Journal of Intelligent & Fuzzy Systems 43 (2022) 6307–6319
DOI:10.3233/JIFS-220279
IOS Press

Multilayer hybrid ensemble machine learning model for analysis of Covid-19 vaccine sentiments

Vipin Jain* and Kanchan Lata Kashyap
VIT University, Bhopal, Madhya Pradesh, India

Abstract. This work presents the analysis of significant sentiments and attitudes of people towards the COVID-19 vaccination. The tweeter messages related to the COVID-19 vaccine is used for sentiment evaluation in this work. The proposed work consists of two steps: (i) natural processing language (NLP) and (ii) classification. The NLP is utilized for text pre-processing, tokenization, data labelling, and feature extraction. Further, a stack-based ensemble machine learning model is used to classify sentiments as positive, negative, or neutral. The stack ensemble machine learning model includes seven heterogeneous machine learning techniques namely, Naive Bayes, Logistic regression, Decision Tree, Random Forest, AdaBoost Classifier, Gradient Boosting, and extreme Gradient Boosting (XGB). The highest classification accuracy of 97.2%, 88.34%, 88.22%, 85.23%, 86.30%, 87.54%, 86.63%, and 88.78% is achieved by ensemble machine learning model, Logistic regression, AdaBoost, Decision Tree, Naive Bayes, Random Forest, Gradient Boosting, and XGB Classifier, respectively.

Keywords: COVID-19 vaccinations, sentiments, social-media, machine learning, ensemble machine learning

1. Introduction

Coronavirus disease (COVID-19) is a lethal virus that afflicts many countries. It is considered as pandemic disease in March 2020 by World Health Organization [29]. COVID-19 pandemic affected the lives of more than 200 nations [30]. Hundreds of thousands died from the unexpected outbreak of COVID-19. Governments and authorities actively combated the disease worldwide through various tactics and policies such as travel restrictions, vaccination, and facility closures. The invention of a COVID-19 vaccine is one of the important technique applied by the government to control its spreading [26]. Many countries such as America, Britain, and Brazil reported fewer COVID cases in 2021 due to the high percentage of vaccination. The Indian government has decided to start a widespread immunization effort to stop spreading of COVID-19. The COVID-19 vaccination was initially available to health care and front-line personnel. Citizens above the age of 18 are now part of the phase-3 vaccination drive [33]. The effectiveness of any immunization campaign is determined by its publicity rate and pace of acceptance [12]. Many misconceptions and doubts emerges in ordinary people's minds about COVID-19 vaccinations. The effectiveness of the vaccination program is determined by public approval. Vaccine development is a slow and time-consuming process which requires multiple test for potency, efficacy, and protection. Furthermore, the acceptability of the newly released vaccine is required for an effective immunization program. Preliminary data indicates that the authorized vaccinations are safe and efficient. Still the long-term efficacy and adverse effects are unclear. The vaccination and immunization of 100% population can avert the pandemic effectively. However, research on the

*Corresponding author. Vipin Jain, VIT University, Bhopal, 466114, Madhya Pradesh, India. E-mail: vipin.jain2020@vitbhopal.ac.in.

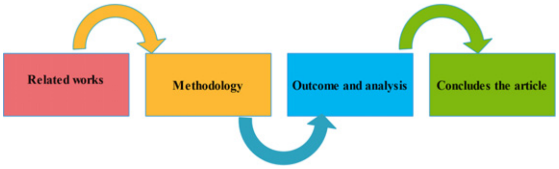


Fig. 1. Organization of the paper.

sentiment analysis of public view towards COVID-19 vaccine has not been done yet.

This study aims to investigate the views of Indians regarding the COVID-19 vaccination. The present work focuses on the sentiment analysis of Twitter messages of Indian people about the COVID-19 vaccine and classify their sentiments into three groups as positive, neutral, and negative. The outcome of the present work can be utilized by the Policymakers to address the people's queries before mass vaccination. Organization of the remaining part of the paper is shown in Fig. 1. Section 2 discusses the literature review. The proposed methodology and algorithm are presented in Section 3. The experimental results and discussions are given in Section 4, followed by conclusions are given in Section 5.

- **Novelty of the proposed work**

The present work investigated the sentiment analysis of Indian people about the COVID-19 vaccine based on Twitter messages. The proposed work broadly includes two steps (i) Natural Language Processing(NLP) and (ii) Sentiment classification. In the first step, natural language processing is applied for data preprocessing. The data preprocessing includes data cleaning, removal of unrelated words, data normalization, and tokenization process. Further, preprocessed data is utilized for the feature extraction and data labelling. Finally, stack ensemble machine learning algorithm is applied for sentiment classification as positive, negative, and neutral. Total seven machine learning techniques listed as Naive Bayes, Random Forest, Logistic regression, Decision Tree, Gradient Boosting, AdaBoost, and Extreme Gradient Boosting (XGB) are utilized to construct ensemble machine learning model.

2. Background information and related work

Alam et al. investigated the public opinion about COVID-19 vaccinations based on the tweets posted between December 2021 to July 2021 [1]. The NLP based tool named as Valence Aware Dictionary for Sentiment Reasoner (VADER) is used to analyze the attitudes about vaccination in their work. The performance of the predictive model is tested using a recurrent neural network, long short-term memory (LSTM), and bidirectional LSTM (Bi-LSTM). The highest 90.59% and 90.83% of accuracy obtained with LSTM and Bi-LSTM model, respectively. Aspect-based sentiment analysis is used by Aygün et al. with six different COVID-19 vaccine-related tweets [2]. Four distinct BERT (Bidirectional Encoder Representations from Transformers) models namely, mBERT-base, BioBERT, ClinicalBERT, and BERTurk is applied in their work with highest accuracy of 87%. COVID-19 Arabic tweets are examined by Baker et al. with 54,065 Twitter posts with four classifiers, namely SVM, k-NN, decision trees, and Naive Bayes [3]. The highest classification accuracy of 89.06% and 86.43% is achieved through Naive Bayes and k-NN, respectively. Bonnevie et al. studied the evolution of vaccine resistance by evaluating the tweets about COVID-19 vaccine posted by citizens of United States [4]. Hou et al. analyzed public interest of COVID-19 by using the Weibo posts [19]. The emotions are classified based on Baidu emotions analysis tool in their work. Hung M et al. utilized a lexicon-based technique to determine the emotional state of COVID-19 [20]. The Latent Dirichlet Allocation (LDA) is utilized to extracts latent semantics patterns from Twitter posts. The three sentiment analysis such as pleasant, neutral, and negative is performed based on Dictionary of Valence Aware and sentiment Reasoner (VADER). The compound score or sentiment score with emotion ratings are also computed by the authors. Jain et al. applied two lexicons namely, SentiwordNet and AFINN for sentiment analysis [21]. Authors applied SVM and Naive Bayes classifier for tweet classification. Lwin et al. utilized 20,325,929 pandemic-related tweets to gauge public emotions using a lexicon technique [25]. The CrystalFeel algorithm is employed by authors to classify four different sentiments such as fear, anger, sorrow, and joy. Total 80,32,78 Persian tweets related to imported vaccines, i.e., Pfizer/BioNTech, AstraZeneca/Oxford, Moderna, and Sinopharm are utilized by Nezhad et al. for sentiment analysis [28]. Deep learning based CNN-LSTM model has been used to determine the sentiments of retrieved tweets by the authors. Valdés et al. utilized a hybrid technique to analyze 1,499,227 vaccine-related tweets from 18 March 2019 and 15 April 2019 with accuracy rate of more than 85% [32]. Praveen et al. used

