# COVID-19 Outbreak: Tweet based Analysis and Visualization towards the Influence of Coronavirus in the World

**Muthusami R[a] *, Bharathi A[b], Saritha K[c]**

[a]Department of Computer Applications, Dr. Mahalingam College of Engineering and Technology, Anna University, Tamilnadu, India, r.muthusami@gmail.com, ORCID id: 0000-0001-7322-8653.
[b]Department of Information Technology, Bannari Amman Institute of Technology, Anna University, Tamilnadu, India.
[c]Department of Mathematics, P.A. College of Engineering and Technology, Anna University, Tamilnadu, India.
*Corresponding author

**Abstract.** A novel coronavirus, SARS-CoV-2, is known as the coronavirus disease pathogen for 2019 (COVID-19). The COVID-19 epidemic has spread across the world and has become an international public health emergency. This paper aims to analyze and visualize the influence of coronavirus (COVID-19) in the world by executing such algorithms and methods of machine learning in sentiment analysis on the tweet dataset to understand very positive and very negative opinions of the public around the world. This reveals that Naive Bayes 'machine learning approach has been produced better execution, and it has been regarded as the basis for basic learning. This also brings out another ensemble technique that uses sentiment score as the input function for the classifiers in machine learning, SVM, Max Entropy, Decision Tree, Boosting, and Random Forest. As a result, the LogitBoost, a blended approach, performed better with accuracy of 74%.

Keywords: COVID-19; Text mining; Sentiment analysis; Machine learning algorithms; Tweets

## 1. Introduction

The advent of person-to-person contact administrations (Social Network Service, SNS) has the paved way for individuals to connect and send messages in a fast and casual manner to their peers, colleagues, and the general public on different issues [1-3]. They always do it, as the cost of transmitting these short messages is small. Furthermore, for example, SNS also implemented rich data about its users, posting user-created text, user logs, and user associations to users. Individuals share their views, suggestions, events, videos, applications, and activities using SNS. With more than five hundred million diverse users on Twitter, SNS has turned out to be a notable communication and data search tool, and Twitter is one of the most popular social media platforms today. The tweet is a message limited to 140 characters (now 240 characters) that elevates users to post information and messages about different interests while traditional blogging appears to take more dedicated time to write a post [4]. Most work has been done on text interpretation and interpretation of emotions, but significant research has been done on Twitter sentiment analysis of products, reviews, breaking news, stance detection and so on [5-7].

Initially, on December 31, 2019, China announced a cluster of cases of unknown cause of pneumonia that would later be identified as extreme acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [8-10]. Patients with the disease, known as coronavirus disease 2019 (Covid-19), frequently experience fever, cough and shortness of breath within 2 to 14 days of exposure [11]. There were 750,890 confirmed cases of Covid-19 reported globally as of 31 March 2020, and 36,405 deaths reported [12]. In acknowledgment of Covid-19's widespread global dissemination, the World Health Organization on 11 March 2020 declared Covid-19 a pandemic [13].

Globally, many people are frightened, frustrated, unsure, and distrustful of their national leadership. Yet amid these bleak opinions there have emerged signs of unity. Health workers have shown a remarkable dedication to their communities, acting with compassion and determination to tackle the virus despite daunting and often hazardous conditions. Neighbors have mobilized to help needy people; companies and national governments have stepped up to provide funding and improve social security and health care to those who need it.

In addition to the threat of a serious and sometimes lethal viral infection, the COVID-19 pandemic is impacting numerous organisations across the globe, such as banking, transport, health care, education, media and other industries as well as most of the countries. In an attempt to minimize the number of people infected and slow the infection rate, several countries have implemented lockdown, social distancing, and self-isolation.

This work was performed on Twitter tweets using sentiment classification methods with opinion lexicon dictionary and machine learning algorithms to analyze and visualize the effect of coronavirus around the globe on dataset tweets that were crawled via Twitter API and Google TAGS between March 26$^{th}$ 2020 and March 31$^{st}$ 2020. The Naive Bayes classifier was used as a base classifier and the five well-known classification strategies are proposed against the current ensemble method, which integrates five individual algorithms into an ensemble-based classification decision-making framework, namely Support Vector Machine (SVM) classifier, MaxEntropy classifier, Decision Tree classifier, Bagging classifier and Random Forest classifier. The classification of the ensemble takes into account the classification outcomes of the five classifiers and the use of the technique Majority-Vote to determine the final goal of the class of opinion. This paper also addresses a detailed exploratory analysis of the five different methods for classifying emotions and critical remarks.

The paper is arranged further as given. The inspirations are explained in section 1, and the intent is addressed. Section 2 offers a short description and review of the important research. Section 3 demonstrates the collection of data and the design of experiments. In section 4, the findings of the assessment are presented along with remarks about the analysis. Section 5 presents the conclusions, which summarize the experiment results and address possible scopes of study.

## 2. Related Work

### 2.1. Sentiment Analysis

Sentiment classification strategies can generally be partitioned into machine learning approach, lexicon based approach and hybrid approach. Machine learning approach (ML) applies the popular ML algorithms and utilizations of phonetic features [14-15]. The lexicon-based approach depends on a sentiment lexicon, a gathering of known and precompiled sentiment terms. It is partitioned into dictionary-based approach and corpus-based approach that utilize measurable or semantic strategies to discover sentiment polarity. The hybrid Approach combines both the approaches and is exceptionally basic with sentiment lexicons assuming a key part in the majority of the methods. The principle of the two techniques of sentiment analysis, lexicon-based strategy (unsupervised approach) and machine learning based technique (supervised approach) is that both depend on the bag-of-words. In the machine learning's supervised method, the classifiers utilize the unigrams or their combinations (N-grams) as features. In the lexicon-based method, the unigrams which are found in the lexicon are dispense to a polarity score and the overall polarity score of the content is then computed as aggregate of the polarities of the unigrams. While many researchers have concentrated on finding the best features, a few endeavors have been made to investigate new strategies for sentiment classification. Wang, et al.[16] have evaluated the performance of ensemble strategies (Bagging, Boosting, Random subspace) and proved that the ensemble models can deliver preferable outcomes

over the base learners. Fersini, et al.[17] have proposed to utilize Bayesian model averaging ensemble technique which outflanked both the conventional classification and the ensemble strategies. Carvalho, et al.[18] have utilized genetic algorithms to discover subsets of words from a set of paradigm words that prompted change of classification precision.

### 2.2. Machine learning approach

A Machine Learning Approach for text classification is a supervised algorithm that analyzes data that was already labelled as positive, negative or neutral, extracts features that model the contrasts between the various classes and construes a function that can be utilized for classifying new cases that are concealed recently [19-21]. The procedure of machine learning and text classification can be broken into the following: Data pre-processing, feature generation, feature selection, learning an algorithm and model evaluation. Before applying any of the sentiment classification technique, it is a typical practice to perform data pre-handling. Data pre-handling permits creating a high quality text classification and reducing the computational complexity. Common pre-preparing method incorporates the Parts-of-Speech tagging (POS), stemming and lemmatisation, stop-words removal, and tokenization. Features are content traits that are valuable for catching certain patterns in information. The most well known feature utilized as part of machine learning classification is the nearness or the recurrence of n-grams extricated amid the pre-processing step. In situations where content length fluctuates incredibly, it may be imperative to utilize term frequency (TF) and inverse document frequency (IDF) measures.

