[1]K. S. Padmashree*,

Dr. P. Velmani,

Dr. S. Loghambal

Classification and the Prediction of Covid-19 by Applying Naive Bayes and Random Forest Algorithms

Journal of Electrical Systems

Abstract: **-** Classification is a supervised learning algorithm in machine learning that categorizes input data into specific labels according to its features. The primary objective of classification is to develop a model that can reliably predict the appropriate label or category to previously unseen data. COVID-19, a highly contagious and severe disease caused by the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), is believed to have originated in bats and transmitted to humans via an unidentified intermediary in Wuhan, China, in late December 2019. This disease can lead to significant organ dysfunction, impacting essential organs such as the heart, liver, and kidneys, as well as disrupting the normal functioning of organ systems, including the cardiovascular and immune systems. The focus of this study aims to classify and predict COVID-19 outcomes using two machine learning algorithms: Naïve Bayes (NB) and Random Forest (RF). The research utilizes the COVID_Data.CSV dataset, which contains 316,800 data points. Of these, 70% are utilized for training the models, while the remaining 30% is allocated for testing. The Naïve Bayes classifier gets an accuracy of 87.39%, while the Random Forest classifier achieves a slightly higher accuracy of 87.47%. A comparative analysis reveals that the Random Forest classifier outperforms Naïve Bayes, establishing it as the more effective model for the classification and prediction of Covid-19 utilizing Machine Learning (ML) techniques.

Keywords: Classification, Prediction, Naive Bayes, Random Forest, Data Pre-processing, Machine Learning, Covid-19, Dataset, Coronavirus.

1. INTRODUCTION

Classification involves the process of grouping objects, information, or data into separate groups or classes according to shared characteristics or attributes. It is a strategy for organizing data or objects into significant groups based on their similarities and contrasts.

In machine learning, classification is a supervised learning approach where a model is trained to assign input data to predefined classes. The objective is to learn a decision boundary that effectively distinguishes between these classes based on the features of the input data. Once trained, the model can be used to predict the classification of new, unidentified data. Classification algorithms are supervised learning techniques that categorize new observations. To train a classification model, the algorithm requires an appropriately labelled dataset that has all examples of input data and their corresponding labels. The algorithm then learns to identify patterns associated with each label, enabling accurate predictions for new, unseen data.

Prediction is estimating or forecasting future events or outcomes based on patterns and trends found in historical data using data and statistical or machine learning techniques. Accurate predictions depend on high-quality, relevant data that pertains to the event or outcome being predicted. This data is utilized in developing predictive models that represent observed patterns and relationships mathematically or statistically. The aim of prediction is to make informed forecasts about future occurrences based on past events and trends.

[1]Research Scholar, PG & Research Department of Mathematics, The M.D.T. Hindu College,
[2]Research Supervisor, Assistant Professor, Department of Computer Science,
The M.D.T. Hindu College,
[3]Research Joint Supervisor, Assistant Professor, Department of Mathematics,
The M.D.T. Hindu College,
(Affiliated to Manonmaniam Sundaranar University, Tirunelveli)
*Corresponding Author E-mail: padmashree.abi@gmail.com, Tel.: +91-6383199533

In data mining, the data is typically prepared by cleaning and pre-processing to ensure its quality, consistency, and relevance. The data is then explored and analyzed using techniques such as clustering, association rules, and regression analysis to identify patterns and relationships. These patterns and relationships serve as the basis for building predictive models that support informed decision-making.

The major objective of this paper is to identify and predict COVID-19 using the Naive Bayes and Random Forest algorithms. Machine Learning is a branch of study that enables computers to learn without being explicitly programmed (Arthur Samuel, 1959). Machine learning is transforming the world. In machine learning, the item we generate from data and labels is referred to as the model. As of January 17, 2023, there were 671,545,004 COVID-19 cases and 6,731,897 deaths globally (https://www.worldometers.info/coronavirus/). The primary goal of the research is to create a machine learning model that can categorize and predict whether a patient has COVID-19 or not based on symptoms. Figure 1.1 depicts the Block Diagram of the Machine Learning Model.

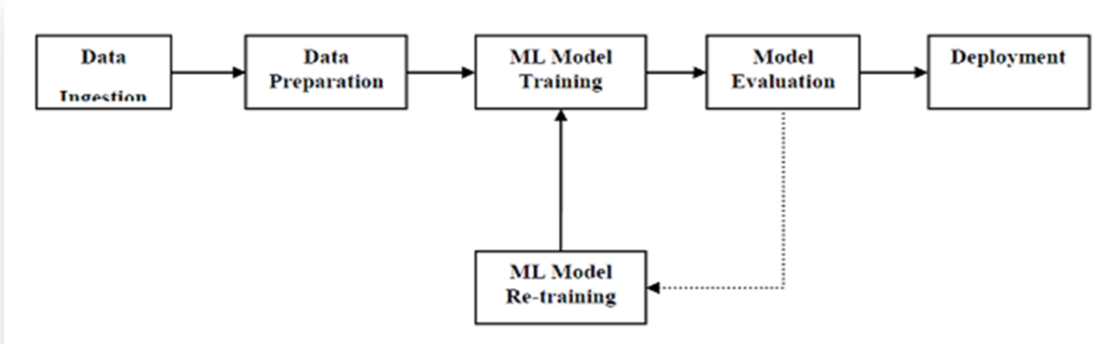


Figure 1.1: The Block Diagram of the Machine Learning Model

1.1 COVID-19 Statistics

As reported by the Worldometers.info platform, the United States has recorded the highest number of COVID-19 cases globally. The COVID-19 case statistics for various countries are presented in Table 1.1.

