MODERN HUMAN CARE SERVICES

Divya M. Chaudhari
Computer Science and Engineering
RNGPIT, Bardoli,

Twinkal A. Mahida
Computer Science and Engineering
RNGPIT, Bardoli, Surat.

Mr. A.D.Prajapati
Assistant Prof.
Computer Science and Engineering
RNGPIT, Bardoli, Surat.

Shreya V. Parekh
Computer Science and Engineering
RNGPIT, Bardoli, Surat.

Vishal B. Prajapat
Computer Science andEngineering
RNGPIT, Bardoli, Surat.

Abstract *-nowaveryfamousepidemicofCOVID-19is* ***going on. The epidemic is most prevalent where humans are overcrowded and come in contact with each other so all people are worried about to current situation. like some people go outside and come to home they are worried about corona virus and second thing is humans are not aware about COVID- 19 medicine so, in our Modern human care services application we propose to develop an application which can predict the vulnerability of a COVID-19 given basic symptoms like fever, cough, diarrhea etc. and we also give COVID-19 related medicine information to take during corona risk.The ensemble learning approach is one of the most reliabletechniquesforpredictingresults.***

Key Words: *classification algorithms, regression algorithms, covid-19,prediction*

I. INTRODUCTION

In December 2019, the novel coronavirus appeared in the Wuhan city of China [1] and was reported to the World Health Organization (W.H.O) on 31st December 2019. The virus created a global threat and was named as COVID-19 by W.H.O on 11th February 2020 [1]. The COVID-19 is the family of viruses including SARS, ARDS. W.H.O declared this outbreak as a public health emergency [1] and mentioned the following; the virus is being transmitted via the respiratory tract when a healthy person comes in contact with the infected person. The virus may transmit between persons through other roots which are currently unclear. The on the incubation period of the middle east respiratory syndrome (MERS), and the severe acute respiratory syndrome (SARS). According to W.H.O the signs and symptoms of mild to moderate cases are dry cough, fatigue and fever while as in severe cases dyspnea (shortness of breath), Fever and tiredness may occur. The persons having other diseases like asthma, diabetes, and heart disease are more vulnerable to the virus and may become severely ill. The person is diagnoses based on symptoms and his travel history. Vital signs are being observed keenly of the client having symptoms. No specific treatment has been discovered as on 10th April 2020, and patients are being treated symptomatically. The drugs like hydroxychloriquine, antipyretic, anti-virals are used for the symptomatic treatment. Currently, no such we may take some precautions to prevent this disease. By washing hands regularly with soap for 20 s and avoiding close contact with others by keeping the distance of about 1m may reduce the chances of getting affected by this virus.While sneezing, Covering the mouth and nose with the helpof disposable tissue and avoiding the contact with the nose,ear and mouth can help in its prevention.Till 10th of April2020, almost 1.6 million confirmed cases of coronavirus aredetected around the globe. Almost 97 K persons have diedand 364 K persons have recovered from this deadly virus[1]. Since no drug or vaccine is made for curing theCOVID-19. Various paramedical companies have claimedof developing a vaccine for this virus. Less testing has alsogiven rise to this disease as we lack the medical resourcesdue to pandemic. Since thousands and thousands are beingtested positive day by day around the globe, it is notpossible to test all the persons who showsymptoms.

Motivated by the event of many machine learning ways for the prediction of covid-19 condition risk, and we have a tendency to propose a sort of homogeneous ensemble learning methodology. The proposed methodology involves the employment of a mean based approach to randomly partition the dataset into smaller subsets and applying the classification and regression algorithmic program to model every partition.

II. RELATED WORK

This section discusses some learning strategies. ASA allows all users to auto-diagnose in their clinical conditions and to determine their levels of risk pattern by fulfilling the type of self-assessment score (SAS) is given in the app ASA.ASA helps people to decide whether they are at risk of coronavirus infection (COVID-19)or the flu without rushing to their 3].

The patient is at a higher level of risk (HLOR) or being diagnosed with COVID-19 and must seek medical support or obey the Union Health Ministry's self-isolation advice or not guided by ASA. The consumer is not unwell and shows no symptoms of COVID-19 then it will show the 0(negative) is the outcome of the self-assessment test (SAT) which is indicating that his pattern of the risk (POTR) for COVID-19.

