INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

RESEARCH ARTICLE

Deep Sentiment Classification using Topic Modeling for Covid-19

P Velvizhy[1]*, **K Arul Deepa**[2], **E Shanmuga Priya**[1]

1 Department of Computer Science and Engineering, College of Engineering Guindy, Anna University, Chennai, Tamil Nadu, India
2 Department of Information Science and Technology, College of Engineering Guindy, Chennai, Tamil Nadu, India

OPEN ACCESS

Received: 26-06-2023

Accepted: 17-08-2023

Published: 27-09-2023

Citation: Velvizhy P, Deepa KA, Priya ES (2023) Deep Sentiment Classification using Topic Modeling for Covid-19. Indian Journal of Science and Technology 16(36): 2929-2937. https://doi.org/10.17485/IJST/v16i36.1464

* **Corresponding author**.

velvizhy@gmail.com

Funding: None

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Abstract

Objectives: To extract various topics related to Covid-19 from Twitter API using LDA topic modelling technique and to analyse the sentiment of the people about the extracted topics. An interactive Q/A system with both voice and text interface is also proposed to guide COVID-19 related decision-making. And also to summarize the tweets containing a query and to suggest suitable solutions. **Method:** The proposed extracts Covid-19 related tweets from twitter API and uses Natural Language Process (NLP) method based on topic modeling to uncover various issues related to COVID-19 from public opinions. The training dataset consists of 3,38,666 COVID 19 related comments and the testing dataset consists of 1,12,888 comments. LSTM recurrent neural network is used for sentiment analysis of the extracted tweets and to produce summary for each topic identified through topic modelling. **Findings:** The accuracy comparison has been done for the existing system against the proposed model with respect various machine learning classifiers. The findings are- LSTM gives an accuracy of 79.5%, the Naïve Bayes classifier gives the accuracy of 74%, the Multinomial Naïve Bayes gives an accuracy of 74.5%, whereas the linear regression classifier achieves an accuracy of 76%, KNN classifier achieves an accuracy of 74.5% and the random forest with an accuracy of 75.5%. **Novelty**: The proposal of interactive Question Answering system is first of its kind. This work sheds light on the importance of using public opinions and suitable computational techniques to understand issues surrounding Covid 19 and to guide related decision-making.

Keywords: COVID19; LDA Topic Modeling; LSTM; Sentiment Analysis; Q/A System

1 Introduction

The COVID-19 pandemic is a major public health issue that has affected different people in many different manners [1]. This has made people lonelier and as a result they have moved towards the use of social media to cast their opinions and sentiments.

Sentiment Analysis is a predictive modeling task where the model is trained to predict the polarity of the textual data that is fed like positive and negative [2]. It is performed by various companies to understand their customer behavior towards the products well. It gives them automatic feedback of the customer that helps them to take actions accordingly. This can be very useful for businesses to label these texts. It can be done to compute feedback, reviews of the movies, etc. Emotion detection is a part of sentiment analysis where we can analyze the emotion of a person being happy, angry, sad, shock, etc. Long-Short Term Memory (LSTM) is a type of recurrent neural network but is better than traditional recurrent neural networks in terms of memory. LSTM networks are similar to RNNs with one major difference that hidden layer updates are replaced by memory cells[3]. This makes them better at finding and exposing long range dependencies in data which is imperative for sentence structures. LSTM network is fed by input data from the current time instance and output of hidden layer from the previous time instance. These two data passes through various activation functions and valves in the network before reaching the output.

In our work, we extract covid-19 related tweets from twitter API. We then use the LDA Topic model for semantic extraction and latent topic discovery of COVID- 19–related comments[4]. After obtaining topics from topic modeling using LDA, the LSTM Recurrent neural network model will be then built to analyze public sentiments about the various topics[5]. Human needs will be analyzed using various machine learning classifiers. The tweets will then be classified as positive, neutral or negative[6]. The opinions of people about various aspects like covid-19 symptoms, vaccination, and quarantine will be then summarized using the T5 summarizer[7,8]. Finally, In the proposed system, we provide an interactive Q/A system in addition to analysis of the tweets which is not present in the existing systems.

