Journalof Cyber Forensics and Security (CYFORS)
Vol.2, No.1 (2025)

Sentiment Analysis Based on Opinion Reviews of Gojek Using the Long Short-Term Memory (LSTM) Method

Muhammad Fahmi [1)], Taufik Iqbal Ramdhani [2)], Agung Budi Prasetio [3)]
[1,2)] Informatics, Institut Teknologi Tangerang Selatan
[3)] Information Technology, Institut Teknologi Tangerang Selatan
email: muhfahmi@students.itts.ac.id [1)], taufik@itts.ac.id [2)], agung@itts.ac.id [3)]

Abstract

Many people are very interested in using online services like GoJek, mainly when using public transportation or traffic jams is challenging. Based on assessments of online transport services, particularly GoJek, we suggest a technique in this study that can identify public attitudes. The 20,000 and 2,500 reviews from motorcycle taxis utilized as the dataset were divided into Positive, Neutral, and Negative categories. To complete this investigation, the LSTM model was used. The LSTM calculation model produced an accuracy value of 0.58% with a dataset of 2,535 positive, negative, and neutral label data; the best accuracy value for positive labels was 75.2%. For the 20,000-data collection, a significantly superior accuracy value of 0.87%, with a positive brand of 91.75%.

Keywords: sentiment analysis, gojek, long-short term memory (LSTM)

I. Introduction

People may now contact others more quickly due to information and communication technology developments. It is advancing faster than ever in the current era, particularly regarding communication gadgets like mobile phones. Today's mobile devices come with various applications that can simplify how we go about our daily lives. A good illustration is the well-liked web application Ojek. This tool is a way to make their daily tasks more accessible, particularly regarding online transportation. One of them is a program made available by the Gojek business.

One of Southeast Asia's top on-demand technology platforms, Gojek gives consumers access to various services, including logistics, payments, food delivery, transportation, and more. This company's fundamental tenet is to use technology to raise people's living standards. The Gojek application was formally introduced in Indonesia in January 2015 after being founded in 2010 with an initial focus on courier services for the delivery of products and two-wheeled transportation [6].

The app allows users to use services such as Go Ride (motorbike ride rental), GoSend (messaging), and GoMart (shopping). Over time, Gojek has transformed into a "Super App," a platform that unites various services under one roof. Gojek connects users with over 2 million registered driver-partners and 500,000 GoFood merchants. The app has been downloaded more than 170 million times across the region. By providing countless experiences in various sectors, Gojek has helped and continues to strive to create added value for society by increasing efficiency, productivity, and financial inclusion [7].

Gojek uses applications on smartphones as an alternate method of locating passengers. There must, however, be both advantages and disadvantages to the development of online motorbike taxis in Indonesia in addition to this occurrence. One benefit is the simplicity with which people can carry out various tasks when they are busy, followed by expanding job prospects, cost reductions, and time and efficiency gains.

Additionally, there are drawbacks, including increasing traffic congestion and conflicts with other modes of public transit, such as Base Ojek and City Transit.

The sentiment analysis or digital text analysis procedure is the main subject of this study. Users of the Gojek program provided review data, which was collected to ascertain whether the message's emotional tone was good, harmful, or neutral. The dataset we used was composed of user reviews, and it was divided into two categories: balanced classification reviews and unbalanced classification reviews, with positive, negative, and neutral values.

Long Short-Term Memory (LSTM) is a Deep Learning model that can be used for sentiment analysis, text translation, and other aspects of Natural Language Processing (NLP). The Recurrent Neural Network (RNN) approach has evolved into LSTM. The vanishing gradient issue in RNN was the motivation behind the development of the LSTM technique [1]. Compared to traditional methods, using LSTM has a better level of accuracy.

II. Literature Reviews

Below is a list of the sentiment analysis and opinion-mining research that is currently available. The sentiment analysis of COVID-19 Tweets using the Long Short-Term Memory (LSTM) technique demonstrated an accuracy rate of 81% and an f-measure of 80% based on the findings of research and testing completed by Rahman et al. [1]. This shows that sentiment can be accurately predicted from COVID-19 Tweets in Indonesian when using LSTM techniques. The limitations of extracting complicated features and identifying pertinent features from short, sparsely populated text input are also overcome using LSTM algorithms. Sentiment analysis utilizing the LSTM approach yields better findings with a higher level of accuracy when compared to the Naive Bayes and RNN methods.

