Decision Support Machine- A Hybrid Model for Sentiment Analysis of News Headlines of Stock Market

Kirti Sharma

School of Computer Application,
Lovely Professional University, India
kirtisharmaa.11@gmail.com

Rajni Bhalla

School of Computer Application,
Lovely Professional University, India
rajni.12936@lpu.co.in

Abstract – *Forecasting and making speculations about the financial market is intriguing and enticing for many of us. Predicting sentiments in the field of finance is a difficult thing as there is a special language that is used in financial markets and the data is unlabeled. Generalized models are not sufficient because the words that are used in financial markets have a completely different meaning when compared to their regular use. This paper represents the study of the stock price fluctuations and forecasting of the future stock prices using financial news about the big IT giants. NLP techniques should be applied to extract the correct sentiments out of the statements. This paper proposes a hybrid Machine Learning model DSM i.e. Decision Support Machine based on Support Vector Machine and Decision Tree. In this study news headlines dataset is preprocessed and then used for making predictions. The results show that the proposed model DSM got an accuracy of 79.75%. Results are then compared with the real-time stock market data for the same time duration, thus giving us a better picture of the actual changes. DSM is also compared with BERT, TextBlob, Decision Tree, Naïve Bayes, NLTK-Vader, SVM and KNN. The proposed model can further be extended if more datasets associated with investors' sentiments can be used for training.*

Keywords: *Sentiment Analysis, Stock Prediction, TextBlob, NLTK-Vader, BERT, SVM, KNN*

1. INTRODUCTION

Finding out the positivity or negativity of sentiments by analyzing the relationship between the sequence of words and building a model on it has become a crucial task. The stock market is highly dynamic. Researchers, Market Professionals, and many other people keep on putting their efforts to find new and accurate ways of finding out the movements of the financial market.

In this study we have used various sentimental analysis techniques on financial news, the focus is only on news headings. Sentiment analysis techniques are used to find out the efficacy of headlines on the financial market.

This study is about finding the impact of financial news headlines related to the NSE IT Giants on stock trends. The Stock Market is very much driven by sentiments of the masses, and government policies, it calls for the application of various Natural Language Processing Techniques. In this paper, a hybrid model DSM is proposed which is giving better results in comparison to other sentiment analyses i.e. TextBlob, NLTK-Vader, BERT models along with Naïve Bayes, SVM, K-NN, and Decision Tree models. The main goal is to identify the impact of financial news headlines on the stock market in real time. This study resulted that few of the big news headlines put an impact on the next day's prices.

Data collection forms the basis of research analysis. Dataset is Gathered from reputed and trustworthy financial news based on different IT Giants, followed by the data preparation steps. Various Sentiment Analysis tools are then applied to the dataset. The summary of this study is as follows:

- To identify the effect of a news headline, a hybrid DSM model is developed.

- Two years dataset with news headlines of six IT Giants of NSE is gathered using a python-script, from moneycontrol.com and two years of histor-

ical data is taken from National Stock Exchange's official website

- A comparative study between different models that were used in past on stocks' news and the hybrid proposed model is performed

- Results are plotted to understand the actual effect of the news on the real market.

Many studies have proposed or developed different techniques for identifying the sentiments of news of the financial market. This section highlights some recent work done in this area.

Nemes and Kiss [1] in their paper have shown the impact of news headlines on the stock market by comparing BERT, RNN, and VADER techniques. As per their research, BERT & RNN gave better accuracy without neutral values when compared to other discussed tools. There is a variation in the results of the stock market when various models are applied to the data. RNN model which was trained gave better results in comparison to other SA tools (neutral values are not considered). Kaczmarek and Perez [2] have used random forest on twenty years of stock data to do the monthly forecasting. They have made use of three portfolio-building approaches and made a comparison of these methods - Hierarchical Risk Parity, mean-variance, and 1/N rule, where the first two methods exceed the result over 1/N. Their research has come up with a portfolio-building approach based on RF along with either HRP or mean-variance.

Ingle and Deshmukh [3]used news data for various BSE-listed companies to make forecasting for the next day. Ensemble Deep Learning framework is used for the same. The results gave better accuracy in comparison to other techniques where high, and low values were giving the exact same values. Authors have suggested that high-frequency trading methods in the future can be used to make improvements, but they have not analyzed the effects of the sentiments of the masses.

