Sentiment Analysis of COVID-19 Tweets Using Voting Ensemble-Based Model

Kamal and Amritpal Singh

1 Introduction

Sentiment analysis is the process of communicating and interpreting feelings, views, ideas, and written emotions to enhance the consistency of various ML and DL methods. With the increased use of digital media for users to express themselves and engage with one another, sentiment analysis is an important technique for extracting their opinions, sentiments, and emotions. It has been one of the most active study fields, using numerous DL and ML models to analyze viewpoints and sentiments. Today, Twitter and social networks discussed what is trending in social media, whether locally or internationally.

The Corona Virus Disease 2019 (COVID-19) Disease has spread exponentially all over the world. By May 9, 2021, there will be 22,295,911 confirmed cases across India [1]. This pandemic has sparked widespread public interest in India. Since the coronavirus emerged so quickly in China, learning the public opinion variations will help the government manage and monitor the pandemic's growth, as well as support more scientific and successful work. For eg, is it feasible and appropriate for people to lock down in India because it is the first time this has occurred in India? Sentiment analysis is a valuable tool for quickly obtaining people's insights from vast amounts of text data. From opinion analysis, direct input from the public on government policy on coronavirus can be found, which is critical knowledge for government decision-making for pandemic prevention. However, since India already has 1.36 billion inhabitants, it is difficult to ask anyone how they feel about coronavirus. At

Kamal (✉) · A. Singh
Department of Computer Science & Engineering, Dr. B.R. Ambedkar, National Institute of Technology, Jalandhar, Punjab, India
e-mail: kamalk.is.19@nitj.ac.in; apsingh@nitj.ac.in

R. Tiwari et al. (eds.), *Image Based Computing for Food and Health Analytics: Requirements, Challenges, Solutions and Practices*,
https://doi.org/10.1007/978-3-031-22959-6_9

the moment, the emotion classification methodology allows one to investigate public sentiment heterogeneity.

According to dsayce.com, in May 2020, there would be an average of 6000 tweets in each second, 350,000 tweets in each minute, 500 million tweets in each day, and 200 billion tweets in each year [3].

In the year 2020, blog.twitter.com announced that #COVID 19 is the most famous hashtag this year. It was tweeted nearly 400 million times, and #StayHome was the third most popular hashtag [blog.twitter.com]. They also stated that this year, 7000 tweets about TV and movies are sent every minute [2].

According to the "DSAYCE" website, 6000 tweets are posted every second on average, with 500 million tweets being reported every day. In Japan, Twitter users set the highest world record of "143,199" tweets in each second when watching Hayao Miyazaki's animation classic "Castle in the Sky" on Saturday third August 2018. In 2016, 6000 tweets are received each second, 350,000 tweets are sent in each minute on average, 500 million tweets are received every day, and 200 billion tweets are in every year [2].

This research paper has organized into five sections in the II-section we have reviewed the research papers that have already been done, in section-III has covered the theoretical background of the research paper, in the section-IV, we have given proposed methods that I have used, and finally, in section-V, we have given the conclusion and future scope of this research paper.

2 Literature Review

In this section, we review the research work that has already been published by the researcher in the area of sentiment predictions and opinion mining, as well as numerous methods and strategies for analyzing sentiment; we have chosen the best research papers in the area of sentiment predictions.

In sentiment analysis, ML-based methods are commonly used; these approaches include training and testing datasets. Supervised classifier learning was used to prepare the dataset [14]. Uni-grams, bi-grams, n-grams, parts of speech (POS) codes, and bag-of-words (BOW), for example, are features of ML-based strategies. It employs three models: NB, SVM, and maximal entropy (ME). Machine learning analyses sentiments using supervised, semi-supervised, and unsupervised methods.

Kumar, Sudhanshu, et al. propose [15] that this paper aims to use the feature extraction method to evaluate the impact of user's emotions based on their gender and age, as well as to create a dataset with user reviews and gender and age information, as well as to present user's expression through extensive experiments, and finally to discuss the comparison of dictionary-based and machine learning methods. This paper employs various sentiment analysis strategies on 900 Facebook users with age and gender records, divides the 900 users into four categories, and then applies sentiment analysis to individual groups. With an accuracy of 78%, [15] used Feature Extraction, Dictionary-Based Classifier, Machine Learning-based Classifier, SVM, NB, and Maximum Entropy to evaluate sentiment analysis.

Kabir, Monika, and colleagues suggest [16] In this case, data sets are mined from various forms of feedback. These data sets are of various sizes and domains, and different methods are used to improve precision in the data sets. For sentiment analysis, [16] employs various ML and classification methods, including RF, DT, SVM, Boosting, Bagging, and DT. For sentiment prediction, three popular data sets such as Amazon, Yelp, and IMDb on various fields such as brands, Movies, and restaurants are used [16] employs data sets to categorize and test emotion polarity.

Jagdale et al. suggest [17] that this paper used a dataset from the Amazon website that included six categories of product reviews: laptops, cameras, televisions, mobile phones, tablets, and video surveillance. Analyze and describe each product review using various ML techniques such as SVM and NB. Using NB and SVM, they achieved accuracy on camera 98.17 percent and 93.54 percent, respectively. They have used a mixed approach, which is a mixture of lexicon-based and machine learning techniques. Machine-learning and lexicon-based methods are used to improve precision.

