Pattern Recognition Letters 158 (2022) 164–170

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

Twitter sentiment analysis using ensemble based deep learning model towards COVID-19 in India and European countries

D. Sunitha[a], Raj Kumar Patra[b], N.V. Babu[c,*], A. Suresh[d], Suresh Chand Gupta[e]

[a] Department of Computer Science & Engineering, Kamala Institute of Technology & Science, Singapur, Telangana 505468, India
[b] CMR Technical Campus, Hyderabad, India
[c] Department of Electrical and Electronics Engineering, SJB Institute of Technology, Bangalore, India
[d] Department of Computer Science and Engineering, Veltech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, India
[e] Department of Computer Science & Engineering, Panipat Institute of Engineering and Technology, Panipat, Haryana, India

A R T I C L E I N F O

Article history:
Received 13 December 2021
Revised 6 April 2022
Accepted 16 April 2022
Available online 18 April 2022

Edited by: Maria De Marsico.

A B S T R A C T

As of November 2021, more than 24.80 crore people are diagnosed with the coronavirus in that around 50.20 lakhs people lost their lives, because of this infectious disease. By understanding the people's sentiment's expressed in their social media (Facebook, Twitter, Instagram etc.) helps their governments in controlling, monitoring, and eradicating the coronavirus. Compared to other social media's, the twitter data are indispensable in the extraction of useful awareness information related to any crisis. In this article, a sentiment analysis model is proposed to analyze the real time tweets, which are related to coronavirus. Initially, around 3100 Indian and European people's tweets are collected between the time period of 23.03.2020 to 01.11.2021. Next, the data pre-processing and exploratory investigation are accomplished for better understanding of the collected data. Further, the feature extraction is performed using Term Frequency-Inverse Document Frequency (TF-IDF), GloVe, pre-trained Word2Vec, and fast text embedding's. The obtained feature vectors are fed to the ensemble classifier (Gated Recurrent Unit (GRU) and Capsule Neural Network (CapsNet)) for classifying the user's sentiment's as anger, sad, joy, and fear. The obtained experimental outcomes showed that the proposed model achieved 97.28% and 95.20% of prediction accuracy in classifying the both Indian and European people's sentiments.

1. Introduction

COVID-19 is a novel viral disease, where the first case is observed in china during December 2019, and it infected over 24.80 crore people worldwide by November 2021, causing an estimated deaths of 50.20 lakhs people [1–3]. Many strategies are followed to decrease the number of infected people such as business closures, self-quarantines, travel bans, and social distancing measures that significantly transforms the society structures all around the world [4]. The vaccination procedure started in several nations for preventing their people from serious ill with COVID-19 disease [5]. In this pandemic circumstance, the social media platforms such as Twitter, Instagram, Facebook, WhatsApp, etc. helps in gathering insightful information related to COVID-19 disease [6–8]. The content about medical services, epidemic sign and the communities affected by COVID-19 disease outbreaks [9]. Compared to other social media sites, the twitter is effective in sharing informative messages with a length of 280 characters. Active users tweets has multiple insightful information about location and travel history of the patients, cases recovered, suspected and confirmed, and the symptoms of the patients like body pains, running nose, headache, fever and cold [10]. The COVID 19 related tweets are labelled as ‘informative’ tweets, and the irrelevant user tweets are labelled as ‘uninformative’ tweets [11,12]. The objectives of this study are given as follows: (i) automatically finds European and Indian people's sentiments expressed on Twitter platform related to COVID-19, and (ii) identifies the most discussed topics by the twitter users while expressing their emotions about COVID-19 [13–15]. The major contributions of this study are determined below:

- After real-time twitter data collection, the data pre-processing is accomplished to eliminate special characters, punctuations, numbers, repeated words, non-English characters, hashtag symbols, un-necessary spaces, tabs and newlines from the tweets for better user's sentiment prediction.
- Further, the exploratory investigation: key word trend investigation and topic modeling is carried out for better understanding of the collected data. In addition, the feature extraction is

* Corresponding author.
E-mail addresses: doddasunitha2004@kitssingapuram.ac.in (D. Sunitha), rajkumarpatra.cse@cmrtc.ac.in (R.K. Patra), nvbabu@sjbit.edu.in (N.V. Babu), drasuresh@veltech.edu.in (A. Suresh), hod.cse@piet.co.in (S.C. Gupta).

https://doi.org/10.1016/j.patrec.2022.04.027
0167-8655/

performed using TF-IDF, GloVe, pre-trained Word2Vec, and fast text embedding's for extracting the discriminative feature vectors from the pre-processed data.
- The extracted discriminative feature vectors are fed to the ensemble classifier (combination of GRU and CapsNet) for classifying the Indian and European people's sentiments as fear, joy, sad and anger. The proposed ensemble classification model reduces the vanishing gradient problem by utilizing the activation function. Here, the activation function is a linear inter-polation between the candidate and the prior candidates.
- In this scenario, the effectiveness of the ensemble based deep learning model is evaluated in terms of f-score, accuracy, recall, Matthews's Correlation Coefficient (MCC), and precision. The proposed ensemble based deep learning model achieved better performance compared to the existing models like support vector machine, logistic regression [16].