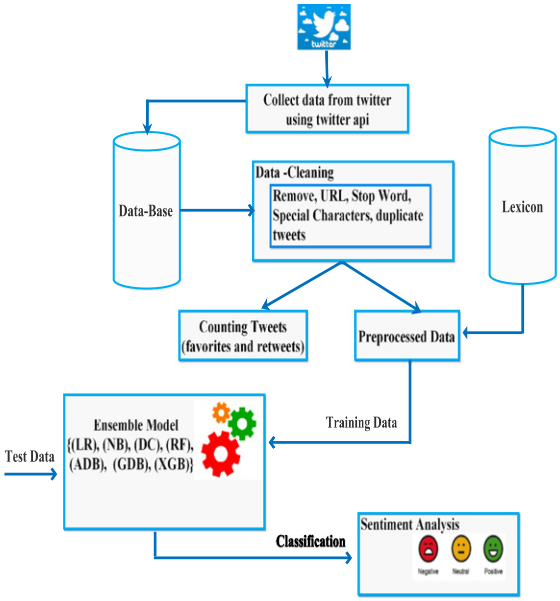

twitter
Collect data from twitter using twitter api
Data-Base
Data -Cleaning
Remove, URL, Stop Word, Special Characters, duplicate tweets
Lexicon
Counting Tweets (favorites and retweets)
Preprocessed Data
Training Data
Ensemble Model {(LR), (NB), (DC), (RF), (ADB), (GDB), (XGB)}
Test Data
Classification
Sentiment Analysis
Negative
Neutral
Positive

Fig. 2. Flow chart of the proposed work.

a text-blob lexicon and Latent Dirichlet Allocation to study Indians attitudes regarding COVID-19 immunisation [34]. Villavicencio et al. suggested Naive Bayes algorithms for analyzing the general sentiments of Filipino citizens about COVID-19 vaccines with 81.77%. accuracy [38]. Liu et al. suggested an attention mechanism and a graph convolutional network for Aspect-based sentiment categorization [24]. The multi-head attention technique is utilized in their work to select input data and achieved 82.72% accuracy with BERT embedding.

3. Methodology used

The proposed methodology for sentiment analysis and classification consist of four steps (i) Data gathering, (ii) Data preprocessing, (iii) Sentiment analysis, and (IV) Sentiment classification as positive, negative or neutral. The flowchart and algorithm of the proposed methodology is illustrated in Fig. 2 and algorithm 1, respectively. The brief description of each step of present work is given in subsequent subsections.

3.1. Data gathering

Social media is the biggest resource of understanding the general viewpoint of public opinions during extraordinary times [5, 7, 13, 40]. The Twitter data related to the COVID-19 vaccine is collected for the sentiment analysis in this work. Total 27,810 tweets related to "COVID-19 vaccination" are collected using Tweepy API.

Algorithm 1 Sentiment analysis of COVID-19 vaccine

```
Require: Twitter data (tweets) about COVID-19 vaccine
Ensure: Sentiment analysis as positive, negative, and neutral)
Twitter scrap using twint.
Collection of tweets as data-set.
Apply data cleaning and preprocessing operation on data-set df1
df1 ⇐ Deletion of null value.
df1 ⇐ conversion of df1 text into lower case.
df1 ⇐ Removal of stop words, @, and URL from data-set.
df1 ⇐ Removal of emoticons and punctuation from df1.
df1 ⇐ Tokenization and lemmatization.
FX ⇐ Feature extraction (TF-IDF, n-grams) from df1.
FX["ps"] ⇐ Calculation of polarity score of tweets stored in the data-set.
Convert polarity score into sentiment categories.
for each $FX[``ps'']_i$, where i =0. . . . . . . . . n
    If $FX[``ps'']_i$ score > 0
        Assign as "Positive".
    else if
        $FX[``ps'']_i$ score < 0
        Assign as "Negative".
    else
        Assign as "Neutral".
end for
Create a filter function filter-by-vaccy(df,vax), filter vaccine namely, 'covaxin', 'moderna', 'pfizer', 'biontech', 'sputnik', 'covishield' according to timeline from the data-set FX.
FX["vaccine-name"] ⇐ Divide all the emotions into three categories: "positive", "negative", and "neutral."
for each FX["vaccinename"] The frequency of each word must be calculated as "positive," "negative," and "neutral".
end for
Analysis of negative sentiment of each vaccine
```

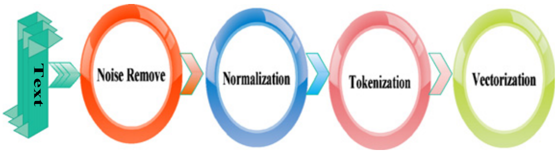


Fig. 3. Steps of data pre-processing.

3.2. Data pre-processing

The data preprocessing process consists of four sub-steps (i) data cleaning, (ii) Normalization of data, (iii) Tokenization of text, and (iv) Vectorization as shown in Fig. 3. As the collected tweets are noisy and unlabelled. Thus, The data cleaning process is applied to remove noises which includes (a) Removal of unnecessary words, (b) Elimination of HTML tags, (c) Elimination of emojis and data numbering patterns, (d) Removal of additional characters in sentences, and (e) Removal of blank spaces, emails, stopping words, URLs, and punctuation. Further, Stemming and Lemmatization techniques are applied for data normalization by using WordNetLemmatize package of natural language tool kit (nltk). The Stemming technique removes terminators from words to determine the root form. The Lemmatization is used for grouping the various forms of similar words for dimensionality reduction. In next step, normalized data-set are tokenized by dividing the text string into tokens and stored as features for labelling process.

- **Data labelling**

The sentences are divided into nouns, verbs, adjectives, and adverbs using the Textblob Python library. Further, sentences are labelled as positive, negative, and neutral based on the polarity score. The mathematical expression for computation of polarity score is given as:

$$Polarity\_score = \frac{\sum(Polarity\_of\_Word)}{\text{Total number of word in a sentence}} \tag{1}$$

here, Polarity_of_Word is a predefined value in the textblob. Labelling of sentences as negative, positive, and neutral is done by using polarity score. the negative, positive, and neutral sentiment is labelled as -1, +1, and 0, respectively .

TF-IDF : The tf-idf vectorization technique converts the text data into the vector form suitable for further classification. It is an features extraction method works with labelled data and based on the frequency of words in the text. TF represents phrase frequency which computes how often the word has appeared in the text and mathematically represented as :

$$TF(G) = \frac{\text{frequency of word G in a document } d}{\text{Total words in a document}} \tag{2}$$

The inverse frequency of words is computed by mathematical expression given as:

$$IDF(G) = log_e \frac{\text{frequency of word G in a document } d}{\text{Total words in a document}} \tag{3}$$

The TF-IDF score is computed by using following mathematical formula :

$$TF - IDF(G) = TF(G) \times IDF(G) \tag{4}$$

3.3. Classification

Sentiment classification as positive, neutral, and negative is performed by applying eight machine learning models namely, Random forest, Logistic regression, Naive Bayes, Decision tree, AdaBoost, Gradient boosting, Extreme Gradient Boosting, and a stack ensemble model. Brief descriptions of each machine learning techniques are given in the subsequent subsections.

