Prediction of Stock Market Behavior Using Financial News and Sentiment Analysis

Ishitva Awasthi[1], M. Pushpalatha[2]

[1,2] Department of Computing Technologies, School of Computing, SRM Institute of Science and Technology, Kattankulathur, India

email: [1]ia9817@srmist.edu.in, [2]pushpalatha.m@ktr.srmuniv.ac.in

Abstract

Recent advancements in computational power and information management have made it easier to anticipate stock market behavior. Data can be examined and intricate patterns found using sophisticated deep learning methods. Two distinct types of inputs are used by the most recent forecast models: both textual information such as news headlines or news content, as well as numerical information like historical prices and technical signs. But developing text models is required when using written data. Due to issues with word sparsity in datasets, conventional techniques like word embedding might not be appropriate for capturing the substance of financial news. In this study, we apply a deep learning approach to textual and numerical data, including financial URLs for finance news stories, to enhance stock market predictions and sentiment analysis. As input, we take into account market history data, event embedding vectors drawn from news headlines, and a number of technical indicators. Among other metrics, they total the net sentiment for each day and demonstrate that it has a strong ability to forecast future stock market movement. We use LSTM-BERT networks in our prediction technique. Loss MAE and RMSE, which are annualised returns based on trading simulation performance metrics, were derived using the GRU model. When losing, RMSE is 0.04735 as well as Losing MAE is 0.0012.

Keywords: Deep learning, GRU Model, Stock Market Prediction, LSTM.

1. Introduction

Given that stock prediction is a crucial part of a nation's economy, it has always been one of the most difficult issues for economists, statisticians, as well as other financial experts. The stock market seems to be a venue for the distribution, sale, and trading of stocks[1]. Initial Public Offerings (IPOs) on the stock market provide businesses with the opportunity to grow and raise capital. If investors can choose when to purchase and dispose of specific stocks, they can invest in the stock of a variety of businesses and profit. Given the constantly shifting atmosphere of stock values as well as how they vary in accordance with the amount of shares purchased and sold, this specific stock market is incredibly unpredictable. the advertising [2][3]. External factors, including social media and financial news, could have a positive or negative impact on stock prices because the market is affected by regional economies, local economies, national policies, and psychological and human factors. The two main types of analysis are technical analysis, which concentrates on historical price data and uses historical price charts to identify patterns and forecast outcomes, as well as sentiment analysis, . In the method we suggest, track market activity and identify the market's favourable and

unfavourable days. This study's main contribution is the creation of a set of training data for supervised sentiment learning based on these important events, or the moments when the performance of the market takes a sharp turn for the better or worse. We are able to create a classifier and instantly forecast the sentiments of tweets using the training data we have gathered, which includes tweets to positive and negative sentiments[4] Stock price forecasting has drawn attention from a number of academic disciplines, including operations research, finance, and computer programming[5].

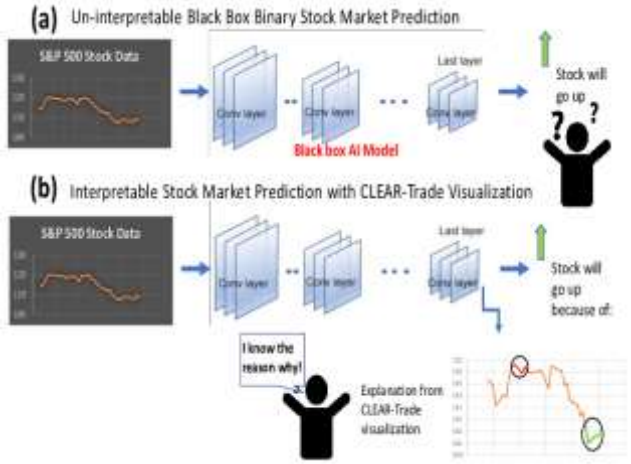

Fig. 1 Stock market Prediction

On the topic of stock prediction, numerous studies are presently underway. To create market prediction models, data scientists first used machine learning algorithms. A computer vision programme was used in earlier studies to predict the stock market using historical, social media network, as well as news data. Long short term memory (LSTM), BERT, and neural networks with artificial intelligence (ANN) have already been used to make stock market predictions. [6][7] Simple trading can't always foresee how different stock markets around the globe will react to a crisis or even other political or financial situations due to how differentially they react to one another. The goal of this study is to create a smart system that can forecast stock market trends by focusing on the perspectives provided in news articles about the financial industry. This study will lower the risk for new investors, as well as experienced traders will gain fresh insights into market movement[8].