Jain Achin et al. suggest [18], and this paper is mostly concerned with analyzing English tweets. This paper used R studio and the Twitter API to compile tweets with blogging hashtags. This paper evaluates and implements sentiment analysis on "RenewableEnergy" using five different forms of ML techniques on 5000 Twitter tweets (SVM, KNN, NB, AdaBoost, Bagging). Using the CfsSubsetEval function selection process and SVM, they finally achieved 92.96 percent precision.

Saad, Shihab Elbagir, and colleagues suggest [19] Data sets are derived from the Twitter API and categorized as positive or negative tweets [19] performs sentiment analysis in five stages. Step 1: Assume that each tweet is classified as positive or negative. Step 2: Determine each tweet's polarity. Step 3: Determine each class's total emotion polarity. Step 4: assign a score of +2 for high positive, +1 for moderate positive, −1 for moderate negative, and − 2 for high negative to each level [5] analyses emotions in the Twitter data collection using ordinal regression and four ML-based algorithms. These algorithms are used for categorization: SVR (support vector regression), SoftMax, RF (Random Forest), and DT (Decision Trees). For sentiment analysis, SVR and RF have almost identical precision but outperform the Multinomial logistic regression classifier. The Decision tree classifier achieved the best precision of 91.81 percent in this case.

Yadav, Nikhil, and colleagues suggest [20]. This paper makes use of a dataset from Twitter as well as the Kaggle repository. Often employs numerous ML-based algorithms such as NB, RF, DT, SVM, and XGBoost, and then compares multiple classifiers to distinguish tweets on a certain product based on positive or negative sentiments [20] assists the corporate company in developing stronger strategic business strategies for their brands. The dataset is evaluated in five measures in this article. The first step is to compile a list of data sets. Step 2 is the preprocessing step. Step 3 is the feature extraction process. Step 4 is the grouping process. Step 5 is the assessment process. This paper relies mostly on English tweets and dissects human thoughts and emotions.

Chaturvedi, Saumya, et al. suggest [21] that this paper uses some ML-based strategies like NB and SVM for exploring business intelligence emotions that are used for business process decisions, as well as Big-data principles. The decision used the following measures to evaluate the emotions for the market process.

(a) Collection of Data
(b) Preprocessing and cleaning the dataset
(c) Prediction of the sentiments
(d) Classification of the sentiments
(e) Final output

Dataset collected from online services and analyzed using machine learning algorithms for market intelligence and sentiment analysis.

Hemalatha et al. suggest [22] that this paper is used to test emotions using ML tools such as Sentiment140, Twittratr, and Tweet feel, as well as ML techniques such as the Naive Bayes (NB) algorithm, the Maximum Entropy (ME) algorithm, and data-sets gathered from the social networking fields such as Twitter and others. Then, on the data collection, apply maximum Entropy and sentiment analysis methods, and finally, equate the Maximum Entropy algorithm to other sentiment analysis tools. Figure 1 shows the comparative analysis of maximum Entropy and sentiment analysis tools.

The Fig. 1 above explicitly reveals that maximum entropy outperforms sentiment analysis methods in terms of precision.

The [23] use over 4000,00 ratings for sentiment predictions by using three ML algorithms, and this paper aims to improve sentiment analysis accuracy using NB, DT, and SVM. The evaluation was carried out using a ten-fold cross-validation procedure. They performed sentiment analysis on cell phone ratings, sorted them into favorable and negative emotions in almost equal proportions, and then analyzed the dataset using machine learning models. SVM has the highest precision of the three versions, at 81.75 percent.

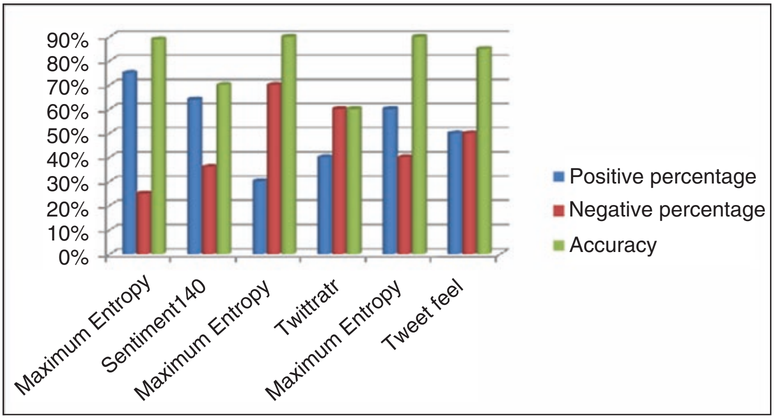


Fig. 1 Comparision of Maximum Entropy and Sentiment Analysis Tools [22]

Many authors use Deep learning models for sentiment analysis, DL is a form of artificial neural network (ANN) that uses several layers of networks to learn tasks. The composition of a neural network works the same as the human physiological brain and its processing units are known as neurons. Neurons can learn to do tasks in the same way as the brain does. Recurrent Neural Networks (RNN) and Feed-forward Neural Networks (FNN) are the two types of neural networks.