2 Methodology

The proposed system consists the modules namely, Data Extraction and Pre-processing, Semantic Processing, Sentiment Analysis, Summarization, and Interactive Q/A System as shown in Figure 1. The proposed system is added with an interactive Q/A system to provide needful information to the user queries.

Data Extraction and Pre-processing involves the acquisition of all the tweets related to Covid19 in India from Twitter API and pre-processing them using Regular Expressions and stop words. Semantic Processing deals with Topic Modelling using Latent Dirichlet allocation as described in Table 1. Sentiment Analysis involves with assigning sentiments using TextBlob, performing sentiment analysis using Long Short — Term Memory Recurrent[9]. The system uses tokenization to remove unwanted tweets. An array named stop containing the unwanted words to be removed was declared. First, we filter the tweets that contain words specified in the 'stop' array are removed. Then we again filter these tweets based on their length. The terms with length6 are only retained. The tokenized tweets are then printed. This forms the final dataset.

Table 1. Summary of Process and its Pseudocode

Process	Pseudocode
Lemmatization	Data = processed tweets Function sent_to_eords(sentences); For sentence in sentences: Convert to words and remove punctuation Data_words = list(sent_to_words(data)) Function lemmatization(texts, allowed_postaga=['NOUN', 'ADJ', 'VERB', 'ADV']); Convert words inti root words
LDA Model	Create document topic matrix For each document: For each topic: Print document.topic weight Print dominant topic For each topic: Print top-30 most relevant terms with adjust relevance metric = 1 Pick topic name Test for sample tweet For tweet in tweets: Predict topic for tweet For each topic: Print tweet count
Topic finding[10,11]	Group together tweet of each topic For each topic: Let a = summarize tweets using T5 summarizer Let b = summarize tweets by comparing sentence similarities with a as a reference and b as modal summary: Calculate rough score With b as reference and a as modal summary: Calculate rough score

