Conference Paper

Contemporary Global Perspectives in Management Environment and Technology (CGPMET-2025)

How to Cite:
Garg, D., Pandey, R. K., & Tiwari, P. (2025). Beyond traditional analysis: Leveraging social media sentiments and market data for nifty50 returns forecasting. *International Journal of Economic Perspectives*, *19*(S1), 162–179. Retrieved from https://ijeponline.org/index.php/journal/article/view/924

Beyond traditional analysis: Leveraging social media sentiments and market data for nifty50 returns forecasting

Dr. Deepshi Garg
Assistant Professor, UIM, Uttaranchal University, Dehradun, Uttarakhand, India
Email: deepshigargdd@gmail.com
8439511702

Dr. Rakesh Kumar Pandey
Head of Department, Department of Petroleum & Energy Studies, School of Engineering and Technology, DIT University, Dehradun – 248009, Uttarakhand, India
Email: rks.ismu@gmail.com
8847525494

Dr. Prakash Tiwari
Professor, Faculty of Management Studies, DIT University, Dehradun – 248009, Uttarakhand, India
Email: drprakash.tiwari@dituniversity.edu.in
9412004879

***Abstract*---**Researchers across various fields are increasingly focused on the influence of social media networks. In finance and economics, particular attention has been given to the relationship between social media sentiments and stock market returns. This study introduces a robust, high-performance deep learning prediction model, by combining long short-term memory (LSTM) networks with densely connected neural networks (DCNN), to forecast Nifty50 index movements. Notably, this approach integrates data from multiple social media platforms (Twitter, StockTwits, Facebook, and YouTube) for sentiment analysis, combined with market position data—a novel application in this context. Extensive tuning of the model's hyperparameters achieved an

Corresponding author: *Garg, D., Email: deepshigargdd@gmail.com*
Submitted: 23 November 2024, Revised: 03 March 2025, Accepted: 13 March 2025

accuracy of over 98%, demonstrating the effectiveness of combining social media sentiments with Nifty50 market positions. The resulting model exhibits strong reliability and robustness in forecasting index returns, making it a valuable tool for market forecasting.

***Keywords*---**Social media, sentiment analysis, deep learning, stock returns, prediction model.

1. Introduction

Over the past few years, social media has evolved into an important platform for organizations and individuals to derive meaningful insights from the data available in the form of posts, texts, messages, and comments on social media channels. It also helps the stock market investors to stay up to date about the currents financial trends and allow then to take well informed decisions. Therefore, provide a as a valuable tool for estimating market sentiments (Sul *et al.,* 2017). Furthermore, investors use multiple social media sites like Facebook, YouTube and Stocktwits to gain knowledge about the investment strategies for long- and short-term returns (Sun *et al.,* 2014). Development in the domain of text-mining and sentiment analysis influence observations and opinions from social media. To develop new investment strategies, stock market experts use the enormous volume of unstructured data to categorize sentiments as either positive, negative, or neutral (Malthouse *et al.,* 2013). Modern innovation in the use of social media sentiment for stock market prediction demonstrated significant potential. Research undertaken by fekrazad et al. examined more than 2 million tweets about twenty-five businesses in the S& P100, find out bidirectional relationship between stock performance and social media sentiments (Fekrazad *et al.,* 2022).

Mokhtari et al. examine the impact of Twitter mood on stock market movements. their analysis of the top 5 S&P 500 equities exposed a strong correlation between stock price and Twitter mood call attention to social media impacts on market dynamics (Mokhtari *et al.,* 2023). Additionally, Jin presented graph CNN a graph based deep learning model that forecast indexes like as the S&P 500, NASDAQ, DJI, NYSE and RUSSELL by combined numerous data sources (Jin, 2024). The model highlights the improvements ranging from 4% to 15% compared to standard technique, emphasize the strength of using deep learning approach for stock market prediction.

In past studies, many investors usually waited before the end of the trading day to anlayze the candle stick of the stock market price of the firms they invested in, this allowed then to develop tactics strategies for future transactions (Gong and Sun, 2009). various techniques and methods have been used to predict the movement of stock market (Kim and Han, 2000; Hagenau et al., 2013).

Conventional strategic frameworks such as auto-regression moving average, generalized autoregressive conditional heteroskedasticity, and linear regression have been abundantly used for financial time series forecasting because of their high generalization ability. The feature engineering module plays a vital role in the prediction process (Ding *et al.,* 2015).

In another study, features were extracted from fact-based market data by using technical analysis. The expansion of features was based on assumption that past movements would have an impact on future price movements. technical analysis-based models depend on certain market conclusions and the significance of these conclusion was an important determinant of the model accuracy (shynkevich *et al.,* 2017). Currently, there is huge amount of information about stock market is available, and investors needs to quickly understand it to make better investment strategies. Stock prices keep changing and this makes things more complex due to which a strong robust model is required to analyze and gather the data in a more meaningful way (Dang & Duong, 2016). Predicting stock market trends has become increasingly challenging due to noisy and uncertain data, making it harder to identify profitable investment opportunities. To address this, machine learning algorithms such as random forest, k-nearest neighbors, logistic regression, and naïve Bayes are used to find patterns between stock indicators and future prices. These methods are easier to interpret, require fewer assumptions, and provide stronger learning capabilities compared to traditional statistical techniques (Patel et al., 2015; Sen and Datta, 2021).

Predictive modeling techniques, such as machine learning and deep learning, are highly effective in identifying important patterns within large datasets from various sources. Researchers use these methods to study different viewpoints in areas like business (Xiong et al., 2015) and politics (Pan et al., 2017). In the stock market, machine learning is commonly used to find relationships between different types of data and predict market movements.

They analyze factors like historical prices (Long et al., 2019), technical indicators (Patel et al., 2015), and social media activity (Nguyen et al., 2015; Galvez and Gravano, 2017). This study aims to develop an accurate and reliable deep learning model to predict NIFTY50 index movements by incorporating social media sentiment from multiple sources along with market data. Information was gathered from platforms such as Twitter, Facebook, YouTube, and StockTwits using web scraping tools. The collected data was then cleaned and processed to determine sentiment—classified as positive, negative, or neutral—based on subjectivity and polarity scores.

A total of 20,930 samples were collected for training and testing. The proposed model is built using long short-term memory (LSTM) networks and densely connected neural networks (DCNNs). Through careful selection of model hyperparameters, it achieved an accuracy of over 98%. Other performance metrics, including precision, recall, F1-score, ROC, and AUC scores, further validated the model's reliability in predicting NIFTY50 returns.

