



ORIGINAL RESEARCH

Aspect based sentiment analysis using multi-criteria decision-making and deep learning under COVID-19 pandemic in India

Rakesh Dutta¹ | Nilanjana Das² | Mukta Majumder³ | Biswapati Jana⁴¹Department of Computer Science and Application, Hijli College, Kharagpur, India²WBSEDCL, Midnapore Zone, Midnapore, West Bengal, India³Department of Computer Science and Application, University of North Bengal, Siliguri, India⁴Department of Computer Science, Vidyasagar University, Midnapore, West Bengal, India**Correspondence**

Mukta Majumder, Department of Computer Science and Application, University of North Bengal, Siliguri, India.

Email: mukta_jgec_it_4@yahoo.co.in**Abstract**

The COVID-19 pandemic has a significant impact on the global economy and health. While the pandemic continues to cause casualties in millions, many countries have gone under lockdown. During this period, people have to stay within walls and become more addicted towards social networks. They express their emotions and sympathy via these online platforms. Thus, popular social media (Twitter and Facebook) have become rich sources of information for Opinion Mining and Sentiment Analysis on COVID-19-related issues. We have used Aspect Based Sentiment Analysis to anticipate the polarity of public opinion underlying different aspects from Twitter during lockdown and step-wise unlock phases. The goal of this study is to find the feelings of Indians about the lockdown initiative taken by the Government of India to stop the spread of Coronavirus. India-specific COVID-19 tweets have been annotated, for analysing the sentiment of common public. To classify the Twitter data set a deep learning model has been proposed which has achieved accuracies of 82.35% for Lockdown and 83.33% for Unlock data set. The suggested method outperforms many of the contemporary approaches (long short-term memory, Bi-directional long short-term memory, Gated Recurrent Unit etc.). This study highlights the public sentiment on lockdown and stepwise unlocks, imposed by the Indian Government on various aspects during the Corona outburst.

KEYWORDS

aspect based sentiment analysis, bi-directional gated recurrent unit, COVID-19, deep learning, k-means clustering, multi-criteria decision-making, natural language processing

1 | INTRODUCTION

Towards the end of 2019, a new viral outbreak called COVID-19 emerged [1]. The virus has spread globally, and all the countries have joined forces to fight it. The World Health Organization has labelled it as one of the deadliest pandemic worldwide and declared guidelines and precautionary measures to control its spread while developing vaccine [2, 3]. A number of studies were carried out for evaluating the impact of the

COVID-19 on people's mental–physical health and economic down fall worldwide [4–6]. National lockdown or cordoning off in several suspected regions were implemented to control community spread in several countries, including China, Italy, Spain, and Australia, by the first week of March 2020. Taking the lead from other countries, the Indian government also declared a 21-day countrywide lockdown on 25th March 2020 [7]. India was at great risk of irreparable harm with its large population, and several guidelines were introduced to contain

Abbreviations: BERT, Bidirectional Encoder Representations from Transformers; CNN, convolutional neural network; FN, false negative; FP, false positive; GDP, gross domestic product; GPU, graphics processing unit; ICMR, Indian Council of Medical Research; LDA, latent Dirichlet allocation; LSTM, long short-term memory; NLP, natural language processing; RNN, recurrent neural network; RTX, Ray Tracing Texel eXtreme; SSD, solid state drive; TP, true positive.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. *CAAI Transactions on Intelligence Technology* published by John Wiley & Sons Ltd on behalf of The Institution of Engineering and Technology and Chongqing University of Technology.

the spread of the virus and to decrease infection rates. The countrywide lockdown came after a 14-hour 'Janata Curfew' on 22nd March from 7 AM to 9 PM [8].

This paper analyses the opinions of common Indians on the government initiatives, nationwide lockdown, and phase-wise unlocks during the Coronavirus outbreak in India. The social media platform Twitter is used for this analysis. The tweets are studied to analyse the feelings of Indians towards the situations. Though lockdown is the only option to slow down the spread of COVID-19 until the vaccination is started, its economic consequences are severe, including a drop in gross domestic product and employment. As a result, industries have shuttered, and jobs have been destroyed. The unemployment, freedom, isolation, and the lack of alternative medical services lead to an increase in dissatisfaction and as a result lot of chaos. The majority of people favour lockdown, yet owing to these concerns, many people have differing views on whether or not to impose a countrywide lockdown. The fast transmission of Coronavirus has necessitated the development of quick analytical techniques for understanding information flow and public opinion in various pandemic scenarios [9]. This opens up the possibility of analysing public opinion on this choice, which is also the purpose of this research.

Recently, scientific interest in sentiment analysis has increased in both industry and academia. Most sentiment analysis researches have so far focussed on determining a text's overall sentiment. But we often want to know what others think about certain aspects of a thing. A new concept called Aspect Based Sentiment Analysis (ABSA) was suggested in 2010 to solve this challenge [10]. Aspect Based Sentiment Analysis system takes a set of texts (like, social media posts) that describe certain facts (such as, new protocols or strategies announced by the government in COVID-19 pandemic situation) and estimate their sentiments. Even though numerous ABSA systems have been suggested, the majority of which are research prototypes [11], there is no such recognised task decomposition for ABSA.

The goal of this study is to find out the public sentiments (feelings) about the lockdown and stepwise unlocks, imposed by the Government of India on various aspects during the Corona outburst in India. The proposed ABSA technique includes three subtasks: aspect aggregation, aspect term extraction, and their polarity estimation. In the first subtask, we perform K-means clustering to aggregate the aspect. The second subtask is to find single or multi-words (phrase) that describe aspects of the subject being addressed. And the final subtask is to estimate the polarity of aspect terms for the individual tweet. For aspect aggregation, K-means clustering has been used to group similar phrases. A Best Worst Method (BWM) of Multi-criteria decision-making (MCDM) technique with different features (Term Frequency-Inverse Document Frequency [TF-IDF], Point-wise mutual information [PMI], Jaccard Coefficient and Dice Coefficient) has been used to extract the aspect term of a cluster. Then we have used Cosine Similarity and Dempster-Shafer Aggregation to measure sentiment per tweet.

Deep learning (DL) techniques are widely employed in natural language processing to classify sentiments. In this article, an advanced DL method called Bi-directional Gated Recurrent Unit (Bi-GRU) has been proposed to classify the public sentiments in the COVID-19 pandemic situation. Word2vec¹ is used to transform the user's tweets into vectors.

The salient features of this article are as follows.

1. Multi-Criteria Decision Making technique is used for extracting trending aspect terms from COVID-19-related tweets.
2. Cosine Similarity and Dempster-Shafer Aggregation are applied to measure sentiment per tweet.
3. The proposed Bi-GRU classifies the public sentiment based on the aspect terms from Twitter during the viral outbreak in India.

The rest of the paper is organised as follows: Section 2 represents the related work. Section 3 describes the proposed methodology for aspect aggregation, aspect term identification, and sentiment measuring. Section 4 describes the DL model for sentiment classification from tweets and compares the outcomes with other models. And the conclusion is drawn in Section 5.

2 | RELATED WORK

Sentiment analysis is one of the most emerging areas in natural language processing (NLP). In medical emergencies like the COVID-19 pandemic, it might play a vital role. Although many studies are currently underway on sentiment analysis using machine learning and DL techniques, some of the allied works are included in the followings.

Medford et al. analysed COVID-19 tweets for extracting keywords like illness prevention, vaccination, and racial prejudice to determine valence and dominant emotions. They used topic modelling to extract and analyse historical subjects. They found that almost half (49.5%) of the data set showed fear, and nearly 30% expressed astonishment [12]. Li et al. analysed 18,865 active Weibo users' postings using machine learning prognostic model based on online ecology identification. They measured word frequency, sentiment indicator scores (e.g. sorrow), and cognitive indicators (such as, risk assessment and lifecycle gratification) [13]. Their results showed that negative emotions (such as anxiety, despair, and anger) and sensitivity to social risks were increased, whereas positive emotions (such as Oxford happiness) and levels of happiness were decreased. Kayes et al. collected 100,000 #coronavirus tweets from Australian users. Three thousand seventy six of these tweets were related to #socialdistancing. They used 8000 tweets to train and evaluate the deep Twitter sentiment detection model. Their accuracy was 83.70%, and their F1-Score was 81.62. They concluded that Australians supported and accepted social

¹<https://code.google.com/archive/p/word2vec/>

distancing [14]. Pastor et al. performed a study to get the opinion of the students regarding the online mode of delivery of instructions because of extreme community quarantine during the COVID-19 pandemic. They performed this study by taking students' opinions in the College of Business and Public Administration of Pangasinan State University, Lingayen Campus. First, they asked all students to answer questions regarding potential online study challenges. They observed that most of the students worried about different possible issues, particularly internet connectivity in the area [15]. Chen et al. studied the controversial and non-controversial subject mentioned in COVID-19 tweets during the outbreak. They used latent Dirichlet allocation to extract topics from both types of tweet. While the non-controversial tweets emphasised the COVID-19 confrontation in the US, they also observed that the controversial tweets largely pertained to China [16]. Barkur et al. dealt with public sentiments in India after the Indian government announced the lockdown from Twitter data. They extracted the tweets using the two frequently used hashtags: #IndiaLockdown and #IndiafightsCorona from 25th March 2020 to 28th March 2020. They examined 24,000 tweets for the study using software R and generated a word cloud that evaluated the sentiments of the tweets. They found that there were sadness, fear, negativity, and disgust about the lockdown; still the positive sentiments were present prominently in the tweets. They concluded that Indians decided to decrease the spreading rate of COVID-19 [17].

