Original Article

Improving the Efficiency in Identification of Sentiments of COVID Patients over Online Social Networks using Novel Naive Bayes Algorithm Comparing Random Forest Algorithm

G. Sri Harshitha[1], Shri Vindhya.A[2]

[1]Research Scholar, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamil Nadu, India. Pincode: 602105.
[2]Project Guide, Corresponding Author, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamil Nadu, India. Pincode: 602105.

Abstract

Aim: The main aim of this research work is to predict the efficiency to recognize emotions of covid patients over social media by comparing novel Naive Bayes algorithm and Random Forest algorithm. **Materials and Methods:** Naive Bayes algorithm with sample size = 10 and Random Forest algorithm with sample size = 10 were evaluated many times to predict the efficiency percentage with confidence interval of 95% and G-power (Value=0.8). **Results and Discussions**: Naive Bayes algorithm has proven better efficiency (54%) when compared to Random Forest efficiency (25%). The results achieved with significance value p=0.776 (p>0.05) shows that two groups are statistically insignificant. **Conclusion:** Naive Bayes algorithm performed significantly better than the Random Forest algorithm.

Keywords: Novel Naive Bayes Algorithm, Random Forest Algorithm, Covid, Emotions, Social Media, Machine Learning.

DOI: 10.47750/pnr.2022.13.S04.071

INTRODUCTION

The aim of this research work is to predict the efficiency to find out people's emotions who faced covid via social media by comparing Novel Naive Bayes algorithm and Random Forest. (Bokaee Nezhad and Deihimi 2022) Nowadays most of the people are affected by covid and it is increasing day by day. This sickness has had profound effects on people in both explicit and tacit ways. Coronavirus infection and fatality rates are increasing by the day. The term "pandemic" refers to (Yoon et al. 2022) a disease that spreads rapidly and engulfs a geographic region, such as a country or the entire planet. One of the major catastrophes of modern times is the never-ending COVID-19 pandemic. COVID-19 is a contagious disease spread through contact, coughing, sneezing, or talking to an affected person (Perez-Cepeda and Arias-Bolzmann 2022). Due to confusing information distributed on social media, it is now regarded as a new source of stress, despair, and worry for people. False information about COVID-19 spread on social media has a direct impact on a person's mental health. As time is passing, a lot of research data and patient (Lanyi et al. 2021) records are available in hospitals. There are a lot of open sources for accessing the patients information and researches can be conducted so that various technologies could be used for doing the correct diagnosis of the patients and detect the diseases to stop it from (Costola et al. 2020) becoming critical. Machine learning is playing a major and huge role in the medical industry. It can use different machine learning algorithms to predict and classify diseases (Cabezas et al. 2021). Algorithms can be trained, knowledge and medical records can be transformed and analyzed for better predictions with the improved efficiency. The techniques and algorithms can be directly used on a dataset. The applications of Naive Bayes are: filtering spam, classifying documents, sentiment prediction. The applications of Random Forest are: Banking Industry, Healthcare and Medicine.

The effect of Individuals' lives are greatly influenced by social media, which connects them to the rest of the globe. It is impossible (Mc Mahon 2020) to operate without access to social media to keep up with all the latest news and updates, such as coronavirus updates, stock market updates, and other topics. Over more than 250 articles are under IEEE xplore are related to covid patients emotions. People nowadays rely more on posts and (Cummins 2022) tweets shared on social media platforms such as Instagram, Facebook, and Twitter. It is expected

that social media posts will direct individuals to receive accurate and reliable information. However, in many cases, as with the (Wang et al. 2020)) material that was propagated on social media, the information prompted people to make incorrect conclusions. People's minds had already been upset by the coronavirus; now, comments and tweets about COVID-19 are unsettling and a source of worry that needs to be addressed in order to deal with misleading information from many sources. The (Omer et al. 2022) key idea of this research is that people should avoid releasing data on social media since it could stifle the impact at some point during an emergency. People must accept responsibility for sharing material that is beneficial to the greater population.

