International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 03 | Mar 2021 www.irjet.net p-ISSN: 2395-0072

GUI BASED COVID-19 PATIENTS PRECONDITION USING SMLT

D.Vetriselvi[1], S.Mohanapriya[2], K.Varsha[3], E.Yamini[4]

[1]*Assistant Professor, Dept. of Computer Science and Engineering, Jeppiaar SRR Engineering College, Chennai.*
[2,3,4]*Final Year Student, Dept. of Computer Science and Engineering, Jeppiaar SRR Engineering College, Chennai.*

Abstract – *The Coronavirus disease 2019 (COVID-19) pandemic, which originated in Wuhan China, has had disastrous effects on the planet wide community and has overburdened advanced healthcare systems throughout the world, WHO is continuously monitoring and responding to this pandemic. The current rapid and exponential rise within the enlargement of patients has efficient and quick prediction of the possible outcome of an infected patient for appropriate treatment using AI techniques. The aim is to predict machine learning based techniques for covid-19 recovery chances are possible or not, prediction results in best accuracy. The analysis of dataset is done by supervised machine learning technique that captures several information's such as variable identification, uni-variate analysis, bi-variate and multi-variate analysis, missing value treatments and analyze the data validation, data cleaning and preparing and data visualization are getting to be done on the entire dataset.*

***Key Words*: Dataset, Machine learning-Classification method.**

1. INTRODUCTION

Machine learning could also be a way of data analysis that automates analytical model building. It's a branch of AI which supported the thought that systems can learn from data, identify patterns and make decisions with minimal of the human intervention. The discipline of machine learning employs various approaches that the computers to accomplish tasks where no fully or partially satisfactory algorithm is out here. In cases where vast numbers of potential answers exist, one approach is to label variety of the proper answers as valid. This will then be used as training data for the pc to reinforce the algorithm(s) it uses to figure out correct answers.

To fill the void of the traditional healthcare system, using supervised machine learning trainer algorithms to process healthcare along with other parameters of COVID-19 positive patients. To predict the most likely outcome of the patient, based on their symptoms and the various conditions of the reporting case by identifying patterns from previous patient data. The work is compared with multiple machine learning classification algorithms that are available for processing patient data and identifies the one which is the best method for predicting accurate results of recovery chances.

The input data and the COVID-19 patient's precondition are compared with the past dataset. The data from the dataset are analysed so as to extract useful information from the data and taking the choice based upon the data analysis. Data cleaning, transforming and modelling the data is done here. Then the analysed data is trained and tested. Training data is that the data you employ to train an algorithm or machine learning model to predict the result you design your model to predict. Test data is employed to live the performance like accuracy or efficiency of the algorithm you're using to train the machine. These data then uses the multiple machine learning classification algorithms to predict the best accuracy results of the model.

2. RELATED WORKS

M. Qjidaa, Y. Mechbal, A. Ben-fares, H. Amakdouf, M. Maaroufi [1] "Early detection of COVID19 by deep learning transfer Model for populations in isolated rural areas"-June 2020, International Conference on Intelligent Systems and Computer Vision (ISCV), Volume-9, July. The spread of COVID 19, the World Health Organization suggests a large-scale implementation of COVID 19 tests. Unfortunately, these tests are expensive and cannot be provided and available for people in rural and remote areas. To remedy this problem, we will develop an intelligent clinical decision support system (SADC) for the early diagnosis of COVID 19 from chest x-rays which are more accessible for people in rural areas. Thus, we collected a total of 566 radiological images classified into 3 classes: a class of COVID19 type, a Class of Pneumonia type and a class of Normal type. In the experimental analysis, 70% of the data set was used as training set and 30% was used as the test set. After preprocessing process, we use some augmentation using a rotation, a horizontal flip, a channel shift and rescale. Our finale classifier achieved the best performance with test accuracy of 99%, f1- score 98%, precision of 98.60% and sensitivity 98.30%. Deep Learning algorithms to classify three classes COVID-19, Pneumonia and normal using Transfer Learning concept and used the pre-trained architectures such as DenseNet121, VGG16, VGG19, InceptionResNetV2, Exception, MobileNet, InceptionV3, they were able to determine characteristic features from chest X-ray images, and also we take advantages of the seven models to build an ensemble model that outperformed all other models. Our finale classifier achieved the best performance with test accuracy of 99%, f1-score 98%, precision of 98.60% and sensitivity 98.30%. In the future work we will

