Sentiment analysis using global vector and long short-term memory

Kusum, Supriya P. Panda
Department of Computer Science and Engineering, Faculty of Engineering and Technology,
Manav Rachna International Institute of Research and Studies, Faridabad, India

Article Info

Article history:

Received Oct 25, 2021
Revised Feb 10, 2022
Accepted Feb 17, 2022

Keywords:

Deep learning
Global vector (GloVe)
Long short-term memory
Recurrent neural network
Sentiment analysis

ABSTRACT

Tweet sentiment analysis is a deep learning study that is beneficial for automatically determining public opinion on a certain topic. Using the long short-term memory (LSTM) algorithm, this paper aims to proposes a Twitter analysis technique that divides Tweets into two categories (positive and negative). The global vector (GloVe) word embedding score is used to rate many selected words as network input. GloVe converts words into vectors by building a corpus matrix. The GloVe outperforms its prior model, owing to its smaller vector and corpora sizes. GloVe has a higher accuracy than the model word embedding word2vec, continuous bag of word (CBoW), and word2vec Skip-gram. The preprocessed term variation was conducted to test the performance of sentiment classification. The test results show that this proposed method has succeeded in classifying with the best results with an accuracy of 95.61%.

Corresponding Author:

Kusum
Department of Computer Science and Engineering, Faculty of Engineering and Technology
Manav Rachna International Institute of Research and Studies
Faridabad, Haryana, India
Email: kusumprerak@gmail.com

1. INTRODUCTION

Background: Due to the increasing growth of social media use, particularly in India, sentiment analysis on social media is currently an intriguing issue to research. One of them is Twitter, a social media platform. By sending a Tweet, which is a short message, Twitter users can express their views on a particular discussion that is happening around them. This Tweet can be analysed for sentiment for a variety of purposes, including as determining an individual's personality or determining people's interest in one or many other things [1], [2].

It entails attempting to judge the emotion expressed in a text, that is, analysing it to determine the emotion that the person is expressing in relation to a product, news, or any other topic. Texts can be classified as positive, negative, or neutral using the most basic opinion tools [3]. According to statista.com [4], [5] the total number of Internet users worldwide is 4,540 million in the year 2021 and on social networks, there are 3.6 billion users in the year 2020. As more people incorporate these modern applications into their daily lives, a demand has arisen for the development of new emerging technologies capable of manipulating and analysing large amounts of data, obtaining patterns and trends from them, and drawing conclusions that aid in better understanding the general public and generating important decisions, whether in commercial or electoral matters or, where appropriate, generating marketing strategies on products or services [6].

Identifying the predominant sentiment of the users is a very complex task even for the human being and is the reason for being of this discipline. In recent years, an enormous number of sentiment analysis studies have been carried out and it has been applied in a wide variety of interdisciplinary fields, such as politics [7], technology [8], medicine [9], companies [10], to name a few, it should be noted that most of the studies in this branch take the social network Twitter as the main source of obtaining data to analyze public opinion. Identifying the predominant sentiment of the users is a very complex task even for the human being and is the reason for being of this discipline. In recent years, an enormous number of sentiment analysis studies have been carried out and it has been applied in a wide variety of interdisciplinary fields, such as politics [11], technology [12], medicine [13], companies [14], to name a few, it should be noted that most of the studies in this branch take the social network Twitter as the main source of obtaining data to analyze public opinion.

That is why it has become a necessity for companies to monitor social networks to analyze the opinions of their customers and obtain feedback on their products/services to improve them, according to [15] this type of study allows companies to carry out market research without the need to resort to surveys of people, obtaining greater quality information. To build this sentiment analysis system, the method will be used word embedding. This method can improve the performance of sentiment analysis, because that word embedding is widely used in research that discusses sentiment analysis. In this study, the model will be used word embedding which name is global vector (GloVe) [16], [17]. Model GloVe is chosen because it has a good level of accuracy compared to the model word embedding more like word2vec (Continuous Bag of Words and Skip-gram) and doc2vec [16]. Word embedding and tweet data will then be analyzed for sentiment classification. In this study, a method will be used deep learning to classify sentiments. The model to be used is long short-term memory (LSTM). In LSTM It has several layers, one of which is used for word embedding and has a good performance for classifying sentiments when used with word embedding model GloVe [16], [17].

