Predicting Indian Sentiments of COVID-19 Using MLP and Adaboost

Jabeen Sultana[a],M. Usha Rani[b], Shabnam Mohammed Aslam[c], Laila AlMutairi[d]

[a,d] Department of Computer Science, [c] Department of Information Technology,
[a,c,d]Majmaah University, Majmaah-11952, Kingdom of Saudi Arabia
[b] Department of Computer Science, Sri Padmavati Mahila Visva Vidyalayam, Tirupati, 517502, India

Article History: Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 28 April 2021

Abstract: Deep Learning is a subset of AI, acknowledged as graded learning. It depends on neural networks and undergoes different stages of representative transforms. It is widely used in image processing,market-crate investigation and classifying student's performance etc. In present times, deep learning is widely used by the researchers to perform COVID data classificationand clustering and further aids in diagnosing COVID. As this deadly virus is spreading rapidly in India and effected life's of many in various ways, therefore sentiments of public towards this are classified and analyzed. This Research work collected tweets related to impact of COVID-19 in India, sentiments of people towards COVID-19 using twitter API. Preprocessing strategies were imparted to COVID tweets in order to well-ordered tweets dependent on the subject of the tweet whether tweet belongs to the class-positive, negative and neutral.Subsequent to Preprocessing, cleaned tweets and its associated class are prepared, sent for training and a perfect model is achieved, which is further tested by feeding testing data. Outcomes are assessed on few considerations like Accuracy, RMSE, Precision, Recall, F-Score, ROC Curve and Kappa Statistics. It was found that conventional deep learning method-MLP outperforms in classifying the tweets in terms of high accuracy with 97%, low root mean square error-0.12, precision and recall-0.97. The outcomes also specify good F-Score-0.95 and ROC Curve area-0.99.

Keywords: Deep Learning, COVID-19, Classification, Naive Bayes, MLP, Decision Trees, Ada Boost and Twitter.

1. Introduction

In the current times, more focus islaid on the impact of novel corona virus across the globe and therefore it became the hot topic for the researchers all over the world. The novel COVID-19 is a fatal virus devising from Wuhan, China in December-2019. It is also called as Corona Virus (Zhu, Zhang, Wang, Li, Yang, Song, Zhao, Huang, Shi, Lu, & Niu, 2020). The virus responsible for this disease is "SARS-Severe Acute Respiratory Syndrome COV-2" which has caused breathing disorders and death in humans. WHO has declared this spread of virus from Wuhan to other parts of the world as global pandemic (WHO, 2020). This spread of harmful virus affected serious effect on economy all over the World. Diagnosing this disease accurately was approved by examining Chest using radiations and by capturing X-Ray's, this deadly disease can be diagnosed and further treatment can be started (Geneva et. Al, 2020). Current trend has been shifted for seeking safety measures to prevent oneself from being infected by this deadly virus. Right from public places, homes, work environments, schools, colleges, parks, cafes and travel places and many more environments, efficient safety measures are being followed but still somehow this virus spreads causing illness, contaminates the body, and spreads to others.

In the recent times, social media platforms have gained huge popularity for numerous users across the world. Now-a-days, heaps of COVID related tweets are generated in social networking databases, but it remains unused. Twitter, a social networking site allows the users to post text up to 280 characters (Twitter.com, 2006). The text once posted terms as a tweet and usually contains the user's sentiments, thoughts, opinion, information, observations and comments towards a particular topic basing on real time incidents. Mining the tweets, processing and further analyzing them leads to sentiments analysis of public towards COVID-19 and helps in making decisions to combat this Pandemic. As there are limited research works to analyze COVID tweets and this area needs to explored, we thought of predicting sentiment analysis by using deep learning technique on COVID tweets. To obtain the public sentiments and impact of COVID from massive tweets, leading tools and technologies are essential. Classifying the tweets based on public sentiments helps in analyzing and combating this pandemic in various sectors like health, education, stock market prediction etc., to name a few.

In recent times, deep learning techniques have yielded promising results to perform data classification, clustering and rule mining pertaining to different fields of data. Deep Learning bases its structure from the human nervous system and is a kind of artificial neural network (LeCun, Bengio & Hinton, 2015). Following the spread of COVID-19, many research works have been carried out for COVID data classification, clustering and mining rules. Various Models have been suggested to overcome this deadly virus in terms of diagnosing COVID-19. The researchers started using AI methods in categorizing this COVID patients and are suggesting proper preventive measures along with treatment. A person once gets infected starts showing symptoms like mild illness, pneumonia, breathlessness, fever, cough, sore throat and mild headaches too. It varies from person to person based

on his immunity and emotional strength to combat it. Symptoms get worse within few days and if left untreated, can lead to mortality state. Therefore, proper diagnosis of this virus and suggesting proper treatment with safety measures can help the person in combating this deadly virus.