This paper is prepared as follows: a few existing articles on the topic "twitter sentiment analysis related to COVID-19" are reviewed in Section 2. Technical description and experimental analysis of the proposed ensemble deep learning model are denoted in the Sections 3 and 4. The conclusion of this work is specified in Section 5.

2. Related works

Gupta et al. [17] developed a new emotion care approach to evaluate multi-modal text data that contain real time COVID-19 tweets. Initially, the collected tweets were transformed into lower case strings and then the punctuation, special characters (except A-Z, a-z), user mentions, retweets, links, and stop-words were eliminated effectively. After data cleaning, tokenization was processed to dis-integrate sentence into words. Further, a multi-modal vector was developed after identifying the frequently used words in the tweets utilizing the Term frequency-Inverse document frequency. Finally, the eight-scale emotions were classified like trust, surprise, sadness, joy, fear, anticipation, anger and disgust. Majumder et al. [16] collected twitter data across India during the time period of March 2020 to June 2020, and then the collected twitter data were converted into lower case to eliminate hyperlinks, punctuations, and abbreviation of retweets. Next, a label encoding technique was used to transform the data labels (negative, neutral, and positive tweets) into numeric form (0, 1, and 2) to ease the process of data classification. After performing lemmatization and Text Blob, the classification was carried out utilizing Support Vector Machine (SVM), and logistic regression to classify the emotions of the tweets. Imran et al. [18] analyzed the people's sentiments about COVID-19 lockdown actions taken by different countries. In this study, the deep Long Short Term Memory (LSTM) network was utilized to estimate the emotions and sentiment polarities from the people tweets. Naseem et al. [19] acquired twitter data from COVID-Senti dataset and then the Text Blob was used to label the emotional sentiments into neutral, negative, and positive. Further, the data cleaning technique was used to pre-process the collected twitter data, because the raw data are often noisy, informal, short and unstructured. In addition, the improved word vector and hybrid word ranking methods were applied to incorporate the context of tweets and sentiments for twitter sentiment analysis. Finally, the sentiment classification was accomplished using different deep and machine learning techniques like LSTM network, SVM, decision tree, Naïve Bayes, Convolutional Neural Network (CNN), and random forest.

Shamrat et al. [20] used supervised K-Nearest Neighbor (KNN) technique to classify the sentiments of people about COVID-19 vaccination. Chintalapudi et al. [21] collected Indian people's tweets from twitter websites during the time period of 23rd March 2020 to 15th July 2020 for sentiment analysis. Initially, the collected data were labelled as anger, joy, fear, and sad. Then, the Bi-directional Encoder Representation from Transformer (BERT) approach was implemented for text analysis. From the experimental examination, the presented model attained high prediction accuracy compared to other classification techniques such as LSTM, SVM, and logistic regression in twitter sentiment analysis related to COVID-19 lockdown. Basiri et al. [22] implemented a new deep learning technique for sentiment analysis by using COVID-19 tweets. In this literature, the people's sentiments were analyzed for eight countries like Canada, England, Australia, Spain, Italy, Iran, China and United States. Firstly, the tweets were collected from eight countries between the time period of 24.01.2020 and 23.04.2020 using coronavirus related key-words. A novel hybrid model was implemented for sentiment analysis, which combines five deep classifiers such as CNN, fast text model, naïve Bayes based SVM, Bi-directional Gated Recurrent Network (Bi-GRU), and BERT. In this study, the presented deep learning model performance was investigated on Stanford sentiment 140 twitter dataset that includes 1600,000 tweets.

Hanschmidt et al. [23] used German language tweets that contain '#corona' and '#cocid-19′ for sentiment analysis, where the twitter data were collected between the time period of 18.03.2020 and 24.04.2020. In this literature, the bi-term topic methods were used to analyze the people's emotions (anger, sad, anxiety, and positive) on sixteen topics: infection related concerns, social contact restrictions, and impact of the pandemic on private and public life. Villavicencio et al. [24] analyzed the Philippines people's sentiments (negative, neutral, and positive polarities) towards COVID-19 vaccines. In this literature, the Naïve Bayes classifier was used to classify the people's sentiments. The obtained experimental outcome indicates that the 8% of the tweets in the Philippines were negative against COVID-19 vaccines, 9% people were neutral, and the remaining 83% of the tweets were positive against COVID-19 vaccines. Malla and Alphonse [25] used a dataset of 226,668 tweets related to COVID-19 that was collected between December 2019 and May 2020. In this literature study, the majority voting based ensemble deep learning model was used for sentimental analysis. The presented model obtained better prediction accuracy and f-score value on the COVID-19 English labelled tweets dataset. Görmez et al. [26] developed a stacked ensemble method for sentiment analysis on the Turkish movie and SemEval-2017 datasets. By reviewing the existing works, the main concerns that exist in the developed models are: incapability in dealing with the complex sentences, which need simple sentiment words for analyzing, inability to perform well in dissimilar domains, and in-adequate performance in sentiment analysis, due to insufficient labelled data. To overcome the aforementioned issues and to improve twitter sentiment analysis, a novel ensemble based deep learning model is proposed in this research article.

3. Methodology

Generally, the sentiment analysis aims in identifying the people's attitudes and opinions from their comments in the social media platforms towards different aspects of events and products [27].