Khattak, Ali, and Rizvi [4] have analyzed the possible determinants of the European Financial market. They have inspected various possible factors that can influence European market. They have made use of the LASSO technique, their research concluded that Germany and France are the biggest predictors. Devlin et al [5] have developed a new model BERT ("Bidirectional Encoder Representations"). BERT is made in such a way that it uses both left to right and right to left contexts to pretrain the model. Fine-tuning this model is possible just by having one extra output layer. Usmani and Shamsi [6], in their paper, have proposed a news headline classification technique without any training data. They have made use of NLP algorithms to fetch the seed terms, which is then used for refinement of classification outcomes. Data validation is done with ANN and shown with a normalized confusion matrix. They have concluded that when a dataset is classified appropriately, model performance improves significantly. Das et al [7] have presented the ways to make economic judgments i.e. Financial market price judgments, and analysed twitter dossier to make predictions about the share prices of a firm. They have used "Spark, Twitter API, Apache Flume", to stream, process and analyse the data.

Lee [8] analyzed early effects of COVID-19 on people's emotions using DNSI and Google trends on corona-related topics about US Financial Market. The aim is to find the interrelation amongst sentiments and eleven chosen US stocks in a said time period. Positivity and negativity of the public's sentiments about financial market crashes can lead to a compound impact on Stockholders. The analysis shows that different companies get affected by variations in the emotions of the public due to Covid-19.

Arras et al.[9] have shown the modest and useful way of using LRP extended method to LSTM approach by laying down the backpropagation axiom. They have made the extension of LRP to a bi-LSTM to make the sentiment forecasting of a given phrase. Using Neural Networks to find the impact of market attributes on the financial market may not do the forecasting accurately, as the randomly chosen problem's weight may lead to inaccurate forecasting. Based on "WORD-VECTOR" in deep learning Pang et al [10] shows the topic of "STOCK VECTOR". They have taken High-Dimensional past data of various stocks as input to the system and have come up with the idea of LSTM with an ingrained layer and auto-encode technique to do the forecasting. L.Lima et al.[11] in their paper have shown better results in forecasting stock prices, they have worked with Support Vector Machine, and have analyzed all public emotional features. They showed the effect of positive tweets is generally positive on the stock market. K.Herng Leong et al. [12] have created an application that classifies the feedback on the basis of positivity, negativity, and neutrality. Key words were determined and analyzed to find the depth of the emotion. This technique helped in enhancing the standards of the services of the healthcare system. J. Chun et al. [13] have come up with a concept of an emotion-based forecasting system and named it ESPS. Their study centered on taking into account the diverse emotions of investors. ESPS is based on DNN and have considered both positive and negative features i.e., joy, anger, sad, scare, shock, etc. Their study concluded that the ESPS's results are better than the models under consideration. Mujahid et al. [14] have analyzed the sentiments of tweets using TF-IDF and WoW, about online learning during covid pandemic. Their study reveals that Random Forest & SVM along with Bow gave good results. Authors have used various Deep learning techniques like CNN, LSTM, Bi-LSTM, CNN-LSTM and topic-modeling to figure out the issues with online learning. Research study done by G. Aditya et al. [15] have studied and implemented a support vector machine on hotel reviews data. Preprocessing, and conversion of the review are done to figure out the +ve or -ve sentiments.

International Journal of Electrical and Computer Engineering Systems

Literature Survey on stock market trend predictions has drawn our attention to some important points. These points are listed below:

- Most of the research work is conducted with a limited amount of dataset

- A low number of news data instances is used, and a reliable news source is required

- In the case of text data, data collected is unlabeled and that's why there is a need to label the data accurately, which should either be done manually or with the help of some NLP technique

- In order to provide solutions to the above-said problems, a new hybrid model DSM is proposed in this study. The model is pre-processed, trained from the scratch, and then applied to the dataset