2. Related work

Opinion mining has increased tremendous consideration and prominence, promising field for research. This area has been chosen up since different interpersonal organizations produced immense information from different organizations, social media and other data driven applications. In these times, specialists are exploring sentiment analysis of students towards learning the subject, foreseeing assumption investigation in learning parcel of understudies by using different methods for information mining, profound learning and NLP procedures. Mining strategies have been utilized to process natural language with few achievements (Sultana, Rani & Farquad, 2020). Finding information from Real time applications, for example, clinical analysis (Sultana & Jilani, 2018) (Sultana, Sadaf, Jilani, & Alabdan, 2019) to associate products using rule mining (Sultana & Nagalaxmi, 2015) and learning environments (Sultana, Rani & Farquad, 2019),require and lean toward information disclosure approaches to comprehend the Classifiers expectation. Conventional method for deep learning was utilized to anticipate understudy's expectations on instructive information. Also, a comparison was carried out with other methods of AI and was found that Multi-Layer Perceptron got the ideal outcomes (Sultana, Nasreen, Yadav, Fayez, 2018). Likewise, informative tweets related to educational sector were broken down in the wake of cleaning the text obtained from twitter and well-ordered and further it was analyzed that deep learning techniques gave best outcomes contrasted with other methods of machine learning (Sultana, Rani & Farquad, 2019) (Sultana, Rani & Farquad, 2020).

In (Ardakani, Kanafi, Acharya, Khadem& Mohammadi, 2020), the authors proposed a fast and legitimate technique for COVID-19 determination utilizing a man-made reasoning based procedure. ResNet-101 and Xception accomplished ideal outcomes and can also identify infected cases with corona virus from non-infected cases resulting an AUC of 0.994. Another model was programmed for COVID-19 discovery utilizing crude chest X-beam pictures. Their model is created to give exact diagnostics to 2-class with an accuracy of 98% and multi-classification with an accuracy of 87% (Ozturk.et al., 2020). The authors distinguished seven huge utilizations of Artificial Intelligence for COVID-19 pandemic. Their work uncovered a significant function of AI in identifying cluster of conditions and to foresee the circumstances influencing because of this deadly virus by assembling and investigating all previous data in the near future (Vaishya et al., 2020).

The authors proposed a determination model prepared on the utilization of multi-see pictures of chest CT. The multi-see profound learning combination model accomplished area under curve with 0.73, precision -0.70, TP-0.73 and TN-0.615 when test set was applied (Wu, Hui, Niu, Li, Wang, He & Zha, 2020).In assessing the model performance, accuracy obtained was 76% followed by area of curve-0.81, TP-0.81 and TN-0.61. This 3-class testing model holds some possibility to im-demonstrate the effectiveness of conclusion and decrease the interest put on radiologists in circumstances such as these. The authors utilized 10 convolutional neural network strategies to recognize a COVID-19 case from a non-COVID-19 case with an incredible precision. ResNet-101 and Xception performed best in the characterization cycle. They additionally tried the presentation of the models' exhibition by their AUC score, Accuracy, TP, TN and accuracy. The consequences of the testing set indicated that ResNet-101 can separate virus infected cases from non-infected cases by yielding accuracy of 99% and area under curve with-0.994 TP and TN with scores of 100% and 99% (Vaishya et al., 2020).

The authors utilized techniques of AI to build up an identification model. They utilized different component extraction strategies like matrix based on gray level, patters based on local directions, discrete wavelet transforms etc. in this work in the wake of creating different arrangements of picture information from the picture set. In an offer to improve the exhibition of the model, 2-crease, 5-overlay and 10-overlap cross approval was utilized in the grouping cycle. The outcome demonstrated that at 10-overlap cross approval and use of GLSZM include extraction strategy, the model attained accuracy of 99.68% (Barstugan, Ozkaya, & Ozturk, 2020).

(Togaçar, Ergen & Zafer, 2020) also developed a AI model to detect the novel corona virus in an earlier manner. They employed Support Vector Machine, a machine learning method, to differentiate virus cases from Pneumonia and COVID-19. Their work was distinct from other related work because they processed the x-ray image data using fuzzy colour technique and imaging stack techniques. The fuzzy colour technique was used to improve the input data (x-ray image files) by reducing the level of blur while the image stack technique was employed to eliminate the noises of the original files hence improving the quality of the image. Deep learning technique Squeezenet was used as well. The achievement of the models was assessed by Precision, F1-score and Accuracy. The three scenarios used were “Squeezenet + original data”, “Squeezenet + structured data (fuzzy technique)” and “Squeezenet + structured data (stacked technique)”. Their overall performance are as follows: 84.5%, 95.58% and 97.06% respectively.

The authors led an exploration dependent on an assessment of COVID-19 asymptomatic disease cases utilizing an AI approach. They received a test system called AI based transmission test system to foresee COVID-19 asymptomatic contaminations. This test system was assembled dependent on the transmission characteristics of COVID-19. In the test system, the contamination was arranged into idle, virus affirmed patients with the inert patients can taint others. The infection progress is isolated into three phases: brooding stage, isolate stage, and confirmation stage. The presentation of the models was assessed by F1-score, Sensitivity, Specificity, Precision and Accuracy. The three situations utilized are "Squeezenet + unique information", "Squeezenet + organized information (fluffy method)" and "Squeezenet + organized information (stacked procedure)". The results attained are as per the following: 84.5%, 95.58% and 97.06% individually (Yu et al., 2020).