In short messages such as tweets, however, words will probably not rehash within one instance, making the double measure of presence as instructive as the counts. Sentiment score computation is the vast process of feature creation. Sentiment computation computes the opinion of a given text from the polarity of the words or phrases present in that text. For this strategy to work, a dictionary of words is alloted to them and polarity is required. The current dictionaries include: Opinion Lexicon [22], SentiWordNet [23], AFINN Lexicon [24], Loughran McDonald Lexicon and NRC- Hashtag [25]. The opinion score of the text can be processed as the normal of the polarities passed on by each of the words in the text. The general thought is to calculate a sentiment score for every tweet, So it would be known how positive or negative the posted message is.

There are diverse approaches to calculate such scores, and one can even make one's own specific equation. A extremely forward yet helpful linear approach has been utilised to characterize score equation suggested by Hu and Liu (KDD-2004). The orientation of a sentiment sentence can be predicted now, i.e., positive or negative. By and large, the overwhelming orientation of the sentiment words are utilized in the sentence to determine the orientation of the sentence. Feature selection is the way toward recognizing a subset of features that have the most elevated prescient power. This progression is urgent for the classification procedure, since disposal of unessential and excess features permits to lessen the span of feature space by expanding the speed of the calculation. It also abstains from over fitting thus contributing to the improved quality of classification.

### 2.3. Machine Learning algorithms

The most well known machine learning algorithms for text classification are Naive Bayes, Support Vector Machines (SVMs), Maximum Entropy (MaxEnt) and Decision Trees [16, 26-27]. As far as the individual classes, some research about classified messages just to classify them as positive or negative, have expected every text to convey an opinion. The Naïve Bayes classifier has been considered as base classifier in this work which has described below.

*2.3.1. Naive Bayes Classifier*

In the BOW (Bag-Of-Words) framework, a document x is represented by $[w_1, \ldots, w_m]$ where $w_k$ denotes the $k^{th}$ word appearing in the document. Naıve Bayes assumes that words are mutually independent. Under this assumption, the conditional probabilities can be simplified as, Eqn. 1.,

$$P(x|y_j) = P([w_1, \ldots, w_m]|y_j) \approx \prod_{k=1}^{m} P(wk|yj) \quad (1)$$

The naive Bayes decision can be described as Eqn. 2.,

$$w^* = \arg\max_{j=1,\ldots,c} \prod_{k=1}^{m} P(wk|yj)\, P(y_j) \quad (2)$$

The probabilities $P(w_k | y_j)$ and $P(y_j)$ can simply be estimated by maximum likelihood. Moreover, Laplace smoothing is necessary in order to prevent infrequently occurring words from being zero probabilities.

*2.3.2. Ensemble learning*

Ensemble learning is chosen over individual classifiers as a result of accuracy where a more reliable mapping can be done by joining the yield of multiple classifiers and efficiency in which an intricate issue can be disintegrated into various sub issues with the goal that it gets to be distinctly simpler to comprehend and tackle (as presented in Fig. 1).

Let $H_m$ :X →{−1,+1} be the *m*-th weak binary classifier (for *m*=1, ...,*M*), and **x** ϵ X some input pattern to be classified, there are many ways to combine the outputs $H_1(\mathbf{x})$, ..., $H_M(\mathbf{x})$ into a single class prediction. For example, assuming that the classifiers independently of each other, a majority vote combination should yield a lower probability of error than any of the individual classifiers. Considering a weighted linear combination of the outputs of the weak classifiers, the ensemble prediction function $H$ : X → {−1,+1} is given by

$H(\mathbf{x}) = \text{sign}\,(\sum_{m=1}^{M} \alpha_m H_m(x))$ ,

where $\alpha_1$, ..., $\alpha_M$ is a set of weights (a simple majority vote results if all the weights are equal).

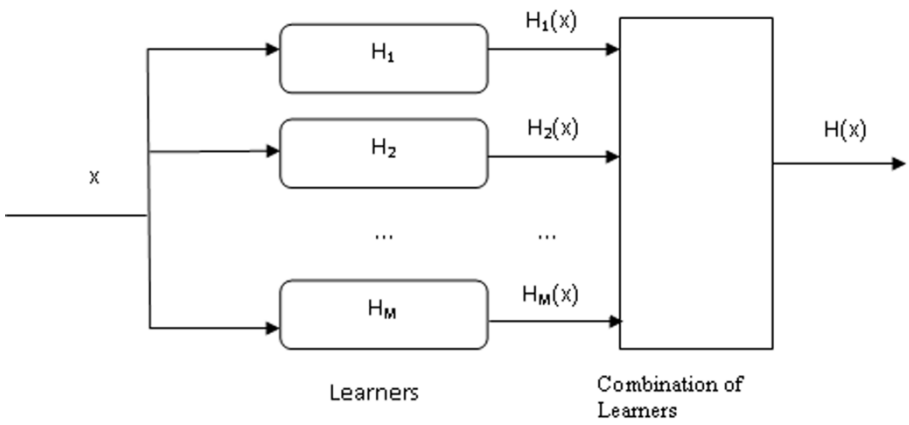


Fig. 1. The overall diagram of ensemble of classifiers

The output of the weak learners $H_m(\mathbf{x})$ with $m \in \{1, \ldots, M\}$ are combined to produce the output of the ensemble of classifiers given by $H(\mathbf{x})$. There are principally three conventional ensemble strategies to be specific and they are bagging, boosting and random subspace. Bagging is an independent learning while boosting is a model guided instance selection dependent

learning. The most surely understood ensemble learning algorithms are talked about in [16, 27-29].

## 3. Dataset and Experimental setting

### 3.1. *Data Preparation*

A Twitter API application was utilized to force tweets from public timeline in real-time. A dataset was made utilizing twitter tweets from a topic that was overwhelming at the season of data gathering; A hashtag (#) is a Twitter tradition used to streamline inquiry, indexing and pattern disclosure. Users incorporate exceptionally designed terms that begin with # into the body of every post. For instance a post about COVID-19 may contain the term #COVID-19. By regarding each hashtag as a label connected just to the post that contain it, sentiment analyzes which words are best connected with each hashtag. A sentence level sentiment analysis was performed on tweets, as many tweets were formed of slang words and incorrect spellings and this was done in three stages. In the principal period of a sentence level sentiment analysis, prehandling was completed. A feature vector was also made by utilizing significant features. A publically accessible sentiment lexicon which comprises of around 6800 words in a list of positive and negative sentiment words or opinion words for English was utilized to isolate the tweets. This list was gathered over many years by Liu and Hu (2004). The tweets are classified into positive and negative classes utilizing diverse classifiers. The final sentiment depends on the quantity of tweets in every class utilizing few sentiment analysis strategies; the bag-of-words approach utilizes accessible lexical assets as observed in sentiment analysis. Machine learning methodologies have been utilized where the tweets dataset has been part of training and testing.

### 3.2. *Dataset: #COVID-19*

Since standard twitter dataset has not been accessible for analysis, another dataset has been made by gathering tweets that have been sent during 26$^{th}$ to 31$^{st}$ March 2020. Tweets have been gathered consequently utilizing Twitter API and Google TAGS with topic specific hastag #COVID-19 in which 18,216 tweets have been gathered from Twitter.