Our team has extensive knowledge and research experience that has translate into high quality publications(Bhansali et al. 2021; Jayanth et al. 2021; Sudhakar, Ravel, and Perumal 2021; Sathiyamoorthi et al. 2021; Deepanraj et al. 2021; Raju et al. 2021; Arun Prakash et al. 2020; Kamath et al. 2020; Shanmugam et al. 2021; Rajasekaran et al. 2020; Adhinarayanan et al. 2020; Rajesh et al. 2020; Aurtherson et al. 2021). The disadvantage of Naive Bayes is that it is less accurate when it comes to detecting emotions in frames. The main aim of our proposed system is to improve efficiency in recognising the emotions of the people who suffered with covid. The implication of the research and the future direction of the research is illustrated in the final section (Terroso-Saenz et al. 2022) in a discussion. One of the potential dataset resources is social media. On social media, people may express their disagreements, emotions, opinions, and stories in a vivid and natural way.

MATERIALS AND METHODS

The research setting of the proposed work is done in Saveetha School of Engineering. The number of groups identified for the research is two. Group 1 is the Novel Naive Bayes algorithm and Group 2 is Random Forest algorithm. The Naive Bayes and Random Forest algorithm are iterated at different times (Kotu and Deshpande 2014) with a sample size of 20, 95% confidence interval and pretest power of 80%. The Affective Tweets dataset was trained using machine learning algorithms. Table 1 demonstrates the dataset profile. (Mc Mahon 2020) Fig. 3. shows (Chidambaram et al. 2022) the procedure of the proposed model. Eliminating noisy data, tokenization, stop-word filtering, token length filtering, stemming, case transformation, and extracting text features from the pre-processed data are used for the following process. (Swetland et al. 2021) After dataset collection, the null values and unimportant content in the datasets were removed by preprocessing and data cleaning steps.

The investigation of sentiment analysis has been carried out in this work using the R programming language (Wang et al. 2020) and tweets created during the COVID-19 time period. COVID-19, coronavirus, deaths, new case, recovered, and other hashtag keywords were used to collect Twitter data. Here Novel Naive Bayes algorithm and Random Forest algorithms are used to classify the sentiment ratings into sad, happy, anger and joy categories. The Novel Naive Bayes algorithm and Random Forest algorithm were evaluated (Broodryk and Robinson, n.d.) a different number of times with a sample size of 34 with confidence interval of 95%, and with pretest power of 80% and maximum accepted error is fixed as 0.05. It shows the efficiency between the Novel Naive Bayes algorithm and Random Forest algorithms.

Naive Bayes

The Naive Bayes algorithm is a probabilistic machine learning technique that may be applied to a wide range of classification applications. Filtering spam, categorizing documents, and predicting sentiment are examples of common applications. It is based on Rev's works. That is, changing the value of one feature has no direct impact on the value of the other features employed in the algorithm. Fig. 1. shows the complete procedure (Omer et al. 2022) of the proposed model. The emotion recognition model follows the following procedure.

Random Forest

Random forest is a supervised machine learning algorithm that is commonly used to solve classification and regression problems. It creates decision trees from various samples, using the majority vote for classification and the (Saxena 2020) average for regression. However, it is usually used to tackle classification problems. Fig. 2. shows the complete procedure of the proposed model. The emotion recognition model follows the following procedure.

Statistical Analysis

For statistical analysis, IBM SPSS is used. User sentiments (Horton 2020) are the independent variable, whereas text emotions are the dependent variable. The T-Test analysis is carried out independently.

RESULTS

The dataset for this work comes from the IEEE data port, where Twitter data regarding COVID-19 is readily available. The IDs and sentiment ratings of tweets on the COVID-19 epidemic on Twitter are included in this dataset.The sentiment analysis was performed on a variety of tweets in order to determine the sentiment scores

associated with the data. The analysis is conducted out of 20 tweets first. Following that, over 50 tweets took 250 tweets, then over 500 tweets, and divided them into several (Melenli and Topkaya 2020) emotion value groups. Table 1. demonstrates the dataset profile of various tweets. Table 2. represents the simulated efficiency analysis of the Novel Naive Bayes algorithm and Random Forest algorithms. Table 3. represents group statistical analysis with the mean values for Novel Naive Bayes algorithm and Random Forest algorithms respectively. Table 4 represents the independent T-test analysis of both the groups with insignificance value p= 0.776 (p>0.05) states that both groups are statistically insignificant.
From the results Fig. 4. shows the bar graph analysis based on efficiency of two algorithms and it is inferred that the Novel Naive Bayes object detection algorithm is more efficient than the Random Forest algorithm.

