International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 04, April-2023

A Machine Learning based System to Serve Critical Patients with A Real-Time Feedback Method

Satyajit J. Pokale [1st]
Department of Electronics and Telecommunication,
MIT, Pune, India

Pratik S. Mule [2nd]
Department of Mechanical Engineering,
UCOER, Pune, India

Pranali J. Khedekar [3rd]
Department of Instrumentation Engineering
AISSMS IOIT, Pune, India

Siddhi Bhand [4th]
Department of Computer Engineering,
PICT, Pune India

***Abstract*— Adaptable Critical Patient Caring system is a primary concern for hospitals in growing countries. Most of the hospitals lack serving proper health services due to the unavailability of appropriate, easy, and scalable smart systems. The focus of this project is to build an adequate system for hospitals to serve critical patients with a real-time feedback method. This system will focus on understanding the user review and organizing them to help future users with their use. This system will help a user make the best decision in terms of the combination of waiting times, travel distance, and find the nearest best hospital. Machine Learning-based health prediction of the patients is the main concept of this project. For our ml models, we have chosen a Random Forest Classifier for prediction. In this project, the mobile application is used for real-time data and information view. The system architecture is planned in such a way that the ML models can train and deploy in a real-time interval by retrieving the data from the dataset. The idea is to create an Android Application where you can provide reviews. The user gets hospital and disease recommendations in a very efficient and user-friendly manner by analyzing the symptoms provided by the user. This is a real-time recommendation system. This provides the prediction about the disease and provides the best suitable hospital to cure the predicted disease.**

Keywords: Machine Learning, Random ForestClassifier, Recommendation system.

I. INTRODUCTION

Most hospitals lack proper health services due to the unavailability of appropriate, easy, and scalable smart systems. At present, we have not found any popular hospital recommendation system based on Popular hospital recommendation system based on the reviews provided by the patients. We found some particular disease prediction systems. Studies and research on disease prediction systems use mostly images to predict the disease. We have not found any disease prediction system based on the symptoms of the patients. The focus of this project is to build a system for users to recommend the best hospitals to serve critical patients with a real-time feedback method. This system is focused on understanding the user review and organizing them to help future users with their use. This system will recommend the best suitable hospital near you on the basis of disease/symptoms. An ideal time for the doctors and patients will be reduced. And increase the throughput of the hospitals. It will helppatients choose the hospital to get the best treatment based on their user reviews. This motivates the hospitals to improve their services. The data will be useful for the hospital administrations to evaluate their performance and decide the regions to improve.

II. CLASSIFICATION ALGORITHM

Random forest classifier selects a subset of the training set and creates a set of decision trees. We are using this classifier to classify diseases on the basis of their symptoms. To decide the final class ofthe test object, a Random forest classifieraggregates the votes from different decision trees. Random decision forests avoid over fitting to theirtraining set. Random forest accuracy is lower thangradient boosted trees. However, datacharacteristics can affect their performance. Toincrease the predictive power of random forest hyper parameters like n estimators, max features, min sample leaf is used. The hyper parameters n jobs, random state, OOB score are used forincreasing the model's speed. For our model we have set an estimator value to 100 i.e. 100 decision trees are created and their average value is taken.

Similarly, the criterion is set to "gini"(calculates the amount of probability of a specific feature that is classified incorrectly when selected randomly), min samples split is set to 2, max features is set to "auto", the random state is set to 0 (to produce the same result every time).

In this paper, they have proposed a generic architecture, associated terminology, and a classificatory model for observing critical patient's health conditions with machine learning and IBM cloud computing as Platform as a service (PaaS). Machine Learning (ML) based health prediction of the patients is the key concept

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 04, April-2023

of this paper. IBM Cloud, IBM Watson studio is the platform for this research to store and maintain data and ml models. For ml models, they have chosen the following Base Predictors: Naïve Bayes, Logistic Regression, KNeighbors Classifier, Decision Tree Classifier, Random Forest Classifier, Gradient Boosting classifiers, and MLP Classifier. For improving the accuracy of the model, the bagging method of ensemble learning has been used. The following algorithms are used for ensemble learning: Bagging Random Forest, Bagging Extra Trees, Bagging KNeighbors, Bagging SVC, and Bagging Ridge [1].