- Problem statement: How to implement GloVe word embedding and long-short-term-memory (LSTM) deep learning algorithm to form a sentiment analysis model on voluminous data.
- Proposed solution: Implement and analyze the performance of a sentiment analysis model using the word embedding GloVe and the LSTM deep learning model for classifying Tweets sentiments as positive or negative. The classification of Tweets by which the scheme can make it easier for companies/groups/individual to find out the perceptions in the form of negative opinions and positive opinions, so that they can be used as a reference in efforts to maintain quality and improve deficiencies, as well as evaluate products and services in a better direction.

2. RELATED WORK

The opinions of Twitter users were examined for the prediction of film industry trends in a study [18]. They used Twitter's application programming interface (API) to download over 500 tweets related to the release of three films, both before and after they were released. To assess the polarity of the tweets, they employed the tool TextBlob to perform lexical analysis and pre-processing operations. What do you mean by positive, negative, and neutral? as a result, they discovered that users' opinions of the films are positive prior to their release and gradually improve after their release. They also discovered that negative opinions become neutral, and that there is a strong relationship between the data analyzed and ticket sales at the box office. They state that this type of study can be useful to develop strategies for marketing in real time.

A sentiment analysis of a collection of tweets was used to evaluate the performance of a television program by Munjal *et al*. [18]. They employed a lexicon-based strategy to determine the polarity of the feelings expressed in the tweets and identify them as positive or negative, and then used this information to train a classifier using support vector machines (SVM). According to the study, a combination of these methodologies may be used to accurately analyze sentiment in television programs with an accuracy rate of 80%.

For the 2016 US presidential elections, Tiara *et al*. [19] conducted a sentiment study on Twitter. They calculated sentiment using two approaches: lexicon-based sentiment analysis using Opinion Finder and sentiment analysis with machine learning using the natural language processing toolkit (NLTK) to implement the algorithm Naive Bayes (NB). The study showed that there is a very high correlation coefficient of 94% with the data from the surveys. Furthermore, they stated that social media surveys may be more heavily incorporated into voting in the future.

Many methods can be used in forming sentiment analysis, as did Imaduddin *et al*. [20] which uses multiple models word embedding Word2vec, continuous bag of word (CBoW), Word2vec Skipgram, and doc2vec. They used LSTM to carry out the sentiment classification procedure for hotel reviews and got the best results when compared to other implementation word embedding methods. According to the study's findings, GloVe outperformed other model word embedding methods [21].

Pennington *et al*. [22] and Sharma *et al*. [23] build a model GloVe of approach CBoW, and Skip-gram in shaping word embedding, GloVe has a higher accuracy than the model word embedding word2vec, CBoW, and word2vec Skip-gram. As a result, it provides benefits in terms of accuracy and computing speed. Apart from that, Li and Qian [24] explained that they chose LSTM to categorize sentiments because it can handle the problem of vanishing/exploding gradient, which is a development of the model recurrent neural network (RNN). Li and Qian [24] stated this in their study, which looked at the sentiment of the text. Because of this, the model will be used for sentiment analysis of disaster Tweets using word embedding GloVe and LSTM [24], [25].

2.1. Literature review

2.1.1. Word2Vec and continuous bag-of-words (CBoW)

According to Mikolov *et al*. [26] introduced Word2Vec, a method for expressing words in vector form or word embedding, which has two architectural models [27]. Continuous bag-of-words (CBoW) and Skip-gram models are two architectural models proposed by Word2vec to build the word representation. Whereas, CBoW is Word2vec's model architecture, which is based on the loglinear model. The CBoW model architecture works by uniformly dividing the projection layer across all words, resulting in evenly dispersed vectors (projected in the same place). To predict words, in the CBoW model words are predicted based on their context.

2.1.2. Skip-gram

Skip-gram model as the second model based on the Word2vec approach. When comparing the models CBoW and Skip-gram, the main distinction is that the CBoW model predicts the probability of a word given its context. The context can have any number of words in it. A context window, which tells how large the neighbour of the provided word as context will be, determines the amount of words in the context. Therefore, CBoW predicts or classifies words depending on their context, whereas Skip- gram predicts words by looking at other words in the same sentence [26], [27]. Figure 1 depicts the architecture of CBoW and Skip-gram.

