2022 3rd International Conference on Intelligent Engineering and Management (ICIEM)

Sentiment Analysis on Twitter Data Using LSTM with Emoji Embedding

Aditya Singh
School of Computing
DIT University
Dehradun, Uttarakhand, India
adiusingh127@gmail.com

Jyoti Rawat
School of Computing
DIT University
Dehradun, Uttarakhand, India
drjyotirawat19@gmail.com

Saurabh Bhandari
School of Computing
DIT University
Dehradun, Uttarakhand, India
saurabh20bhandari@gmail.com

Umangdeep Negi
School of Computing
DIT University
Dehradun, Uttarakhand, India
umangdeep18@gmail.com

Shreshtha Pahwa
School of Computing
DIT University
Dehradun, Uttarakhand, India
shreshthapahwa21@gmail.com

2022 3rd International Conference on Intelligent Engineering and Management (ICIEM) | 978-1-6654-6756-8/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICIEM54221.2022.9853058

***Abstract*—Data analysis in today's data-driven world has proved to be the most essential thing. The majority of the data generated online is unstructured and needs to be processed. Social media plays a vital role in sharing data on the internet in large amounts, this data needs to be handled properly and information from this unstructured data must be extracted. By using sentiment analysis, the intentions of textual data can be figured out and can classify properly on the internet. In the proposed approach, the implementation of sentiment analysis is done using the combination of various artificial intelligence technologies and LSTM including some key symbols like emojis with special characters to increase the accuracy of pre-built models. That can make them more practical in today's world.**

Keywords— Artificial Intelligence; Sentiment Analysis; Machine Learning; Twitter data; Text Mining; emojis; special characters.

I. Introduction

In the past few years, the production of textual information from various social media sites has increased exponentially. Sentiment analysis with the assistance of social media sites has additionally gained quality among a good variety of individuals having different interests. As users everywhere on the planet has the likelihood to specify their opinion concerning totally different subjects associated with politics, surroundings, movies, travel, industrial products, culture thus extracting data from that information becomes a theme of nice significance and importance for varied business for his or her profit [1].

Sentiment analysis is performed on textual information to assist businesses to monitor whole and products sentiment in client feedback and perceive client wants. Within industrial field, sentiment analysis has a massive result, like government organizations, large corporations, their need is to understand what folks believe their product is and their market price [2].

Sentiment evaluation is used to locate the mood, opinion and behaviour of someone using numerous data. For appearing sentiment evaluation, various social media sites make use of diverse sentiment evaluation methods to require the overall public information. Sentiment evaluation is broadly applied in numerous domain names like politics, defence, finance. The statistics handy on numerous social media sites may be structured or unstructured. Nearly eighty percentage of information found in net is unstructured. [2] Sentiment evaluation methods are used to finish user's evaluations on numerous social media sites. Besides information concerning users visited web sites, getting options etc., understanding their emotions as they are expressed via means of their messages on numerous social platforms, clothed to be a completely critical thing for the prediction of people's mind on few precise subjects. A broadly used method is to categorise the polarity of a textual content in phrases of the user's satisfaction, discontent or neutrality. The polarity will range in phrases of labelling or variety of tiers from effective to impartial and impartial to negative, but usually it denotes the feelings of a textual content numerous from a cheerful to sad mode [3][4].

The techniques used for performing sentiment analysis are varied, they are majorly based on NLP (Natural Language Processing) and ML (Machine Learning) techniques for extracting valuable vectors and then classifying data in acceptable polarity categories. This work is focused mainly on reading and cleaning the unstructured data and converting it into structured format and to clean that data using various ETL (Extraction, Transform, and Load) techniques and then categorizing it. To accomplish this, NLTK (Natural Language Toolkit) and various other sentiment analysis tools are used in this paper.