Table 1.1: Covid-19 Statistics from Worldometers (as of on 17-01-2023)

Sl. No.	Country	Total Cases	Total Deaths	Total Recovered	Active Cases	Critical Cases
1	USA	103,583,738	1,125,558	100,422,095	2,036,085	4,491
2	India	44,681,884	530,726	44,148,309	2,849	698
3	France	39,453,006	163,463	39,056,473	233,070	869
4	Germany	37,605,135	163,775	37,063,800	377,560	1,281
5	Brazil	36,661,526	695,461	35,580,516	385,549	8,318
6	Japan	31,471,011	62,963	21,370,395	10,037,653	687
7	South Korea	29,821,035	32,984	28,898,191	889,860	510
8	Italy	25,363,742	185,993	24,824,106	353,643	310
9	UK	24,243,393	202,157	23,926,710	114,526	146
10	Russia	21,860,902	394,438	21,278,106	188,358	2,300
11	China	503,302	5,272	379,053	118,977	7,557

Source: https://www.worldometers.info/coronavirus/#countries

As stated by the World Health Organization (WHO), These are the five primary symptoms that are commonly associated with COVID-19: fever, tiredness, difficulty breathing, dry cough, and sore throat. A detailed summary of these key symptoms is presented in Table 1.2.

Table 1.2: Symptoms of Covid-19 suggested by WHO

Sl.No.	Symptoms
1	Fever
2	Tiredness
3	Difficulty in Breathing
4	Dry Cough
5	Sore Throat

In the present study, patients will be diagnosed based on the attributes outlined in Table 1.2. When individuals exhibit symptoms as described in Table 1.3, these indicators will be used to classify and predict the patient's chance of contracting Covid-19.

Table 1.3: Attributes for the Detection of COVID-19

Sl. No.	Symptoms									COVID-19 or Not (YES / NO)
	Fever	Tiredness	Dry Cough	None Experiencing	Sore Throat	Diarrhoea	Runny Nose	Difficulty in Breathing	Nasal Congestion	
1	y	y	y	y	y	y	y	y	y	y
2	y	y	y	y	y	y	y	y	n	y
3	y	y	y	y	y	y	y	n	n	y
4	y	y	y	y	y	y	n	n	n	y
5	y	y	y	y	y	n	n	n	n	y
6	n	n	n	n	n	n	n	n	n	n
7	n	n	n	n	n	n	n	n	y	n
8	n	n	n	n	n	n	n	y	y	n
9	n	n	n	n	n	n	y	y	y	n

2. LITERATURE REVIEW

This section provides an overview of prior research efforts focused on the classification of COVID-19 using Naive Bayes and Random Forest machine learning algorithms.

Sugandh Bhatia et al. (2021) introduced the use of the Naïve Bayes Classifier for predicting COVID-19 cases through machine learning techniques. This paper explores various data mining methods aimed at predicting the occurrence of COVID-19 cases, with the primary goal of developing a predictive model for identifying coronavirus infections.

Nehal Mansour et al. (2022) discussed the effective identification of COVID-19 patients using the Feature Correlated Naïve Bayes (FCNB) classification method. This study presents FCNB as an innovative diagnostic strategy for COVID-19 detection. Additionally, the authors highlight the critical role of chest radiological imaging techniques, including Computed Tomography (CT) scans and X-ray images, in facilitating early detection and subsequent management of COVID-19 cases.

L. J. Muhammad et al. (2021) introduced supervised machine learning models for predicting COVID-19 infections utilizing an epidemiological dataset. In their study, 80% of the data was allocated for model training, while the remaining 20% served as the testing set. Among the evaluated models, the Naive Bayes model demonstrated the highest specificity, achieving a value of 94.30%.

H. Zakiyyah et al. (2021) conducted a study aimed at predicting COVID-19 infections in Indonesia using machine learning techniques. The research explored three predictive models: Gaussian Naive Bayes (GNB), Support Vector Machine (SVM), and Decision Tree (DT), focusing on forecasting the total number of cases and deaths due to COVID-19 in the country. The analysis, which utilized the CITSAD dataset from ten Indonesian provinces, revealed that the Decision Tree model achieved the highest accuracy rate of 70%, while also offering the quickest processing time of 60 seconds.

3. FRAMEWORK OF THE PROPOSED SYSTEM

This section provides an in-depth overview of the framework underlying the proposed model for COVID-19 classification. A block diagram illustrating the structure of the proposed classification model is presented in Figure 3.1.

3.1.