employ our deep learning method with a large dataset to give accurate and efficient results.

Alaa, A. R. Alsaeedy, Edwin K. P. Chong [2] "Detecting Regions At Risk for Spreading COVID-19 Using Existing Cellular Wireless Network Functionalities"-15 June 2020 , IEEE Open Journal of Engineering in Medicine and Biology ,Volume- 1. The article is to introduce a new strategy to identify areas with high human density and mobility, which are at risk for spreading COVID-19. Crowded regions with actively moving people (called at-risk regions) are susceptible to spreading the disease, especially if they contain asymptomatic infected people together with healthy people. Methods: Our scheme identifies at-risk regions using existing cellular network functionalities—handover and cell (re)selection—used to maintain seamless coverage for mobile end-user equipment (UE). The frequency of handover and cell (re)selection events is highly reflective of the density of mobile people in the area because virtually everyone carries UEs. Introduced a new strategy for identifying areas that potentially contribute to the spread of COVID-19. Our strategy exploits existing cellular network procedures, HO and CS, required to maintain connectivity for mobile UEs. The frequency of HO/CS events reflects how the UEs move and gather within the coverage area. High HO/CS rates imply at risk areas—those with high UE density and mobility over time. Measuring HO/CS rates allows distinguished high- from low-risk areas, enabling prioritization of further risk mitigation.

Shuvankar Roy, Mahua Nandy Pal, Sonali Bhattacharyya [3] "Implementation of an Informative Website"–"Covid19 Predictor", Highlighting COVID-19 Pandemic Situation in India,-IEEE Conference Proceedings (IEEE Conf Proc), Volume-2020. They represented implementation of an informative online platform- Covid-19 Predictor which is capable of disseminating accurate prediction of confirmed, deceased and affected Covid-19 cases in India on the basis of the data available in a reliable online repository. The work characterizes proper utilization of advanced technologies for web scrapping, model prediction, implementation of web application framework and cloud hosting. Information dissemination to the people allows them to have proper visualization power about the spreading of the disease which helps them in establishing their own opinions during pandemic situation. This, they become aware of the requirements of different containment measures also such as travel ban, border closing, trade control etc. as well as sanitization measures as per WHO's directives. and provide support to the implementation authority spontaneously. This in turn, provides more successful imposition of human beneficiary actions during pandemic.

Xuejun Qian, Robert Wodnicki, Haochen Kang [4] "Current Ultrasound Technologies and Instrumentation in the Assessment and Monitoring of COVID-19 Positive Patients"-IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, Volume-67, Issue- 11, Nov. 2020. The COVID-19 pandemic in December of 2019, clinicians and scientists all over the world have faced overwhelming new challenges that not only threaten their own communities and countries but the world at large. These challenges have been enormous and debilitated, as the infrastructure of many countries including developed ones had little, or no resources to deal with the crisis. They have reviewed existing and future technologies for the potential use of US for diagnosis and management of the novel COVID-19 lung illness. This condition is especially challenging for physicians and health systems due to its multiple comorbidities and refractory primary disease. Clinicians are forced to consistently balance the requirements for obtaining useful information for critical treatment decisions, with the need to maintain isolation of themselves and COVID-19-negative patients. US holds great promise in this regard, due to the fact that is it highly versatile, portable, and inexpensive, while at the same time being applicable across a broad spectrum of conditions. With multiple expected waves of this disease pandemic, there exists a critical need to streamline the work of physicians, providing tools to increase the ability to quickly diagnose and treat the primary condition as well as comorbidities. New technologies such as US arrays with a wide frequency range, computer aided diagnosis, and 2D arrays producing volumetric images should be further investigated to establish their utility for COVID-19.