In this paper, they are focused on understanding the user review and organizing them to help future users with their use. On online platforms, the data is in an unorganized format contributed by the users. Users spend a lot of time on the various sites for user reviews for their interesting products, only some are useful to them. This paper talks about analyzing user reviews previously submitted and purchase decisions made by a user and searching products best fit for users. Natural Language Processing tools to parse and analyze sentences in the review. The online review system is not the formal way users write the reviews at their convenience; there can be many more spelling mistakes in words and not the proper sentence formats which lead to low accuracy. To resolve this problem this paper discusses the two supervised learning techniques, Class association recluse, and Naive Bayes classifier [2].

In this paper, they have used the unsupervised methodology and exploratory data analysis to find out the hidden features of a dataset of patient reviews to find out the cause of dissatisfaction [3].

This paper says that the Patients that remain in the hospital system longer than necessary (overstay patients) represent a sizable operational cost and contribute to hospital waiting times and bed shortages. Patient data from four hospitals were analyzed to build the classifier that would identify patients that are likely to overstay. The patients that overstay often require special assistance, such as nursing home placement or home care arrangements, and need to be identified early in admission so as to schedule a timely discharge from the hospital. Age, co morbidity, and activities of daily living scores (such as the ability to dress and feed oneself) were the major factors in determining if a patient is likely to overstay while waiting for special dispensation. The aim of the research is to develop a decision support system using machine learning strategies. A decision tree classifier achieved an F-Measure of 0.826 identifying overstay patients from a tertiary teaching hospital and an F-Measure of 0.784 at a community hospital [4].

In this paper, the author has investigated the risk factors that lead to severe retinopathy of prematurity using statistical analysis and logistic regression as a form of generalized additive model (GAM) with pair wise interaction terms (GA2M). In this process, they discuss the trade-off between accuracy and interpretability of these machine learning techniques on clinical data. They have also confirmed the intuition of expert neonatologists on a few risk factors, such as gender, that were previously deemed as clinically not significant in RoP prediction [5].

III. PROPOSED SYSTEM

In today's digital world *Healthcare* is one core area of the medical domain. There is no popular recommendation system that recommends the best suitable hospital near you on the basis of symptoms. This system also helps users to make decisions in less time. This system can help a user make the best decision and find the nearest best hospital. Most hospitals lack proper health services due to the unavailability of appropriate, easy, and scalable smart systems. The main focus of this project is to build a system for users to recommend the best hospitals to serve patients with a real-time feedback method.

This system is focused on understanding the user review and organizing them to help future users with their use. The idea is to create an Android Application where you can provide reviews. The user gets hospital recommendations in a very efficient and user-friendly manner by analyzing the symptoms provided by the user.

This is a real-time recommendation system. This system provides the best suitable hospital to cure the predicted disease on the basis of the symptoms or disease name.

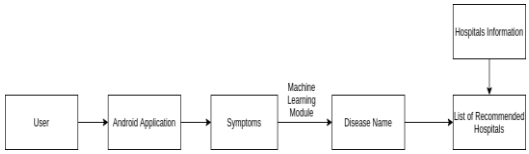

Fig. 1 System Architecture

I. Dataset

We have referred to the "Disease Symptom Prediction" dataset from Kaggle to predict disease using symptoms. In this dataset, we had added our dataset of covid-19 disease. To make the dataset, we had collected the information related to symptoms faced by covid-19 infected patients. Our covid-19 disease dataset is combined with the kaggle dataset. Now, It consists of 41 unique disease names and 143 unique symptoms. Dataset consists of 5101 entries.