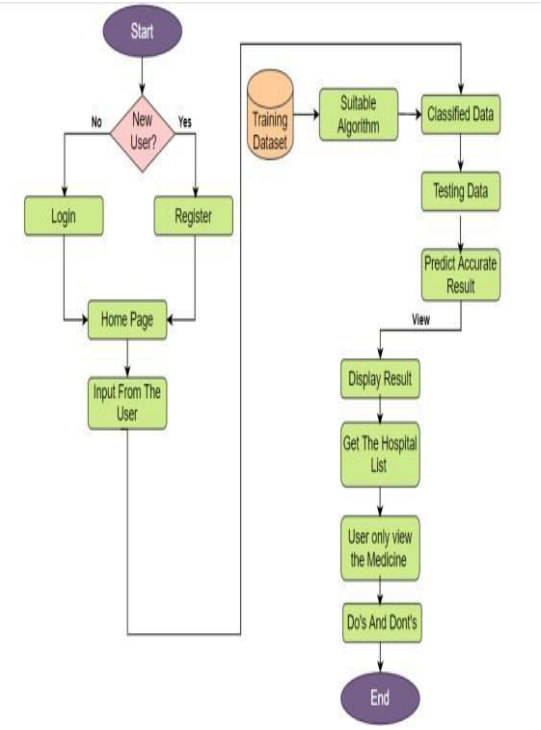
Machine learning and natural language processing use big dels for pattern recognition, explanation, and prediction. Classification is one of the major task in text mining and can be performed using different algorithms[1].

Once the COVID-19 is detected in a person, the question is whether and how intensively that person will be affected. Not all COVID-19 positive patients will need rigorous attention. Being able to prognosis who will be affected more severely can help in directing assistance and planning

medical resource allocation and utilization. [1] used machine stic prediction algorithm to predict the mortality risk of a plearning to develop a prognost used machine learning to develop a prognoerson that has been infected.

Fig : Block diagram of the system

This is the flow of the our application. In the initial state system can check it is a new user or not if user is new then first user is register after that login and user is already register then directly log in after the login user can show the home page on the user can show the various options 1st user can input to the application and then machine learning model 1st check training dataset and use algorithm and then user data classified and then give the accurate results and then user go to another options like hospital list in this option user can show the hospital list. after that user can go to the medicine option in this option user can show the medicine name after click on the medicine name user can



show the description about the medicine go to the next option like do's and don'ts in this option user only show related how to care aboutcovid-19.

III. PROPOSEDMETHODOLOGY

3.1 Overview:

To proceed with our proposed method of implementation, we need to follow few basic steps. The implementation starts with collection of symptoms based data. We are predicting the symptoms and extract from the data set. Then pre-processed the data and then give the final results using various kind of machine learning algorithm.

3.2 Machine learningclassification

Classification is a supervised learning process used to predict outcomes based on available data. This article proposes a method for diagnosing heart disease using classification algorithms and using many classifiers to improve classification accuracy. Dataset is divided into two parts: training set and test set. Every classifier is trained using training data set. Use the test data set to check the efficiency of the classifier. The machine learning algorithms like support vector machine (SVM), logistic regression, decision tree, random forest, gradient boosting were used for performing thistask.

3.2.1 Logistic regression

Logistic regression is a statistical and machine learning technique that can classify data sets in a data set based on the values of input fields. Predict the dependent variable based on one or more sets of explanatory variables to predict the outcome. It can be used for both binary classifications and multi-class classification. Logistic regression is one of the machine learning algorithms, which is relatively widely used in research to assess the risk of complex diseases. Therefore, the purpose of this study is to determine the most important predictors ofcardiovascular disease and use logistic regression analysis to predict the overall risk. [20]

3.2.2 Support vector machine(SVM)

Support vector machine (SVM) is a supervised machine learning algorithm for classifying text into different categories [1]. It takes 'n' number of features for the particular text with the given label. Here we have taken 20 features that are of nature unigram and bigram as the dataset is small. Here the data points of the training set n 1, where n is the number of features taken. The 20 features that have been selected in feature engineering with values are represented in the form of a table and are supplied as aninput.

3.2.3 Decisiontrees

An alternative approach for classification it partitions the input space into regions and classifies every region independently [7]. The 20 features that have been selected in feature engineering with values are represented in the form of a table and are supplied as an input. It splits the space recursively according to the inputs and classifies at the bottom of the tree. The leaf nodes classify the text into four classes. While building a decision tree, a vital function needs to be considered which is known as the splitting criterion. The function defines how data should be split in order to maximize performance.