Xinggang Wang, Xianbo Deng, Qiang Zhou [5] "A Weakly supervised Framework for COVID-19 Classification and Lesion Localization from Chest"-IEEE Transaction On Medical Imaging, Volume- 39, No. 8, August 2020. Accurate and rapid diagnosis of COVID-19 suspected cases plays a crucial role in timely quarantine and medical treatment. Developing a deep learning-based model for automatic COVID-19 diagnosis on chest CT is helpful to counter the outbreak of SARS-CoV-2. A weakly supervised deep learning framework was developed using 3D CT volumes for COVID-19 classification and lesion localization. For each patient, the lung region was segmented using a pre-trained UNet; then the segmented 3D lung region was fed into a 3D deep neural network to predict the probability of COVID-19 infectious; the COVID19 lesions are localized by combining the activation regions in the classification network and the unsupervised connected components. They reviewed existing and future technologies for the potential use of US for diagnosis and management of the novel COVID-19 lung illness. This condition is especially challenging for physicians and health systems due to its multiple comorbidities and refractory primary disease. Clinicians are forced to consistently balance the requirements for obtaining useful information for critical treatment decisions, with the need to maintain isolation of themselves and COVID-19-negative patients. US holds great promise in this regard, due to the fact that is it highly versatile, portable, and inexpensive, while at the same time being applicable across a broad spectrum of conditions. With multiple expected waves of this

disease pandemic, there exists a critical need to streamline the work of physicians, providing tools to increase the ability to quickly diagnose and treat the primary condition as well as comorbidities. New technologies such as US arrays with a wide frequency range, computer aided diagnosis, and 2D arrays producing volumetric images should be further investigated to establish their utility for COVID-19.

3. PROPOSED SYSTEM

The proposed method is built a machine learning model to classify the patient recovery chances. To overcome this method to implement machine learning approach by interface of GUI application. The dataset is first preprocessed and the columns are analyzed to see the dependent and independent variable and then different machine learning algorithms would be applied to extract patterns and to obtain results with maximum accuracy. These reports to the classification of supervised machine learning techniques for covid-19 patient precondition prediction. Finally, it highlights some of the observations on future research issues, challenges and needs.