We have created our own dataset consisting of information about the hospital. The features of this dataset are hospital name, hospital address, disease name, staff rating, medical facilities rating, parking ratings, ward facility ratings, overall ratings, and any suggestions. In these features, ward facility rating is optional as every patient may not admit to the hospital. Dataset consists of 558 entries. This dataset is used to recommend the

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 04, April-2023

hospital list to users based on the various features as well as nearest from a user by distance.

II. SYSTEM IMPLEMENTATION

In this system, we have provided two options to the user to get the best hospital list by selecting various symptoms or disease names. When users select the symptoms predict the disease based on those symptoms using a random forest classifier. In this classifier, we had to provide hyper parameters. We have set an estimator value to 100 i.e. 100 decision trees are created and their average value is taken to predict disease. To predict the disease name we are using the "Disease Symptom Prediction" dataset. We have taken 70% of the data for the training purpose of the model and 30% data for testing purposes. When we use a decision tree as a classifier we got 90% accuracy. But due to the over fitting problem of the decision tree, we have used the random forest classifier. We got 100% accuracy in the random forest classifier.

Once we get the disease name from a machine learning model or directly from the user we apply our algorithm to provide the best hospital to the user. Our algorithm calculates the distance between the current user and the hospital and calculates the average for all the ratings. Using all these calculations our hospital recommendation model provides a list of hospitals best for the user. These hospitals are provided by considering their reviews as well as the distance from the user.

We have worked on all the real-time hospitals and their locations but due to the copyright issue. Here, we are providing dummy hospital names as well as addresses.