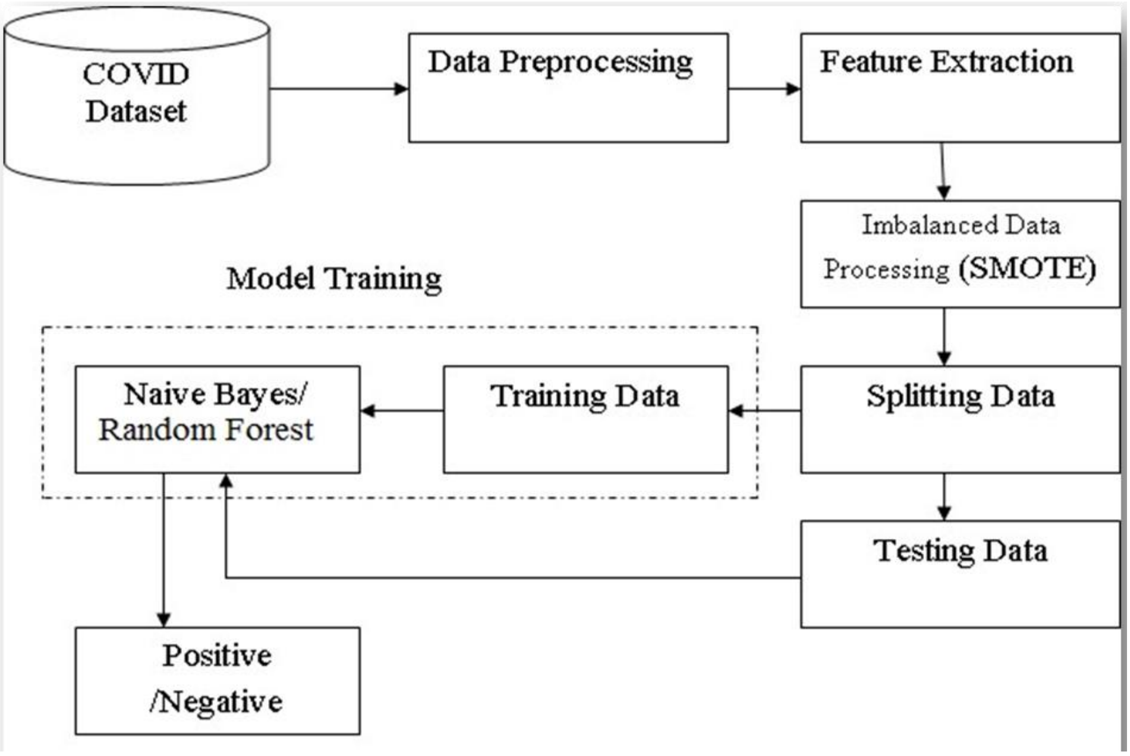


Figure 3.1: The Block Diagram of the Proposed COVID-19 Classification Model

3.1 DATA PREPROCESSING

In the COVID-19 Classification Model process, preprocessing is considered an important step in COVID-19 classification. Today, most of the data in the real world are incomplete, meaningless, and noisy data. The main aim of this data preprocessing is to remove the unnecessary noisy data which does not give meaningful information.

3.2 FEATURE EXTRACTION

Feature selection refers to the process of identifying and selecting the most relevant variables from a dataset. This procedure is crucial in enhancing the effectiveness and accuracy of predictive models. By reducing dimensionality, feature selection facilitates the extraction of essential information from large datasets, and reduces processing time with better performance.

3.3 SMOTE

The Synthetic Minority Over-sampling Technique (SMOTE) is a statistical technique that increases the number of cases in your dataset in a balanced manner. It is a prominent approach for addressing the problem of imbalanced datasets in machine learning. Imbalanced datasets are a frequent problem in many real-world applications, where one class has much less instances than the others. This can provide biased models that are more likely to predict the majority class, resulting in poor performance in recognizing minority class cases. Overall, SMOTE is an effective tool for dealing with imbalanced datasets, and it is commonly utilized in machine learning applications. The goal of SMOTE is to generate synthetic minority samples that are identical to existing minority samples in order to balance the class distribution and improve classification performance.

3.3.1 The SMOTE algorithm works as follows:

Split the original data into minority and majority class samples.Determine the imbalance ratio between the minority and majority classes. This is the ratio of the number of samples from the majority and minority classes. Set your desired imbalance ratio. This is the desired ratio of majority to minority class samples in the balanced dataset. Calculate how many minority class samples will be generated. This is the difference between the number of minority class samples in the desired balanced dataset and those in the original dataset.

From each minority class sample, choose k (default k=5) nearest neighbours. Create a synthetic sample for each of the k neighbours by randomly selecting a point on the line connecting the minority sample and the neighbours, then adding a random fraction of the difference between the two points to the minority sample. The synthetic sample adds to the feature space. To generate a new balanced dataset, combine the original minority class samples with the synthetic minority class samples.Train a classifier or another machine learning model using the balanced dataset.

By creating synthetic minority class samples that are similar to the existing minority class samples, the SMOTE algorithm helps to reduce the bias towards the majority class and improve classification performance.

