

Data Driven Machine Learning System for Optimization of Clinic Activities

Paresh Devlekar*¹, Prajwal Gawande*², Chinmay Chalke*³, Shruti Agrawal*⁴

*^{1,2,3}(Student, Department of Information Technology, Vidyalkar Institute of Technology, Mumbai, Maharashtra, India)

*⁴(Prof., Department of Information Technology, Vidyalkar Institute of Technology, Mumbai, Maharashtra, India)

ABSTRACT

At present, a person sick from any disease needs to visit the hospital/clinic to book an appointment. The patients need to wait in a queue while getting an appointment. Just in case a doctor calls off an appointment due to some emergency reasons then the patient remains unaware about the cancellation of the appointment till the time he or she visits the clinic. Also whenever the patient visits the doctor is asked to carry their prescriptions. Doctors don't keep a track record of when the patient last visited him for an appointment. Doctors are unaware of who all will visit him on a given day and hence are not able to manage their appointments accordingly. Doctors are unable to access the patient details and medical history until and unless he/she checks the medical report. Also, now-a-days major problems based on symptoms humans predict the disease and that creates a panic situation. Thus, our integrated machine learning system helps the user to find its disease based on symptoms and also helps the user to redirect to the specialists for seeking help. This paper deals with an automated virtual clinic system which consists of an automated appointment booking facility, disease prediction system and electronic prescription section which will help doctors along with the patients.

Keywords - automated, machine learning, automated appointment booking, disease prediction.

Date of Submission: 05-04-2022

Date of Acceptance: 20-04-2022

I. INTRODUCTION

Disease is one of the most common issues in a person's life. This paper aims to cover all the aspects of management and operations of a clinic and take them online so any one can access it from remote location. It will help healthcare provider to improve its operational effectiveness, reduce time consumption, and enhancing delivery of quality of care. This project implements all the activities of a clinic in a computerized automated way to reduce the human load, interference and also to fasten the performance.

The appointment-scheduling process has become a necessary burden in medical offices, healthcare facilities and wellness centers. An automated Appointment and scheduling software have been developed to minimize some of the challenges faced by medial workers. The advantages of implementing this technology would benefit everyone involved in the scheduling process, as administrators and staff can conduct their tasks more efficiently and accurately, while customers and clients have the ability to book their appointments and reservations quickly and more comfortably. The

proposed system is aimed at simplifying the task of the patient and the doctor. It will reduce long waiting time for patients and eliminate long queue. Patient also have privilege to fix their appointment and also book an appointment according to their preference. The system will carry out timely and convenient access to health services for all patients.

Another problem arises when the patients who are in need to know the condition of their health and at sometimes if the person has been observing few symptoms and he/she is not sure about the disease he/she is suffering from will lead to various diseases in future. So, to avoid that and get to know the disease in early stages of the symptoms this disease prediction will help a lot to various people ranging from children to teenagers to adults and also the senior citizens. The disease is predicted using various techniques and algorithms and the user has to enter the symptoms from the given drop-down menu.

II. LITERATURE SURVEY

In paper [1], ShankarBirajdar, Snehal Dalvi, Gaurav Kumavat, Kavita Devadiga,, Prachi Kshirsagar in the research work proposed a

dependable online appointment booking system. The research proposes a web based medical appointment booking system. The scope of the research work is to design a web-based appointment booking system where patients can register themselves, login to the system, book an appointment with a doctor and view their appointments. The system also incorporates the doctor logging into the system, cancel an appointment, generate appointment and view appointments. However, the research was limited and laboratory results were not included in the system due to technical constrains. And more so, the system was not able to diagnose or prescribe drug for usage.

In paper [2], S. B. Choudhari, Chaitanya Kusrkar, Rucha Sonje, Parag Mahajan in the research work proposed Patient Appointment and Scheduling System. The proposed system is aimed at simplifying the task of the patient and the doctor. It will reduce long waiting time for patients and eradicate long queue. Patient also have freedom to fix their appointment and as well book an appointment according to their preference. This paper targets on developing a system to improve upon the efficiency and quality of delivering a web based appointment system to reduce waiting time. In this paper, a patient appointment and scheduling system is designed using technologies such as Angular JS for the frontend, Ajax framework for handling client-server request and Sqlite3 and MYSQL for the backend.