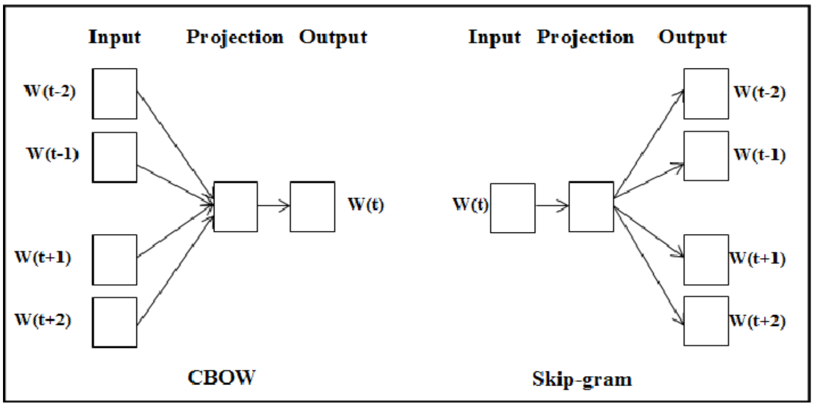

Figure 1. CBOW and Skip-gram: source [28]

2.1.3. Global vector (GloVe)

The RNN is a deep learning method based on architectural neural networks that can represent sequential input. RNN can store information about the previous state in order to determine the potential or provide output based on the prior state. However, RNN has the drawback of experiencing vanishing gradient/exploding gradient when too many sequences are executed. The term "vanishing/exploding gradient" refers to a circumstance in which the gradient value might be very tiny or equal to zero, as well as highly large [29]. As a result, a method known as LSTM was created to solve these flaws by using a gate system.

2.1.4. Recurrent neural network (RNN)

The RNN is a deep learning method based on architectural neural networks that can represent sequential input. RNN can store information about the previous state in order to determine the potential or provide output based on the prior state. However, RNN has the drawback of experiencing vanishing gradient/exploding gradient when too many sequences are executed. The term "vanishing/exploding gradient" refers to a circumstance in which the gradient value might be very tiny or equal to zero, as well as highly large [29]. As a result, a method known as LSTM was created to solve these flaws by using a gate system.

2.1.5. Long short-term memory (LSTM)

The LSTM is the development of RNN to solve the problem vanishing/exploding gradient. In the LSTM architecture addedgate or gate that serves to regulate what information to remember. There are additional three gates, each of which serves as an input gate, a gate to delete previous information and an output gate. With the addition of these three gates, the LSTM can better manage the stored information so that it doesn't happenvanishing/exploding gradient [30]. Figure 2 is an example of a sentiment classification application model with LSTM.

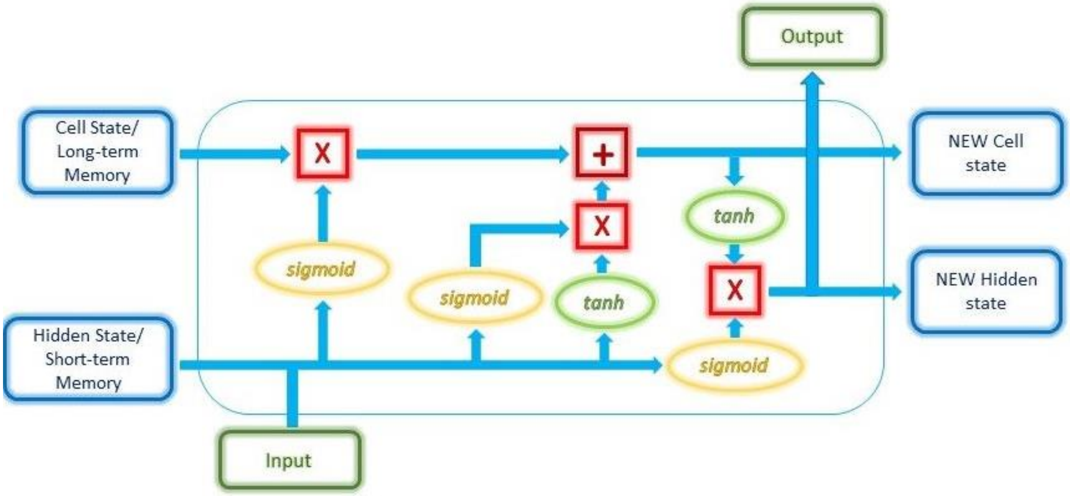

Figure 2. LSTM architecture: source [31]

2.1.6. Pre-processing

Pre-processing is a process that must be done before the data can be used for analysis. Because text data obtained from Twitter usually contains many errors such as unstructured words, writing errors, unnecessary characters, abbreviations, and other things that can make the process of word extraction and sentiment analysis unable to provide good performance [32].