2. Related Work

Fazlija 2022 et al. Using data on financial market sentiment gleaned from news articles, one can predict changes in the price of the Standard and also Poor's 500 stock market indicator by using the estimated sentiment value. The best success in sentiment classification is achieved by using contemporary bidirectional encoder representations from transducers (BERT) models. The pre-trained transducer systems are enhanced on a labelled financial text dataset to predict this same sentiment scores of news stories from reliable sources of financial news content. Historical data of a stock market indicator is compared to sentiment ratings produced for the titles of the supplied news articles, as well as for the (text) content of those news stories and the combined title-content consideration. A simple method for predicting the future price movements of such a stock market indicator uses the forecasted sentiment values as features in a random forest classifier. The results show the worth of sentiment scores when used with news content to forecast stock price direction[9].

Darapaneni 2022 et al. Using past prices and the availability of sentiment data forecast future share price movement. LSTM became the initial model in the experiment, with historical prices acting also as the independent variable. For the second part's Random Forest Model, the sentiment analysis obtained with the Intensity Analyzer served as the primary parameter, along with some macroparameters like the price of gold, the price of oil, the USD exchange rate, and the Indian government. In order to increase the model's accuracy, securities yields also were incorporated. In the end, the two models were used to predict the rates of four stocks: Reliance, Hdfc, TCS, as well as SBI. The results were evaluated utilising the RMSE measure[10].

Albahli 2022 et al. It's critical to analyse financial news and comments in order to gauge public opinion, predict how financial markets will act in the future, and calculate both systematic and idiosyncratic risks. In this study, we suggest a novel StockSentiWordNet (SSWN) method to evaluate Twitter posts and Google Finance data to forecast the future activity of a stock markets (one of the major financial markets) over a given time period, including such hourly, daily, weekly, etc. In order to train the advanced learning machine (ELM) as well as recurrent neural net (RNN) to predict stock prices, the suggested SSWN model enriches the SentiWordNet (SWN) lexicon to words that are particularly pertinent to the stock markets. Using the Sentiment140 as well as Twitter datasets, the experiments were effective in achieving an overall accuracy of 86.06%. Results indicate that with us work is generally more precise than cutting-edge methods. In the future, we hope to improve the utility of our approach by incorporating more well-known social networks, such as Google News and Facebook, etc[11].

Rekha 2022 et al. A deep autoencoder, LSTM/GRU levels for prediction, and a lexicon-based method for sentiment analyzing news headlines are all included in the proposed cooperative deep-learning design. The deep learning model is updated using the denoised historical market data and headline mood analysis after the autoencoder has denoised the data. The sentiment score was added before the stock data was sent to an LSTM/GRU algorithm for output prediction. Utilizing recognised metrics which are found in the books, the model's efficacy is evaluated. Results reveal that the LSTM/GRU models perform better than the hybrid model that incorporates both news emotions and deep autoencoding. Additionally, our model's performance is favourable when compared to other contemporary models in the literature[12].

Koukaras 2022 et al. using information from StockTwits as well as SA on Twitter, a model was developed to predict market behaviour. Utilizing market movement and sentiment data, the efficacy of this approach was evaluated before being confirmed with Microsoft stock. Along with comments from Twitter & StockTwits, we gathered financial data from Finance Yahoo. After SA was used to classify tweets, the following seven ML categorization models were chosen: Decision Tree (DT), Naive Bayes (NB), the Support Vector Machine (SVM), Logistic Regression (LR), Randomised Forest (RF), and Multilayer Perceptron (MLP). In order to improve stock prediction accuracy, this study's main innovation involves the combination of different SA and ML techniques with an emphasis on the extraction of extra features from social networking sites (such as public sentiment). The most insightful results came from the study of twitter using VADER & SVM. The highest Area Under Curve (AUC) value was 67%, and the greatest F-score was 76.3%[13]