3.3.1. Logistic regression

A logistic regression classifier is adopted by machine learning from the field of statistics which establishes the relation between independent and dependent variables. The logistical function of the classifier is used to obtain the input, set of weighted functions, and the correlation between event classes [17]. The accuracy and generalizability of the model can be increased by proper selection of the features. The feature vector *i* is categorized as positive, neutral, or negative with mathematical expression represented as:

$$S(\text{f} = 1 \mid \text{i}) = l(i) = \frac{1}{1 + h^{zw_i}} \tag{5}$$

here, S denotes the probability of text *i* which belongs to class f, z represents the feature weight.

3.3.2. Naive Bayes

The Naive Bayes is a probabilistic based classifier that estimates the group probability [27]. This classifier requires minimal amount of training data.

It gives better results due to its strong foundation and simplicity [31]. The Bayes theorem is based on the mathematical equation given as [23]:

$$T(G \mid B) = \frac{T(B \mid G) \cdot T(G)}{T(B)} \qquad (6)$$

Here, T(G) represents prior probability of class G. T(B) is the knowledge from the text itself to be categorised. T(B | G) is the probability of document B having a distribution in the class space G.

3.3.3. Decision tree

The decision tree classifier provides provision of multi-stage decisions by distributing a complicated problem into a union of a smaller one [10, 18]. Data classification is performed by using entropy technique represented as:

$$L = -\sum u(c) \log u(c) \qquad (7)$$

u(c) denotes the probability of an occurrence c in state L.

3.3.4. Random forest classifier

Random forest classifier uses many decision trees to resolve the regression and classification problems [6]. This classifier operates by generating many decision-making models during training and anticipating the most common classes of decision-makers. It uses the Gini Index and Entropy for the classification of data which are mathematically defined as:

$$\text{Gini Impurity } = \sum_{n=1}^{i} Kn(1 - Kn), \qquad (8)$$

$$\text{Entropy } = \sum_{n=1}^{i} -K \log(Kn) \qquad (9)$$

here, the total number of classes is denoted by i and the probability of selecting a data point inside a class is denoted by K(n).

3.3.5. AdaBoost algorithm

AdaBoost is widely used boosting algorithm based on an ensemble approach for building a strong classifier from several weak classifiers [15]. It integrates the several classifiers in each cycle with training set selection and delivers the final vote.

3.3.6. Gradient boosting classifier

Gradient boosting classifiers combine many weak learning models to generate a potent prediction model for categorization of huge data-set. The bias error can be reduced by this model. The gradient boosting model create an approximation, $\widehat{H}(\mathbf{b})$ of the function $H^*(\mathbf{b})$, that converts instances b to their output values z. Function approximation $H^*(\mathbf{b})$ can be represented as a weighted sum of functions as :

$$H_c(\mathbf{b}) = H_{b-1}(\mathbf{b}) + \rho_c q_c(\mathbf{b}) \qquad (10)$$

here, ρ_c denotes the weight of the c^{th} function $q_c(\mathbf{b})$

3.3.7. Extreme gradient boosting classifier

XGB includes the collection of gradient boosting techniques designed for current data science challenges and tools. It is an ensemble model of classification and regression tree sets (CART). XGB is highly scalable parallelizable, faster, and regularised to control over-fitting. The mathematical description of this model is given as:

$$\hat{A}_l = \sum_{p=1}^{P} q_p\,(m_l)\,, q_p \in Q \qquad (11)$$

here, P denotes the total number of trees, the q_p for p^{th} tree denotes a function in functional space Q, and Q represents set of all possible CARTs.

3.3.8. Stack ensemble machine learning model

The results of the machine learning model can be improved by an ensemble machine learning technique for training and testing [8]. Stack ensemble machine learning technique enables a better predictive model as compared to a single model by combining predictions from several models [14, 22, 36]. Block diagram of stacking machine learning model is shown in Fig. 4. A stacking model consists of two or more base models, in the first layer. Second layer known as Meta-Classifier combines the predictions of all base models. The multi-layer stacking ensemble technique with seven base models namely, Naive Bayes, Logistic regression, Decision Tree, Random Forest, AdaBoost, GradientBoosting, and Extreme GradientBoosting, has been used to construct an ensemble model as shown in Fig. 5. Whole training data-set is used to train all base models. Meta-Classifier model has been trained by using the output generated by base models. All base-models used in stacking are distinct and fit to the same data-set. The XGB classifier is used as a Meta-Classifier in this work. The working process of the stack ensemble model as:

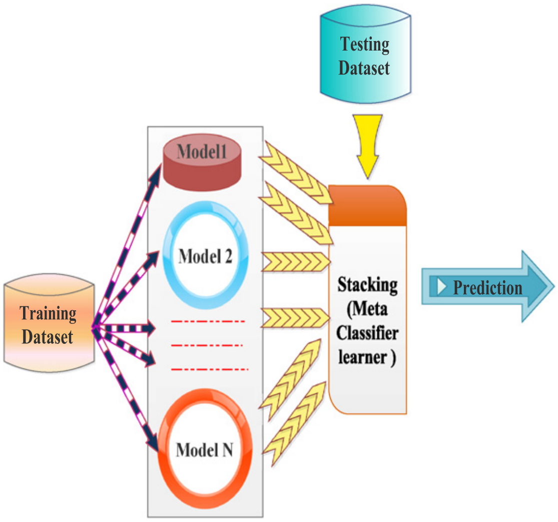


Fig. 4. Stacking ensemble technique.

i. Divide the data-set into two sets: training and testing. Afterward, the training data is broken down into K-folds.
ii. A base model is fitted into K-1 pieces and given the predictions for the K^{th} part.
iii. This approach is repeated until anticipated all folds.
iv. After that, the base model is fitted to the entire training data-set in order to compute performance of test data-set .
v. Repetition of steps ii-iv for other base model
vi. Predictions of the base models are used as input features for the Meta-Classifier.
vii. The second-level model is used to classify the test data-set as positive, neutral, and negative.