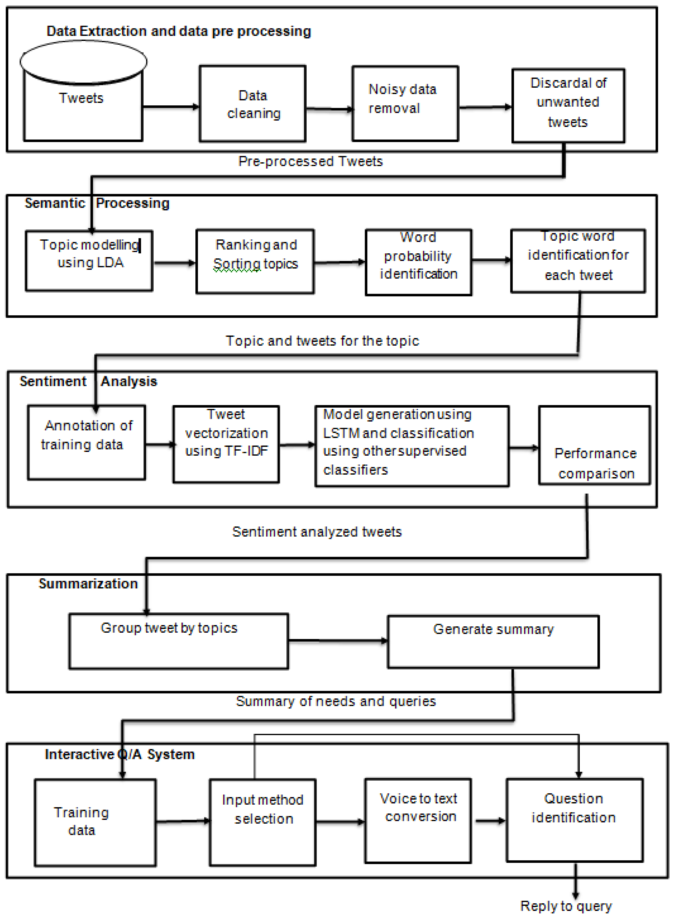

Fig 1. Proposed System Architecture

2.1 Semantic Pre-processing

The second module in this paper is Semantic Processing. After extracting tweets from the Twitter API and after pre-processing the same using regex we have used the dataset in the second module to find the topics on which people have tweeted about [12]. The topic modelling was done using Latent Dirichlet allocation. Based on the LDA model, we considered a collection of documents, such as COVID-19 related comments and words, as topics (K), where the discrete topic distributions are drawn from a symmetric Dirichlet distribution. The possible topics are listed. We have considered 10 topics. The tokens which represent each topic are printed. Each topic is labelled with the token which has the highest probability. The probability of each word under a topic is printed and the cumulative probability of a document determines the topic under which it falls. Each tweet is labelled with a topic name. Number of tweets under each topic is analysed.

2.2 Sentiment Analysis

Training data annotated using VADER. Text Blob can also be used for annotating the dataset. The advantage of using VADER using Text Blob is that VADER considers the compound polarity of the sentence whereas Text Blob considers a particular word to decider the polarity. LSTM Recurrent neural network model is built to analyse public sentiments about the various topics obtained from topic modeling using LDA. The formula for each LSTM cell can be formalized as in Equations (1), (2) and (3).

$$ft = (Wfxt + Ufht1 + bf) \tag{1}$$

$$it = (Wfxt + Uiht1 + bi) \tag{2}$$

$$ot = (Woxt + Uoht1 + bo) \tag{3}$$

The tweets are classified as positive, neutral or negative. Machine Learning classifiers like Multinomial Naive Bayes, Bernoulli Naive Bayes, K-Nearest Neighbours, Random Forest were used to identify human needs and human satisfaction on aspects like healthcare system, social life, government and education. The performance of various classifiers is compared with LSTM model[13].

2.3 Summarization

People's opinions about various topics obtained from LDA topic modeling are summarized using abstractive summarizers[14]. Two summarizers used are T5 summarizer and Pipeline. User's queries, comments and opinions regarding the obtained COVID-19 related topics namely 'impact on people', 'lock- down', 'vaccine', 'economy', 'health', 'cases' and 'work' are summarized. As there was no specific reference document, the performance of both the summarizers are compared with each other[15,16].

2.4 Interactive Q/A System

The user can ask questions in either text or voice format. The given query will be compared with all the question in the dataset and the one with the highest similarity score will be considered as the predicted question.[17] Answers to queries about COVID-19 will be provided in text or voice format. If the question was irrelevant to COVID then the system tells the user that it was not able to comprehend the question.

3 Result and Discussion

3.1 Data Extraction and Pre-processing

Developer account was created in twitter. Four keys were obtained for authentication purposes. Tweets related to covid19 are extracted from Twitter API. The dataset is created by extracting tweets from Twitter using python's twitter API 'tweepy'[18]. The feature retained in the dataset is only the Tweet Text. The acquired tweets contain a large amount of information about covid19 virus, its impact, the symptoms, treatment options, vaccination details such as dosage gap, side effects. Thus, it is ideal for implementing approaches to extract useful and important data from the large set of tweets.

A) Data Extraction: The tweets are extracted from twitter using tweepy.

Step 1: Collection of tweets from Twitter API

Step 2: Extraction of tweets using search words, using Cursor API

Step 3: Storage to tweets to .csv file

B) Data Pre-processing:

Step 1: Removal of unnecessary characters, URLs, retweets.

Step 2: Removal of tweets containing the stop words.

Step 3: Storing the cleaned tweets.

Output: Pre-processed Tweets

All unwanted characters (hashtags, emoticons, URLs, non-ascii sequences) are eliminated and all tweets containing the set stop words are removed. Resulting cleaned tweets are stored in the same csv file.

The extracted tweets after cleaning are pre-processed to remove the hashtags, mentions, special characters etc. This step is crucial because only useful information is necessary to perform sentiment analysis.