3. Research Methodology

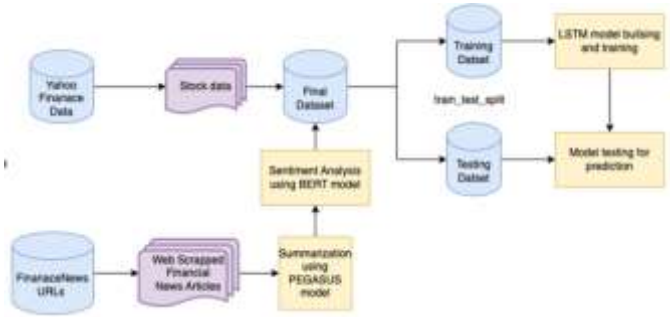

Fig.2 Proposed Model

Stock market predictions is covered in this section. such as the use of a list of predefined mood words, or manually generated training data are methods for sentiment analysis. To create a sentiment classifier, we produce training data automatically. The ability to label as much data as required without the need for any manual labelling is a significant benefit of this approach. The first step in this process is to collect data from the Yahoo dataset, which contains URLs for financial news articles. Next, web scraped financial news articles are used, and the Pegasus model is used to summarise them. The second stage, which also involves data preparation, comprises cleaning and removing information from the data[14]. The third step involves selecting a model, running the research using the LSTM-BERT model for the sentiment Analysis[3], and then dividing the information into testing and training sections.

Data collection

The action of collected from multiple sources information from various sources as well as analysing it to discover trends, probabilities, and other solutions to investigate problems is known as data collection. To learn more, scroll down. the description in the Yahoo Dataset gathered from.https://finance.yahoo.com/quote/WIT?p=WIT&.tsrc=fin-srch&guccounter=1 in which Wipro & USO statistics are featured, along with finance news URLs.

Pre-processing

Pre-processing includes trimming the review's text as well as lemmatizing the dataset as well as eliminating extra columns from it. Unprocessed data needs to be transformed and tidied up as part of a "data preparation" procedure before processing and analysis. The preparation, the first stage in the processing pipeline, entails duties like preparing the information, ensuring its accuracy, and integrating it into additional datasets to add more complexity. Most likely, the dataset that's been initially made accessible for training is not yet suitable to be used[15][16]. The information might be incomplete, contain null numbers, or be disorganised, for instance. Another option is that there are duplicate values in the dataset. Data planning is one method for addressing each of these issues.

Data Splitting

A 70:30 number has been created from the data. 70% of the time is used for teaching, and 30% is used for assessment. Overfitting can be avoided by using the machine learning technique for data splitting (ML). Overfitting occurs when a machine learning algorithm matches the training data so well that it is unable to accurately fit any new data. This situation fits into that category. Before being added to an ML model, this initial data is commonly divided into three to four different subgroups[17]. Common datasets include the training and

testing groups. It is suggested that the data be separated into different categories in order to increase the amount of training data for any specific data collection[18].

Feature Selection

Machine-learning models can only be created using a small subset of such a dataset's parameters; the remainder were unnecessary or redundant. It may be detrimental to the model's general performance and accuracy if we add any of these unnecessary and redundant characteristics to the dataset. [19][20] Finding and selecting the most relevant traits from the data while removing the irrelevant or unimportant ones is essential when using variable selection through computer vision.

Model Selection

The process of choosing a model from a sizable collection of possible models is referred to as a "predictive modelling problem." When selecting a model, one may also take into consideration aspects besides model efficacy, such as complexity, maintainability, and the availability of resources. Assessment and forecasting have been done using the LSTM-BERT algorithm[21][22] Recurrent neural networks include LSTM models as a subclass. They are employed for the purpose of identifying patterns in data sequences, like those found in sensor data, market prices, or natural language. [23][24]The decision of whether to keep prior data in short-term memory as well as discard it is made possible by a unique architecture used by the LSTM model[25]. Language models can be pre-trained using BERT. Pre-training describes how BERT is initially taught on a significant amount of text, like Wikipedia. Afterward, you can use the training findings for additional Natural Language Processing (NLP) activities like sentiment analysis and question answering. Training on the Intelligence Platform and BERT. The LSTM-BERT Model is depicted in fig.3.