In past few years, users on social media sites have become conversant in using a group of symbols in their conversations to precise their emotions. These characters or graphical symbols, usually referred to as emojis [3], have become an universal language that is used by all the apps and platforms [4–7]. In general, the use of emojis is to supply emotions that is otherwise missing from typewritten or spoken communication. Compared with texts, emojis helps us to express emotions a lot precisely and directly. While analyzing day to day texts, like tweets, or comments or blogs, emojis offer crucial data relating to user sentiment [8–13]. In

978-1-6654-6756-8/22/$31.00 ©2022 IEEE

the year 2019, there were around 815 million users who were using the Instagram monthly. For the year 2023, this number is projected to reach around 1.2 billion users [14].

As per the primary quarter of 2019, Twitter averaged around 330 million MAU (monthly active users), a decline from its incomparable high of 336 MAU within the half-moon of 2018 [15]. Twitter is a micro blogging site wherever shoppers post 'tweets' for their devotees or a different shopper. In 2016, Twitter has over 313 million dynamic shoppers within a given month, as well as one hundred million shoppers daily [16]. Shopper origins are widespread, with seventy-seven set outside of the North American nation, manufacturing over five hundred million tweets daily [17]. Twitter was positioned 12th universally for activity in the year 2017 [18] and served more than fifteen billion API calls every day [19]. Under this monumental development, Twitter was recently the topic of abundant scrutiny, as Tweets often categorise users sentiments on sensitive problems. In the social media context, sentiment evaluation and mining critiques are notably difficult job, this is because of huge amount of data generated through humans [20].

The main aim of current research is to analyse the currently present sentiment analysis techniques used for tweets, perform theoretical comparisons between different approaches and find a more accurate method for sentiment analysis. The paper is divided as follows: the second section contains related work done by other researchers in this field. The next section is about methods and various classification techniques used for analysing sentiment. Various other sentiment-analysis methods used for Twitter are also compared that includes supervised, unsupervised and deep learning. In the next section, the result of the research is presented. Finally, a brief discussions and comparisons of few previous research papers are highlighted.

II. Literature Review

The basic idea of processing tweets for sentiment analysis is proposed by Gorkhi et al. [21] where the data is first collected and all those features which cannot contribute to finding accurate results is eliminated. A sentiment classification engine powered by a naive bayes algorithm is then used to calculate and pick the most accurate emotion that the text is conveying as an output.

An approach is known as "Quantification" is proposed by M. bouazizi et al. [22] in which scores of more than one sentiments are extracted and the sentiment with the highest score is selected as an output. This is done using multi-class sentiment analysis and the SENTA tool.

G. Gautam et al. [23] worked on product review classification using labelled data. They used various algorithms of machine learning (like SVM, maximum entropy, naïve bayes). They concluded that naive bayes gives satisfactory results as its accuracy is better than various other like SVM, Maximum entropy and accuracy can be improved by using WordNet.

Goldberg et al. [24] suggested a semi-supervised category set of rules which scored from 0 to 4 for superb and bad remarks based on graphs. Wang et al. [25] analyzed the sentiment evaluation of short texts using couple of features which includes sentiment capabilities, bad capabilities and emoji, a high-dimensional combined function sentiment evaluation version primarily based totally on SVM become proposed [26]. The sentiment evaluation approach of machine learning is more reliable, but it mainly focuses on the quality of the text labelled with their polarity. Many scholars have also used deep learning techniques for sentiment evaluation and have achieved satisfactory output. The Recursive Neural Tensor Network model proposed via way of means of Socher et al. [27] delivered a sentiment tree library that received accurate sentiment evaluation outcomes within side the statistics set of film reviews.

The character to sentence CNN (Convolutional Neural Network) [28] version makes use of 2 convolutional layers to improve the sentiment analysis. Irsoy et al. [29] used the RNN (Recurrent Neural network) based totally on-time series records for representing sentences, which helped in getting a better accuracy for sentiment detection.

In this work, LSTM (Long Short-Term Memory) model is used for obtaining the text information and dictionary approach for emojis of the Twitter tweet dataset. Table 1 summarizes the literature work related to the sentiment analysis using different methods.