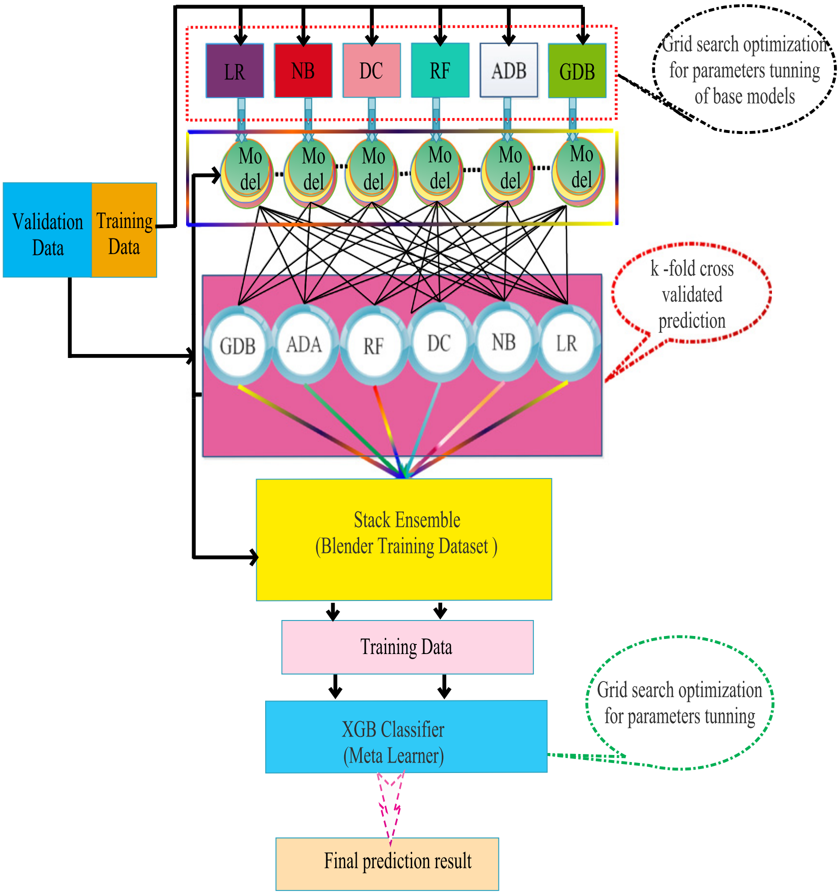

LR
NB
DC
RF
ADB
GDB
Grid search optimization for parameters tunning of base models
Mo del
Mo del
Mo del
Mo del
Mo del
Mo del
Validation Data
Training Data
k -fold cross validated prediction
GDB
ADA
RF
DC
NB
LR
Stack Ensemble (Blender Training Dataset)
Training Data
XGB Classifier (Meta Learner)
Grid search optimization for parameters tunning
Final prediction result

Fig. 5. Proposed hybrid machine learning architecture using multi-layer stacking ensemble.

Table 1
Dataset contents

Id	User-name	Date	Text
340539111-971516416	Rachel Roh	2020-12-20 06:06:44	The same sources said that daikon paste may be used ..
133785121-5875608579	Gunther Fehlinger	2020-12-12 20:06:00	It is a little depressing to attribute success to âŁ¦
133784229-5857623042	Ch.Amjad Ali	2020-12-12 19:30:33	States will start getting C
133784193-4170255365	Tamer Yazar	2020-12-12 19:29:07	while deaths are closing in on the 300,000 mar...

Table 2
Hashtags

Hashtags 1	Hashtags 2	Hashtags 3
#vaccination	#AstraZeneca	#VaccinesSaveLives
#CovidVaccine	#Pfizervaccine	#Moderna
#PfizerBioNTech	#COVID19Vaccine	#COVIDvaccine'

4. Experimental results and discussions

The outcome of sentiment analysis using NLP and performance evaluation of the various applied machine learning techniques are discussed in this section. In this work, total 27,810 tweets of year 2020 and 2021 are collected for sentiment analysis of Indian people about vaccines. The Python Tweepy API is used for collection of tweets. Total 15 extracted fields such as *user i_d, user_name, user_location, user_description, user_created, user_ followers, user_friends, user_fav- ourites, user_verified, date, text, has_htags, source, re_tweets, favorites,re_tweet* are stored as database. The sample data of the collected data-set is shown in Table 1.

The hashtags linked to tweets of COVID-19 vaccination trend denotes the people opinions about it. Sample tweets collected from hashtags are shown in Table 2.

Figure 6(a) and (b) illustrated the outcome of the preprocessed data-set and polarity score of sentence. Polarity score value is utilized for labelling of data-set as positive, negative, and neutral. Figure 7 shows negative sentiment tweets about the COVID-19 vaccine. The negative, neutral, and positive sentiments of the user towards Pfizer or BioNTech COVID-19 vaccines are shown in Fig. 8. The word cloud for the Moderna vaccine is shown in Fig. 9, which reveals similar information as the Pfizer tweets. The side effects of this vaccine such as arm cramps, chilliness, nervousness, nausea, tiredness sensations, and other minor signs can be seen. Positive Moderna-specific tweets are similar to positive Pfizer-specific tweets posted by users who have completed their first dose. Lastly, the word cloud for Covaxin is illustrated in Fig. 10. The words "propaganda" and "political" are associated with negative tweets which require further investigation. The distribution of sentiment after data labelling is shown in Table 3. Highest 50.39%, 39.84%, and 9.77%, input tweets are analyzed as neutral, positive, and negative tweets, respectively. The COVID-19 cases decreases as positive attitude increases towards the vaccination, whereas negative attitude towards vaccines increases the COVID-19 cases.

4.1. Classification results and discussion

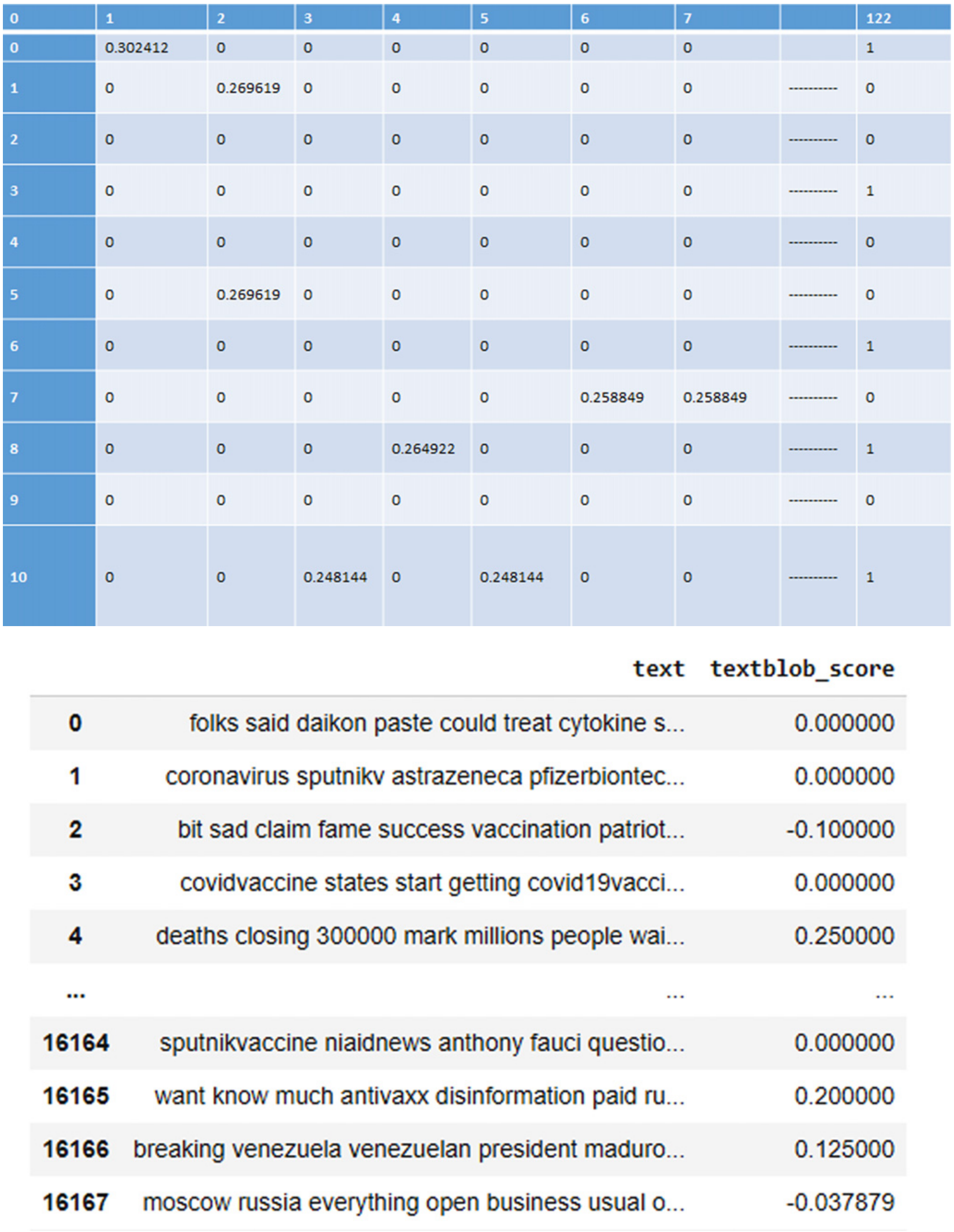
The 5-fold cross-validation technique is used for training and testing of all seven machine learning models namely, linear regression, decision tree, random forest, AdaBoost, GradientBoost, XGB, and stack ensemble. The various parameters values initialized for the stacked ensemble classifier are presented in Table 4. The various components used for training and testing with 5-fold cross-validation technique are presented in Fig. 11. First, training data-set is randomly divided into five sub-parts. The first four sub-part and last remaining sub-part is used as training and testing data, respectively.