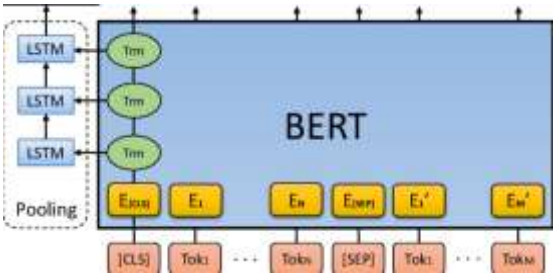

Fig.3 LSTM-BERT Model

Perform EDA

Exploratory data analysis (EDA) is among the techniques used for extracting crucial features as well as trends in use by deep learning and machine learning models throughout Data Science. EDA has consequently become a significant turning point for anyone involved in data science. To fully inform a beginner looking to start a job in data science, this article discusses the idea, significance, tools, and methods of EDA. The article also lists the industries that routinely use EDA to effectively promote their commercial endeavours.

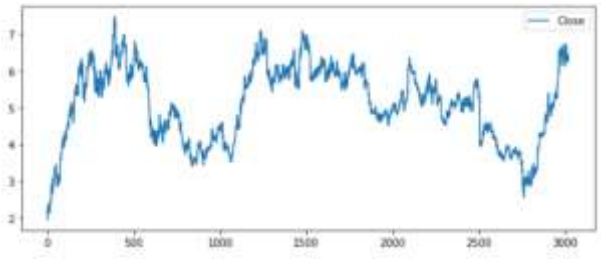

Fig.4 Close of dataset

Figure 4 displays the dataset's Closure for the Stock Market Prediction.

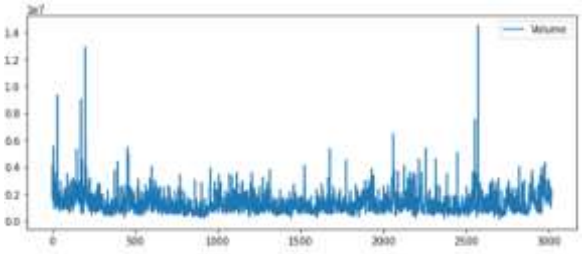

Fig. 5 Volume of dataset

Figure 5 displays the size of a stock market forecast dataset.

Fig.6 Close price of wipro

Figure 6 displays the wipro closure price from the stock market forecast dataset.

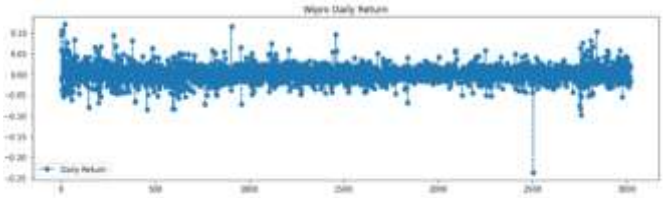

Fig.7 Daily return of wipro

The wipro daily return is shown in Figure 7 from of the stock market forecasting models.

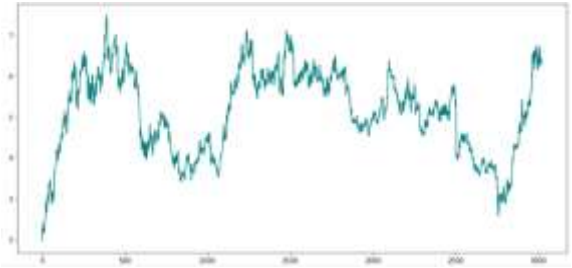

Fig.8 Close price of the data

Its dataset's Close price for such Stock Market Forecast is shown in Figure 8.

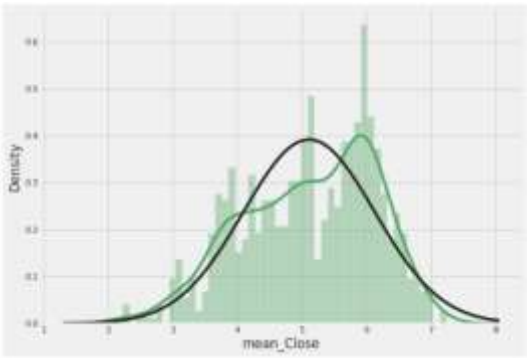

Fig.9 Mean close, density Graph

Figure.9 displays the dataset's mean proximity and density, and the black line's value is 0.4.