III. RESULT

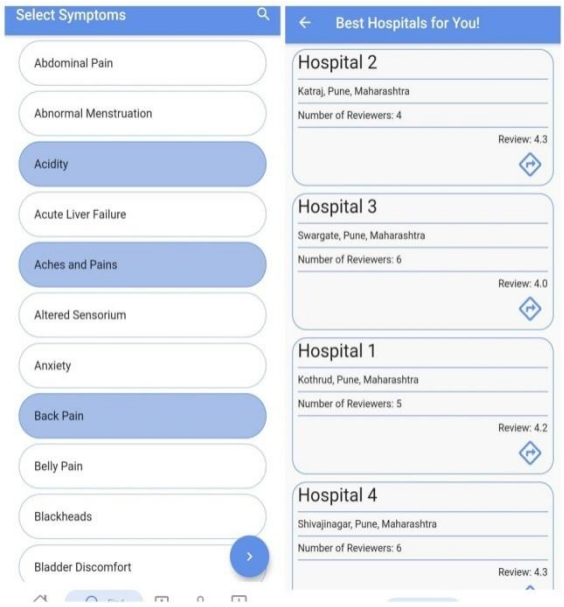

Fig. 2 Select Symptoms and List of Hospitals

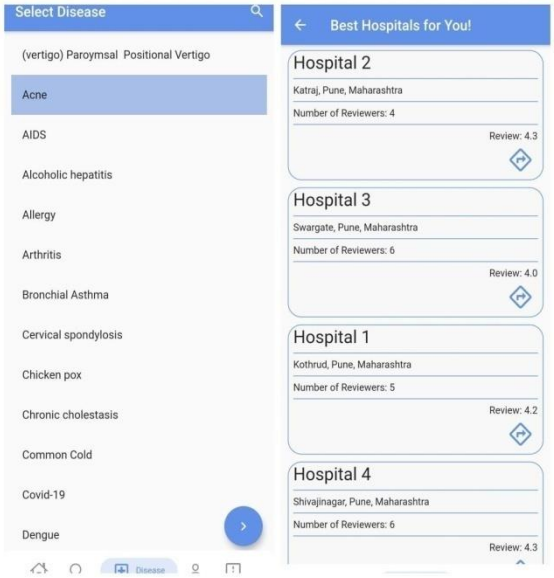

Fig. 3 Select Disease and List of Hospitals

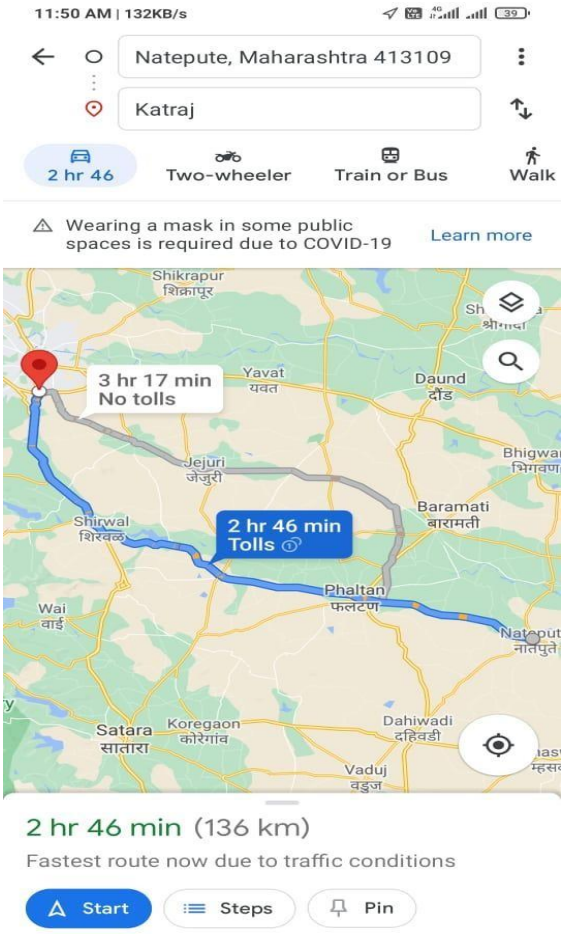

Fig. 6 Map from Current Location to HospitalLocation

IV. CONCLUSION

This system will recommend the best suitable hospital near you on the basis of disease/ symptoms. An ideal time for the doctors and patients will be reduced. And increase the throughput of the hospitals. It will help patients choose the hospital to get the best treatment on their user reviews. This motivates the hospitals to improve their services.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 04, April-2023

V. DECLARATION

Funding: No funds, grants or other support was received from any organization to assist with the preparation of this manuscript.

Conflicts of Interest/Competing interests:

The Authors have no relevant financial or non-financial interests to disclose.

VI. REFERENCE

[1] A. A. Neloy, S. Alam, R. A. Bindu and N. J. Moni, "Machine Learning based Health Prediction System using IBM Cloud as PaaS," 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), 2019, pp. 444-450, doi: 10.1109/ICOEI.2019.8862754.

[2] Yang, C.C., Tang, X., Wong, Y.C. and Wei, C.P., 2010. Understanding online consumer review opinions with sentiment analysis using machine learning. Pacific Asia Journal of the Association forInformation Systems, 2(3), p.7.

[3] T. S. Tabrizi, Mohammad Reza Khoie, E. Sahebkar, S. Rahimi and N. Marhamati, "Towards a patient satisfaction based hospital recommendation system," 2016 International Joint Conference on Neural Networks (IJCNN), 2016, pp. 131-138, doi: 10.1109/IJCNN.2016.7727190.

[4] R. Vivanco and D. Roberts, "Predicting Patients Likely to Overstay in Hospitals," 2011 10th International Conference on Machine Learning and Applications and Workshops, 2011, pp. 168-171, doi: 10.1109/ICMLA.2011.115.

[5] T. Karatekin et al., "Interpretable Machine Learning in Healthcare through Generalized Additive Model with Pairwise Interactions (GA2M): Predicting Severe Retinopathy of Prematurity," 2019 International Conference on Deep Learning and Machine Learning in Emerging Applications (Deep-ML), 2019, pp. 61-66, doi: 10.1109/Deep-ML.2019.00020.