DISCUSSION

In this research paper it resulted that the Naive Bayes algorithm has better efficiency (54%) when compared to Random Forest efficiency (25%). The model proposed resulted in achieving more than 33% increase of efficiency compared to the existing model. The Naive Bayes algorithm appears to give outcomes that are more consistent and with lower standard deviation.
In this research process the Novel Naive Bayes algorithm and Random Forest algorithms-based on sentiment analysis which clarifies the tweets based on their sentiment analysis. The proposed model for this (Abdelminaam et al. 2021) paper is during the experiment, a sample of 20, 50, and 250 tweets are considered. Bar graphs are used to display the collected results. During the COVID-19 pandemic, (Pratama and Tjahyanto 2022) the total number of tweet sentiments is categorized for the most often occurring term in the Twitter data displayed below (table 1). The results are shown in Table 1. A similar work is used to predict the patients emotions by (To et al. 2021) using the algorithms. (Bangyal et al. 2021) On sparse and unstructured data, the Random Forest does not perform as well. The two algorithms, the Novel Naive Bayes algorithm and Random Forest, to find the efficiency for the algorithms. The average success rate for the classification of (Aljameel et al. 2020) people's emotions was found. There are no such opposite findings of existing tweets detection for predicting sentiments of people.
As a result, the findings of the inquiry reveal improved performance in both experimental and statistical analyses. The limitations are that the model will give it zero probability and will be unable to forecast anything in this regard. In the future, this work could be improved by developing a model that recognises people's emotions.

CONCLUSION

In this research work, prediction of efficiency to recognize emotions of covid patients over social media. Overall results, The Novel Naive Bayes algorithm achieved the best efficiency (54%). In contrast, the lowest performance was achieved by the Random Forest efficiency score (25%).

Declarations
Conflict of Interest
The author declares no conflict of interest.

Authors Contribution
Author GSH was involved in data collection, data analysis, and manuscript writing. Author SVA was involved in conceptualization, data validation, and critical review of manuscript.

Acknowledgement
The authors would like to express their gratitude towards Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences (Formerly known as Saveetha University) for providing the necessary infrastructure to carry out this work successfully.

Funding
We thank the following organizations for providing financial support that enabled us to complete the research.
1.Cyclotron Technologies, Chennai, Tamilnadu.
2.Saveetha University.
3.Saveetha Institute of Medical and Technical Sciences.
4.Saveetha School of Engineering.

REFERENCES

1. Abdelminaam, Diaa Salama, Fatma Helmy Ismail, Mohamed Taha, Ahmed Taha, Essam H. Houssein, and Ayman Nabil. 2021. "CoAID-DEEP: An Optimized Intelligent Framework for Automated Detecting COVID-19 Misleading Information on Twitter." IEEE Access :