2. Literature Review

2.1 Social media sentiments for stock returns

Social media platforms have become essential for networking and content sharing. Predicting stock prices accurately is crucial for business planning, but building a reliable prediction model remains difficult (Garg & Tiwari, 2021).

This study used a logistic regression model to analyze differences in social media engagement between technical and non-technical investors. It also examined how this activity impacted market risk, measured by the VIX index. The findings showed that social media interactions affect stock market movements, with the impact varying based on the investor type. Data for the study was collected from StockTwits.com (López-Cabarcos et al., 2017).

According to research conducted by Sul, tweet sentiments about a particular company from individuals with fewer than one hundred seventy-one followers had a substantial impact on the stock's returns in the next 10 and next 20 days. Interestingly, sentiment in tweets that were neither retweeted nor had more than 171 followers exerted the most significant influence on future stock gains (Sul *et al.*, 2017). Renault examined the relationship between intraday stock returns and online investor sentiments using a vast dataset of approximately sixty million messages sent by online investors over five years, from January 2012 to December 2016. The study aggregated individual StockTwits messages at 30-minute interval to create five different intraday investor sentiment indices (Renault, 2017). Liu explored whether earnings announcements were positive or negative, noting that social media attention remained consistently high. This observation was made using the residual methodology for the attention proxy on Twitter data (Liu *et al.*, 2019). Another study analyses S&P 100 companies using computational linguistics to analyse daily stock-related messages. It discover correlations between tweet sentiment, stock returns, message and trading volume, and disagreement and volatility (Sprenger *et al.*, 2014). Analysing social media provides a unique opportunity to uncover valuable insights into social structure characteristics (Kennedy *et al.*, 2020). It enables both qualitative and quantitative patterns of human behaviour and interaction. Moreover, at times, it empowers us to predict forthcoming human-related events, thus underscoring the profound impact and potential of social media in shaping our understanding of society and enhancing our ability to anticipate future developments.

2.2 Machine learning model for stock market returns

Coyne's research examined stock price prediction using three different machine learning models, utilizing data from StockTwits. The study introduced a model that analyzed social media sentiment to forecast stock prices (Coyne et al., 2018).
Similarly, in 2023, Thara and Sidharth conducted a study on sentiment classification. They compared different classification techniques and found that SVM, polynomial, and RBF algorithms delivered the best performance (Salsabila et al., 2023).

Jain and Dandannavar applied machine learning to analyze sentiment in Twitter data. Their study followed a step-by-step process, starting with data collection and NLP techniques to clean and refine tweets. After identifying key sentiment-related features, they used classifiers like Naive Bayes, SVM, and decision trees to develop a predictive model (Dandannavar, 2016).

Malla Reddy developed a framework to analyze tweets related to the Bombay Stock Exchange (BSE) index using sentiment analysis. The study included text-processing steps such as removing punctuation and stop-words, followed by tokenization. Sentiment classification was carried out using a combination of a text blob and a predefined sentiment lexicon.

Historical stock prices were also factored into the analysis, and a stacked LSTM model with two layers was used for prediction. The final model selected for forecasting was SVM, which predicted daily closing prices based on sentiment data and past stock trends (Venkata Malla Reddy, 2019).

Sun's research delved into the connection between stock price fluctuations and social media sentiment within the Chinese stock market. Data was gathered from multiple online sources, including web forums, chat rooms, and microblogs. The findings suggested that factors such as post length, user activity, and correlation with stock trends played a role in shaping this relationship (Sun et al., 2017). Another study focused on sentiment analysis of Twitter data, revealing that a support vector classifier-based model demonstrated a stronger correlation between market trends and tweet sentiment compared to the Naive Bayes classifier (Li et al., 2018). Meanwhile, Kim et al. utilized a unique dataset from the Korean market, applying the Fama-MacBeth regression model to investigate how investor emotions and trading behavior—both individually and collectively—impacted stock returns (Kim et al., 2019).

2.3 Deep learning for stock market prediction

Traditionally, financial analysts were the go-to experts for stock market predictions. However, with the rise of advanced learning techniques, data scientists have stepped in to tackle the challenges of stock price forecasting. In recent years, computer scientists have leveraged machine learning to improve both the accuracy and efficiency of predictive models. As research in this area has progressed, deep learning has emerged as a powerful tool for refining stock market predictions (Deepshi Garg *et al.,* 2023). To forecast stock prices from the NSE and NYSE, Hiransha explored the effectiveness of four deep learning models: multilayer perceptron (MLP), recurrent neural networks (RNN), long short-term memory (LSTM), and convolutional neural networks (CNN). Their findings revealed that deep learning models, assessed using the mean absolute percentage error (MAPE), performed better than traditional linear models like ARIMA when applied to univariate time series data. Among the tested deep learning approaches, CNN exhibited the highest accuracy, reinforcing the potential of deep learning in stock price forecasting (Hiransha et al., 2018).

In another study, Zou introduced an innovative generative model that combined both textual and numerical price signals for stock market predictions. The research demonstrated the viability of deep generative models by implementing a neural network architecture called StockNet, designed to address stock price forecasting challenges (Zou et al., 2022). Additionally, a separate study focused on LSTM networks for financial time series forecasting, emphasizing their ability to identify intricate dependencies within stock price data, ultimately enhancing prediction accuracy. The integration of sentiment analysis with deep learning has also gained traction, with research by Hiransha demonstrating that incorporating financial news sentiment into deep learning models led to improved stock price predictions (Hiransha et al., 2018).

Further research by Chong assessed the effectiveness of deep learning algorithms for stock market forecasting, with results showing a notable improvement in prediction accuracy. The study suggested that deep neural networks could extract more insights compared to autoregressive models (Chong et al., 2017). Similarly, long investigated a deep neural network model that utilized transaction records and open market data to analyze price fluctuations. Experimental results indicated that bidirectional LSTM outperformed other predictive models in forecasting stock prices, proving to be a reliable tool for financial decision-making (Long et al., 2020). Nemes and Kiss explored the role of Natural Language Processing and sentiment classification, using recurrent neural networks to analyze Twitter data, including comments, hashtags, posts, and tweets related to trending topics during the pandemic (Nemes and Kiss, 2020). In another study, Gangopadhyay and Majumder examined how news reports influence stock market closing prices. Their research evaluated various classification techniques that combined statistical approaches with deep learning, leveraging both price data and textual content to predict stock price movements effectively (Gangopadhyay and Majumder, 2023).