Alhajji et al. used the Naïve Bayes model and Natural Language Toolkit (NLTK) in Python and found that the positive feelings were apparent on Saudi tweets in the religious practice-related steps of COVID-19 [18]. They analysed the sentiment about the seven public health measures (such as closure of Grand Mosque, Qatif lockdown, closure of school, college and university, suspension of sports competition, congregational and weekly Friday prayers, nationwide curfew, and closure of shopping malls, parks, and restaurants) imposed by the Saudi government. Out of the seven measures, public sentiment on most of these were positive, except one. Samuel et al. presented the public sentiment analysis from the Twitter data about the spread of fear related to COVID-19 infections [9]. They used two machine learning methods, such as Naive Bayes (NB) and Logistic Regression, for classifying public sentiment. Their outcomes demonstrated that the classification of short tweets by the NB and Logistic Regression methods provided an accuracy of 91% and 74%, respectively, and the classification was poor for long tweets. Pota et al. used neural networks to analyse political tweets on UK General Election. Their CNN-based technique showed classification accuracy of 63.59%. They compared their proposed model with a lexicon-based technique that used a list of English opinion words. They observed that the proposed model outperformed the lexicon-based technique for classifying phrases into positive and negative polarities [18]. Li et al. addressed method to detect and classify the public opinion about COVID-19 situation from the Sina Weibo [19]. They used supervised machine learning techniques such as Support Vector Machines (SVM), NB, and Random Forest (RF) to classify the sentiment. They

observed that the SVM, NB, and RF classifiers achieved accuracies of 54%, 45%, and 65% respectively. Afroz et al. proposed a Tweet Sentiment Analysis method which focussed on public opinion during the lockdown in India [20]. They built a real-time TextBlob sentiment analysis tool based on the NLTK. Their model was trained on the Twitter data and classified the opinions based on the subjectivity and polarity of words or phrases. They tested with seven primary keywords: Lockdown1.0, Migrant Workers, Indian Economic, Indian Council of Medical Research (ICMR), Lockdown5.0, Medical Facilities, and Policies. The result showed that Lockdown 1.0 got the most positive sentiments, followed by ICMR and Medical Facility. Chandra et al. presented a DL based framework for sentiment analysis during the rise of COVID-19 cases in India [21]. They employed long short-term memory (LSTM) and Bidirectional LSTM models with global vector (GloVe) for word representation while constructing their language model. In addition, they applied the Bidirectional Encoder Representations from Transformers model to compare the findings of LSTM and Bi-directional long short-term memory (Bi-LSTM) and then adopted the best model for sentiment analysis during the pandemic in India. Sunitha et al. proposed a sentiment analysis technique to assess coronavirus-related tweets [22]. Term Frequency-Inverse Document Frequency, GloVe, Word2Vec, and fastText embeddings were used for feature extraction. The collected feature vectors were given to the ensemble classifier (Gated Recurrent Unit [GRU] and Capsule Neural Network [CapsNet]) in order to classify the sentiments. Their proposed model was able to classify the feelings of Indian and European people with 97.28% and 95.20% accuracies respectively.

Our objective is to identify the Indian public sentiment as well to develop a DL model trained on the Twitter data set to classify public opinion and understand the circumstances during COVID-19 pandemic to take precautionary measures and make appropriate policies in this stage.

3 | PROPOSED METHOD FOR ASPECT TERM EXTRACTION AND SENTIMENT AGGREGATION

In this section, we discuss data set, the data pre-processing, and the structure of the proposed model for analysing the sentiment of the Indian people during the lockdown and unlock phases in the COVID-19 pandemic. The overall block diagram of the proposed method is shown below (Figure 1).

3.1 | Data set

The Indian government implemented the lockdown into four phases until 31st May 2020; and announced that the process of unlocking would start from then onwards. Services would be resumed in a phased manner starting from 8th June. The IEEE data port provides data set ("Tweets Originating from India During COVID-19 Lockdowns") [23] which consists of 700

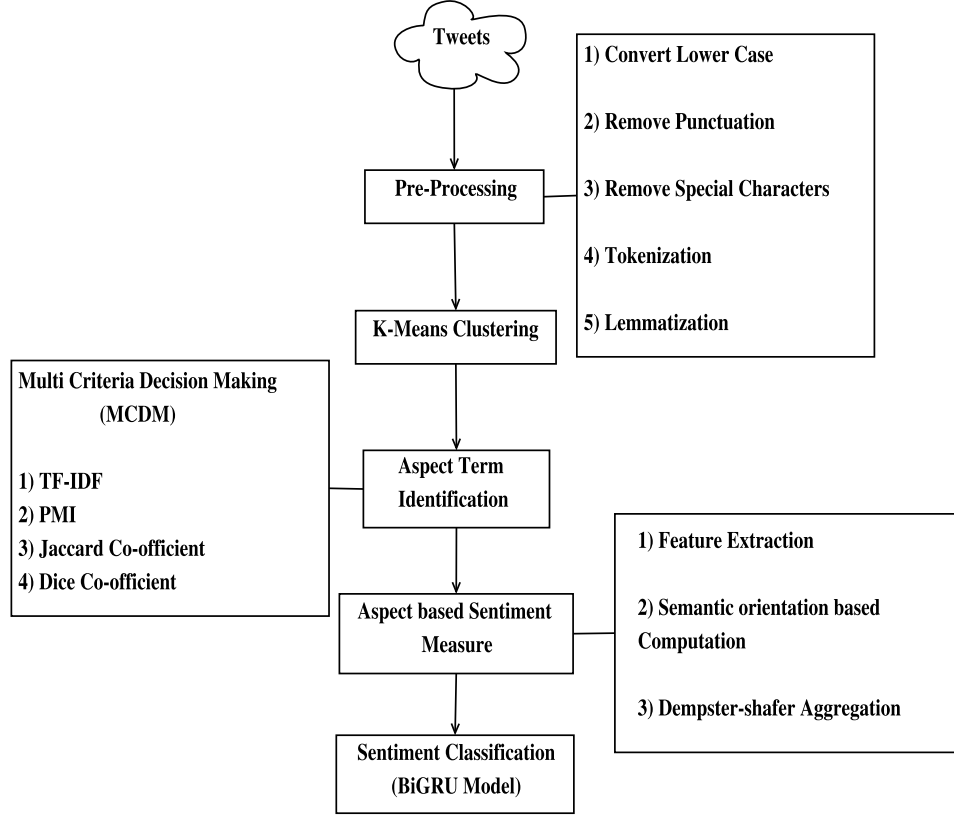


FIGURE 1 Block diagram of the proposed method.

million tweets from 25th March to 15th September, 2020, during the pandemic. We have separated the tweet data set as the Lockdown (LOC) and Unlock (UNLOC) based on the dates mentioned in the government notification² announced in India during the pandemic.

3.2 | Pre-processing

Data pre-processing comprises preparing data into an appropriate or acceptable form to NLP tools and DL model. Data collected from open-source or social media platform (such as Twitter) is often unstructured and may not be directly useable for NLP task [24–26]. The unstructured data may also include various noises. The data set must be cleaned up to enhance classification efficiency. Our pre-processing work includes converting lower case, removing punctuation, removing special characters, tokenisation, and lemmatisation.

3.3 | Clustering

In K-means clustering data points are grouped into K clusters. In the resulting clusters, the intra cluster similarity is high and the inter cluster similarity is low (Algorithm 1).

Algorithm 1: K-means Algorithm.

Require: Set of tweets $T = \{T_1, T_2, T_3, \dots, T_n\}$, K
Ensure: Set of clusters $C = \{C_1, C_2, C_3, \dots, C_K\}$

```

while  $j \leftarrow 1$  to  $K$  do
     $C_j \leftarrow \emptyset$ 
    while Each set  $T_i \in T$  do
        find  $\overline{T}_i$ ;  $\triangleright$  mean of set  $T_i$ 
    while  $j \leftarrow 1$  to  $K$  do
        randomly choose centroid  $x_j$  among  $\overline{T}_i$ 
        belongs to  $C_j$  repeat
        while each set  $T_i \in T$  do
            Assign  $T_i$  to the cluster  $C_j$  with nearest  $x_i$ 
            (i.e.,  $|\overline{T}_i - x_{j*}| \leq |\overline{T}_i - x_j|$ ;  $j \in \{1, \dots, K\}$ )
        while Each cluster  $C_j$ , where  $j \in \{1, \dots, K\}$  do
            Update the centroid  $x_i$  to be the centroid
            of all sets currently in  $C_j$ , so that  $x_j = \frac{1}{|C_j|}$ 
             $\sum_{i \in C_j} \overline{T}_i$ 
        until clusters memberships no longer
        changes
    return C
  
```

K-means clustering is a well-known and well-studied exploratory data analysis technique [27]. K is a positive integer based on which the number of clusters are defined. The main idea of K-means is to define K centroids, one for each cluster.

²<https://www.mha.gov.in/notifications/circulars-covid-19>

K-means clustering calculates the means of data points, and randomly chooses a centroid for a cluster. The technique allocates each data point to the nearest cluster based on the distance of it from the centroid of the cluster. It repeatedly reduced the intra-cluster variance to optimise the clusters' centroid.

3.3.1 | Elbow method

In our Elbow approach, the size of clusters (K) is increasing from 1 to 20. We calculate Within-Cluster Sum of Square (WCSS) for every K value. The WCSS represents the sum of the square distance of each point and the centroid in a cluster. The graph is like an Elbow when we draw the WCSS with the K value (Figure 2). At $K = 1$, the WCSS value is the highest. We have observed that the diagram changes swiftly and therefore creates an elbow when WCSS value is decreasing and K value is increasing.

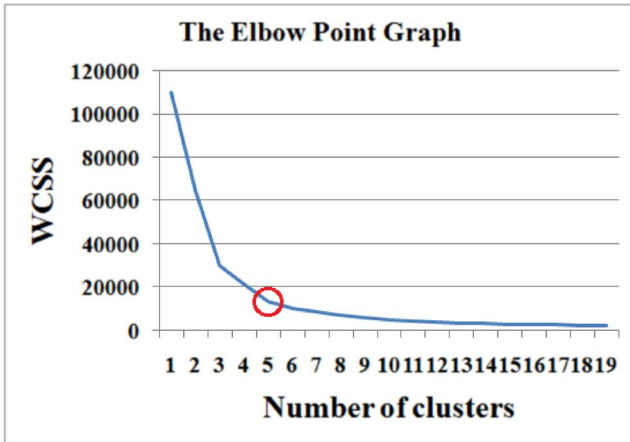


FIGURE 2 Number of clusters obtained by Elbow scheme for Lockdown data set.