3.2 Semantic Pre-processing

The first step is to perform lemmatization. The reason for this is to convert a word to its root form so as to reduce the total number of unique words in the dictionary corpus. The next step is to vectorize the data to be fed to the LDA model. Vectorization was done using Count Vectorizer. It considered only English words with length at least 3. Next process is to train the LDA model. The model was tested with learning rates 0.5, 0.7 and 0.9 for different number of topics such as 10, 15, 20, 25 and 30. It was estimated that the best model was with learning rate 0.7 and 10 topics. The document probability matrix was found. Each topic had some meaningful terms and the most accurate one was chosen as the topic label[19]. Each tweet was assigned a topic based on the cumulative probability score of each token. Number of tweets under each topic was listed.

A) Lemmatization:
Step 1: Tokenize each tweet Step
Step 2: Remove all English stop words
Step 3: Convert word to its root form
B) Topic Modelling with LDA:
Step 1: Vectorize data
Step 2: Train LDA model
Step 3: Choose best learning rate and number of topics
C) Ranking and Choosing Topics:
Step 1: Extract document frequency matrix
Step 2: Choose most relevant term for each topic
Step 3: Annotate each tweet with the highest probability topic

10 topics were extracted in total and they were labelled. Each tweet was given the topic with the highest cumulative probability [20,21].

The lemmatization process is carried out to identify the root form of the word. Finding the root form of the word eliminates duplicates in the word dictionary. Different training rates and number of topics were taken into consideration. Multiple terms were listed under each topic and the most relevant term is considered as the topic label. Seven such topics were found and each tweet is annotated with a topic. Number topics and their values counts are presented in Table 2. For each topic class wise count is presented in Table 3.

Table 2. The Topics and Tweets under each Topic

Topics	Value Counts
Impact on people	325
Lockdown	318
Vaccine	314
Economy	291
Health	274
Cases	225
Work	164

Table 3. Topic and number of Positive and Negative Tweets under each Topic

Topics	Sentiment Class	Count
Cases	Negative	165
	Positive	37
Covid warriors	Negative	224
	Positive	84
Economy	Negative	220
	Positive	58
Impact on people	Negative	231
	Positive	73
Lockdown	Negative	214
	Positive	87
Vaccine	Negative	164
	Positive	82
Work	Negative	135
	Positive	28

3.3 Sentiment Analysis

The first step is to perform lemmatization. The reason for this is to convert a word to its root form so as to reduce the total number of unique words in the dictionary corpus. The next step is to vectorize the data to be fed to the LDA model. Vectorization was done using Count Vectorizer. It considered only English words with length at least 3. Next process is to train the LDA model. The model was tested with learning rates 0.5, 0.7 and 0.9 for different number of topics such as 10,15,20,25,30. It was estimated that the best model was with learning rate 0.7 and 10 topics [22,23]. The document probability matrix was found. Each topic had some meaningful terms and the most accurate one was chosen as the topic label. Each tweet was assigned a topic based on the cumulative probability score of each token. Number of tweets under each topic is listed. Figure 2 presents the working of LSTM.

```
# Hyperparameters of the model
vocab_size = 200 # choose based on statistics
oov_tok = ''
embedding_dim = 90
max_length = 165 # choose based on statistics, for example 150 to 200
padding_type='post'
trunc_type='post'
# tokenize sentences
tokenizer = Tokenizer(num_words = vocab_size, oov_token=oov_tok)
tokenizer.fit_on_texts(train_sentences)
word_index = tokenizer.word_index
# convert train dataset to sequence and pad sequences
train_sequences = tokenizer.texts_to_sequences(train_sentences)
train_padded = pad_sequences(train_sequences, padding='post', maxlen=max_length)
# convert Test dataset to sequence and pad sequences
test_sequences = tokenizer.texts_to_sequences(test_sentences)
test_padded = pad_sequences(test_sequences, padding='post', maxlen=max_length)
```

Fig 2. Performing Sentiment Analysis using LSTM

As a first step the sentiment is assigned to each tweet in the dataset. Then, positive tweets are encoded as 1 and negative tweets are encoded as 0. The LSTM model is then used to perform sentiment analysis on dataset. The Accuracy of LSTM model is obtained as 78 percent. The accuracy of the existing system is compared with the proposed system as shown in Table 4. The comparison of LSTM model with other ML classifiers are depicted in the form of a Bar chart as given in Figure 3. Figure 4 presents the graph depicting different learning rates. And the inter-topic distance matrix is presented in Figure 5.