4. PROPOSED ARCHITECTURE

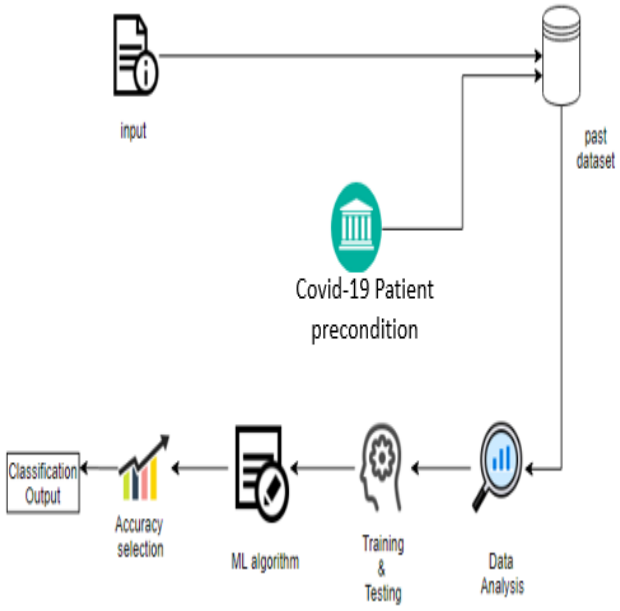

Fig –1: Proposed Architecture

5. MODULES AND DESCRIPTION

5.1 Data Validation and Pre-processing Technique

Validation techniques in machine learning are used to get the error rate of the Machine Learning (ML) model, which can be considered as on the brink of true error rate of the dataset. If the data volume is large enough to be representative of the population, you'll not need the validation techniques. To finding the missing value, duplicate value and outline of knowledge type whether it's float variable or integer. Pre-processing of data refers to transformations applied to your data before feeding it to the algorithm, whenever the data is gathered from different sources it is collected in raw format which isn't feasible for the analysis. For achieving better results from the applied model in Machine Learning projects the format of the data has got to be during a proper manner. The validation set is used to guage a given model, but this is often for frequent evaluation.

5.2 Exploration Data Analysis of Visualization and Training a Model by Given Attributes

Data visualization is a critical skill in applied statistics and machine learning. Statistics does indeed specialize in quantitative descriptions and estimations of data. Data analysis and visualization provides a crucial suite of tools for gaining a qualitative understanding of the data. This can be helpful when exploring and getting to know a dataset and will help with identifying patterns, corrupt data, outliers, and much more. With touch domain knowledge, data visualizations are often used to express and demonstrate key relationships in plots and charts that are more visceral and stakeholders than measures of association or significance.

5.3 Performance Measurements of Logistic Regression and Naive Bayes Algorithm

It is important to compare the performance of multiple different machine learning algorithms consistently and it will discover to make a test harness to match multiple different machine learning algorithms in Python with scikit-learn. Each model will have different performance characteristics.

In this module below 2 algorithms are compared

- Logistic Regression
- Naive Bayes Algorithm

5.3.1 Logistic Regression

First, we'd like to import the libraries that we'll use to create our logical regression model. We'll use the Pandas library to load within the CSV or the dataset, and Numpy to convert the info frame into arrays. We need to separate the data from the dataset into dependent and independent variables. Now that we've defined the target variable(Y) and therefore the independent variables, we'd like to separate the info set into the training set and therefore the test set. We will use the training set to coach our logistic regression algorithm. Similarly, the test data set are going to be wont to validate the logistic regression model. To split the info into two sets, we'll use Sklearn. To get better accuracy for our model, we need to rescale the data to bring value which may have extremely varying values into alignment with one another.

Now, we'd like to create the logistic regression model and that fit it to the training and testing data set.

5.3.2 Naive Bayes Algorithm

A Naive Bayes classifier assumes that the presence of a selected feature during a category is unrelated to the presence of the opposite feature. This model is straightforward to create and particularly useful for very large data sets. It calculates the prior probability for the given class labels. And, it finds the Likelihood probability with each attribute for every class. Then put these values into the Bayes Formula and calculates the posterior probability. To see which class features has a higher probability, given the input belongs to the upper probability class.

5.4 Performance Measurements of Random Forest and Support Vector Machines

When have a replacement dataset, it's an honest idea to see the info with various parameters using different techniques so as to seem at the info from different perspectives. The same idea applies to model selection. You should use variety of various ways of watching the estimated accuracy of your machine learning algorithms so as to settle on the one or two to finalize. A way to try to to this is often to use different visualization methods to point out the typical accuracy, variance and other properties of the distribution of model accuracies.

In this module below 2 algorithms are compared

- Random Forest
- Support Vector Machines

5.4.1 Random Forest

The essential steps involved in performing the random forest algorithm are listed. Pick N random records from the dataset. Build a choice tree supported these N records. Firstly, choose the amount of trees you would like to have in your algorithm and repeat steps 1 and a couple of. In case of a regression problem, for a replacement of the record, each tree within the forest predicts a worth for Y (output). The final values are often calculated by taking the typical values of all the values predicted by all the trees in the random forest. Or, in case of a classification problem, each tree within the random forest predicts the category to which the new record belongs to. Finally, the new record will be assigned to the category that wins the bulk votes among all the trees used.