4. Result & Discussion

This section examines the effectiveness of the algorithm to use for Yahoo Datasets, which also uses data from Wipro and USO.We assessed the Loss MAE & RMSE measures. An assessment model is used GRU,ADAM optimizer, a deep learning model, was used and implemented over 500 iterations.

Performance Evaluation of Proposed Model LSTM-BERT & Base MOdels SVM,Decision TRee,Logistic Regression

S.No	**Performance evaluation of Model**		
	Model	***Loss MAE***	***RMSE***
1.	BaseModel SVM	0.0435	0.06874
2	Decision Tree	0.0376	0.05763
3	Logistic Regression	0.0366	0.05345
2.	ProposedModel LSTM-BERT	0.022	0.03405

As according Table I's performance assessment of a base model and suggested model, the proposed model, LSTM-BERT, outperforms comparisons of the SVM, DT, and LR. The LSTM-loss BERT's MAE was 0.0012 and its RMSE was 0.04735.

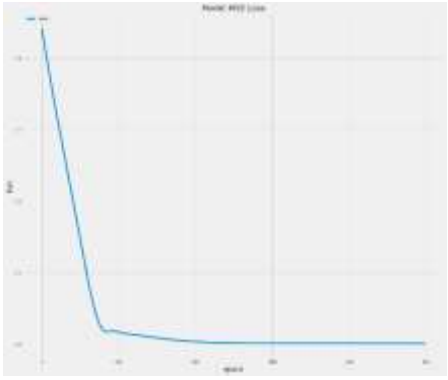

Fig.10 MSE Loss graph

Figure.10 displays the MSE Loss Graph, which displays the outcomes following the model's assessment.

Fig.11 Forecasting Plot

With the aid of yahoo dataset, Figure 11 depicts the stock market forecasting figure.

5. Conclusion

It is now easier to anticipate the behaviour of a stock market thanks to recent advancements in computing power and information management. Patterns inside the data can be looked into and found with the aid of advanced deep learning methods. Two distinct types of inputs are used by the greatest recent forecast models: Along with textual information like news headlines as well as news articles, I also use numerical data like past prices and technical indicators. However, creating text models is necessary when using written material. Due to issues with word sparsity in datasets, traditional techniques like word embedding might not be appropriate for capturing the substance of financial news. To enhance sentiment analysis and stock market predictions, we use textual and numerical data in this study, such as finance URLs for finance news articles. Only a variety of technical analysis are used as input, along with market history data and event embedding vectors drawn from news headlines. Our forecast method makes use of LSTM-BERT architectures. Using the GRU model, the loss MAE as well as RMSE, which are yearly based on trading scenario success measures, were generated. Losing results in RMSE of 0.04735 as well as Losing MAE of 0.0012.

References

[1] X. Li, H. Xie, R. Y. K. Lau, T. L. Wong, and F. L. Wang, “Stock Prediction via Sentimental Transfer Learning,” IEEE Access, vol. 6, pp. 73110–73118, 2018, doi: 10.1109/ACCESS.2018.2881689.

[2] M. R. Bouadjenek, S. Sanner, and G. Wu, “A User-Centric Analysis of Social Media for Stock Market Prediction,” ACM Trans. Web, vol. 1, no. 1, pp. 1–22, 2022, doi: 10.1145/3532856.

[3] J. L. Wu, M. T. Huang, C. S. Yang, and K. H. Liu, “Sentiment analysis of stock markets using a novel dimensional valence–arousal approach,” Soft Comput., vol. 25, no. 6, pp. 4433–4450, 2021, doi: 10.1007/s00500-020-05454-x.

[4] R. S. Deepak, S. I. Uday, and D. D.Malathi, “Machine Lerning Approach in Stock Market Prediction,” vol. 115, no. 8, pp. 71–77, 2017.

[5] X. Li, H. Xie, Y. Song, S. Zhu, Q. Li, and F. L. Wang, “Does Summarization Help Stock Prediction? A News Impact Analysis,” IEEE Intell. Syst., vol. 30, no. 3, pp. 26–34, 2015, doi: 10.1109/MIS.2015.1.