4. DATASET

Data is the key to unlocking machine learning as much as machine learning is the very important key to unlocking the insight hidden in data. Machine learning technology helps to detect COVID-19 positive cases based on the data and its features.

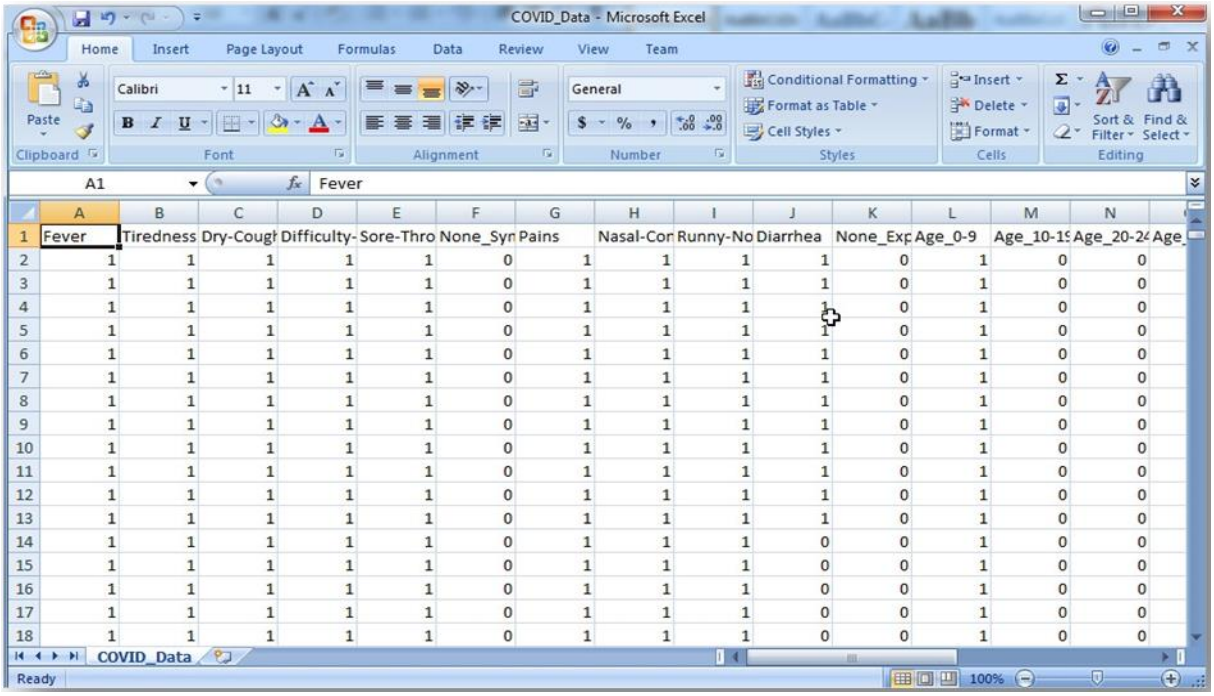
4.1 DATASET COLLECTION

For the data collection, the researchers used a dataset available from Kaggle entitled “COVID_Data”. This dataset has 27 features with 3,16,800 records. The descriptions of the dataset are shown in Figure 4.1.

```
Data columns (total 27 columns):
 #   Column                   Non-Null Count   Dtype
---  ------                   --------------   -----
 0   Fever                    316800 non-null  int64
 1   Tiredness                316800 non-null  int64
 2   Dry-Cough                316800 non-null  int64
 3   Difficulty-in-Breathing  316800 non-null  int64
 4   Sore-Throat              316800 non-null  int64
 5   None_Sympton             316800 non-null  int64
 6   Pains                    316800 non-null  int64
 7   Nasal-Congestion         316800 non-null  int64
 8   Runny-Nose               316800 non-null  int64
 9   Diarrhea                 316800 non-null  int64
 10  None_Experiencing        316800 non-null  int64
 11  Age_0-9                  316800 non-null  int64
 12  Age_10-19                316800 non-null  int64
 13  Age_20-24                316800 non-null  int64
 14  Age_25-59                316800 non-null  int64
 15  Age_60+                  316800 non-null  int64
 16  Gender_Female            316800 non-null  int64
 17  Gender_Male              316800 non-null  int64
 18  Gender_Transgender       316800 non-null  int64
 19  Severity_Mild            316800 non-null  int64
 20  Severity_Moderate        316800 non-null  int64
 21  Severity_None            316800 non-null  int64
 22  Severity_Severe          316800 non-null  int64
 23  Contact_Dont-Know        316800 non-null  int64
 24  Contact_No               316800 non-null  int64
 25  Contact_Yes              316800 non-null  int64
 26  Country                  316800 non-null  object
dtypes: int64(26), object(1)
```

