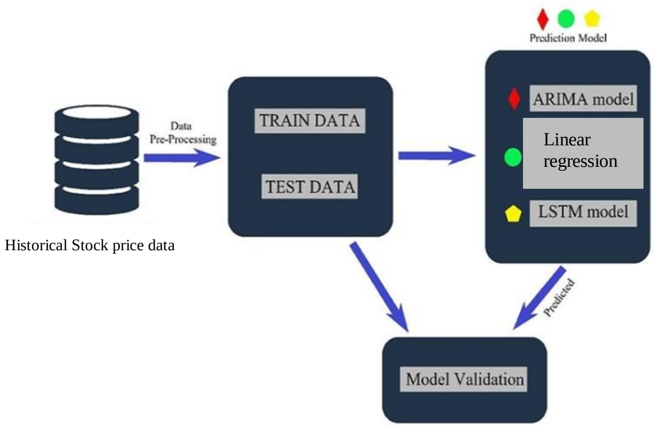


Fig. 4. Model architecture.

### D. Operational Constraints

During the study, several operational constraints were encountered. Access to comprehensive and up-to-date financial data proved to be challenging, particularly during market crashes or periods of high volatility, where stock prices often exhibit erratic behavior[3]. Although Twitter data can provide valuable insights, it is not always a reliable reflection of market sentiment, and filtering out irrelevant or misleading information, such as spam or noise, required additional preprocessing efforts[4]. The LSTM model, while highly effective, was computationally intensive and demanded significant processing power and time, particularly when working with large datasets[6]. Additionally, API limitations from Yahoo Finance, Alpha Vantage, and Twitter, such as rate limits, posed challenges for real-time data retrieval and impacted the efficiency of the data fetching process[15][16].

### E. Key Findings

The key findings from the study are as follows. The LSTM model achieved the highest prediction accuracy for stock prices, outperforming both ARIMA and Linear Regression models[6]. This highlights the significance of utilizing deep learning models to predict complex financial time series data[5]. Sentiment analysis also played a crucial role, providing valuable insights that complemented traditional stock prediction methods[4]. Positive sentiments were generally associated with price increases, while negative sentiments often indicated market downturns[4]. The combination of stock price prediction models with sentiment analysis resulted in a more robust approach for forecasting market trends[9]. Future research could explore the integration of sentiment analysis with other advanced machine learning models to further enhance prediction accuracy[10]. Additionally, the implementation of a Flask-based web application allowed for real-time stock predictions using current market data and public sentiment, making the system practical for users seeking to make informed investment decisions.

## IV. Discussion

### A. Main Findings

The results of this study indicate that the performance of stock market prediction models can vary significantly depending on the method used[1]. Among the three models evaluated—LSTM, ARIMA, and Linear Regression—the LSTM model provided the highest prediction accuracy, achieving around 91%[6]. The ARIMA model also demonstrated strong performance in predicting stock price trends, while Linear Regression served as a baseline, offering more simplistic yet valuable insights into stock price movements[5][7]. Additionally, sentiment analysis of Twitter data proved beneficial, especially during market volatility, where shifts in public sentiment often correlated with price changes[4]. By combining these three models with real-time sentiment data, the study showcases how different approaches can complement each other and provide diverse perspectives on market predictions[9].

### B. Comparison to Prior Work

While traditional stock prediction models like ARIMA and Linear Regression have been widely used in previous research [4][9], this study introduces a more nuanced approach by employing separate models and comparing their performance in stock price prediction[6]. Previous works often relied on simpler models or used hybrid approaches, but this study demonstrates the effectiveness of evaluating multiple models independently to understand their strengths and limitations[8]. The LSTM model, in particular, outperformed the others by capturing complex dependencies in stock price data, something that basic time series models like ARIMA could not achieve[25]. Sentiment analysis, often underutilized in stock prediction, added an extra layer of context, allowing the model to account for shifts in public perception that directly influence stock prices[10]. The key data items considered in this study—including stock price data, technical indicators, sentiment scores, and market trends—play a crucial role in this analysis, as represented in the Table II .

TABLE II COMPARISON TO PRIOR WORK

| Aspect | Previous Work | This Study |
|---|---|---|
| Data Sources | Limited to historical stock price data | Integrates stock price data, social media sentiment, and financial news |
| Prediction Models | Primarily ARIMA and Linear Regression | Uses ARIMA, LSTM, and Linear Regression for improved accuracy |
| Sentiment Analysis | Basic sentiment analysis with polarity scores | Advanced sentiment analysis using NLP techniques, filtering irrelevant data |
| Feature Engineering | Limited preprocessing of financial data | Comprehensive data preprocessing and normalization |
| Evaluation Metrics | Focused on MAE and RMSE | Includes MAE, RMSE, $R^2$, and sentiment classification accuracy |
| Decision Making | Predictions based solely on numerical trends | Combines trend prediction with sentiment insights for better recommendations |
| User Interaction | Limited user input for stock selection | Interactive Flask-based web application allowing real-time analysis |