Practical Innovations, Open Solutions 9 (February): 27840–67.
2. Adhinarayanan, Rajesh, Aravindh Ramakrishnan, Gopal Kaliyaperumal, Melvinvíctor De Poures, Rajesh Kumar Babu, and Damodharan Dillikannan. 2020. "Comparative Analysis on the Effect of 1-Decanol and Di-N-Butyl Ether as Additive with diesel/LDPE Blends in Compression Ignition Engine." Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, June, 1–18.
3. Aljameel, Sumayh S., Dina A. Alabbad, Norah A. Alzahrani, Shouq M. Alqarni, Fatimah A. Alamoudi, Lana M. Babili, Somiah K. Aljaafary, and Fatima M. Alshamrani. 2020. "A Sentiment Analysis Approach to Predict an Individual's Awareness of the Precautionary Procedures to Prevent COVID-19 Outbreaks in Saudi Arabia." International Journal of Environmental Research and Public Health 18 (1). https://doi.org/10.3390/ijerph18010218.
4. Arun Prakash, V. R., J. Francis Xavier, G. Ramesh, T. Maridurai, K. Siva Kumar, and R. Blessing Sam Raj. 2020. "Mechanical, Thermal and Fatigue Behaviour of Surface-Treated Novel Caryota Urens Fibre–reinforced Epoxy Composite." Biomass Conversion and Biorefinery, August. https://doi.org/10.1007/s13399-020-00938-0.
5. Aurtherson, P. Babu, Bhanu Teja Nalla, Karthikeyan Srinivasan, Kulmani Mehar, and Yuvarajan Devarajan. 2021. "Biofuel Production from Novel Prunus Domestica Kernel Oil: Process Optimization Technique." Biomass Conversion and Biorefinery, May. https://doi.org/10.1007/s13399-021-01551-5.
6. Bangyal, Waqas Haider, Rukhma Qasim, Najeeb Ur Rehman, Zeeshan Ahmad, Hafsa Dar, Laiqa Rukhsar, Zahra Aman, and Jamil Ahmad. 2021. "Detection of Fake News Text Classification on COVID-19 Using Deep Learning Approaches." Computational and Mathematical Methods in Medicine 2021 (November): 5514220.
7. Bhansali, Karan J., Kamlesh R. Balinge, Subodh U. Raut, Shubham A. Deshmukh, M. Senthil Kumar, C. Ramesh Kumar, and Pundlik R. Bhagat. 2021. "Visible Light Assisted Sulfonic Acid-Functionalized Porphyrin Comprising Benzimidazolium Moiety for Photocatalytic Transesterification of Castor Oil." Fuel 304 (November): 121490.
8. Bokaee Nezhad, Zahra, and Mohammad Ali Deihimi. 2022. "Twitter Sentiment Analysis from Iran about COVID 19 Vaccine." Diabetes & Metabolic Syndrome 16 (1): 102367.
9. Broodryk, Terise, and Kealagh Robinson. n.d. "Locking down Behavioural Compliance: The Role of Emotion and Emotion Regulation in Breaches of COVID-19 Mitigation Measures." https://doi.org/10.31234/osf.io/mvt4e.
10. Cabezas, Javier, Daniela Moctezuma, Alberto Fernández-Isabel, and Isaac Martín de Diego. 2021. "Detecting Emotional Evolution on Twitter during the COVID-19 Pandemic Using Text Analysis." International Journal of Environmental Research and Public Health 18 (13). https://doi.org/10.3390/ijerph18136981.
11. Chidambaram, Vignesh, Amudha Kumar, Giuseppe Calcaterra, and Jawahar L. Mehta. 2022. "Persistent Cardiac Injury - An Important Component of Long COVID-19 Syndrome." EBioMedicine 77 (February): 103892.
12. Costola, Michele, Michael Nofer, Oliver Hinz, and Loriana Pelizzon. 2020. Machine Learning Sentiment Analysis, Covid-19 News and Stock Market Reactions.
13. Cummins, Jack Alexander. 2022. "Getting a Vaccine, Jab, or Vax Is More Than a Regular Expression. Comment on 'COVID-19 Vaccine-Related Discussion on Twitter: Topic Modeling and Sentiment Analysis.'" Journal of Medical Internet Research.
14. Deepanraj, B., N. Senthilkumar, D. Mala, and A. Sathiamourthy. 2021. "Cashew Nut Shell Liquid as Alternate Fuel for CI Engine—optimization Approach for Performance Improvement." Biomass Conversion and Biorefinery, February. https://doi.org/10.1007/s13399-021-01312-4.
15. Horton, Richard. 2020. The COVID-19 Catastrophe: What's Gone Wrong and How to Stop It Happening Again. John Wiley & Sons.
16. Jayanth, Bellappu Venkat, Melvin Victor Depoures, Gopal Kaliyaperumal, Damodharan Dillikannan, Dilipsingh Jawahar, Kumaran Palani, and Ganesha Prasad Meravanigee Shivappa. 2021. "A Comprehensive Study on the Effects of Multiple Injection Strategies and Exhaust Gas Recirculation on Diesel Engine Characteristics That Utilize Waste High Density Polyethylene Oil." Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, June, 1–18.
17. Kamath, Manjunath, Subha Krishna Rao, Jaison, Sridhar, Kasthuri, Gopinath, Sivaperumal, and Shantanu Patil. 2020. "Melatonin Delivery from PCL Scaffold Enhances Glycosaminoglycans Deposition in Human Chondrocytes – Bioactive Scaffold Model for Cartilage Regeneration." Process Biochemistry 99 (December): 36–47.
18. Kotu, Vijay, and Bala Deshpande. 2014. Predictive Analytics and Data Mining: Concepts and Practice with RapidMiner. Morgan Kaufmann.
19. Lanyi, Katherine, Rhiannon Green, Dawn Craig, and Christopher Marshall. 2021. "COVID-19 Vaccine Hesitancy: Analysing Twitter to Identify Barriers to Vaccination in a Low Uptake Region of the UK." Frontiers in Digital Health 3: 804855.
20. Mc Mahon, Ciarán. 2020. Psychological Insights for Understanding COVID-19 and Media and Technology. Routledge.
21. Melenli, Sadettin, and Aylin Topkaya. 2020. "Real-Time Maintaining of Social Distance in Covid-19 Environment Using Image Processing and Big Data." October 2020. https://ieeexplore.ieee.org/document/9259891.
22. Omer, Mazin, Shadi Al-Afif, Kathrin Machetanz, Christoph Bettag, Marcos Tatagiba, Veit Rohde, and Joachim K. Krauss. 2022. "Impact of COVID-19 on the Neurosurgical Resident Training Program: An Early Experience." Journal of Neurological Surgery. Part A, Central European Neurosurgery, February. https://doi.org/10.1055/s-0042-1743108.
23. Perez-Cepeda, Maximiliano, and Leopoldo G. Arias-Bolzmann. 2022. "Sociocultural Factors during COVID-19 Pandemic: Information Consumption on Twitter." Journal of Business Research 140 (February): 384–93.
24. Pratama, Rivanda Putra, and Aris Tjahyanto. 2022. "The Influence of Fake Accounts on Sentiment Analysis Related to COVID-19 in Indonesia." Procedia Computer Science 197 (January): 143–50.
25. Rajasekaran, S., D. Damodharan, K. Gopal, B. Rajesh Kumar, and Melvin Victor De Poures. 2020. "Collective Influence of 1-Decanol Addition, Injection Pressure and EGR on Diesel Engine Characteristics Fueled with diesel/LDPE Oil Blends." Fuel 277 (October): 118166.
26. Rajesh, A., K. Gopal, De Poures Melvin Victor, B. Rajesh Kumar, A. P. Sathiyagnanam, and D. Damodharan. 2020. "Effect of Anisole Addition to Waste Cooking Oil Methyl Ester on Combustion, Emission and Performance Characteristics of a DI Diesel Engine without Any Modifications." Fuel 278 (October): 118315.
27. Raju, P., K. Raja, K. Lingadurai, T. Maridurai, and S. C. Prasanna. 2021. "Glass/Caryota Urens Hybridized Fibre-Reinforced nanoclay/SiC Toughened Epoxy Hybrid Composite: Mechanical, Drop Load Impact, Hydrophobicity and Fatigue Behaviour." Biomass Conversion and Biorefinery, March. https://doi.org/10.1007/s13399-021-01427-8.
28. Sathiyamoorthi, Ramalingam, Gomathinayakam Sankaranarayanan, Dinesh Babu Munuswamy, and Yuvarajan Devarajan. 2021. "Experimental Study of Spray Analysis for Palmarosa Biodiesel-diesel Blends in a Constant Volume Chamber." Environmental Progress & Sustainable Energy 40 (6). https://doi.org/10.1002/ep.13696.
29. Saxena, Shailendra K. 2020. Coronavirus Disease 2019 (COVID-19): Epidemiology, Pathogenesis, Diagnosis, and Therapeutics. Springer Nature.
30. Shanmugam, Rajasekaran, Damodharan Dillikannan, Gopal Kaliyaperumal, Melvin Victor De Poures, and Rajesh Kumar Babu. 2021. "A Comprehensive Study on the Effects of 1-Decanol, Compression Ratio and Exhaust Gas Recirculation on Diesel Engine Characteristics Powered with Low Density Polyethylene Oil." Energy Sources, Part A: Recovery, Utilization, and Environmental Effects