TABLE I. Related Studies For Sentiment Analysis Using Different Methods

Author	Dataset	Method / Tools	Accuracy (%)	Output	Summary
Yuliyanti et al. [30]	Activities of the community development program	SVM, Lexicon method	88.64%	0 for positive and 1 for negative	It generated effective evaluation on community development program and link them with best action.
Martin et al. [31]	34,528 Tweets downloaded from London Heathrow Airport (LHR) using the official Twitter account	Machine learning	82.4%	Negative if < 0, neutral if 0 and positive if >0	It helps to conclude how different areas of Airport like WIFI, boarding is performing and made twitter data to be a reliable option for companies.

Mansour et al. [32]	Twitter data related to Eastern and Western countries view ISIS	Lexicon-based	70%	0 is flagged for neutral activity. 1, flagged for the terrorist activity of any kind	It explains about the clash between ISIS and the world. It also helps us to show the importance of such data from social medias for geopolitics.
Husain et al. [33]	twitter dataset and 20 news groups	Naïve Bayes Maximum Entropy	83% 80%	Levels from 0 to 5 to measure the extent of depression.	It compares various classifiers like SVM, Naïve bayes, and maximum entropy. The accuracy of SVM is found to be maximum.
Joyce et al. 34]	Gathered tweets from the 2012 presidential election that contained the phrase “McCain” or “Obama.”	Machine learning and lexicon-based method	74.98%	The popularity of Trump and Clinton is measured on a graph.	Lexicon is compared with a machine learning algorithm and Lexicon based is found to be more effective.
G. Gautam et al. [23]	twitter dataset	SVM, Semantic Analysis (WordNet),Naive Bayes	85.5% 83.8% 88.2% 89%	Negative if < 0, neutral if 0 and positive if >0	Naive bayes gives satisfactory results in terms of accuracy then various other like SVM, Maximum entropy and accuracy can be improved by using WordNet.
Seyed-Ali et al.[35]	Twitter dataset on customer review	Naive Baye ,Unigram featur Hybrid Approacl SVM	89.78%	0 for positive and 1 for negative	Hybrid method outperforms the Unigram baseline.
Neethu M. S. et al. [36]	Twitter data on Electronic products	Naïve Bayes, SVM Max Entropy Ensemble	89%,90%,90 %, 90%	0 for positive and 1 for negative	Various classifiers are compared and Naïve bayes lacks in giving a decent amount of accuracy.
A. Agarwal et al. [37]	10,000 manually Annotated Tweets	Senti-features, Kernel, Unigram Senti features Kernel+Senti features, Unigram	56.58% 60.60% 56.31% 60.50% 60.83%	Negative if < 0, neutral if 0 and positive if >0	Implemented tree kernel and feature based models and demonstrate that both these models outperform the unigram baseline
D. Gurkhe et al. [38]	Twitter Data	Unigram Bigram Unigram+Bigram	81.2% 15% 67.5%	0 for positive and 1 for negative	Unigram + Bigram show the best result bi grams alone shows weak results because not all data items consist of bigrams indicating their sentiments.
Monali Bordoloi et al. [39]	Tweets collected using Twitter API	MaxEnt, and SVI classifiers.	83%, 82 %	0 for positive and 1 for negative	Max entropy and SVM are used to classify a dataset in which both have generated same accuracy level and are much better.
Malhar et al. [40]	Collecting tweets using Twitter API	NB, SVM and AN classifiers.	92%	0 for positive and 1 for negative	Social network based behavioral analysis parameters can increase the prediction accuracy along with sentiment analysis
Anton et al. [41]	Using online system to collect tweets Sentiment140	NB, Support vect classifiers	81% 74%	0 = negative, 4 = positive	Naive Bayesian approach out-performs SVM in certain areas and is required to consider.
Goldberg et al. [24]	Semi-supervised learning for sentiment categorization using graph	SVM	92%	0 for positive and 1 for negative	Graph-primarily based on semi-supervised category set of rules which scored from 0 to 4 for superb and bad remarks.

III. METHODOLOGIES

A. Dataset

For this work, dataset is taken from Kaggle which consists of 162980 tweets [25].