Figure4.1: Descriptions of the Dataset

The sample dataset of COVID-19 is shown in figure 4.2. COVID-19 Data is stored in .CSV format.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Fever	Tiredness	Dry-Cougl	Difficulty-	Sore-Thro	None_Syn	Pains	Nasal-Cor	Runny-No	Diarrhea	None_Exp	Age_0-9	Age_10-1!	Age_20-2
2	1	1	1	1	1	0	1	1	1	1	0	1	0	0
3	1	1	1	1	1	0	1	1	1	1	0	1	0	0
4	1	1	1	1	1	0	1	1	1	1	0	1	0	0
5	1	1	1	1	1	0	1	1	1	1	0	1	0	0
6	1	1	1	1	1	0	1	1	1	1	0	1	0	0
7	1	1	1	1	1	0	1	1	1	1	0	1	0	0
8	1	1	1	1	1	0	1	1	1	1	0	1	0	0
9	1	1	1	1	1	0	1	1	1	1	0	1	0	0
10	1	1	1	1	1	0	1	1	1	1	0	1	0	0
11	1	1	1	1	1	0	1	1	1	1	0	1	0	0
12	1	1	1	1	1	0	1	1	1	1	0	1	0	0
13	1	1	1	1	1	0	1	1	1	1	0	1	0	0
14	1	1	1	1	1	0	1	1	1	0	0	1	0	0
15	1	1	1	1	1	0	1	1	1	0	0	1	0	0
16	1	1	1	1	1	0	1	1	1	0	0	1	0	0
17	1	1	1	1	1	0	1	1	1	0	0	1	0	0
18	1	1	1	1	1	0	1	1	1	0	0	1	0	0

Figure 4.2: Sample Dataset

5. NAÏVE BAYES CLASSIFIER FOR THE CLASSIFICATION OF COVID-19

Naïve Bayes is a supervised probabilistic Machine Learning Algorithm completely based on the Bayes theorem. Bayes algorithm is mainly used for classification purposes. According to Bayes theorem, where A and B are events and

P(B) is not zero, the conditional probability can be calculated using the given mathematical equation 1.

$$P(A|B) = \frac{P(B|A)*P(A)}{P(B)} \quad \text{.. Equation (1)}$$

The Bayes theorem is used to calculate the conditional probability, which is nothing but the probability of an event occurring based on information about the events in the past. In the above equation 1,

$P(A|B)$:-Conditional probability of event A occurring, given the event B (Posterior Probability)

$P(A)$: Probability of event A occurring (Prior Probability)

$P(B)$: Probability of event B occurring (Marginal Probability)

$P(B|A)$: Conditional probability of event B occurring, given the event A (Likelihood Probability)

The Posterior Probability ($P(A|B)$ can be calculated first by creating a frequency table.

5.1 HOW NAÏVE BAYES CLASSIFIER WORKS IN COVID-19?

For simplifying prior and posterior probability we can use the two tables such as frequency table and likelihood table. These two tables are used to calculate the prior and posterior probability. The frequency table contains the occurrence of labels for all features. In this research work, Naive Bayes uses multiple features. An example of multiple features is listed in table 5.1.

Table 5.1: Multiple Features Used in Covid-19 Classification

Sl.No.	Symptoms
1	Fever
2	Tiredness
3	Dry Cough
4	NoneExperiencing
5	Sore Throat
6	Diarrhoea
7	Runny Nose
8	Difficulty in Breathing
9	Nasal Congestion

Now we should try this by applying the above equation (1) automatically on the COVID-19 dataset. For applying this, we took the first 25 data from the dataset. The probabilities of symptoms are listed in table 5.2.

Table 5.2: Probabilities of Symptoms

Sl.No.	Symptoms	Yes	No	P (yes)	P(no)
1	Fever	20	5	20/25	5/25
2	Tiredness	19	6	19/25	6/25
3	Dry Cough	21	4	21/25	4/25
4	NoneExperiencing	10	15	10/25	15/25
5	Sore Throat	22	3	22/25	3/25
6	Diarrhoea	18	7	18/25	7/25
7	Runny Nose	21	4	21/25	4/25
8	Difficulty in Breathing	22	3	22/25	3/25
9	Nasal Congestion	21	4	21/25	4/25
		174	**51**		

For example, the probability of occurring COVID-19 given that the symptom is fever, P (probability=fever | occurrence=yes) =20/25. Also, we have to find the class probabilities P(y) from above table 5.2.

For example, P (occurrence=yes) =174/ 225 and P (occurrence=no) = 51/ 225

So, the probability of occurring of COVID-19 is given by:

P (yes| occurrence) = P (Fever | yes) * P (Tiredness | yes) * P (Dry Cough | yes)* P (None Experiencing | yes)* P (Sore Throat | yes)* P (Diarrhoea | yes)* P (Runny Nose | yes)* P (Difficulty in Breathing | yes)* P (Nasal Congestion | yes) * P (yes) / P (current occurrence)

P(yes|occurrence)=20/25 *19/25*21/25*10/25*22/25*18/25*21/25*22/25*21/25*174/225 / 1

P (yes | occurrence) = 0.0621

The probability of not occurring of COVID-19 given by:

P (no | occurrence) = P (Fever | no) * P (Tiredness | no) * P (Dry Cough | no)* P (None Experiencing | no)* P (Sore Throat | no)* P (Diarrhoea | no)* P (Runny Nose | no)* P (Difficulty in Breathing | no)* P (Nasal Congestion | no) * P (no) / P (current occurrence)

P (no | occurrence) = 5/25 *6/25*4/25*15/25*3/25*7/25*4/25*3/25*4/25*51/225 / 1

P (no | occurrence) =0.000749

Since P (Yes | occurrence)> P (No | occurrence)

So, the prediction that COVID-19 occurs is 'Yes'.

