Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2023 SCPE. Volume 24, Issues 3, pp. 315–326, DOI 10.12694/scpe.v24i3.2138

A DEEP LSTM-RNN CLASSIFICATION METHOD FOR COVID-19 TWITTER REVIEW BASED ON SENTIMENT ANALYSIS

JATLA SRIKANTH* AND AVULA DAMODARAM SHANMUGAM†

Abstract. In today's world, advanced internet technologies have significantly increased people's affinity towards social networks to stay updated on current events and communicate with others residing in different cities. Social opinion analyses helped determine the optimal public health response during the COVID-19 pandemic. Analysis of articulating tweets from Twitter can reveal the public's perceptions of social distance. Sentiment Analysis is used for classifying text data and analyzing people's emotions. The proposed work uses LSTM-RNN with the SMOTE method for categorizing Twitter data. The suggested approach uses increased characteristics weighted by attention layers and an LSTM-RNN-based network as its foundation. This method computes the advantage of an improved information transformation framework through the attention mechanism compared to existing BI-LSTM and LSTM models. A combination of four publicly accessible class labels such as happy, sad, neutral, and angry, is analyzed. The message of tweets is analyzed for polarization and subjectivity using TextBlob, VADER (Valence Aware Dictionary for Sentiment Reasoning), and SentiWordNet. The model has been successfully built and evaluated using two feature extraction methods, TF-IDF (Term Frequency-Inverse Document Frequency) and Bag of Words (BoW). Compared to the previous methodologies, the suggested deep learning model improved considerably in performance measures, including accuracy, precision, and recall. This demonstrates how effective and practical the recommended deep learning strategy is and how simple it is to employ for sentiment categorization of COVID-19 reviews. The proposed method achieves 97% accuracy in classifying the text whereas, among existing Bi-LSTM, achieves 88% maximum in the text classification.

Key words: Sentiment Analysis, Covid-19, deep learning, Twitter reviews, social networks, classification, SentiWordNet, TextBlob, Bag of Words.

1. Introduction. The worldwide pandemic of coronavirus disease (COVID-19) has adverse effects on human health [29]. From its first detection in Wuhan, China, it has already spread to multiple nations on all continents, and on March 11, 2020, the World Health Organization (WHO) declared it a pandemic. On numerous social media platforms, there has been a lot of COVID-19 information shared. Moreover, misleading news might propagate on social media platforms like Twitter; thus, it's crucial to comprehend people's emotions from such textual resources [26]. The identification of attitudes from Twitter data on material about COVID-19 can be made with the help of deep learning algorithms. Moreover, given the difficulties involved in textual analysis, there still needs to be a technological issue regarding how deep learning networks may be adapted and adjusted to obtain high accuracy.

At the beginning of the COVID-19 outbreak, social media quickly became a vital communication method for creating, disseminating, and consuming data. The discovery and characterization of infectious disease outbreaks and understanding people's sentiments, behaviors, and views have all benefited from including social media data in numerous research. Social media user-generated material can be opinionated or inaccurate, frequently including false data and conspiracy theories [17]. The practice of categorizing feelings in qualitative information utilizing machine learning (ML) and natural language processing is known as sentiment analysis, commonly referred to as opinion mining (NLP) [4].

Different methods for evaluating consumer emotion in social media data have been presented in previous research [27, 31, 20, 1, 25]. Propose a hybrid machine learning technique, for instance, to categorize user emotions as either positive or negative. In a collection of customer reviews, the [12] authors employ natural language processing (NLP) to decipher user sentiment. A Bayesian graphical method is used to assess data

*Department of Computer Science and Engineering, Aurora's Technological and Research Institute, Parvathapur, Uppal, Hyderabad, TS, India (jatlasrikanthres1@outlook.com)
†School of Information Technology (SIT), JNTUH, Hyderabad, TS, India

from Twitter. Such practices may be computationally challenging, time-consuming, and not always offer high precision.

This work provides a deep learning method for categorizing the emotion of tweets about COVID-19. Positive and negative feelings can be distinguished through an analysis of people's opinions, and efforts are being made to detect texts in the literature expressing negative emotions [11, 7]. The rectified linear activation function (ReLU) is frequently employed because of the low complexity of training and the potential for higher results. Most critically, nonlinear connections are commonly studied using ReLU. Because ReLU can help convolution neural networks better capture complicated patterns, it has been used in several research neural network models [8].

Several frameworks, such as the Long-Short Term Model (LSTM), have been suggested to sidestep the constraints of neural networks. LSTM has been recognized as a crucial component because it effectively solves time-series data and sequentially issues [19]. Due to its capacity to understand text sequence and identify relationships among words or phrases in sentiment analysis, the suggested deep learning strategy incorporates the LSTM non-linear activation [14]. The main objective of the research is to classify the sentiment of the text with high accuracy. The research question relies on

1. What is the effectiveness of using LSTM-RNN with SMOTE method and attention layers for sentiment analysis of COVID-19-related tweets, compared to previous methodologies using TextBlob, VADER, and SentiWordNet?
2. How the sentimental analysis is performed effectively on the covid 19 twitter data?

The proposed method uses LSTM- recurrent neural network-based SMOTE for classifying the Covid-19 Twitter reviews.