In paper [3], Raj H. Chauhan, Daksh N. Naik, Rinal A. Halpati, Sagarkumar J. Patel, Mr. A.D.Prajapati in the research work proposed a disease prediction system. This system is employed to predict disease on the basis of symptoms. This system uses decision tree algorithm for evaluating the model. This system is used by end-users. However, the research included only Decision Tree Algorithm for predicting disease. And more so, the accuracy was 90.12%.

In paper [4], Anuj Kumar, Mr. Analp Pathak in the research work proposed a disease prediction system which takes down five symptoms from the user and evaluate them by applying algorithms such as Decision Tree, Random Forest, Naïve Bayes and KNN which will help in getting accurate prediction. Also developed an elegant GUI to provide interaction with system. Apart from these, system also comprises of a database for storing the data entered by the users and the name of the disease the patient is suffering from which can be used as a reference in future for further treatment. The proposed system achieved the mean accuracy of more than 95% which shows remarkable rectification and high accuracy than the previous

work and also makes this system more reliable than the existing one.

In paper [5], Mesfin Leranso Betalo in the research work proposed Electronic Prescription System. The paper-based prescription process is inefficient, costly and resource-intensive. Most of the restrictions of paper prescription could be eliminated or reduced by electronic prescription. Features of E-Prescription System - Producing prescription electronically, generating a complete active medication list, printing prescriptions.

III. PROBLEM STATEMENT

Currently, the scenario is such that the patient who is suffering from any disease has to personally go to the doctor's clinic in order to book an appointment or sometimes has to make a call to the receptionist to book an appointment. Also, there is no such software that can tell the patient what disease is he/she suffering from based on the symptoms given. In some cases, the patients also lose their prescription which is given by the doctor and in that case the patient will not come to know which medicines to take at what time. So our proposed web application will reduce this workload of the patients as well as the doctors by automating the consultation process.

IV. SYSTEM ARCHITECTURE

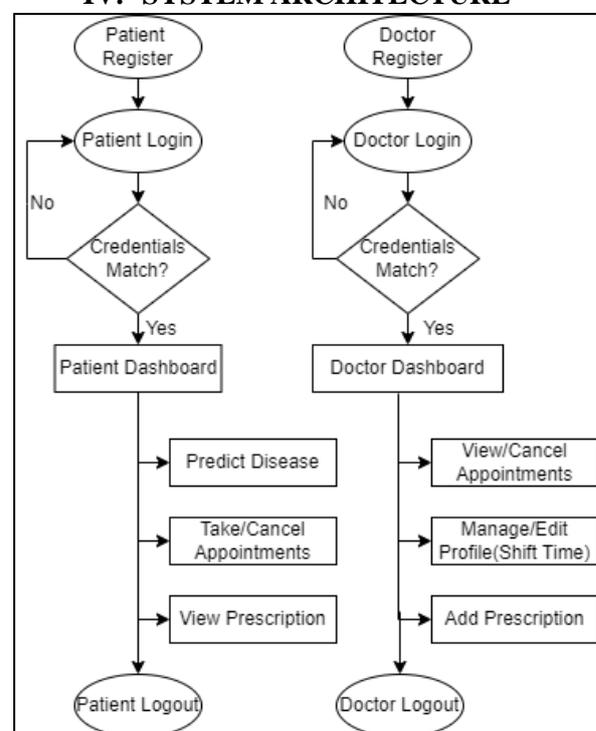


Fig.1 System Architecture

V. METHODOLOGY

Our proposed system consists of the following modules –

1. Patient Appointment Scheduling System.
2. Disease Prediction System Using Machine Learning.
3. Electronic Prescription & Medical History.

[1] Patient Appointment Scheduling System

Patient login or register on a web application to enter the system. Patient needs to create his/her profile by entering basic information. After creating a profile, patient can use our disease predictor functionality if he/she is not aware about the disease they are suffering from by providing symptoms as input. After the disease is predicted, the patient can take an appointment with the concerned department (specialist). Patient needs to enter the preferred department, date and time at which he/she wants the appointment to be booked. If the desired slot is already booked, the system will check for the other slots which are available for that day and book an appointment. If all the slots of a doctor are booked for the day, the system will book an appointment with another doctor of the same department. But if not, then the system will book an appointment in earliest available slot. The time slot given for each appointment is 60 minutes.