3. Research Objective

The objective of this study is to build a strong and accurate deep learning model to predict NIFTY50 index movements by combining the multiple social media sentiments with NIFTY50 market positions. To address the research gap from existing literature, which mostly relied on the single social media platform for sentiment-based stock predictions, often leading to inaccuracy. By exploring how the investor decision is influenced by the sentiments of multiple social media platforms. ultimately it seeks to add value to existing knowledge while offering investors a data-driven method for making informed decisions based on stock market trends.

4. Research Methodology

This section discusses the data collection, cleaning, pre-processing, preparation, and model paradigm.

4.1. Data collection

Data for the proposed study was extracted from Yahoo finance and four social media platforms, including Twitter, StockTwits, Facebook, and Youtube. This data collection spanned a period of four years, starting from 2018 and continuing through 2021. These platforms were chosen as they regularly generate a multitude of daily feeds related to Nifty50 by various market experts. Therefore, the feeds have been aggregated and aligned day-wise, corresponding to the market return date for the sentiment score calculation. The daily returns based on the closing price of the Nifty50 were taken as the dependent variable to observe the trends (positive or negative) corresponding to derived sentiments from social media platforms and other measures of Nifty50, including open, high, low, adj. close, and volume as the independent variables. The feeds from well-known domain influencers were analyzed based on their participative leadership numbers, average posting frequency, viewing frequency, and the number of subscribers in the financial market. These are the most trustworthy profiles that provide investors instant and reliable news. Moreover, they present helpful insights about the stock market, business, and economy to viewers effectively and efficiently.

A total of 8448 relevant feeds about Nifty50 have been collected from eight recognized users in the financial market using the Twitter API. The StockTwits is another platform where a leading community of investors, traders, and market participants exchange ideas and opinions. In the year 2013, it had more than 230000 active members; by 2020, the number increased to three million; by 2021, the homepage dashboard reflected five million users. The StockTwits data was gathered using the Octoparse web scrapping tool, which is a modern visual web data extraction tool. A total of 8,994 StockTwits feeds have been collected over the four years using relevant tickers in a message, particularly Nifty50.

The Facebook data was gathered via the publicly accessible application Facepager, which enables collecting Facebook status updates, metadata, and engagement metrics from any public page. The Facebook pages selected for comment scrapping are dynamic regarding stock-related news and provide services that target investors and readers of business news who are highly engaged in their investments, particularly CEOs, Equity Analysts, Brokers, Investment Experts, Fund Managers, and retail investors. A total of 5279 posts have been extracted from this platform over four years.

Since its establishment in 2005, YouTube, the widely used video-sharing platform, has undergone substantial enhancements. It is now an excellent resource for individuals interested in investment-related topics. Investors can explore a wide range of channels that provide valuable investment information. Furthermore, fund companies utilize these channels to communicate their investment strategies and options, while renowned value investors share their insights, introducing their philosophies and methodologies. We compiled 240 video scripts corresponding to the given duration from expert market analysts.

4.2. Data cleaning and pre-processing

The text data collected from various sources were cleaned and pre-processed through natural language processing using Python libraries such as Pandas, Numpy, NLTK, and Regex to exclude the noises, including special characters, lowercasing, stopwords, symbols, spaces, punctuations, HTML/XML tags, stemming and lemmatization. The sentiments were attained from the cleaned text data through subjectivity and polarity scores using a hybrid and combination of rule-based and corpus-based classification methods (Liu, B., 2012). The subjectivity score in the range of [0, 1] represents the degree to which a text expresses an opinion. The polarity score measures the positivity, negativity, or neutrality of the opinion within the range of [-1, 1]. The attained sentiments have been assigned numeric values, as shown in **Table 1.**

Figure 1 displays illustrates the frequency distribution of input and output features using histograms. The X-axis represents the scores for various factors, including open, high, low, adjusted close, trade volume, subjectivity score, polarity score, sentiment value, and Nifty50 trends. Meanwhile, the Y-axis indicates the count of samples within each category.

Table 1: Sentiments associated with the polarity score values

Polarity Score	Sentiment	Sentiment Value
[-1, 0)	Negative	-1
[0]	Neutral	0
(0, 1]	Positive	1