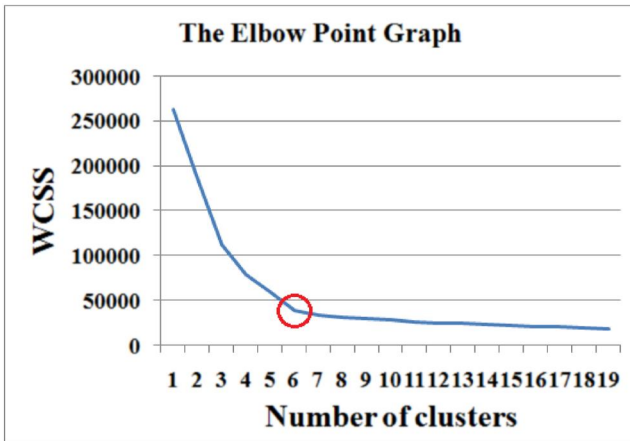


FIGURE 3 Number of clusters obtained by the Elbow scheme for Unlock data set.

Our K-means clustering works well when the K value is selected appropriately. Figure 2 shows that the optimal K value for the Lockdown data set is five, whereas the optimal K value for the Unlock data set is six as show in Figure 3. After that, the Silhouette technique is used to validate the clustering.

3.3.2 | Silhouette method

The silhouette technique studies the quality of clusters formed by K-means clustering algorithms by calculating the Silhouette score for each data point to find the mean intra-cluster and mean nearest-cluster distance. Closer data points improve the Silhouette score for a cluster. A high score indicates that, data is well matched to its own cluster and poorly matched to nearby clusters. In general, the Silhouettes score greater than 0.5 indicates that the clustering outcomes have been validated [28, 29].

According to Figure 4, the K-means clustering approach presented in this article produces effective clustering when K is set to 5 for the Lockdown data set and the Silhouette score is 0.58. Figure 5 shows that when K is set to 6 for the Unlock data set and the Silhouette score is 0.52.

3.4 | Aspect term identification

We select a word or collection of words (N -gram) that have the potential to become the correct sentiment aspect term. The aspect term of a tweet is either Unigram or Bigram or N -gram. The N -gram topic gives more information about the tweet cluster than the Unigram. In our experiments, N -gram has only three words. Several keyword extraction techniques, such as

The silhouette plot for the various clusters.

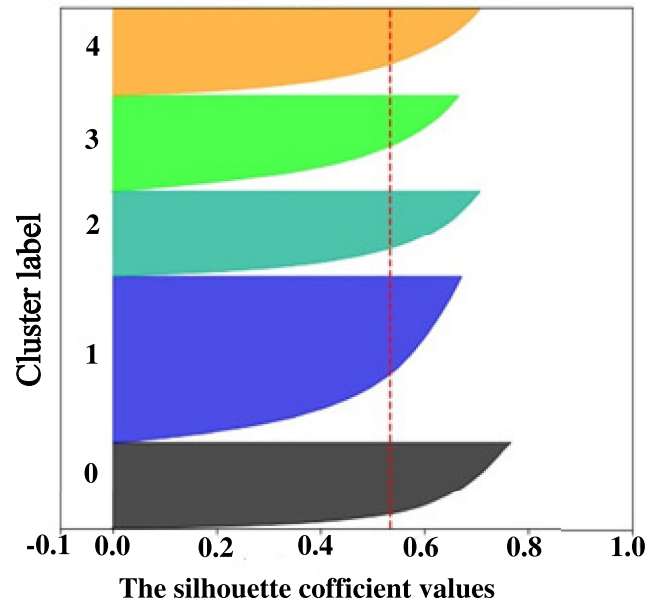


FIGURE 4 Silhouette analysis for K-means clustering on Lockdown data set.

The silhouette plot for the various clusters.

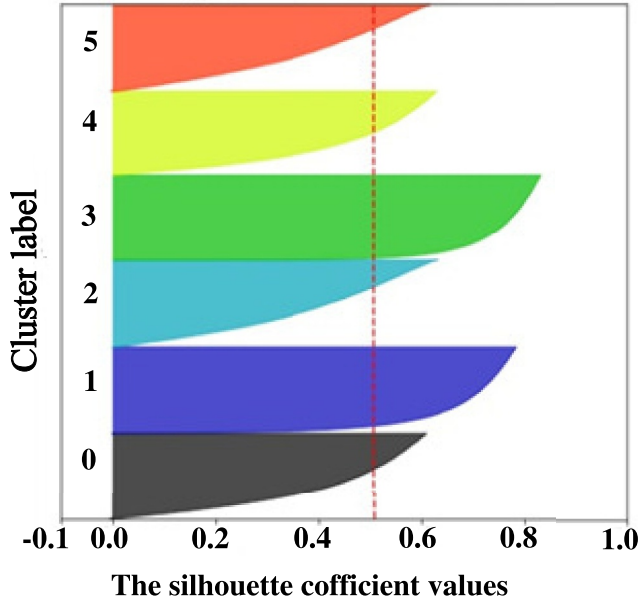


FIGURE 5 Silhouette analysis for K-means clustering on Unlock data set.

TF-IDF, PMI, Jaccard Coefficient, and Dice Coefficient are applied to identify the most relevant aspect term. Next, we have used the MCDM technique on the aforementioned keyword extraction methods to accumulate the appropriate weightage for choosing the best possible aspect term for each cluster.

3.4.1 | TF-IDF

Term Frequency-Inverse Document Frequency is a popular approach for keyword extraction. The TF-IDF is defined as

$$TF_{ij}IDF_{ij} = \frac{n_{ij}}{\sum_k n_{kj}} \times \log \frac{|T|}{|t_j \in T : w_i \in t_j|} \quad (1)$$

where n_{ij} symbolises the number of occurrences of the word w_i in texts t_j and $\sum_k n_{kj}$ represents the total number of words in texts t_j . k and T are the numbers of keywords and texts respectively. In our experiment, we have used TF-IDF of Unigram, Bigram, and Trigram.

3.4.2 | Point-wise mutual information (PMI)

The PMI association technique is used for identifying the set of multiword for obtaining topics from the texts. Higher the association score of a multiword has more potential to become a topic.

The PMI is described for Bigram as follows

$$PMI(t_i, t_j) = \log_2 \left[\frac{P(t_i, t_j)}{\{P(t_i).P(t_j)\}} \right] \quad (2)$$

where $P(t_i, t_j)$ is the joint probability of two words t_i and t_j coming sequentially in a text and $P(t_i)$, and $P(t_j)$ are the probabilities of t_i and t_j appearing individually in the text respectively. $P(t_i, t_j) = P(t_i).P(t_j)$ signifies the two words are independent of each other, and $PMI(t_i, t_j) = 0$ indicates that these two words are not good as the aspect term. Point-wise mutual information score is greater than 0.5 indicates the Bigram as an aspect topic. Similarly, PMI for Trigram (t_i, t_j , and t_k) is given by

$$PMI(t_i, t_j, t_k) = \log_2 \left[\frac{P(t_i, t_j, t_k)}{\{P(t_i).P(t_j).P(t_k)\}} \right] \quad (3)$$

3.4.3 | Jaccard coefficient

We have applied the 'Jaccard similarity coefficient' for extracting multiword aspect term [30]. This is a statistical technique well known for comparing the similarity of finite sets. The Jaccard coefficient similarity between two sets P and Q is defined by Equation (4), where $0 \leq J(P, Q) \leq 1$. If P and Q are both empty, then we defined $J(P, Q) = 1$.

$$J(P, Q) = \frac{|P \cap Q|}{|P \cup Q|} = \frac{|P \cap Q|}{|P| + |Q| - |P \cap Q|} \quad (4)$$

Here, $|P|$ is the frequency of first word (P), $|Q|$ is the frequency of second word (Q) and $|P \cap Q|$ is the joint frequency of two co-occurrence word-pair (P, Q).

3.4.4 | Dice coefficient

We have applied the Dice coefficient for extracting a set of N -gram aspect terms from the document [31]. The Dice Coefficient of Bigram is described as follows:

$$Dice(1, r) = 2 * \frac{y_{11}}{(y_{1r} + y_{r1})} \quad (5)$$

where y_{11} is the joint frequency of Bigram and y_{1r} and y_{r1} are the marginal totals of the Bigram. This can be also extended for N -grams of any size.

3.4.5 | Non-linear best worst method

Multi-criteria decision-making is considered as a complex decision-making technique where one has to identify the best possible alternative from a set of options considering a set of criteria [32]. Multi-criteria decision-making has been applied to a variety of applications [33–36]. In this article, we have used MCDM technique for finding out the best aspect term from a set of alternative. We have used BWM which is a modern technique of MCDM for our purpose [37, 38]. It is best suited when all alternatives are expressed in the same unit.

The following steps have been used to determine the optimal weight of different alternatives using BWM.

Step 1: A set of decision-making criteria $C = \{c_1, c_2, c_3, \dots, c_n\}$.

Step 2: We determine the best (B) and the worst (W) alternatives before conducting the pair-wise comparisons among the set of criteria.

Step 3: We perform pair-wise comparisons between the best alternative with the remaining alternatives (best-of-others). Here, one indicates less important and nine indicates high importance.

$$A_B = (a_{B1}, a_{B2}, a_{B3}, \dots, a_{Bn})$$

Where, a_{Bj} represents the best preference from the criterion $c_j \in C$.

Step 4: In the same way, pair-wise comparisons between the worst alternative with the other criteria (worst-of-others) have been done.

$$A_W = (a_{1W}, a_{2W}, a_{3W}, \dots, a_{nW})^T$$

Where, a_{jW} represents the worst (W) from the criterion $c_j \in C$.

Step 5: To achieve the optimal weights, the ratio of 'best-of-others' and 'worst-of-others' pair-wise maximum distance and their corresponding weight should be minimum.

$$\begin{aligned} \min_w \max_j \left\{ \left| \frac{w_B}{w_j} - a_{Bj} \right|, \left| \frac{w_j}{w_W} - a_{jW} \right| \right\} \\ \text{s.t. } \sum_{j=1}^n W_j = 1, W_j \geq 0, \forall j. \end{aligned} \quad (6)$$

Solving model (6) the optimal weights ($w_1^*, w_2^*, \dots, w_n^*$) are obtained.

