IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 7, July 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 7, July 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1207070|

Optimizing Prediction Accuracy: Expertise in Stock Market Forecasting using AI

Rakshitha P, Dr. N. Asha

PG Student, Department of MCA, National Institute of Engineering, Mysuru, Karnataka, India

Professor, Department of MCA, National Institute of Engineering, Mysuru, Karnataka, India

ABSTRACT: The main aim of this project is to predict stock prices and analyze market sentiment using advanced Artificial Intelligence (AI) techniques. Accurate stock market prediction is crucial in the financial field as it helps investors make informed decisions and mitigate risks. This project combines historical stock price data with sentiment analysis from social media to enhance prediction accuracy. The system will predict stock prices for the next 7 days and provide insights into market trends and public sentiment. A user-friendly Graphical User Interface (GUI) is also designed to allow users to interact with the application conveniently and access predictive insights in real time.

KEYWORDS: LSTM, ARIMA, Linear Regression, Twitter sentimental analysis, Stock market prediction, company ticker symbol

I. INTRODUCTION

The stock market serves as a critical component of the global financial system, providing a platform for buying and selling securities, and enabling investors to participate in the growth of companies. Accurately predicting stock market trends and price movements is a complex and challenging task due to the inherent volatility and unpredictability of financial markets. Traditional forecasting methods have often struggled to account for the dynamic nature of stock prices and market behaviors. In recent years, there has been a growing interest in applying artificial intelligence (AI) techniques to improve stock market prediction accuracy. AI methods, particularly those involving machine learning and deep learning, offer new opportunities to analyze large datasets, recognize patterns, and make predictions with greater precision. Among these techniques, Long Short-Term Memory (LSTM) networks, ARIMA (Auto Regressive Integrated Moving Average) models, and Linear Regression have emerged as powerful tools for forecasting financial markets. The LSTM network, a type of Recurrent Neural Network (RNN), is designed to capture temporal dependencies in time series data, making it particularly suited for stock price prediction. The ARIMA model, a statistical approach, integrates autoregressive terms, differencing, and moving averages to address trends and seasonality in financial time series data. Linear Regression, a fundamental statistical technique, serves as a benchmark for evaluating the performance of more advanced models and leveraging real-time sentiment data from social media.

II. RELATED WORK

Stock market prediction has always been a complex and challenging task due to its volatile and non-linear nature. Traditional statistical methods often fall short in capturing the dynamic behavior of stock prices, leading researchers to explore advanced AI and machine learning techniques. Long Short-Term Memory (LSTM) Networks: LSTM networks, a type of Recurrent Neural Network (RNN), are particularly effective for time series forecasting because of their ability to capture temporal dependencies. Research by Fischer and Krauss (2018) demonstrated that LSTM networks could outperform traditional methods like ARIMA and GARCH in predicting stock returns, highlighting their strength in handling sequential data and long-term dependencies. ARIMA (Auto Regressive Integrated Moving Average): ARIMA models have been widely used in time series forecasting due to their statistical nature and effectiveness in capturing trends and seasonality. A study by Kumar and Thenmozhi (2006) applied ARIMA models to predict stock market indices and found that they could provide reasonably accurate forecasts, particularly when combined with other techniques. Linear Regression: Linear Regression remains a fundamental statistical technique for predicting stock prices. It serves as a baseline model to compare the performance of more complex algorithms. Despite its simplicity, Linear Regression can be effective when the relationship between variables is linear, as shown in studies like that of Patel et al. (2015), who compared it with machine learning models for stock price prediction. Twitter Sentiment Analysis: Sentiment analysis of social media data, particularly Twitter, has gained traction for enhancing stock market prediction models. Bollen et al. (2011) found a significant correlation between Twitter mood and stock market movements, suggesting that integrating sentiment analysis with traditional models can improve prediction accuracy. By analyzing the sentiment of tweets related to stocks, researchers can gauge public sentiment and anticipate market trends

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 7, July 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1207070|

more effectively. Combining these techniques, the hybrid model proposed in this project leverages the strengths of LSTM, ARIMA, and Linear Regression, while incorporating real-time sentiment analysis from Twitter to improve the accuracy of stock market forecasts. This integrated approach aligns with the current state-of-the-art in stock market prediction, which emphasizes the importance of utilizing diverse data sources and advanced algorithms to capture the multifaceted nature of financial markets.