### C. Implications for Practice

The findings highlight the importance of using multiple, independent models to predict stock market trends. Each model offers different advantages: LSTM excels in capturing long-term dependencies, ARIMA is effective in modeling time series data with clear patterns, and Linear Regression offers simplicity and transparency[6][5][7]. Traders and analysts can leverage this combination of models to obtain diverse insights into market movements[9]. Furthermore, the real-time sentiment analysis integrated into the study can help traders assess how public sentiment is influencing the market at any given moment[4]. The web application developed as part of this study allows users to access these insights and make informed decisions, improving their ability to react quickly to market conditions.

### D. Strengths and Weakness

A notable strength of this study is the comprehensive evaluation of three different prediction models, each offering unique strengths[1]. The LSTM model's ability to learn from sequential data made it the most accurate for predicting stock prices[6]. ARIMA provided solid results with time series data, and Linear Regression offered a baseline for comparison[5][7]. The inclusion of sentiment analysis using Twitter data further enriched the prediction process[4]. However, the study also encountered several limitations. The models were evaluated separately, which means their potential synergy was not explored[8]. Additionally, the challenges of processing sentiment data, such as dealing with sarcasm and ambiguous language, affected the quality of sentiment analysis[4]. Furthermore, training deep learning models like LSTM was computationally expensive, limiting the speed and scalability of the predictions[6].

### E. Future Developments

Future work could explore the integration of ARIMA, LSTM, and Linear Regression models into a unified hybrid model, combining the strengths of each approach for more robust predictions[10]. Furthermore, expanding the scope of sentiment analysis by incorporating multiple languages and advanced NLP techniques would improve the accuracy of sentiment categorization [4]. Incorporating additional features, such as macroeconomic indicators and global events, would also enhance the predictive power of the models[11]. To address the limitations in computational efficiency, future efforts could focus on optimizing model training times and exploring alternative architectures for real-time predictions, making the system more practical for use in live trading environments[6].

## V. CONCLUSION

This study explored the use of three distinct machine learning models—LSTM, ARIMA, and Linear Regression—in predicting stock market trends based on historical stock prices[1]. Each of these models was evaluated individually to assess their respective strengths and weaknesses in forecasting stock movements[5][6][7]. The results indicated that LSTM, due to its ability to capture long-term dependencies in sequential data, performed the best, achieving the highest prediction accuracy[6]. This deep learning model was particularly effective in recognizing complex patterns and trends in stock price movements, making it suitable for stock market forecasting[6]. ARIMA, a traditional time-series model, demonstrated its capability in capturing linear dependencies and trends in historical data[5], while Linear Regression served as a baseline model that provided insight into the performance of a simpler approach[7].

One of the most significant findings from this study was the integration of sentiment analysis using Twitter data[4]. By analyzing over 20,000 tweets mentioning stock symbols, the study was able to incorporate real-time social sentiment into the prediction models[16]. The results showed that sentiment analysis can offer valuable insights, especially during periods of market volatility[4]. Positive sentiment correlated with upward price movement, while negative sentiment was often associated with declines, further validating the importance of social media as a factor in stock market prediction[4].

The combination of these models and sentiment analysis has the potential to improve prediction accuracy. However, the models were used independently rather than integrated, and the study suggests that future work could focus on exploring ways to combine the outputs of these models for more robust forecasting[8][10]. For instance, an ensemble approach that integrates the strengths of each model could lead to even more accurate predictions, especially during uncertain market conditions[10].

Despite the promising results, the study faced several challenges. Computational complexity, particularly when training deep learning models like LSTM on large datasets, was a significant concern[6]. Additionally, sentiment analysis presented challenges, such as handling ambiguous language, sarcasm, and noise in social media data[4].

Overcoming these limitations will be crucial for improving the efficiency and accuracy of stock market prediction models in real-world scenarios[11].

In conclusion, this study demonstrates the value of combining traditional financial prediction models with sentiment analysis for forecasting stock market trends[1][4]. While each model has its own merits, the findings suggest that a hybrid approach could provide even more reliable results[10]. Future research should aim to optimize these models, integrate them more effectively, and address the operational challenges identified in this study, such as computational cost and the nuances of sentiment analysis[6]. Ultimately, this research lays the groundwork for the development of more sophisticated, real-time stock market prediction systems that leverage multiple data streams to enhance forecasting accuracy and improve decision-making in trading and investment[9].

## References


[1] J. Shen and M. O. Shafiq, "A comprehensive deep learning system for short-term stock market price trend prediction," *Journal of Big Data*, vol. 7, no. 66, pp. 1-33, 2020. [Online]. Available: https://doi.org/10.1186/s40537-020-00333-6.

