International Journal of Engineering Applied Sciences and Technology, 2024
Vol. 9, Issue 01, ISSN No. 2455-2143, Pages 141-146
Published Online May 2024 in IJEAST (http://www.ijeast.com)

STOCKSIGHT: A NOVEL APPROACH TO STOCK PRICE PREDICTION WITH SENTIMENT ANALYSIS AND LSTM NETWORKS

Dr. Jothimani K
Professor,
Department of Master of Computer Applications
Vivekananda College of Engineering & Technology
Puttur, India

Mahammad Afnan M
II MCA,
Department of Master of Computer Applications
Vivekananda College of Engineering & Technology
Puttur, India

Abstract—**StockSight, a predictive model integrating sentiment analysis and Long Short-Term Memory (LSTM) networks to enhance the accuracy of stock price predictions. By leveraging sentiment data from the financial news, the intention of the model is to provide more reliable forecasts of stock market movements. This study demonstrates the effectiveness of combining sentiment analysis with advanced machine learning techniques in stock market prediction.**

Keywords: **Beautiful Soup, LSTM Model, Sentiment Analysis, Stock Price Prediction, Web Crawling.**

I. INTRODUCTION

Stock market prediction is a challenging task due to its volatile nature and various factors influencing the market. Traditional prediction methods often find it difficult in capturing the complexity of market dynamics. StockSight provides solution to get market insight, a comprehensive model that combines sentiment analysis with LSTM networks to improve the accuracy of prediction. Sentiment data provides additional context that enhances the model's predictive capabilities. By using the real-time sentiment data from the news articles, the aims of the model is to provide more reliable predictions. The main focus is on the five prominent stocks: TCS, Tata Motors, Infosys, Asian Paints, and Tech Mahindra Ltd.

Development of a novel stock price prediction system that leverages the power of sentiment analysis and LSTM networks. The intention is to establish a robust data foundation by comprehensively collecting news headlines from several trusted sources and combining that with the historical stock prices. The another objective is to combine sentiment analysis along with LSTM model. Finally, the research seeks to translate these insights into actionable information for investors through the development of a user-friendly web application.

II. RELATED WORK

The literature reveals several approaches to stock market prediction, predominantly focusing on either technical analysis or sentiment analysis. Derakhshan and Beigy [1] demonstrated the use of sentiment analysis to get sentiment score from stock social media for predicting stock price movements. Jin et al. [2] integrated sentiment analysis with LSTM to predict stock closing prices, showing significant improvements over traditional methods. Mittal and Goel [3] explored Twitter sentiment analysis to get public sentiment for performing stock price prediction, highlighting the potential of social media as a valuable data source. Bollen et al. [4] analyzed Twitter mood for economic indicators, demonstrating a correlation between public sentiment and market trends. Zhang et al. [5] used LSTM networks for forecasting the stock movement, combining it with convolutional neural networks to improve performance. Schumaker and Chen [6] utilized textual analysis of news articles for stock prediction, developing the AZFin text system that showed promising results. Vu et al. [7] combined Twitter sentiment along with the advanced ML tools for stock prediction, achieving high accuracy rates. This study builds on these foundations by integrating sentiment analysis with LSTM networks, aim to enhance prediction accuracy further by leveraging both historical stock price data and the public sentiment data.

International Journal of Engineering Applied Sciences and Technology, 2024
Vol. 9, Issue 01, ISSN No. 2455-2143, Pages 141-146
Published Online May 2024 in IJEAST (http://www.ijeast.com)

III. METHODOLOGY

The methodology used involves a systematic approach to collecting data, processing it, and using advanced ML technologies to get insight into the stock market. The key phases in the methodology includes Data Collection, Data Preprocessing, Sentiment Analysis, Model Building, Training and Testing, and Development of a User Interface.

A. Data Collection

Data collection is the foundational step for this research. We gathered historical stock price data and news headlines for five stocks: TCS, Tata Motors, Infosys, Asian Paints, and Tech Mahindra Ltd. The sources for the data include Finance for stock prices and NewsAPI for news headlines. Historical stock prices were fetched using the Finance API, which provides comprehensive historical market data. News articles and headlines relevant to the chosen stocks were collected from various financial news sources using the NewsAPI and scraping.