Table 4. Accuracy comparison

Classifiers	Accuracy	
	Existing system [1]	Proposed system
Naïve Bayes	72.38	74
Logistic Regression	78.72	79
KNN	56.18	75

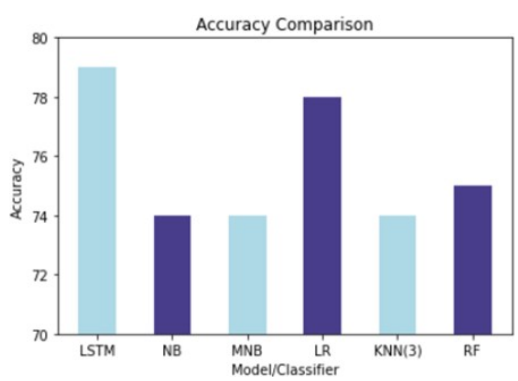

Fig 3. Comparison

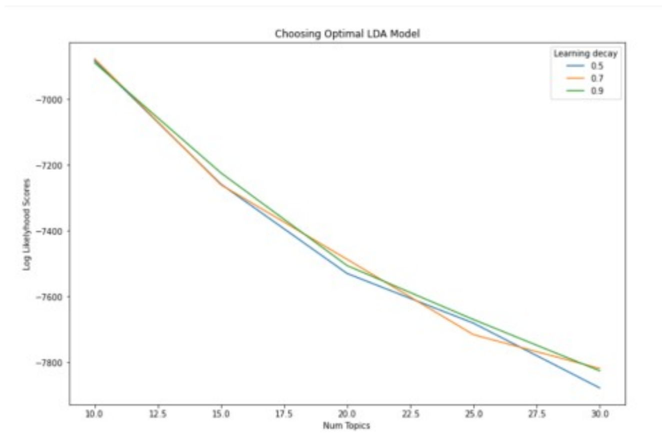

Fig 4. Graph depicting the different learning rates

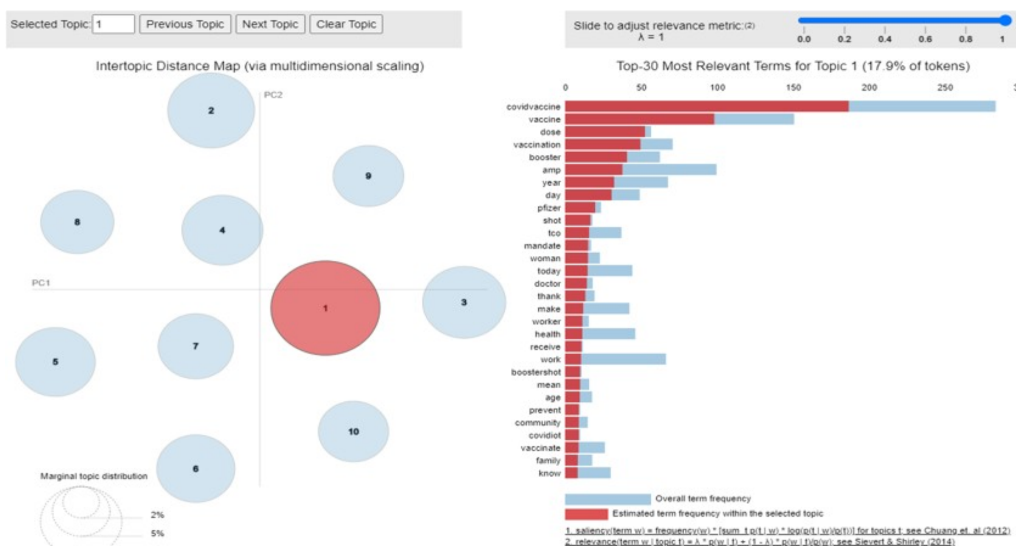

Fig 5. Inter topic Distance Matrix

3.4 Summarization

People's opinion about various aspects obtained from LDA topic modeling are summarised using T5 summarizer and Pipelining. User's queries, comments and opinions regardingcovid-19likesymptoms, quarantining, and vaccines are also summarised and a solution is provided for the same using an interactive chatbot.