IV. DATASET AND PERFORMANCEMETRICES

In this paper, one covid-19 dataset is used, the symptoms based dataset obtained from the Kaggle web site [10]. The previous contains 303 instances and 20 attributes, whereas the latter consists of 5000 instances and 20 attributes. The Symptoms based dataset contains missing attributes, and it's preprocessed to form it appropriate for machine learning. This dataset health records like heart disease, fever, dry cough, running nose, diabetes, etc. For our experiments, the 75-25 train- test holdout validation scheme is used; this is often to enable us to form a good and higher comparison between our proposed technique and former studies that used an analogous dataset. To adequately assess the performance of the proposed methodology, numerous performance indices are used, including accuracy, precision, sensitivity, specificity and F1 score. The accuracy is the proportion of the total number of predictions that were correct, and precision is the magnitude relation of correct positive predictions to the number of positive results predicted. At the similar time, Sensitivity (True Positive

rate) measures the proportion of positives that are correctly identified, while Specificity (True Negative rate) measures the proportion of negatives that are correctly identified and F- score is the harmonic mean between precision and sensitivity. The mathematical representations of these performance metricsare:

Accuracy = (TP+TN)/(TP+FP+FN+TN) (1)
Precision=TP/(TP+FP) (2)
Sensitivity=TP/(TP+FN) (3)
Specificity=TN/(TN+FP) (4)
F1 Score=2*(Sensitivity*Precision)/(Sensitivity + Precision)

=2TP/(2TP+FP+FN) (5)

Where, TP represents the number of true positives,
TN represents the number of true negatives,
FP represents the number of false positives and FN represents the number of falsnegatives.

V. RESULTS ANDDISCUSSION

To validate the effectiveness of the proposed method, a comparative study is conducted with other well-known machine learning methods. The methods include k-nearest neighbor (KNN), Logistic regression (LR), support vector machine(SVM), classification and regression tree (CART), gradient boosting (GB), and random forest (RF).

	Model	Score
5	Decision Tree	97.148114
0	Support Vector Machines	96.964121
2	Logistic Regression	96.504140
1	KNN	96.320147
3	Random Forest	89.625661
6	Gradient Boosting Classifier	87.252723
4	Naive Bayes	77.368905

Fig. summarize the test results of the various methods on the symptoms based test sets

Fig : Accuracy Result

From Fig. it is evident that the proposed method achieved superior classification performance on the Symptoms based test sets with accuracy of 96%. Also, from the results, i.e., the Logistic Regression and Decision Tree, performed better than the other algorithms. The performance of these classifier, together with the proposed method, is further studied with the receiver operating characteristic (ROC) curves. The ROC curves are useful for evaluating the predictive ability of the various ensemble models. They are created by plotting the true positive rate against the false positive rate at various threshold settings. The ROC curves for the ensemble approach is shown in Fig.

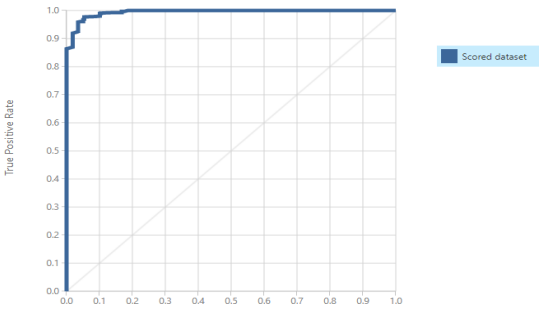
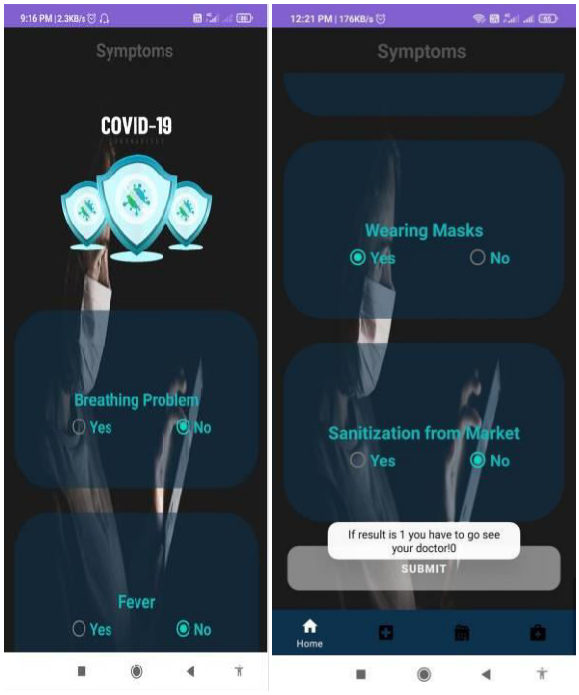