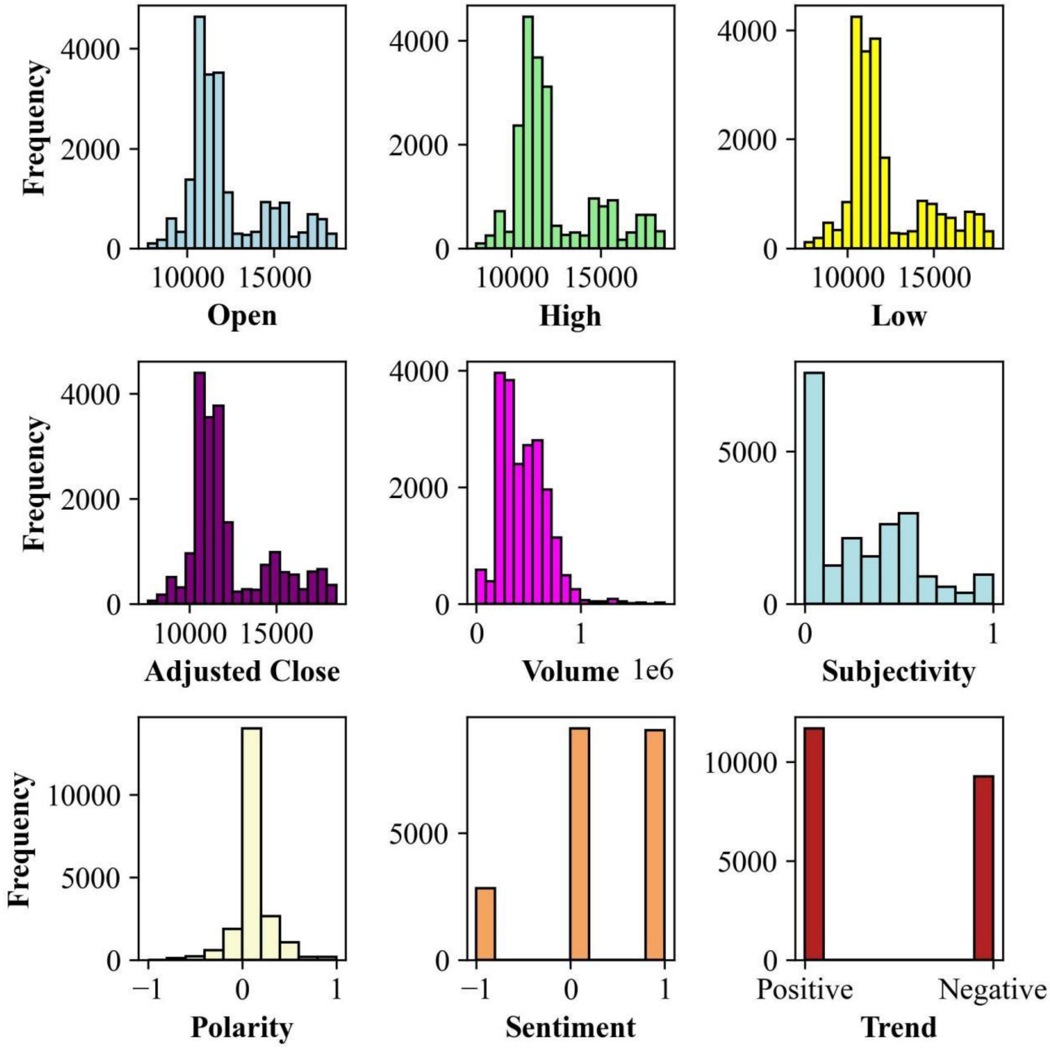


Figure 1: Frequency distribution of input and output features.

4.3 Data preparation

A total of 20930 samples have been accumulated to form a (20930, 9) matrix consisting of open, high, low, adjusted close, trade volume, subjectivity score, polarity score, sentiment value as input features, and Nifty50 returns value as the output feature for the deep-structured classifier. The (20930, 8) input features have been scaled using the z-score normalization, which is mathematically expressed using Equation 1:

$$I_s = (I_a - \mu) / \chi \quad (1)$$

Where I_a and I_s represent the actual and scaled data, the terms μ and x signify the mean and standard deviation of the samples.
The scaled dataset has been split into 80% train set, 10% validation set, and 10% test set. At last, the scaled data was further reshaped as (BS, 8, 1) to make it a suitable input to the LSTM layer. The output labels include two categories denoted as "0" for negative and "1" for the positive trend of the Nifty50 stock.

4.4 Model paradigm

The proposed framework follows a structured flow, where input data along with their corresponding output labels are processed through a predictive model. This model is built using a sequentially stacked architecture consisting of two Long Short-Term Memory (LSTM) layers followed by two Deep Neural Network (DNN) layers. A total of 20,930 pre-processed and labeled samples are fed into the first LSTM layer in batches of 64 samples. LSTMs, a specialized form of Recurrent Neural Networks (RNNs), are designed to efficiently handle sequential data. Unlike standard RNNs, LSTMs incorporate memory cells that retain important information over time. These memory cells are regulated by gates, which manage the flow of information—determining what to store, update, or discard. This gated mechanism enables LSTMs to capture long-term dependencies, making them particularly useful for processing sequential patterns in data (Pandey et al., 2022).

An LSTM memory cell, represented in Figure 2 manipulates, and retains the information through the gates, including forget, input, and output. The input gate controls the flow of information into the memory cell, the forget gate controls which information is to be discarded, and the output gate decides the flow of information out of the memory cell. The first LSTM layer with 64 memory units takes the input features in reverse order and returns the output shape (BS, 8, 64). The second LSTM layer with 32 units has been assigned with dropout and recurrent dropout of 20%. This helped to prevent the LSTM layer from becoming too dependent on any specific input, recurrent connection, or output, which can lead to overfitting. The LSTM layer shown in **figure 2** is followed by a rectified linear unit (ReLU) activated DNN layer with 32 neurons, which accept the input and returns the output of shape (BS, 32). The DNN layer creates dense connections between the input and output layers, allowing the network to learn complex representations from the data (Pandey *et al.*, 2023). Finally, the second DNN layer has a single neuron and sigmoid activation, which generates the final prediction output.

5. Results and Discussion

The developed sequential model's parameters are learned from the train samples, and the validation set tuned the adjustable hyperparameters. The test set has been utilized to assess the generalization ability of the trained model. The model has been trained for 200 iterations. Figure 3 illustrates the use of Binary Cross-Entropy (BCE) loss, accuracy, and the number of training epochs as key metrics for monitoring the training progress and assessing the model's performance. These metrics are crucial in evaluating how well the model is learning and making predictions during the training process. The binary cross-entropy (BCE) and accuracy have been utilized as the loss and metric functions.

$$BCE = (-1/N) \times \sum_{i=1}^{N} \left[Z_j \times log\left(\hat{Z}_j\right) + \left(1 - Z_j\right) \times log\left(1 - \hat{Z}_j\right) \right] \quad (2)$$

$$Accuracy(\%) = 100 \times \frac{Number\ of\ true\ detections}{Number\ of\ total\ detections} \quad (3)$$

where N represents the number of cases; Z_i and $\hat{Z}_i$ are the j^{th} target and predicted values.

The BCE decreased gradually and became constant at values of 0.045 and 0.039 for training and validation data. The accuracy of 98.21 % and 98.61 % have been observed for the train and validation sets. The performance of the trained model has been investigated using the validation and test set. The confusion matrix shown in **Figure 4** reports the predictions of the negative labelled as 0, and the positive labelled as 1, samples.

To address the class imbalance issue between the number of samples with positive and negative trends, additional statistical measures, including accuracy, recall, precision, F1-score, and receiver operating characteristic-area under curve (ROC-AUC) score (Pandey, Mandal and Kumar, 2021), were employed to assess the reliability of the developed model. **Figure 5** displays error bars, which signify the level of uncertainty associated with the metric values at a 95% confidence level. This indicates a 95% probability that the predicted value falls within the specified range. The significantly high scores for these measures suggest that the model is robust for Nifty50 trend predictions. In addition, it has been observed that using the sentiment scores along with the market positions can provide enhanced insights into market behaviour and provide highly reliable index returns.