Chandra, Yogesh, et al. suggest [24]. A machine learning model, a deep learning model, and polarity-based approaches are used in this article. To determine the percentage of trust in classification systems, tweet data is first obtained from Twitter and then transmitted through a machine learning classifier. The ratio of positive and negative tweets is calculated using polarity-based techniques. Deep learning algorithms are then used to interpret the tweets. The tweets were classified using CNN-RNN, LSTM, and RNN templates. Deep learning models, such as CNN-RNN and LSTM, outperformed machine learning models. This paper has used the Twitter API to retrieve live tweets and identify them using voting-based classification and polarity-based classification. The dataset was obtained from the Twitter API/Kaggle repositories and the Kaggle kernel. Using machine learning classifiers, we achieved an accuracy of 81 percent to 90 percent for training and testing datasets.

Alarifi, Abdulaziz, et al. suggest [25] This paper gathered data from Amazon and employs a greedy approach since it uses the best classifier, a cat swarm optimization-based long short-term memory neural network (CSO-LSTMNN). The CSO-LSTMNN algorithm, the greedy method, and the particle swarm optimization (PSO) algorithm are used to compare results. The CSO-LSTMNN algorithm produces better results in terms of multifaceted nature and proficiency. CSO-LSTMNN employs five polarity levels: negative, relational negative (RN), neutral, positive, and relational positive (RP), with each level's threshold value balanced.

Mohamed Neha et al. suggest [26] and uses four deep learning methods for processing movie reviews datasets. This paper uses a data collection from IMDB that contains 50 k movie ratings. The data set is also split into half positive and half negative data sets. The dataset is first pre-processed with Word2Vec before applying deep learning algorithms such as LSTM, RNN CNN, Multilayer Perceptron (MLP), and a hybrid model that combines CNN and LSTM. Finally, the Hybrid model had an accuracy of 89.2 percent, while CNN, LSTM, and MLP had an accuracy of 87.7 percent, 86.64 percent, and 86.74 percent, respectively.

Alharbi et al. [27] provides sentiment analysis for the Arabic language. Many Arabic datasets are gathered from various sources and then subjected to sentiment analysis and comparison with other languages' datasets. The first small dataset is taken from Opinion Corpus for Arabic (OCA), and it has 500 movie comments. The second dataset is taken from Large Scale Arabic Book Reviews (LABR), and it contains book reviews from consumers on a scale of 5 to 1, and it is taken from www.goodreads.com, and it contains 63,000 Arabic reviews [27] considered three major levels: text, statement, and entity aspect. Find the issue and articulate the dataset with positive or negative sentiments at the document level. They have calculated individual sentences at the sentence level.

Ramadhani et al. suggest [28] for sentiment analysis using deep learning models; they retrieved tweets in both Korean and English from the Twitter API and used three stages for text mining from the Twitter API: information acquisition, information extraction, and data mining. Finally, they ran data into two DL models, Multilayer Perceptron (MLP) and Deep Neural Network (DNN), and obtained 67.45 percent and 77.45 percent accuracy on the training dataset, respectively.

Heikal et al. propose [29] for sentiment prediction of the Arabic dataset, taking 10,000 tweets from the Arabic social sentiment prediction dataset (ASTD). They used an ensemble-based CNN approach and the LSTM deep learning model. First, they used an individual variant on the sample and obtained 64.30 percent and 64.75 percent accuracy using CNN and LSTM, respectively. Eventually, they used an ensemble-based approach using a soft voting classifier and obtained 65.05 percent accuracy.

Wang, Jian, et al. recommend [30] for Chinese online shopping product reviews. They collected 13,000 votes, 7000 of which were positive and the rest were negative. They used three models for sentiment analysis: LSTM, LSTM dependent on Nadam with L2 regularisation, and Bi-directional LSTM. After reviewing the performance of both models, they discovered that the LSTM based on Nadam with L2 had greater accuracy than the other two models, which provided about 98 percent accuracy.

3 Theoretical Background

ML: For sentiment predictions, we have taken many machine learning techniques SVM, RF, DT, K-Nearest Neighbor (KNN), Logistic Regression (LR), and ensemble voting classifier.

SVM: Support vector machines are supervised machine learning algorithms that are both efficient and scalable. They are used for classification and regression. However, they are most often found in classification problems. SVMs were initially developed in the 1960s, but they were refined in 1990. As opposed to other machine learning algorithms, SVMs have a distinct implementation process. Because of their ability to manage several continuous and categorical variables, they are highly common [4].

An SVM model is essentially a representation of various groups in a hyperplane in a multidimensional space. SVM creates a hyperplane iteratively to minimize error. SVM aims to classify datasets to find the maximal marginal hyperplane (MMH) [4].

Support Vectors: Support vectors are data points that are in marginal lines which are parallel to the hyperplane. These data points are mainly used for separating the data sets into classes.

Hyperplane: As seen in the above Fig. 2, it is a decision plane or space that divides the dataset into different classes.