I. First four fold training data is used to train a model.
II. Validity of the resultant model is tested for the rest of the data-set component.

4.2. Performance assessment parameters of classifiers

Performance of classifiers are measured in terms of accuracy, precision, recall, f1- score, and receiver operating curve (ROC). Definition and mathematical expression of each parameters are given as:
Accuracy is defined as the number of data correctly classified by the classifier out of the total number of data instance. The mathematical representation of accuracy is given as:

$$Accuracy = \frac{\sum(M\_Q, M\_X)}{\sum(M\_Q, W\_Q, M\_X, W\_X)} \quad (12)$$



0	1	2	3	4	5	6	7		122
0	0.302412	0	0	0	0	0	0		1
1	0	0.269619	0	0	0	0	0	----------	0
2	0	0	0	0	0	0	0	----------	0
3	0	0	0	0	0	0	0	----------	1
4	0	0	0	0	0	0	0	----------	0
5	0	0.269619	0	0	0	0	0	----------	0
6	0	0	0	0	0	0	0	----------	1
7	0	0	0	0	0	0.258849	0.258849	----------	0
8	0	0	0	0.264922	0	0	0	----------	1
9	0	0	0	0	0	0	0	----------	0
10	0	0	0.248144	0	0.248144	0	0	----------	1

(a)

	text	textblob_score
0	folks said daikon paste could treat cytokine s...	0.000000
1	coronavirus sputnikv astrazeneca pfizerbiontec...	0.000000
2	bit sad claim fame success vaccination patriot...	-0.100000
3	covidvaccine states start getting covid19vacci...	0.000000
4	deaths closing 300000 mark millions people wai...	0.250000
...	...	...
16164	sputnikvaccine niaidnews anthony fauci questio...	0.000000
16165	want know much antivaxx disinformation paid ru...	0.200000
16166	breaking venezuela venezuelan president maduro...	0.125000
16167	moscow russia everything open business usual o...	-0.037879

(b)

Fig. 6. (a) Preprocessing outcome. (b) Polarity value.

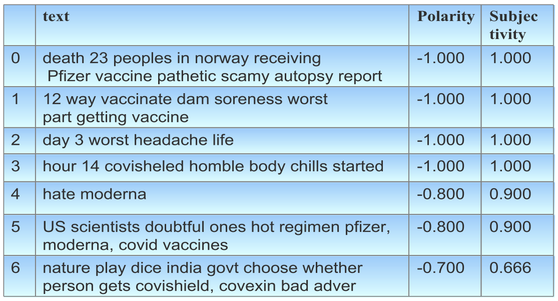
Precision, defines as the ratio of true positive samples over the total number of predicted positive samples [37]. Mathematically it is represented as:

$$Precision = \frac{(M\_Q)}{\sum(M\_Q, W\_Q)} \qquad (13)$$

A recall is defined as accurate positive predictions divided by the total number of positive [39]. It is mathematically denoted as:

$$Recall = \frac{(M\_Q)}{\sum(M\_X, W\_X)} \qquad (14)$$

The F1-Score needs to provide W_Q weighting and occasionally W_X weighting. F1 is a weighted average precision and recall, which implies that W_Q and



	text	Polarity	Subjectivity
0	death 23 peoples in norway receiving Pfizer vaccine pathetic scamy autopsy report	-1.000	1.000
1	12 way vaccinate dam soreness worst part getting vaccine	-1.000	1.000
2	day 3 worst headache life	-1.000	1.000
3	hour 14 covisheled homble body chills started	-1.000	1.000
4	hate moderna	-0.800	0.900
5	US scientists doubtful ones hot regimen pfizer, moderna, covid vaccines	-0.800	0.900
6	nature play dice india govt choose whether person gets covishield, covexin bad adver	-0.700	0.666

Fig. 7. Popular negative sentiment of input data-set.

W_X have equal significance.

$$F1score = \frac{2 * (Precision * Recall)}{\sum(Precision, Recall)} \quad (15)$$

here, M_Q, M_X, W_Q, W_X denotes the truly identified, wrongly identified, truly rejected, and wrongly rejected, respectively. The classification result in terms of accuracy is shown in Fig. 12. Highest classification accuracy **97.2%** is achieved by ensemble learning classifier.

Accuracy obtained by ensemble learning model is compared with all seven machine learning model as shown in Fig. 12 The precision, recall, and F1-score achieved from stack ensemble model along with seven basic models are also presented in Table 5. The value of precision, recall, and F1-score should be 1 (high) which indicates the good results. The value of precision, recall, and F1-score obtained by the proposed model near to 1 indicates the higher accuracy level of this model. The area under curve (AUC) achieved by the ROC curve and confusion matrix of all eight classifiers are shown in Figs. 13 and 14, respectively. The highest 97.2%, 0.99%, 96.86%, 95.41%, and 96.63% of classification accuracy, AUC, precision, recall, and F1-score have been achieved, respectively, by the stack ensemble model. It can be observed from the confusion matrix that the proposed model correctly identified 97% of the positive sentiment. The 96.23% and 96.71% of sentiments are correctly classified as neutral and negative, respectively.