B. Pre-processing of the datasets

Data pre-processing or data preparation simply filters the data to remove useless unwanted features or inappropriate data. The following tasks are performed in pre-processing:

- Removing URLs
- Removing Stop words
- Lemmatizing
- Tokenization

The twitters dataset used for research is already labelled. It contains a positive, negative, and neutral polarity. The unprocessed records having polarity is extraordinarily vulnerable to redundancy and inconsistency. The quality of the records affects the output and consequently to enhance the quality, data pre-processing is required. Datasets contain various features, dropping the unnecessary and repeated words that do not add any value in analysis of the dataset to improves the efficiency of the data.

C. Feature Extraction

The improved or processed dataset after pre-processing has a lot of distinctive words (tokens). The feature extraction method extracts tokens from dataset. Later these tokens are assigned a value which is the frequency or the importance of the token in the dataset. It is used to show the negative, neutral, and positive polarity in the sentence which is then used for determining the polarity of the tweet.

Term frequency-inverse document frequency tells us about the significance of a word in sentence depending on how frequently it occurs in a document. Term frequency-inverse document frequency is a type of text vectorizer that converts normal textual data into a numerical vector. It mainly consists of two concepts:

- *Term Frequency:* Frequency of a specific term in a document given in Eq. 1.
- *Document Frequency:* The number of documents containing a specific term given in Eq. 2.

where,

w= w is a word in the document
D= D is combination of whole documents
N= N is the number of documents included in database
d=d is a specified document in database
f=f is a frequency

$$tf(w,d) = \log\big(1 + f(w,d)\big) \quad (1)$$

$$idf(w,D) = \log(N/f(w,D)) \quad (2)$$

$$tfidf(w,d,D) = tf(w,d) * idf(w,d) \quad (3)$$

Final step is to process the tf-idf (Inverse Document Frequency) score with the following formula given in Eq. 3 [26].

D. Sentiment Detection

Sentiment detection or emotion identification is a crucial task in many sentiments' evaluation and opinion mining applications such as tweet mining, opinion detection, and tweet classification. The most important job in sentiment evaluation is to classify the polarities of the characteristics of a particular tweet. Polarity can be divided into three classes: Positive, Neutral and Negative. Polarity is identified by using a variety of lexicons such as tf-idf vectorizer and the count vectorizer, which helps to calculate emotional polarity.

E. Training and Classification

LSTM is known for its effectiveness in memorizing important information which can be of great use in a classification problem. In this too, LSTM is applied to get more accurate result for sentiment analysis.

F. Long Short-Term Memory

It is a special recurrent neural network model that can process entire sequence of data. LSTM contain feedback connections. An LSTM incorporates a cell state, a forget gate, an output gate and an input gate which gives them freedom to learn, retain or unlearn information from each unit on their own. Cell state of LSTM is just like a conveyor belt that runs straight down the entire chain. It helps the information to flow along it unchanged and allowing only some linear interactions. Three gates can add or remove the knowledge to the cell state. The forget gate uses sigmoid function and is used to decides which information from the previous cell state should be forgotten. The input gate uses both sigmoid and tanh functions to controls the knowledge flow to this cell state. The output gate then passes the information to the subsequent hidden state.

G. Emoji Lexicon

A specific Emoji character is a Unicode representation in a program. A lexicon consists of 18 different signs for around 843 unique Emoji characters that supports Emoji notations including Java, R and other forms. In today's world, these emojis have become an essential component of writing. People share their feelings and use these characters in plain text sentences to enhance the messages delivered by that sentence. It is very important to include such symbols or characters during sentiment analysis. Fig. 1 contain example of few emojis.

Figure 1. Example of emoji lexicon

IV. PROPOSED MODEL

A. Implementation

In the implemented model LSTM is used to analyze the text data along with emojis. A twitter dataset is taken from Kaggle for model training. The dataset has three different sentiments namely, negative (-1), neutral (0), and positive (+1).

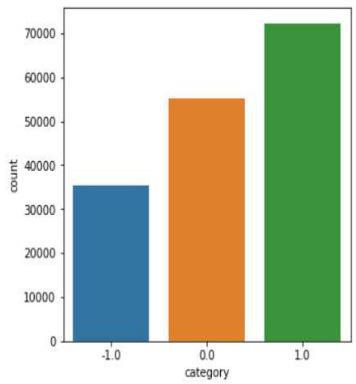

Figure 2. The category wise tweet label.