The main contribution of the proposed method is given below:

1. Utilizing attentive learning and LSTM, the primary aim is to maximize effective weight by the semantic relevance of terms.
2. To evaluate the effectiveness of TF-IDF, BoW, and feature extraction methods, as well as the results of various deep learning models for sentiment analysis using different annotation techniques like TextBlob, VADER, and SentiWordNet.
3. SMOTE helps to reduce the imbalanced issues of datasets. also, with the help of the majority class, it balances the datasets.

The rest of our research article is written as follows: Section 2 discusses the related work on COVID-19 Twitter reviews and Deep learning classification methods. Section 3 shows the general working methodology of the proposed work. Section 4 evaluates the implementation and results of the proposed method. Section 5 concludes the work and discusses the result evaluation.

2. Related Works. The COVID-19 pandemic has caused massive deaths of innocent people throughout the globe and constitutes a severe risk to the food supply, the workplace, and public health. Numerous people have died due to the COVID-19 epidemic, which also poses a severe threat to workplace safety, food production, and the promotion of healthy lifestyles. The characteristics of COVID-19 comprise death, transmission characteristics, period of national viruses, and initial fatalities. This is exposed by the disparities in social networking and global financial responses due to the extreme viral spread's residual effects. As users utilized the channels to vent their feelings during the lockdown, social networking sites were crucial in sharing data well about epidemics around the globe. Given this critical scenario, it is essential to look at how people are responding on Twitter while considering standard terms relevant to the outbreak [5].

To read or understand all that's been stated on Twitter about COVID-19 immunizations would only be possible for a person. Moreover, we can utilize natural language processing (NLP) techniques to examine an extraordinarily complex and wide-ranging discourse utilizing word cloud visualizations, sentiment analysis, and linguistic extraction of features. Using deep learning classifiers, the author conducted a study on the sentiment classification of COVID-19 tweets [6]. This study found that tweets regarding the COVID-19 pandemic did not influence people. The results demonstrate that neither the number of words in tweets nor WordCloud includes helpful comments. The assertions are validated by a suggested deep learning classifier model, which has an accuracy of up to 81 percent. According to the researchers, a fuzzy approach built on Gaussian participation may accurately identify tweet attitudes.

Using the BERT model, the author has suggested conducting a sentiment analysis of how the coronavirus

Table 2.1: Previous studies based onCovid-19 Twitter reviews

Author	Methodology Used	Dataset	Objective
Relucio, F.S.; Palaoag, T.D [22]	Web analytics strategy	There are 1717 tweets, all were gathered on Twitter	Observe the attitude of informative posts.
Bhagat, K.K.; Mishra, S. [3]	Using a thesaurus as a guide	There are 154 documents gathered from Google	Public perception of COVID-19's online education.
Ashwitha, R.; Jeevitha [1]	NLP methods and a classification using logistic regression	Data collection is done using Google Forms.	Evaluation of Public Opinion on the COVID-19 Epidemic's Teaching
Althagafi, A.; Althobaiti, [2]	Random forest, KNN, and Naive Bayes	The Twitter API was used to collect 10,445 Twitter posts.	Coronavirus sentiment classification in online learning
Mamtesh, M.; Mehla, [16]	Logistic Regression and KNN	Evaluations of movies with emotional commentary	The information was gathered from a multitude of sources.
Imran, A.S.; Daudpota, S.M. [24]	In-depth learning (Multilayer LSTM)	There are 27,357 tweets in total relating to COVID-19	COVID-19 sentiment classification
Chintalapudi, N.; Battineni, G. [10]	In-depth learning (BERT and LSTM)	There were 3090 tweets about COVID-19 overall.	COVID-19 sentiment classification

affects social interactions. They assert that since Twitter has grown to be one of the most well-known social media platforms, the authors used the BERT model to conduct sentiment analysis on Twitter to comprehend people's emotions and psychological states better [28]. In this study, the authors performed a sentiment classification on two variables: the first group had Twitter from users around the globe, and the other set contained tweets from users in India. They checked the reliability of the emotion classification using the GitHub repository. The experimental results show that the accuracy rate is 94%.

3. Proposed Twitter Analysis Methodology . The proposed method uses LSTM-RNN with SMOTE for classifying the twitter text data. Initially, it collects the data from the Twitter reviews and then the collected data is preprocessed by using tokenization, URL removal, stop word removal, stemming, and lemmatization. Next, the TextBlob is used to compute the polarity score, and then the feature engineering process is done. Next, the LSTM-RNN with SMOTE method is used for assigning efficient weights. Figure 3.1 shows the architecture of the proposed method. first the twitter data is collected and preprocessed by removing URL, Noises, spaces, etc from the dataset. Text Blob is based on the Naive Bayes algorithm and is suitable for analyzing short and informal texts, such as social media posts and product reviews. This helps to classify text polarity of the context. further the features are extracted and analyzed using frequency of the terms and bag of words. Feature matrix is computed by the training and testing the features. Finally, classification is done using LSTM-RNN with SMOTE to obtain the optimal results.