Output: Summary (textformat)

The output for both voice and text interface is provided. The rouge score is calculated between the summaries obtained using T5 summarizer and pipelining.

3.5 Interactive Q/A System

The user can ask questions in either text or voice format. The given query will be compared with all the question in the dataset and the one with the highest similarity score will be considered as the predicted question. Answers to queries about covid19 will be provided in text or voice format. If the question was irrelevant to covid-19 then the system tells the user that it was not able to comprehend the question.

Output: Reply in text(voice) form

For Interactive Q/A system, the 'COUGH' dataset containing FAQ questions about COVID is used. 960 questions and answers related to the topics found from LDA topic modeling are present which is presented in Figure 6.

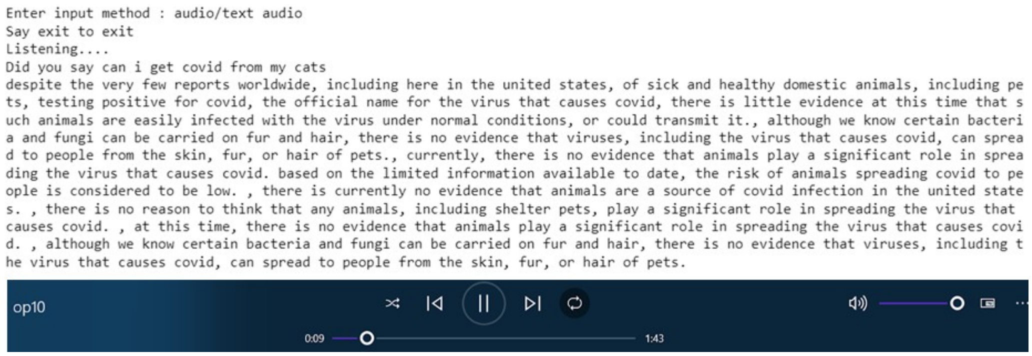

Fig 6. Sample Input and output for interactive Q/A system using Voice Interface

3.6 Comparison with Existing System

In the proposed system 10 topics have been extracted from topic modelling using LDA whereas the existing work used only 4 [4] and 6 [2] topics. In this work the dataset size has been increased by extracting tweets using keywords when compared to existing system [3], which was not sufficient there by limits the analysis. The proposed work considers emoticons while annotating the input data. But the existing work [5] has not considered the sentiments in emoticons and hashtags of the tweets. Hence, the classifier efficiency is hampered. The existing works considered only positive and negative class labels where the proposed work considered positive, negative and neutral classes [6]. The summary of comparison is given in Table 5.

Table 5. Existing vs proposed work

Comparison criteria	Existing work [citation]	Proposed work
Number of topics covered	6 [2] , 4 [4]	10
Dataset size	1300 tweets [3]	2500 tweets
Emoticons & hashtags	Not considered [5]	considered
Class labels	Positive/negative [6]	Positive/negative/neutral

4 Conclusion

Tweets related to COVID-19 were generated from Twitter API. The tweets were pre-processed using regular expressions. LDA topic modelling was done and the tweets were classified into 7 topics. LSTM, RNN was used to analyse the sentiment of tweets under each topic. The tweets under each topic are then summarized using T5 Summarizer and Pipelining. Queries related to the found topics can be cleared using an interactive system which can accept input in voice or text format. The output can also be presented in text or voice format. The performance of LSTM was compared with supervised machine learning classifiers like NaiveBayes, MNB, Linear Regression, KNN and random forest and achieved accuracy of 74%, 74.5%, 76%, 74.5% and 75.5% respectively. The proposed system achieved an accuracy of 79.5%.