Fig. 2 SVM Working [4]

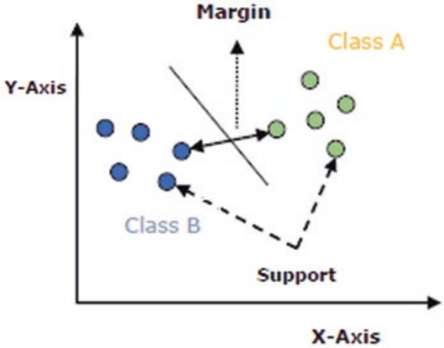
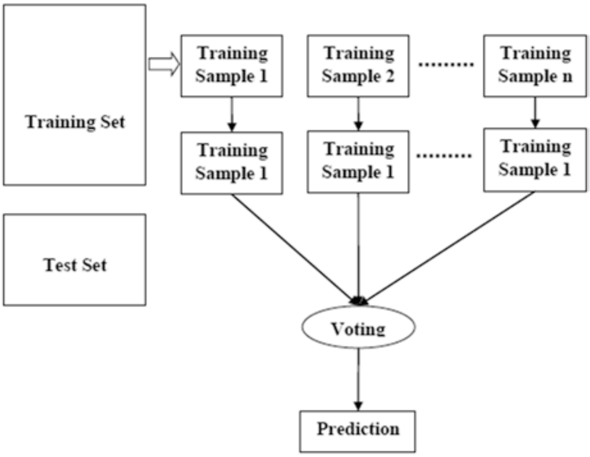


Fig. 3 Working of RF [6]



Margin: It can be described as the distance of the hyperplane and marginal plane that separates groups of closet data points. It is computed as the distance between the hyperplane and the marginal line. A maximum margin is regarded as a positive margin, whereas a small margin is regarded as a poor margin.

RANDOM FOREST: Tin Kam Ho prepared a random forests decision algorithm in 1995, with the help of the random subspace technique, in which Ho described how to apply Eugene Kleinberg's "stochastic discrimination" approach to classification.

RF, also known as random decision forests, is an ensemble-based technique for classification, regression, and other tasks that works like training a large number of decision trees and then outputting the class that is the mode of the classes (classification) or mean/average estimation (regression) of the individual trees [5] (Fig. 3).

Random Forest Algorithm took the following four steps to classify the dataset.

(a) Selection of Random sample
(b) Construct a decision tree and predict the result

(c) Perform ensemble technique
(d) Select the best result as the final prediction

DECISION TREE: these are statistical modeling techniques used in ML and DT for analytics and data processing. It employs a DT to progress from assumptions about the item (represented by nodes) to judgments about the items and target value (represented in the leaf nodes). Classification trees are tree structures in which the target nodes have a distinct range of values [7] (Fig. 4).

Root Node: The decision tree begins at the root node. It contains the whole dataset, which is then split into two or more homogeneous sets.
Leaf Node: Leaf nodes are the tree's last output node, and the tree cannot be further separated until obtaining a leaf node.

The decision tree uses three main methods to make a decision which are the Entropy, Gini index, and Information gain.

ENTROPY (E): It is used to assess a dataset's impurity or randomness.

$$\mathbf{E}(\mathbf{X}) = \mathbf{I}_{\mathbf{E}}(\mathbf{P1},\mathbf{P2},\ldots\ldots,\mathbf{Pj})$$

$$\mathbf{E}(\mathbf{X}) = -\sum_{\mathbf{i}=\mathbf{1}}^{\mathbf{j}} \mathbf{p}_{\mathbf{i}}\ {}_{\log_2 \mathbf{P}_{\mathbf{i}}.}$$

INFORMATION GAIN (IG): To identify the best function that can be used as a root node in terms of information gain.

$$\mathbf{IG} = \mathbf{E}(\mathbf{X}) - \mathbf{E}(\mathbf{X}|\mathbf{a})$$

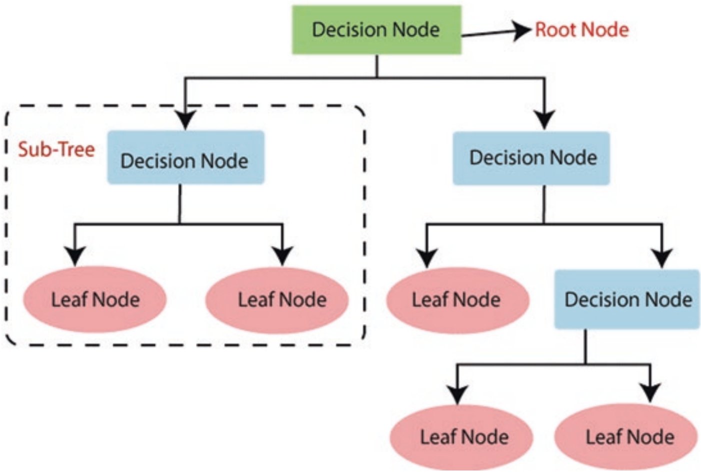


Fig. 4 Working of DT [8]

$$\mathbf{IG} = -\sum_{i=1}^{j} p_i \log_2 P_i - \sum_{i=1}^{j} p_{(i|a)} \log_2 P_{(i|a)}$$

Where E(X) Entropy of X and E(X|a) Sum of all entropy children P is the probability.

GINI INDEX (GI): It is calculated by subtracting the total of every class's squared percentages from one. It always considers the largest partitions that are simple to enforce, whereas information gain always considers smaller partitions that have distinct values.

KNN: k-NN is a non-parametric classification technique introduced in 1951 by Evelyn Fix and Joseph Hodges and then enhanced by Thomas Cover. It is used in both classification and regression problems. In all circumstances, the response consists of the k closest training instances in the data set [9]. KNN predicts the new data points based on 'feature similarities,' and the new data point will predict a value based on how exactly it resembles the points in the training collection [10] (Fig. 5).