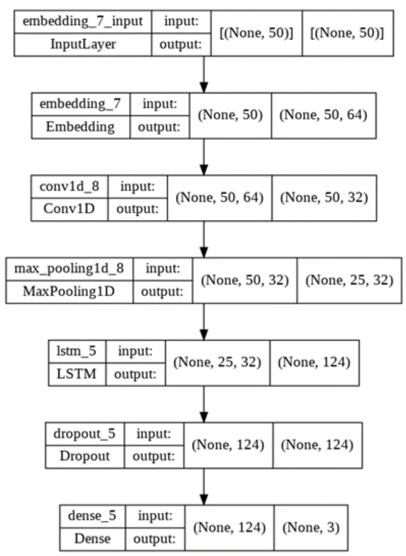

Figure 3. The summary of model.

The dataset consists of 162980 tweets and contains two fields for the tweet and label. Fig. 3 shows the category wise tweet label. If the category is -1.0 it is a negative tweet, if 0 then it's a neutral tweet and if 1 then it's a positive tweet. The stop-words are removed, and the dataset is cleaned in the process of pre-processing. Tokenization and Lemmatization are carried out.

The list of all the emojis present in the dataset is extracted. These emojis will be essential in building the model, their polarity must be taken and added to the rest of the polarity of the text data. An overall sentiment score will be generated then to predict whether the given tweet is positive (1), negative (-1) or neutral (0). In this model embedding layer, LSTM and activation functions such as ReLu and Sigmoid is used. The summary for the model used is shown in Fig. 3.

B. Emoji Confusion

The confusion between the two tweets mentioned in Fig. 4, must be resolved where one tweet express happy sentiment whereas another one with the same text content but a different kind of emoji express the sentiment of sadness must be resolved by the polarity concept where an average of the sentiment score of emojis and text content is taken to conclude.

Figure 4. Similarity between two tweets model

C. Results

An accuracy of 87.26% is obtained by using LSTM of which the report is mentioned below. Table 2 shows the precision score , recall and f1 score.

TABLE II. THE DETAILS OF DIFFERENT PERFORMANCE PARAMETERS

	Precision	**Recall**	**F1-score**	**Support**
-1.0 (Negative)	0.84	0.75	0.79	2893
0.0 (Neutral)	0.84	0.96	0.90	4488
1.0 (Positive)	0.92	0.87	0.89	6200
Accuracy			0.87	13581
Macro avg	0.86	0.86	0.86	13581
Weighted avg	0.87	0.87	0.87	13581

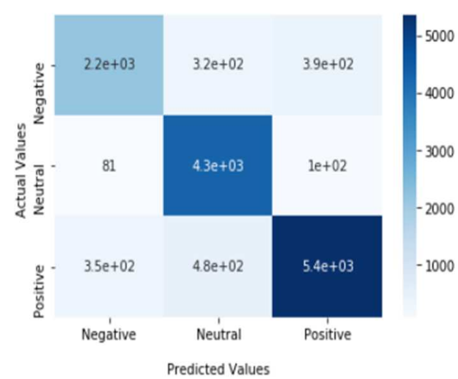

Figure 5. Confusion matrix of the performance.

Fig. 5 is the confusion matrix which tells about how many datapoints are correctly classified by the model.

In Table 2 summarized about the results predicted by the model, where 1 belongs to positive, 0 is for neutral and -1 is for negative.

Fig. 6 explain about polarity of text used with emojis.

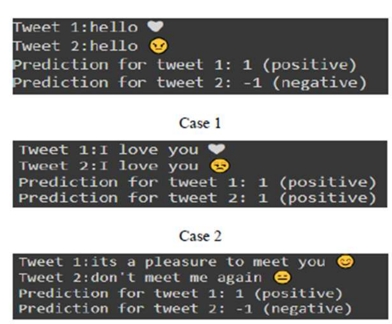

Figure 6. Polariy of text used in emojis.