As per reviewing the above papers, we thought of classifying and analyzing Indian COVID-19 sentiments and problem faced by public in different walks of life in this research work.

Data Collection and Preprocessing

COVID tweets of India are gathered from a social networking site named as twitter using application programming interface. The extracted text from the tweets are assembled and are in unstructured format as 80% of the text is in unstructured form. Therefore, tweets were extracted and preprocessed well by removing urls, white spaces, hash tags, quotations, duplicate words. Tokenization and stemming is performed using Python in windows systems (Python). Cleaned text is further classified into proper class by calculating TF-IDF values for the sentences in the text.

Suggested Methodology

A method was suggested for classifying and analyzing COVID sentiments of Indian public using Adaboost and Multi-Layer Perceptron. After necessary preprocessing and classifying the text, MLP, Adaboost, Naïve Bayes and decision tree were applied to classify the text comprising Indian Covid-19 data information. MLP attained outstanding classification results in terms of accuracy, precision, recall and F-Score. Adaboost attained promising results followed by decision tree and naive bayes in terms of accuracy, Root Mean Square Error, Precision, Recall, Kappa Statistic and ROC Curve.

Steps for Suggested Methodology

The below five steps describe the suggested framework

1. Indian COVD tweets are extracted using Twitter API
2. Tweets are well preprocessed using python libraries
3. Machine Learning Classifiers like Decision Tree, Naive Bayes, Adaboost and Deep Learning Classifier- are used for training and testing the data with 70% of training data, a model is obtained.
4. Test data set with 30% of actual data is used for testing the obtained model and the results are obtained.
5. Results obtained by the classifiers are evaluated.

The below Figure1 describes the framed work for classifying Indian COVID tweets and analyzing them for suggesting reforms to public.

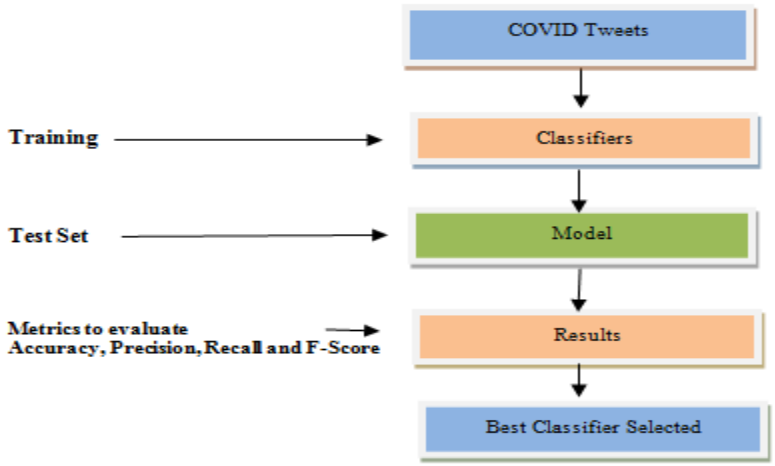

Figure 1. The Proposed framed work for classifying Indian COVID tweets

Methods Used

1. MLP: A multi-layer perceptron is a perceptron with many layers. It has input layer, output layer and hidden layer. The cardinality of hidden layer is n and the cardinality of input and output layer is 1.Each neuron has

its own activation function. It uses back propagation algorithm and passes the input values into the network. Hidden layer performs classification of features, two layers are sufficient to solve any problem, and more layers can be added as per feature selection required when more features exist in a particular data set. Just like hidden layer, output layer passes the classification results and the basic technique for training multilayer perceptron using the back propagation algorithm to count variations using logistic and hyperbolic tangent sigmoid functions [6].

2. Decision Trees: An essential classifier in Machine Learning. C4.5 makes use of the split-and-overcome technique to develop choice trees. These trees start at the base of the tree and go to its leaf hubs. The J48 calculation that utilizes Decision Tree execution is utilized in the trials detailed here. DT is generally acknowledged in dynamic frameworks and is utilized due to its human reasonable structure. A test thing for the class name begins from the base of the tree and travels through it to the leaf hub, which gives the grouping of the example of the instance(Delashmit & Manry, 2005).

3. Decision Tree: C4.5, an essential classifier makes use of the divide and conquer technique in develop building the decision tree. Tree is constructed from the root node, also called as mother node or basic node by selecting the most relevant feature and traverses until the bottom node called as leaf node by selecting subsequent features. It is generally used in dynamic frameworks because of its easy understandable nature among humans. Rules can also be extracted from the trees built by calculating from the top node until bottom nodes. It yields promising results on many of the data sets and is widely used among the researchers (Quinlan, 2014).