Topic modeling and sentiment analysis using LSTM was done only for tweets related to Covid-19. It can be done for any other dataset also. All the outputs are presented in Jupyter Notebook. The proposed models are integrated with a novel user-friendly interface in the form of a web application to present the results of LDA Topic Modeling, Sentiment Analysis and Summarization.

References

1) Jelodar H, Wang Y, Orji R, Huang S. Deep Sentiment Classification and Topic Discovery on Novel Coronavirus or COVID-19 Online Discussions: NLP Using LSTM Recurrent Neural Network Approach. *IEEE Journal of Biomedical and Health Informatics*. 2020;24(10):2733–2742. Available from: https://doi.org/10.1109/JBHI.2020.3001216.
2) Naseem U, Razzak I, Khushi M, Eklund PW, Kim J. COVIDSenti: A Large-Scale Benchmark Twitter Data Set for COVID-19 Sentiment Analysis. *IEEE Transactions on Computational Social Systems*. 2021;8(4):1003–1015. Available from: https://doi.org/10.1109/TCSS.2021.3051189.
3) Wang N, Lv X. Research on emotional analysis of netizens and topic distribution under public health emergencies : ——A Case Study of COVID-19. In: 2020 International Conference on Public Health and Data Science (ICPHDS), 20-22 November 2020, Guangzhou, China. IEEE. 2020;p. 76–80. Available from: https://doi.org/10.1109/ICPHDS51617.2020.00023.
4) Wahid JA, Hussain S, Wang H, Wu Z, Shi L, Gao Y. Aspect oriented Sentiment classification of COVID-19 twitter data; an enhanced LDA based text analytic approach. In: 2021 International Conference on Computer Engineering and Artificial Intelligence (ICCEAI), 27-29 August 2021, Shanghai,

China. IEEE. 2021. Available from: https://doi.org/10.1109/ICCEAI52939.2021.00054.