Based on findings from research by Widantoa et al. [2], it was found that using LSTM models with Adam optimizer and batch size 64 and trained for five epochs resulted in the best performance with an accuracy rate of 77.11%. In contrast, LSTM models using RMSprop optimizers with a batch size of 128 and trained for five epochs achieved the best performance with an accuracy rate of 80.07%.

The LSTM model with Adam optimizer showed satisfactory results with a precision level of 76% for negative sentiment and 79% for positive emotion, recall of 79% for negative view and 76% for positive outlook, an F1 score of 78% for negative sentiment and 78% for positive sentiment, and accuracy of 77.11%. On the other hand, LSTM models with RMSprop optimizers provide better results with a precision level of 77% for negative sentiment and 84% for positive emotion, recall of 85% for negative view and 75% for positive sentiment, F1 score of 81% for negative sentiment and 79% for positive sentiment, and accuracy of 80.07%.

Based on research conducted by Windasari et al. [3] presented a system designed to identify public sentiment based on Twitter posts related to GoJek, the conclusion can be drawn that sentiment analysis of posts about GoJek can predict sentiment towards the service. This study used the n-gram unigram and TF-IDF feature extraction approaches and SVM algorithms to perform classification. The results showed an accuracy rate of 86% in predicting the sentiment of the analyzed tweets. The study recommends conducting experiments involving more data and comparing different classification methods and features to improve the system's accuracy.

Research [4] proposes a method of analyzing hotel review sentiment using Latent Dirichlet Allocation (LDA), Semantic Similarity, and Long-short Term Memory (LSTM). Such plans involve preprocessing reviews, determining hidden topics using LDA, categorizing studies into hotel aspects using Semantic Similarity, and classifying sentiment using LSTM. The results suggest that the proposed method can accurately organize reviews into hotel aspects and determine their opinion. The comfort aspect received more negative views than other aspects.

This study [5] focuses on user reviews of Gojek and Grab applications on the Google Play Store. The results showed that the SVM algorithm worked well in classifying review sentiment, with the Gojek app achieving higher performance scores than the Grab app. The Gojek application achieved a higher performance score with an accuracy score of 89%, precision of 94%, recall of 86%, and an F1 score of 90%. At the same time, the Grab application has an accuracy score of 87%, precision of 94%, recall of 85%, and F1-score of 89%. This study emphasizes the importance of sentiment analysis in understanding user opinions and improving the quality of ride-hailing services.

III. Methodology

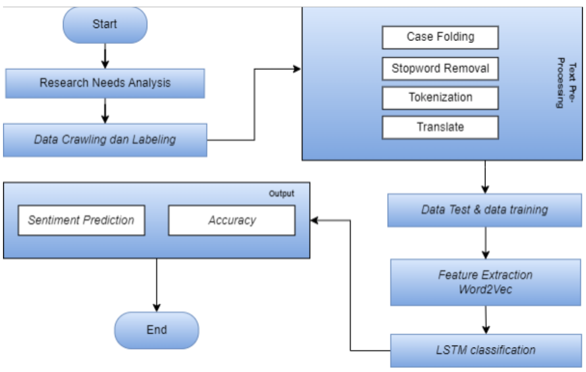

Figure 1. Research Flow

In this study, user review sentiment analysis of the GoJek application was obtained through several processes, namely data crawling & manual labeling for training datasets, preprocessing text, feature extraction, and LSTM classification. The research flow can be seen in Figure 1.

III. Research Method

A. Needs Analysis

Needs are related to selecting research topics, how much data will be used, and the need for data collection.

B. Data Crawling dan Labeling

The dataset was obtained from Gojek application user review data from January-March 2024, with a total number of reviews received of 20,000 review data. There were 12,914 positive reviews, 6,241 negative reviews, and 845 neutral reviews. The labeling process is based on the Gojek user rating, where ratings 1-2 are included in the negative label, rating 3 is in the neutral category, and ratings 4-5 are in the Positive title.

Table 1. Dataset

No.	Data	Class Sentiment		
		Positive	Negative	Neutral
1.	20.000	12914	6241	845
2.	2.500	845	845	845

With the data gap between Positive, Negative, and Neutral reviews, the author divided into dataset 2 with the same data composition with a total of 2,535 review data, namely 845 positive review data, 845 negative thoughts, and 845 neutral reviews.

C. Text Preprocessing

Text preprocessing is preparing initial text data into data that is easier to further process into basic text form and remove noise. This stage aims to get more optimal calculation results. This study's text preprocessing steps are converting emoticons, cleansing, case folding, stemming, convert negation, stopword removal, and tokenizing.