4. ADA Boost: It is a broadly utilized method to get results that are accurate and builds an ensemble method by accomplishing several repetitions respectively. It is also called as adaptive boosting. It combines multiple weak learners into a single strong learner and will create different decision trees with a single split called as decisions stumps and they depend upon the number of features present in the data set. For N features, there will be N decision stumps and best decision stump is selected by calculating the Gini index. The best decision stumps evaluated for its accurateness with algorithm. Misclassified instances error rate is calculated, weights are increased, and the correctly classified instances weights are reduced. To normalize the weights, second decision stump is calculated and this process repeats until the best decision stump obtains the best results (Schapire, 2013).

5. Naive Bayes Tree: It is like DT aside at the leaves the probability exists and probabilities of all the leaf nodes will be equivalent to one. It is a crossover of decision tree ordering and uses basics of Bayes rule to from the naive bayes tree. The Bayes rule ascertains the probabilities of each class utilizing the given models. Each property estimation at a given name requires a gauge of the contingent likelihood. The Bayesian classifier finishes arrangement at leaf hubs (Kohavi, 1996).

3. Discussion Of Results

In this section, we analyze the results. The below table-1 describe the results of MLP, Ada Boost, Naïve Bayes and Decision Tree.The results are evaluated based on few parameters like Accuracy, Error rate, Precision, Recall, F-Score, Kappa Statistic and ROC Curve. Result obtained by machine learning classifiers like Decision tree, Adaboost and Naive Bayes are compared with MLP-a classifier of deep learning. Results are well analyzed and shown in the form of graphs.

Table-1 Results of Naive Bayes, Decision Tree, Adaboost and MLP

Performance Measures	**Naive Bayes**	**Decision tree**	**Ada Boost**	**MLP**
Accuracy	73	75	89	97
RMSE	0.32	0.35	0.21	0.12
Precision	0.72	0.74	0.88	0.97
Recall	0.73	0.75	0.89	0.97
F-Score	0.72	0.73	0.89	0.95
ROC	0.98	0.57	0.95	0.99
Kappa Statistic	0.53	0.52	0.80	0.94

Figure 2. describes the comparison of accuracies obtained among the classifiers.MLP outperformed in classifying COVID tweets with highest accuracy of 97% followed by Adaboost, Naive Bayes and Decision tree with accuracies of 89%, 75% and 73%.

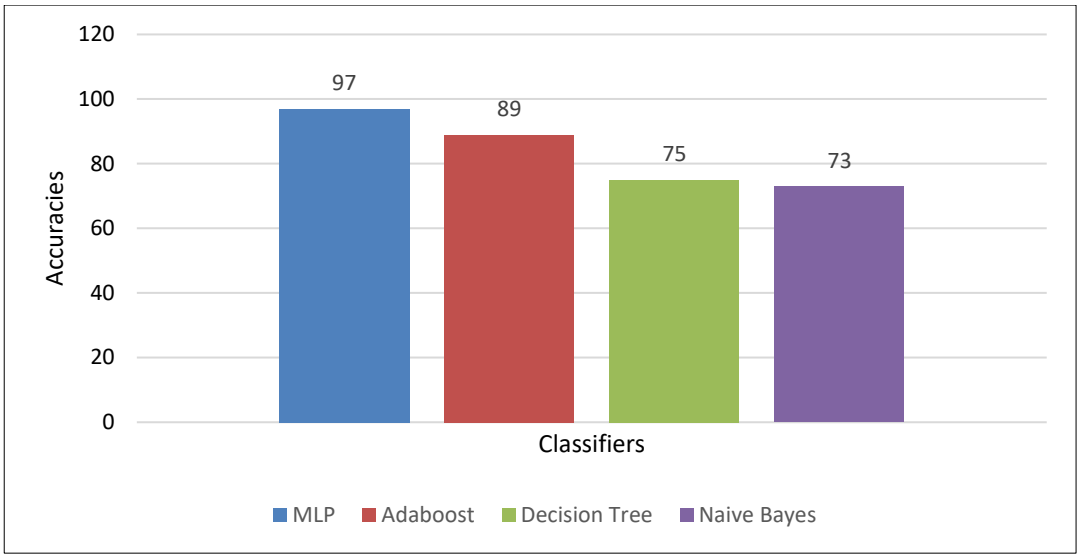

Figure 2. Classification accuracy on COVID Tweets by different classifiers

Figure 3. shows Precision obtained on COVID Tweets from different classifiers. It shows that the MLP obtained highest precision value with 0.97 followed by Adaboost, Naive Bayes and Decision tree with precision values of 0.88, 0.74 and 0.72. Figure 4. presented the recall obtained from COVID Tweets. It shows that the highest recall value was obtained for MLP-0.97 followed by Ada boost, Naive Bayes and decision trees with recall values of-0.89, 0.75 and 0.73.