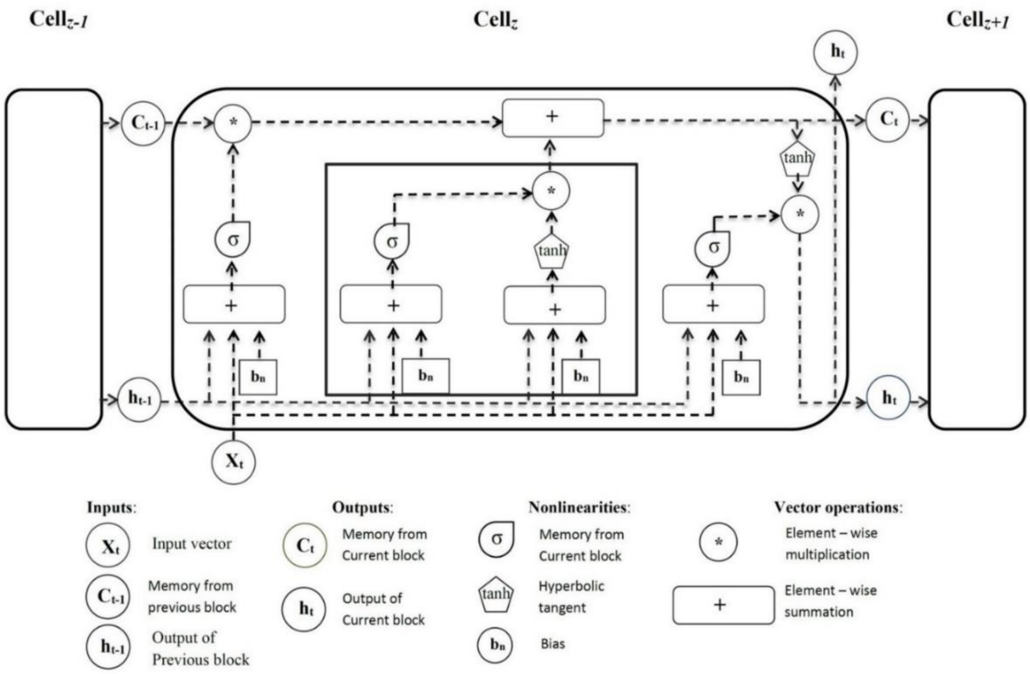


Figure 2: Schematic representation of memory cells in the LSTM layer

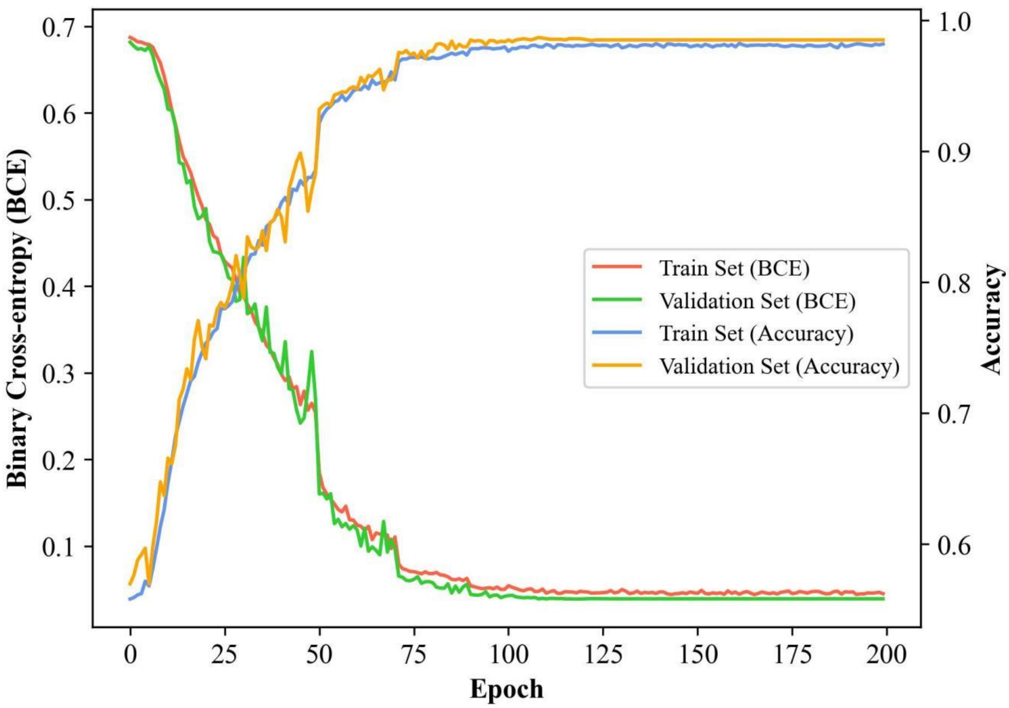
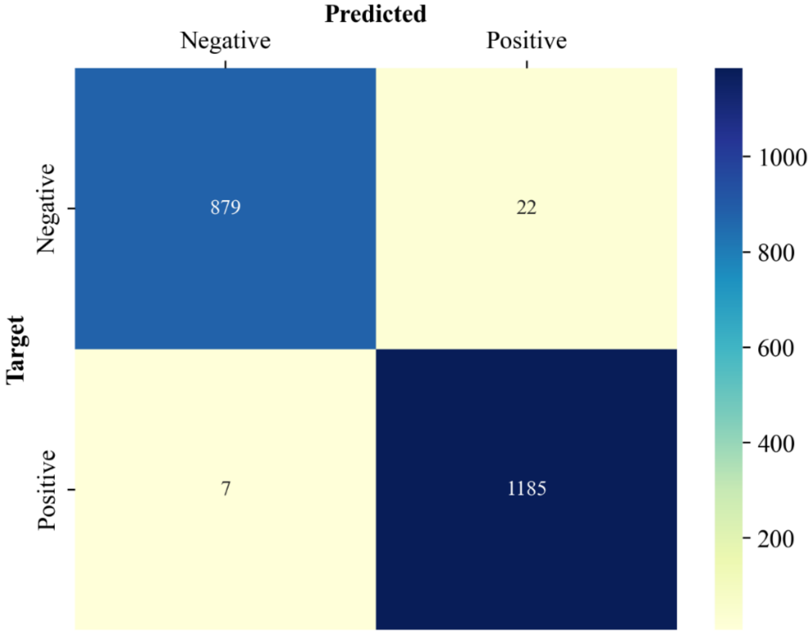
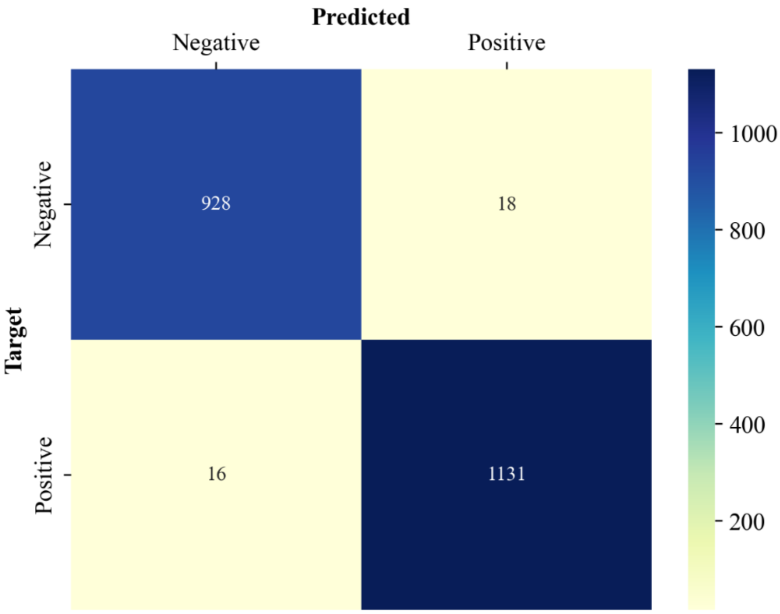