6. RANDOM FOREST CLASSIFIER FOR THE CLASSIFICATION OF COVID-19

The Random Forest (RF) classification method was introduced by Leo Breiman in the year 2001. Random forest is a supervised learning algorithm. The "forest" it builds is an ensemble of decision trees, usually trained with the "bagging" method. The general idea of the bagging method is that a combination of learning models increases the overall result.

Random Forest is an ensemble classifier formed by the fusion of many decision trees. It calculates the result on the basis of the majority voting method. Random forest is superior then the decision tree as it overcomes the problem of over-fitting. As a tree grows deep, they start to over-fit, i.e., they have low bias and high variance.

Step 1: Training Dataset is converted into Bootstrapped Dataset (Random sampling with replacement)

Step 2: Construct a Decision Tree using the Bootstrapped Dataset. Bootstrapping is the process of random sampling with replacement. We can select the same value in multiple times.

Step 3: Each decision tree will generate an output.

Step 4: Voting Phase: This is the final phase of the RF method which basically deals with the voting or pooling phase. This phase helps in determining correct and incorrect features for each of the trees in the forest. The decision tree of the COVID-19 Random Forest classification is shown in figure 6.1.

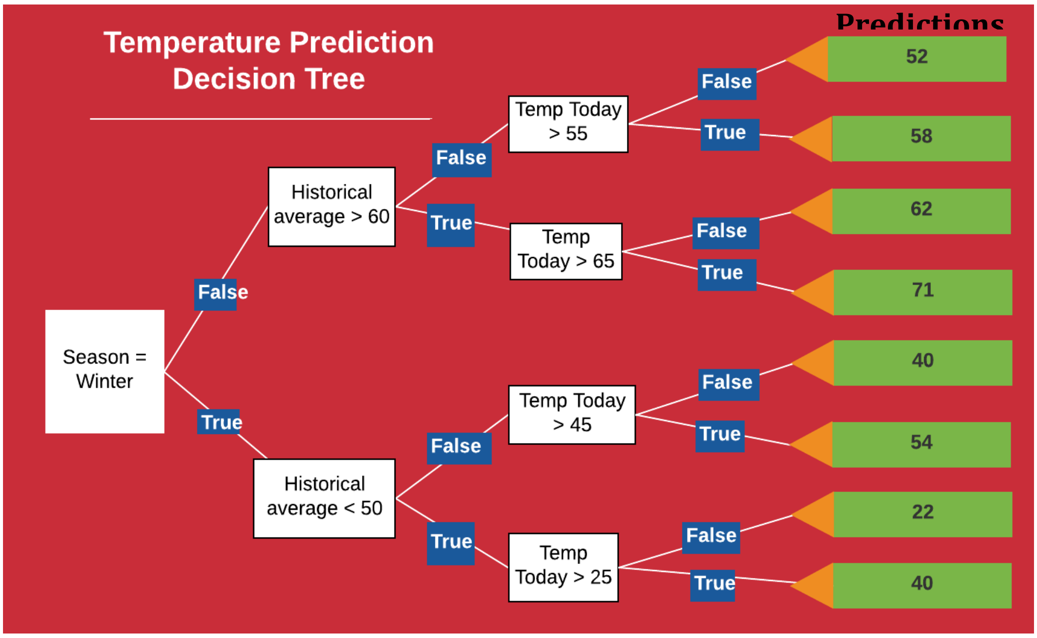


Figure 6.1: Decision Tree of COVID-19 Random Forest Classification

Using decision trees, we can build a random forest. One problem that might occur with one big (deep) single DT is that it can overfit. That is the DT can "memorize" the training set the way a person might memorize an Eye Chart.

The point of Random Forest is to prevent overfitting. It does this by creating random subsets of the features and building smaller (shallow) trees using the subsets and then it combines the sub-trees. The downside of RF is it can be slow if you have a single process but it can be parallelized. The working principle of Random Forest is shown in figure 6.2.