Therefore, it is necessary to do pre-processing steps, to minimize or eliminate errors in the data. So that when the data is processed it will produce maximum results both accuracy and classification process. There are several schemes commonly applied for pre-processing on Twitter namely:

- Case folding, uniform characterization, so that the words in the sentence are all lowercase or uppercase. In this study the letters will be converted into lowercase.
- Stop word removal, removes words that have a weak influence or meaning in a sentence, such as the word “yang”.
- Symbol removal, removes unnecessary symbols such as URLs, @, # or extra spaces that are the most dense in tweets.
- Tweet Tokenization, at this stage the tweet in the form of a sentence will be cut every word into a token form.

2.1.7. Confusion matrix

One way to evaluate the results of sentiment analysis is to use an evaluation matrix known as the confusion matrix. In the confusion matrix, there are several things that can be evaluated from the classification results of sentiment analysis, namely accuracy, precision, sensitivity and f-score which is a combination of two sensitivity and precision evaluations [33].

Table 1 is an example of a confusion matrix, where the results of the classification each has a value of TP, FN, FP and TN. To calculate the value of each evaluation, as shown in (1)-(3):

$$Sensitivity = \frac{TP}{TP+FN} \ \dots \quad (1)$$

$$Specificity = \frac{TN}{TN+FP} \ \dots \quad (2)$$

$$Accuracy = \frac{TN+TP}{T+F} \ \dots \quad (3)$$

Table 1. Confusion matrix

	Predicted class positive	Predicted class negative
Positive	True positive (TP)	False negative (FN)
Negative	False positive (FP)	True negative (TN)

3. RESEARCH METHOD

The suggested approach is depicted in Figure 3 that includes steps used in LSTM methods such as data collection, dataset Twitter, data pre-processing techniques like stopword removal, removal of URL and mentions from tweets, removal of punctuation and digits, case folding, lemmatization and tokenization using natural language processing. Feature extraction using GloVe model training, and finally classification using the LSTM is done for result evaluation.

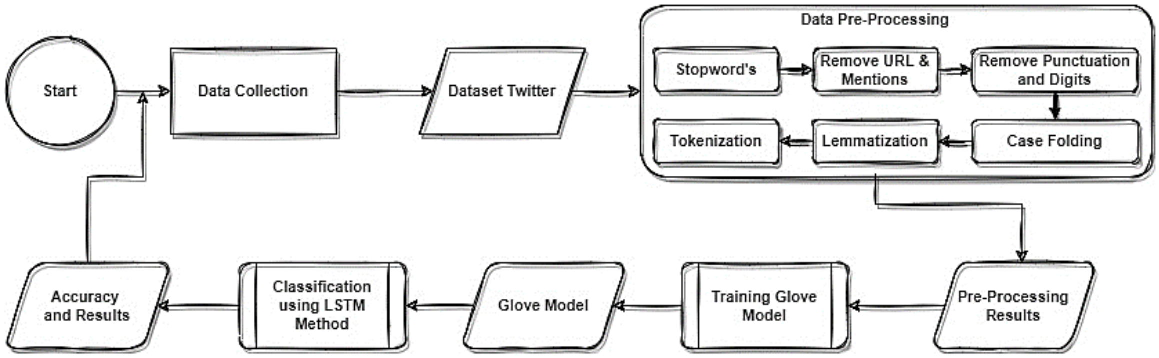

Figure 3. Proposed model: source (Self)

3.1. Data collection

In general, the data collection process is carried out using the requesting access from the application programming interface (API) provided by Twitter. To gain access to the Twitter API, one must first register with the Twitter account at Twitter Developers' resource. Thereafter, the resources grant the key and token to access the Twitter API. The Algorithm 1 listed below serves to get Tweet data from certain keywords, the results of which will be written in a CSV file containing a list of Tweets from the keywords being searched for as under:

Algorithm 1

```
Begin
Initialize Secret Key API
Define Search Keyword
Initialize CSV File to Save Tweets based on Search Query
Preserve Results into CSV File
End
```

3.2. Pre-processing

The data gathering findings will not be used right away, but will be pre-processed first. This is because the data collected on Twitter still contains a lot of characters that aren't needed, as well as faults in writing, the usage of acronyms, and other factors that can skew sentiment accuracy and classification results. Table 2 shows the phases of pre-processing that will be used in this study.