[6] S. Jangir, “Finetuning BERT and XLNet for Sentiment Analysis of Stock Market

Tweets using Mixout and Dropout Regularization," 2021.

[7] C. Gondaliya, A. Patel, and T. Shah, "Sentiment analysis and prediction of Indian stock market amid Covid-19 pandemic," IOP Conf. Ser. Mater. Sci. Eng., vol. 1020, no. 1, 2021, doi: 10.1088/1757-899X/1020/1/012023.

[8] Z. Hu, Z. Wang, S. B. Ho, and A. H. Tan, "Stock Market Trend Forecasting Based on Multiple Textual Features: A Deep Learning Method," Proc. - Int. Conf. Tools with Artif. Intell. ICTAI, vol. 2021-Novem, pp. 1002–1007, 2021, doi: 10.1109/ICTAI52525.2021.00160.

[9] B. Fazlija and P. Harder, "Using Financial News Sentiment for Stock Price Direction Prediction," Mathematics, vol. 10, no. 13, 2022, doi: 10.3390/math10132156.

[10] N. Darapaneni et al., "Stock Price Prediction using Sentiment Analysis and Deep Learning for Indian Markets," pp. 1–15, 2022, [Online]. Available: http://arxiv.org/abs/2204.05783

[11] S. Albahli, A. Irtaza, T. Nazir, A. Mehmood, A. Alkhalifah, and W. Albattah, "A Machine Learning Method for Prediction of Stock Market Using Real-Time Twitter Data," Electron., vol. 11, no. 20, 2022, doi: 10.3390/electronics11203414.

[12] K. Rekha and M. Sabu, "A cooperative deep learning model for stock market prediction using deep autoencoder and sentiment analysis," PeerJ Comput. Sci., vol. 8, p. e1158, 2022, doi: 10.7717/peerj-cs.1158.

[13] P. Koukaras, C. Nousi, and C. Tjortjis, "Stock Market Prediction Using Microblogging Sentiment Analysis and Machine Learning," Telecom, vol. 3, no. 2, pp. 358–378, 2022, doi: 10.3390/telecom3020019.

[14] E. Ris and A. Sjöberg, "Index prediction on the Swedish stock market using natural language processing methods on Swedish news," Master's thesis Mat. Scieces, 2021, [Online]. Available: https://lup.lub.lu.se/student-papers/record/9059706

[15] K. Sharma and S. Member, "An Artificial Intelligence Based Mechanism For Stock Market Analysis And Prediction," Elem. Educ. Online, vol. 20, no. 3, pp. 3182–3188, 2021, doi: 10.17051/ilkonline.2021.03.329.

[16] L. Nemes and A. Kiss, "Prediction of stock values changes using sentiment analysis of stock news headlines," J. Inf. Telecommun., vol. 5, no. 3, pp. 375–394, 2021, doi: 10.1080/24751839.2021.1874252.

[17] M. Al Ridhawi, "Stock Market Prediction Through Sentiment Analysis of Social-Media and Financial Stock Data Using Machine Learning," 2021, [Online]. Available: https://ruor.uottawa.ca/handle/10393/42828

[18] J. J. Duarte, S. Montenegro González, and J. C. Cruz, "Predicting Stock Price Falls Using News Data: Evidence from the Brazilian Market," Comput. Econ., vol. 57, no. 1, pp. 311–340, 2021, doi: 10.1007/s10614-020-10060-y.

[19] P. Koukaras, V. Tsichli, and C. Tjortjis, "Predicting Stock Market Movements with Social Media and Machine Learning," Int. Conf. Web Inf. Syst. Technol. WEBIST - Proc., vol. 2021-Octob, no. Webist, pp. 436–443, 2021, doi: 10.5220/0010712600003058.

[20] X. Li, P. Wu, and W. Wang, "Incorporating stock prices and news sentiments for stock market prediction: A case of Hong Kong," Inf. Process. Manag., vol. 57, no. 5, p. 102212, 2020, doi: 10.1016/j.ipm.2020.102212.