### 3.3. *Experimental work*

This section deals with an experiment account on Twitter dataset implementing the machine learning approach. Analysis of sentiment in Twitter tweets, adding the subassignment given a message to choose from, whether the tweet given was positive, negative or neutral. The more grounded one is to be selected for messages that express both a positive and a negative feeling. To tokenize the corpus and return the sparse matrix was done the following stage. Research of word recurrence was carried out from here. The data has been split into a training dataset and test dataset for Naive Bayes, so that the classifier can evaluate the data. Sentiment scores have been derived by counting positive and negative messages. Positive and negative words have been defined by the subjectivity lexicon from Opinion Lexicon (Hu and Liu, KDD-2004), a word list containing about 2,006 and 4,783 words have marked as positive and negative respectively. A message is defined as positive if it contains any positive word and negative if it contains any negative word. (This allows the messages to be either positive or negative.) This ended up in similar results as simply counting positive and negative words on a given document collection because Twitter messages are short. A major issue when using such short messages is the presence of misspellings, emoticons, links and other unnecessary contents. The pre-processing stage helped us remove these shortcomings considerably and made the resultant text clean. In order to compute the sentiment score for a new word, the other way around has

simply been worked out. Once there was the score of the text, the words were associated with the general feel of the entire text. After wards, the ensemble model has been implemented with five classifiers. The pursuit of ensemble has been motivated by the intuition that an appropriate integration of different participants might leverage distinct strengths. In multiple classifier combination, the scores generated by contributing classifiers on component feature set are taken as inputs to the combination function. Assuming that we combine D component models for a C class classification task, the ensemble model can be formulated as

$$O_j(x) = F\begin{pmatrix} O_{11}(x_1), \dots, O_{1j}(x_1), \dots, O_{1C}(x_1) \\ \dots \\ O_{D1}(x_D), \dots, O_{Dj}(x_D), \dots, O_{DC}(x_D) \end{pmatrix}$$

where $o_{kj}(x_k)$ is the predicted score of classification model k for class j and F(.) indicates the combining function. The ensemble components can be generated by different classification algorithms on different feature set.

In this ensemble model, the fixed ruels ensemble method has been adopted. The fixed rules combine the individual outputs in a fixed manner, such as the sum rule and voting rule. The common rules for fixed combination include voting rule, sum rule, max rule, and product rule. The voting rule counts the predictions of component classifiers and then assigns test sample x to class i with the most component predictions, Eqn. 3., [16,27].

$$O_j = \sum_{k=1}^{D} I\left(\underset{i}{\arg max}\,(o_{kj})=j\right), \tag{3}$$

where I(.) means the indicator function.
The sum rule combines component outputs by Eqn. 4.,

$$O_j = \sum_{k=1}^{D} o_{kj}\,, \tag{4}$$

which is equivalent to the averaging of outputs over classifiers (average rule).

## 4. Results and Discussion of Dataset #COVID-19

The accompanying step is to make visual histograms and different plots to imagine the conclusions of the Twitter user.

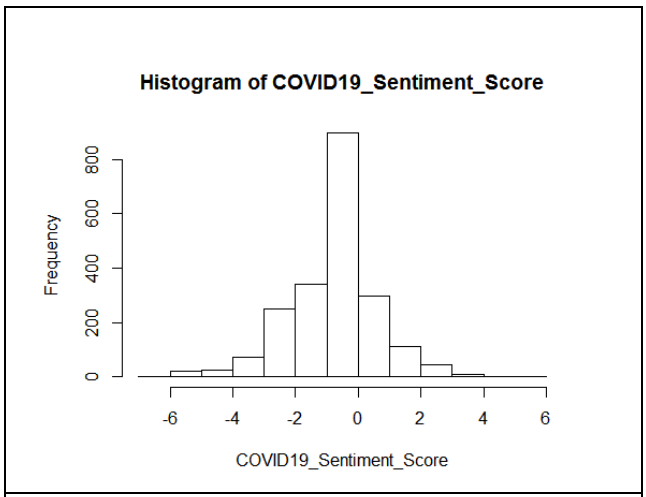


Fig. 2. Histogram of Sentiment Score of dataset #COVID-19

The above histogram (Fig. 2) demonstrates the recurrence of tweets with respect to the scores designated to every tweet. The x-pivot demonstrates the score of every tweet as a negative and positive integer or zero. A positive score speaks to positive or great sentiments connected with that specific tweet and a negative score speaks to negative or terrible sentiments connected with that tweet. A score of zero demonstrates a neutral opinion. The more positive the score, the more positive the notions of the individual tweeting and vice-versa. The snapshot (Table 1) of the score record demonstrates the score of every tweet as a whole number infront of each tweet.

Table 1. Sample Tweets with Sentiment Score on Dataset #COVID-19

| Tweet | Sentiment score |
|---|---|
| Ontario's public schools will remain closed to students until May 4th in response to the ongoing COVID-19 pandemic, the provincial government confirmed. | -1 |
| What does it look to go to online learning where not every student has a computer, or a home or a quiet space to study. Elizabeth Martinez on online learning during Covid 19. | 1 |
| Due to the COVID-19 crisis L/Cpl Brodie Gillon sadly, will not have a full military funeral. Brodie was sadly killed in a rocket attack in Iraq. So as a mark of your respect it is requested that everyone should light a candle at 9pm tonight 31 March 2020, in her memory RIP Brodie. | -5 |
| Rapid response lateral flow test for #COVID19 announces yesterday an LOI for the sale of 2M units for a #G20 country. Quickly gaining traction and international recognition for immediate, low cost testing. | 3 |
| 22-year-old from Nyanga breaks down as she is explains about the threats and treatment she has received. She was turned away by the police after she went to report that someone sent a picture of her saying she has Covid-19. The picture is currently making rounds on social media. | -2 |
| The entire Reliance team has been making effective contributions in the fight against COVID-19. Be it in healthcare or assisting people, they have been active. | 4 |
| My aunt died yesterday...her daughter died today. My cousins lost their mother and grandmother to COVID-19....in less than 24 hours. | -3 |
| We all need to do what we can to reduce the spread of COVID-19. | 0 |
| LIC Contributes Rs.105 Crore to PM CARES Fund to support India's fight against Covid-19. LIC reaffirms its solidarity in supporting the efforts of Government of India in combating the global pandemic. | 3 |
| This is unconscionable. Americans are already afraid of the impact the deadly COVID-19 pandemic is having on their lives they don't need the added stress of losing their health insurance. Drop the lawsuit. | -4 |
| D-Wave gives anyone working on responses to the COVID-19 free cloud access to its quantum computers (Frederic Lardinois/TechCrunch) | 0 |
| COVID-19 is going to be like Rise of Skywalker. It’s a nightmare. It’s going to be a struggle. But then we are going to get to come out the other side and pretend it never happened. | -2 |
| More Americans have now died with Covid-19 than died on 9/11. By the end of this week (if trend lines hold) more will have died than died in the Iraq War. By Monday the chances are high that more Americans will have died in this pandemic than in the decade long Vietnam War. | -5 |

| | |
|---|---|
| Well done! #Uber is offering NHS staff 200,000 free trips and 100,000 free meals. Each NHS staff member is able to claim up to 10 rides and five meals a week. | 4 |