3.1. Data Collection. On Twitter, companies may engage with customers in much more intimate ways. But choosing which tweets to be responded to first can be challenging for marketers because Twitter contains so much information. Sentiment analysis has become necessary as a technique in social media management initiatives. Using sentiment analysis technology, social media activities are automatically monitored for emotion. The initial stage in solving the topic of sentiment classification of a Covid 19 tweets is information gathering [18]. We took advantage of Tweepy and Twitter's streaming API to follow specific popular phrases and profiles during information gathering. Utilizing the Twitter API to create the large dataset entails a sequence of iterations. This involves setting up an identity, installing Tweepy, running a small test, looking at a tweet's JSON file,

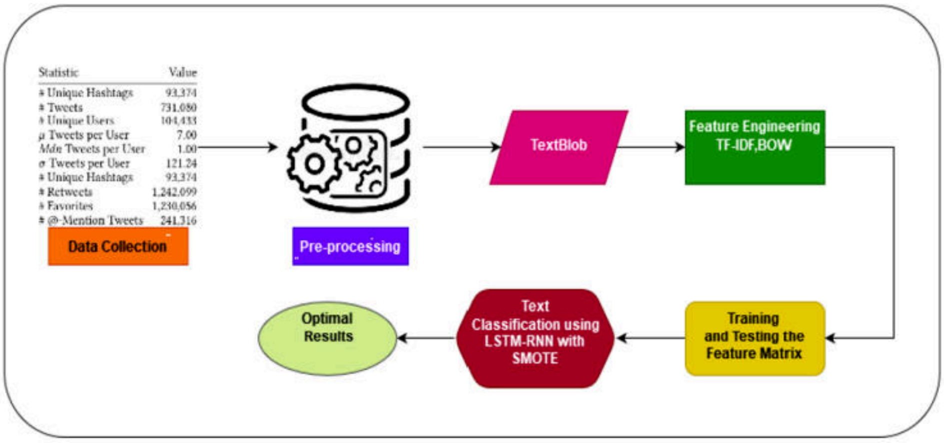

Fig. 3.1: Architecture of Proposed Method

extracting the information, and gathering the data. Following data collection, preprocessing must be done before moving on.

3.2. Data Pre-processing. Applications that analyze the data must pre-process the data to eliminate extraneous information and speed up the learning of categorization algorithms. Any material that makes almost no contributions to determining the desired class is referred to as redundant data, it expands the feature representation and adds needless computing complexity. As a result, classification techniques perform worse when preprocessing is neglected or done incorrectly. So, before preprocessing, data screening or pretreatment is done.

The second phase includes both exploratory data analysis and data pre-processing. Once the data had been cleaned to extract meaningful information, the raw tweets could not produce objective findings for Sentiment predictions. The major obstacles were #tags, @mentions, URLs, and stop words in tweets. The #tags, @mentions, and URLs again from tweets are substituted using regular expressions. Stop words are eliminated using the Python NLTK library.

3.2.1. URL removals. Although they have yet to add new definitions for training systems to simplify the feature space, URL links, tags, punctuation, and numerals do not improve classifier performance. By eliminating them, the feature representation is made simpler.

3.2.2. Stop words removal. Stop words are often used yet provide no meaningful data for the research. Stop words like "the," "is," "a," and "an" are eliminated.

3.2.3. Stemming and Lemmatization. While stemming and lemmatization focus on reducing a word's inflectional forms and, occasionally, its derivationally related forms to a basic shape. In this way, words like "walks," "walking," and "walked" are changed to the core word "walk."

3.3. TextBlob. The lexicon-based TextBlob method can be applied to various Natural language processing tasks, such as part-of-speech labeling, sentiment classification, noun word retrieval, paraphrasing, and grouping [30]. It was employed in this investigation for sentimental reasons. A polarization score between one and one is provided by the TextBlob emotion method. Less than zero indicates a negative comment, zero indicates a reasonable statement, and more than zero shows a positive assertion in a tweet [23].VADER uses a rule-based approach to analyze the polarity of text and can handle the complexity of social media language, such as slang, emoticons, and abbreviations. VADER provides not only the text's polarity but also the sentiment's intensity, making it useful for fine-grained sentiment analysis. SentiWordNet uses a hybrid approach that combines machine learning and lexicon-based methods to analyze the sentiment of text. SentiWordNet helps analyze longer texts and can provide a more detailed analysis of the sentiment by considering the sentiment of each word in context.

3.4. Feature Engineering. The second most prominent feature extraction techniques, BoW and TF-IDF, are employed to collect features from tweets.

3.4.1. Bag of Words. The boW is a straightforward method for extracting characteristics from condensed texts or information and is frequently used in information retrieval and natural language processing. The boW should be used to estimate the appearance of a phrase in a word and create a feature representation that includes the frequency of each particular comment. The primary purposes of the BoW are to expand the dictionary of all unique terms and to train the teaching methods using their frequency distributions.

3.4.2. Term Frequency-Inverse Document Frequency. For extracting the features, a TF-IDF method is employed to collect weighted features using textual information. To help learning models perform much better, it offers the value of every phrase in the library. The result of TF and IDF is TF-IDF. TF is calculated by using the following:

$$TF\,(t,d) = \frac{n_t}{N_{T,d}} \tag{3.1}$$

Here $N_{T,d}$ denotes the set of terms T in the documents, and n_t is the number of times the term t appears in content d. The IDF of a word reveals its importance over the entire set and could be determined as follows:

$$IDF = \log \frac{Doc}{n_d} \tag{3.2}$$

Here n_d is the range of resources where its term t occurs, and Doc is the maximum number of documents in the collection. The formula for calculating TF-IDF utilizing TF and IDF is

$$\text{TF-IDF} = TF \cdot IDF \tag{3.3}$$

3.5. Text Classification using LSTM-RNN with SMOTE method.

3.5.1. SMOTE. By rebalancing the number of observations for each class in a database, SMOTE (Synthetic Minority Oversampling Technique) is utilized to overcome the challenges associated with imbalanced datasets. Through creating synthetic examples of minority classes, it is possible to establish balance by matching the quantity of minority class samples with those of the majority class label.