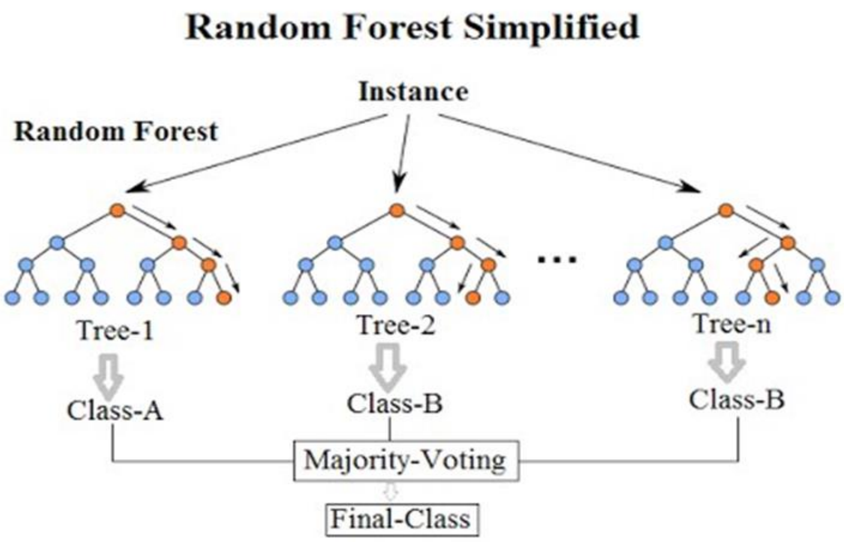


Figure 6.2: The Working Principle ofRandom Forest Structure

7. IMPLEMENTATION

Implementation of COVID-19 classification has been done using Machine Language with Python. This paper was successfully completed with the implementation of the Machine Learning Technique.

8. RESULTS AND DISCUSSIONS

This section presents a detailed description of the proposed COVID-19 classification model, accuracy result and their comparisons.

8.1 EVALUATION METRICS

During the experiments, the evaluation parameters such as Accuracy, Recall, Precision and F1-score will be calculated. The confusion matrix is used to calculate the above-said parameters.

True positive (TP): correctly predicted positive class data points.

False positive (FP): negative class data points predicted as a positive class. True negative (TN): correctly predicted negative class data points.

False negative (FN): negative class data points predicted as a positive class.

P: total positive class data points.

N: total negative class data points.

True positive rate (TPR) = TP/P

False positive rate (FPR) = FP/N

Performance measures are defined using the following equations: The equations are shown in figure 8.1.

$$\text{Accuracy} = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{FP} + \text{TN} + \text{FN})},$$

$$\text{Precision} = \frac{\text{TP}}{(\text{TP} + \text{FP})},$$

$$\text{Recall} = \frac{\text{TP}}{(\text{TP} + \text{FN})},$$

$$\text{F} - \text{measure} = \frac{2 * \text{precision} * \text{recall}}{(\text{precision} + \text{recall})}.$$

Figure 8.1: The Equations of Performance Measures

Based on the experiments done in this research, the overall accuracy result and comparisons are listed in table 8.1.

Table 8.1: Accuracy Result and Comparisons

Sl.No.	Number of Patients Records	Naive Bayes (NB)-Accuracy in Percentage	Random Forest (RF)-Accuracy in Percentage
1	3,16,800	87.39%.	87.47%

The confusion matrix of both the Naive Bayes and Random Forest algorithms of COVID-19 classification is shown in figure 8.2 and figure 8.3.