D. Data Training dan Testing

Table 2. Split Training dan Testing Data Dataset A

No.	Label	Class Sentiment		
		Positive	Negative	Neutral
1.	Label Test	2583	1249	169
2.	Label Train	10331	4992	676

Table 3. Split Training dan Testing Data Dataset B

No.	Label	Class Sentiment		
		Positive	Negative	Neutral
1.	Label Test	254	254	254
2.	Label Train	591	591	591

Text preprocessing is preparing initial text data into data that is easier to further process into basic text form and remove noise. This stage aims to get more optimal calculation results. This study's text preprocessing steps are converting emoticons, cleansing, case folding, stemming, convert negation, stopword removal, and tokenizing.

E. Word2vec Model Creation

Creating a word2vec model helps make every word in our data into a vector. The number of features can specify the vector in the word2vec model. In this study, a word2vec model was made with 100 parts, 1000 Embedding dimensions.

Table 3. Word2vec Process

Model Word2Vec	
Embedding Matrix Shape:	(10000, 1000)

Embedding is used in a model's Embedding layer to embed a token into its vector representation, which contains information about the ticket or word. Vocabulary is obtained from the tokenizer and the corresponding vector of the Embedding, Word2Vec model.

F. Model LSTM

At this stage of LSTM classification, the test data that will be entered into the LSTM classification method must be numbers so that LSTM can process the test data. Itself. After that, the test data will be tested using the LSTM method.

Long Short-Term Memory (LSTM) is one method in Deep Learning that can be used for Natural Language Processing (NLP), such as speech recognition, text translation, and sentiment analysis. LSTM is a development of the Recurrent Neural Network (RNN) method; this LSTM method was created to solve the vanishing gradient problem in RNN. Research (Hassan &; Mahmood, 2017) and (Wang &; Liu, 2018) proves that using LSTM has a higher level of accuracy than conventional methods.

IV. Result and Discussion

A. Dataset Preprocessing

Long Short-Term Memory (LSTM) is one method in Deep Learning that can be used for Natural Language Processing (NLP), such as speech recognition, text translation, and sentiment analysis. LSTM is a development of the Recurrent Neural Network (RNN) method; this LSTM method was created to solve the vanishing gradient problem in RNN. Research (Hassan &; Mahmood, 2017) and (Wang &; Liu, 2018) proves that using LSTM has a higher level of accuracy than conventional methods.

B. Word2Vec Model

```
[ ] embedding_matrix = np.zeros((vocab_length, Embedding_dimensions))
    for word, token in tokenizer.word_index.items():
        if word2vec_model.wv.__contains__(word):
            embedding_matrix[token] = word2vec_model.wv.__getitem__(word)

    print("Embedding Matrix Shape:", embedding_matrix.shape)

    Embedding Matrix Shape: (10000, 1000)
```

Figure 2. Embedding Matrix Shape

The value of the matrix size embediding_dimensions to create an embedding matrix by filling it with words from the word2vec model. Then every word in the tokenizer vocabulary is checked against the word2vec model.

C. Model LSTM

```
Layer (type)                 Output Shape              Param #
=================================================================
embedding_1 (Embedding)      (None, 100, 1000)         10000000

lstm_2 (LSTM)                (None, 100, 128)          578048

dropout_2 (Dropout)          (None, 100, 128)          0

lstm_3 (LSTM)                (None, 128)               131584

batch_normalization_1 (Batc  (None, 128)               512
hNormalization)

dense_2 (Dense)              (None, 64)                8256

dropout_3 (Dropout)          (None, 64)                0

dense_3 (Dense)              (None, 3)                 195

=================================================================
Total params: 10,718,595
Trainable params: 718,339
Non-trainable params: 10,000,256
```

Figure 3. LSTM Architecture

A model created for word order classification tasks is represented as an index with 10,718,595 parameters, 718,339 of which are trainable and 10,000,256 of which are not. A dropout layer is added to decrease overfitting, and batch normalization promotes quicker convergence during training.