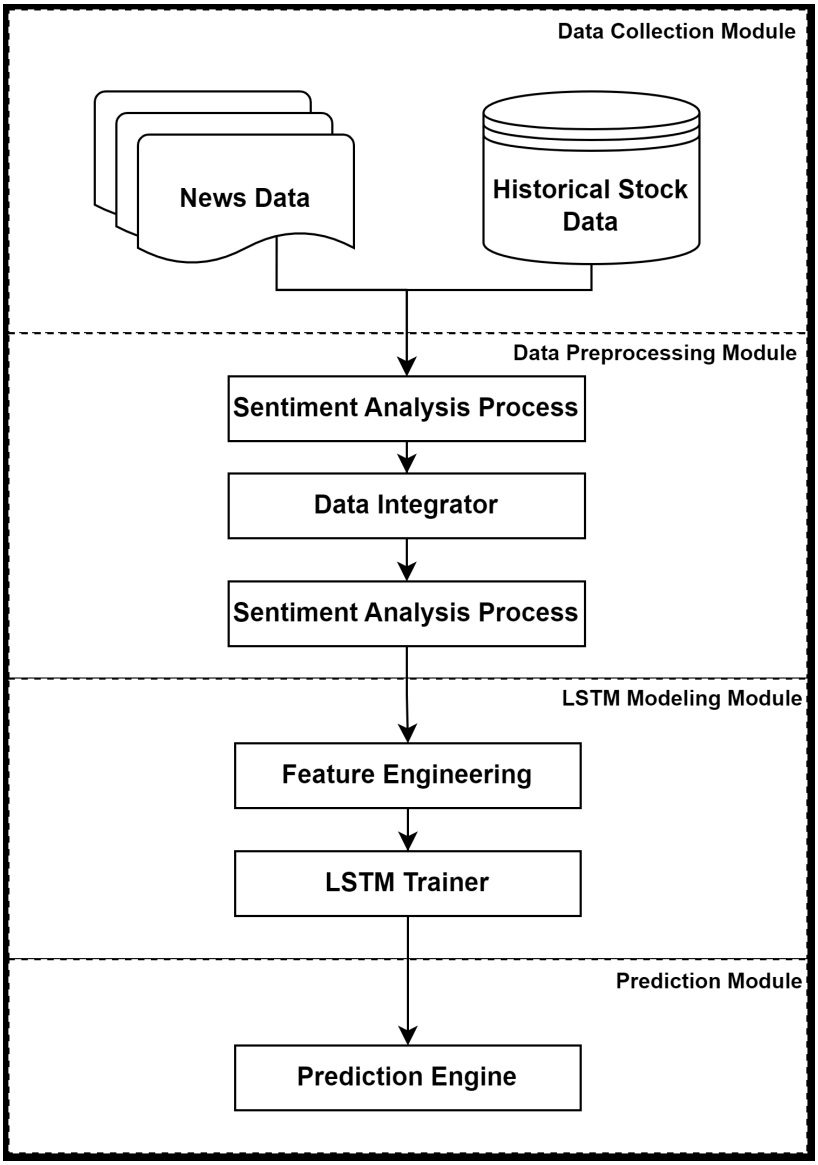

Fig. 1: System Architecture

B. Data Preprocessing

The preprocessing steps involved removing irrelevant data, handling missing values, and filtering out noise from both stock prices and news headlines. Stock price data was normalized, and news headlines were converted into a structured format suitable for sentiment analysis. The final

International Journal of Engineering Applied Sciences and Technology, 2024
Vol. 9, Issue 01, ISSN No. 2455-2143, Pages 141-146
Published Online May 2024 in IJEAST (http://www.ijeast.com)

step was aligning news headlines with corresponding stock prices to create a unified dataset, involving timestamp alignment to match news events with stock market reactions.

C. Model Building

The core of our methodology is the development of a predictive model using LSTM networks. LSTM is chosen for its effectiveness in handling sequential data and capturing temporal dependencies. Feature engineering involves creating input features by combining historical stock prices and sentiment scores. The LSTM model architecture includes input layers, LSTM layers, dropout layers, and output layers.