V. CONCLUSION

Sentiment analysis nowadays is more than just text mining and is getting improved drastically. The proposed model helps the existing sentiment analysis techniques to predict better results and helps to include various symbols especially the new emerging language of emojis which can be now used every day in messages and posts on various social media sites by the society.

REFERENCES

[1] S. Stieglitz, M. Mirbabaie, B. Ross, and C. Neuberger, "Social media analytics – Challenges in topic discovery, data collection, and data preparation". International Journal of Information Management, 2018, 39(39), pp.156–168.

[2] www.computerweekly.com. (n.d.). IBM: 80 per cent of our global data is unstructured (so what do we do?) - CW Developer Network.

[3] Description of emoji. [Cited 2021 20 May]. Available from: https://en.wikipedia. Org/wiki/Emoji.

[4] N. Ljubesic, D. Fiser, "A global analysis of emoji usage", in Proceedings of the 10th Web as Corpus Workshop, 2016, pp. 82–89.

[5] W. Ai, X. Lu, X. Liu, et al., "Untangling emoji popularity through semantic embeddings", Eleventh International AAAI Conference on Web and Social Media, 2017.

[6] P. K. Novak, J. Smailovi, B. Sluban, et al., "Sentiment of emojis", PloS One 10 (12) (2015) 1–22.

[7] M. Shiha, S. Ayvaz, "The effects of emoji in sentiment analysis", Int. J. Comput. Electr. Eng. (IJCEE.) 9 (1) (2017) 360–369.

[8] M. Boia, B. Faltings, CC. Musat, P. Pu, "A:) is worth a thousand words: How people attach sentiment to emoticons and words in tweets", Intl. Conf. Soc. Comput. (SocialCom). (2013) 345–350.

[9] A. Hogenboom, D. Bal, F. Frasincar, M. Bal, Jong F. de, U. Kaymak, "Exploiting emoticons in polarity classification of text", J. Web Eng. 14 (1–2) (2015) 22–40.

[10] D. Davidov, O. Tsur, A. Rappoport, "Enhanced sentiment learning using Twitter hashtags and smileys", in Proc. 23rd Intl. Conf. on Computational Linguistics: Posters., ACL, 2010, pp. 241–249.

[11] K. Liu, W. Li, M. Guo, "Emoticon smoothed language models for Twitter sentiment analysis", in Proc. 26th AAAI Conf. on Artificial Intelligence, 2012, pp. 1678–1684.

[12] J. Read, "Using emoticons to reduce dependency in machine learning techniques for sentiment classification", in Proc. ACL Student Research Workshop, ACL, 2005, pp. 43–48.

[13] J. Zhao, L. Dong, J. Wu, K. Xu, "Moodlens: An emoticon-based sentiment analysis system for Chinese tweets", in Proc. 18th ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining, ACM, 2012, pp. 1528–1531.

[14] Statista article "Instagram: number of global users 2019-2023"

[15] Statista article "Twitter users worldwide from 1st quarter 2010 to 1st quarter 2019".

[16] I. Twitter, "Second Quarter 2016 Report," ed, 2016.

[17] I. Twitter, "Twitter IPO Prospectus," ed, 2013

[18] Alexa.com, "Website Traffic Ranking," ed, 2017.

[19] A. DuVander, "Which APIs are handling billions of requests per day?," Programmable Web, 2012.

[20] A. Giachanou and F. Crestani, "Like It or Not: A Survey of Twitter Sentiment Analysis Methods," ACM Comput. Surv., vol. 49, no. 2, pp. 1-41, 2016.

[21] Gurkhe, Dhiraj, Niraj Pal, and Rishit Bhatia. "Effective sentiment analysis of social media datasets using Naive Bayesian classification." International Journal of Computer Applications 975.8887 (2014): 99.

[22] Bouazizi, Mondher, and Tomoaki Ohtsuki. "A pattern-based approach for multi-class sentiment analysis in Twitter." IEEE Access 5 (2017): 20617-20639.