Fig : ROC Curve

We build one android application for the detection of covid-19 . User can use the app by login to the application and registration page is also provided. After that user can check self-assessment application will predict that user have a covid-19 or not. And user also view the medicine and covid-19 hospital list in to the application.

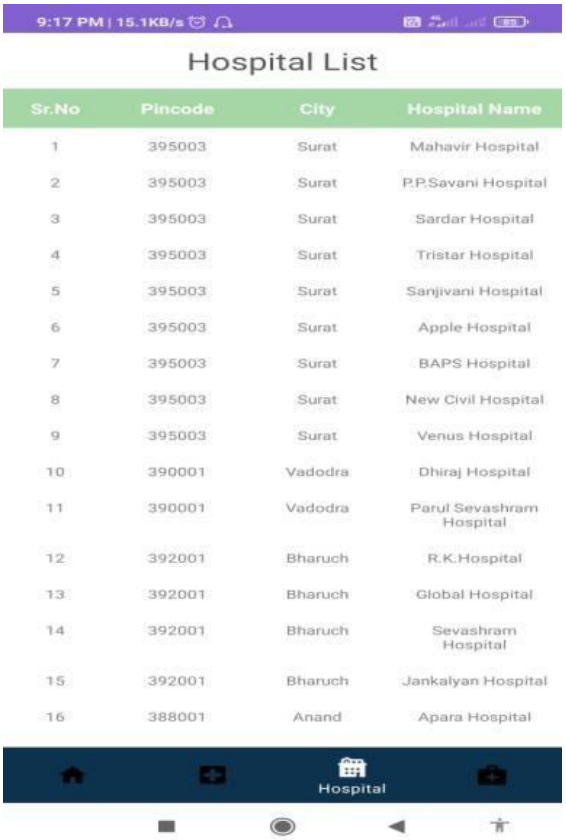
1. *Registration andLogin*

2. *Self-Assessment*



3. *View medicine*

1. *HospitalList*



Sr.No	Pincode	City	Hospital Name
1	395003	Surat	Mahavir Hospital
2	395003	Surat	P.P.Savani Hospital
3	395003	Surat	Sardar Hospital
4	395003	Surat	Tristar Hospital
5	395003	Surat	Sanjivani Hospital
6	395003	Surat	Apple Hospital
7	395003	Surat	BAPS Hospital
8	395003	Surat	New Civil Hospital
9	395003	Surat	Venus Hospital
10	390001	Vadodra	Dhiraj Hospital
11	390001	Vadodra	Parul Sevashram Hospital
12	392001	Bharuch	R.K.Hospital
13	392001	Bharuch	Global Hospital
14	392001	Bharuch	Sevashram Hospital
15	392001	Bharuch	Jankalyan Hospital
16	388001	Anand	Apara Hospital

2.Do &Dont's

CONCLUSION

The Modern Human Care System using Machine learning Algorithm, COVID-19 has shocked the world due to its non-availability of vaccine or drug. Various researchers are working for conquering this deadly virus. The proposed ensemble achieved accuracy of 96% on Symptoms based test sets using Heroku cloud services. More feature engineering is needed for better results and deep learning approach can be used in future.

ACKNOWLEDGMENT

We thank Mr. Ankit D. Prajapati for his assistance with paper and for helping out as our mentor/guide throughout the process.