KNN uses the many steps to classify the datasets which are.

(a) 'K' number of neighbors selection.
(b) Euclidean distance calculation for 'K'.
(c) Taken k as per the Euclidean distance
(d) Count the number of data points in each group among these k neighbors.
(e) Assign the latest data points to the group with the highest number of neighbors.

LR: it is a classification technique that uses supervised learning to predict the likelihood of a target value. Because the target variable is dichotomous, means there are only two possible classes [12].

An LR model forecasts P(Y = 1) as a function of X mathematically. It is a very basic ML technique that can be used to solve various classification problems such as spam diagnosis, diabetes prediction, cancer detection, Sentiment analysis, and so

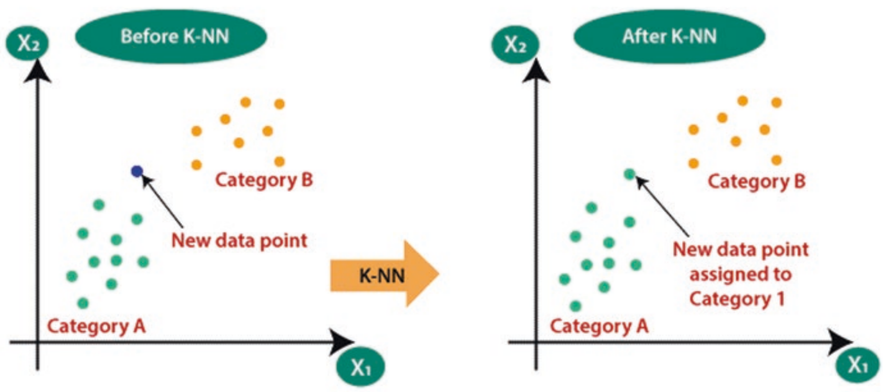
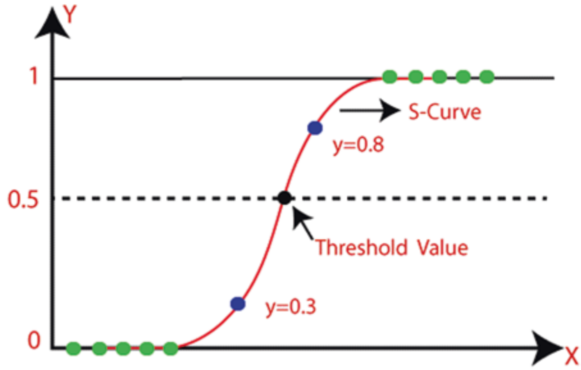


Fig. 5 working of KNN [11]

Fig. 6 working of LR [13]



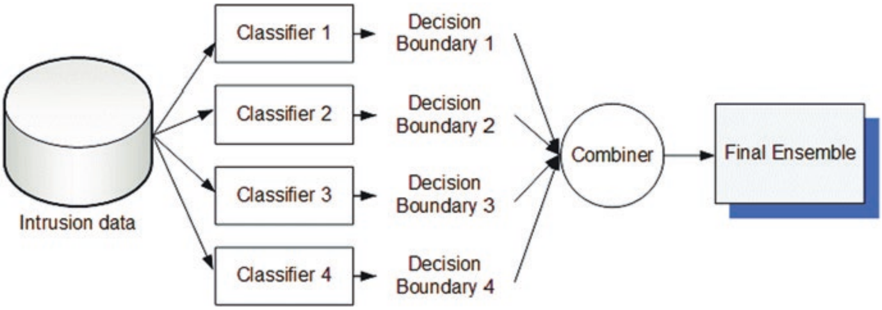


Fig. 7 Architecture of voting classifier [33]

on [12]. We fit an "S" form logistic function that predicts two maximum values in LR (0 or 1) (Fig. 6).

The logistic feature curve shows the probability of anything like whether the cells are cancerous or not, when a mouse is a fat or not depending on its weight, and so on [13].

ENSEMBLE VOTING CLASSIFIER: Ensemble approaches are strategies for developing different models and then combining them to achieve better performance. Ensemble models usually yield more reliable results than a single model [31]. A Voting Classifier is an ML technique that trains on various ML and DL models and predicts an outcome (class) based on the highest likelihood of the taken class as the output [32].

It essentially merges the results of each classifier that passed into Voting Classifier and then predicts the output class depending on the voting majority. Instead of developing individual dedicated models and determining their precision, we can develop a single model that trains on these models and predicts performance based on their cumulative plurality of voting for every output class [32] (Fig. 7).

The voting classifier has two methods to ensemble the models.

Hard Voting Classifier: The predicted performance class in hard voting is the class with the maximum majority of votes, i.e. the class with the high probability of being analyzed by any of the classifiers. Assume that we have three classifiers that have their output classes (A, A, B), and the majority predicted A. As a result, A will be the final prediction [32].

Soft Voting Classifier: The performance class of soft voting is the prediction technique that depends on the average of the probabilities assigned to that class. Assume that provided any input to three models, the estimation chances for each class A = (0.30, 0.47, 0.53) and B = (0.30, 0.47, 0.53). (0.20, 0.32, 0.40). So, if the average for class A is 0.4333 and the average for class B is 0.3067, class A is the winner since it has the highest likelihood averaged by each classifier [32].