Ticker	Top	Date	Time
COFORGE	Wait for Nifty to break above 15,370 before going long: Ashish Biswas of CapitalVia Global Research	21 Feb 2021	8.04 am
COFORGE	Is the current consolidation the right time to go long on Coforge?	10 Feb 2021	8.28 am
COFORGE	Buy Coforge; target of Rs 3051: Prabhudas Lilladher	05 Feb 2021	8.57 pm
COFORGE	Neutral Coforge; target of Rs 2690: Motilal Oswal	02 Feb 2021	9.09 pm
COFORGE	COFORGE LTD. Consolidated December 2020 Net Sales at Rs 1,190.60 crore, up 10.92% Y-o-Y	29 Jan 2021	9.22 am
COFORGE	COFORGE LTD. Standalone December 2020 Net Sales at Rs 612.30 crore, up 3.67% Y-o-Y	29 Jan 2021	9.11 am
COFORGE	Wait for Nifty to break above 15,370 before going long: Ashish Biswas of CapitalVia Global Research	21 Feb 2021	8.04 am
COFORGE	Is the current consolidation the right time to go long on Coforge?	10 Feb 2021	8.28 am
COFORGE	A strong quarter from Coforge, but can the stock rally further?	23 Oct 2020	10.14 am
COFORGE	Why should investors hold on to Coforge despite the stupendous rally?	28 Sep 2020	9.27 am
COFORGE	Buy NIIT Technologies; target of Rs 2200: ICICI Direct	03 Aug 2020	11.37 am
COFORGE	Sell NIIT Technologies; target of Rs 1540: Dolat Capital	30 Jul 2020	7.09 pm
COFORGE	Buy NIIT Technologies; target of Rs 2190: Prabhudas Lilladher	30 Jul 2020	6.26 pm
COFORGE	NIIT Technologies Q1 profit dips 21% at Rs 80 crore	28 Jul 2020	8.15 pm
COFORGE	Hot Stocks | NIIT Tech, Bajaj Auto, Berger Paints three intraday trading ideas	16 Jul 2020	7.40 am
COFORGE	NIIT Technologies – life beyond buyback	01 Jun 2020	9.15 am

Fig. 1. Part from the economics news headlines data frame

2. THE PROPOSED MODEL AND ITS IMPLEMENTATION

Proposed Model DSM works with the data collected from two sources, first from moneycontrol.com, which is a very reputed and reliable source of financial news, and second the historical data from the official website of NSE. Data is scraped using Beautiful-Soup, a python library. The scraping technique was capable enough to fetch the news data for specific symbols for the past two years. This data is then extensively pre-processed with the help of tokenization, data standardization, removal of stop-words and stemming, abbreviation pro-

cessing, and token filtering. The historical data used in this study was available on NSE's official website. This numerical data was also pre-processed with the required normalization and transformation techniques.

The proposed model integrates the feature set of both news and historical dataset. It's joined based on the date. The final feature-set is sentiment, opening price, high, low, and closing price. Then the final step is to do the prediction based on the derived feature set. Data sampling i.e., splitting the dataset into training and testing sets is done. The ratio in which splitting was done is 30.
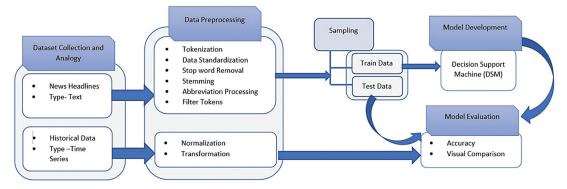
Fig. 2. The overall system architecture of the proposed DSM model

Then hybrid approach for the acquired data set based on the Decision Tree and Support vector machine is applied. Results show that the Decision tree, when applied to our data set gave better classification results for the pos category whereas the Support Vector Machine gave much better accuracy in predicting the neu category. Following these results, when both of these techniques were combined, we achieved much better accuracy. Further, to increase performance, the Tune-sklearn package was used, which is an alternative for Scikit-model Learn's selection module (GridSearchCV, RandomizedSearchCV) with cutting-edge hyperparameter tuning approaches.

The base models for DSM are Support Vector Machine and Decision Tree, which have been optimized further to make predictions with better accuracy. The DSM model is an optimized approach that is trained and tested with well extracted and pre-processed feature set. The accuracy achieved by DSM is 79.75% (Fig. 8.), which is better when compared to the other state-of-the-art models. As shown in Fig. 9., it can be observed that the predictions made by the proposed model for three-class classification achieve better results. It is clear from the table1 that the accuracy of DSM is the highest.