Following using TextBlob, the proportion of feelings is not identical, which makes it possible for algorithms to generalize the skewed information. SMOTE is used to equalize the dataset by producing inaccurate numbers for the minority class to prevent this over-fitting issue.

3.5.2. Training and Testing the text classification using LSTM-RNN with SMOTE. This study aims to investigate sentiment classification accuracy using tweet data using a deep learning architecture with attention levels. There are many additional processing stages for the LSTM-RNN with SMOTE architecture that is presented. The convolution of features comes next. The LSTM-RNN model is fed the input COVID-19 tweets throughout this stage. This phase aims to identify significant lexical elements from the sentence construction. Additionally, the LSTM-RNN method generates feature vectors and determines the temporal link between features. To minimize the features overlap, the weights provided by the LSTM-RNN mapping characteristics are modified after the feature mapping has been completed. Utilizing the learning algorithm to improve weights aids in the selection of pertinent details. Figure 3.2 shows the workflow of the proposed method.

To train the dataset, we also take into account the semantic information of the tweets and produce a second set of keywords with numbers attributed to the emotions conveyed in those tweets: afraid = 0, sadness = 1, angry = 2, and happiness = 3. Recognizing that this allocation of emotion labels is arbitrary, the goal is to produce a matrices level that can be input to LSTM-RNN to produce results for sentiment classification.

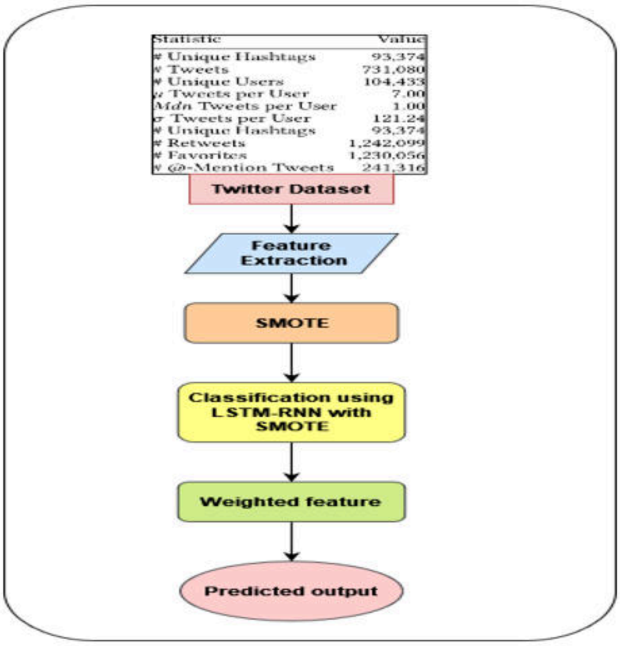

Fig. 3.2: Workflow of Proposed Method

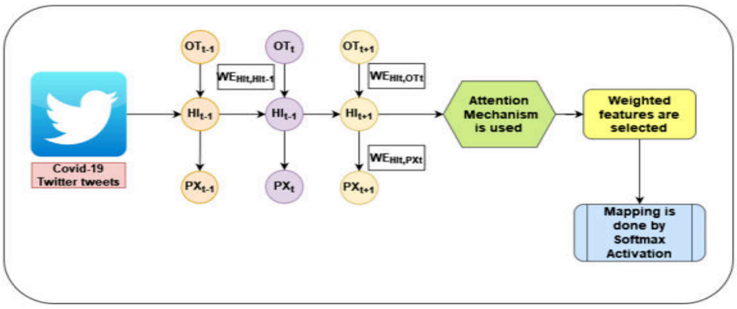

Fig. 3.3: Structure of RNN

The RNN model recognizes the series of pixels $P_x = PX_1, PX_2,......, PX_n$, generates hidden layer $HI = HI_1, HI_2,......., HI_n$, and generates output states $OT = OT_1, OT_2,......, OT_n$ in the manner shown in Figure 3.3.

$$OT_t = \sigma(WE_{HI_t OT_t} + b_t) \tag{3.4}$$

$$HI_t = \sigma(WE_{HI_{t-1}\ HI_t} HI_{t-1} + WE_{PX_t HI_t PX_t} + b_{HI_t}) \tag{3.5}$$

In this case, WE $_{HItOTt}$ stands for the vector from the hidden unit HI $_t$ and the output unit OT_t, HI_{t1} for the hidden unit for a t-1 pixel series, $WE_{HIt1HIt}$ for the sequencing period t, and b_{HIt} and b_t for bias.