Recently, the twitter sentiment analysis has been carried out on the different topics such as political events, product reviews, movie reviews, drug reviews, classification of twitter streams during out-breaks and numerous other subjects [28,29]. In the past few decades, the twitter sentiment analysis has gained great attention among the researcher's communities, due to the advance in machine and deep learning techniques. In this research, the European and Indian people's opinions are examined towards COVID-19 during the specified time period. The experimental outcome

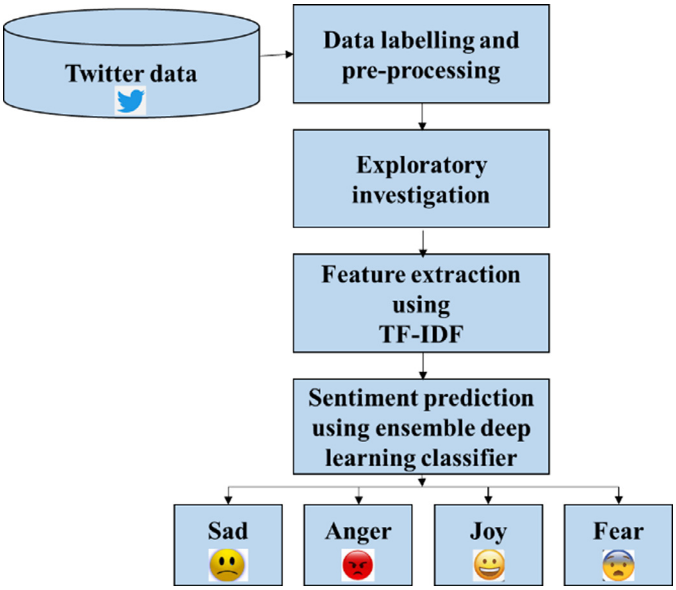

Fig. 1. Flowchart of the proposed model.

superiorly reflects the European and Indian people's opinions and sentiments towards COVID-19. In this article, the key-words used for collecting the tweets are Quarantine, social distancing, lockdown, coronavirus, corona, Covid-19, corona outbreak, pandemic, stay home, and coronavirus outbreak. The proposed twitter sentiment analysis framework contains five phases such as twitter data collection, data pre-processing, exploratory analysis, feature extraction, and sentiment prediction. The flowchart of the proposed model is graphically stated in Fig. 1.

3.1. Twitter data collection

In this research, the data of Indian and European users' tweets are collected separately from the different twitter websites during the lockdown of COVID-19. The generated dataset consists of 3100 tweets (2000 European tweets and 1100 Indian tweets), which are extracted from github.com (https://github.com/gabrielpreda/CoViD-19-tweets). Hence, the generated dataset comprises of relevant tweets on the topics of lockdown, coronavirus, covid, etc. The generated dataset consists of the extracted tweets from the Indian and European platforms that are considered for experimental analysis. As mentioned earlier, the tweets concentrated on the topics of social distance, COVID-19, corona outbreak, quarantine, lockdown, coronavirus, corona, stay home, pandemic, and coronavirus outbreak. After data collection, every tweet is annotated as 'sad', 'joy', 'fear', and 'anger'. In this research, the Text Blob tool is utilized to label the emotional sentiments into 'sad', 'joy', 'fear', and 'anger'. The Text Blob tool represents the sentence attitude by estimating the polarity score between −1 and 1. The sentiments are regarded as joy, if the polarity score is higher than 0.1. Correspondingly, the sentiments are regarded as 'fear', 'sad', and "anger", if the polarity scores are in the range of −0.1 to −0.3, −0.4 to −0.7, and −0.8 to −1. The polarity score estimation is mathematically depicted in Eq. (1).

$$L_{T_i} = \begin{Bmatrix} Joy & P > 0.1 \\ Fear & -0.1 \geq P \leq -0.3 \\ Sad & -0.4 \geq P \leq -0.7 \\ Anger & -0.8 \geq P \leq -1 \end{Bmatrix} \quad (1)$$

Where, P_i indicates polarity of tweet T_i.

Table 1
Sample pre-processed tweets.

Pre-processed tweets
India's Vaccination Drive crosses crore.
When scares end.
A gentle kind honest man has passed away. A very sweet man full of life love and joy. My condolences to his family. Rest in peace.
My head is like a prison cell.
Hackers uses fake coronavirus maps to infect visitors with malware.

Pseudocode of tweets labeling
Input: Unlabeled tweet T_U
Output: Labelled tweet T_L
Calculate:
  Sad: $T_{Sad} = []$;
  Joy:$T_{Joy} = []$;
  Fear:$T_{Fear} = []$;
  Anger:$T_{Anger} = []$;
Steps:
For t is English in $T(t)$ do:
  If (t):
  Perform labeling using Text Blob tool
    If ($-0.4 \geq Polarity\ of\ t \leq -0.7$):
    Labelled as 'sad'
    If ($-0.1 \geq Polarity\ of\ t \leq -0.3$):
    Labelled as 'fear'
    If ($-0.8 \geq Polarity\ of\ t \leq -1$):
    Labelled as 'anger'
    Else
    Labelled as 'joy'
End For
Output: Labelled tweets: $T_L = [T_{joy}, T_{sad}, T_{fear}, T_{anger}]$.

3.2. Data pre-processing

After collecting the tweets, the quality of the raw labelled data is enhanced by performing the following pre-processing operations.