[2] Disease Prediction System Using Machine Learning

This module implements a robust machine learning model that can efficiently predict the disease of a human, based on the symptoms that he/she possess. Below is the figure which depicts the workflow of the implementation.

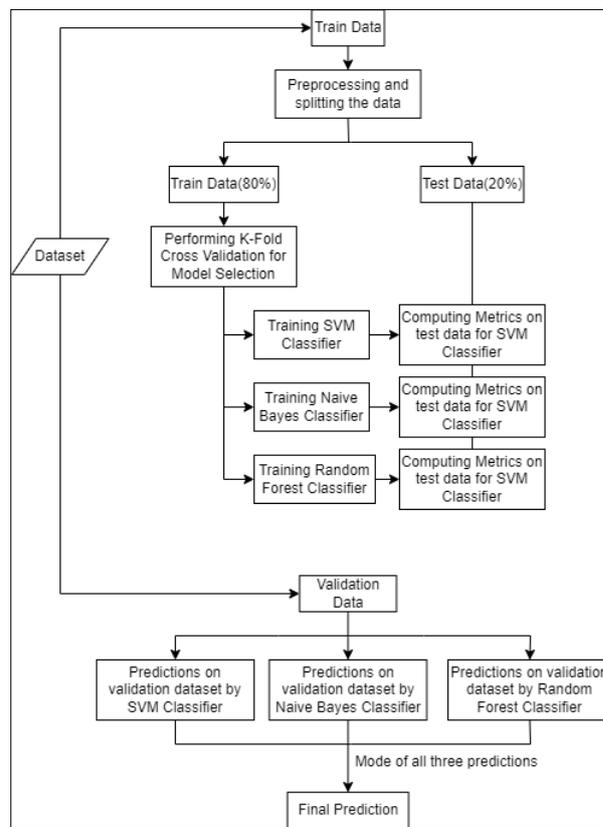


Fig.2 Workflow of the Implementation.

Collecting the Data: Data preparation is the primary step for any machine learning problem. We have used a dataset from Kaggle for this problem. This dataset consists of two CSV files, one for training and one for testing. There are a total of 133 columns in the dataset out of which 132 columns represent the symptoms and the last column is the prognosis.

Cleaning the Data: Cleaning is the most vital step in a machine learning project. The quality of our machine learning model totally depends on the quality of our data. So it is always mandatory to clean the data before feeding it to the model for training. In our dataset all the columns are numerical except the target column i.e. prognosis is a string type (categorical variable) and hence is encoded to numerical form using a label encoder.

Splitting the data for training and testing the model: Now that we have cleaned our data by removing the Null values and converting the labels to numerical format, it's time to split the data to train and test the model. We have split the data into 80:20 format i.e. 80% of the dataset has been used for training the model and 20% of the data been used to evaluate the performance of the models.

Model Building: After collecting and cleaning the data, the data is ready and can be used to train a machine learning model. We have used **K-Fold cross-validation** to evaluate the machine learning models. We have used this cleaned data to train the **Support Vector Classifier, Naive Bayes Classifier, and Random Forest Classifier**. We have used a confusion matrix to determine the quality of the models.

- **K-Fold Cross-Validation:** K-Fold cross-validation is one of the cross-validation techniques in which the whole dataset is split into k number of subsets, also known as folds, then training of the model is performed on the k-1 subsets and the remaining one subset is used to evaluate the model performance [6].

The steps for k-fold cross-validation are:

- Split the input dataset into K groups
- For each group:
- Take one group as the reserve or test data set.
- Use remaining groups as the training dataset
- Fit the model on the training set and evaluate the performance of the model using the test set.
- **Support Vector Classifier:** Support Vector Classifier is a discriminative classifier i.e. when given a labelled training data, the algorithm tries to find an optimal hyperplane that accurately separates the samples into different categories in hyperspace. The goal of the SVM algorithm is to create the best line also known as decision boundary that can segregate n-dimensional space into classes so that we can easily put the new data point in the correct category in the future [7].
- **Gaussian Naive Bayes Classifier:** It is a probabilistic supervised machine learning algorithm that internally uses Bayes Theorem and used for solving classification problems. A Naïve Bayes algorithm has a parallel performance with decision tree and other selected classifiers. It is very simple to build and useful for multiclass classification. Naive Bayes classify the data by calculating the probability of independent variable. After the probability of each class is computed, complete transaction is assigned to high probability class. Naive Bayes works excellent in various complex real world problem. The benefit of using Naïve Bayes is that it needs very less amount of training dataset to evaluate the parameters necessary for classification [8].
- **Random Forest Classifier:** It is an ensemble classifier using many decision tree models. Random Forest is a classifier that contains