5.4.2 Support Vector Machines

Support Vector Machines are one among the foremost popular and talked about machine learning algorithms. It works rather well when a transparent margin of separation it's effective in high dimensional spaces. It is effective in cases where the quantity of dimensions is greater than the quantity of samples. Explore the info to work out what they appear like. Pre-process the data. Split the data into attributes and labels. Divide the data into training and testing sets. Train the SVM algorithm. Make some predictions. Evaluate the results of the algorithm.

5.5 Graphical User Interface Output Prediction of Covid-19 Precondition

After the comparison of above 4 algorithms, the best algorithm based on all the attributes like the accuracy, precision, F1 score and the error rates (Sensitivity and Specificity) is selected and the chances of recovery is being predicted using that model. The prediction result will be displayed in GUI application which is built by using tkinter package.

6. RESULTS

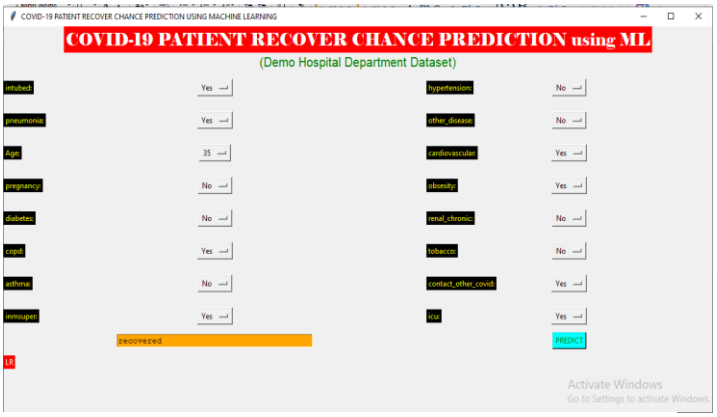

Fig –2: GUI Prediction

7. CONCLUSIONS

In this work, the clinical data from COVID-19 patients are used to measure the prediction accuracy of the recovery chances of hospitalized patients by implementing different analysis models. The results indicate the most accurate recovery chances prediction for each algorithm and compared with the other algorithms. And finally, the best accuracy among the results of all the algorithms is used to analyze and predict the recovery chances of the patients.

REFERENCES

[1] Yi-Cheng Chen and Ping-En-Lu, "A Time-Dependent SIR Model for COVID-19 with Undetectable Infected Persons", IEEE Transactions on Network Science and Engineering, Vol.7, No.4, pp.3279-3292, 2020.

[2] I.Nesteruk, "Statistics Based Predictions of Coronavirus 2019-ncov Spreading in Mainland China", MedRxiv, 2020.

[3] M.Qjidaa and Y.Mechbal," Early Detection of COVID19 by Deep Learning Transfer Model for Populations in Isolated Rural Areas", Vol.1, pp.1-7, 2020.

[4] VM Corman, O Landt, M Kaiser, R Molenkamp, A Meijer, DK Chu et al., "Detection of 2019 Novel Coronavirus (2019-nCoV) by Real-time RT-PCR", Euro Surveillance: Bulletin European Sur les maladies Transmissibles = European Communicable disease bulletin, vol. 25, no. 3, 2020.

[5] B. S. Dayal and J. F. MacGregor, "Identification of Finite Impulse Response Models: Methods and Robustness Issues", Ind. Eng. Chem. Res., vol. 35, no. 11, pp. 4078-4090, 1996.

[6] M. Mandal, S. Jana, S. K. Nandi, A. Khatua, S. Adak and T. Kar, "A Model Based Study on the Dynamics of COVID-19: Prediction and Control", Chaos Solitons & Fractals, vol. 136, pp. 109889, 2020.

[7] C. Huang et al., "Clinical Features of Patients Infected with 2019 Novel Coronavirus in Wuhan China", Lancet, vol. 395, no. 10223, pp. 497-506, 2020.