- Eliminate numbers, punctuations, and special characters from the dataset, where it majorly won't improve the prediction performance.
- Eliminate repeated words. For instance: "sooooo boring" is converted as "so boring".
- Eliminate non-English characters, because this study mainly concentrated on the analysis of information in English language.
- Eliminate hashtag symbols (#china, #lockdown, #Wuhan, etc.), uniform resource locators, and @users from the tweets, because it won't contribute in analyzing the messages.
- Eliminate un-necessary newlines, tabs, and spaces from the tweets.
- The emoji's are converted into short textual description using python emoji2 library. The sample pre-processed tweets are represented in Table 1.

3.3. Exploratory investigation

After pre-processing the data, the exploratory investigation is carried-out for obtaining a more comprehensive view of the datasets. The exploratory investigation includes two steps such as key word trend investigation and topic modeling.

3.3.1. Keyword trend investigation

Firstly, the keyword trend investigation is carried-out on the pre-processed twitter data for identifying the frequently mentioned words. The European and India people are commonly talking about social distancing, staying in home, coronavirus cases, coronavirus pandemic, covid outbreak and crisis due to coronavirus.

3.3.2. Topic modeling

In this section, the topic distribution is accomplished using the Latent Dirichlet Allocation (LDA) technique for quantitatively analyzing the topics in the generated dataset [30]. The LDA technique is an effective topic model, which captures the topics from the weighted features and then each tweet is classified based on the concepts. The LDA technique creates a topic for related words or tweets as dirichlet distributions, and it describes each tweet with a probability distribution function *pr* that is mathematically stated in the Eqs. (2)–(4).

$$pr(\aleph|\pi) = \frac{\Gamma\left(\sum_{i=1}^{k}\pi_i\right)}{\prod_{i=1}^{k}\Gamma(\pi_i)}\aleph_1^{\pi_1-1}\ldots\ldots\aleph_k^{\pi_k-1} \tag{2}$$

$$pr(\aleph, x, y|\pi, \mu) = pr(\aleph|\pi)\prod_{n=1}^{N} pr(x_n|\aleph)pr(y_n|x_n, \beta) \tag{3}$$

$$pr(D|\pi, \mu) = \prod_{d=1}^{M}\int pr(\aleph_d|\pi) \times \left(\prod_{n=1}^{N_d}\sum_{x_{dn}} pr(x_{dn}|\aleph_d)pr(y_{dn}|x_{dn}, \mu)\right)d\aleph_d \tag{4}$$

Where, M specifies text review, D denotes Dirichlet distribution, Γ represents gamma function, π indicates Dirichlet parameter, μ states topics, x represents topic assignment upto k^{th} text, $\aleph$ states document level topic vectors, N denotes number of tweets, and y specifies observed text. In LDA, the number of topics is fixed as 6, and the respective topics are represented by a word distribution.

3.4. Feature extraction

After performing exploratory investigation, the feature extraction is carried out using word embedding's, and vectorization techniques. In this research, the TF-IDF technique is used for text vectorization that extracts meaningful feature vectors from the twitter data. The TF-IDF technique estimates how often a term t arises in a tweet [31], where the mathematical expressions of TF-IDF technique are given in the Eqs. (5) and (6).

$$TF = \frac{number\ of\ times\ the\ term\ (t)\ arises\ in\ a\ tweet}{total\ number\ of\ terms\ in\ a\ tweet} \tag{5}$$

$$IDF = log\frac{total\ number\ of\ tweets}{number\ of\ tweets\ with\ \ term\ (t)} \tag{6}$$

Similarly, the word embedding is accomplished using GloVe, pre-trained Word2Vec, and fast text embedding's. The word embedding techniques address two issues in twitter sentiment analysis (i) improves semantic relationship between the words by reflecting the words in the direction and distance of the vectors, and (ii) word embedding helps in obtaining dense feature vectors with low dimensionality. The GloVe does not rely on the local context information of words (local statistic), where it incorporates word co-occurrence (global statistic) to achieve better word vectors [32,33]. The Word2Vec and fast text embedding's utilizes a neural network model for learning the word's relation from a large text corpus. Once the network is trained, the synonymous words are detected and suggested extra words for a partial tweets. Hence, the Word2Vec includes two models such as continuous bag of words model and skip gram for learning word embedding [34]. Additionally, a hybrid model such as improved word vector and hybrid ranking are used to incorporate context of the tweets and sentiments for better twitter sentiment analysis [35].