[21] I. Publication, "Use Sentiment Analysis To Predict Future Price Movement in the Stock Market," vol. 11, no. 11, pp. 1123–1130, 2020.

[22] K. Kothari, A. Shah, S. Khara, and H. Prajapati, "a Novel Approach in User Reviews Analysis Using Text Summarization and Sentiment Analysis: Survey," vol. 8, no. 3, p. 3067, 2020, [Online]. Available: www.ijcrt.org

[23] S. Mohan, S. Mullapudi, S. Sammeta, P. Vijayvergia, and D. C. Anastasiu, "Stock price prediction using news sentiment analysis," Proc. - 5th IEEE Int. Conf. Big Data Serv. Appl. BigDataService 2019, Work. Big Data Water Resour. Environ. Hydraul. Eng. Work. Medical, Heal. Using Big Data Technol., pp. 205–208, 2019, doi: 10.1109/BigDataService.2019.00035.

[24] P. Oncharoen and P. Vateekul, "Deep Learning for Stock Market Prediction Using Event Embedding and Technical Indicators," ICAICTA 2018 - 5th Int. Conf. Adv. Informatics Concepts Theory Appl., pp. 19–24, 2018, doi: 10.1109/ICAICTA.2018.8541310.

[25] S. Chaudhari, K. Rajeswari, and S. Vispute, "A Review on using Long-Short Term Memory for Prediction of Stock Price," Int. J. Eng. Res. Technol., vol. 10, no. 11, pp. 251–256, 2021.

[1] X. Li, H. Xie, R. Y. K. Lau, T. L. Wong, and F. L. Wang, "Stock Prediction via Sentimental Transfer Learning," *IEEE Access*, vol. 6, pp. 73110–73118, 2018, doi: 10.1109/ACCESS.2018.2881689.

[2] M. R. Bouadjenek, S. Sanner, and G. Wu, "A User-Centric Analysis of Social Media for Stock Market Prediction," *ACM Trans. Web*, vol. 1, no. 1, pp. 1–22, 2022, doi: 10.1145/3532856.

[3] J. L. Wu, M. T. Huang, C. S. Yang, and K. H. Liu, "Sentiment analysis of stock markets using a novel dimensional valence–arousal approach," *Soft Comput.*, vol. 25, no. 6, pp. 4433–4450, 2021, doi: 10.1007/s00500-020-05454-x.

[4] R. S. Deepak, S. I. Uday, and D. D.Malathi, "Machine Lerning Approach in Stock Market Prediction," vol. 115, no. 8, pp. 71–77, 2017.

[5] X. Li, H. Xie, Y. Song, S. Zhu, Q. Li, and F. L. Wang, "Does Summarization Help Stock Prediction? A News Impact Analysis," *IEEE Intell. Syst.*, vol. 30, no. 3, pp. 26–34, 2015, doi: 10.1109/MIS.2015.1.

[6] S. Jangir, "Finetuning BERT and XLNet for Sentiment Analysis of Stock Market Tweets using Mixout and Dropout Regularization," 2021.

[7] C. Gondaliya, A. Patel, and T. Shah, "Sentiment analysis and prediction of Indian stock market amid Covid-19 pandemic," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1020, no. 1, 2021, doi: 10.1088/1757-899X/1020/1/012023.

[8] Z. Hu, Z. Wang, S. B. Ho, and A. H. Tan, "Stock Market Trend Forecasting Based on Multiple Textual Features: A Deep Learning Method," *Proc. - Int. Conf. Tools with Artif. Intell. ICTAI*, vol. 2021-Novem, pp. 1002–1007, 2021, doi: 10.1109/ICTAI52525.2021.00160.

[9] B. Fazlija and P. Harder, "Using Financial News Sentiment for Stock Price Direction Prediction," *Mathematics*, vol. 10, no. 13, 2022, doi: 10.3390/math10132156.

[10] N. Darapaneni *et al.*, "Stock Price Prediction using Sentiment Analysis and Deep Learning for Indian Markets," pp. 1–15, 2022, [Online]. Available:

http://arxiv.org/abs/2204.05783

[11] S. Albahli, A. Irtaza, T. Nazir, A. Mehmood, A. Alkhalifah, and W. Albattah, "A Machine Learning Method for Prediction of Stock Market Using Real-Time Twitter Data," *Electron.*, vol. 11, no. 20, 2022, doi: 10.3390/electronics11203414.