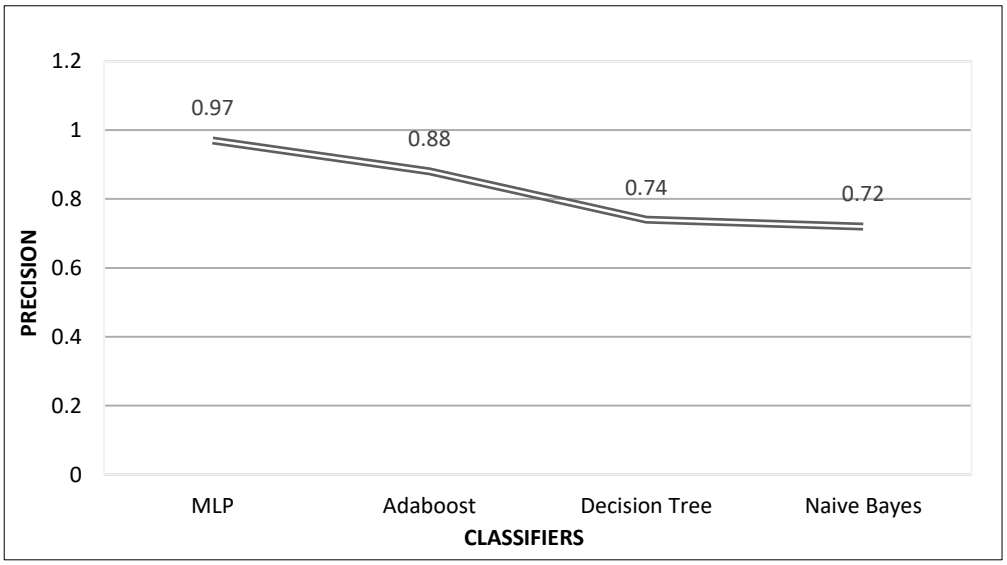

Figure 3. Precision obtained on COVID Tweets by different classifiers

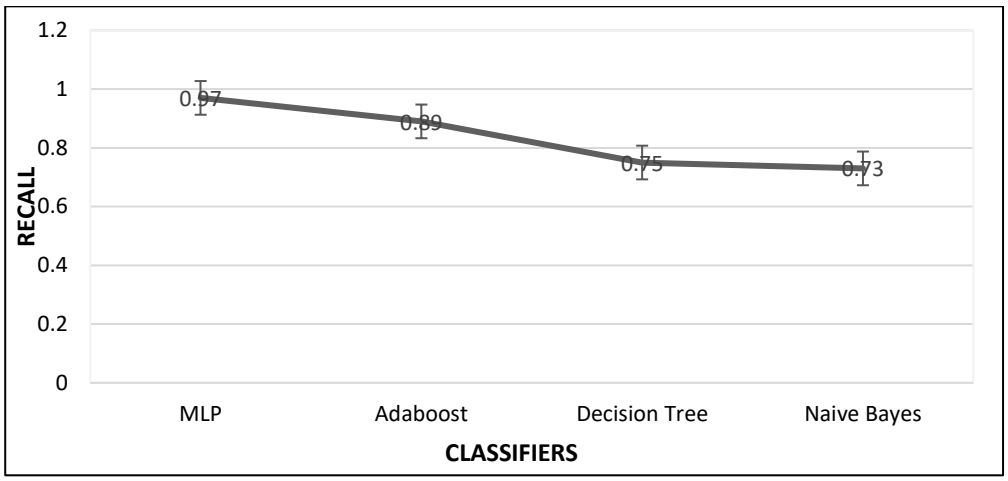

Figure 4. Recall obtained on COVID Tweets by different classifiers

Figure 5below describes the comparison of F-Score values obtained among the classifiers.Highest F-Score was obtained for MLP-0.95 followed by Ada boost, Naive Bayes and decision trees with recall values of-0.89, 0.75 and 0.72.

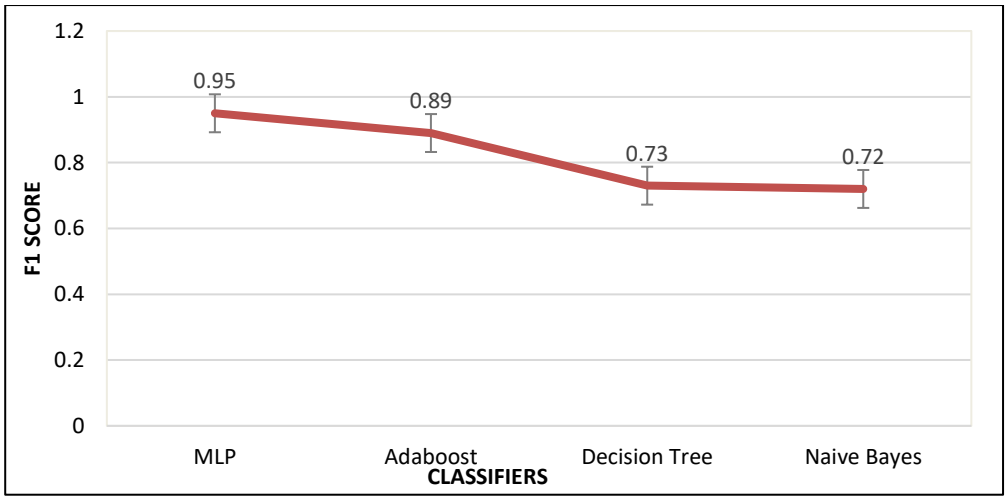

Figure 5. F-Score obtained on COVID Tweets by different classifiers

Figure 6 describes the comparison of Root Mean Square Error obtained among the classifiers.Root mean square error was very low for MLP with 0.12 followed by Adaboost, Naive Bayes and Decision tree with 0.21, 0.35 and 0.32. Figures 7 and 8describe the comparison of ROC Curve values and kappa statistic obtained among the classifiers.