Model (6) is remodelled as follows:

$$\begin{aligned} \min \xi, \\ \text{s.t. } \left| \frac{w_B}{w_j} - a_{Bj} \right| \leq \xi, \forall j, \\ \left| \frac{W_j}{W_W} - a_{jW} \right| \leq \xi, \forall j, \\ \sum_{j=1}^n W_j = 1, \\ W_j \geq 0, \forall j. \end{aligned} \quad (7)$$

The following formula is used to get the Consistency Ratio (CR):

$$\text{Consistency Ratio} = \frac{\xi^*}{CI},$$

where ξ^* is the optimal weight of model (7), and CI is the consistency index which is obtained from Table 1 [37, 38].

TABLE 1 CI table

a_{BW}	1	2	3	4	5	6	7	8	9
CI	0.00	0.44	1.00	1.63	2.30	3.00	3.73	4.47	5.23

The optimal weights are acceptable if the CR does not exceed the threshold value 0.3409 [39]; else, the pair-wise comparisons should be modified.

Model (8) and model (9) are used to determine the minimum and maximum optimum weights for each criterion [38].

$$\begin{aligned} \min W_j, \\ \text{s.t. } |W_B - a_{Bj} W_j| \leq \xi^* W_j, \forall j, \\ |W_j - a_{jW} W_W| \leq \xi^* W_W, \forall j, \\ \sum_{j=1}^n W_j = 1, \\ W_j \geq 0, \forall j. \end{aligned} \quad (8)$$

$$\begin{aligned} \max W_j, \\ \text{s.t. } |W_B - a_{Bj} W_j| \leq \xi^* W_j, \forall j, \\ |W_j - a_{jW} W_W| \leq \xi^* W_W, \forall j, \\ \sum_{j=1}^n W_j = 1, \\ W_j \geq 0, \forall j. \end{aligned} \quad (9)$$

We have chosen the optimal weight (W_j^*) from the interval $[W_j^{\min*}, W_j^{\max*}]$, represented as follows:

$$W_j^* = \frac{(W_j^{\min*} + W_j^{\max*})}{2}. \quad (10)$$

where $W_j^{\min*}$ and $W_j^{\max*}$ are calculated in Equations (8) and (9) respectively.

3.5 | Sentiment measure

The proposed sentiment polarity measuring system is depicted in Figure 6. First, we have collected those words or phrases which carry sentiment from the clusters using N -gram, part-of-speech (POS) and Dependency parsing. Next, semantic associations of these words or phrases are selected using Cosine Similarity and word embedding. Finally, the polarities of these tweets are aggregated by the Dempster-Shafer Aggregation Method [40].

3.5.1 | Feature extraction methods

Unigram Features: Unigram is treated as feature if confined to a specific POS tag, such as, JJ (adjective), RB (adverb), NN (noun), and VB (verb).

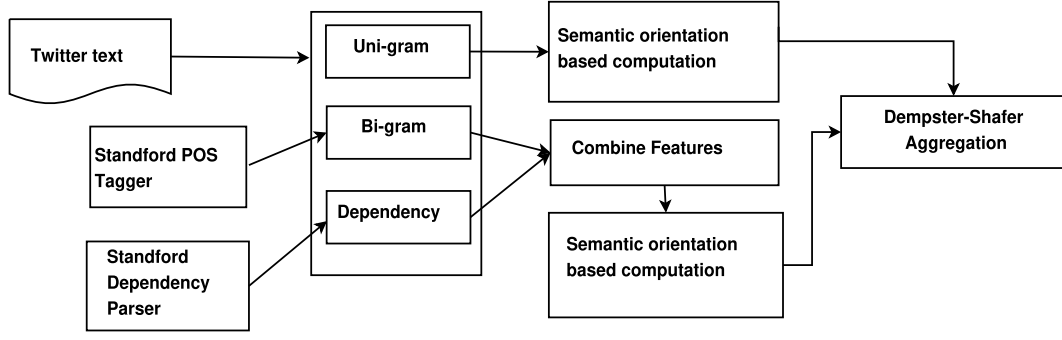


FIGURE 6 Flow diagram of semantic orientation approach for sentiment analysis.

TABLE 2 Dependency relationship features are selected according to the dependency rules for sentiment analysis

Sl. no	Relation	Meaning
1	Acomp	Adjectival complement
2	Advmod	Adverbial complement
3	Amod	Adjectival modifier
4	Dobj	Direct object
5	Neg	Negation modifier
6	Nsubj	Nominal subject
7	Rcmmod	Relative clause modifier
8	Ccomp	Clausal complement

Bigram Features: A Bigram is a pair of contiguous words in the text. Bi-tagged phrases are naturally sentiment-rich phrases. Bigram phrases may include contextual and sentiment information.

Dependency Features: Syntactic connections between words in a sentence are essential during sentiment analysis. Wiebe and Riloff showed that syntactic patterns can effectively influence sentiment analysis [41]. In our studies, we extract sentiment-related characteristics that adhere to the dependence relationships shown in Table 2. The Stanford Dependency Parser³ is used to build the dependency connection.

Combine Features: Bigram and dependency characteristics are combined to generate unique composite feature. The Bigram feature extraction approach uses POS-tagged information; however, it cannot extract all sentiment strong phrases in a sentence. Bigram features may extract contextual information but not grammatical information. Combining these two approaches allows us to extract additional sentiment information from the text in the form of Syntactic and grammatical relationships. Composite features combine Bigram and dependency characteristics to provide more information.

Semantic orientation-based computation: Word phase representation or word embedding can help to obtain semantic orientation. We have used Word2Vec for this purpose.

Word2Vec presents skip-gram model along with continuous bag-of-words (CBOW) architecture, introduced by Mikolov et al. at Google [42]. Skip-gram predicts the neighbouring words or context where a single word is specified.

The semantic polarity of a Unigram and dependency-based Bigram is determined by comparing to a positive or negative referent word ('Good', and 'Bad' respectively). The semantic polarity has been calculated by subtracting the cosine similarity between the supplied phrase and the word 'good' to the cosine similarity between the provided phrase and the word 'bad'.

$$SO(w) = \cosine(\vec{vec}_{good}, \vec{vec}_w) - \cosine(\vec{vec}_{bad}, \vec{vec}_w). \quad (11)$$

If the mean of the sentiment orientations in a tweet is less than zero, it is categorised as negative, otherwise as positive. Unlike Turney's phrase-oriented method, our strategy allows us to identify the orientation of all terms in the vocabulary [43].

Dempster-Shafer Aggregation Method (DS): Finally, the total polarity score of a tweet is computed by aggregating all of its similarity scores using Dempster-Shafer method [40]. The core of every DS-method is based on orthogonal sum of the evidence from different sources. DS is applied to two pieces of evidence, P_1 and P_2 , as follows:

$$P_{1,2}(A) = \frac{\sum_{X \cap Y = A} P_1(X)P_2(Y)}{1 - \sum_{X \cap Y = \emptyset} P_1(X)P_2(Y)}. \quad (12)$$

where A is a subset of a finite set of mutually exclusive hypotheses such as belongingness of a sentence to positive or negative classes and $P(A)$ is a mass function or basic probability assignment used to show the strength of evidence supporting A. This rule can be used for n distinct basic probability assignments, P_i for $i \in \{1, 2, \dots, n\}$ as follows:

$$P_{1,\dots,n}(A) = \frac{\sum_{\cap_{i=1}^n x_i = A} (\prod_{j=1}^n P_j(X_i))}{1 - \sum_{\cap_{i=1}^n x_i = \emptyset} (\prod_{j=1}^n P_j(X_i))}. \quad (13)$$

Since the above relation is associative, it may be used iteratively as follows:

³<https://nlp.stanford.edu/software/lex-parser.shtml>

$$P(A) = \frac{\sum_{i=1}^n x_i = AP_n(X)P_o(Y)}{1 - \sum_{i=1}^n x_i = \emptyset P_n(X)P_o(Y)} \quad (14)$$

where, P_n and P_o correspond to preceding and succeeding sentence in the tweet.

3.6 | Result and discussion on aspect term extraction and sentiment aggregation

In this article, we have applied MCDM technique to identify the aspect terms from the COVID-19 data set of lockdown and unlock phases. To calculate the optimum weights for every criterion or feature (PMI, TF-IDF, Jaccard coefficient and Dice coefficient) in the MCDM, we have used non-linear BWM. According to the non-linear BWM, the best (PMI) and the worst (TF-IDF) are selected. The PMI is chosen as best criterion because it measures the association between a pair of words or phrases and aggregates all the possible weights [44]. The TF-IDF is the worst criterion because it is only based on the word frequency and it is not useful for extracting multiword aspect terms [45]. After finding the lowest and maximum optimum weights for every criterion, we calculate the centre of the interval which is treated as the optimum weight to extract the high level information. Our CR is 0.0689, which is less than the maximum thresholds value of 0.3409 for these four criteria (see Table 3). The optimum weights for PMI, TF-IDF, Jaccard coefficient and Dice coefficient are 0.6206, 0.0689, 0.1724, and 0.1379 respectively. Using PMI, TF-IDF, Jaccard coefficient and Dice coefficient, we have selected Unigram, Bigram, and N-gram aspect terms from the cluster of Lockdown data set, such as *safety, home, government, prevent, step, supply, safety government, step prevent, number increase, government home, government crisis, government police, rule government, step protect, government India, government central, order proper, relief public, support government, lockdown supply, Italian businessman, government takes strategy, first fight corona, difficult India situation etc.* Finally, the MCDM technique has selected the Unigram, Bigram, and N-gram aspect terms which are ‘GOVERNMENT TAKES STRATEGY’, ‘PREVENT VIRUS’, ‘INDIAN ECONOMIC GROWTH’, ‘INDIAN

POLITICS’ and ‘SPREAD PANDEMIC’ from the Lockdown data set. Similarly from Unlock data set, we have extracted the aspect terms, ‘HEALTH CHECKUP’, ‘ONLINE EDUCATION INDIA’, ‘ECONOMIC PROBLEM’, ‘SPREAD VIRUS’, ‘POLITICS INDIA’ and ‘HUMAN VACCINE’.

After finding the aspect-based topics, we assess the tweet's sentiment polarity using Cosign and Dempster-Shafer Aggregation Method. We have computed the average semantic relationships of the terms in the supplied tweets and label these as positive and negative. Tables 4 and 5 show positive and negative phrases from the Lockdown and Unlock data set.

The cosine similarity method has determined the sentiment polarity of a Unigram and dependency-based Bigram (see Equation (11)). Dempster-Shafer Theory of Aggregation has been used to compute the overall sentiment polarity score of a tweet.