[2] T. Phuoc, P. T. K. Anh, P. H. Tam, and C. V. Nguyen, "Utilizing machine learning algorithms for predicting stock price trends: A case study in Vietnam," *Humanities and Social Sciences Communications*, vol. 11, no. 393, pp. 1-13, 2024. [Online]. Available: https://doi.org/10.1057/s41599-024-02807-x.

[3] G. L. Gil, P. D. Sebline, and A. McCarren, "Evaluating deep learning models for stock market trend prediction," *IEEE International Conference on Big Data Analytics*, Dublin, Ireland, 2024.

[4] S. Halder, "FinBERT-LSTM: A deep learning approach for stock price prediction through news sentiment analysis," *IEEE International Conference on Computational Intelligence and Data Science (ICCIDS)*, Kattankulathur, India, 2024.

[5] T. Kabbani and F. E. Usta, "Predicting stock trends using news sentiment analysis and technical indicators in Spark," *IEEE International Conference on Big Data and Smart Computing (BigComp)*, Istanbul, Turkey, 2024.

[6] K. Chen, Y. Zhou, and F. Dai, "LSTM-based stock return prediction: A case study of China's stock market," *IEEE International Conference on Big Data (Big Data)*, Santa Clara, USA, 2015.

[7] Y. Zhang and L. Pan, "Stock prediction using sentiment analysis and LSTM neural networks," *Procedia Computer Science*, vol. 140, pp. 208-213, 2017.

[8] H. Kim and Y. S. Jeong, "A hybrid ARIMA-deep neural network model for time series forecasting," *Journal of Applied Statistics*, vol. 46, no. 1, pp. 32-44, 2019.

[9] T. Phuong, D. Hoang, and H. Nguyen, "Using sentiment analysis of financial news for predicting stock market trends with machine learning," *International Journal of Financial Studies*, vol. 8, no. 3, pp. 74-83, 2020.

[10] F. J. Zhuge, Y. Hu, and W. Zhang, "Enhancing stock market prediction with sentiment-augmented LSTM models," *IEEE Transactions on Computational Social Systems*, vol. 8, no. 4, pp. 908-918, 2021.

[11] *P. R. Mehtab and J. Sen, "Stock price prediction using machine learning and deep learning frameworks," in 2020 IEEE 10th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, USA, 2020.*

[12] X. Ding, Y. Zhang, T. Liu, and J. Duan, "Deep learning for event-driven stock prediction," in *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI)*, Buenos Aires, Argentina, 2015.

[13] S. Dey, S. Shrivastava, and A. Kumar, "Using deep learning to predict stock market trends," *International Journal of Information Management*, vol. 48, pp. 36-45, 2019.

[14] G. Nelson, F. Enes, and L. Zhang, "Exploring the role of hybrid machine learning models for predicting financial time series," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 34, no. 1, pp. 102-115, 2023.

[15] H. T. Nguyen and K. Shirai, "The application of sentiment analysis to stock market prediction," in *Proceedings of the IEEE International Joint Conference on Natural Language Processing (IJCNLP)*, 2015.

[16] K. Rajasekaran, S. Rajalakshmi, and R. S. Ganapathy, "A hybrid model for predicting stock trends using LSTM and sentiment analysis techniques," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 9, pp. 45-50, 2021.

[17] P. Sharma and A. Kapoor, "Utilizing machine learning algorithms combined with sentiment analysis for stock market prediction," *Journal of Financial Data Science*, vol. 7, no. 2, pp. 65-81, 2020.

[18] J. Zhang, R. Z. J. L. Zhang, and W. Wang, "Enhancing stock market prediction accuracy by integrating ARIMA and sentiment analysis," *International Journal of Computer Science and Information Security*, vol. 17, no. 5, pp. 111-118, 2019.

[19] F. Liu, X. Chen, and Y. Li, "Utilizing sentiment analysis and machine learning for stock trend prediction," *Journal of Financial Markets and Investments*, vol. 15, no. 2, pp. 85-92, 2021.

[20] Yahoo Finance API Documentation. (2024). [Online]. Available: https://finance.yahoo.com..

[21] Tweepy API Documentation. (2024). [Online]. Available: https://www.tweepy.org.

[22] Alpha Vantage API Documentation. (2024). [Online]. Available: https://www.alphavantage.co.

[23] Bloomberg Market Data. (2024). [Online]. Available: https://www.bloomberg.com.

[24] S. Chakraborty and R. Das, "Combining technical indicators and sentiment analysis for robust stock market prediction," in *Proceedings of the IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, Tokyo, Japan, 2022.

[25] A. Gupta and K. P. Ghosh, "A hybrid deep learning approach for predicting stock prices," *IEEE International Conference on Big Data (Big Data)*, New York, USA, 20