3. METHODOLOGY

3.1. DATA FRAME BUILDING

We have used the financial headlines of the news to do the analysis. Two years of financial News is gathered from 'moneycontrol.com' [16]. Six IT Giants of NSE i.e. COFORGE, HCLTech, Wipro TCS, Tech Mahindra, and Infosys are collected. Dataset is then separated as per the tickers and is stored in CSV. When there is a need to do sentimental analysis, data should be identified as either having positive or negative impacts. As the gathered data was not labeled, part of the data is manually labeled to prepare a labeled training set, which was then used to train the undertaken models. Fig. 1. shows the sample of a raw dataset containing the first few headlines. Part of the news headlines was assigned labels manually to get the best accuracy. Moreover, there was more than one news on the same date, that was also considered and combined based on the date. After that extensive pre-processing i.e. tokenization, standardization, stop-word removal, stemming, and abbreviation processing was done. For historical price data, features i.e. opening price, high, low, the closing price is used. Discretization of this dataset is done. Comparing the values with the previous day's close price, if a value is bigger than the pos, if equal then neu, and in case the value is smaller, the neg value is assigned to the closing price column. Then each row is assigned with a trend value "Up" or "Down", which acts as the input in the proposed model.
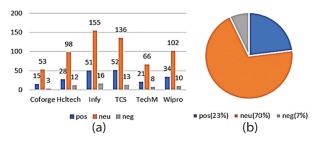
The aim of this study is to do a comparison of undertaken companies' stock prices [17]in the pre-decided period, analyze the sentimental effect of financial news headlines on stock price fluctuations, and determine the actual impact of news headlines. Proposed Model DSM and other sentiment analysis techniques like Text-Blob, NLTK-Vader, and BERT, K-NN is used and detailed results are presented in this paper.
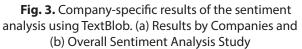
3.2. MODELS TRAINED ON DATASETS

Various pre-existing models i.e. TextBlob, NLTK-Vader, Naïve Bayes BERT, K-NN, SVM, and DT are trained on the manually labeled and preprocessed news dataset. The models are listed below with descriptions and their accuracy.

3.2.1. TextBlob

[18] Python's API to process the text dossier makes it easier to perform NP tasks. PoS tagging, deriving nouns, categorization, conversion, and a lot more are provided by this API. Depending on the subjectivity, the higher the score is lesser objective the statement is. TextBlob is applied to the dataset and it has shown an accuracy of only 25%. Fig. 3. shows the results of Text-Blob on the undertaken dataset. It is visible that the neutral category is the dominating one. Considering INFY, it can be seen that 155 news headings are termed as neutral, 51 positive and only 12 are negative. Fig. 3(b) shows positive, negative, and neutral as 23%, 7%, and 70% respectively. As the neutral category is 70%, the trend is pushing towards positivity, which is leading to distorted results.

Fig. 3. Company-specific results of the sentiment analysis using TextBlob. (a) Results by Companies and (b) Overall Sentiment Analysis Study

3.2.2. Vader

[19][20] "Valence Aware Dictionary and sentiment Reasoner" is a lexicon and rule-based SA, specially signed to work with sentiments of social networking dossier. Values lie between -1 and +1, negative score means negative statement and positive statement means positive statement. To label the data, the same approach as TextBlob is used. VADER, when applied to the dataset, gave an accuracy of 30% which is again not satisfactory, but a little better than TextBlob which is 25 in this case. Fig. 4. shows that the neutral category is fallen down to 57% and there is an increase in the positive values. This has led to an improvement in the prediction by 5%.
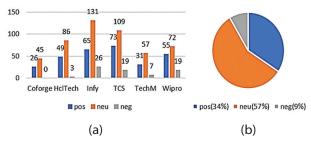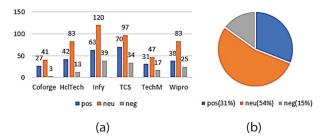
Fig. 4. Company-wise outcome of the sentiment analysis using VADER. (a) Results by Companies and (b) Overall Sentiment Analysis Study

3.2.3. BERT

[5]"Bidirectional Encoder Representations from Transformers" is a pre-trained model in which all words of the statement undertaken are read, and using encoder & decoder models the predictions are made. Text is read by the encoder whereas the decoder makes the predictions. BERT has given an accuracy of 51.36%, which is significantly better than the previous two models. Fig. 5(b) shows that the neu category is reduced to 54%, which is giving a clearer picture of being a sentiment either falls into the pos or neg category. These improved categorizations gave almost double the accuracy of the previous ones.