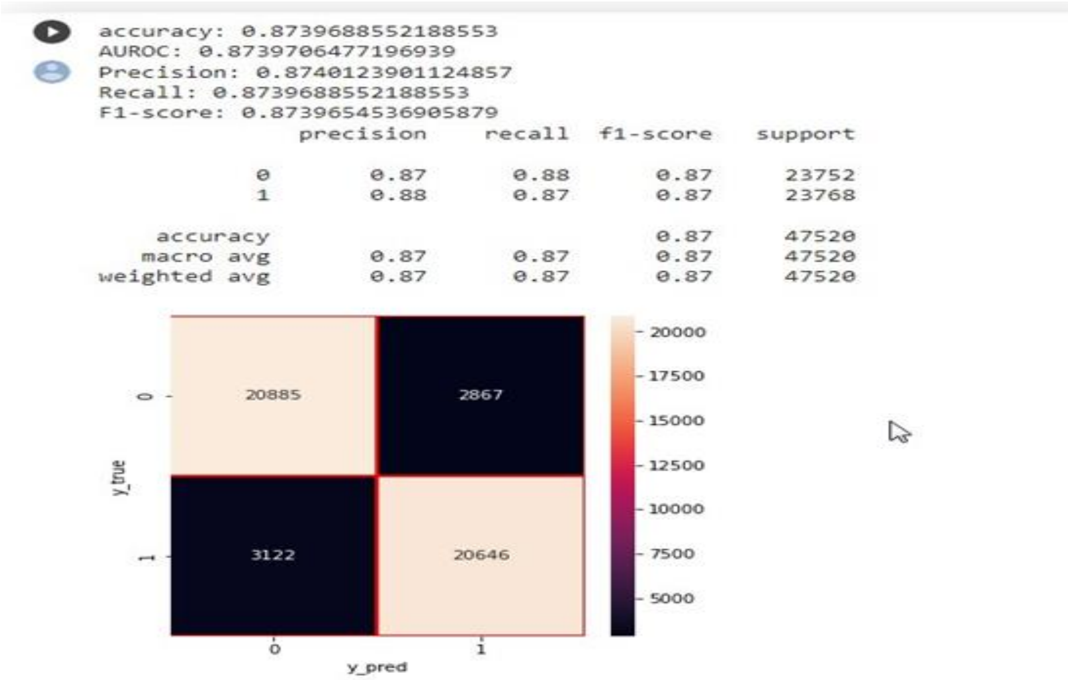


Figure 8.2: The Confusion Matrix of Naive Bayes Algorithm

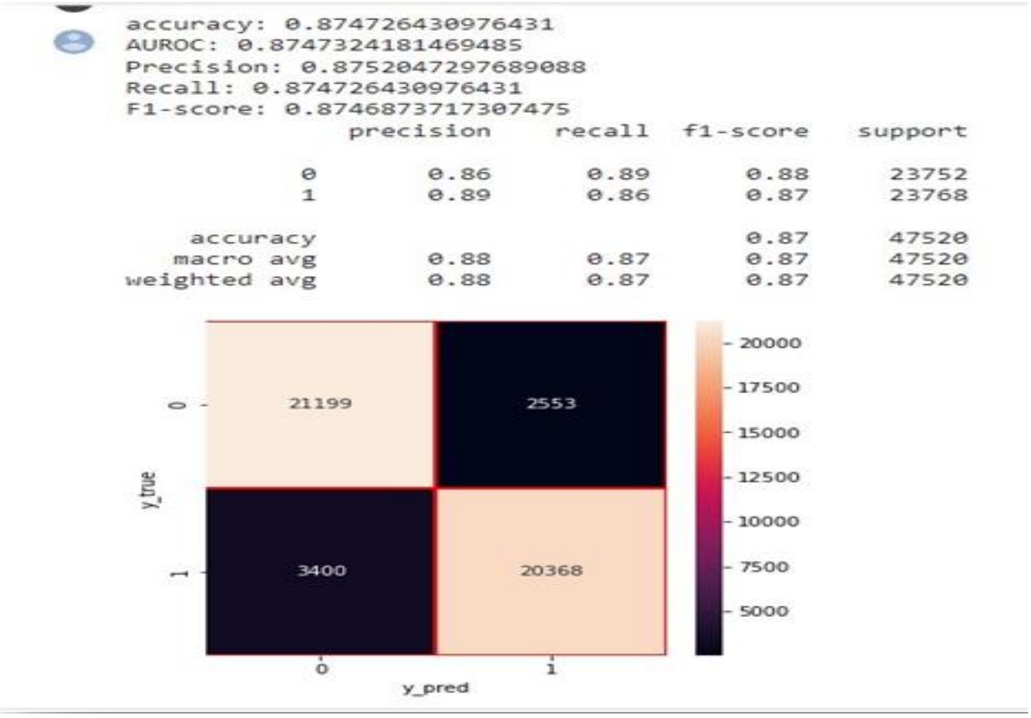


Figure 8.3: The Confusion Matrix of Random Forest Algorithm

9. CONCLUSION

This paper proposes a work to classify & predict COVID-19 using Naive Bayes and Random Forest classifier. The accuracy of the classification model using Naïve Bayes is 87.39%. Random forest classifier gives an accuracy of 87.47%. From the table 8.1, Random Forest classifier is the best classifier for COVID-19 classification.

REFERENCES

[1] ZuZ Jiang MXuP Chen W et al.,(2020) Corona Virus disease 2019(COVID-19) a perspective from china radiology 296(2): 15-25

[2] Worldometer COVIS Live update29 June 2021 Available online: https://www.worldometer.info/coronovirus/ (accessed on 29 June 2021)

[3] https://www.worldometers.info/coronavirus/

[4] https://www.worldometers.info/coronavirus/#countries

[5] Sugandh Bhatia and Jyoteesh Malhotra (2021) “Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV 2021)" IEEE Xplore Part Number: CFP21ONG-ART; 978-0-7381-1183-4.

[6] Nehal A. Mansour, Ahmed I. Saleh, Mahmoud Badawy, Hesham A. Ali (2022) "Journal of Ambient Intelligence and Humanized Computing (2022)" Volume 13 Pages: 41–73.

[7] L. J. Muhammad, Ebrahem A. Algehyne, Sani Sharif Usman, Abdulkadir Ahmad, Chinmay Chakraborty and I. A. Mohammed (2021) "SN Computer Science (2021)" 2:11.

[8] H Zakiyyah and S Suyanto (2020) “2nd International Conference on Science & Technology (2020 2nd ICoST)-Journal of Physics: Conference Series" 1844 (2021) 012002.

AUTHORS PROFILE

K.S.Padmashree Research Scholar, PG & Research Department of Mathematics, M.D.T. Hindu College, Tirunelveli.

Dr.P.Velmani MCA., M.Phil. Ph.D is currently working as an Assistant Professor of Department of Computer Science, The M.D.T. Hindu College, Tirunelveli since 2007. Earlier she had worked as a Lecturer in G.Venkataswamy Naidu College (SF) Kovilpatti from 1994 to 2007. Her main research area focuses on Wireless Sensor Network, Big Data, Digital Image Processing. She has 30 years of experience in teaching.

Dr.S.Loghambal M.Sc., M.Phil. Ph.D is currently working as an Assistant Professor of Research Department of Mathematics, The M.D.T. Hindu College, Tirunelveli since 2007. Her main research area focuses on Mathematical modeling of ordinary and Partial Differential Equations, Approximation Techniques, Numerical and Mathematical Analysis and Computational Mathematics.