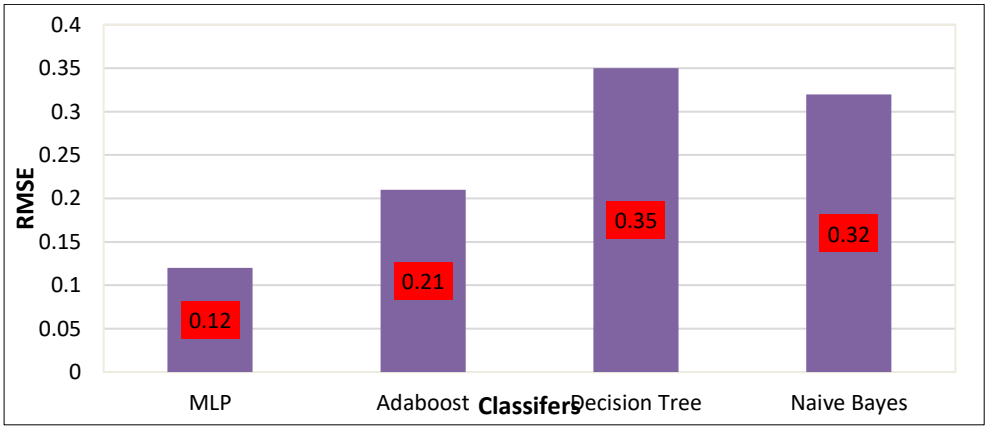

Figure 6. RMSE obtained on COVID Tweets by different classifiers

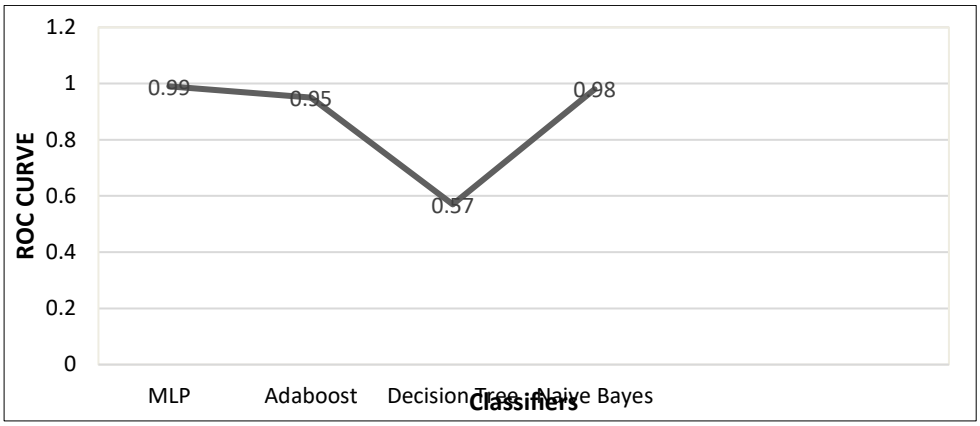

Figure 7. ROC Curve values obtained on COVID Tweets by different classifiers

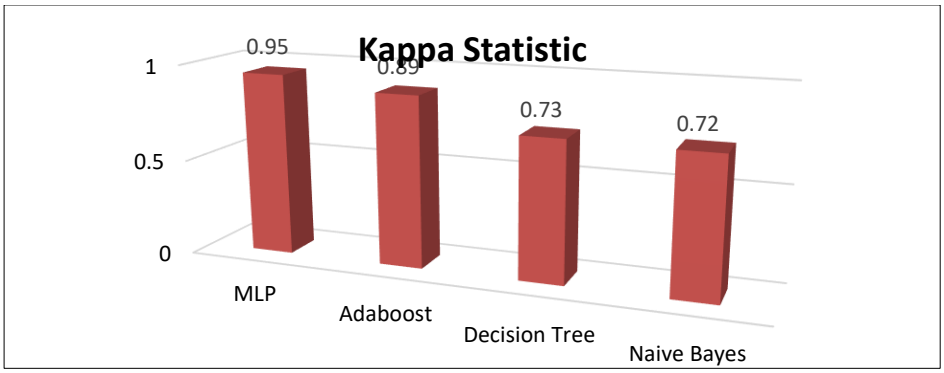

Figure 8. Kappa Statistics obtained on COVID Tweets by different classifiers

Other parameters like ROC curve area and kappa statistic was obtained high by MLP with values of 0.99 and 0.94. Adaboost obtained Roc Curve and kappa statistic as 0.95 and 0.80. Decision tree obtained low ROC curve and kappa statistic with values of 0.57 and 0.52. Naive Bayes obtained ROC curve and kappa statistic with values of 0.98 and 0.53. Overall results were analyzed and found that MLP obtained highest accuracy, precision, recall, F-Score, RMSE, Kappa Statistic and ROC curve followed by Adaboost, Naive Bayes and Decision tree.