3.5. Sentiment prediction

After feature extraction, the twitter sentiment analysis is carried out using an ensemble classifier, which combines CapsNet and GRU model [36]. The GRU is an updated version of recurrent neural network, where it integrates forget gate and input gate into "update gate" and includes an additional gate named "reset gate". The GRU model has few tensor operations, and parameters compared to recurrent neural network and LSTM, so it has lesser prediction time and faster convergence speed [37]. Firstly, the GRU modulates the extracted feature information inside the unit without memory cell. In the GRU model, the activation function is a linear interpolation between the candidate and the previous candidate activation function that is mathematically defined in Eq. (7).

$$h_t^j = \left(1 - z_t^j\right) h_{t-1}^j + z_t^j\tilde{h}_t^j \tag{7}$$

Where, h_t^j indicates activation function of GRU model, h_{t-1}^j states previous candidate activation function, t denotes time, and $\tilde{h}_t^j$ specifies present candidate activation, which is defined in Eq. (8). In addition, the update gate in the GRU model decides how much the unit is required to update its activation function. The update gate is mathematically determined in Eq. (9).

$$\tilde{h}_t^j = tanh\left(wA_t + U_r\left(r_t^j \times h_{t-1}\right)\right)^j \tag{8}$$

$$z_t^j = \sigma\left(w_zA_t + U_zh_{t-1}\right)^j \tag{9}$$

Where, r_t^j represents reset gates that is mathematically expressed in Eq. (10).

$$r_t^j = \sigma\left(w_rA_t + h_{t-1}\right)^j \tag{10}$$

Where, σ indicates sigmoid function, *tanh* denotes hyperbolic tangent function, and w represents parameter or weight function. In the GRU model, the Stochastic Gradient Descent (SGD) iterative method U is used to optimize the stochastic objective function on the basis of low order moments. By using gradient function $\partial L/\partial w$, the SGD iterative method updates the present weight function w and multiplied with the learning rate $\alpha = 0.025$. Further, the reset gate is updated, as mentioned in Eq. (11).

$$r_t^j = \sigma\left(w_{r+1}A_t + h_{t-1}\right)^j \tag{11}$$

Where, A_t states extracted feature information, $\partial L/\partial w_r$ denotes gradient loss function and$w_{r+1} = w_r - \alpha\partial L/\partial w_r$. The hyper-parameter settings of GRU model is denoted as follows: decay is 0.9, dropout rate is 0.6, number of epoch is 200, and batch size is 30.

Additionally, the CapsNet is an effective deep neural network that consists of several capsules (group of neurons) for predicting the user's sentiment's [38]. In the CapsNet model, every capsule is responsible to find an individual component of the objects, and all the capsules jointly finds the overall object structure [39]. The input and output of the CapsNet model are feature vectors, where the feature vectors directions encodes different characteristics (position, size, etc.) of an individual component and the output length of the feature vector u_i represents the existence probability of an individual component. The predictive vector indicates the belief $\hat{u}_{j|i}$ that encodes the relationship between the j^{th} capsule in the high level capsules and i^{th} capsule in the low level capsules utilizing a linear transformation matrix M_{ij}, which is mathematically expressed in Eq. (12).

$$\hat{u}_{j|i} = M_{ij} \times u_i \tag{12}$$

In the low level capsules, s_j is represented as the sum values of $\hat{u}_{j|i}$ with weight function c_{ij} and in the high level capsules, s_j and v_j are indicated as input and output value of the capsule i. In the CapsNet model, an iterative dynamic routing approach is used to

Table 2
Performance analysis of the ensemble based deep learning model with different feature extraction techniques on the European tweets.

Features	Classifiers	Accuracy (%)	Recall (%)	MCC (%)	Precision (%)	F-score (%)
GloVe	CapsNet	89.02	90.91	87.87	88.34	79.82
Word2Vec		89.27	89.27	86.02	87.72	82.20
Fast text embedding		90.38	88.90	83.20	90.86	74.02
Hybrid		92.82	90.67	89.09	93.09	88.88
GloVe	GRU	90.21	92.91	92.18	93.20	83.49
Word2Vec		90.28	93.07	87.02	93.44	84.65
Fast text embedding		91.20	94.04	93.92	95.90	89.90
Hybrid		92.03	95.92	93.36	95.55	91.91
GloVe	Ensemble	94.44	96.64	95.01	96.30	92.04
Word2Vec		94.80	96.22	95.98	97.02	93.30
Fast text embedding		94.92	97.30	96.11	97.90	94.43
Hybrid		95.20	97.78	97.70	98.32	96.65

find the coupling coefficient, which is mathematically expressed in Eq. (13).

$$s_j = \sum_i c_{ij} \times \hat{u}_{j|i} \tag{13}$$

If $c_{ij} = 1$, the capsule information is transmitted to the high level capsule j, and there is no information is transmitted between the capsules i and j, if $c_{ij} = 0$. In addition to this, a non-linear squash function is used in the CapsNet model that compress the longer feature vectors to 1 and the shorter feature vectors to 0. The undertaken non-linear squash function is mathematically denoted in Eq. (14).

$$v_j = \frac{s_j^2}{1 + s_j^2} \frac{s_j}{s_j^{2'}} \tag{14}$$

The c_{ij} value becomes large, if the higher and lower level capsules are consistent with the predictions and the c_{ij} value becomes small, if the higher and lower level capsules are inconsistent. The dynamic routing approach in the CapsNet ensures that the higher level capsules consistently send their prediction feature vectors to the lower level capsules for an effective sentiment analysis. The hyper-parameter setting of CapsNet model is listed as follows: number of nodes in the hidden layer is 256, batch size is 148, number of network iteration is 100, learning rate is 0.05, and routing time is 2 s.