The LSTM-enhanced RNN's portion is shown in a modeling framework for RNN in Figure 3.3 above, along with suggestions made. Additionally, the LSTM stacking will be used to train the timing series characteristics in issues containing a single range of views, which is necessary for a model to learn from in ability to forecast the subsequent value in a sequence.

$$ig_t = \tanh(WE_{PX_t ig_t} PX_t + WE_{HI_{t-1} ig_t} HI_{t-1} + b_{ig_t}) \tag{3.6}$$

$$PX_t = \sigma(WE_{PX_tPE_t}PX_t + WE_{HI_{t-1}PE_t}HI_{t-1} + b_{pX_t}) \tag{3.7}$$

$$f\mathrm{g}_t = \sigma(WE_{PX_tf\mathrm{g}_t}PX_t + WE_{HI_{t-1}f\mathrm{g}_t}HI_{t-1} + b_{f\mathrm{g}_t}) \tag{3.8}$$

$$OTPX_t = \sigma(WE_{PX_tOTPX_tPX_t} + WE_{HI_{t-1OTPX_t}}HI_{t-1} + b_{OTPX_t}) \tag{3.9}$$

$$CE_t = ce_{t-1}\Theta f\mathrm{g}_t + i\mathrm{g}_t\Theta PX_t \tag{3.10}$$

$$HI_t = \tanh(CE_t)\Theta OTPX_t \tag{3.11}$$

Here $i\mathrm{g}_t$ stands for the input nodes, PX_t for predictions in the first levels, $f\mathrm{g}_t$ for forget gate, Ht for output data, $b_{i\mathrm{g}_t}$, b_{pX_t}, $b_{f\mathrm{g}_t}$, and b_{OTPX_t} for bias vectors, Cet for cell state, and $WE_{PX_tf\mathrm{g}_t}$ for weight matrices. To retrieve linguistic information of the input tweets, it is possible to mix both RNN and LSTM systems.

Importantly, as mentioned in [15], we use attention layer in the research to enhance learning from features and feature weights. The LSTM-RNN is being used to generate features that are weighted by the attentiveness process to understand phrase sequences. Additionally, the usage of secondary labeling in conjunction with LSTM-RNN makes it easier for learners to gain more technical knowledge. The concentration functional of the attention layer analyzes distribution of weight and calculates arrays for the various levels, according to the research [21].

As a result, given an input PX_i, the characteristics $\mathrm{f}(\mathrm{PX}_i, \mathrm{PX}_{i+1})$ are those produced from the second layer, and $\mathrm{f}(\mathrm{PX}_i, \mathrm{PX}_{i+1}$,……, PX_i + L-1) are those produced from the L-th layer. These features values represent the reactions of multi-scale n-grams namely the unigram PXi, bigram PX_i PX_{i+1}, and L-gram f(PX_i PX_{i+1} …… PX_i + L-1). The filtering ensembles and reweight scale were utilized together in the focusing process. Additionally, taking the descriptors as input, scale reweight is employed to calculate the SoftMax distribution of attention weights and produces weighed feature values for suggests that the market [9, 13].

$$S_l^i = FEL_{ensm}(PX_l^i) \tag{3.12}$$

$$PX_{atten}^i = \sum_{j=1}^{L} \alpha_L^i PX_L^i \tag{3.13}$$

$$\partial_i^L = Softmax(MLP(PX_{atten}^i)) \tag{3.14}$$

Accuracy, precision, and recall are the three performance criteria used to evaluate the effectiveness of the suggested deep learning model. The originality of the suggested strategy lies in the enhancement of feature weighting achieved through the attention layering method.

The suggested method extracts textual information from an LSTM-RNN series that has been SMOTE-mapped; the LSTM creates a series of comments for every inputs. The vectors employed in this study are essentially a concatenated of the encoder's hidden layers, as well as the attention layer technique then refines the characteristics. The softmax activation mechanism significantly enhances the feature representation assistance provided by the attention mechanism.

4. Result Analysis. The Python language and the TextBlob and Tweepy libraries are used in the study to examine the emotion of Twitter posts. Employing phrases, tags, tweets, patterns, or geolocation, Tweepy enables us to locate pertinent information. The following command is used to search tweets that have the COVID hash tag. For instance the twitter posts are given below [29].

api = tweepy.API(auth, wait_on_rate_limit=True) #important

Table 4.1: Description of the Covid19 tweets dataset's

Characteristics Used	Total number of Tweets	Unique Tweets	Tweets (%)
Hashtag #	4317928	683108	40%
Mention @	6942379	1351963	50%
Entities	12826437	385273	80%

Table 4.2: Instances based on Covid-19 tweets

Instances based on Covid-19 tweets	Emotional Kind
Medical centre for Good Visions is moving all of its patients to make a way for healthy COVID-19 situations.	Mixed opinions
Anyone who supports Boris getting the Corona Virus is a total cunt, even a fellow nationalist.	Negative or Sad
Hello, Twittizens. I hope your day is free of coronas.	Positive or Happy
The coronavirus might be prevented from entering if I shut my front doors.	Angry
Attackers make fake coronavirus maps to spread malware to visitors.	Afraid
The thought of Jacob's Nashville show getting postponed makes my heart ache so deeply. Leave now, please.	Negative / Sad

Table 4.3: Experimental Results using Various Classifiers

Classifiers Used	Accuracy (%)	Precision (%)	Recall (%)
Bi-LSTM	88.65%	83%	81%
LSTM-RNN	81.5%	92%	79%
CNN-LSTM	78%	80%	75%
Proposed Method	**97.5 %**	**98%**	**89%**

$$self.tweets = tweepy.Cursor(api.search,\ q='COVID19,\ lang=''en'').items \tag{4.1}$$

The trials made use of the prepared dataset, which had classes that were positive, negative, and neutral. 50,000 tweets in all, 15,000 of which were used for the testing cycle, were used during the training phase. The suggested method combines a number of classifiers such as LSTM-RNN, Bi-LSTM and CNN-LSTM. Table 2.1 shows the total number of Covid-19 tweets used for training and testing.