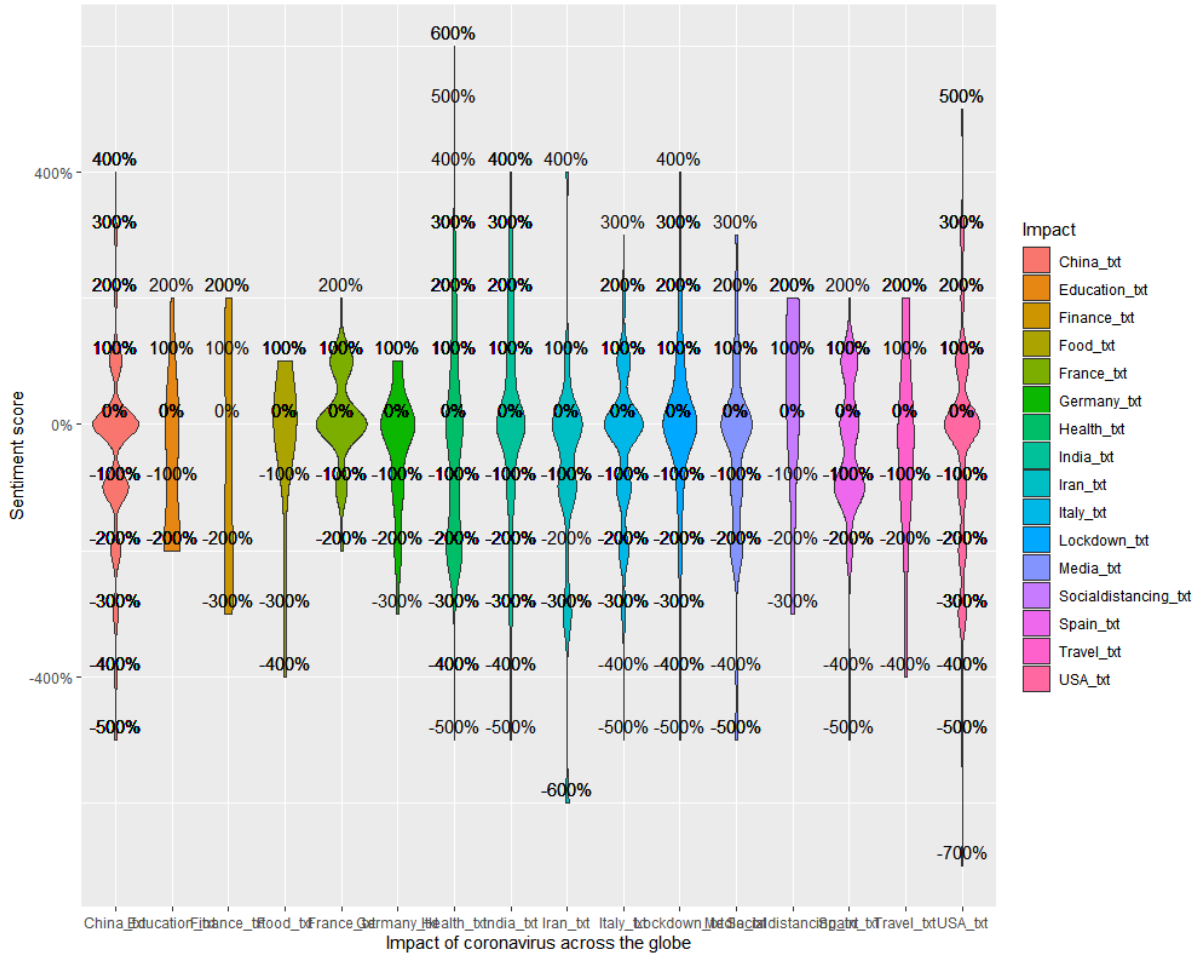


Fig. 3. Distribution of data with sentiment score in dataset #COVID-19

The chart, Figure 3 marginally skewed towards positive score which demonstrates that the sentiments of individuals in regards to COVID-19 are more positive, with a slight skew towards extremely positive sentiments on impact of coronavirus across the globe such as lockdown, media, health, education, food, finance, travel, social distancing, USA, India, Iran, Spain, Italy, Germany, France, and China. Likewise the marginal and high bitterness have been shown in negative side of chart.

Figure 4 showed the average very positive sentiment score of impact of coronavirus across the globe such as Iran, education, finance, media, China, Spain, health, Germany, Italy, USA, food, travel, lockdown is -80%, -80.49%, -76.19%, -62.50%, -54.17%, -53.09%, -43.28%, -42.59%, -35.71%,-35.25%, -33.88%, -29.03%, -13.64%, and -4.17% respectively. And the average very negative sentiment score of impact of coronavirus across the globe such as India, France, and social distancing is 6.95%, 8.09%, and 35.71% respectively. After the sentiment analysis, a histogram of the quantity of tweets with each impact classification was glanced at. There are tweets, for example, impact of coronavirus across the globe such as "lockdown, media, health, education, food, finance,travel, social distancing, USA, India, Iran, Spain, Italy, Germany, France, and China", Figures 4, 5 and 6. We should get more straightforward plots, which clearly show to us whether the tweets are in opinion positive or negative.