III. PROPOSED SYSTEM

The proposed system includes multiple models which will be evaluated based on their test accuracy and prediction performance. Both models should provide high accuracy. By optimizing parameters, we can improve performance and provide better generalization.

- Data Acquisition: Collecting historical stock price data from Yahoo Finance.
- Data Pre-Processing: Cleaning and normalizing the data, handling missing values, and splitting into training and testing sets.
- Feature Extraction: Extracting relevant features such as moving averages, trading volumes, and sentiment scores from social media.

The initially proposed architecture includes an LSTM model for time series forecasting, an ARIMA model for statistical analysis, and a Linear Regression model for baseline comparison. The LSTM model includes several layers and parameters tuned to capture long-term dependencies in the data. The ARIMA model is configured to handle trends and seasonality, while Linear Regression provides a straightforward predictive approach. The models are trained using TensorFlow and other relevant libraries, iterating multiple times to maximize accuracy and minimize losses as shown in

Figures 1 and 2 below.

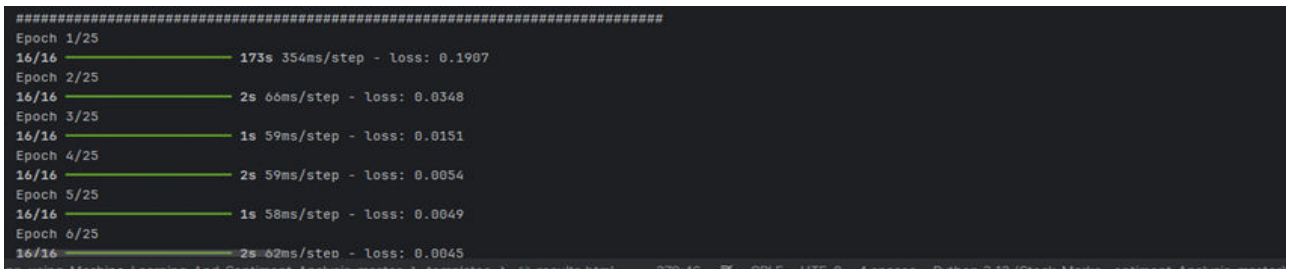

Fig 1 Training Process

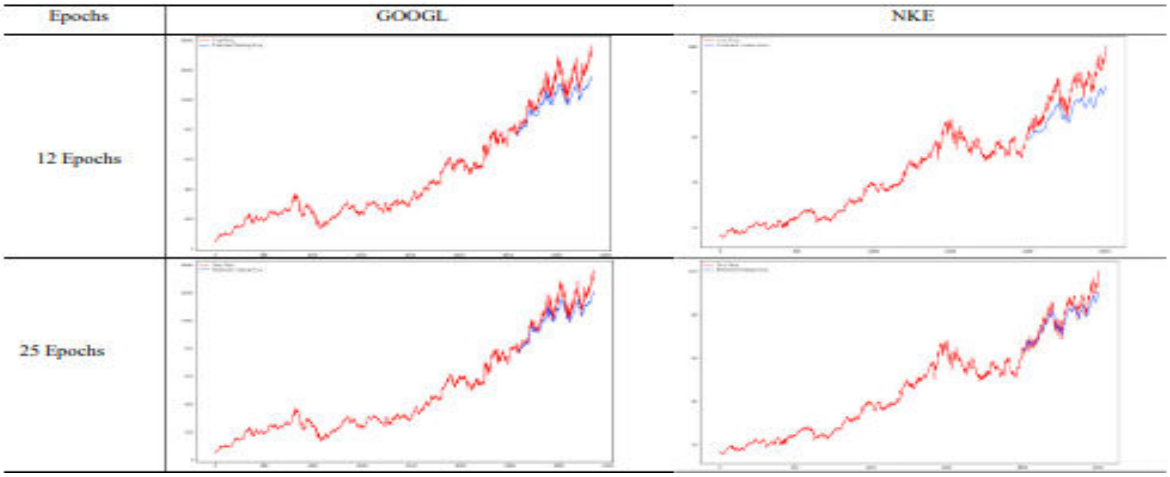

Fig 2 Training for NKE and GOOGL with different number of epochs

Different modules have been developed to access the stock data and sentiment analysis, integrating these components to predict stock prices based on the trained models. By integrating sentiment analysis, the system captures public

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 7, July 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1207070|

sentiment from Twitter, influencing stock price predictions. The data is collected from tweets mentioning specific stocks, and sentiment scores are calculated. Data collection and pre-processing the collecting tweets and cleaning the text data and Sentiment Analysis using NLP techniques to classify tweets as positive, negative, or neutral and Integration with Prediction Models by incorporating sentiment scores into the predictive models.