4. Experimental results

In this application, the proposed ensemble based deep learning model performance is validated using Python environment on a computer with 128 GB random access memory, 4 TB hard disk, i9 Intel Core processor, and windows 10 operating system. In this research article, the ensemble based deep learning model performance is valuated individually for Indian and European nations in light of prediction accuracy, recall, MCC, precision, and f-score. The positive identification that belongs to the positive class is defined as “precision”, and the “recall” is defined as the ratio of total positive predictions, which are correctly determined from the all positive scenarios. Similarly, the weighted harmonic mean of recall and precision value is represented as “f-score”. The “prediction accuracy” is stated as the ratio of correctly predicted samples and the overall samples. The “MCC” is a more reliable performance measure that delivers a high score, if the prediction obtained better results in the twitter sentiment analysis. The mathematical expressions of prediction accuracy, recall, MCC, precision, and f-score are mathematically defined in the Eqs. (15)–(19).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \tag{15}$$

$$Recall = \frac{TP}{TP + FN} \times 100 \tag{16}$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \times 100 \tag{17}$$

$$Precision = \frac{TP}{TP + FP} \times 100 \tag{18}$$

$$F - score = \frac{2TP}{FP + 2TP + FN} \times 100 \tag{19}$$

Where, True Positive (TP) represents that the number of informative tweets are predicted correctly, False Positive (FP) denotes that the number of informative tweets are predicted incorrectly, True Negative (TN) indicates that the number of uninformative tweets are predicted correctly and False Negative (FN) represents that the number of uninformative tweets are predicted incorrectly.

4.1. Quantitative study on the European tweets

In this scenario, the quantitative study on the European tweets is carried out using proposed ensemble based deep learning model and the individual classifiers with different feature extraction techniques. Among 3100 collected tweets, around 2000 tweets are commented by European people in that 1600:400 tweets are utilized for proposed model training and testing. By investigating Table 2, the performance valuation is carried out with different feature extraction techniques like GloVe, Word2Vec, fast text embedding, and hybrid feature and different classification techniques such as CapsNet, GRU, and Ensemble. As seen in Table 2, the combination: ensemble classifier with hybrid features obtained superior performance in twitter sentiment analysis in light of prediction accuracy, recall, MCC, precision, and f-score. Whereas, the ensemble based deep learning model achieved 95.2% of prediction accuracy, 97.78% of recall, 97.7% of MCC, 98.32% of precision, and 96.65% of f-score, where the achieved experimental results are superior compared to other combinations.

In addition, the proposed ensemble based deep learning model performance is validated with different cross-folds like 3 folds, 5 folds, and 10 folds. By inspecting Table 3, the proposed model attained effective performance in 10 folds in terms of prediction accuracy, recall, MCC, precision, and f-score. By doing cross validation, the proposed model is protected against overfitting concern, while the amount of data is limited.

4.2. Quantitative study on the Indian tweets

As similar to the previous section, the quantitative study is performed on the Indian tweets using proposed ensemble based

Table 3
Performance of the ensemble based deep learning model with different cross-folds on the European tweets.

Ensemble based deep learning model					
Cross-folds	Accuracy (%)	Recall (%)	MCC (%)	Precision (%)	F-score (%)
3-folds	92.39	94.20	94.09	95.34	95.27
5-folds	94.34	95.37	96.96	96.02	96.11
10-folds	95.20	97.78	97.70	98.32	96.65

Table 4
Performance analysis of the ensemble based deep learning model with different feature extraction techniques on the Indian tweets.

Features	Classifiers	Accuracy (%)	Recall (%)	MCC (%)	Precision (%)	F-score (%)
GloVe	CapsNet	86.12	89.94	81.68	87.38	81.85
Word2Vec		87.29	88.28	84.07	88.78	86.27
Fast text embedding		91.88	87.95	83.40	91.87	84.42
Hybrid		93.42	92.68	85.82	92.12	86.89
GloVe	GRU	88.93	91.95	90.17	92.23	89.99
Word2Vec		89.23	92.72	90.02	92.45	89.95
Fast text embedding		91.90	94.78	92.99	93.60	90.87
Hybrid		94.45	94.82	93.96	94.15	91.97
GloVe	Ensemble	95.46	95.55	94.09	95.80	92.54
Word2Vec		95.88	95.27	94.99	95.92	93.80
Fast text embedding		95.98	96.39	95.19	96.80	95.93
Hybrid		97.28	96.98	95.90	97.77	96.29

Table 5
Performance of the ensemble based deep learning model with different cross-folds on the Indian tweets.

Ensemble based deep learning model					
Cross-folds	Accuracy (%)	Recall (%)	MCC (%)	Precision (%)	F-score (%)
3-folds	94.55	94.92	94.92	95.39	95.68
5-folds	96.78	95.97	95.48	96.78	96.10
10-folds	97.28	96.98	95.90	97.77	96.29

deep learning model with different feature extraction techniques. Among 1100 Indian tweets, 80:20% of the tweets are utilized for proposed model training and testing. As seen in Table 4, the proposed ensemble based deep learning model achieved prediction accuracy of 97.28%, recall of 96.98%, MCC of 95.90%, precision of 97.77%, and f-score of 96.29% in twitter sentiment analysis towards COVID-19 in India. The obtained experiment results are better related to individual classifiers, and feature extraction techniques. Whereas, the proposed ensemble based deep learning model almost showed 2% to 10% improvement in prediction compared to individual classifiers and feature extraction techniques.