Figure 3: Sequential model performance over train and validation set data



(a)



(b)

Figure 4: Confusion matrix indicating predictions for (a) validation set (Correct predictions: 2064/2093) and (b) test set (Correct predictions: 2059/2093)

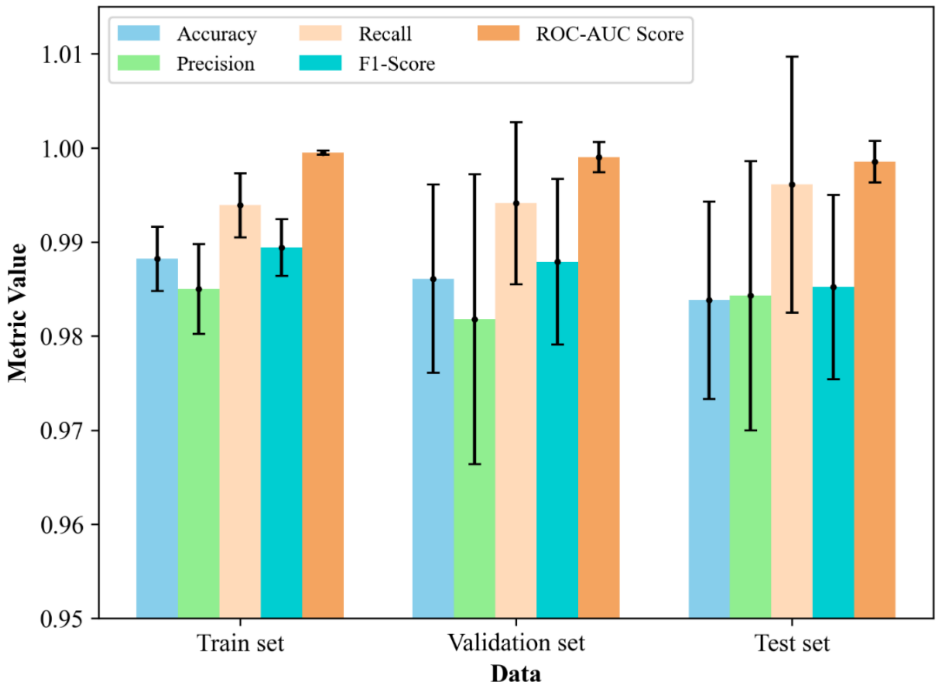


Figure 5: Error bars of the statistical measures at 95% confidence interval

6. Implications

This study has significant implications for various stakeholders in financial and technological sectors. Investors and traders, both retail and institutional, can utilize the developed model to make more informed investment decisions by incorporating social media sentiment analysis alongside traditional market indicators. Financial analysts and market researchers can leverage its predictive capabilities to identify trends and improve forecasting strategies, while algorithmic trading firms can enhance their automated systems by integrating sentiment-driven insights. Hedge funds and portfolio managers can use the model for better risk assessment and portfolio optimization. Additionally, stock market regulators and policymakers may find value in understanding how social media sentiment influences market movements, helping them to design policies to mitigate misinformation risks. FinTech companies developing AI-based trading platforms can adopt similar models to offer advanced predictive tools to their users. Moreover, businesses and corporations listed in stock markets monitor sentiment trends to gauge public perception and strategize accordingly.

7. Scope of the study

This research focuses on the Nifty50 index, which includes the top 50 companies listed on India's National Stock Exchange (NSE). However, the approach can be adapted for sector-specific indices or other global stock markets to test its broader applicability. Future studies can expand by incorporating additional macroeconomic factors, financial news, geopolitical events, and market volatility indices to improve stock market prediction models. Researchers can also explore methods for detecting and reducing sentiment manipulation in financial markets. This could involve developing fake news detection algorithms or using explainable AI (XAI) techniques to ensure transparency in predictions. Beyond textual sentiment analysis, future research can investigate image-based sentiment analysis, video content analysis, and voice sentiment recognition from news broadcasts and investor speeches. These approaches could provide deeper insights into market sentiment and investor behavior.

8. Conclusion

The use of social media as a source of information and sentiment analysis has been growing in popularity, especially in the stock market. This is because investors and traders alike have found that social media sentiment can serve as a valuable indicator of market sentiment and future market trends. Factors like company financials, industry trends, and macroeconomic indicators, should also be considered along with social media sentiment while making investment decisions. This study focused on the Nifty50 index between the sample period of 2018 to 2021 and sentiments from an investor perspective because investors are primarily active on such platforms. A robust, high-performance deep learning model has been developed to predict the Nifty50 index movement using social media sentiments and Nifty50 market positions. The proposed model consisted of LSTM and DCNN networks whose hyperparameters have been carefully tuned, and more than 98% prediction accuracies have been achieved. We find that the hybrid approach combining the market positions with social media sentiments yielded an improved prediction performance. The proposed model is robust and dependable to forecast NIFTY50 returns. In future work, we intend to investigate the performance of the proposed model to examine the returns of the sector-specific market indices.

Declaration of conflicting interests

There are no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

9. References

Aichner, T., Grünfelder, M., Maurer, O., & Jegeni, D. (2021). Twenty-Five Years of Social Media: A Review of Social Media Applications and Definitions from 1994 to 2019. *Cyberpsychology, Behavior, and Social Networking*, 24(4), 215-222. https://doi.org/10.1089/cyber.2020.0134

Castro, J. W., & Luque, A. J. (2022). The Evolution of the Internet and Social Media: A Literature Review on the Opportunities and Challenges in Organizations. *Journal of Business Research*, 143, 563-579. Retrieved from https://www.academia.edu/66740790

Chong, E., Han, C. and Park, F.C. (2017) 'Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies', *Expert Systems with Applications*, 83, pp. 187–205. Available at: https://doi.org/10.1016/j.eswa.2017.04.030.