[23] Gautam, Geetika, and Divakar Yadav. "Sentiment analysis of Twitter data using machine learning approaches and semantic analysis." 2014 Seventh international conference on contemporary computing (IC3). IEEE, 2014.

[24] A. B. Goldberg and X. Zhu, "Seeing stars when there aren't many stars: Graph-based semi-supervised learning for sentiment categorization," in Proc. Workshop Graph-Based Methods Natural Lang. Process., Sydney, VIC, Australia, Jun. 2006, pp. 45–52.

[25] Hussain, Sherif ,"Twitter Sentiments Dataset", Mendeley Data, V1, doi: 10.17632/z9zw7nt5h2.1, (2021).

[26] Liu, C., Fang, F., Lin, X., Cai, T., Tan, X., Liu, J. and Lu, X. (2021). "Improving sentiment analysis accuracy with emoji embedding. Journal of Safety Science and Resilience", 2(4), pp.246–252.

[27] R. Socher, ''Recursive deep models for semantic compositionality over a sentiment treebank,'' in Proc. Conf. Empirical Methods Natural Lang. Process., Seattle, WA, USA, 2013, pp. 1631–1642.

[28] C. D. Santos and M. Gattit, ''Deep Convolutional Neural Networks for Sentiment Analysis of Short Texts,'' in Proc. 25th Int. Conf. Comput. Linguistics: Tech. Papers, Dublin, Ireland, 2014, pp. 69–78.

[29] O. Irsoy and C. Cardie, ''Opinion Mining with Deep Recurrent Neural Networks,'' in Proc. Conf. Empirical Methods Natural Lang. Process., Doha, Qatar, 2014, pp. 720–728.

[30] Yuliyanti, Siti, Taufik Djatna, and Heru Sukoco. "Sentiment mining of community development program evaluation based on social media." TELKOMNIKA (Telecommunication Computing Electronics and Control) 15.4 (2017): 1858-1864.

[31] Martin-Domingo, Luis, Juan Carlos Martín, and Glen Mandsberg. "Social media as a resource for sentiment analysis of Airport Service Quality (ASQ)." Journal of Air Transport Management 78 (2019): 106-115.

[32] Mansour, Samah. "Social media analysis of user's responses to terrorism using sentiment analysis and text mining." Procedia Computer Science 140 (2018): 95-103.

[33] Hassan, Anees Ul, et al. "Sentiment analysis of social networking sites (SNS) data using machine learning approach for the measurement of depression." 2017 international conference on information and communication technology convergence (ICTC). IEEE, 2017.

[34] Joyce, Brandon, and Jing Deng. "Sentiment analysis of tweets for the 2016 US presidential election." 2017 ieee mit undergraduate research technology conference (urtc). IEEE, 2017.

[35] Bahrainian, Seyed-Ali, and Andreas Dengel. "Sentiment analysis and summarization of Twitter data." 2013 IEEE 16th International Conference on Computational Science and Engineering. IEEE, 2013.

[36] Neethu, M. S., and R. Rajasree. "Sentiment analysis in Twitter using machine learning techniques." 2013 fourth international conference on computing, communications and networking technologies (ICCCNT). IEEE, 2013.

[37] Agarwal, Apoorv, et al. "Sentiment analysis of Twitter data." Proceedings of the workshop on language in social media (LSM 2011). 2011.

[38] Gurkhe, Dhiraj, Niraj Pal, and Rishit Bhatia. "Effective sentiment analysis of social media datasets using Naive Bayesian classification." International Journal of Computer Applications 975.8887 (2014): 99.

[39] Bordoloi, M. and Biswas, S.Kr. (2020). "Graph-based sentiment analysis using keyword rank based polarity assignment". Multimedia Tools and Applications.

[40] M. Anjaria and R. M. R. Guddeti, "Influence factor-based opinion mining of Twitter data using supervised learning," in 2014 Sixth International Conference on Communication Systems and Networks (COMSNETS), 2014, pp. 1-8.

[41] A. Barhan and A. Shakhomirov, "Methods for Sentiment Analysis of Twitter messages," in 12th Conference of FRUCT Association, 2012.