Similarly, the ensemble based deep learning model performance with different cross-folds on the Indian tweets is depicted in Table 5. By inspecting Table 5, the ensemble based deep learning model obtained better performance, while performing 10 fold cross-validation. In addition, the standard deviation of the results are better compared to individual features and classifiers on both European and Indian datasets.

5. Conclusion

In the recent periods, the COVID-19 or coronavirus is the biggest human challenge, where all the nations' governments and researchers are trying to decrease the mortality rate of this pervasive disease. In this article, an experiment is conducted to determine the general opinion (sentiment) of the people in India and the European countries. Firstly, the Indian and European people tweets are collected between 23.03.2020 to 01.11.2021, and then data pre-processing is carried out to improve the quality of the collected data. Additionally, the exploratory investigation and feature extraction are performed for extracting the discriminative feature vectors that are finally fed to the ensemble classifier for user's sentiment prediction. The experimental investigation showed that the proposed ensemble based deep learning model obtained better performance in sentiment prediction related to individual feature extraction techniques and classifiers by means of MCC, prediction accuracy, recall, precision, and f-score. In the twitter sentiment analysis, the proposed ensemble based deep learning model achieved 97.28% and 95.20% of accuracy in classifying both Indian and European people's sentiments. However, the computational complexity of the proposed method is high, while performing the experiment with large feature length from GloVe, Word2Vec, Fast text embedding features. Therefore, as a future extension, a novel hybrid optimization technique can be included in the proposed model to select relevant features for further improving the prediction accuracy with limited computational complexity.

Funding

This research received no external funding.

Declaration of Competing Interest

The authors declare that they have no conflict of interest.

References

[1] W. Zhai, Z.R. Peng, F. Yuan, Examine the effects of neighborhood equity on disaster situational awareness: harness machine learning and geotagged Twitter data, Int. J. Disaster Risk Reduct. 48 (2020) 101611.
[2] R. Mahajan, V. Mansotra, Predicting geolocation of tweets: using combination of CNN and BiLSTM, Data Sci. Eng. 6 (4) (2021) 402–410.
[3] A. Arora, P. Chakraborty, M.P.S. Bhatia, P. Mittal, Role of emotion in excessive use of Twitter during COVID-19 imposed lockdown in India, J. Technol. Behav. Sci. 6 (2) (2021) 370–377.
[4] P. Gupta, S. Kumar, R.R. Suman, V. Kumar, Sentiment analysis of lockdown in India during COVID-19: a case study on twitter, IEEE Trans. Comput. Soc. Syst. (2020).