Cooke, M. and Buckley, N. (2008) 'Web 2.0, Social Networks and the Future of Market Research', *International Journal of Market Research*, 50(2), pp. 267–292. Available at: https://doi.org/10.1177/147078530805000208.

Coyne, S., Madiraju, P. and Coelho, J. (2018) 'Forecasting stock prices using social media analysis', *Proceedings - 2017 IEEE 15th International Conference on Dependable, Autonomic and Secure Computing, 2017 IEEE 15th International Conference on Pervasive Intelligence and Computing, 2017 IEEE 3rd International Conference on Big Data Intelligence and Compu*, 2018-Janua, pp. 1031–1038. Available at: https://doi.org/10.1109/DASC-PICom-DataCom-CyberSciTec.2017.169.

Dandannavar, P. (2016) 'Application of Machine Learning Techniques to Sentiment Analysis', in *Conference: 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT)At: SJBIT, Bengaluru, Kaarnataka, India*. Available at: https://doi.org/10.1109/ICATCCT.2016.7912076.

Dang, L.M. and Duong, D. (2016) 'Improvement methods for stock market prediction using financial news articles', in *2016 3rd National Foundation for Science and Technology Development Conference on Information and Computer Science* , pp. 125–129. Available at: https://doi.org/10.1109/NICS.2016.7725636.

Ding, X. *et al.* (2015) 'Deep Learning for Event-Driven Stock Prediction', in *Twenty-Fourth International Joint Conference on Artificial Intelligence*, pp. 2327–2333.

Fekrazad, A., Harun, S. M., & Sardar, N. (2022). Social media sentiment and the stock market. *Journal of Economics and Finance*, *46*(2), 397–419. https://doi.org/10.1007/s12197-022-09575-x

Gálvez, R.H. and Gravano, A. (2017) 'Assessing the usefulness of online message board mining in automatic stock prediction systems', *Journal of Computational Science*, 19, pp. 43–56. Available at: https://doi.org/https://doi.org/10.1016/j.jocs.2017.01.001.

Gangopadhyay, S. and Majumder, P. (2023) 'Text representation for direction prediction of share market', *Expert Systems with Applications*, 211, p. 118472. Available at: https://doi.org/https://doi.org/10.1016/j.eswa.2022.118472.

Garg, D. and Tiwari, P. (2021) 'Impact of social media sentiments in stock market predictions: A bibliometric analysis', *Business Information Review*, 38(4), pp. 170–182. Available at: https://doi.org/10.1177/02663821211058666.

Gong, J. and Sun, S. (2009) 'A New Approach of Stock Price Trend Prediction Based on Logistic Regression Model', in *New Trends in Information and Service Science, 2009. NISS '09. International Conference on*, pp. 1366–1371. Available at: https://doi.org/10.1109/NISS.2009.267.

Garg, D., Tiwari, P., & Jain, V. K. (2025). Do social media sentiments affect investment decisions? A moderated mediation analysis of the relationship between social media sentiments, trust, and investment decisions. *Global Business and Economics Review*, *32*(1), 67-87.

H. Maqsood, I.M.M.M.M.Y.S.A.F.A. et al. , (2020) 'A local and global event sentiment based efficient stock exchange forecasting using deep learning ', *International Journal of Information Management*, 50, pp. 432–451.

Hagenau, M., Liebmann, M. and Neumann, D. (2013) 'Automated news reading: Stock price prediction based on financial news using context-capturing features', *Decision Support Systems*, 55(3), pp. 685–697. Available at: https://doi.org/https://doi.org/10.1016/j.dss.2013.02.006.

Hiransha, M. *et al.* (2018) 'NSE Stock Market Prediction Using Deep-Learning Models', *Procedia Computer Science*, 132(Iccids), pp. 1351–1362. Available at: https://doi.org/10.1016/j.procs.2018.05.050.

Garg, Deepshi, Pandey, R. K., & Tiwari, P. (2023). A systematic deep learning approach to forecast Nifty50 index trend. *2023 11th International Conference on Intelligent Systems and Embedded Design (ISED)*.

Jin, Y. (2024). GraphCNNpred: A stock market indices prediction using a Graph based deep learning system. In *arXiv [q-fin.CP]*. http://arxiv.org/abs/2407.03760

Kaplan, A.M. and Haenlein, M. (2010) 'Users of the world, unite! The challenges and opportunities of Social Media', *Business Horizons*, 53(1), pp. 59–68. Available at: https://doi.org/10.1016/j.bushor.2009.09.003.

Kennedy, H. *et al.* (2020) 'Using Predictive Analytics to Measure Effectiveness of Social Media Engagement: A Digital Measurement Perspective', *Measurement in sport marketing* [Preprint]

Kim, K. and Han, I. (2000) 'Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index', *Expert Systems with Applications*, 19(2), pp. 125–132. Available at: https://doi.org/https://doi.org/10.1016/S0957-4174(00)00027-0.

Kim, K., Ryu, D. and Yang, H. (2019) 'Investor sentiment, stock returns, and analyst recommendation changes: The KOSPI stock market', *Investment Analysts Journal*, 48(2), pp. 89–101. Available at: https://doi.org/10.1080/10293523.2019.1614758.

Li, M. *et al.* (2018) *Stock market analysis using social networks*, *ACSW '18: Proceedings of the Australasian Computer Science Week Multiconference*. Available at: https://doi.org/10.1145/3167918.3167967.

Liu, B. (2012) *Sentiment Analysis and Opinion Mining*. Morgan & Claypool Publishers.

Liu, H., Wu, L. and Li, X. (Robert) (2019) 'Social Media Envy: How Experience Sharing on Social Networking Sites Drives Millennials' Aspirational Tourism Consumption', *Journal of Travel Research*, 58(3), pp. 355–369. Available at: https://doi.org/10.1177/0047287518761615.

Long, J. *et al.* (2020) 'An integrated framework of deep learning and knowledge graph for prediction of stock price trend: An application in Chinese stock exchange market', *Applied Soft Computing*, 91, p. 106205. Available at: https://doi.org/https://doi.org/10.1016/j.asoc.2020.106205.