PIPELINING: Pipelines are used to divide the machine learning workflows into separate, interchangeable, scalable sections that can then be pipelined together to continually increase the model's accuracy and produce a reliable algorithm. Generally pipelining includes the following steps.

Preprocessing and Vocab building: Unwanted texts (stop words), punctuation, URLs, handles, and so on that have little sentimental meaning are removed. Then adding unique preprocessed words to a vocabulary.

Feature Extraction: Iteratively going through and data illustration to remove features using a frequency dictionary and then creating a function matrix.

4 Proposed Method

In this part, we will go over the methodology we used in this paper. We build a Voting ensemble of ML models for the sentiment prediction of Covid-19 tweets. In the Voting ensemble-based model, we have included various ML classifiers for classifying the dataset into positive as well as negative tweets. It is divided into four major stages. The first stage is data preprocessing, which includes word sorting and tokenization, as well as stop words elimination and cleaning the dataset. The second stage is pipelining, in which we have passed two methods first one is TFIDF vectorizer and the second one is a machine learning classifier. TFIDF vectorizer and classifier model are used in the third stage for making a single model. In the fourth stage, we have applied a hard voting classifier (Fig. 8).

Figure 1 shows the working of the ensemble-based technique, in the preprocessing stage of the technique we have used the following steps.

Remove special characters: in this step, we have removed special characters from the tweets like _,? $, etc.

Remove Accented characters: In this, we have removed accented characters from the tweets like â, î or ô, etc.

Remove Emails: In this step, we have removed emails from the tweets.

Remove html tags: In this steps we have removed html tags like <html>, </html>, etc.

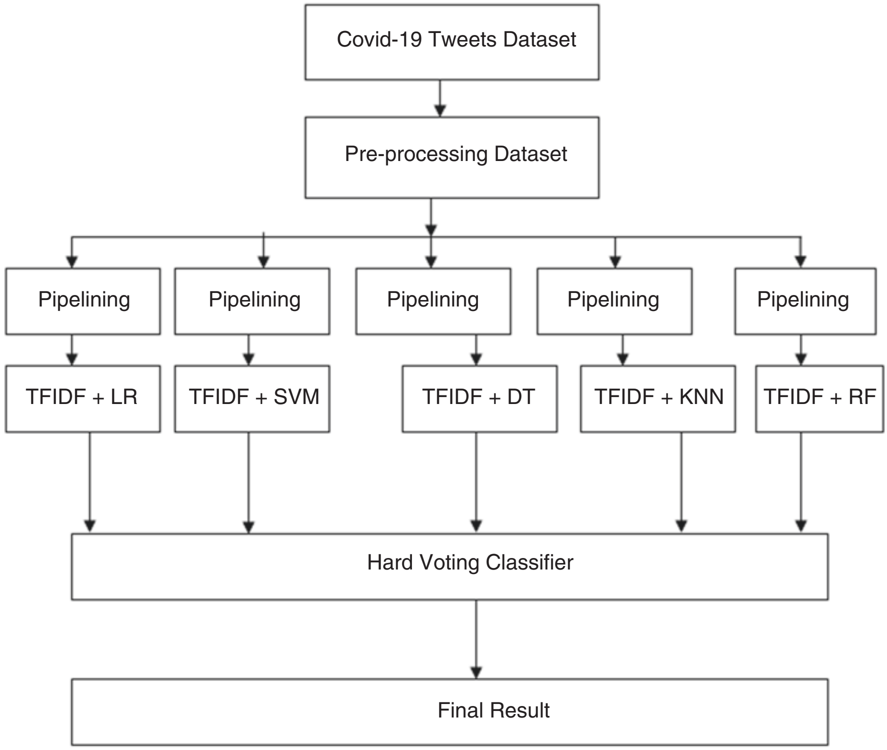


Fig. 8 Working of the ensemble-based model

Remove URLs: In this step, we have removed all URLs from the tweets.
Lower case: In this step, we have converted all upper case tweets into lower case.

After preprocessing we have used pipelining in that we have passed TFIDF vectorizer and machine learning classifier. Then prepare a single pipeline model. Finally, we have compared all models with their accuracy.

After preparing all models, we used an ensemble voting classifier to make average accuracy. By using the Hard voting classifier we got 94.7% accuracy on the testing dataset.

5 Experiment and Results

For the experiment, we have used Covid-19 tweets which have been downloaded from the Kaggle repository that has 1.4 million tweets related to covid-19. Dataset has four types of sentiments which are positive, extremely positive, negative, and extremely negative. For the experiment purpose, we have converted extremely positive as positive and extremely negative as negative. Then by using an individual

Table 1 Individual model accuracy

ML - Model	Accuracy
Decision tree	81.11%
CNN	72.33%
SVM	81.10%
Logistic regression	76.81%
Random Forest	90%