The aspect topics bear the most information with respect to public sentiment in the COVID-19 pandemic situation. People have tweeted their views on various issues during the lockdown phase; the proposed method have identified the following aspect terms as most trending, ‘GOVERNMENT TAKES STRATEGY’, ‘PREVENT VIRUS’, ‘INDIAN ECONOMIC GROWTH’, ‘INDIAN POLITICS’ and ‘SPREAD PANDEMIC’. Figure 7 shows the polarity of public sentiment on these aspect terms. Similarly, Figure 8 shows the polarity of public sentiment on the aspect terms (‘HEALTH CHECKUP’, ‘ONLINE EDUCATION INDIA’,

TABLE 4 Positive and negative phrases extracted from Lockdown data set

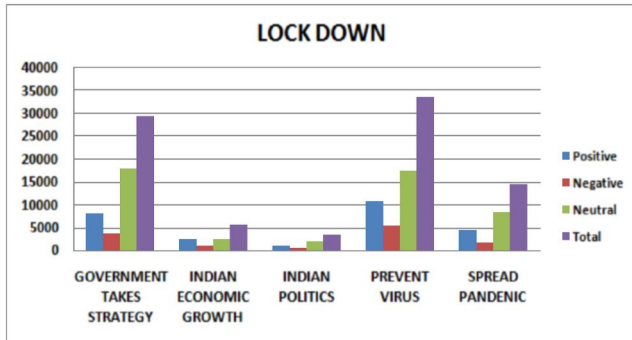
Extracted phase	Dependency relationship	Semantic orientation
Positive increase	Amod	0.249
High risk	Amod	0.056
Fresh vegetables	Amod	0.165
Lockdown prove	Nsubj	0.073
Urgently propose	Advmod	0.126
Necessary precautions	Amod	0.259
Migrant workers	Amod	0.158
Safe families	Amod	0.246
Social distance	Amod	0.155
Complete isolation	Amod	0.151
Vehicles lock	Nsubj	0.010
Die infection	Amod	−0.104
Coronavirus spread	Nsubj	−0.064
China conceal	Nsubj	−0.018
Infectious disease	Amod	−0.025
India suffer	Nsubj	−0.049
Prisoners jail	Nsubj	−0.013
Restrictions impose	Nsubj	−0.067

TABLE 3 Consistency Ratio maximum thresholds for different pairwise comparison system

Number of criteria, n							
aBW	3	4	5	6	7	8	9
3	0.2087	0.2087	0.2087	0.2087	0.2087	0.2087	0.2087
4	0.1581	0.2352	0.2738	0.2928	0.3102	0.3154	0.3273
5	0.2111	0.2848	0.3019	0.3309	0.3479	0.3611	0.3741
6	0.2164	0.2922	0.3565	0.3924	0.4061	0.4168	0.4225
7	0.2090	0.3313	0.3734	0.3931	0.4035	0.4108	0.4298
8	0.2267	0.3409	0.4029	0.4230	0.4379	0.4543	0.4599
9	0.2122	0.3653	0.4055	0.4225	0.4445	0.4587	0.4747

TABLE 5 Positive and negative phrases extracted from Unlock data set

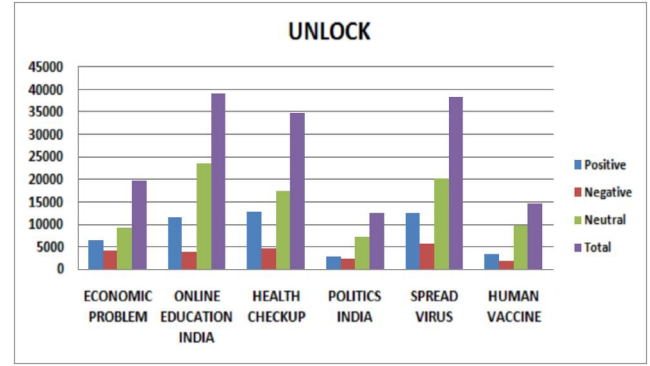
Extracted phase	Dependency relationship	Semantic orientation
Fake factories	Amod	-0.083
Fake news	Amod	-0.011
Pregnant women	Amod	0.136
Chronic patients	Amod	-0.164
Poor workers	Amod	0.064
Patients recover	Nsubj	-0.045
Well recover	Advmod	0.195
News spread	Nsubj	0.070
Positive politicians	Amod	0.065
Immediately elections	Advmod	0.086
Coronavirus spread	Nsubj	-0.064
Infectious disease	Amod	-0.025
Medical team	Amod	0.211
Nation need	Nsubj	0.309
Active case	Amod	0.242
Nifty sensenx	Amod	0.065
Sorry irresponsible	Advmod	-0.247
Inevitable vaccine	Amod	-0.060
Existent unemployment	Amod	-0.103

**FIGURE 7** Aspect wise peoples' sentiment analysis and compare our method with the existing model on the nationwide Lockdown Twitter data set.

‘ECONOMIC PROBLEM’, ‘SPREAD VIRUS’, ‘POLITICS INDIA’ and ‘HUMAN VACCINE’) during the unlock phase.

4 | PROPOSED DEEP LEARNING MODEL FOR ASPECT BASED SENTIMENT CLASSIFICATION

In sentiment analysis, we look at how people feel about objects [46]. Now there are three types of text categorisation

**FIGURE 8** Aspect wise peoples' sentiment analysis and compare our method with the existing model on the Unlock Twitter data set.

or sentiment classification methods which are rule-based, statistical, and neural network-based method [47, 48]. Conventional sentiment analysis focuses on language signals to determine a text's polarity (positive, negative, or neutral) [11, 43, 49]. The text's sentiment is frequently complex or random, making it difficult to discern from individual sentences or words. A recent study found DL [50] to be highly effective for such task [51, 52]. Deep learning includes both convolutional and recurrent neural network based (CNN and RNN) techniques [53, 54]. CNN can learn local responses but not sequential connections from spatial data. Unlike CNN, RNN can only extract features sequentially but cannot be simultaneous. Sentiment analysis is a sequential problem. RNN is widely used for this. However, simple RNN does not handle the long sequential data because it suffers from vanishing gradient problem. Long short-term memory is used as a hidden layer to combat vanishing gradient and gradient explosion [55]. The LSTM model learns long-term dependencies. Bi-directional long short-term memory combines the forward and backward hidden layers, enabling access to past and future contexts [56]. K. Cho designed GRU, a kind of RNN model [57]. Gated Recurrent Unit is simpler and more efficient than LSTM [58]. We have proposed a Bi-GRU based model, consisting of Bi-LSTM and GRU. We compare it to LSTM, Bi-LSTM, and GRU to show its superiority.

In this paper, we present a DL-based classification strategy to classify tweets by their sentiment. Five aspect topics are considered for Lockdown Twitter data set with three sentiments (Positive, Negative and Neutral). We assume that T is a set of tweets $\{t_1, t_2, t_3, \dots, t_n\}$ and t_1 consists of two of the five aspect topics: ‘prevent virus’ (PV_1) and ‘spread pandemic’ (SP_1). Suppose t_1 tweet carries positive sentiment for ‘prevent virus’, negative sentiment for ‘spread pandemic’, and null for others. We have labelled positive, negative, neutral as T, F, N , and for others as O . So, the tweet t_1 is labelled as TFOOO. This is how we have labelled each Lockdown tweet. Similarly, we have labelled the Unlock Twitter data set. This is a multi-class classification problem, which is one of the crucial tasks in NLP.

4.1 | Long short-term memory

The LSTM model has three gates and a storage unit, as shown in Figure 9. The input, forget, and output gates govern the flow of input data and storage respectively. In Figure 9, f represents tanh function and sigmoid activation function is denoted as σ . When the matrix x_t at time t is supplied, we first examine the forget gate f_t , which determines how many units of the previous instant are transferred to the present.

$$f_t = \sigma(W_f \cdot [h_{t-1} \oplus x_t] + b_f). \quad (15)$$

where \oplus symbolises the concatenation of matrix, W_f and b_f represent weighted forget matrix and the bias respectively. σ is

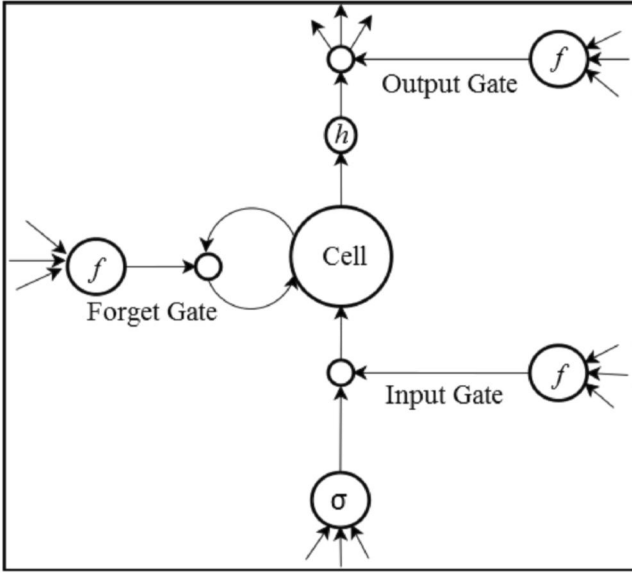


FIGURE 9 Illustration for the architecture of long short-term memory.

the activation function. The input gate is i_t , which largely in charge of sequence input at the present time.

$$i_t = \sigma(W_i \cdot [h_{t-1} \oplus x_t] + b_i). \quad (16)$$

$$\tilde{c}_t = \tanh(W_c \cdot [h_{t-1} \oplus x_t] + b_c). \quad (17)$$

$$c_t = f_t * c_{t-1} + i_t * \tilde{c}_t. \quad (18)$$

First, i_t determines the amount of data from the present input kept in the storage unit. \tilde{c}_t indicates the quantity of new information stored in the memory unit. c_t represents the current memory unit, which varies on the input and output of forget gate. h_{t-1} determines the output of the previous state.

$$o_t = \sigma(W_o \cdot [h_{t-1} \oplus x_t] + b_o). \quad (19)$$

$$h_t = o_t * \tanh(c_t). \quad (20)$$

The typical LSTM network can only use the previous context. However, the absence of the near future may result in an inadequate comprehension. Bi-directional long short-term memory can retrieve the prior and following contexts by merging a forward and backward hidden layer, as shown in Figure 10.