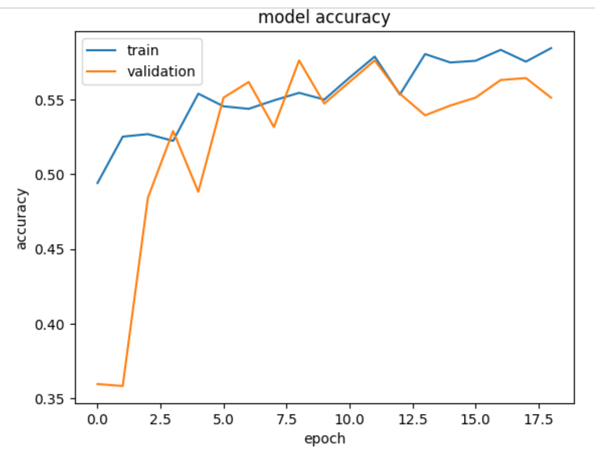

Figure 4. Dataset A Accuracy Model: 2.535 Data

A high accuracy value with positive, negative, and neutral sentiment aspects achieves excellent learning polarity values in training and data validation.

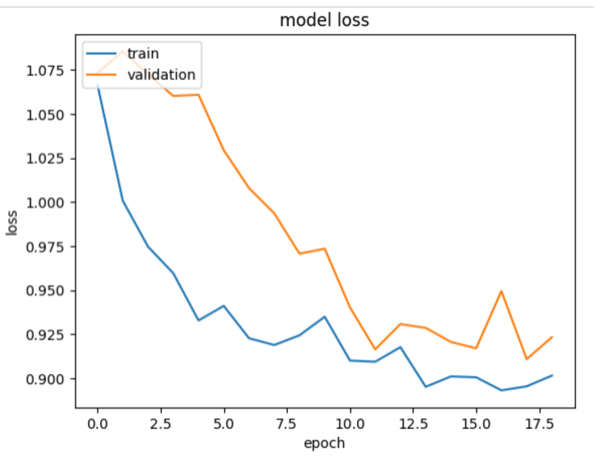

Figure 5. Dataset A Loss Model: 2.535 Data

Has a loss value that can be used for LSTM training and validation data analysis models reaching a value of 17.5 epochs.

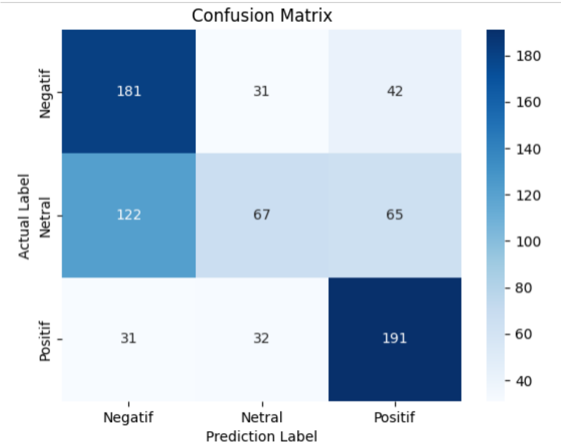

Figure 6. Confusion Matrix Dataset A

Each label class's values have all been sorted in the same order, totaling 254 for each matrix variable used.

	precision	recall	f1-score	support
0	0.54	0.71	0.62	254
1	0.52	0.26	0.35	254
2	0.64	0.75	0.69	254
accuracy			0.58	762
macro avg	0.57	0.58	0.55	762
weighted avg	0.57	0.58	0.55	762

Figure 7. Evaluation Matrix

Able to provide good accuracy scores from the same number of classes.

Table 4. Accuracy Value

Label Accuracy Value		
Negative	Neutral	Positive
71.26 %	26.38 %	75.2 %

The accuracy findings demonstrate that the LSTM approach can classify Gojek sentiment with a high level of accuracy, demonstrating that this method is the best algorithm to employ to analyze Gojek's assessment sentiment toward a service. Online transportation service businesses can use Gojek's sentiment analysis to see how the general public feels about their offerings.

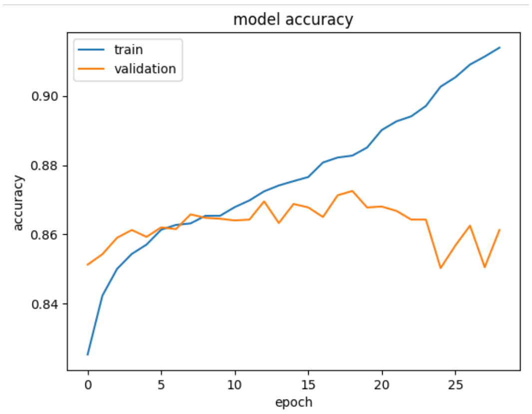

Figure 7. Dataset B Accuracy Model: 20.000 Data

With many datasets and an increasing number of trained epochs, it can provide a much better accuracy value than training and validation data.