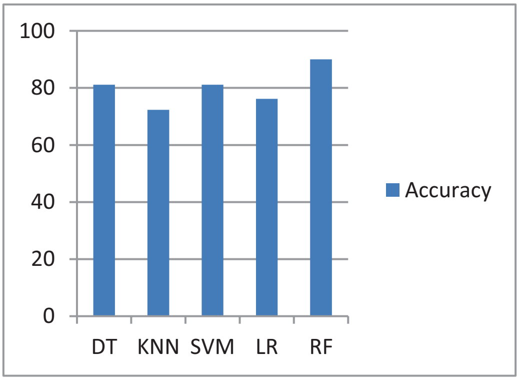


Fig. 9 Comparative analysis of classifiers

Table 2 Pipeline model accuracy

Pipeline model	Accuracy
TFIDF + LR	89%
TFIDF + SVM	95%
TFIDF + DT	91%
TFIDF + KNN	77%
TFIDF + RF	94%
Adaboost + DT	87%

machine learning classifier we got the following results (Table 1) by using Random Forest ML – classifier we got 90% accuracy which is better than other classifiers. Figure 9 shows the comparative analysis of individual models.

After calculating the individual accuracy we got that some classifiers perform well and some are not performing well so we have used the concept of pipelining with the help of TFIDF vectorizer. In the pipeline, we have passed one ML-classifier with TFIDF vectorizer and prepared a model, After preparing all models we have compared their results using Table 2 which shows the accuracy of each model.

After preparing models using the concept of pipeline techniques, we got that some models perform well and some are performing poorly so to overcome this problem we have used an ensemble model for the average accuracy. Figure 10 shows the comparative analysis of pipelining models.

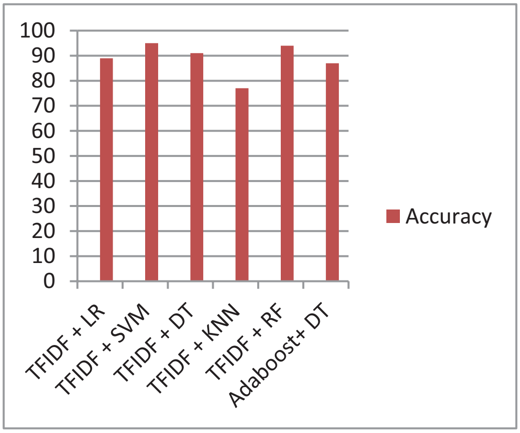


Fig. 10 Comparative analysis of pipelining models

Figure 10 shows that the pipeline model using TFIDF + SVM has the highest accuracy and TFIDF + KNN has the worst accuracy compared to other models.

To overcome the above problem we have used the concept of ensemble model using a Hard Voting Classifier. Finally, we have passed all models in the voting machine and got an accuracy of 94.70% which is better than other classifiers and pipeline models.

6 Conclusion and Future Scope

In this research paper, we introduced an ensemble-based model for sentiment prediction. This paper aims to support the government in this pandemic situation for taking the right decision using public opinions, it will also be helpful for the researcher to research in the area of sentiment predictions. This paper uses the concept of pipelining in which we have passed a machine learning classifier with a TFIDF vectorizer.

This research paper implementation part is divided into four levels, The first level is data preprocessing, which involves word sorting and tokenization, as well as stop words removal and dataset cleaning. The second level is pipelining, by which we have passed two methods: a TFIDF vectorizer and a machine learning classifier. In the third level, a TFIDF vectorizer and classifier model was used to create a single model. We used a hard voting classifier in the fourth level. We have used covid-19 tweets which have been downloaded from the Kaggle. We have also calculated the accuracy of the individual model and then finally by using hard voting we got an accuracy of 94.70% which is better than all models.

For future studies, we have observed that if we train our model using the BERT model and some neural network models, we can improve accuracy. I have also found that no work has used ensemble-based models using BERT and Neural networks.