The below figures, Figures 9-12) represent ROC Threshold curve where X-axis represents False Positive Rate and Y-axis represents True Positive Rate.

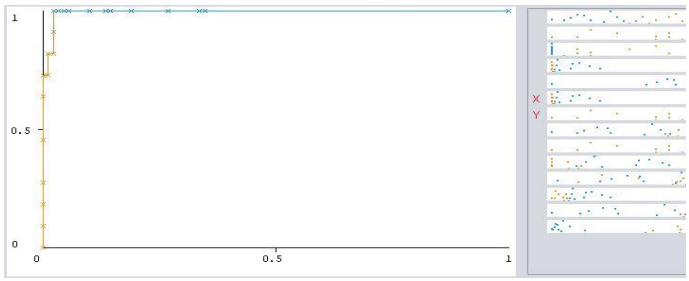

Figure 9.ROC Threshold curve for MLP

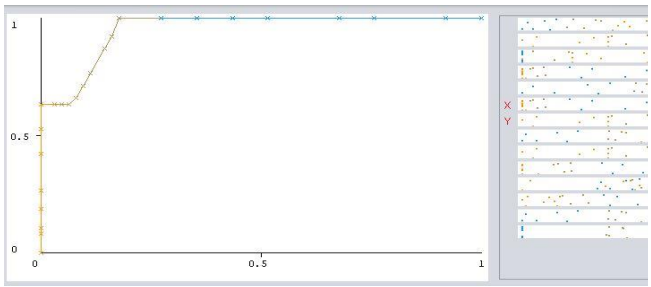

Figure 10. ROC Threshold curve for Adaboost

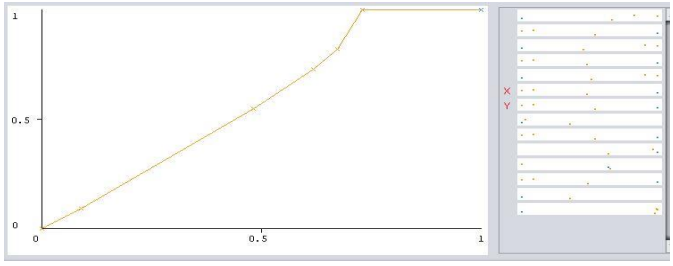

Figure 11. ROC Threshold curve for Decision Tree

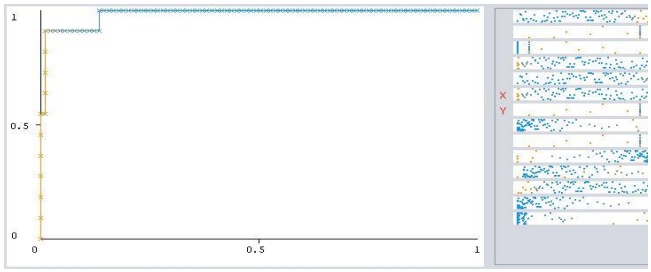

Figure 9.ROC Threshold curve for Naïve Bayes

4. Conclusion

Due to widespread of social media during this COVID-19 pandemic, most of the public have preferred Twitter to share their sentiments related to COVID disease and its spread right form diagnosis to getting cure. Artificial Intelligence is a collection of tools and techniques that can greatly influence the impact of COVID-19 on humanity. The call to integrate AI into the medical front to combat COVID-19 is imperative. In this research work, corona virus related tweets of Indians to analyze the sentiments of Indian people towards this deadly virus and its impact on people. Twitter was accessed using keys and tokens to retrieve Indian sentiments on this deadly virus and how public has overcome over this pandemic. Raw tweets were well preprocessed and well classified using MLP, Adaboost, Naive Bayes and Decision Tress. Outcomes obtained signifies that MLP attains good classification accuracy with 97%, low error rate of 0.12% and with high precision-0.97, recall-0.97, F-Score-0.95 and ROC Curve area-0.99 followed by Adaboost with accuracy of 89%, and Decision Tree and Naive Bayes with accuracies of 75 and 73%. As a result, it is suggested to use these proposed classifiers to classify and diagnose Indian tweets related to COVID-19.

This study shows deep learning techniques can greatly improve the classification rate in classifying the COVID-19 public sentiments. Well analyzed tweets reveal the diagnosis rate and regulate proper treatment in health centers. Well analyzed tweets reveal that the public is avoiding social distancing at rural places, not sanitizing well due to lack of amenities. Private schools in India face fee collection problem affecting the salaries of teachers and other school activities. Some students are unable to learn the full concepts. Students at rural areas of India face technical difficulties and special needs students lack in learning the concepts. In future, we plan to improve the accuracy of Decision Tree, Naive Bayes in classifying COVID tweets..