D. Prediction

The developed application integrates the trained LSTM model and provides a web-based interface for users. Fronted development involves designing and implementing the user interface using web technologies such as HTML, CSS, and JavaScript. Backend development is carried out using Flask to handle user requests and serve predictions. Finally, the frontend is integrated with the backed to enable seamless communication between the end user and the prediction model system. By following this methodology, the aim of the research is to provide a sturdy instrument for estimating stock prices, thereby empowering investors with actionable insights derived from both historical data and market sentiment.

The System architecture is shown in Figure 1 illustrates the overall methodology of the StockSight system, encompassing data collection, processing, model training, and prediction phases.

IV. RESULTS AND DISCUSSION

The StockSight focuses on the five stocks: TCS, Tata Motors, Infosys, Asian Paints, and Tech Mahindra Ltd. For brevity, we present detailed results and discussions for Infosys only. For Infosys, the outcome features two charts: one shows sentiment breakdown and another comparing the training and testing sets' stock price forecasts. Additionally, the accuracy scores for all five stocks are summarized in Table I.

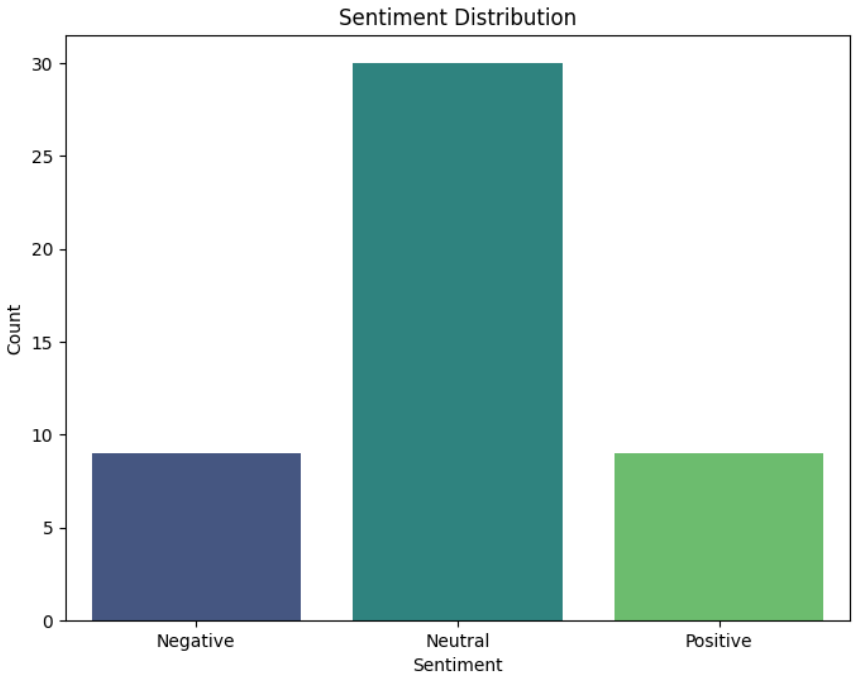

Fig. 2: Sentiment Distribution for Infosys

Figure 2 shows the sentiment distribution for Infosys based on the analyzed news headlines. Sentiments are classified into positive, negative, or neutral categories, providing an brief understanding of the sentiment towards Infosys during the analysis period.

International Journal of Engineering Applied Sciences and Technology, 2024
Vol. 9, Issue 01, ISSN No. 2455-2143, Pages 141-146
Published Online May 2024 in IJEAST (http://www.ijeast.com)