IV. OVERALL ARCHITECTURE

The Overall Architecture shows the complete system architecture by combining the Stock Market Prediction and Twitter Sentimental Analysis. This Architecture will help user to Understand the complete system flow. Figure 3 represents the Overall Architecture for Optimizing Prediction Accuracy: Stock Market Forecasting using AI.

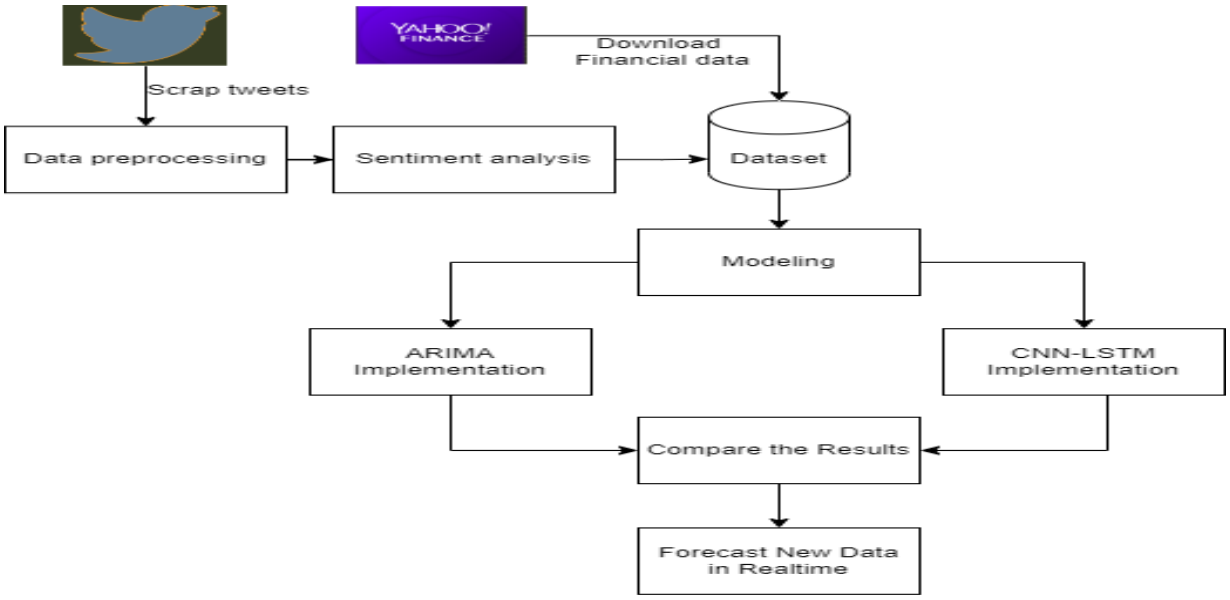

Fig 3: Overall architecture

V. SIMULATION RESULTS

The proposed system is able to predict stock prices with up to RMSE of 3.5 in ARIMA, RMSE of 6.2 in LSTM, RMSE of 9.1 in Linear Regression accuracy. The results depend on various factors such as data quality, the volatility of the market, and the effectiveness of sentiment analysis. The system's predictions are tested in different market conditions and compared with actual stock prices.