Table 2. Pre-processing of Tweets

Pre-Processing Step	Output
Original Tweet	@twista202 I motionlessly haven't examine the 9th and 10th Princess chronicle. Reduction Francesca prepared me to weep at the ending. Hmm those are straightforward books. http://ff.im/1XTTi
Stopwords	@twista202 motionlessly haven't examine9th10th Princess chronicle. Reduction Francesca prepared me weep ending Those are straightforward books. http://ff.im/1XTTi
Remove URLs and Mentions	Motionlessly haven't examined 9th10th Princess chronicle. Reduction Francesca prepared me weep ending Those are straightforward books.
Remove Punctuation and Digit	Motionlessly not examinePrincess chronicle. ReductionFrancesca prepared weep ending those are straightforward books
Case folding	Motionlessly not examine princess chronicle reductionFrancesca prepared weep ending those are straightforward books
Remove White Space	Motionlessly not examine princess chronicle reduction Francesca prepared weep ending those are straightforward books
Tokenization	[Motionlessly, not, examine, princess, chronicle, reduction, Francesca, prepared, weep, ending, those, are, straightforward, books]

3.3. GloVe model training

At this stage, the Tweets training is carried out, which will be formed as a model global vector. In this study, several vector dimensions will be used, namely dimensions 150, 200, 250, 300, and 350. The Tweets are represented as a vector by forming a matrix of word occurrences in a certain context; it can be seen in Table 3.

Table 3. Co-occurrence GloVe

	Still	Not	Read	Diaries	Saving	Cry	Easy	Books
Still	2	0	2	1	1	1	2	1
Not	0	0	9	0	0	1	0	0
Read	2	0	1	1	0	0	1	2
Diaries	1	0	0	2	0	0	0	1
Saving	1	0	0	0	2	0	2	0
Cry	1	0	0	0	0	2	0	0
Easy	2	0	0	1	0	0	2	0
Books	1	0	1	1	1	0	0	2

The calculation of the probability of the occurrence of the word is in the form as shown in (4). Then the results obtained will be calculated by the values of cost function:

$$f(X_{ik}) = \begin{cases} \left(\frac{X_{ik}}{Xmax}\right)^{\alpha} & if(X_{ik}<Xmax \\ & 1;others \end{cases} \ldots \quad (4)$$

for word weighting, which in (4) is denoted byX_{ik}, where i is vector of data used for prediction or training and k is the weight, that depicts the result of the matrix in Table 3. In this study, GloVe parameters such as $Xmax$ and alpha α values will be used in accordance with $Xmax$=100,α=0.75 and iterations are adjusted to the size of the vector, in order to get good performance. After the model is formed, the GloVe model will be embedded into the LSTM layer for sentiment classification.

3.4. Sentiment classification with LSTM

In the sentiment classification process in the LSTM model, the form of the input and its magnitude is needed. In this study, the input sequence will be used. Then in the next layer, the word embedding model will be embedded in the LSTM embedding layer. By adjusting the size of the embedded vector and also in this layer, feature extraction will be carried out where each word in the dataset will be searched for its vector weight, which will then be classified as negative and positive in sentiment. The LSTM layer in this study uses 25 LSTM units and default dropout 0.5 with a layer using ReLU with 256 activation units and 1 sigmoid activation unit after that, which can be seen in Figure 4.

The data process is classified by LSTM, starting with data that has been processed. Pre-processing looks for the form of numbers or their representations in vector form that has been created using word embedding global vector. Then enter one by one sequentially in a set to the LSTM, the results of each hidden layers will be distributed to other hidden layers along with subsequent entries. After the final result comes out, it will then be forwarded to dense layers to change the output according to rectified linear activation unit (ReLU) and Sigmoid activation functions.