multiple number of decision trees on various subsets of the given dataset which are formed using row sampling with replacement technique. In Random Forest, we end up with trees that are not only trained on different sets of data (a.k.a bootstrapping) but also use different features to make decisions. Instead of relying on one decision tree output, the random forest takes the prediction from each tree (a.k.a aggregation) and based on the majority votes of predictions, it predicts the final output. The greater the number of trees in the forest leads to higher accuracy and prevents the problem of overfitting [9].

Inference: After training the three models we will be predicting the disease for the input symptoms by combining the predictions of all three models i.e. by taking the mode of the predictions of all three models so that even one of the models makes wrong predictions and the other two make correct predictions then the final output would be the correct one. This approach makes our overall prediction more robust and accurate.

Model	Accuracy(%)
Accuracy on test data by SVM Classifier	90.26
Accuracy on test data by Naive Bayes Classifier	86.67
Accuracy on test data by Random Forest Classifier	88.72
Accuracy on Test dataset by the combined model	93.5

[3] E-Prescription and Medical History

Electronic Prescription section has also been implemented so that doctors can write prescription and even update prescription if needed. This helps in maintaining a soft copy of prescription which will be visible to both doctors as well as patients in the medical history tab. So there is no need for the patients to carry prescription file every time they visit clinic. Doctors can view the medical history of their patients which will help doctors in providing suitable treatment.

VI. IMPLEMENTATION AND RESULTS

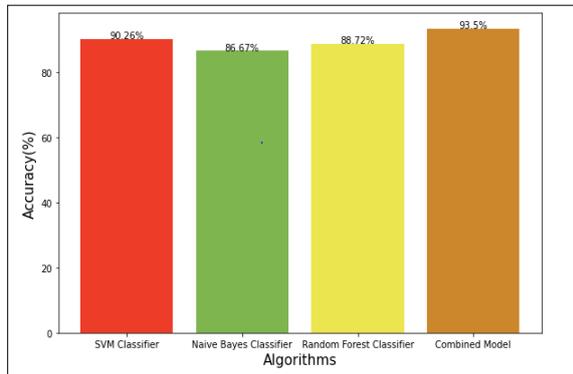


Fig.3 Accuracy representation in graph with their respective algorithms.

Below figure shows the appointment list of a doctor in which appointments are getting booked after an interval of 60 minutes. After the doctor's shift ends, appointment gets booked with another doctor of same specialty.

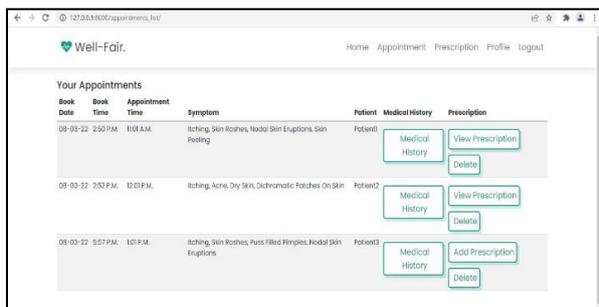


Fig.4 Doctor Appointment List Section.

Below figure shows the disease predictor functionality which provides a drop-down menu where a list of symptoms is passed. The user can select a minimum of two to a maximum of five symptoms. More accuracy will be attained if more number of symptoms are provided as input. Also shows the respective specialist from whom consultation needs to be taken.

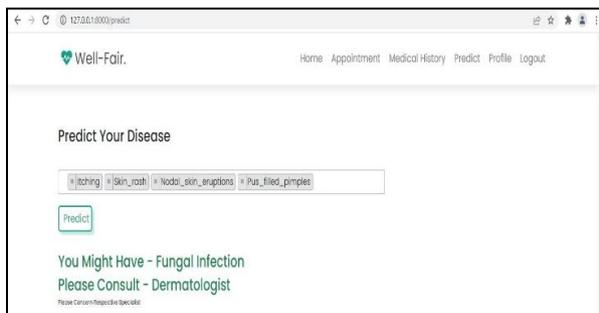


Fig.5 Disease Predictor Functionality.