Long, W., Lu, Z. and Cui, L. (2019) 'Deep learning-based feature engineering for stock price movement prediction', *Knowledge-Based Systems*, 164, pp. 163–173. Available at: https://doi.org/https://doi.org/10.1016/j.knosys.2018.10.034.

López-Cabarcos, M.Á., Piñeiro-Chousa, J. and Pérez-Pico, A.M. (2017) 'The impact technical and non-technical investors have on the stock market: Evidence from the sentiment extracted from social networks', *Journal of Behavioral and Experimental Finance*, 15, pp. 15–20. Available at: https://doi.org/https://doi.org/10.1016/j.jbef.2017.07.003.

Mokhtari, M., Seraj, A., Saeedi, N., & Karshenas, A. (2023). The impact of Twitter sentiments on stock market trends. In *arXiv [cs.LG]*. http://arxiv.org/abs/2302.07244

M, H. *et al.* (2018) 'NSE Stock Market Prediction Using Deep-Learning Models', *Procedia Computer Science*, 132, pp. 1351–1362. Available at: https://doi.org/https://doi.org/10.1016/j.procs.2018.05.050.

Malthouse, E.C. *et al.* (2013) 'Managing Customer Relationships in the Social Media Era: Introducing the Social CRM House', *Journal of Interactive Marketing*, 27(4), pp. 270–280. Available at: https://doi.org/10.1016/j.intmar.2013.09.008.

Nemes, L. and Kiss, A. (2020) 'Social media sentiment analysis based on COVID-19', *Journal of Information and Telecommunication*, 0(0), pp. 1–15. Available at: https://doi.org/10.1080/24751839.2020.1790793.

Nguyen, T.H., Shirai, K. and Velcin, J. (2015) 'Sentiment analysis on social media for stock movement prediction', *Expert Systems with Applications*, 42(24), pp. 9603–9611. Available at: https://doi.org/https://doi.org/10.1016/j.eswa.2015.07.052.

Pan, Y. *et al.* (2017) 'A multiple support vector machine approach to stock index forecasting with mixed frequency sampling', *Knowledge-Based Systems*, 122, pp. 90–102. Available at: https://doi.org/https://doi.org/10.1016/j.knosys.2017.01.033.

Pandey, R., Mandal, A. and Kumar, A. (2021) 'Identifying Applications of Machine Learning and Data Analytics Based Approaches for Optimization of Upstream Petroleum Operations', *Energy Technology*, 8. Available at: https://doi.org/10.1002/ente.202000749.

Pandey, R.K. *et al.* (2023) 'Genetic Algorithm Optimization of Deep Structured Classifier- Predictor Models for Pressure Transient Analysis', *Journal of Energy Resources Technology, Transactions of the ASME*, 145(2). Available at: https://doi.org/10.1115/1.4054896.

Pandey, R.K., Kumar, A. and Mandal, A. (2022) 'A robust deep structured prediction model for petroleum reservoir characterization using pressure transient test data', *Petroleum Research*, 7(2), pp. 204–219. Available at: https://doi.org/https://doi.org/10.1016/j.ptlrs.2021.09.003.

Patel, J. *et al.* (2015) 'Predicting stock market index using fusion of machine learning techniques', *Expert Systems with Applications*, 42(4), pp. 2162–2172. Available at: https://doi.org/https://doi.org/10.1016/j.eswa.2014.10.031.

Renault, T. (2017) 'Intraday online investor sentiment and return patterns in the U.S. stock market', *Journal of Banking & Finance*, 84, pp. 25–40. Available at: https://doi.org/https://doi.org/10.1016/j.jbankfin.2017.07.002.

Salsabila, S. *et al.* (2023) 'Aspect-based Sentiment and Correlation-based Emotion Detection on Tweets for Understanding Public Opinion of Covid-19', *Journal of Information Systems Engineering and Business Intelligence*, 9, pp. 84–94. Available at: https://doi.org/10.20473/jisebi.9.1.84-94.

Sen, J. and Datta Chaudhuri, T. (2021) 'A Robust Predictive Model for Stock Price Forecasting', in *Conference on Business Analytics and Intelligence* . Available at: https://doi.org/10.36227/techrxiv.16778611.v1.

Shynkevich, Y. *et al.* (2017) 'Forecasting price movements using technical indicators: Investigating the impact of varying input window length', *Neurocomputing*, 264, pp. 71–88. Available at: https://doi.org/https://doi.org/10.1016/j.neucom.2016.11.095.

Sprenger, T. *et al.* (2014) 'Tweets and Trades: the Information Content of Stock Microblogs', *European Financial Management*, 20, pp. 926–957. Available at: https://doi.org/10.1111/j.1468-036X.2013.12007.x.

Sul, H.K., Dennis, A.R. and Yuan, L. (Ivy) (2017) 'Trading on Twitter: Using Social Media Sentiment to Predict Stock Returns', *Decision Sciences*, 48(3), pp. 454–488. Available at: https://doi.org/https://doi.org/10.1111/deci.12229.

Sun, T. *et al.* (2017) 'Predicting Stock Price Returns Using Microblog Sentiment for Chinese Stock Market', *Proceedings - 2017 3rd International Conference on Big Data Computing and Communications, BigCom 2017*, pp. 87–96. Available at: https://doi.org/10.1109/BIGCOM.2017.59.

Sun, X.-Q., Shen, H.-W. and Cheng, X.-Q. (2014) 'Trading Network Predicts Stock Price', *Scientific Reports*, 4(1), p. 3711. Available at: https://doi.org/10.1038/srep03711.

Jishag, A. C., Athira, A. P., Shailaja, M., & Thara, S. (2020). Predicting the stock market behavior using historic data analysis and news sentiment analysis in R. In *First International Conference on Sustainable Technologies for Computational Intelligence: Proceedings of ICTSCI 2019* (pp. 717-728). Springer Singapore.

Venkata Malla Reddy, A. (2019) *Stock market prediction using RNN and sentiment analysis*, *International Journal of Advance Research*. Available at: www.IJARIIT.com.

Xiong, T. *et al.* (2015) 'A combination method for interval forecasting of agricultural commodity futures prices', *Knowledge-Based Systems*, 77, pp. 92–102. Available at: https://doi.org/https://doi.org/10.1016/j.knosys.2015.01.002.

Zou, J. *et al.* (2022) 'Stock Market Prediction via Deep Learning Techniques: A Survey'. Available at: http://arxiv.org/abs/2212.12717.