Fig. 5. Company specific results of the sentiment analysis using BERT. (a) Results by Companies and (b) Overall Sentiment Analysis Study

3.2.4. SVM

Multiclass classification is not native to SVM as it does it in a binary way. To achieve multi-classification, there is a need for mapping points in high dimensions. We have used Python and scikit-learn to do this classification.[21] SVM is a supervised learning method and is useful in both classification and regression. [22] It uses a hyperplane to do the classification. When providing the labeled dataset, it outputs an optimal hyperplane that discriminatively classifies the data. It does the prediction based on(1) the pre-defined categories of input. It uses feature space, which is divided by a hyperplane into vectors.

$$min_{\beta,b,\tau} \frac{1}{2} a^T a + c \sum_{i=1}^{n} \tau_i$$
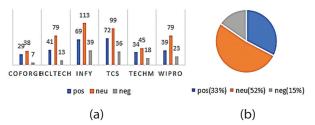$$y_i(a^T \emptyset(x_i) + b) \geq 1 - \tau_i \qquad (1)$$

Where

τ_i=distance to the correct margin with $\tau_i \geq 0, i=1,....,n$
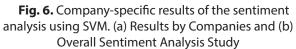
c=regularization parameter,

$a^T a = |(|a^2|)|$=normal vector,

$\emptyset(x_i)$=transformed input space vector

b=bias parameter, y_i=ith target value

Where vector is given as can be used for high dimensional feature space. [11] SVM is applied to this dataset and it achieves an accuracy of 54.12%.

Fig. 6. Company-specific results of the sentiment analysis using SVM. (a) Results by Companies and (b) Overall Sentiment Analysis Study
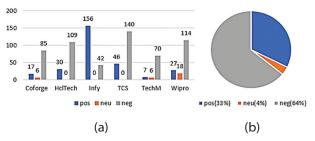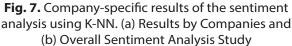
As shown in Fig. 6., SVM is able to identify the positive category, thereby bringing neutral values down and increasing the possibility of more accurate predictions.

3.2.5. K-NN

K-Nearest Neighbour effectively works where there is a need to classify[23]. It does classification by analogy and computes the equality on the basis of Euclidean Distance. To balance the attributes normalization is applied. More unfamiliar data is given more prevalent class out of its nearest neighbours.

Fig. 7. Company-specific results of the sentiment analysis using K-NN. (a) Results by Companies and (b) Overall Sentiment Analysis Study
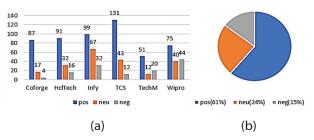
After the application of K-NN on the news dataset, we got an accuracy of 57.58%. K-NN has identified a majority of the news as negative but reduced the neutral category values drastically, which when compared to the actual data set is giving better classification.

3.2.6. DT

[14]In the Decision tree, data collection is done in a tree structure or discrete rules. DT has the ability to work with big data. It does splitting based on rules and classifies the data. [24] Recursion results in the addition of new nodes, which is performed till the time criteria

for predicting values are satisfied. The decision tree achieved an accuracy of 61.36%, which can be considered better as it's the highest amongst all the previously mentioned algorithms. Fig. 8. shows how the positive category has substantially increased, which gives the indication of the market moving in a positive direction.

(a)   (b)

Fig. 8. Company-specific results of the sentiment analysis using DT. (a) Results by Companies and (b) Overall Sentiment Analysis Study

4. RESULTS AND DISCUSSION

This section represents the experimental results to do the predictions of the Stock Market sentiments using DSM and other state-of-the-art models. There are two levels of experimentation, in the first level.
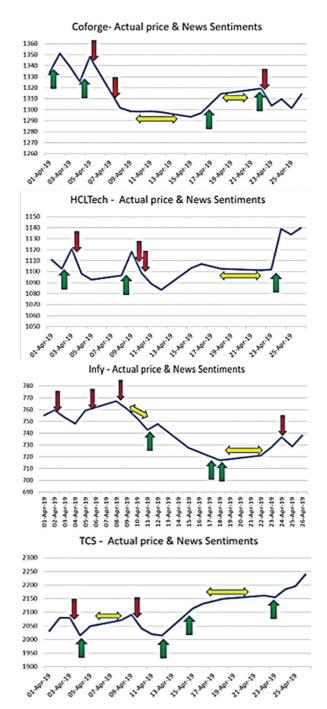
Table 1. Summarised Accuracy Comparison.