VII. CONCLUSION

The main aim of this paper is to provide the doctors and patient with a very easy and simple user interface. We have developed a web application which provides the patients with automated appointment booking facility. This paper also deals with disease identification based on the symptoms provided as an input by the user which requires an integrated machine learning system. We have also provided electronic prescription section which reduces paper based prescription system. This research paper concludes that this virtual clinic system with ML is a boon for upcoming and today's world.

VIII. FUTURE SCOPE

Portal can be made for multiple clinics / hospitals. Filters can be included based on doctor's specialty, location as well as search feature can be added. Alert system can be added for reminding patients to take their medicines on time. Chemist can be added as another user who will deliver medicines after viewing the prescription uploaded by patient.

REFERENCES

- [1]. ShankarBirajdar, Snehal Dalvi, Gaurav Kumavat, Kavita Devadiga,, Prachi Kshirsagar, "Efficient Doctor-Patient Web Portal", International Journal of Engineering Research in Computer Science and Engineering (IJERCSE), Vol. 5, Issue. 3, March 2018.
- [2]. S. B. Choudhari, Chaitanya Kusurkar, Rucha Sonje, Parag Mahajan, "Design And Implementation Of A Patient Appointment And Scheduling System", International Advanced Research Journal in Science, Engineering and Technology (IARJSET), Vol. 4, Issue 12, December 2017.
- [3]. Raj H. Chauhan, Daksh N. Naik, Rinal A. Halpati, Sagarkumar J. Patel, Mr. A.D.Prajapati, "Disease Prediction using Machine Learning", International Research Journal of Engineering and Technology (IRJET), Volume 07, Issue 05, May 2020.
- [4]. Anuj Kumar, Mr. Anulp Pathak, "A Machine Learning Model for Early Prediction of Multiple Diseases to Cure Lives", Turkish Journal of Computer and Mathematics Education, Vol. 12, No. 6, 2021.
- [5]. Mesfin Leranso Betalo, "A Theoretical Approach to Electronic Prescription System", International Journal of Advancements in Research & Technology, Volume 7, Issue 11, November 2018.
- [6]. Chutisant Kerdvibulvech, Nwe Ni Win, "The Dentist Online Reservation System Design

- and Implementation Web Based Application and Database Management System Project” International Conference on Education Technology and Computer (ICETC2012) IPCSIT vol.43 (2012) © (2012) IACSIT Press, Singapore
- [7]. Archana Chaudhary, Savita Kolhe & Raj Kamal, “Machine Learning Classification Techniques: A Comparative Study”, International Journal on Advanced Computer Theory and Engineering (IJACTE), Volume-2, Issue-4, 2013.
- [8]. Prof. Swati Dhabarde, Rohit Mahajan, Satyam Mishra, Sanjiv Chaudhari, Satish Manelu, Prof. Dr. N S Shelke, “DISEASE PREDICTION USING MACHINE LEARNING ALGORITHMS”, International Research Journal of Modernization in Engineering Technology and Science (IRJMETS), Volume:04/Issue:03/March-2022.
- [9]. J.V. Joshua , G.A. Onanuga , K.N. Kodi, “Online Medical Prescription System (OMPS)”, International Journal of Latest Technology in Engineering, Management & Applied Science (IJLTEMAS), Volume VIII, Issue VIII, August 2019.
- [10]. “K-Fold Cross Validation in Machine Learning” -<https://www.javatpoint.com/cross-validation-in-machine-learning>.
- [11]. “Support Vector Machine Algorithm” - <https://www.javatpoint.com/machine-learning-support-vector-machine-algorithm>.
- [12]. “Naive Bayes Classifier Algorithm” - <https://www.javatpoint.com/machine-learning-naive-bayes-classifier>.
- [13]. “Random Forest Classifier Algorithm” - <https://www.javatpoint.com/machine-learning-random-forest-algorithm>.
- [14]. Dataset - <https://www.kaggle.com/datasets/kaushil268/disease-prediction-using-machine-learning>.

Paresh Devlekar, et. al. “Data Driven Machine Learning System for Optimization of Clinic Activities.” *International Journal of Engineering Research and Applications (IJERA)*, vol.12 (04), 2022, pp 40-45.