4.2 | Gated recurrent unit

The GRU model has two gates, as shown in Figure 11. Unlike the LSTM model, the GRU model excludes storage units. The mathematical representation of GRU is as follows:

$$r_t = \sigma(W_r \cdot [h_{t-1} \oplus x_t]). \quad (21)$$

$$z_t = \sigma(W_z \cdot [h_{t-1} \oplus x_t]). \quad (22)$$

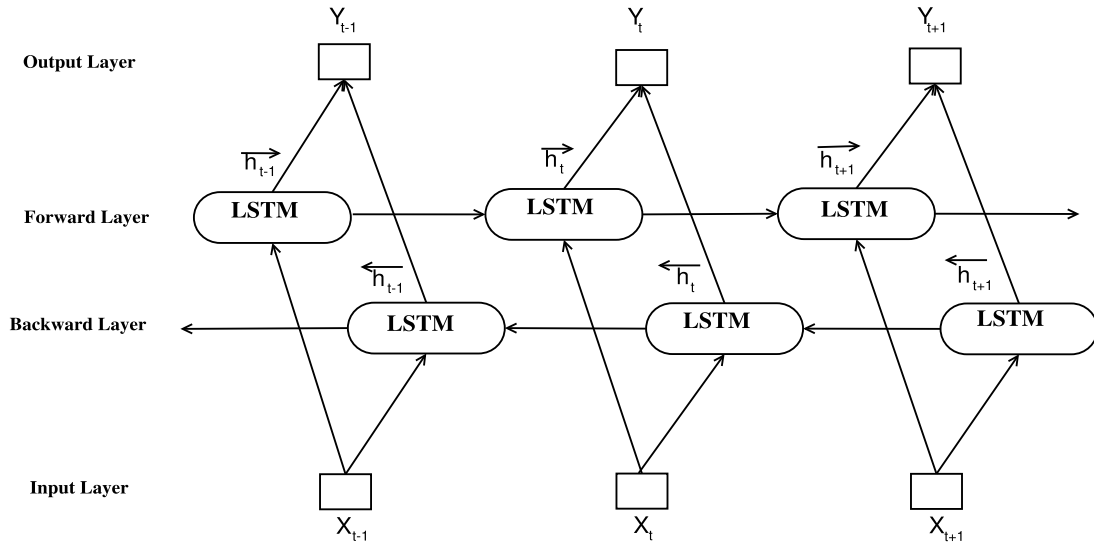


FIGURE 10 Illustration for the architecture of Bi-LSTM.

$$\tilde{h}_t = \tanh\left(W_{\tilde{h}} \cdot [r_t \otimes h_{t-1} \oplus x_t]\right). \quad (23)$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t. \quad (24)$$

where \otimes is the pairwise multiplication operator. r_t represents the reset unit which determines the update units from the previous activation h_{t-1} . z_t represents the update gate which decides the number of units updated from its activation.

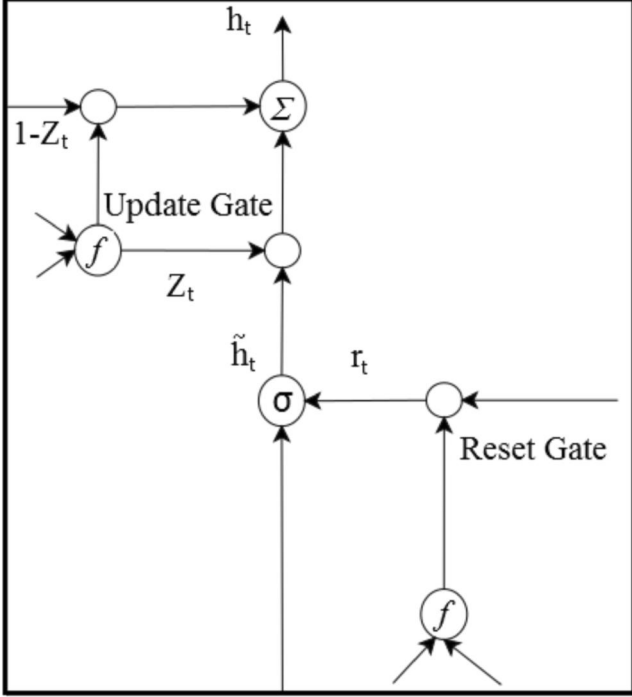


FIGURE 11 Illustration for the architecture of Gated Recurrent Unit.

We use a Bi-directional GRU unit, which is an extension of the GRU, to track records in both historical and farther information simultaneously. Here, \vec{h}_t and \overleftarrow{h}_t are two hidden states of Bi-GRU, for forward and backward calculation respectively. After that, we connect forward and backward hidden states into the final one ($h_t = [\vec{h}_t, \overleftarrow{h}_t]$), and use this state as the output of the Bi-GRU to obtain context information. The Bi-GRU is illustrated in Figure 12.

A softmax classifier has been used to predict the semantic relationship label \hat{y} from a discrete range of classes Y . Finally, the loss function is as follows:

$$j(\theta) = -\frac{1}{k} \sum_{k=1}^k y_k \log(\hat{y}_k) + \lambda \theta^2. \quad (25)$$

Here, K is the number of target classes and $y_k \in \mathbb{R}^k$ is labelled in one-hot representation, the estimated probability of every class is calculated by the activation function softmax $\hat{y}_k \in \mathbb{R}^k$ and L2 regularisation hyper-parameter is symbolised as λ . Our proposed model has used Adam Optimiser in the training period.

4.3 | Experimental setup

Hyper-parameter settings: The memory dimension of Bi-LSTM is set to 128. The batch size of Bi-LSTM and Bi-GRU is set as 50. The learning rate of our proposed model with Adam is 0.01, and the L2 penalty with coefficient is 10^{-3} . The dropout rate is 0.5 for LSTM, Bi-LSTM, and Bi-GRU models. We have randomly selected 10% of the data as the test set. Text vectors were trained using a continuous bag of words architecture with a dimensionality of 300. The simulation has been carried out in a system with

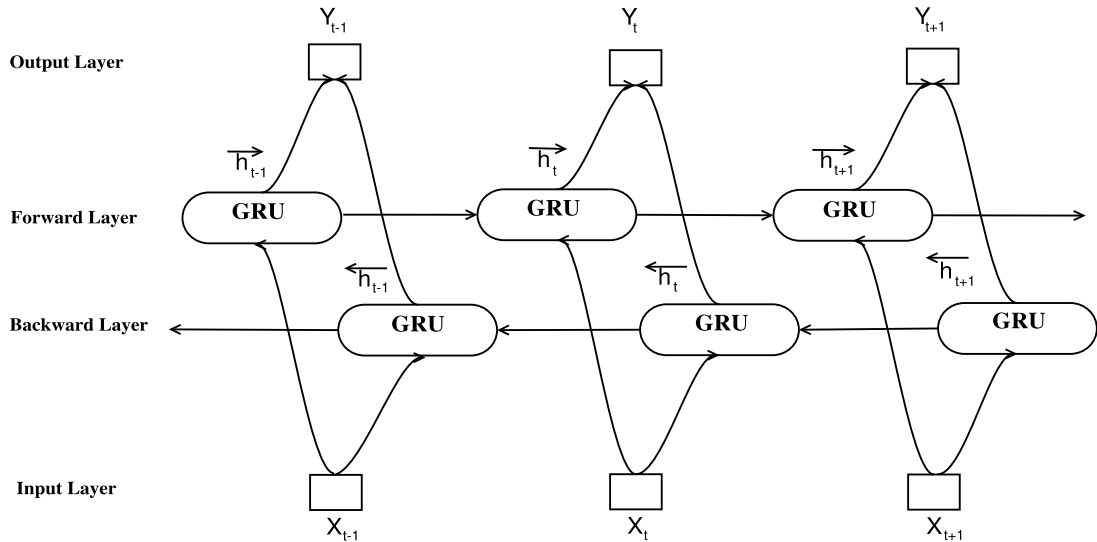


FIGURE 12 Illustration for the architecture of Bi-GRU.

configuration; Processor: Intel Core i7, solid state drive capacity: 512 GB, RAM: 16 GB, No. of CPU Cores: 8, graphics processing unit: NVIDIA GeForce RTX 3050, dedicated graphic memory capacity: 4 GB, clock speed: 4.60 GHz with max turbo frequency.

4.4 | Result and discussion on aspect based sentiment classification

This section presents our experimental results and compares our proposed model with the existing models on the same data set. We intend to examine the public opinion on trending topics during lockdown and unlock phases in COVID-19 pandemic from Twitter. Automatic aspect tagging with sentiment measuring and classification are important in this scenario. We have proposed a DL-based Bi-GRU with word2vec to classify the public sentiment. We have calculated precision, recall, f -measure, and accuracy of our proposed sentiment classification model on lockdown and unlock data set of Twitter corpus. To evaluate the model, we have used the following evaluation metrics: accuracy, recall, precision, and f -measure in the Equations (26–29) respectively. The terms: TP, TN, FP, and FN indicate true positive, true negative, false positive and false negative respectively in the Equations (26–28). Next, we have compared our proposed model with LSTM, Bi-LSTM, and GRU with the same feature vector and the comparative results are shown in Table 6.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (26)$$

$$Recall = \frac{TP}{TP + FN} \quad (27)$$

$$Precision = \frac{TP}{TP + FP} \quad (28)$$

$$F - measure = 2 * \frac{precision * recall}{precision + recall} \quad (29)$$

Gated Recurrent Unit learns faster than LSTM and generalises with fewer data [58]. Bi-directional GRU can learn both the previous and present sequential data at the same time.