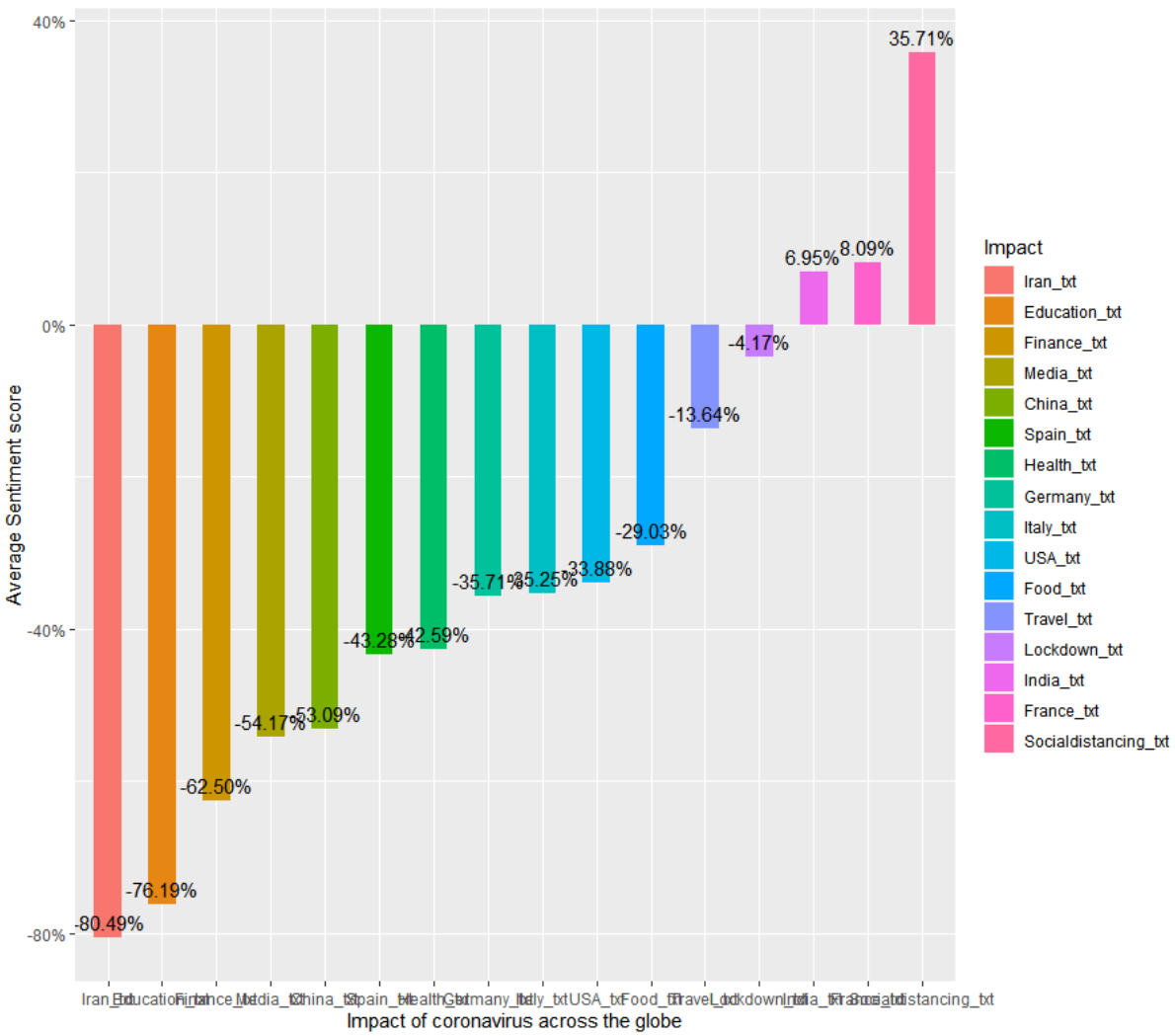


Fig. 4. Average Sentiment score on #COVID-19 dataset

From Figure 5, it was noted that the impact of coronavirus, 'social networking' has been secured an average very high positive i.e. 50% and other impacts such as finance, India, travel, France, USA, Spain, health, food, lockdown, Italy, media, China, Germany, education, and Iran are secured 38%, 30%, 27%, 26%, 25%, 24%, 23%, 23%, 21%, 20%, 17%, 16%,14%, 14%, and10% respectively.

From Figure 6, it was observed that the impact of coronavirus, 'education' has been secured an average very high negative i.e. 52.38% and other impacts such as Spain, finance, Iran, Health, China, Media, Travel, Italy, USA, Germany, Indai, Social distanceing, lockdown, food, finance are secured 50.75%, 50%, 46.34%, 46.03%, 41.98%, 39.58%, 36.36%, 34.43%,33.06%, 32.14%, 25.11%, 21.43%, 20.83%, 19.35%, and 16.18% respectively.

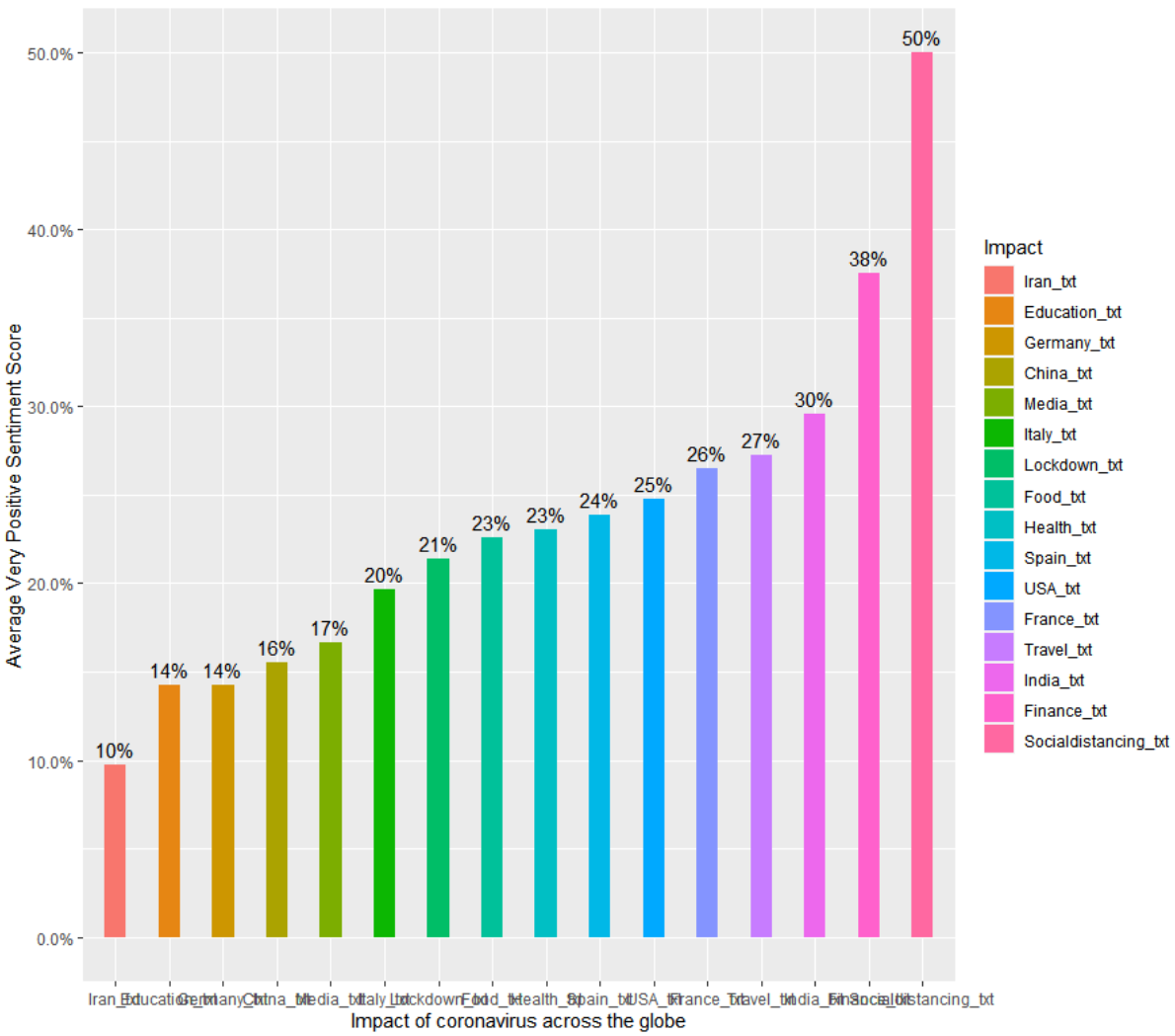


Fig. 5 Average very positive sentiment score on dataset #COVID-19

## 5. Base Machine Learning Classifier: Naive Bayes

We used 7000 tweets for training in # COVID-19 dataset, and the 3000 tweets that remained were used for testing, a mixture of positive, negative, incredibly positive and highly negative tweets. These tweets were then used with linear Naive Bayes classifier for training and testing. The existence of a few miscategorized tweets that indicate that, however, the training model was completely fit when linked to the test data that the model was in fit. The Naive Bayes classifier has accomplished 72% of accuracy in #COVID-19 dataset. The Figure 7 shows a scatterplot matrix for the Naïve Bayes classification on dataset. A scatterplot plots two variables together, one on each of the x and y axes with points showing the interaction. The spread of the points indicates the relationship between the attributes. The points in a scatterplot matrix has been colored by the class label, i.e., triangle symbol with green colored has been labeled as 'Yes' and plus symbol with red colored has been labeled as 'No' in classification problem. Note the clear separation of the points by class label on most pairwise plots. Each symbol is a data point and its position is determined by the values that data point has for a pair of variables. The class determines the symbol and color of the data point. A scatterplot matrix demonstrates a grid of scatterplots where each attribute is plotted against the other attribute. It can be read by column or row, and each plot appears twice, allowing you to consider the spatial relationships from two perspectives. From the plot take note of that (Fig. 7) dataset have comparatively does the better classification.