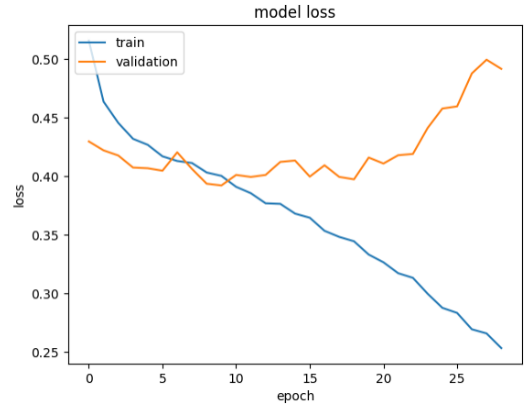

Figure 8. Dataset B Loss Model: 20.000 Data

Reaches a loss value of 25 epochs and can be utilized to train data analysis models and validate LSTM.

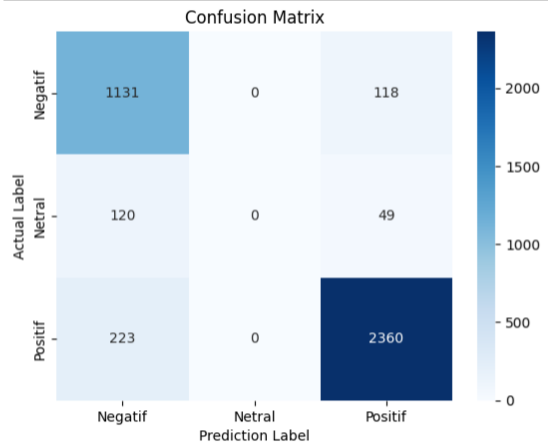

Figure 9. Confusion Matrix Dataset B

The confusion matrix does not display the value since there are an uneven number of neutral label prediction values from negative and positive labels; nevertheless, all the actual label values can appear.

	precision	recall	f1-score	support
0	0.77	0.91	0.83	1249
1	0.00	0.00	0.00	169
2	0.93	0.91	0.92	2583
accuracy			0.87	4001
macro avg	0.57	0.61	0.58	4001
weighted avg	0.84	0.87	0.86	4001

Figure 10. Evaluation Matrix

Enables a broader range of accuracy and recall values for the classification process using the LSTM model.

Table 5. Accuracy Value

Label Accuracy Value		
Negative	Neutral	Positive
87.83 %	-	91.75 %

Online transportation business operators can use Gojek sentiment analysis to discover how customers feel about their offerings. According to the accuracy results, the LSTM technique can categorize Gojek sentiment with a high degree of accuracy, demonstrating that this algorithm is the best to employ when examining Gojek's assessment sentiment regarding a service.

CONCLUSION

Containing 2,535 datasets containing positive, negative, and neutral labels, the accuracy value obtained from the LSTM calculation model yields a precision of 0.58%. Therefore, 75.2% is the best accuracy value from the positive label. 20,000-dataset A dataset has an accuracy value of 0.87% and a positive label of 91.75%, which is significantly superior. To ensure that the modeling's value can produce the best results, you can employ various techniques and label increasingly complicated data classes in future studies.

REFERENCES

[1] Rahman, M. Z., Sari, Y. A., & Yudistira, N. (2021). Analisis Sentimen Tweet COVID-19 menggunakan Word Embedding dan Metode Long Short-Term Memory (LSTM). Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer, 5(11), 5120-5127.

[2] Witantoa, K. S., ERa, N. A. S., Karyawatia, A. E., Arya, I. G. A. G., Kadyanana, I., & Astutia, L. G. Implementasi LSTM pada Analisis Sentimen Review Film Menggunakan Adam dan RMSprop Optimizer. Jurnal Elektronik Ilmu Komputer Udayana p-ISSN, 2301, 5373.

[3] Windasari, I. P., Uzzi, F. N., & Satoto, K. I. (2017, October). Sentiment analysis on Twitter posts: An analysis of positive or negative opinions on GoJek. In 2017 4th international conference on information technology, computer, and electrical engineering (ICITACEE) (pp. 266-269). IEEE.

[4] Priyantina, R. A., & Sarno, R. (2019). Sentiment analysis of hotel reviews using Latent Dirichlet Allocation, semantic similarity, and LSTM. Int. J. Intell. Eng. Syst, 12(4), 142-155.