We conducted several experiments to find the parameters for the pre-processing methods and classifiers that would produce most precise findings. The selected pre-processing techniques are fully detailed in Section 3. In Table 4.2 the sample instances based on Covid-19 tweets and the emotional category. Like these instances were used for training and testing the datasets.

The proposed LSTM-RNN-SMOTE method uses following metrics for evaluating the Covid-19 twitter tweets. The metrics are accuracy, precision, recall, ROC curve and Pie chart for displaying the emotions of people.

Table 4.3 shows the evaluation results of various deep learning classifiers such as LSTM-RNN, Bi-LSTM, CNN-LSTM and Proposed methods are given. Among all the experimental results the proposed method outperforms better.

In figure 4.1 shows the accuracy of various deep learning classifiers. It uses Covid-19 tweets for classifying the emotions. The proposed method LSTM-RNN-SMOTE achieves 97.5% of accuracy in classifying the emotions of people comparing with other classifiers. The Bi-LSTM achieves 88.65%, LSTM-RNN achieves 81.5% and CNN-LSTM achieves 78%.

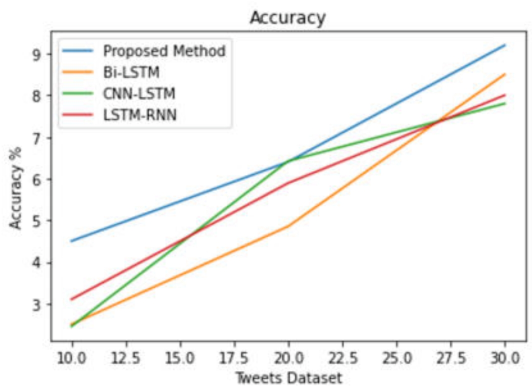

Fig. 4.1: Accuracy

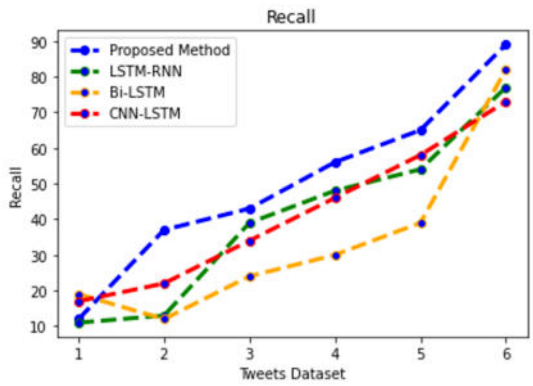

Fig. 4.2: Recall

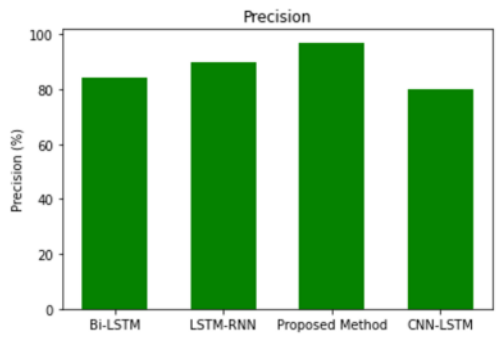

Fig. 4.3: Precision

In figure 4.2 shows the recall of various deep learning classifiers. It uses Covid-19 tweets for classifying the emotions. The proposed method LSTM-RNN-SMOTE achieves 89% of precision in classifying the emotions of people comparing with other classifiers. The Bi-LSTM achieves 81%, LSTM-RNN achieves 79% and CNN-LSTM achieves 75%.

In figure 4.3 shows the precision of various deep learning classifiers. It uses Covid-19 tweets for classifying

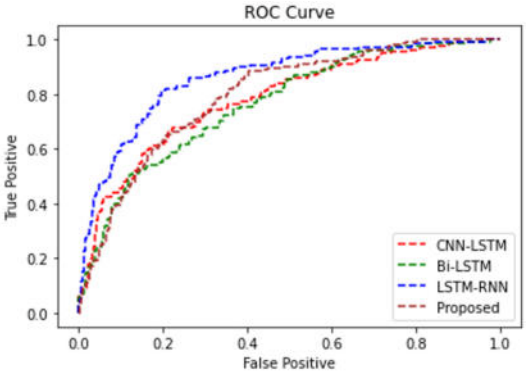

Fig. 4.4: ROC Curve

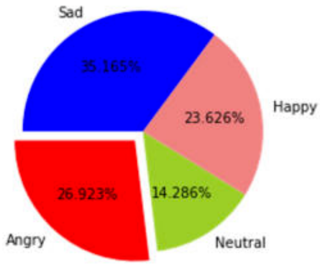

Fig. 4.5: The emotions are classified and displayed in Pie Chart

Table 4.4: Comparative Analysis of Various Activation Function

Activation Function	**Accuracy**	**Precision**	**Recall**
Sigmoid	84.75%	83.3%	82.65%
RELU	85.36%	82.75%	84.25%
TANH	83.69%	80.95%	80
LEAKY RELU	87.32%	83.57%	85.63%

the emotions. The proposed method LSTM-RNN-SMOTE achieves 98% of precision in classifying the emotions of people comparing with other classifiers. The Bi-LSTM achieves 83%, LSTM-RNN achieves 92% and CNN-LSTM achieves 80%.