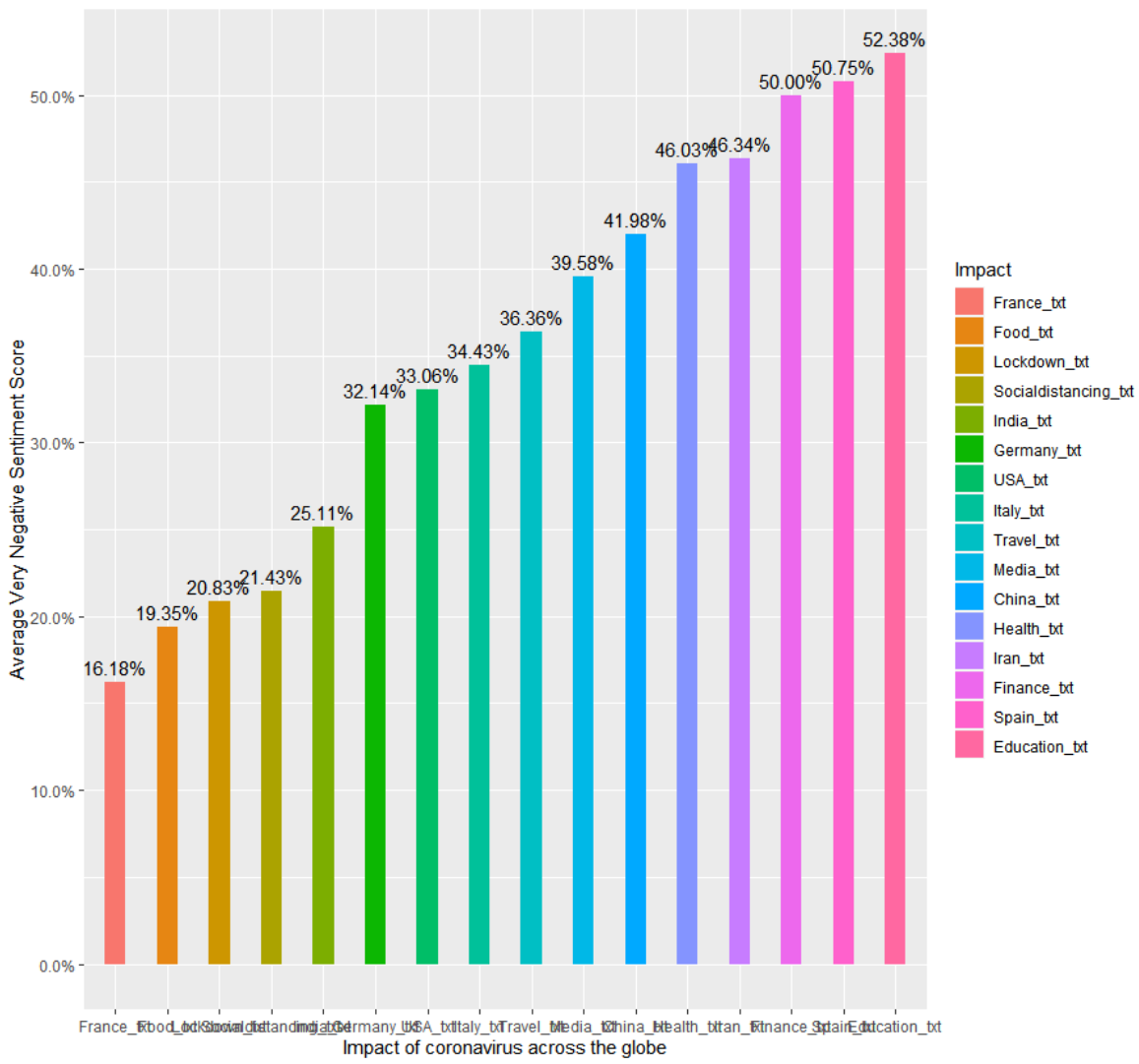


Fig. 6 Average very negative sentiment score on dataset #COVID-19

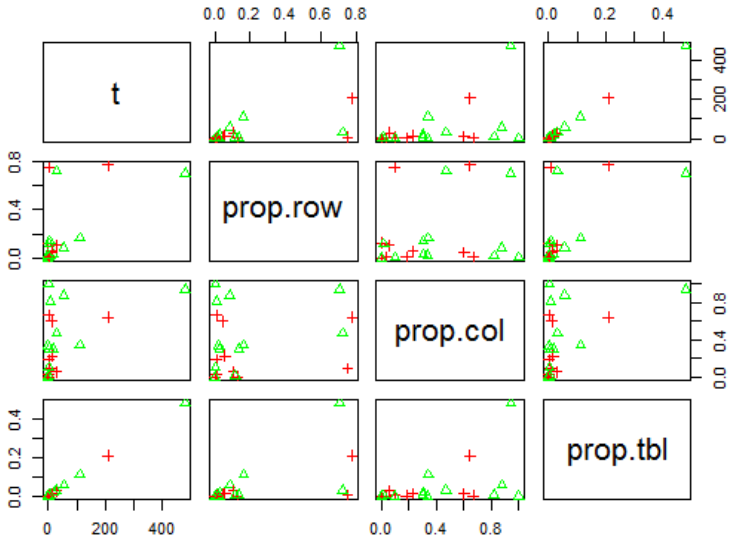


Fig. 7. Scatter plot matrix of Naïve Bayes classification on dataset #COVID-19

### 4.2. The Ensemble Classifier and Validation

Tweets data were taken in this research work, which is typically not a significant sum, but rather it is still a good decision to enforce cross validation. Cross-validation is a model validation method in which samples of a subset of data are used to show training and another subset of data is used to test the experiment. As part of this experiment, 5-fold validation technique has been used in the area of data mining research.

The 5-fold validation plan was utilized to assess the machine learning classification approaches including: the SVM classifier, the Decision Tree, Random Forest classifier, Boosting classifier and the MaxEntropy classifier. The test outcomes for the three-class experiment analysis is given in Table 2.

Table 2. Performance Metrics of Classifiers on Dataset #COVID-19

| Classifier | Performance metrics | Dataset #COVID-19 |
|---|---|---|
| Naive Bayes | Precision | 74% |
| | Recall | 54% |
| | F1 | 56% |
| SVM | Precision | 44% |
| | Recall | 35% |
| | F1 | 37% |
| Decision Tree | Precision | 39% |
| | Recall | 28% |
| | F1 | 30% |
| MaxEntropy | Precision | 39% |
| | Recall | 37% |
| | F1 | 38% |
| LogitBoost | Precision | 36% |
| | Recall | 35% |
| | F1 | 35% |
| Random Forests | Precision | 52% |
| | Recall | 31% |
| | F1 | 33% |

From the result of classifier performance metrics, we can find that recall values are comparatively lower regardless of classifier and dataset as there are some missing values in data set interpretation, and therefore relatively smaller dataset size. Especially, the SVM classifier and RandomForests classifier has been examined the median for the computation.