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
embedding (Embedding)        (None, 25, 50)            20000050
_________________________________________________________________
bidirectional (Bidirectional (None, 25, 256)           183296
_________________________________________________________________
bidirectional_1 (Bidirection (None, 256)               394240
_________________________________________________________________
dense (Dense)                (None, 1)                 257
=================================================================
Total params: 20,577,843
Trainable params: 577,793
Non-trainable params: 20,000,050
_________________________________________________________________
```

Figure 4. LSTM model

3.5. Evaluation of classification results

The next stage is to analyse the results once the model has classified the sentiment analysis. The confusion matrix approach is used to calculate the level of accuracy, which is how accurate the model can classify correctly, by entering the value from the classification results into the equation. Then there's the amount of precision, or how to express the degree of accuracy between the desired data and the model's predicted findings.

4. RESULTS AND DISCUSSION

In this study, tests were conducted using data with a balanced number of negative and positive labels with a total of 800000, 807078 and 1607078 tweets of data. Also, data with an unbalanced number of negative and positive labels (800000 negative labels and 800000 positive labels) was used to test the hypothesis. Model GloVe is used in the test using dimensions 150, 200, 250, 300, 350, 400, and 450 using 25 epoch and the number of sequences is 50. To see the success of the test, in this study, Confusion Matrix is used to see the results of the test in the form of accuracy.

4.1. Sentiment analysis using GloVe model with balanced amount of data

In this test, testing is carried out with the Tweet data that has a balanced sentiment label with different vector dimensions. Tweet data is divided into train and test data, with percentages of test data as much as 20% of each total amount of the Tweet data. The test results can be seen in Figure 5 whereas the Sentiment 0 depicts the positive and 1 depicts the negative, respectively:

```
Expected sentiment: 0. Input: mood gray weather ... need help cheer
Expected sentiment: 0. Input: soccer game cancel today hittin gym
Expected sentiment: 0. Input: back last minute shopping town family holiday unk
Expected sentiment: 0. Input: morning think sick ehe
Expected sentiment: 1. Input: cool yeah come like hmmm ... afford thanks
Expected sentiment: 0. Input: sad miss flash rave library last night
Expected sentiment: 1. Input: handle 290 psd file unk php thats worried
Expected sentiment: 0. Input: sleep 9pm unk body need rush get photo print get class 8am
Expected sentiment: 0. Input: want one need unk al night
Expected sentiment: 0. Input: bless really well unk good tired though
Expected sentiment: 0. Input: headed home ... unk vet
Expected sentiment: 0. Input: since obviously live alaska radio station get
Expected sentiment: 0. Input: phone conversation bore bedtime
```

Figure 5. Sentiment analysis using GloVe model with balanced amount of data

4.2. Validation accuracy

Based on Figure 5, it can be seen that the level of performance for each amount of data and the vector dimensions is capable to identify the sentiments. For the best results in this test, the number of tweets is 16000000 and the vector dimensions are 300 with an accuracy rate of 95.61% as in Figure 6.

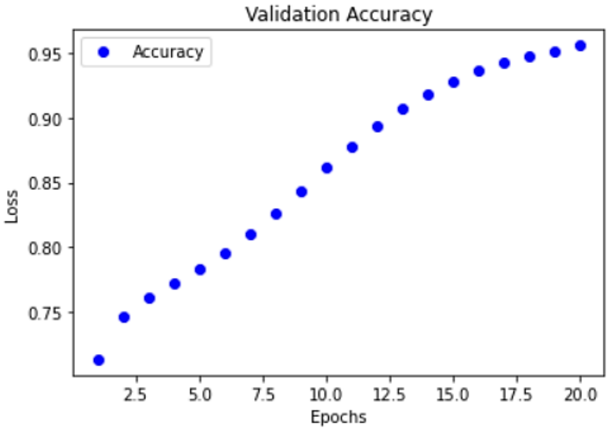

Figure 6. Validation accuracy

5. CONCLUSION

The outcomes of tests that were carried out utilizing the GloVe model as word embedding on the LSTM for sentiment analysis of Tweets can be deduced based on the aforementioned discussion. As a result, the scheme produces the best categorization results, with an accuracy rate of 95.61 percent. Following that, a level of accuracy of 1600000 Tweets was achieved, as well as GloVe 300 vector dimensions of 800000 positive and 800000 negative, respectively.

REFERENCES

[1] A. S. Imran, S. M. Daudpota, Z. Kastrati, and R. Batra, "Cross-cultural polarity and emotion detection using sentiment analysis and deep learning on covid-19 related tweets," *IEEE Access*, vol. 8, pp. 181074–181090, 2020, doi: 10.1109/ACCESS.2020.3027350.

[2] P. Vedurumudi, "Twitter sentiment analysis," *i-manager's Journal on Computer Science*, vol. 8, no. 4, p. 1, 2021, doi: 10.26634/jcom.8.4.18269.

[3] M. F. Çeliktuğ, "Twitter sentiment analysis, 3-way classification: positive, negative or neutral?," in *Proceedings - 2018 IEEE International Conference on Big Data, Big Data 2018*, Dec. 2019, pp. 2098–2103, doi: 10.1109/BigData.2018.8621970.

[4] J. Johnson, "Number of internet users worldwide 2005-2019 | Statista," *Statista*, 2005. Accessed: Feb. 15, 2020. [Online]. Available: https://www.statista.com/statistics/273018/number-of-internet-users-worldwide/

[5] Statista, "Number of social media users worldwide 2010-2021," *Statista*, Sep. 10, 2021. Accessed: Feb. 15, 2020. [Online]. Available: https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/

[6] J. Prodanova and A. Van Looy, "How beneficial is social media for business process management? a systematic literature review," *IEEE Access*, vol. 7, pp. 39583–39599, 2019, doi: 10.1109/ACCESS.2019.2903983.

[7] M. Vergeer, "Politics, elections and online campaigning: Past, present and a peek into the future," *New Media and Society*, vol. 15, no. 1, pp. 9–17, Feb. 2013, doi: 10.1177/1461444812457327.

[8] K. Hiroshi, N. Tetsuya, and W. Hideo, "Deeper sentiment analysis using machine translation technology," in *Proceedings of the 20th international conference on Computational Linguistics - COLING '04*, 2004, pp. 494-es, doi: 10.3115/1220355.1220426.

[9] K. Denecke and Y. Deng, "Sentiment analysis in medical settings: New opportunities and challenges," *Artificial Intelligence in Medicine*, vol. 64, no. 1, pp. 17–27, May 2015, doi: 10.1016/j.artmed.2015.03.006.

[10] C. Zong, R. Xia, and J. Zhang, "Sentiment analysis and opinion mining," in *Text Data Mining*, Singapore: Springer Singapore, pp. 163–199, 2021.

[11] F. A. Wenando, R. Hayami, Bakaruddin, and A. Y. Novermahakim, "Tweet sentiment analysis for 2019 indonesia presidential election results using various classification algorithms," in *Proceeding - 1st International Conference on Information Technology, Advanced Mechanical and Electrical Engineering, ICITAMEE 2020*, Oct. 2020, pp. 279–282, doi: 10.1109/ICITAMEE50454.2020.9398513.

[12] N. H. Khun and H. A. Thant, "Visualization of Twitter sentiment during the period of US banned Huawei," in *2019 International Conference on Advanced Information Technologies, ICAIT 2019*, Nov. 2019, pp. 274–279, doi: 10.1109/AITC.2019.8921014.

[13] S. Garg, "Drug recommendation system based on sentiment analysis of drug reviews using machine learning," in *Proceedings of the Confluence 2021: 11th International Conference on Cloud Computing, Data Science and Engineering*, Jan. 2021, pp. 175–181, doi: 10.1109/Confluence51648.2021.9377188.

[14] S. Ernawati, E. R. Yulia, Frieyadie, and Samudi, "Implementation of the Naïve Bayes algorithm with feature selection using genetic algorithm for sentiment review analysis of fashion online companies," in *2018 6th International Conference on Cyber and IT Service Management, CITSM 2018*, Aug. 2019, pp. 1–5, doi: 10.1109/CITSM.2018.8674286.