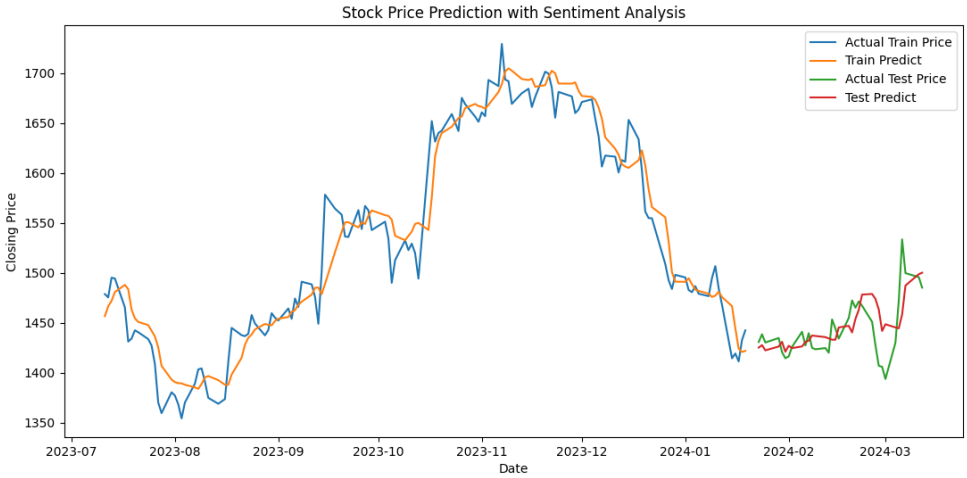

Fig. 3: Stock Prediction (Train vs Test) for Infosys

Figure 3 illustrates the outcome of the LSTM networks model when it is predicting Infosys's prices. The graph com pares the predicted stock prices with the actual stock prices for both the train and test datasets individually, showcasing the model's predictive accuracy. A user-friendly application is developed to provide real-time stock price predictions and in sights to investors. The UI screenshots in Figure 4 demonstrate the successful completion of the research.

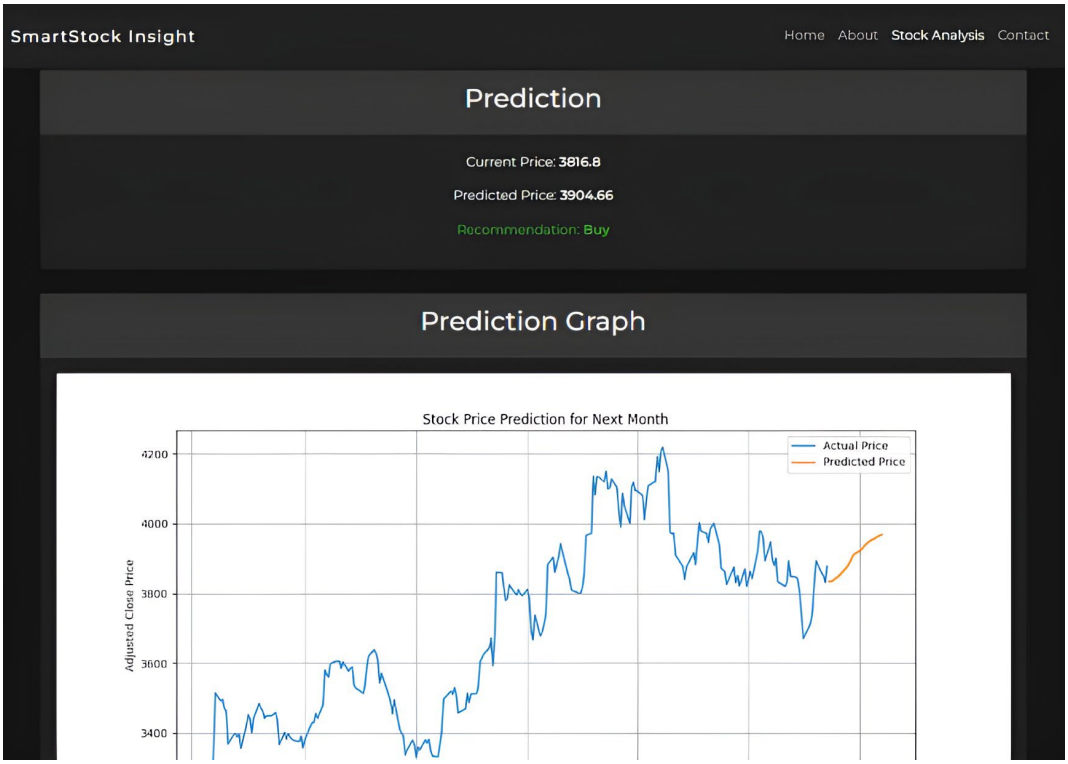

Fig. 4: UI Screenshots of StockSight

Table I provides a summary of the accuracy scores for the predicted prices of all the five analyzed stocks. These scores shows the efficiency of the StockSight research in capturing the patterns in the stock market.