References

1. Worldometers.info. (2021). India COVID: 22,295,911 Cases and 242,398 Deaths - Worldometer. [online] Available at: <https://www.worldometers.info/coronavirus/country/india/>
2. Blog.twitter.com. (2021). Spending 2020 Together on Twitter. [online] Available at: <https://blog.twitter.com/en_us/topics/insights/2020/spending-2020-together-on-twitter.html>
3. Sayce, D., & Ltd., P., (2021). The number of tweets per day in 2020 | David Sayce. [online] David Sayce. Available at: <https://www.dsayce.com/social-media/tweets-day/>
4. "ML - Support Vector Machine(SVM)," Tutorialspoint.com. [Online]. Available: https://www.tutorialspoint.com/machine_learning_with_python/machine_learning_with_python_classification_algorithms_support_vector_machine.htm
5. Wikipedia contributors, "Random forest," Wikipedia, The Free Encyclopedia, 06-May-2021. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Random_forest&oldid=1021839899
6. "Classification Algorithms - Random Forest," Tutorialspoint.com. [Online]. Available: https://www.tutorialspoint.com/machine_learning_with_python/machine_learning_with_python_classification_algorithms_random_forest.htm
7. Wikipedia contributors, "Decision tree learning," Wikipedia, The Free Encyclopedia, 06-May-2021. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Decision_tree_learning&oldid=1021669314
8. "Decision tree classification algorithm," Javatpoint.com. [Online]. Available: https://www.javatpoint.com/machine-learning-decision-tree-classification-algorithm
9. Wikipedia contributors, "K-nearest neighbors algorithm," Wikipedia, The Free Encyclopedia, 06-May-2021. [Online]. Available: https://en.wikipedia.org/w/index.php?title=K-nearest_neighbors_algorithm&oldid=1021698330
10. "KNN algorithm - Finding Nearest Neighbors," Tutorialspoint.com. [Online]. Available: https://www.tutorialspoint.com/machine_learning_with_python/machine_learning_with_python_knn_algorithm_finding_nearest_neighbors.htm
11. "K-Nearest Neighbor(KNN) Algorithm for Machine Learning," Javatpoint.com. [Online]. Available: https://www.javatpoint.com/k-nearest-neighbor-algorithm-for-machine-learning
12. "Machine learning - Logistic Regression," Tutorialspoint.com. [Online]. Available: https://www.tutorialspoint.com/machine_learning_with_python/machine_learning_with_python_classification_algorithms_logistic_regression.htm
13. "Logistic Regression in Machine Learning," Javatpoint.com. [Online]. Available: https://www.javatpoint.com/logistic-regression-in-machine-learning
14. Jain, A. P., & Dandannavar, P. (2016, July). Application of machine learning techniques to sentiment analysis. In *2016 2nd international conference on applied and theoretical computing and communication technology (iCATccT)* (pp. 628–632). IEEE.
15. Kumar, S., Gahalawat, M., Roy, P. P., Dogra, D. P., & Kim, B. G. (2020). Exploring the impact of age and gender on sentiment analysis using machine learning. *Electronics, 9*(2), 374.
16. Kabir, M., Kabir, M. M. J., Xu, S., & Badhon, B. (2019). Empirical research on sentiment analysis using machine learning approaches. *Int J Comput Appl*, 1–9.
17. Jagdale, R. S., Shirsat, V. S., & Deshmukh, S. N. (2019). Sentiment analysis on product reviews using machine learning techniques. In *Cognitive informatics and soft computing* (pp. 639–647). Springer.
18. Jain, A., & Jain, V. (2019). Sentiment classification of twitter data belonging to renewable energy using machine learning. *Journal of Information and Optimization Sciences, 40*(2), 521–533.
19. Saad, S. E., & Yang, J. (2019). Twitter sentiment analysis based on ordinal regression. *IEEE Access, 7*, 163677–163685.
20. Yadav, N., Kudale, O., Rao, A., Gupta, S., & Shitole, A. (2021). Twitter sentiment analysis using supervised machine learning. In *Intelligent data communication technologies and internet of things: Proceedings of ICICI 2020* (pp. 631–642). Springer.

21. Chaturvedi, S., Mishra, V., & Mishra, N. (2017, September). Sentiment analysis using machine learning for business intelligence. In *2017 IEEE international conference on power, control, signals, and instrumentation engineering (ICPCSI)* (pp. 2162–2166). IEEE.
22. Hemalatha, I., Varma, G. S., & Govardhan, A. (2013). Sentiment analysis tool using machine learning algorithms. *International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), 2*(2), 105–109.
23. Singla, Z., Randhawa, S., & Jain, S. (2017). Sentiment analysis of customer product reviews using machine learning. In *2017 international conference on intelligent computing and control (I2C2)* (pp. 1–5). IEEE.
24. Chandra, Y., & Jana, A. (2020, March). Sentiment analysis using machine learning and deep learning. In *2020 7th international conference on computing for sustainable global development (INDIACom)* (pp. 1–4). IEEE.
25. Alarifi, A., Tolba, A., Al-Makhadmeh, Z., & Said, W. (2020). A big data approach sentiment analysis using greedy feature selection with cat swarm optimization-based long short-term memory neural networks. *The Journal of Supercomputing, 76*(6), 4414–4429.
26. Ali, N. M., Abd El Hamid, M. M., & Youssif, A. (2019). Sentiment analysis for movie reviews dataset using deep learning models. *International Journal of Data Mining & Knowledge Management Process (IJDKP), 9*.
27. Alharbi, A., Taileb, M., & Kalkatawi, M. (2021). Deep learning in Arabic sentiment analysis: an overview. *Journal of (Information Science), 471*, 129–140.
28. Ramadhani, A. M., & Goo, H. S. (2017, August). Twitter sentiment analysis using deep learning methods. In *2017 7th international annual engineering seminar (InAES)* (pp. 1–4). IEEE.
29. Heikal, M., Torki, M., & El-Makky, N. (2018). Sentiment analysis of Arabic tweets using deep learning. *Procedia Computer Science, 142*, 114–122.
30. Wang, J., & Cao, Z. (2017, October). Chinese text sentiment analysis using LSTM network based on L2 and Nadam. In *2017 IEEE 17th international conference on communication technology (ICCT)* (pp. 1891–1895). IEEE.
31. Demir, N. "Ensemble methods: Elegant techniques to produce improved machine learning results," Toptal.com, 04-Feb-2016. [Online]. Available: https://www.toptal.com/machine-learning/ensemble-methods-machine-learning
32. "ML," Geeksforgeeks.org, 23-Nov-2019. [Online]. Available: https://www.geeksforgeeks.org/ml-voting-classifier-using-sklearn/
33. Adhi, T. B., & Rhee, K. H. (2017). An extensive empirical evaluation of classifier ensembles for the intrusion detection task. *The Computer Systems Science and Engineering, 32*, 149–158.