43 (23): 3064–81.
31. Sudhakar, M. P., Merlyn Ravel, and K. Perumal. 2021. "Pretreatment and Process Optimization of Bioethanol Production from Spent Biomass of Ganoderma Lucidum Using Saccharomyces Cerevisiae." Fuel 306 (December): 121680.
32. Swetland, Sarah B., Ava N. Rothrock, Halle Andris, Bennett Davis, Linh Nguyen, Phil Davis, and Steven G. Rothrock. 2021. "Accuracy of Health-Related Information Regarding COVID-19 on Twitter during a Global Pandemic." World Medical & Health Policy, July. https://doi.org/10.1002/wmh3.468.
33. Terroso-Saenz, Fernando, Andres Muñoz, Francisco Arcas, and Manuel Curado. 2022. "An Analysis of Twitter as a Relevant Human Mobility Proxy: A Comparative Approach in Spain during the COVID-19 Pandemic." GeoInformatica, February, 1–30.
34. To, Quyen G., Kien G. To, Van-Anh N. Huynh, Nhung T. Q. Nguyen, Diep T. N. Ngo, Stephanie J. Alley, Anh N. Q. Tran, et al. 2021. "Applying Machine Learning to Identify Anti-Vaccination Tweets during the COVID-19 Pandemic." International Journal of Environmental Research and Public Health 18 (8). https://doi.org/10.3390/ijerph18084069.
35. Wang, Chen, Peter W. Horby, Frederick G. Hayden, and George F. Gao. 2020. "A Novel Coronavirus Outbreak of Global Health Concern." The Lancet. https://doi.org/10.1016/s0140-6736(20)30185-9.
36. Yoon, Sunmoo, Peter Broadwell, Carmela Alcantara, Nicole Davis, Haeyoung Lee, Amanda Bristol, Dante Tipiani, Joo Young Nho, and Mary Mittelman. 2022. "Analyzing Topics and Sentiments from Twitter to Gain Insights to Refine Interventions for Family Caregivers of Persons with Alzheimer's Disease and Related Dementias (ADRD) During COVID-19 Pandemic." Studies in Health Technology and Informatics 289 (January): 170–73.