[12] K. Rekha and M. Sabu, "A cooperative deep learning model for stock market prediction using deep autoencoder and sentiment analysis," *PeerJ Comput. Sci.*, vol. 8, p. e1158, 2022, doi: 10.7717/peerj-cs.1158.

[13] P. Koukaras, C. Nousi, and C. Tjortjis, "Stock Market Prediction Using Microblogging Sentiment Analysis and Machine Learning," *Telecom*, vol. 3, no. 2, pp. 358–378, 2022, doi: 10.3390/telecom3020019.

[14] E. Ris and A. Sjöberg, "Index prediction on the Swedish stock market using natural language processing methods on Swedish news," *Master's thesis Mat. Scieces*, 2021, [Online]. Available: https://lup.lub.lu.se/student-papers/record/9059706

[15] K. Sharma and S. Member, "An Artificial Intelligence Based Mechanism For Stock Market Analysis And Prediction," *Elem. Educ. Online*, vol. 20, no. 3, pp. 3182–3188, 2021, doi: 10.17051/ilkonline.2021.03.329.

[16] L. Nemes and A. Kiss, "Prediction of stock values changes using sentiment analysis of stock news headlines," *J. Inf. Telecommun.*, vol. 5, no. 3, pp. 375–394, 2021, doi: 10.1080/24751839.2021.1874252.

[17] M. Al Ridhawi, "Stock Market Prediction Through Sentiment Analysis of Social-Media and Financial Stock Data Using Machine Learning," 2021, [Online]. Available: https://ruor.uottawa.ca/handle/10393/42828

[18] J. J. Duarte, S. Montenegro González, and J. C. Cruz, "Predicting Stock Price Falls Using News Data: Evidence from the Brazilian Market," *Comput. Econ.*, vol. 57, no. 1, pp. 311–340, 2021, doi: 10.1007/s10614-020-10060-y.

[19] P. Koukaras, V. Tsichli, and C. Tjortjis, "Predicting Stock Market Movements with Social Media and Machine Learning," *Int. Conf. Web Inf. Syst. Technol. WEBIST - Proc.*, vol. 2021-Octob, no. Webist, pp. 436–443, 2021, doi: 10.5220/0010712600003058.

[20] X. Li, P. Wu, and W. Wang, "Incorporating stock prices and news sentiments for stock market prediction: A case of Hong Kong," *Inf. Process. Manag.*, vol. 57, no. 5, p. 102212, 2020, doi: 10.1016/j.ipm.2020.102212.

[21] I. Publication, "Use Sentiment Analysis To Predict Future Price Movement in the Stock Market," vol. 11, no. 11, pp. 1123–1130, 2020.

[22] K. Kothari, A. Shah, S. Khara, and H. Prajapati, "a Novel Approach in User Reviews Analysis Using Text Summarization and Sentiment Analysis: Survey," vol. 8, no. 3, p. 3067, 2020, [Online]. Available: www.ijcrt.org

[23] S. Mohan, S. Mullapudi, S. Sammeta, P. Vijayvergia, and D. C. Anastasiu, "Stock price prediction using news sentiment analysis," *Proc. - 5th IEEE Int. Conf. Big Data Serv. Appl. BigDataService 2019, Work. Big Data Water Resour. Environ. Hydraul. Eng. Work. Medical, Heal. Using Big Data Technol.*, pp. 205–208, 2019, doi: 10.1109/BigDataService.2019.00035.

[24] P. Oncharoen and P. Vateekul, "Deep Learning for Stock Market Prediction Using Event Embedding and Technical Indicators," *ICAICTA 2018 - 5th Int. Conf. Adv.*

Informatics Concepts Theory Appl., pp. 19–24, 2018, doi: 10.1109/ICAICTA.2018.8541310.

[25] S. Chaudhari, K. Rajeswari, and S. Vispute, "A Review on using Long-Short Term Memory for Prediction of Stock Price," *Int. J. Eng. Res. Technol.*, vol. 10, no. 11, pp. 251–256, 2021.