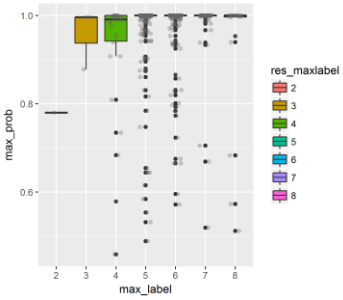


Fig. 8(a). Ensemble MaxEntropy classifier

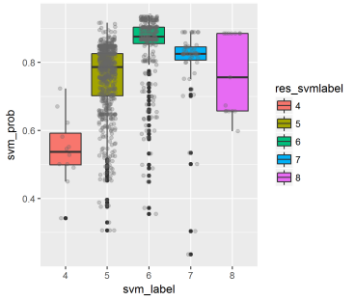


Fig. 8(b). Ensemble SVM classifier

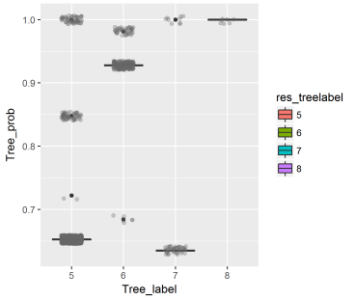


Fig. 8(c). Ensemble Decision Tree classifier

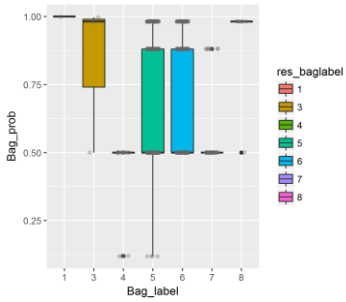


Fig. 8(d). Ensemble Bagging classifier

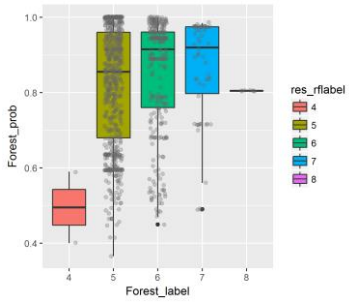


Fig. 8(e). Ensemble RandomForests classifier

Fig. 8. Classification of ensemble model on dataset #COVID-19

From Figure 8 we can use box and whisker plots to look at the distribution of the data in a different way. The box collects the middle 50 percent of the data, the line shows the median and the plot whiskers show the appropriate data range. Any dots outside the whiskers are good candidates for outliers. Again, each attribute can be summarized in terms of their observed class value, given an idea of how attribute values and class values relate each other. Especially, from the Figure 8(d), we can see the better distribution, i.e., LogitBoost (Ensemble Bagging Classifier) has been given better distribution on #COVID-19 dataset.

Table 3. Accuracy of 5-Fold Cross Validation of Ensemble Classifier on Dataset #COVID-19

| Classifier | Accuracy on dataset #COVID-19 |
|---|---|
| Naive Bayes | 72% |
| SVM | 55% |
| Decision Tree | 51% |
| MaxEntropy | 67% |
| LogitBoost | **74%** |
| Random Forests | 59% |

In the experiment with three classes, the LogitBoost ensemble classifier gets the most noteworthy accuracy which is 74% in #COVID-19 dataset, which are shown in Table 3.

### 4.3. Performance Evaluation

In performance evaluation of classification, there are three measures to assess the overall accuracy of the classifier and they are Precision, Recall and F-measure.

Precision is the fraction of the correctly classified instances for one class of the overall instances which are classified to this class,

$$\text{Precision} = \frac{True\ Positives}{True\ Positives + False\ Positives}$$

Recall is the fraction of the correctly classified instances for one class of the overall instances in this class,

$$\text{Recall} = \frac{True\ Positives}{True\ Positives + False\ Negatives}$$

A measure that combines the harmonic mean of precision and recall is the traditional F-measure or balanced F-score:

$$F = 2 * \frac{Precision * Recall}{Precision + Recall}$$

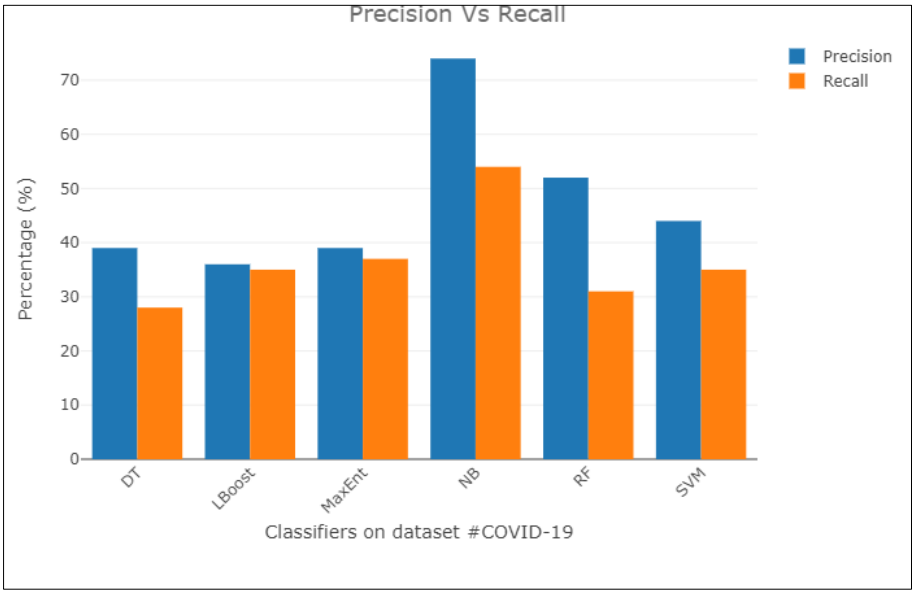


Fig. 9. Precision vs Recall of various classifiers on dataset #COVID-19

Figure 9 has shown the resultant examination amongst precision and recall on dataset #COVID-19 which has been obtained from the execution of ensemble classifier. It tells us that the dataset #COVID-19 have been done better information retrieval task.

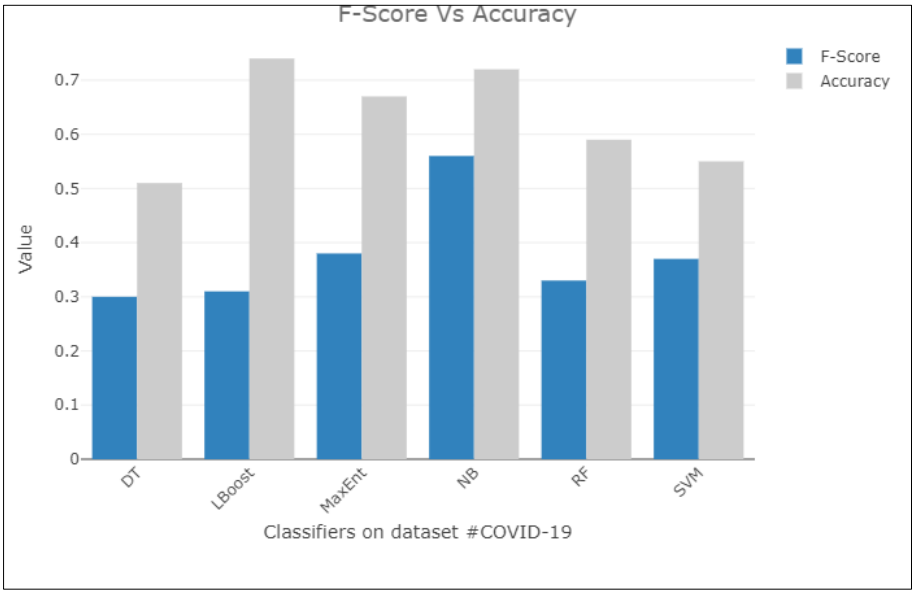


Fig. 10. F-Score vs Accuracy of various classifiers on dataset #COVID-19

From the Figure 10, we can see that the correlation of F-Score which has been obtained in view of precision and recall amid ensemble classifier and accuracy which has been acquired by cross validation. Since the size of dataset are comparatively smaller we had attempted 5-fold cross validation which has been produced high percentage of mean accuracy than F-Score which has computed based on precision and recall.