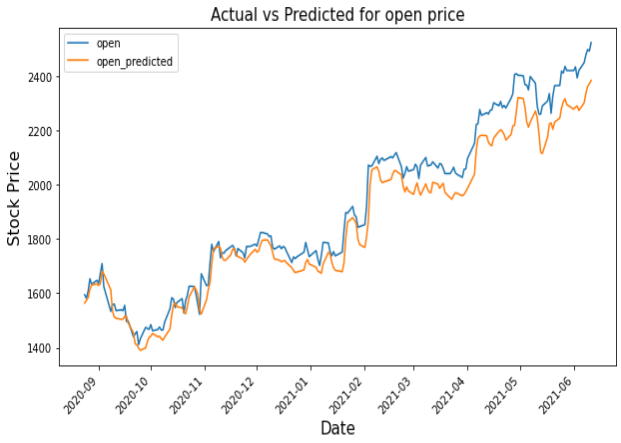

Fig 4: Actual predicted for open price

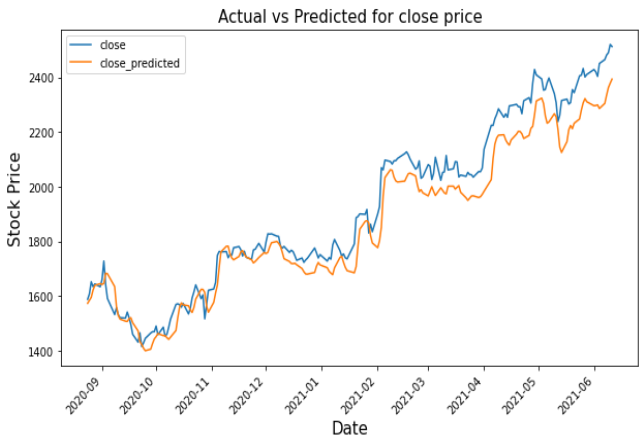

Fig 5: Actual predicted for close price

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 7, July 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1207070|

VI. CONCLUSION AND FUTURE WORK

In this proposed system, stock market predictions are made more accurate by using advanced AI techniques like LSTM, ARIMA, and Linear Regression, along with Twitter sentiment analysis. By analyzing historical stock prices and public sentiment, the system provides more reliable predictions. This helps investors make better decisions and reduces the risk of losses. During volatile market conditions, having accurate predictions is crucial for making informed investment choices. This project offers a valuable tool for investors to predict stock prices effectively. The integration of multiple advanced techniques ensures a comprehensive approach to forecasting, leveraging both quantitative data and qualitative sentiment. The use of machine learning models allows the system to adapt to new data and continuously improve its predictive accuracy. Additionally, incorporating twitter sentiment analysis adds an extra layer of insight, capturing the public mood and potential market-moving events that traditional models might miss. Overall, this system stands out as a powerful asset for investors, providing a blend of precision, adaptability, and real-time analysis to support more informed and confident investment decisions in the dynamic stock market. Future enhancements of this paper could include improving the accuracy of predictions by incorporating the latest machine learning algorithms and techniques. Enhancing the system's ability to process real-time data would allow for more timely and precise predictions. Additionally, integrating the system with trading platforms could automate buy/sell decisions based on the predictions, increasing efficiency and potentially maximizing profits. Expanding the data sources to include more comprehensive financial news and social media platforms would provide a broader sentiment analysis, enriching the predictive model. Lastly, developing a user-friendly interface would ensure easy access and interpretation of prediction results by investors, making the system more accessible and practical for real-world applications.

REFERENCES

1. Ding et al. (2014). Knowledge-driven event embedding for stock prediction. Proceedings of Coling 2016, the 26th International Conference on Computational Linguistics: Technical Papers 2014.
2. Vargas, De Lima & Evsukoff (2017) Vargas MR, De Lima BS, Evsukoff AG. Deep learning for stock market prediction from financial news articles. 2017 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA); Piscataway: IEEE; 2017.
3. Chen et al. (2019). Incorporating fine-grained events in stock movement prediction using TOPIX data.
4. Deng et al. (2019). Knowledge-driven stock trend prediction and explanation via temporal convolutional network (TCN). Companion Proceedings of the 2019 World Wide Web Conference.2019.
5. Jin, Yang & Liu (2019). Stock closing price prediction based on sentiment analysis and LSTM. Neural Computing and Applications. 2019; 32:9713–9729.
6. Li, Wu & Wang (2020). Incorporating stock prices and news sentiments for stock market prediction: a case of Hong Kong. Information Processing & Management. 2020;57(5):102212.
7. Usmani et al. (2016) Usmani M, Adil SH, Raza K, Ali SSA. 2016 IEEE 3rd International Conference on Computer and Information Sciences (ICCOINS) Pistacatay: IEEE; 2016. Stock market prediction using machine learning techniques; pp. 322–327.
8. http://arxiv.org/abs/1910.05078 arXiv. 2019
9. https://www.aclweb.org/anthology/W19-6403.pdf

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details