[5] P. Aggrawal, B.L.K. Jolly, A. Gulati, A. Sethi, P. Kumaraguru, T. Sethi, Psychometric analysis and coupling of emotions between state bulletins and twitter in India during COVID-19 infodemic, Front. Commun. 6 (2021) 695913.
[6] Z.T. Osakwe, I. Ikhapoh, B.K. Arora, O.M. Bubu, Identifying public concerns and reactions during the COVID-19 pandemic on twitter: a text-mining analysis, Public Health Nurs. 38 (2) (2021) 145–151.
[7] R. Lamsal, Design and analysis of a large-scale COVID-19 tweets dataset, Appl. Intell. 51 (5) (2021) 2790–2804.
[8] M.A. Al-Shaher, A hybrid deep learning and NLP based system to predict the spread of Covid-19 and unexpected side effects on people, Period. Eng. Nat. Sci. (PEN) 8 (4) (2020) 2232–2241.
[9] S. Sharma, S. Sharma, Analyzing the depression and suicidal tendencies of people affected by COVID-19's lockdown using sentiment analysis on social networking websites, J. Stat. Manag. Syst. 24 (1) (2021) 115–133.
[10] J.S.L. Kwan, K.H. Lim, TweetCOVID: a system for analyzing public sentiments and discussions about COVID-19 via twitter activities, in: Proceedings of the 26th International Conference on Intelligent User Interfaces, 2021, pp. 58–60.
[11] F. Rustam, M. Khalid, W. Aslam, V. Rupapara, A. Mehmood, G.S. Choi, A performance comparison of supervised machine learning models for Covid-19 tweets sentiment analysis, PLoS One 16 (2) (2021) e0245909.
[12] J. Xue, J. Chen, C. Chen, C. Zheng, S. Li, T. Zhu, Public discourse and sentiment during the COVID 19 pandemic: using latent Dirichlet allocation for topic modeling on twitter, PLoS One 15 (9) (2020) e0239441.
[13] N. Afroz, M. Boral, V. Sharma, M. Gupta, Sentiment analysis of COVID-19 nationwide lockdown effect in India, in: Proceedings of the 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS) IEEE, 2021, pp. 561–567.
[14] S. Das, A. Dutta, Characterizing public emotions and sentiments in COVID-19 environment: a case study of India, J. Hum. Behav. Soc. Environ. 31 (1–4) (2021) 154–167.
[15] J. Cabezas, D. Moctezuma, A. Fernández-Isabel, I. Martín de Diego, Detecting emotional evolution on twitter during the COVID-19 pandemic using text analysis, Int. J. Environ. Res. Public Health 18 (13) (2021) 6981.
[16] Majumder, S., Aich, A. and Das, S., 2021. Sentiment analysis of people during lockdown period of COVID-19 using SVM and logistic regression analysis. Available at SSRN 3801039.
[17] V. Gupta, N. Jain, P. Katariya, A. Kumar, S. Mohan, A. Ahmadian, M. Ferrara, An emotion care model using multimodal textual analysis on COVID-19, Chaos Solitons Fractals 144 (2021) 110708.
[18] A.S. Imran, S.M. Daudpota, Z. Kastrati, R. Batra, Cross-cultural polarity and emotion detection using sentiment analysis and deep learning on COVID-19 related tweets, IEEE Access 8 (2020) 181074–181090.
[19] U. Naseem, I. Razzak, M. Khushi, P.W. Eklund, J. Kim, Covidsenti: a large-scale benchmark Twitter data set for COVID-19 sentiment analysis, IEEE Trans. Comput. Soc. Syst. 8 (4) (2021) 1003–1015.
[20] F.J.M. Shamrat, S. Chakraborty, M.M. Imran, J.N. Muna, M.M. Billah, P. Das, M.O. Rahman, Sentiment analysis on Twitter tweets about COVID-19 vaccines using NLP and supervised KNN classification algorithm, Ind. J. Electric. Eng. Comput. Sci. 23 (1) (2021) 463–470.
[21] N. Chintalapudi, G. Battineni, F. Amenta, Sentimental analysis of COVID-19 tweets using deep learning models, Infect. Dis. Rep. 13 (2) (2021) 329–339.
[22] M.E. Basiri, S. Nemati, M. Abdar, S. Asadi, U.R. Acharrya, A novel fusion-based deep learning model for sentiment analysis of COVID-19 tweets, Knowl. Based Syst. 228 (2021) 107242.
[23] F. Hanschmidt, A. Kersting, Emotions in Covid-19 twitter discourse following the introduction of social contact restrictions in central Europe, J. Public Health (2021) 1–14.
[24] C. Villavicencio, J.J. Macrohon, X.A. Inbaraj, J.H. Jeng, J.G. Hsieh, Twitter sentiment analysis towards COVID-19 vaccines in the Philippines using naïve Bayes, Information 12 (5) (2021) 204.
[25] S. Malla, P.J.A. Alphonse, COVID-19 outbreak: an ensemble pre-trained deep learning model for detecting informative tweets, Appl. Soft Comput. 107 (2021) 107495.
[26] Y. Görmez, Y.E. Işık, M. Temiz, Z. Aydın, FBSEM: a novel feature-based stacked ensemble method for sentiment analysis, Int. J. Inf. Technol. Comput. Sci. 6 (2020) 11–22.
[27] W. Budiharto, M. Meiliana, Prediction and analysis of Indonesia presidential election from twitter using sentiment analysis, J. Big Data 5 (2018) 51.
[28] G.A. Ruz, P.A. Henríquez, A. Mascareño, Sentiment analysis of twitter data during critical events through Bayesian networks classifiers, Future Gener. Comput. Syst. 106 (2020) 92–104.
[29] N. Öztürk, S. Ayvaz, Sentiment analysis on Twitter: a text mining approach to the Syrian refugee crisis, Telemat. Inf. 35 (1) (2018) 136–147.
[30] N. Jamal, C. Xianqiao, J. Hussain Abro, D. Tukhtakhunov, Sentimental analysis based on hybrid approach of latent dirichlet allocation and machine learning for large-scale of imbalanced twitter data, in: Proceedings of the 2020 3rd International Conference on Algorithms, Computing and Artificial Intelligence, 2020, pp. 1–7.
[31] A.I. Kadhim, Term weighting for feature extraction on Twitter: a comparison between BM25 and TF-IDF, in: Proceedings of the 2019 international conference on advanced science and engineering (ICOASE), 2019, pp. 124–128.
[32] S.M. Rezaeinia, R. Rahmani, A. Ghodsi, H. Veisi, Sentiment analysis based on improved pre-trained word embeddings, Expert Syst. Appl. 117 (2019) 139–147.
[33] A.K. Sharma, S. Chaurasia, D.K. Srivastava, Sentimental short sentences classification by using CNN deep learning model with fine tuned Word2Vec, Procedia Comput. Sci. 167 (2020) 1139–1147.
[34] K.W. Church, Word2Vec, Nat. Lang. Eng. 23 (1) (2017) 155–162.
[35] U. Naseem, I. Razzak, S.K. Khan, M. Prasad, A comprehensive survey on word representation models: from classical to state-of-the-art word representation language models, ACM Trans. Asian Low Resour. Lang. Inf. Process. 20 (5) (2021) 74.
[36] Y. Wang, M. Liu, Z. Bao, S. Zhang, Short-term load forecasting with multi–source data using gated recurrent unit neural networks, Energies 11 (5) (2018) 1138.
[37] X. Cao, Y. Liu, J. Wang, C. Liu, Q. Duan, Prediction of dissolved oxygen in pond culture water based on K-means clustering and gated recurrent unit neural network, Aquacult. Eng. 91 (2020) 102122.
[38] W. Zhang, P. Tang, L. Zhao, Remote sensing image scene classification using CNN-CapsNet, Remote Sens. 11 (5) (2019) 494.
[39] H. Chao, L. Dong, Y. Liu, B. Lu, Emotion recognition from multiband EEG signals using CapsNet, Sensors 19 (9) (2019) 2212.