4.3. *Comparison with existing work*

Outcome of the present work is compared with the existing work also. The comparative result of proposed model with some existing work is shown

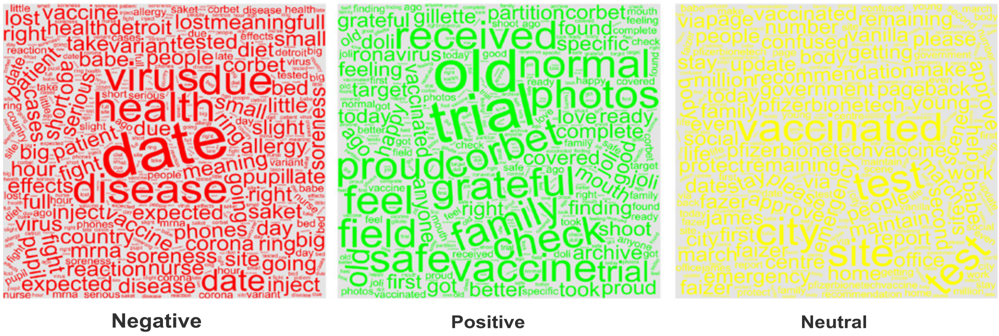


Fig. 8. Negative, positive, and neutral sentiment about Pfizer and BioNTech COVID-19 vaccine.

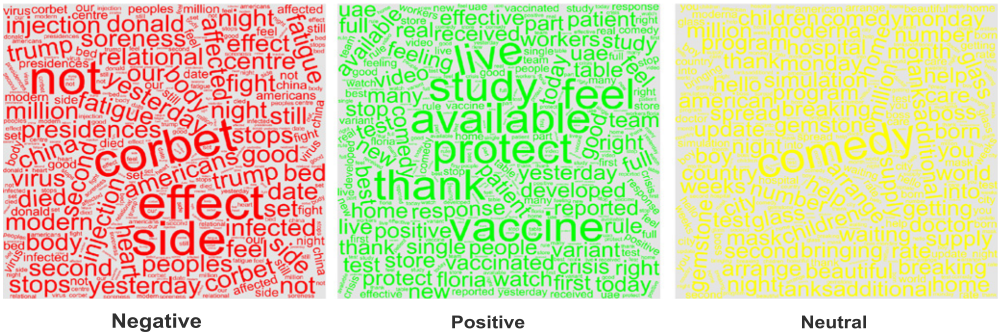


Fig. 9. Negative, positive, and neutral sentiment about Moderna COVID-19 vaccine.

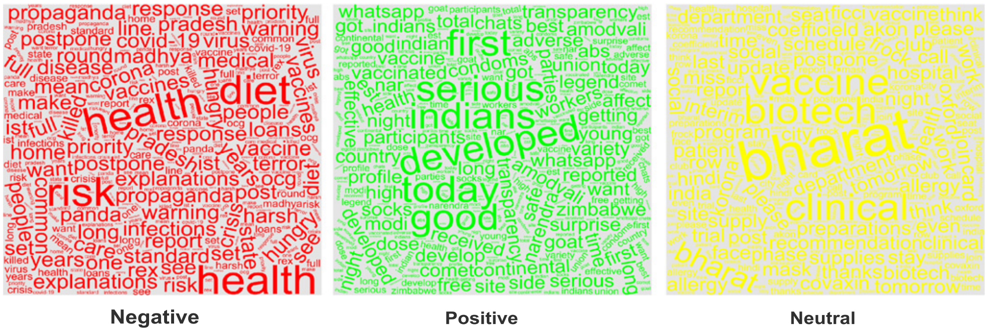


Fig. 10. Negative, positive, and neutral sentiment about Covaxin COVID-19 vaccine.

Table 3
Distribution of sentiments in data-set after data labelling

Total number of tweets	Neutral	Positive	Negative
27,810	14,013 (50.39%)	11,078 (39.84%)	2,718 (9.77%)

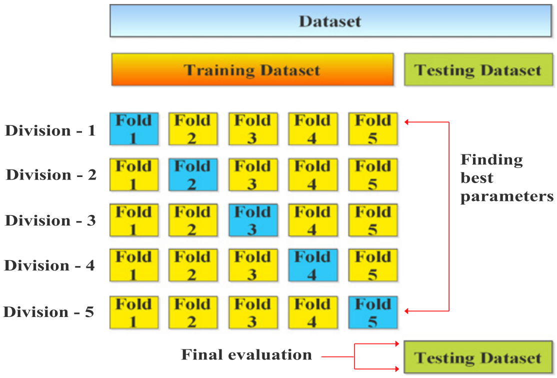


Fig. 11. Five-fold cross-validation architecture.

Table 4
Set of parameters and its value taken for stack ensemble model

Parameters	Value
max_depth	15
colsamplebytree	0.5
gamma	0.3
min_childweight	3
learning_rate	0.1

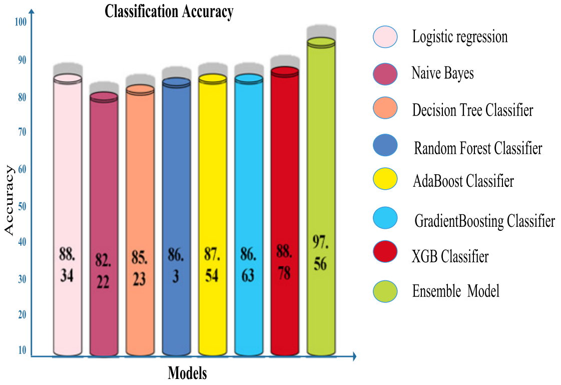


Fig. 12. Accuracy classification report.

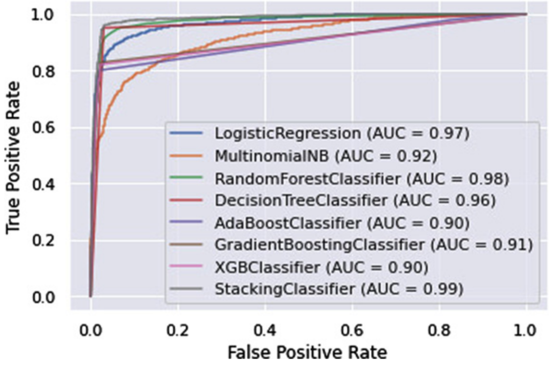


Fig. 13. ROC curve obtained from all classifiers.

in Table 6. A Gaussian membership function-based fuzzy rule is presented by Chakraborty et al. for sentiments detection of COVID-19 Twitter messages and achieved highest accuracy of 79% [9]. Garcia et al. examined a large number of COVID-19 tweets of Brazil and USA [16]. The Gibbs sampling algorithm and Topic modelling for the Dirichlet Multinomial Mixture are used for sentiment analysis in their work with 87% accuracy. In the context of text analytics, Samuel et al. has been given a methodological

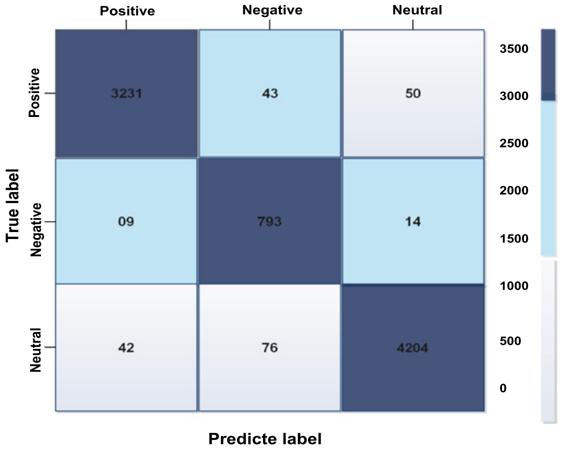


Fig. 14. Confusion matrix obtained by stack ensemble model.

analysis of two machine classification algorithms and compared their efficacy for categorizing COVID-19 tweets [35]. They achieved a high rating accuracy of 91% and 74% using the Naive Bayes and logistic regression techniques, respectively. The highest classification accuracy of 97.2% is achieved by the proposed stack ensemble machine learning model. The proposed model outperforms the existing work.