Tables and Figures

Table 1. Dataset Name, Dataset Emotions, Dataset Source

S.NO	DATASET NAME	DATASET EMOTIONS	DATASET SOURCE
0	3204	sad .	agree the poor in india are treated badly their poors seek a living in singapore and are treated like citizens they are given free medical treatment given food daily sim cards to call home to tell their family that they are fine if covid 19 case treated foc in hospitals
1	1431	joy	if only i could have spent the with this cutie vc sakshi_s i n g h coast crossing republik
2	4624	fear	joe biden's coronavirus web address lands on a donation page for his campaign he is profiting from fear demopos
3	4646	anger	maybe if i bolt my front door shut coronavirus will stay out
4	2530	sad	The coronavirus disappearing in Italy shows this to "intellectuals" who say lockdowns do not work.

Table 2. Efficiency of Naive Bayes and Random Forest. The Naive Bayes algorithm is 29% more Efficient than the Random Forest algorithm.

ITERATION NO	Naive Bayes	Random Forest
1	68.00	38.00
2	65.30	35.00

3	61.00	31.20
4	59.30	29.40
5	56.20	26.10
6	51.00	23.00
7	49.00	19.00
8	47.00	17.30
9	46.20	15.20
10	39.10	13.10

Table 3. Group Statistics of Naive Bayes and Random Forest algorithm with the mean value of 54.2% and 24.7 %

Group	**N**	**Mean**	**Std.deviation**	**Std.Error Mean**
Naive Bayes	10	54.2100	9.25580	2.92694
Random Forest	10	24.7300	8.59083	2.71666

Table 4. Independent sample T-test is performed for the two groups for significance and standard error determination. The significance value p=0.776 (p>0.05) shows that two groups are statistically insignificant.

		Levene's Test for Equality of Variance		**T-test for Equality of Means**						
									95% Confidence Interval of the Difference	
	Equal Variance	**F**	**Sig**	**t**	**df**	**Sig (2-tailed)**	**Mean Difference**	**Std.Error Difference**	**Lower**	**Upper**

Efficiency	**Assumed**	.083	.776	7.382	18	.001	29.48000	3.93340	21.0901	37.86982
	Not Assumed			7.382	17.901	.001	29.48000	3.93340	21.0868	37.87315

1.Start Program
2.Input the dataset
3.Load the data
4.Give path to Naive Bayes to configure and store it in a variable
5.visualize the data
6. Pre-process the data
1)Remove nan values and unwanted values
2)clearing the stop words using lemitizing
3)Removing the unwanted text by cleaning and stemming process
4)calculating frequent words of emotions(fear,anger,sad,joy)
7.Import Naive Bayes from sklearn.linear_model
8.Divide the data into training and testing sets
9.Train the SVM algorithm
10.Release the tweets
11.End program

Fig. 1. Pseudocode for novel Naive Bayes algorithm.