Considering all of these difficulties, the classifiers demonstrated excellent results in differentiating COVID-19 tweets, indicating that current deep learning techniques can generate very efficient and convenient classifiers when used with a Covid-19 tweet dataset. Traditional classifiers do better with greater ROC values than machine learning does. The empirical findings clearly show that classification algorithm can be a useful method for enhancing the efficiency of conventional classifications. In figure 4.4 shows the ROC curve performance.

In figure 4.5 shows the pie chart with the percentage of classified emotions such as Happy, Sad, Neutral and Angry by using Covid-19 tweets. It shows that Happy emotions are 3.626%, Sad emotions are 35.165%, Neutral emotions are 14.286% and Angry emotions are 26.923%. From the results we understand during pandemic situation the people mostly tweets sad emotions and then the angry emotion percentage is higher.

Also, the proposed method uses activation function in attention layer. The activation function findings are shown in Table 4.4, which unmistakably demonstrates that four attention layers with activation function

Leaky ReLU offered greater accuracy, precision and recall in compared to certain other activation functions. The accuracy, precision and recall for the Leaky ReLU activation function are shown to be 87.32 percent, 83.57 percent and 85.63 percent respectively. This is significantly greater than the accuracy displayed by ReLU, which is 85.36 percent with a precision and recall of 82.75 percent and 84.25 percent.

As can be shown, Leaky ReLU outperforms other activation functions taken into account TANH, sigmoid, and ReLU in terms of accuracy, precision and recall values. Leaky ReLU is a backpropagation-based activation function that affects both forward and backward learning in LSTM-RNN. The interpretation of the results would depend on the proportion and distribution of each sentiment category in the analyzed text. if the analysis finds a high proportion of negative sentiment, it may suggest that the public's perceptions of social distancing measures are largely negative and that there is a need for improved communication and public engagement around the issue. Similarly, if there is a high proportion of positive sentiment, it may suggest that the public is generally supportive of social distancing measures and that there may be opportunities to build on this positive sentiment to promote greater compliance with public health guidelines.

5. Conclusion. Social networking websites have risen in popularity and are now a powerful way to influence people and educate the common person. Therefore, sentiment analysis for short communications like Twitter is especially difficult because the texts lack relevant information. As an outcome, many methods are constantly being developed to produce the best sentiment model analysis outcome. To finish the classification procedure, a preparatory stage of text pre-processing and extraction of features is necessary. We do numerous tests on various produced databases because pre-processing actions have an effect on classification performance. As a result, this research provides a deep learning method for analyzing the emotions in Twitter messages on COVID-19 opinions. The technique uses enhanced featured weighting from an attention layer and is dependent on an LSTM-RNN-SMOTE network. Through the attention mechanism, this algorithm takes advantage of an improved feature transformation framework. In this investigation, four class labels from a publicly accessible Twitter database (sad, happy, neutral, and angry) were employed. The suggested deep learning strategy improved significantly the performance metrics when compared to existing methods, with gains of 25% in accuracy, 11%–14% in precision, but only 15% in recall. Therefore, the suggested deep learning method for classifying the emotion of COVID-19 evaluations is proven to be effective and practical. The suggested approach is also compared with LSTM-RNN, CNN-LSTM, and Bi-LSTM. The proposed method classifies emotions more efficiently.

In the future, the for-sentiment analysis can be applied to other health-related applications, such as analyzing the sentiment of patient reviews of healthcare providers or identifying negative sentiment around certain health-related products or services.

REFERENCES

[1] M. Aflakparast, M. de Gunst, and W. van Wieringen, *Analysis of twitter data with the bayesian fused graphical lasso*, PloS one, 15 (2020), p. e0235596.

[2] A. Althagafi, G. Althobaiti, H. Alhakami, and T. Alsubait, *Arabic tweets sentiment analysis about online learning during covid-19 in saudi arabia*, International Journal of Advanced Computer Science and Applications, 12 (2021).

[3] K. K. Bhagat, S. Mishra, A. Dixit, and C.-Y. Chang, *Public opinions about online learning during covid-19: a sentiment analysis approach*, Sustainability, 13 (2021), p. 3346.

[4] M. Bhat, M. Qadri, M. Kundroo, N. Ahanger, B. Agarwal, et al., *Sentiment analysis of social media response on the covid19 outbreak*, Brain, behavior, and immunity, 87 (2020), p. 136.

[5] A. K. Chakraborty, S. Das, and A. K. Kolya, *Sentiment analysis of covid-19 tweets using evolutionary classification-based lstm model*, in Proceedings of Research and Applications in Artificial Intelligence: RAAI 2020, Springer, 2021, pp. 75–86.

[6] K. Chakraborty, S. Bhatia, S. Bhattacharyya, J. Platos, R. Bag, and A. E. Hassanien, *Sentiment analysis of covid-19 tweets by deep learning classifiers—a study to show how popularity is affecting accuracy in social media*, Applied Soft Computing, 97 (2020), p. 106754.

[7] S. Chetanpal, W. Santoso, and G. Srimannarayana, *A deep learning approach for human face sentiment classification*, in 2021 21st ACIS International Winter Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD-Winter), IEEE, 2021, pp. 28–32.

[8] J. Han, K. Qian, M. Song, Z. Yang, Z. Ren, S. Liu, J. Liu, H. Zheng, W. Ji, T. Koike, et al., *An early study on intelligent analysis of speech under covid-19: Severity, sleep quality, fatigue, and anxiety*, arXiv preprint arXiv:2005.00096, (2020).