[5] Styawati, S., Nurkholis, A., Aldino, A. A., Samsugi, S., Suryati, E., & Cahyono, R. P. (2022, January). Sentiment analysis on online transportation reviews using Word2Vec text embedding model feature extraction and support vector machine (SVM) algorithm. In 2021 International Seminar on Machine Learning, Optimization, and Data Science (ISMODE) (pp. 163-167). IEEE.

[6] Pressrelease.kontan.co.id. (2021, 13 Oktober). GoCorp: Solusi Teknologi Terbaru dari Gojek Bantu Pelaku Usaha Atur Tunjangan Transportasi Karyawan. Diakses pada 15 Juli 2023, dari https://pressrelease.kontan.co.id/release/gocorp-solusi-teknologi-terbaru-dari-gojek-bantu-pelaku-usaha-atur-tunjangan-transportasi-karyawan?page=all

[7] Yusuf, K. K., & Pradekso, T. (2021). Pengaruh Terpaan Iklan dan Terpaan Konten Celebrity Endorsement Terhadap Minat Menggunakan Jasa Go-Food (Melalui Brand Attitude). Interaksi Online, 9(3), 217-229.

[8] Adilah, M. T., Supendar, H., Ningsih, R., Muryani, S., & Solecha, K. (2020, November). Sentiment analysis of online transportation service using the naïve Bayes methods. In Journal of Physics: Conference Series (Vol. 1641, No. 1, p. 012093). IOP Publishing.

[9] Handani, S. W., Saputra, D. I. S., Arino, R. M., & Ramadhan, G. F. A. (2019, March). Sentiment analysis for go-jek on google play store. In Journal of Physics: Conference Series (Vol. 1196, No. 1, p. 012032). IOP Publishing.

[10] Prananda, A. R., & Thalib, I. (2020). Sentiment analysis for customer review: A case study of GO-JEK expansion. Journal of Information Systems Engineering and Business Intelligence, 6(1), 1.

[11] AL-JARRAH, I., Mustafa, A. M., & Najadat, H. Aspect-Based Sentiment Analysis for Arabic Food Delivery Reviews. ACM Transactions on Asian and Low-Resource Language Information Processing.

[12] Somboon, T., Gleeson, M. P., & Hannongbua, S. (2012). Insight into the reaction mechanism of cis, cis-muconate platonizing enzymes: a DFT QM/MM study. Journal of molecular modeling, pp. 18, 525–531.

[13] Bitto, A. K., Bijoy, M. H. I., Arman, M. S., Mahmud, I., Das, A., & Majumder, J. (2023). Sentiment analysis from Bangladeshi food delivery startup based on user reviews using machine learning and deep learning. Bulletin of Electrical Engineering and Informatics, 12(4), 2282-2291.

[14] Chakravarthi, B. R., Priyadharshini, R., Muralidaran, V., Suryawanshi, S., Jose, N., Sherly, E., & McCrae, J. P. (2020, December). Overview of the track on sentiment analysis for Dravidian languages in code-mixed text. In Proceedings of the 12th Annual Meeting of the Forum for Information Retrieval Evaluation (pp. 21–24).

[15] El Rahman, S. A., AlOtaibi, F. A., & AlShehri, W. A. (2019, April). Sentiment analysis of Twitter data. In 2019 international conference on computer and information sciences (ICCIS) (pp. 1–4). IEEE.

[16] Elbagir, S., & Yang, J. (2019, March). Twitter sentiment analysis using natural language toolkit and VADER sentiment. In Proceedings of the international multiconference of Engineers and computer scientists (Vol. 122, p. 16).

[17] Nandwani, P., & Verma, R. (2021). A review on sentiment analysis and emotion detection from text. Social Network Analysis and Mining, 11(1), 81.

[18] Bonta, V., Kumaresh, N., & Janardhan, N. (2019). A comprehensive study on lexicon-based approaches for sentiment analysis. Asian Journal of Computer Science and Technology, 8(S2), pp. 1–6.

[19] Vashishtha, S., & Susan, S. (2019). Fuzzy rule-based unsupervised sentiment analysis from social media posts. Expert Systems with Applications, p. 138, 112834.

[20] Wang, T., Lu, K., Chow, K. P., & Zhu, Q. (2020). COVID-19 sensing: negative sentiment analysis on social media in China via BERT model. Ieee Access, p. 8, 138162–138169.