TABLE I: Prediction Accuracy for Different Stocks

Stock	MAE	RMSE
TCS	2.1	2.9
Tata Motors	1.9	2.6
Infosys	2.3	2.9
Asian Paints	1.8	2.5
Tech Mahindra Ltd	2.2	2.8

The results demonstrate that the LSTM Network, enhanced with sentiment analysis, can effectively analyse the stock prices. The detailed analysis for Infosys serves as a representative case, illustrating the model's capability to incorporate market sentiment and historical price data for accurate predictions. Future work could involve further refining the model and extending the analysis beyond the stocks and market conditions.

V. CONCLUSION

StockSight is a novel approach to stock price prediction by integrating sentiment analysis with LSTM networks. The results shows the advantages of this approach in improving prediction accuracy. Future work will focus on expanding the data sources and refining the model to further enhance its predictive capabilities.

VI. FUTURE WORK

Enhancing the model's accuracy through automated processes, incorporating additional data sources, and refining sentiment analysis techniques. These improvements will further refine prediction accuracy and expand the application's capabilities.

REFERENCES

[1]. Derakhshan, A., & Beigy, H. (2019). Sentiment analysis on stock social media for stock price movement prediction. Engineering applications of artificial intelligence, 85, 569-578.

[2]. Jin, Z., Yang, Y., & Liu, Y. (2020). Stock closing price prediction based on sentiment analysis and LSTM. Neural Computing and Applications, 32, 9713-9729.

[3]. Mittal, A., & Goel, A. (2012). Stock prediction using twitter sentiment analysis. Standford University, CS229 (2011 http://cs229. stanford. edu/proj2011/GoelMittal StockMarketPredictionUsingTwitterSentimentAnalysis. pdf), 15, 2352.

[4]. Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. Journal of computational science, 2(1), 1-8.

[5]. Li, Y., Wen, H., Wang, W., Li, X., Yuan, Y., Liu, G., ... & Liu, Y. (2024). Personal llm agents: Insights and survey about the capability, efficiency and security. arXiv preprint arXiv:2401.05459.

[6]. Schumaker, R. P., & Chen, H. (2009). Textual analysis of stock market prediction using breaking financial news: The AZFin text system. ACM Transactions on Information Systems (TOIS), 27(2), 1-19.

[7]. Qasem, M., Thulasiram, R., & Thulasiram, P. (2015, August). Twitter sentiment classification using machine learning techniques for stock markets. In 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI) (pp. 834-840). IEEE.

[8]. Jadhav, R., Sinha, S., Wattamwar, S., & Kosamkar, P. (2021, October). Leveraging Market Sentiment for Stock Price Prediction using GAN. In 2021 2nd Global Conference for Advancement in Technology (GCAT) (pp. 1-6). IEEE.

[9]. Aasi, B., Imtiaz, S. A., Qadeer, H. A., Singarajah, M., & Kashef, R. (2021, April). Stock price prediction using a multivariate multi step LSTM: a sentiment and public engagement analysis model. In 2021 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS) (pp. 1-8). IEEE.

[10]. Soni, P., Tewari, Y., & Krishnan, D. (2022). Machine learning ap proaches in stock price prediction: a systematic review. In Journal of Physics: Conference Series (Vol. 2161, No. 1, p. 012065). IOP Publishing.

[11]. Madeeh, O. D., & Abdullah, H. S. (2021, February). An efficient prediction model based on machine learning techniques for prediction of the stock market. In Journal of Physics: Conference Series (Vol. 1804, No. 1, p. 012008). IOP Publishing.

[12]. Obthong, M., Tantisantiwong, N., Jeamwatthanachai, W., & Wills, G. (2020). A survey on machine learning for stock price prediction: Algorithms and techniques.