From Table 6, we have observed that our proposed model Bi-GRU achieves precision, recall, f -measure, and accuracy of 0.90%, 0.83%, 0.86%, and 82.35%, respectively which are much better than other models in most of the cases like LSTM (precision: 0.70, recall: 0.78, f -measure: 0.74 and accuracy: 70.59%), Bi-LSTM (precision: 0.70, recall: 0.88, f -measure: 0.78, and accuracy: 77.78%), and GRU (precision: 0.74, recall: 0.89, f -measure: 0.81, and accuracy: 79.49%) for the Lockdown data set. Again the same Bi-GRU model obtains precision, recall, f -measure, and accuracy of 0.81%, 0.85%, 0.83%, and 83.33%, respectively which are much higher than the other contemporary approaches in most of the cases like LSTM (precision: 0.63, recall: 0.71, f -measure: 0.67, and accuracy: 72.22%), Bi-LSTM (precision: 0.70, recall: 0.78, f -measure: 0.74, and accuracy: 75%) and GRU (precision: 0.89, recall: 0.73, f -measure: 0.80, and accuracy: 76.47%) for the Unlock data set. Hence, we can conclude that for both data set Bi-GRU performs better.

We have compared the proposed Bi-GRU model with GRU, Bi-LSTM, and LSTM based on accuracy and loss function on the validation data set as shown in Figures 13–

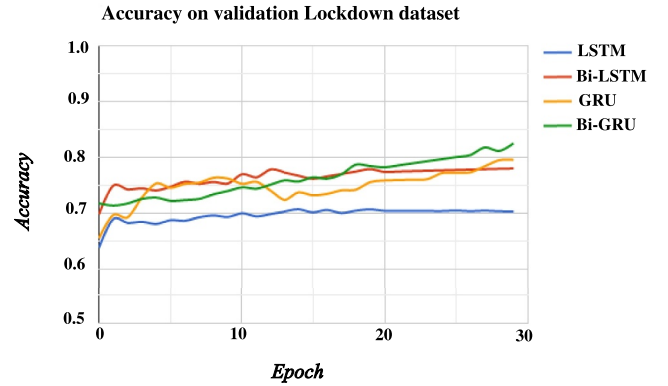


FIGURE 13 Comparison of Bi-directional gated recurrent unit (Bi-GRU), gated recurrent unit (GRU), Bi-directional long short-term memory (Bi-LSTM), and long short-term memory (LSTM) in the form of accuracy on Lockdown data set.

TABLE 6 Comparison of our proposed model with long short-term memory (LSTM), Bi-directional long short-term memory (Bi-LSTM), and gated recurrent unit (GRU)

Data set	Model name	Feature name	Precision	Recall	F -measure	Accuracy (%)
Lockdown data set	LSTM	W2V	0.70	0.78	0.74	70.59
	Bi-LSTM	W2V	0.70	0.88	0.78	77.78
	GRU	W2V	0.74	0.89	0.81	79.49
	Bi-GRU (proposed model)	W2V	0.90	0.83	0.86	82.35
Unlock data set	LSTM	W2V	0.63	0.71	0.67	72.22
	Bi-LSTM	W2V	0.70	0.78	0.74	75
	GRU	W2V	0.89	0.73	0.80	76.47
	Bi-GRU (proposed model)	W2V	0.81	0.85	0.83	83.33

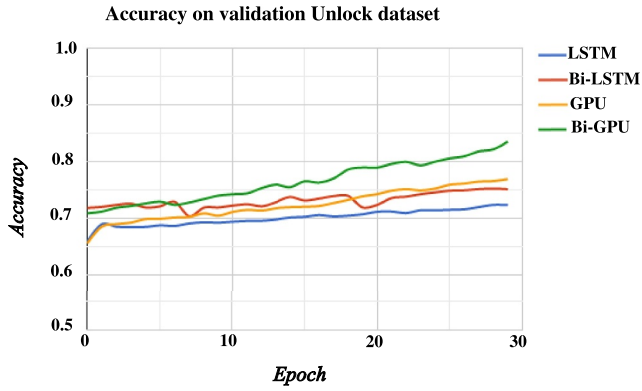


FIGURE 14 Comparison of Bi-directional gated recurrent unit (Bi-GRU), gated recurrent unit (GRU), Bi-directional long short-term memory (Bi-LSTM), and long short-term memory (LSTM) in the form of accuracy on Unlock data set.

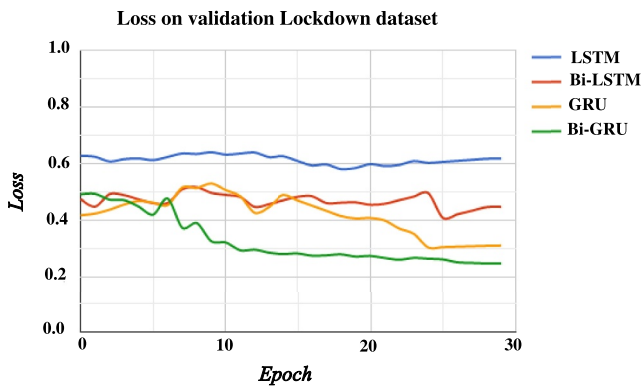


FIGURE 15 Comparison of Bi-directional gated recurrent unit (Bi-GRU), gated recurrent unit (GRU), Bi-directional long short-term memory (Bi-LSTM), and long short-term memory (LSTM) in the form of loss on Lockdown data set.

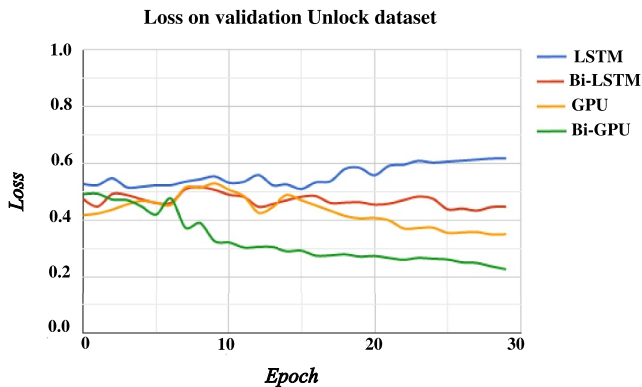


FIGURE 16 Comparison of Bi-directional gated recurrent unit (Bi-GRU), gated recurrent unit (GRU), Bi-directional long short-term memory (Bi-LSTM), and long short-term memory (LSTM) in the form of loss on Unlock data set.

16 respectively. It is evident from Figures 13 and 14 that our proposed model, Bi-GRU, shows higher accuracy than the other contemporary approaches such as LSTM, Bi-

LSTM, and GRU during most of the epochs. In an experiment, if the training loss is higher than the validation loss, then the model is overfitting. To reduce the overfitting, the network parameters and dropout can be adjusted. If the validation loss and training loss are equal, then the model is underfitting. To obtain a successful model the dropout value must be in between 0 and 1. This study considers different dropout values and maintains a fair balance between training loss and validation loss. Moreover, from Figures 15 and 16, it is clear that the value of the loss function of the Bi-GRU is lower than the LSTM, Bi-LSTM, and GRU.

5 | CONCLUSION

Social media platforms have seen a significant expansion in number of users per day. People prefer to discuss their candid ideas on social media rather than in persons. Issues during this pandemic situation and lockdown are nothing else than this. People have shared their views, opinions or feelings on this viral outbreak and lockdown-unlock strategies taken by government on social media platform, such as Twitter, Facebook etc. We have proposed an ABSA technique for identifying the public sentiment during the spread of COVID-19 from the Twitter data set. In the proposed method, tweets are clustered using the k-means approach. Aspect terms are identified using the MCDM technique based on different features (TF-IDF, PMI, Jaccard Coefficient, and Dice Coefficient) from every cluster. Then, based on the aspect terms public sentiment is extracted from each tweet. And the sentiment polarity is aggregated based on the Dempster-Shafer Aggregation Method. Next, a Bi-GRU model has been proposed for sentiment classification. The proposed model has achieved the best classification accuracies of 82.35% for Lockdown and 83.33% for Unlock data set. Our findings indicate that people support the strategies and precautionary measures taken by the Indian Government during this pandemic. Moreover, the proposed technique can help to implement practical strategies for public healthcare services, treatment and prevention of COVID-19. The comparative study highlights that the proposed model outperforms many of the contemporary sentiment classification models.

Emojis are commonly used symbols having specific meanings and often included with Twitter posts (tweets) to more explicitly express opinion. In the current study, sentiments carried by Emojis have not been taken care. This can be considered as a future direction of this research by incorporating techniques which can extract sentiment/opinion information from Emojis. Another future direction of this research could be identification of psychological/mental health of common people and hate speech by analysing public opinion/sentiment from tweets during COVID-19 pandemic.

CONFLICT OF INTEREST

The author declares that they have no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in 'IEEE Dataport' at <https://dx.doi.org/10.21227/k8gw-xz18>, reference number [23].