5) Ji LI, Xueyan T, Daoli D. Identification of Public Opinionon COVID-19 in Microblogs. In: 2021 16th International Conference on Computer Science & Education (ICCSE), 17-21 August 2021, Lancaster, United Kingdom. IEEE. 2021. Available from: https://doi.org/10.1109/ICCSE51940.2021.9569649.
6) Long Z, Alharthi R, Saddik AE. NeedFull – a Tweet Analysis Platform to Study Human Needs During the COVID-19 Pandemic in New York State. *IEEE Access*. 2020;8:136046–136055. Available from: https://doi.org/10.1109/ACCESS.2020.3011123.
7) Mudassir MA, Mor Y, Munot R, Shankarmani R. Sentiment Analysis of COVID-19 Vaccine Perception Using NLP. In: 2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA), 02-04 September 2021, Coimbatore, India. IEEE. 2021;p. 516–521. Available from: https://doi.org/10.1109/ICIRCA51532.2021.9544512.
8) Gupta P, Kumar S, Suman RR, Kumar V. Sentiment Analysis of Lockdown in India During COVID-19: A Case Study on Twitter. *IEEE Transactions on Computational Social Systems*. 2021;8(4):992–1002. Available from: https://doi.org/10.1109/TCSS.2020.3042446.
9) D'Andrea E, Ducange P, Marcelloni F. Monitoring negative opinion about vaccines from tweets analysis. In: 2017 Third International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN), 03-05 November 2017, Kolkata, India. Kolkata, India. IEEE. 2017;p. 186–191. Available from: https://doi.org/10.1109/ICRCICN.2017.8234504.
10) Chandrasekaran R, Mehta V, Valkunde T, Moustakas E. Topics, Trends, and Sentiments of Tweets About the COVID-19 Pandemic: Temporal Infoveillance Study. *Journal of Medical Internet Research*. 2020;22(10):1–12. Available from: https://www.jmir.org/2020/10/e22624/.
11) Wang M, Mengoni P. How Pandemic Spread in News: Text Analysis Using Topic Model. In: 2020 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT), 14-17 December 2020, Melbourne, Australia. IEEE. 2020;p. 764–770. Available from: https://doi.org/10.1109/WIIAT50758.2020.00118.
12) Mishra RK, Urolagin S, Jothi JAA, Neogi AS, Nawaz N. Deep Learning-based Sentiment Analysis and Topic Modeling on Tourism during Covid-19 Pandemic. *Frontiers in Computer Science*. 2021;3:1–14. Available from: https://doi.org/10.3389/fcomp.2021.775368.
13) Krajah A, Almadani YF, Saadeh H, Sleit A. Analyzing Covid-19 Data Using Various Algorithms. In: 2021 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT), 16-18 November 2021, Amman, Jordan. IEEE. 2021;p. 66–71. Available from: https://doi.org/10.1109/JEEIT53412.2021.9634124.
14) Shurrab S, Shannak Y, Almshnanah A, Khazaleh H, Najadat H. Attitudes Evaluation Toward COVID-19 Pandemic: An Application of Twitter Sentiment Analysis and Latent Dirichlet Allocation. In: 2021 12th International Conference on Information and Communication Systems (ICICS), 24-26 May 2021, Valencia, Spain. IEEE. 2021;p. 265–272. Available from: https://doi.org/10.1109/ICICS52457.2021.9464558.
15) Arbane M, Benlamri R, Brik Y, Alahmar AD. Social media-based COVID-19 sentiment classification model using Bi-LSTM. *Expert Systems with Applications*. 2023;212:1–9. Available from: https://doi.org/10.1016/j.eswa.2022.118710.
16) Mishra RK, Urolagin S, Jothi JAA, Neogi AS, Nawaz N. Deep Learning-based Sentiment Analysis and Topic Modeling on Tourism during Covid-19 Pandemic. *Frontiers in Computer Science*. 2021;3:1–14. Available from: https://doi.org/10.3389/fcomp.2021.775368.
17) Farouk M. Measuring Sentences Similarity: A Survey. *Indian Journal of Science and Technology*. 2019;12(25):1–11. Available from: https://doi.org/10.17485/ijst/2019/v12i25/143977.
18) Jang H, Rempel EC, Carenini G, Janjua N. Exploratory Analysis of COVID-19 Related Tweets in North America to Inform Public Health Institutes. In: Proceedings of the 1st Workshop on NLP for COVID-19 (Part 2) at EMNLP 2020. Association for Computational Linguistics. 2020;p. 1–6. Available from: https://aclanthology.org/2020.nlpcovid19-2.18.
19) Velvizhy P, Pravi A, Selvi M, Ganapathy S, Kannan A. Fuzzy-based review rating prediction in e-commerce. *International Journal of Business Intelligence and Data Mining*. 2020;17(1):101–116. Available from: https://doi.org/10.1504/IJBIDM.2020.108034.
20) Kaila P, Dr, Rajesh, Prasad DAV, Krishna. Informational Flow on Twitter - Corona Virus Outbreak - Topic Modelling Approach. *International Journal of Advanced Research in Engineering and Technology (IJARET)*. 2020;11(3):128–134. Available from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3565169.
21) Huang B, Carley KM. Disinformation and Misinformation on Twitter during the Novel Coronavirus Outbreak. 2020. Available from: https://arxiv.org/abs/2006.04278.
22) Baldha T, Mungalpara M, Goradia P, Bharti S. Covid-19 Vaccine Tweets Sentiment Analysis and Topic Modelling for Public Opinion Mining. In: 2021 International Conference on Artificial Intelligence and Machine Vision (AIMV), 24-26 September 2021, Gandhinagar, India. IEEE. 2021;p. 1–6. Available from: https://doi.org/10.1109/AIMV53313.2021.9671000.
23) Khan K, Yadav S. Sentiment analysis on covid-19 vaccine using Twitter data: A NLP approach. In: 2021 IEEE 9th Region 10 Humanitarian Technology Conference (R10-HTC), 30 September 2021 - 02 October 2021, Bangalore, India. IEEE. 2021. Available from: https://doi.org/10.1109/R10-HTC53172.2021.9641515.