Model Name	Accuracy(%)	Recall	Precision	F1 Score
BERT	51.36	0.142	0.142	0.1428
TextBlob	25	0.193	0.46	0.6365
NLTK-VADER	30	0.453	0.492	0.4469
Decision Tree	61.36	0.438	0.609	0.5748
SVM	54.12	0.372	0.556	0.5193
Naïve Bayes	40.91	0.402	0.325	0.3217
K-NN	57.58	0.513	0.602	0.5546
DSM	79.75	0.585	0.7424	0.6467

Data from six NSE -IT giants have been used for both levels. News headlines with the help of listed techniques are classified and labeled in three classes: pos, neu, and neg. The accuracy in % is displayed in table1. The higher accuracy of the DSM model is achieved in comparison to past studies. Other models have not achieved an accuracy of more than 62%, whereas DSM has reached the level of 79.75% accuracy. Fig. 9. shows the actual stock price along with the derived sentiment from the DSM algorithm. It can be observed that the maximum times the indications given by DSM i.e. calls to buy(↑), sell(↓), or hold(↔) a particular stock is correct except a few times, for example when the actual stock price of Infy was in declining phase, sentiments were not giving clear indications of selling, instead it was showing neutral sentiment that means to hold the stock. But overall, a good clear indication can be found, which can help the investors to take decisions on their holdings.

For the purpose of evaluation of DSM with BERT, Text-Blob, NTLK-Vader, Decision Tree, Support Vector Machine, Naïve Bayes, and K-NN, table1 can be referred. The comparison shows that our DSM is outperforming the other models with text data. Further to the combined results, the actual price along with the buy/sell/hold calls indications as represented in Fig. 9. DSM is capable to enhance the trend predictions on the basis of sentiment analysis and historic data.

5. CONCLUSION

Various studies have drawn our attention towards sentimental analysis in today's era, the hybrid approach proposed in this study can be taken as a complemental approach for decision making.
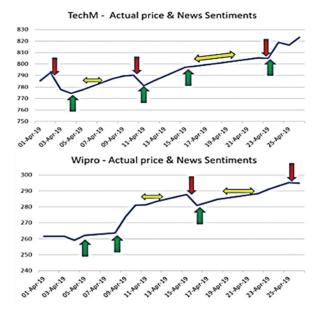
Fig. 9. Buy, sell or hold indications on Coforge, Infy, HclTech, TCS, TechM, and Wipro

The proposed DSM explored the combined effect of news headlines along with historical data in predicting the stock market's trend. The proposed model has shown significant improvement in SM predictions, this study worked on various Sentiment calculation tools to find out the actual sentiment of the undertaken news heading that too without looking into the news details. Sentiments are classified in neg, neu, and pos. The proposed Model DSM works at two levels, first with news headlines and finding polarity and compound scores to classify the headlines in pos, neu and neg. And in the second level, the historical data set is used to find the "Up" and "fall" in the SM for the upcoming day. As shown in Table 1 F1-score of the proposed DSM is 0.6467, which is the best among all the undertaken models. Precision and recall values 0.585 and 0.7424 respectively. Combining both Decision Tree and SVM techniques gave us the advantage of getting better accuracy when it came to pos and neu categories, moreover, the results of neg improved greatly. When applied Tune-sklearn package hyperparameter tuning results improve significantly. The results show that there is a strong co-relation amongst news and SM future trends. These indicators including the full news articles along with annual reports of the companies can also be considered for evaluating the future SM prices.

6. REFERENCES:

[1] L. Nemes, A. Kiss, "Prediction of stock values changes using sentiment analysis of stock news headlines", Journal of Information and Telecommunication, 2021, pp. 1-20.

[2] T. Kaczmarek, K. Perez, "Building portfolios based on machine learning predictions", Economic research - Ekonomska istraživanja, 2021, pp. 1-19.