5. Conclusions

The sentimental analysis of Indian people about various COVID-19 vaccine is analyzed in this work. The analysis has been done on the sample data collected from Twitter platform. Various natural language processing operations has been applied for text preprocessing of raw text messages. Total 50.39%, 39.84%, and 9.77% tweets are labelled as neutral, positive, and negative, respectively. The proposed stack ensemble model is used for classification of sentiment as neutral, positive, and negative. Proposed model gives highest 97.2%, 0.99%, 96.86%, 95.41%, and 96.63% of classification accuracy, AUC, Precision, Recall, and F1-score, respectively. The obtained results proved that the ensemble model outperforms the basic machine learning techniques. The classification of other sentiments such as happy, sad, and fear will be performed in future work.

Q 1: How do the general people feel about the COVID-19 vaccine?

While most people's views on social media about vaccinations and their consequences are neutral, just 39.84 % are optimistic, which should be a source of concern for the government and policymakers. Vaccination will not be effective unless the government can persuade the majority of the population that the vaccine's results and consequences will be positive. The government should concentrate on overcoming vaccine anxiety before introducing mass vaccination.

Q 2: What are Indian citizens' core questions in the case of the COVID-19 vaccine?

Although nearly 11 million Indians have been infected with COVID-19. A sizable proportion of Indian people believe the pandemic has been exaggerated, according to our research. Citizens would oppose the vaccine as a result of this mindset. Aside from that, skepticism about the vaccines, skepticism about vaccine trials since receiving the vaccine, skepticism about one's health, fear of vaccination mortality, vaccine-induced allergic reactions, skepticism against pharmaceutical firms, Concerns about data from vaccine companies, a large number of vaccines are available, and there are questions over which ones are the safest. Although the Indian general public has legitimate doubts about the COVID-19 vaccine, our research has shown that certain conspiracy theories based on superstition have not been proved, COVID-19 is inflated. Certain vaccinations are viewed with scorn or disbelief due to nationality and have also been echoed when discussing the COVID-19 vaccine. According to our research, a significant portion of the Indian population does not

Table 5
Performance evaluation of different classifiers

Models	Precision_score	Recall_score	F1_score
Logistic Regression	95.46	72.46	82.38
Naive Bayes	88.44	69.78	71.34
Random Forest	91.43	85.13	88.17
Decision Tree	88.92	89.00	88.96
AdaBoost Classifier	92.33	87.83	84.46
GradientBoosting Classifier	95.04	87.67	89.05
XGB Classifier	95.12	85.49	87.57
Stacking Classifier	96.86	95.41	96.63

Table 6
Comparison of proposed work with existing work in terms of accuracy

Author/year	Technical approach	Accuracy (in %)	Dataset
Garcia et al. [16]	Topic modeling with Gibbs sampling algorithm for the Dirichlet Multinomial Mixture.	87%	COVID-19-related tweets from Brazil and USA.
Chakraborty et al. [9]	Gaussian membership function based fuzzy rule	79	COVID-19-related tweets.
Aygün et al.[2]	BERT models	87	COVID-19 vaccines .
Alam et al.[1]	Bidirectional LSTM (Bi-LSTM)	90.83	Collected from Kaggle .
Villavicencio et al.[38]	Naïve Bayes and RapidMiner	81.77	Gathered tweets related to COVID-19 vaccines in the Philippines .
Valdés et al.[32]	SVM classifier	86	COVID-19 vaccines related tweets had been Crawled through a API.
Samuel et al. [35]	Naïve Bayes and Logistic Regression classifiers.	91 and 74	COVID-19-related Tweets from Twitter.
Proposed Model	Bi-gram+Tf-idf+classifier (Logistic regression, Naive Bayes, Decision Tree, Random Forest, ADB, GDB, XGB) then ensemble all of them with XGB as Meta-Classifier.	**97.2**	Scrap twitter tweets related to COVID-19 vaccine.

trust the government or pharmaceutical companies on social media. According to recent studies, even COVID-19 survivors' immunity is only expected to last for eight months [11]. Policymakers and the government must engage in Vaccination education for the general public and the importance of returning to normal life. Governments, pharmaceutical firms, and Non-Governmental Organizations (NGOs) should make significant efforts to educate people about the vaccine program and tell them how essential it to regain normalcy. It is necessary to pay special attention to people's fears and misconceptions about COVID-19 vaccines to inspire and persuade people to get vaccines. With just over 39.84%of people saying they are optimistic about vaccinations, the government should overcome vaccine anxiety before introducing mass vaccination. It is crucial to understand why people hesitate to get a COVID-19 vaccination. It can assist health officials in raising vaccination awareness and reduce the spread of the illness. According to studies, vaccine uptake is a complex decision-making process. Using Random Forest, decision trees, logistic regression, and Naive Bayes categorize feelings as positive, neutral, or negative. The proposed ensemble model achieved 97.2% accuracy and outperforms to all other classifiers in terms of classification accuracy.

References

[1] K.N. Alam, M.S. Khan, A.R. Dhruba, M.M. Khan, J.F. Al-Amri, M. Masud and M. Rawashdeh, Deep learningbased sentiment analysis of covid-19 vaccination responses from twitter data, *Computational and Mathematical Methods in Medicine*, 2021.

[2] I. Aygun, B. Kaya and M. Kaya, Aspect based twitter sentiment analysis on vaccination and vaccine types in covid-19 pandemic with deep learning, *IEEE Journal of Biomedical and Health Informatics*, (2021).

[3] Q.B. Baker, F. Shatnawi, S. Rawashdeh, M. Al-Smadi and Y. Jararweh, Detecting epidemic diseases using sentiment analysis of arabic tweets, *J Univers Comput Sci* **26** (2020), 50–70.

[4] E. Bonnevie, J. Goldbarg, A.K. Gallegos-Jeffrey, S.D. Rosenberg, E. Wartella and J. Smyser, Content themes and influential voices within vaccine opposition on twitter, 2019, *American Journal of Public Health* **110** (2020), S326–S330.

[5] U. Brajawidagda and A.T. Chatfield, Twitter tsunami early warning network: A social network analysis of twitter information flows, (2012).

[6] L. Breiman, Random forests, *Machine Learning* **45** (2001), 5–32.