1.Start Program
2.Input the dataset
3.Explore the data to figure out what they look like
4.Give path to Random Forest to configure and store it in a variable
5.Pre-process the data
6.Import Random Forest from sklearn.metrics
7.Fit train set and test set to random forest
8.Split the data into attributes and tables
9. Divide the data into training and testing sets
10.Train the Naive Bayes algorithm
11.Destroy all windows
12.End program

Fig. 2. Pseudocode for Random Forest algorithm.

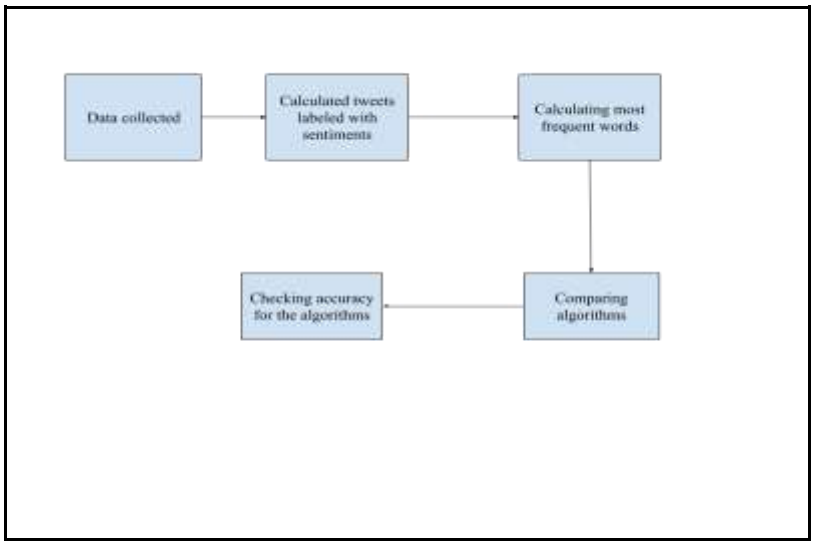

Fig. 3. Architecture for emotions of a covid patient using novel Naive Bayes algorithm and Random Forest from dataset processing to output of each frame.

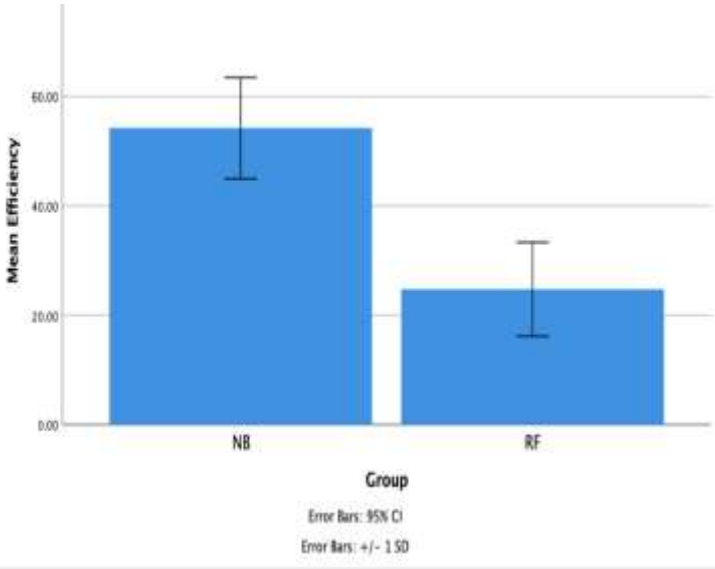

Fig. 4. Bar graph analysis of novel Naive Bayes algorithm and Random Forest algorithm. Graphical representation shows the mean efficiency of 54.2% and 24.7% for the proposed algorithm Naive Bayes and Random Forest respectively. X-axis : Naive Bayes vs Random Forest, Y-axis : Mean precision ± 1 SD.