[15] B. Joyce and J. Deng, "Sentiment analysis of tweets for the 2016 US presidential election," in 2017 *IEEE MIT Undergraduate Research Technology Conference, URTC 2017*, Nov. 2018, vol. 2018-January, pp. 1–4, doi: 10.1109/URTC.2017.8284176.

[16] P. Preethi Krishna and A. Sharada, "Word embeddings - skip gram model," in *ICICCT 2019 – System Reliability, Quality Control, Safety, Maintenance and Management*, Singapore: Springer Singapore, 2020, pp. 133–139.

[17] R. Ni and H. Cao, "Sentiment Analysis based on GloVe and LSTM-GRU," in *Chinese Control Conference, CCC*, Jul. 2020, vol. 2020-July, pp. 7492–7497, doi: 10.23919/CCC50068.2020.9188578.

[18] P. Munjal, M. Narula, S. Kumar, and H. Banati, "Twitter sentiments based suggestive framework to predict trends," *Journal of Statistics and Management Systems*, vol. 21, no. 4, pp. 685–693, Jul. 2018, doi: 10.1080/09720510.2018.1475079.

[19] Tiara, M. K. Sabariah, and V. Effendy, “Sentiment analysis on Twitter using the combination of lexicon-based and support vector machine for assessing the performance of a television program,” in *2015 3rd International Conference on Information and Communication Technology, ICoICT 2015*, May 2015, pp. 386–390, doi: 10.1109/ICoICT.2015.7231456.
[20] H. Imaduddin, Widyawan, and S. Fauziati, “Word Embedding Comparison for Indonesian Language Sentiment Analysis, *International Conference of Artificial Intelligence and Information Technology (ICAIIT)*, Mar. 2019, doi: 10.1109/icaiit.2019.8834536.
[21] M. Bilgin and İ. F. Şentürk, “Sentiment analysis on Twitter data with semi-supervised Doc2Vec,” in *2nd International Conference on Computer Science and Engineering, UBMK 2017*, Oct. 2017, pp. 661–666, doi: 10.1109/UBMK.2017.8093492.
[22] J. Pennington, R. Socher, and C. D. Manning, “GloVe: Global vectors for word representation,” in *EMNLP 2014 - 2014 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference*, 2014, pp. 1532–1543, doi: 10.3115/v1/d14-1162.
[23] Y. Sharma, G. Agrawal, P. Jain, and T. Kumar, “Vector representation of words for sentiment analysis using GloVe,” *ICCT 2017 - International Conference on Intelligent Communication and Computational Techniques*, vol. 2018-January, pp. 279–284, 2018, doi: 10.1109/INTELCCT.2017.8324059.
[24] D. Li and J. Qian, “Text sentiment analysis based on long short-term memory,” in *2016 1st IEEE International Conference on Computer Communication and the Internet, ICCCI 2016*, Oct. 2016, pp. 471–475, doi: 10.1109/CCI.2016.7778967.
[25] J. H. Wang, T. W. Liu, X. Luo, and L. Wang, “An LSTM approach to short text sentiment classification with word embeddings,” *Proceedings of the 30th Conference on Computational Linguistics and Speech Processing, ROCLING 2018*, pp. 214–223, 2018.
[26] T. Mikolov, K. Chen, G. Corrado, and J. Dean, “Efficient estimation of word representations in vector space,” *1st International Conference on Learning Representations, ICLR 2013 - Workshop Track Proceedings*, Jan. 2013, [Online]. Available: http://arxiv.org/abs/1301.3781.
[27] W. K. Sari, D. P. Rini, and R. F. Malik, “Text classification using long short-term memory with GloVe features,” *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika*, vol. 5, no. 2, p. 85, Feb. 2020, doi: 10.26555/jiteki.v5i2.15021.
[28] T. Mikolov, Q. V. Le, and I. Sutskever, “Exploiting similarities among languages for machine translation,” *arxiv*, Sep. 16, 2013. http://arxiv.org/abs/1309.4168 (accessed Feb. 15, 2020).
[29] N. Ketkar and J. Moolayil, “Recurrent neural networks,” in *Deep Learning with Python*, Berkeley, CA: Apress, 2021, pp. 243–285.
[30] J. Korstanje, “LSTM RNNs,” in *Advanced Forecasting with Python*, Berkeley, CA: Apress, 2021, pp. 243–251.
[31] Blog.floydhub.com, “Figure of LSTM Architecture,” blog.floydhub.com, 2019. Accessed: Feb. 15, 2020. [Online]. Avalable: https://blog.floydhub.com/content/images/2019/06/Slide19.JPG
[32] Z. Jianqiang and G. Xiaolin, “Comparison research on text pre-processing methods on twitter sentiment analysis,” *IEEE Access*, vol. 5, pp. 2870–2879, 2017, doi: 10.1109/ACCESS.2017.2672677.
[33] W. W. Piegorsch, “Confusion Matrix,” in *Wiley StatsRef: Statistics Reference Online*, Wiley, 2020, pp. 1–4.

BIOGRAPHIES OF AUTHORS

Kusum is currently pursuing Ph.D at Manav Rachna International Institute of Research and Studies, Faridabad, India. She completed her B.E. from M.D.U. University, Rohtak in 2002 and M.Tech from Banasthali Vidyapith, Jaipur in 2006.She has Over 11 years of experience in Academics and Research and published papers in various National and International Journals. She can be contacted at email: kusumprerak@gmail.com.

Supriya P. Panda is working as a professor (CSE) and HoD (CSE) at Manav Rachna International Institute of Research & Studies at Faridabad (India) since 2016. She started her stint at BITS, Pilani, and Rajasthan, India wherein she served for ten years. She has over three decades of experience in academics across various organizations in India and abroad. She has numerous research papers in reputed national and International Journals to her credit. She can be contacted at email: supriya.fet@mriu.edu.in.