[7] C. Buntain, J. Golbeck, B. Liu and G. LaFree, Evaluating public response to the boston marathon bombing and other acts of terrorism through twitter, in: *Proceedings of the international AAAI conference on web and social media*, (2016).

[8] E. Cambria, D. Das, S. Bandyopadhyay and A. Feraco, Affective computing and sentiment analysis, in: A practical guide to sentiment analysis, Springer, (2017), pp. 1–10.

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

[10] J.R.B. Cockett, Discrete decision theory: Manipulations, *Theoretical Computer Science* **54** (1987), 215–236.

[11] R. COVID, 19. patients last immunity for 8 months, raise hopes for vaccine: Study.

[12] A.A. Dror, N. Eisenbach, S. Taiber, N.G. Morozov, M. Mizrachi, A. Zigron, S. Srouji and E. Sela, Vaccine hesitancy: the next challenge in the fight against covid-19, *European Journal of Epidemiology* **35** (2020), 775–779.

[13] A. Earle, A. Jagerskog and J. Ojendal, Transboundary WaterManagement: Principles and Practice, (2010).

[14] M. Fayaz, A. Khan, J.U. Rahman, A. Alharbi, M.I. Uddin and B. Alouffi, Ensemble machine learning model for classification of spam product reviews, *Complexity* **2020** (2020).

[15] Y. Freund and R.E. Schapire, A decision-theoretic generalization of on-line learning and an application to boosting, *Journal of Computer and System Sciences* **55** (1997), 119–139.

[16] K. Garcia and L. Berton, Topic detection and sentiment analysis in twitter content related to covid-19 from brazil and the usa, *Applied Soft Computing* **101** (2021), 107057. doi: https://doi.org/10.1016/j.asoc.2020.107057

[17] M. Ghosh and G. Sanyal, Analysing sentiments based on multi feature combination with supervised learning, *International Journal of Data Mining, Modelling and Management* **11** (2019), 391–416.

[18] R.M. Haralick, The table look-up rule, *Communications in Statistics-Theory and Methods* **5** (1976), 1163–1191.

[19] K. Hou, T. Hou and L. Cai, Public attention about covid-19 on social media: An investigation based on data mining and text analysis, *Personality and Individual Differences* **175** (2021), 110701.

[20] M. Hung, E. Lauren, E.S. Hon, W.C. Birmingham, J. Xu, S. Su, S.D. Hon, J. Park, P. Dang and M.S. Lipsky, Social network analysis of covid-19 sentiments: Application of artificial intelligence, *J Med Internet Res* **22** (2020), e22590. URL: http://www.jmir.org/2020/8/e22590/, doi:10.2196/22590.

[21] V.K. Jain and S. Kumar, Effective surveillance and predictive mapping of mosquito-borne diseases using social media, *Journal of Computational Science* **25** (2018), 406–415.

[22] G. Kyriakides and K. Margaritis, Hands-On Ensemble Learning with Python: Build highly optimized ensemble machine learning models using scikit-learn and Keras, *Packt Publishing* (2019). URL: https://books.google.co.in/books?id=N4mkDwAAQBAJ

[23] D.V. Lindley, Fiducial distributions and bayes' theorem, *Journal of the Royal Statistical Society, Series B (Methodological)* (1958), 102–107.

[24] J. Liu, P. Liu, Z. Zhu, X. Li and G. Xu, Graph convolutional networks with bidirectional attention for aspect-based sentiment classification, *Applied Sciences* **11** (2021), 1528.

[25] M.O. Lwin, J. Lu, A. Sheldenkar, P.J. Schulz, W. Shin, R. Gupta and Y. Yang, Global sentiments surrounding the covid-19 pandemic on twitter: analysis of twitter trends, *JMIR Public Health and Surveillance* **6** (2020), e19447.

[26] A. Malik, S. McFadden, J. Elharake and S. Omer, Determinants of covid-19 vaccine acceptance in the us, *Eclinicalmedicine* **26** (2020), 100495.

[27] K. McKeown, A. Agarwal and F. Biadsy, Contextual phrase-level polarity analysis using lexical affect scoring and syntactic n-grams, (2009).

[28] Z.B. Nezhad and M.A. Deihimi, Twitter sentiment analysis from iran about covid 19 vaccine, *Diabetes & Metabolic Syndrome: Clinical Research & Reviews* **16** (2022), 102367.

[29] W.H. Organization, Global situation report-55, (2020). https://www.who.int/publications/m/item/situation-report–55

[30] M. Pal, G. Berhanu, C. Desalegn and V. Kandi, Severe acute respiratory syndrome coronavirus-2 (sars-cov-2): an update, *Cureus* **12** (2020).

[31] T.R. Patil, Msss performance analysis of naive bayes and j48 classification algorithm for data classification, intl. *Journal of Computer Science and Applications* **6** (2013).

[32] H. Piedrahita-Valdes, D. Piedrahita-Castillo, J. Bermejo-Higuera, P. Guillem-Saiz, J.R. Bermejo-Higuera, J. Guillem-Saiz, J.A. Sicilia-Montalvo and F. Machio-Regidor, Vaccine hesitancy on social media: Sentiment analysis from june 2011 to april 2019, *Vaccines* **9** (2021), 28.

[33] K. Pogue, J.L. Jensen, C.K. Stancil, D.G. Ferguson, S.J. Hughes, E.J. Mello, R. Burgess, B.K. Berges, A. Quaye and B.D. Poole, Influences on attitudes regarding potential covid-19 vaccination in the united states, *Vaccines* **8** (2020), 582.

[34] S. Praveen, R. Ittamalla and G. Deepak, Analyzing the attitude of indian citizens towards covid-19 vaccine–a text analytics study, *Diabetes & Metabolic Syndrome: Clinical Research & Reviews* **15** (2021), 595–599.

[35] J. Samuel, G.G.M.N. Ali, M.M. Rahman, E. Esawi and Y. Samuel, Covid-19 public sentiment insights and machine learning for tweets classification, *Information* **11** (2020). URL: https://www.mdpi.com/2078-2489/11/6/314, doi:10.3390/info11060314.

[36] D. Sarkar and V. Natarajan, Ensemble Machine Learning Cookbook: Over 35 practical recipes to explore ensemble machine learning techniques using Python, *Packt Publishing*, (2019). URL: https://books.google.co.in/books?id=dCWGDwAAQBAJ

[37] F. Shen, X. Zhao, Z. Li, K. Li and Z. Meng, A novel ensemble classification model based on neural networks and a classifier optimisation technique for imbalanced credit risk evaluation, *Physica A: Statistical Mechanics and its Applications* **526** (2019), 121073.

[38] C. Villavicencio, J.J. Macrohon, X.A. Inbaraj, J.H. Jeng and J.G. Hsieh, Twitter sentiment analysis towards covid-19 vaccines in the philippines using naive bayes, *Information* **12** (2021), 204.

[39] D. Vujović, Classification model evaluation metrics, *Int J Adv* **12** (2021), 6.

[40] B. Wang and J. Zhuang, Crisis information distribution on twitter: a content analysis of tweets during hurricane sandy, *Natural Hazards* **89** (2017), 161–181.