References

1. Ardakani, A. A., Kanafi, A. R., Acharya, U. R., Khadem, N., & Mohammadi, A. (2020). Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images: Results of 10 convolutional neural networks. Computers in Biology and Medicine, 103795.
2. Barstugan, M., Ozkaya, U., & Ozturk, S. (2020). Coronavirus (covid-19) classification using ct images by machine learning methods. arXiv preprint arXiv:2003.09424.
3. Delashmit, W. H., & Manry, M. T. (2005, May). Recent developments in multilayer perceptron neural networks. In Proceedings of the seventh Annual Memphis Area Engineering and Science Conference, MAESC.
4. Kohavi, R. (1996, August). Scaling up the accuracy of naive-bayes classifiers: A decision-tree hybrid. In Kdd (Vol. 96, pp. 202-207).
5. LeCun, Y., Bengio, Y. and Hinton, G. (2015) Deep Learning. Nature , 521, 436-444.
6. Ozturk, T., Talo, M., Yildirim, E.A., Baloglu, U.B., Yildirim, O. and Rajendra, U.A. (2020) Automated Detection of COVID-19 Cases Using Deep Neural Networks with X-Ray Images. Computers in Biology and Medicine , 121, Article ID: 103792.
7. Python Software Foundation. https://www.python.org/.
8. Quinlan, J. R. (2014). C4. 5: programs for machine learning. Elsevier.
9. Schapire, R. E. (2013). Explaining adaboost. In Empirical inference (pp. 37-52). Springer, Berlin, Heidelberg.
10. Sultana, J., Rani, M. U., & Farquad, M. A. H. (2020). An Extensive Survey on Some Deep-Learning Applications. In Emerging Research in Data Engineering Systems and Computer Communications (pp. 511-519). Springer, Singapore.
11. Sultana, J., & Jilani, A. K. (2018). Predicting Breast Cancer using Logistic Regression and Multi-Class Classifiers. International Journal of Engineering & Technology, 7(4.20), 22-26.
12. Sultana, J., Sadaf, K., Jilani, A. K., & Alabdan, R. (2019, December). Diagnosing Breast Cancer using Support Vector Machine and Multi-Classifiers. In 2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE) (pp. 449-451). IEEE.
13. Sultana, J., & Nagalaxmi, G. (2015). How Efficient is Apriori: A Comparative Analysis. International Journal of Current Engineering and Scientific Research, ISSN (PRINT): 2393-8374,(ONLINE): 2394-0697, 2(8), 91-99.
14. Sultana, J., Rani, M. U., & Farquad, M. A. H. (2019, November). Knowledge Discovery from Recommender Systems using Deep Learning. In 2019 International Conference on Smart Systems and Inventive Technology (ICSSIT) (pp. 1074-1078). IEEE.
15. Sultana, J., Sultana, N., Yadav, K., & AlFayez, F. (2018, April). Prediction of sentiment analysis on educational data based on deep learning approach. In 2018 21st Saudi Computer Society National Computer Conference (NCC) (pp. 1-5). IEEE.

16. Sultana, J., Rani, M. U., & Farquad, M. A. H. (2019, December). Deep Learning Based Recommender System Using Sentiment Analysis to Reform Indian Education. In International Conference On Computational and Bio Engineering (pp. 143-150). Springer, Cham.
17. Sultana, M. J., Rani, M. U., & Farquad, M. A. H. (2020). Sentiment Analysis based Recommender System for Reforming Indian Education using Multi-Classifiers.
18. Toğaçar, M., Ergen, B., & Cömert, Z. (2020). COVID-19 detection using deep learning models to exploit Social Mimic Optimization and structured chest X-ray images using fuzzy color and stacking approaches. Computers in Biology and Medicine, 103805.
19. Twitter.Com, Twitter Developer. https://dev.twitter.com/streaming/overview, Founded on 21st March, 2006.
20. Vaishya, R., Javaid, M., Khan, I. H., & Haleem, A. (2020). Artificial Intelligence (AI) applications for COVID-19 pandemic. Diabetes & Metabolic Syndrome: Clinical Research & Reviews.
21. Vaishya, R., et al . (2020) Artificial Intelligence (AI) Applications for COVID-19 Pandemic. Diabetology & Metabolic Syndrome, 14, 337-339.
22. WHO (2020) Novel Coronavirus. (2019-nCoV) Technical Guidance: Laboratory Guidance, Situation Report1, Jnauary 2020.
23. WHO, Geneva.et al . (2020) A Novel Coronavirus from Patients with Pneumonia in China, 2019. The New England Journal of Medicine , 382, 727-733.
24. Wu, X., Hui, H., Niu, M., Li, L., Wang, L., He, B., ... & Zha, Y. (2020). Deep learning-based multi-view fusion model for screening 2019 novel coronavirus pneumonia: a multicentre study. European Journal of Radiology, 109041.
25. Yu, Y., Liu, Y. R., Luo, F. M., Tu, W. W., Zhan, D. C., Yu, G., & Zhou, Z. H. (2020). COVID-19 Asymptomatic Infection Estimation. medRxiv.
26. Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Zhao, X., Huang, B., Shi, W., Lu, R. & Niu, P. (2020). A novel coronavirus from patients with pneumonia in China, 2019. New England Journal of Medicine.