## 6. Conclusions

In this research work, machine learning methodology has been utilized as part of a request to give a careful examination of the tweets which have been extracted from Twitter. A publicly accessible sentiment lexicon which comprises of 6800 words with a list of positive and negative sentiment words or opinion words for English has been utilized to isolate the tweets. This list has been composed over numerous years by (Hu and Liu, KDD-2004). Tweets have then been classified into positive, negative and neutral classes with the help of the machine learning classifier, Naive Bayes. Furthermore, it makes exact commitments to this research area by comparing the execution of various well known sentiment classification methodologies and building up an ensemble approach, which additionally enhances the execution of the sentiment classification.

Research work has been done in the field of twitter sentiment analysis on the classification of coronavirus effects around the globe. These past works have addressed few diverse conventional classification strategies and choose the most precise classification strategy for performing classification of sentiments. In either case, the approach to the ensemble is added, improves the accuracy by entering certain classifiers of emotions. The sentiments classification accuracy is high enough to achieve public loyalty perusal for the impact of coronavirus all over the globe. This methodology is important for the government and the public to inspect the information on twitter in order to understand public opinion about the impact of coronavirus around the world. The other impacts of coronavirus that can be identified over the time as well as the users who tweet them, the followers and the seasons of the re-tweets are to be considered for further research.

## References

[1] Haewoon Kwak, Changhyun Lee, Hosung Park, and Sue Moon (2010). What is Twitter, a Social Network or a News Media?, International World Wide Web Conference Committee (IW3C2), WWW 2010, USA.

[2] Wu. S, Hofman. J. M, Mason. W. A, and Watts. D. J (2011). Who Says What to Whom on Twitter, Proceedings of 20th International Conference on World Wide Web (WWW'11), pp. 705-714.

[3] Mike Ember Isaac (2016). For Election Day Influence, Twitter Ruled Social Media, The New York Times.

[4] Muthusami. R, Bharathi. A (2019). Stance detection and mobile app recommendation discourse on tweets, Computational Intelligence, 35:1043–1060. https://doi.org/10.1111/coin.12231

[5] Agarwal. R, Xie. A, Vovsha. B, Rambow. I, and Passonneau. O (2011). Sentiment Analysis of Twitter Data, Proceedings of the Workshop on Language in Social Media (LSM 2011), pp.30-38, Portland.

[6] Pak. P, Paroubek. A (2011). Twitter as a Corpus for Sentiment Analysis and Opinion Mining, Proceedings of Seventh International Conference on Language Resources and Evaluation (LREC'10).

[7] Sara Rosenthal, Noura Farra, Preslav Nakov (2017). SemEval-2017 Task 4: Sentiment Analysis in Twitter, Proceedings of the 11th International Workshop on Semantic Evaluations (SemEval-2017), pp. 502–518, Canada.

[8] World Health Organization. Pneumonia of unknown cause - China. January 5, 2020. (https://www.who.int/csr/don/ 05-january-2020-pneumonia-of-unkown-cause-china/en/).

[9] World Health Organization. Novel Coronavirus — China. January 12, 2020 (https://www .who .int/ csr/ don/ 12 -january-2020-novel-coronavirus-china/en/).

[10] Zhu. N, Zhang. D, Wang. W (2020). A novel coronavirus from patients with pneumonia in China, 2019. N Engl J Med, 382:727-33.

[11] Centers for Disease Control and Prevention. Symptoms of coronavirus disease 2019 (COVID-19). 2020 (https://www.cdc.gov/coronavirus/2019-ncov/about/symptoms .html).
[12] World Health Organization. Coronavirus disease 2019 (COVID-19): situation report - 71. March 31, 2020 (https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200331-sitrep-71-covid-19.pdf?sfvrsn=4360e92b4)
[13] World Health Organization. Rolling updates on coronavirus disease (COVID-19). 2020 (https://www.who.int/emergencies/diseases/novel–coronavirus-2019/events-as-they-happen).
[14] Bing Liu (2012). Sentiment Analysis and Opinion Mining, Morgan & Claypool Publishers.
[15] Hoda Korashy Walaa Medhat, Ahmed Hassan (2014). Sentiment analysis algorithms and applications: A survey, Ain Shams Engineering Journal, pp.1093-1113.
[16] Jibao Gu Gang Wang., Jianshan Sun., Jian Ma, and Kaiquan Xu (2013). Classification: The contribution of ensemble learning, Decision Support Systems, Elsevier, pp.77-93.
[17] Federico Alberto Pozzi, Elisabetta Fersini, and Vincenzina Messina (2013). Bayesian Model Averaging and Model Selection for Polarity Classification, Proc. of the 18th International Conference on Application of Natural Language to Information Systems, Volume: LNCS 7934, pp. 189-200.
[18] Carvalho. J., Prado. A, Plastino. A (2014). A Statistical and Evolutionary Approach to Sentiment Analysis, Proceedings of the 2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT) , Washington, DC, USA. IEEE Computer Society, 02, pp. 110–117.
[19] Olga Kolchyna., Tharsis. T., Souza. P, Philip. C., Treleaven, and Tomaso Aste (2015). Twitter Sentiment Analysis: Lexicon Method, Machine Learning Method and Their Combination, arXiv:1507.00955v3 [cs.CL].
[20] Khaled Mohammad Alomari, Hatem M. ElSherif, and Khaled Shaalan (2017). Arabic Tweets Sentimental Analysis Using Machine Learning, IEA/AIE 2017, Part I, LNAI 10350, pp. 602–610.
[21] Tapan Sahni, Chinmay Chandak, Naveen Reddy and Manish Singh (2017). Efficient Twitter Sentiment Classification using Subjective Distant Supervision, arXiv:1701.03051v1 [cs.SI].
[22] Liu. B, and Hu. M. (2004). Opinion lexicon, (http://www.cs.uic.edu/liub/FBS/sentiment-analysis.html).
[23] Esuli. A, Sebastiani. F (2006). SentiWordNet: A Publicly Available Lexical Resource For Opinion Mining, (www.bibsonomy.org/bibtex/25231975d0967b9b51502fa03d87d106b/mkroell)
[24] Nielsen. F (2011), A new ANEW: Evaluation of a word list for Sentiment Analysis in Microblogs, The ESWC2011 Workshop on 'Making Sense of Microposts': Big things come in small, pp. 93-98.
[25] Mohammad. S., Kiritchenko. M, and Zhu (2013). Nrc-Canada: Building the state-of- the-art in sentiment analysis of tweets, Proceedings of seventh international workshop on Semantic Evaluation Exercises (SemEval-2013).
[26] Mitchell. T. M (2013). Machine Learning, McGrw-Hill, New York,USA.
[27] Rui Xia, Chengqing Zong, and Shoushan Li (2011), Ensemble of Feature Sets and Classification Algorithms for Sentiment Classification, Information Sciences, 181:1138–1152, Elsevier.
[28] Martin Sewell (2011), Ensemble Learning, Research Note, UCL Department of Computer Science.
[29] Yun Wan, and Qigang Gao (2015). An Ensemble Sentiment Classification System of Twitter Data for Airline Services Analysis, Proceedings of IEEE 15th International Conference on Data Mining Workshops, pp.1318- 1325.