ORCID

Mukta Majumder  <https://orcid.org/0000-0003-2608-5762>

REFERENCES

- Wang, H., et al.: Phase-adjusted estimation of the number of coronavirus disease 2019 cases in Wuhan, China. *Cell Discov.* 6(1), 1–8 (2020). <https://doi.org/10.1038/s41421-020-0148-0>
- World Health Organization: Coronavirus (2020). <https://www.who.int/healthtopics/coronavirus.html>. Accessed 3 April 2020
- Yu, N., et al.: From sars to Covid-19: a previously unknown sars related coronavirus (sars-cov-2) of pandemic potential infecting humans call for a one health approach. *One Health.* 9, 100124 (2020). <https://doi.org/10.1016/j.onehlt.2020.100124>
- Boldog, P., et al.: Risk assessment of novel coronavirus Covid-19 outbreaks outside China. *J. Clin. Med.* 9(2), 571 (2020). <https://doi.org/10.3390/jcm9020571>
- Goyal, K., et al.: Fear of Covid 2019: First Suicidal Case in India! (2020)
- Bhat, R., et al.: Covid 2019 outbreak: the disappointment in Indian teachers. *Asian J. Psychiatr.* 50, 102047 (2020). <https://doi.org/10.1016/j.ajp.2020.102047>
- Bbc: India's 1.3bn Population Told to Stay at Home (2020). Retrieved from <https://www.bbc.com/news/world-asia-india-52024239>
- The Economic Times (2020). Retrieved from <https://economictimes.indiatimes.com/news/politics-and-nation/india-to-observe-janata-curfew-on-sunday-amid-spurt-incoronaviruscases/articleshow/74750784.cms?from=mdr>
- Samuel, J., et al.: Covid-19 public sentiment insights and machine learning for tweets classification. *Information.* 11(6), 314 (2020). <https://doi.org/10.3390/info11060314>
- Thet, T.T., Na, J.C., Khoo, C.S.: Aspect-based sentiment analysis of movie reviews on discussion boards. *J. Inf. Sci.* 36(6), 823–848 (2010). <https://doi.org/10.1177/0165551510388123>
- Liu, B.: Sentiment analysis and opinion mining. *Synth. Lect. Hum. Lang. Technol.* 5(1), 1–167 (2012). <https://doi.org/10.2200/s00416ed1v01y201204hlt016>
- Medford, R., et al.: An “infodemic”: leveraging high-volume twitter data to understand public sentiment for the Covid-19 outbreak. *medrxiv*. Preprint Posted Online April 7 (2020). <https://doi.org/10.1093/ofid/ofaa258>
- Li, S., et al.: The impact of Covid-19 epidemic declaration on psychological consequences: a study on active Weibo users. *Int. J. Environ. Res. Publ. Health.* 17(6), 2032 (2020). <https://doi.org/10.3390/ijerph17062032>
- Kayes, A., et al.: Automated Measurement of Attitudes towards Social Distancing Using Social Media: A Covid-19 Case Study (2020)
- Pastor, C.K.L.: Sentiment analysis on synchronous online delivery of instruction due to extreme community quarantine in the Philippines caused by Covid-19 pandemic. *Asian J. Multidiscip. Stud.* 3(1), 1–6 (2020)
- Chen, L., et al.: In the Eyes of the Beholder: Analyzing Social Media Use of Neutral and Controversial Terms for Covid-19 (2020). *arXiv preprint arXiv:200410225*
- Barkur, G., Vibha, G.B.K., Kamath, G.B.: Sentiment analysis of nationwide lockdown due to Covid 19 outbreak: evidence from India. *Asian J. Psychiatr.* 51, 102089 (2020). <https://doi.org/10.1016/j.ajp.2020.102089>
- Pota, M., et al.: A subword-based deep learning approach for sentiment analysis of political tweets. In: 2018 32nd International Conference on Advanced Information Networking and Applications Workshops (WAINA), pp. 651–656. IEEE (2018)
- Li, L., et al.: Characterizing the propagation of situational information in social media during Covid-19 epidemic: a case study on Weibo. *IEEE Trans. Comput. Soc. Syst.* 7(2), 556–562 (2020). <https://doi.org/10.1109/tcss.2020.2980007>
- Afroz, N., et al.: Sentiment analysis of Covid-19 nationwide lockdown effect in India. In: 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), pp. 561–567. IEEE (2021)
- Chandra, R., Krishna, A.: Covid-19 sentiment analysis via deep learning during the rise of novel cases. *PLoS One.* 16(8), e0255615 (2021). <https://doi.org/10.1371/journal.pone.0255615>
- Sunitha, D., et al.: Twitter sentiment analysis using ensemble based deep learning model towards Covid-19 in India and European countries. *Pattern Recogn. Lett.* 158, 164–170 (2022). <https://doi.org/10.1016/j.patrec.2022.04.027>
- Lamsal, R.: Tweets Originating from India during Covid-19 Lockdowns. *IEEE Dataport.* (2020). <https://dx.doi.org/10.21227/k8gw-xz18>
- Salloum, S.A., et al.: A survey of text mining in social media: Facebook and twitter perspectives. *Adv. Sci. Technol. Eng. Syst. J.* 2(1), 127–133 (2017). <https://doi.org/10.25046/aj020115>
- Xue, J., et al.: The hidden pandemic of family violence during Covid-19: unsupervised learning of tweets. *J. Med. Internet Res.* 22(11), e24361 (2020). <https://doi.org/10.2196/24361>
- Naseem, U., et al.: Covidsent: a large-scale benchmark twitter data set for Covid-19 sentiment analysis. *IEEE Trans. Comput. Soc. Syst.* 8(4), 1003–1015 (2021). <https://doi.org/10.1109/tcss.2021.3051189>
- Xu, R., Wunsch, D.: Clustering, vol. 10. John Wiley & Sons (2008)
- Oliveira, G.V., et al.: Improving k-means through distributed scalable metaheuristics. *Neurocomputing.* 246, 45–57 (2017). <https://doi.org/10.1016/j.neucom.2016.07.074>
- Rousseeuw, P.J.: Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* 20, 53–65 (1987). [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7)
- Tan, P.N., Steinbach, M., Kumar, V.: Introduction to Data Mining, Pearson Education. Inc, New Delhi (2006)
- Dice, L.R.: Measures of the amount of ecologic association between species. *Ecology.* 26(3), 297–302 (1945). <https://doi.org/10.2307/1932409>
- Mardani, A., et al.: Multiple criteria decision-making techniques and their applications—a review of the literature from 2000 to 2014. *Economic Research-Ekonomska istraživanja.* 28(1), 516–571 (2015). <https://doi.org/10.1080/1331677x.2015.1075139>
- Peng, Y., et al.: An incident information management framework based on data integration, data mining, and multi-criteria decision making. *Decis. Support Syst.* 51(2), 316–327 (2011). <https://doi.org/10.1016/j.dss.2010.11.025>
- Dursun, M., Karsak, E.E., Karadayi, M.A.: Assessment of health-care waste treatment alternatives using fuzzy multi-criteria decision making approaches. *Resour. Conserv. Recycl.* 57, 98–107 (2011). <https://doi.org/10.1016/j.resconrec.2011.09.012>
- Rezaian, S., Jozi, S.A.: Health-safety and environmental risk assessment of refineries using of multi criteria decision making method. *Apchee Procedia.* 3, 235–238 (2012). <https://doi.org/10.1016/j.apchee.2012.06.075>
- Abd AlAziz, A.M., Gheith, M., Eldin, A.S.: Lexicon based and multi-criteria decision making (MCDM) approach for detecting emotions from Arabic microblog text. In: 2015 First International Conference on Arabic Computational Linguistics (ACLing), pp. 100–105. IEEE (2015)
- Rezaei, J.: Best-worst multi-criteria decision-making method. *Omega.* 53, 49–57 (2015). <https://doi.org/10.1016/j.omega.2014.11.009>
- Rezaei, J.: Best-worst multi-criteria decision-making method: some properties and a linear model. *Omega.* 64, 126–130 (2016). <https://doi.org/10.1016/j.omega.2015.12.001>
- Rezaei, J.: A concentration ratio for nonlinear best worst method. *Int. J. Inf. Technol. Decis. Making.* 19(03), 891–907 (2020). <https://doi.org/10.1142/s0219622020500170>
- Buchanan, B.G., Shortliffe, E.H.: Rule-based Expert Systems: The Mycin Experiments of the Stanford Heuristic Programming Project (1984)
- Riloff, E., Wiebe, J.: Learning extraction patterns for subjective expressions. In: Proceedings of the 2003 Conference on Empirical Methods in Natural Language Processing, pp. 105–112 (2003)

42. Mikolov, T., et al.: Distributed representations of words and phrases and their compositionality. In: *Advances in Neural Information Processing Systems*, pp. 3111–3119 (2013)
43. Turney, P.D.: Thumbs up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews (2002). arXiv preprint cs/0212032
44. Das, B., et al.: Automatic generation of fill-in-the-blank question with corpus-based distractors for e-assessment to enhance learning. *Comput. Appl. Eng. Educ.* 27(6), 1485–1495 (2019). <https://doi.org/10.1002/cae.22163>
45. Das, B., et al.: Multiple-choice question generation with auto-generated distractors for computer-assisted educational assessment. *Multimed. Tool. Appl.* 80(21), 31907–31925 (2021). <https://doi.org/10.1007/s11042-021-11222-2>
46. Liu, B.: *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*. Cup, *Přejít k prvodnímu zdroji*, Cambridge (2015)
47. Ullah, H., et al.: Comparative study for machine learning classifier recommendation to predict political affiliation based on online reviews. *CAAI Trans. Intell. Technol.* 6(3), 251–264 (2021). <https://doi.org/10.1049/cit2.12046>
48. Fang, F., et al.: A fusion method of text categorization based on key sentence extraction and neural network. In: *2017 2nd International Conference on Knowledge Engineering and Applications (ICKEA)*, pp. 166–172. IEEE (2017)
49. Pang, B., Lee, L.: A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts (2004). arXiv preprint cs/0409058
50. Schmidhuber, J.: Deep learning in neural networks: an overview. *Neural Network.* 61, 85–117 (2015). <https://doi.org/10.1016/j.neunet.2014.09.003>
51. Tai, K.S., Socher, R., Manning, C.D.: Improved Semantic Representations from Tree-Structured Long Short-Term Memory Networks (2015). arXiv preprint arXiv:150300075
52. Basavegowda, H.S., Dagnev, G.: Deep learning approach for microarray cancer data classification. *CAAI Trans Intell Technol.* 5(1), 22–33 (2020). <https://doi.org/10.1049/trit.2019.0028>
53. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. *Commun. ACM.* 60(6), 84–90 (2017). <https://doi.org/10.1145/3065386>
54. Funahashi, K.i., Nakamura, Y.: Approximation of dynamical systems by continuous time recurrent neural networks. *Neural Network.* 6(6), 801–806 (1993). [https://doi.org/10.1016/s0893-6080\(05\)80125-x](https://doi.org/10.1016/s0893-6080(05)80125-x)
55. Hochreiter, S., Schmidhuber, J.: Long short-term memory. *Neural Comput.* 9(8), 1735–1780 (1997). <https://doi.org/10.1162/neco.1997.9.8.1735>
56. Graves, A., Schmidhuber, J.: Framewise phoneme classification with bidirectional LSTM and other neural network architectures. *Neural Network.* 18(5-6), 602–610 (2005). <https://doi.org/10.1016/j.neunet.2005.06.042>
57. Cho, K., et al.: On the Properties of Neural Machine Translation: Encoder-Decoder Approaches (2014). arXiv preprint arXiv:14091259
58. Gao, S., et al.: Short-term runoff prediction with GRU and LSTM networks without requiring time step optimization during sample generation. *J. Hydrol.* 589, 125188 (2020). <https://doi.org/10.1016/j.jhydrol.2020.125188>

How to cite this article: Dutta, R., et al.: Aspect based sentiment analysis using multi-criteria decision-making and deep learning under COVID-19 pandemic in India. *CAAI Trans. Intell. Technol.* 1–16 (2022). <https://doi.org/10.1049/cit2.12144>