REFERENCES

[1].Khanday, A.M.U.D., Rabani, S.T., Khan, Q.R. *et al.* Machine learning basedapproaches for detecting COVID-19 using clinical text data. *Int. j. inf. tecnol.* **12,** 731–739 (2020). https://doi.org/10.1007/s41870-020-00495-9
[2].Kabi A, Mohanty A, Mohanty AP, Vijalaxmi, Kumar N, Kumar S. Medical management of COVID-19: treatment options under consideration. Int J Adv Med 2020;7:xxx-xx.
[3].Venkat Ananth (April 15, 2020). Aarogya Setu's not all that healthy for a person's privacy". The Economic Times. 27.
[4]. Khanday, A.M.U.D., Amin, A., Manzoor, I., & Bashir, R., ''Face Recognition Techniques: A Critical Review'' 2018
[5]. Kumar A, Dabas V, Hooda P (2018) Text classification algorithms for mining unstructured data: a SWOT analysis. Int J Inf Technol. https://doi.org/10.1007/s41870-017-0072-1
[6]. Verma P, Khanday AMUD, Rabani ST, Mir MH, Jamwal S (2019) Twitter Sentiment Analysis on Indian Government Project using R. Int J Recent Tech Eng. https://doi.org/10.35940/ijrte. C6612.098319
[7]. Description of Logistic Regression Algorithm. https://machine learningmastery.com/logistic-regression-for-machine-learning/. Accessed 15 May 2019
[8]. H. Li et al., "Ensemble Learning for Overall Power Conversion Efficiency of the AllOrganic Dye-Sensitized Solar Cells," IEEE Access, vol. 6, pp. 34118– 34126, 2018, doi: 10.1109/ACCESS.2018.2850048.
[9]. Tin Kam Ho, "The random subspace method for constructing decision forests," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 20, no. 8, pp. 832–844, Aug. 1998, doi: 10.1109/34.709601.
[10]. https://www.kaggle.com/hemanthhari/symptoms-and-covid-presence
[11]. B. Zhang et al., "Ensemble Learners of Multiple Deep CNNs for Pulmonary Nodules Classification Using CT Images," IEEE Access, vol. 7, pp. 110358– 110371, 2019, doi: 10.1109/ACCESS.2019.2933670.
[12]. C. B. C. Latha and S. C. Jeeva, "Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques," Informatics in Medicine Unlocked, vol. 16, p. 100203, Jan. 2019, doi: 10.1016/j.imu.2019.100203.
[13]. Jiang X, Coffee M, Bari A, Wang J, Jiang X, Huang J, Shi J, Dai J, Cai J, Zhang T, Wu Z, He G, Huang Y (2020) Towards an artificial intelligence framework for data-driven prediction of coronavirus clinical severity. Compu Mater Contin63(1):537–551
[14]. Description of Logistic Regression Algorithm. https://machine learningmastery.com/logistic-regression-for-machine-learning/. Accessed 15 May 2019
[15]. Description of Multinomial Naı ¨ve Bayes Algorithm https://www. 3pillarglobal.com/insights/document-classification-using-multi nomial-naive- bayes-classifier. Accessed 15 May 2019
[16]. Khanday AMUD, Khan QR, Rabani ST. SVMBPI: support vector machine based propaganda identification. SN Appl. Sci. (accepted)
[18]Description of Decision Tree Algorithm: https://dataspirant.com/ 2017/01/30/how_decision_tree_algorithm_works/. Accessed 10 July 2019
[19]. Katuwal R, Suganthan PN (2018) Enhancing Multi-Class Classification of Random Forest using Random Vector Functional Neural Network and Oblique Decision Surfaces, Arxiv:1802.01240v1
[20]. Friedman JH (2002) Stochastic gradient boosting. Comput. Stat. Data Anal. 38(4):367–378. https://doi.org/10.1016/S01679473(01)00065-
[21]Description of Logistic Regression Algorithm. https://machine learningmastery.com/logistic-regression-for-machine-learning/.Accessed 15 May 2019
[22]. Description of Multinomial Naı ¨ve Bayes Algorithm https://www. 3pillarglobal.com/insights/document-classification-using-multi nomial-naive- bayes-classifier. Accessed 15 May 2019
[23]. Khanday AMUD, Khan QR, Rabani ST. SVMBPI: support vector machine based propaganda identification. SN Appl. Sci. (accepted)
[24]. Description of Decision Tree Algorithm: https://dataspirant.com/ 2017/01/30/how_decision_tree_algorithm_works/. Accessed 10 July 2019
[25]. Description of Boosting Algorithm: https://towardsdatascience. com/boosting. Accessed 10 July2019
[26]. Description of Adaboost Algorithm: https://towardsdatascience. com/boosting-algorithm-adaboost-b673719ee60c. Accessed 10 July 2019
[27]. Katuwal R, Suganthan PN (2018) Enhancing Multi-Class Classification of Random Forest using Random Vector Functional