[3] V. Ingle, S. Deshmukh, "Ensemble deep learning framework for stock market data prediction (EDLF-DP)", Global Transitions Proceedings, Vol. 2, No. 1, 2021, pp. 47-66.

[4] M. A. Khattak, M. Ali, S. A. R. Rizvi, "Predicting the European stock market during COVID-19: A machine learning approach", MethodsX, Vol. 8, 2021, p. 101198.

[5] J. Devlin, M. W. Chang, K. Lee, K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding", Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2019, pp. 4171-4186.

[10] X. Pang, Y. Zhou, P. Wang, W. Lin, V. Chang, "An innovative neural network approach for stock market prediction", Journal of Supercomputing, Vol. 76, No. 3, 2020, pp. 2098-2118.

[11] M. L.Lima et al. "Using Sentiment Analysis for Stock Exchange Prediction", International Journal of Artificial Intelligence Applications, Vol. 7, No. 1, 2016, pp. 59-67.

[13] J. Chun, J. Ahn, Y. Kim, S. Lee, "Using Deep Learning to Develop a Stock Price Prediction Model Based on Individual Investor Emotions", Journal of Behavioral Finance, 2020, pp. 1-10.

[14] M. Mujahid et al. "Sentiment analysis and topic modeling on tweets about online education during covid-19", Applied Sciences, Vol. 11, No. 18, 2021.

[23] S. Hota, S. Pathak, "KNN classifier based approach for multi-class sentiment analysis of twitter data", International Journal of Engineering & Technology, Vol. 7, No. 3, 2018, pp. 1372-1375.

[24] B. Devipriya, Y. Kalpana, "Evaluation of sentiment data using classifier model in rapid miner tool", International Journal of Engineering and Advanced Technology, Vol. 9, No. 1, 2019, pp. 2966-2972.

[6] S. Usmani, J. A. Shamsi, "News Headlines Categorization Scheme for Unlabelled Data", 2020 Int. Conf. Emerg. Trends Smart Technol. ICETST 2020, 2020.

[7] S. Das, R. K. Behera, M. Kumar, S. K. Rath, "Real-Time Sentiment Analysis of Twitter Streaming data for

Stock Prediction", Procedia Computer Science, Vol. 132, 2018, pp. 956-964.

[8] H. S. Lee, "Exploring the initial impact of COVID-19 sentiment on US stock market using big data", Sustainability, Vol. 12, No. 16, 2020.

[9] L. Arras, G. Montavon, K.-R. Müller, W. Samek, "Explaining Recurrent Neural Network Predictions in Sentiment Analysis", 2018, pp. 159-168.

[12] K. H. Leong, D. P. Dahnil, "Classification of Healthcare Service Reviews with Sentiment Analysis to Refine User Satisfaction", International Journal of Electrical and Computer Engineering Systems, Vol. 13, No. 4, 2022, pp. 323-330.

[15] G. Aditya, T. Priyanka, C. Tanupriya, S. Mohammad, "Sentiment Analysis Using Support Vector Machine", 2019.

[20] C. Hutto, E. Gilbert, "VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text", Proceedings of the International AAAI Conference on Web and Social Media, Vol. 8, No. 1, 2014, pp. 216-225.

[21] V. S. Vineeth, H. Kusetogullari, A. Boone, "Forecasting Sales of Truck Components: A Machine Learning Approach", Proceedings of theIEEE 10th International Conference on Intelligent Systems, 2020, pp. 510-516.

[22] M. H. Krishna, K. Rahamathulla, A. Akbar, "A feature based approach for sentiment analysis using SVM and coreference resolution", Proceedings of the International Conference on Inventive Communication and Computational Technologies, 2017, pp. 397-399.

[16] "Business News | Stock and Share Market News | Finance News | Sensex Nifty, NSE, BSE Live IPO News." https://www.moneycontrol.com/ (accessed: 2022)

[17] N. S. Exchange, https://www.nseindia.com/ (accessed: 2022)

[18] TextBlob: Simplified Text Processing — TextBlob 0.16.0 documentation, https://textblob.readthedocs.io/en/dev/ (accessed: 2022)

[19] NLTK :: nltk package, https://www.nltk.org/api/nltk.html (accessed: 2022)