[9] A. Hernández and J. M. Amigó, *Attention mechanisms and their applications to complex systems*, Entropy, 23 (2021), p. 283.

[10] R. S. Jagdale, V. S. Shirsat, and S. N. Deshmukh, *Sentiment analysis on product reviews using machine learning techniques*, in Cognitive Informatics and Soft Computing: Proceeding of CISC 2017, Springer, 2019, pp. 639–647.

[11] Z. Jianqiang, G. Xiaolin, and Z. Xuejun, *Deep convolution neural networks for twitter sentiment analysis*, IEEE access, 6 (2018), pp. 23253–23260.

[12] N. Jnoub, F. Al Machot, and W. Klas, *A domain-independent classification model for sentiment analysis using neural models*, Applied Sciences, 10 (2020), p. 6221.

[13] S. Kardakis, I. Perikos, F. Grivokostopoulou, and I. Hatzilygeroudis, *Examining attention mechanisms in deep learning models for sentiment analysis*, Applied Sciences, 11 (2021), p. 3883.

[14] X. Li, Y. Rao, H. Xie, R. Y. K. Lau, J. Yin, and F. L. Wang, *Bootstrapping social emotion classification with semantically rich hybrid neural networks*, IEEE Transactions on Affective Computing, 8 (2017), pp. 428–442.

[15] N. Liu, B. Shen, Z. Zhang, Z. Zhang, and K. Mi, *Attention-based sentiment reasoner for aspect-based sentiment analysis*, Human-centric Computing and Information Sciences, 9 (2019), pp. 1–17.

[16] M. Mamtesh and S. Mehla, *Sentiment analysis of movie reviews using machine learning classifiers*, Int. J. Comput. Appl, 182 (2019), pp. 25–28.

[17] M. Mujahid, E. Lee, F. Rustam, P. B. Washington, S. Ullah, A. A. Reshi, and I. Ashraf, *Sentiment analysis and topic modeling on tweets about online education during covid-19*, Applied Sciences, 11 (2021), p. 8438.

[18] U. Naseem, I. Razzak, M. Khushi, P. W. Eklund, and J. Kim, *Covidsenti: A large-scale benchmark twitter data set for covid-19 sentiment analysis*, IEEE transactions on computational social systems, 8 (2021), pp. 1003–1015.

[19] M. Parimala, R. Swarna Priya, M. Praveen Kumar Reddy, C. Lal Chowdhary, R. Kumar Poluru, and S. Khan, *Spatiotemporal-based sentiment analysis on tweets for risk assessment of event using deep learning approach*, Software: Practice and Experience, 51 (2021), pp. 550–570.

[20] D.-H. Pham and A.-C. Le, *Learning multiple layers of knowledge representation for aspect based sentiment analysis*, Data & Knowledge Engineering, 114 (2018), pp. 26–39.

[21] A. U. Rehman, A. K. Malik, B. Raza, and W. Ali, *A hybrid cnn-lstm model for improving accuracy of movie reviews sentiment analysis*, Multimedia Tools and Applications, 78 (2019), pp. 26597–26613.

[22] F. S. Relucio and T. D. Palaoag, *Sentiment analysis on educational posts from social media*, in Proceedings of the 9th international conference on E-education, E-business, E-management and E-learning, 2018, pp. 99–102.

[23] F. Rustam, M. Khalid, W. Aslam, V. Rupapara, A. Mehmood, and G. S. Choi, *A performance comparison of supervised machine learning models for covid-19 tweets sentiment analysis*, Plos one, 16 (2021), p. e0245909.

[24] J. Samuel, G. M. N. Ali, M. M. Rahman, E. Esawi, and Y. Samuel, *Covid-19 public sentiment insights and machine learning for tweets classification*, Information, 11 (2020), p. 314.

[25] A. Schmidt and M. Wiegand, *A survey on hate speech detection using natural language processing*, in Proceedings of the fifth international workshop on natural language processing for social media, 2017, pp. 1–10.

[26] C. Singh, T. Imam, S. Wibowo, and S. Grandhi, *A deep learning approach for sentiment analysis of covid-19 reviews*, Applied Sciences, 12 (2022), p. 3709.

[27] C. Singh, S. Wibowo, and S. Grandhi, *An integrated non-linear deep learning method for sentiment classification of online reviews*, in The International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery, Springer, 2020, pp. 889–896.

[28] M. Singh, A. K. Jakhar, and S. Pandey, *Sentiment analysis on the impact of coronavirus in social life using the bert model*, Social Network Analysis and Mining, 11 (2021), p. 33.

[29] J. Srikanth, A. Damodaram, Y. Teekaraman, R. Kuppusamy, A. R. Thelkar, et al., *Sentiment analysis on covid-19 twitter data streams using deep belief neural networks*, Computational intelligence and neuroscience, 2022 (2022).

[30] K. Vijayaprabakaran and K. Sathiyamurthy, *Towards activation function search for long short-term model network: A differential evolution based approach*, Journal of King Saud University-Computer and Information Sciences, 34 (2022), pp. 2637–2650.

[31] Y. Wang, M. Wang, W. Xu, et al., *A sentiment-enhanced hybrid recommender system for movie recommendation: a big data analytics framework*, Wireless Communications and Mobile Computing, 2018 (2018).

Edited by: Sathishkumar V E
Special Issue on: Scalability and Sustainability in Distributed Sensor Networks
Received: Mar 9, 2023
Accepted: May 2, 2023