

AUTOMATED DISEASE PREDICTION AND MEDICINE RECOMMENDATION SYSTEM

Shrinidhi M

shrin7885@gmail.com

Department of ISE,

Don Bosco Institute of Technology,
Kumbalagodu Bengaluru, Karnataka , India

Shreas S

shreas12@gmail.com

Department of ISE,

Don Bosco Institute of Technology,
Kumbalagodu Bengaluru, Karnataka , India

Shreyas S

ishshreyass22@gmail.com

Department of ISE,

Don Bosco Institute of Technology,
Kumbalagodu Bengaluru, Karnataka , India

Abstract

The continuous pandemic has transformed the globe, resulting in a situation where the majority of previously manual tasks are now completed online. However, the method for determining the severity of an illness has remained unchanged. We want to develop an application that processes data, forecasts diseases, and assists in providing disease details before an emergency or catastrophe arises. The main goal of our project is to design a Disease Prediction Application. The application detects the diseases based on symptoms. The system proposes a detailed classification about the fetched disease and provides the fetched disease details. For reliable disease prediction, We practice the system getting to know Decision Tree (DT) and Logistic Regression (LR).A collection of disease symptoms is required for disease prediction. The method is simple to use even for those with less medical understanding, and it can help with early disease identification and diagnosis. It can also help people who are hesitant to go to the doctor when they have minor symptoms. This will give them a general picture of the disease's severity.

Keywords: Decision tree(DT), Logistic Regression(LR), Disease prediction, Machine Learning, Healthcare, Supervised Learning.

1. Introduction

While healthcare has not always been completely reliant on computers and technology, advances in recent years such as Machine Learning and Data Mining have made it much easier to anticipate and detect diseases in patients. Advanced computers have been developed that are capable of performing accurate surgery. Medical diagnosis has always been a difficult task for clinicians who must diagnose a large number of patients. In addition, with a little aid from technology, traditional physical disease diagnosis can be substantially streamlined, if not improved. Machine learning and data mining can be utilised in this case to effectively diagnose patients using historical data from real patients. Then, with the usage of ML, sufferers can diagnose and expect what situation they have. They can do it from the comfort of their own homes, by just selecting symptoms that they are experiencing. This study employs unique machine learning algorithms to detect disease based on symptoms and provide additional information about the disease.

2. MACHINE LEARNING TECHNIQUES

ML is used to remedy troubles wherein the connection among the enter and output is unknown. The automated acquisition of structural descriptions is referred to as learning.

2.4 Decision tree(DT) :

It is a supervised getting to know set of rules used typically for classification problems. DT constantly divide statistics into or extra homogenous sets, relying on a few pre-determined extra significant/essential variables to make as distinct/separate corporations as viable at each division.

It is vital to realize that at the same time as pretty clean to recognize and use, decision tree are clean to get over-fitted, and also, strictly require the information to be divided into 2 parts.

ID3 Algorithm:

The ID3 algorithm, created by J.R. Quinlan, is one of the most important algorithms we've employed in our research.

ID3 performs a greedy top-down search across the given columns, testing each column (attribute=symptoms) at each node and selecting the best attribute (symptom) for classification of a given set. ID3 uses Entropy and Information Gain to determine which symptom is ideal for building a decision Tree.

Entropy is the measure of how much uncertainty or unpredictability there is in a system. That is, it denotes the likelihood of a specific occurrence occurring.

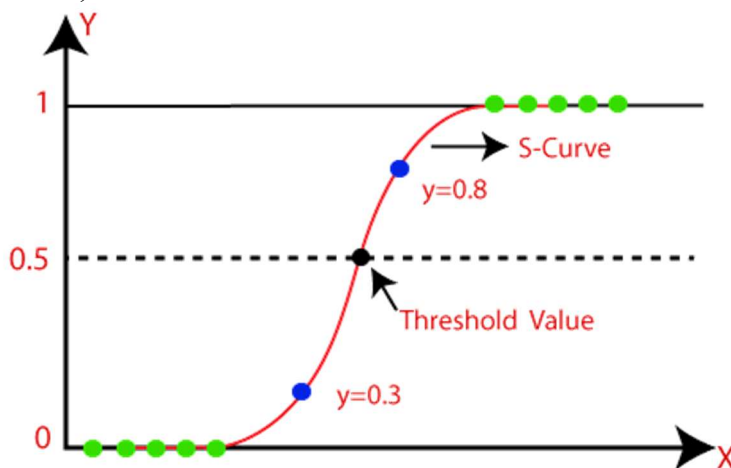
Information Gain: For a state, $IG(C, A)$ represents information gain (also known as Kullback-Leibler divergence). After finishing an attribute A, C is the effective change in entropy. It calculates the entropy change (drop) in relation to the symptoms.

ADVANTAGES

- 1) It is easy to apprehend because it follows the equal technique which a human observe even as making any choice in real life.
- 2) It may be very beneficial for fixing decision-associated problems.
- 3) It allows to reflect on consideration on all of the viable effects for a problem.
- 4) There is much less requirement of information cleansing in comparison to different algorithms.

2.5 Logistic regression(lr) :

The supervised learning technique consists of logistic regression, that is one in all the foremost common ML algorithms. It's a technique for predicting a specific structured variable from a hard and fast of impartial variables. A specific structured variable's output is expected the use of logistic regression. As a end result, the end result have to be a discrete or specific value. It is probably yes or no, zero or one, true or false, and so on.



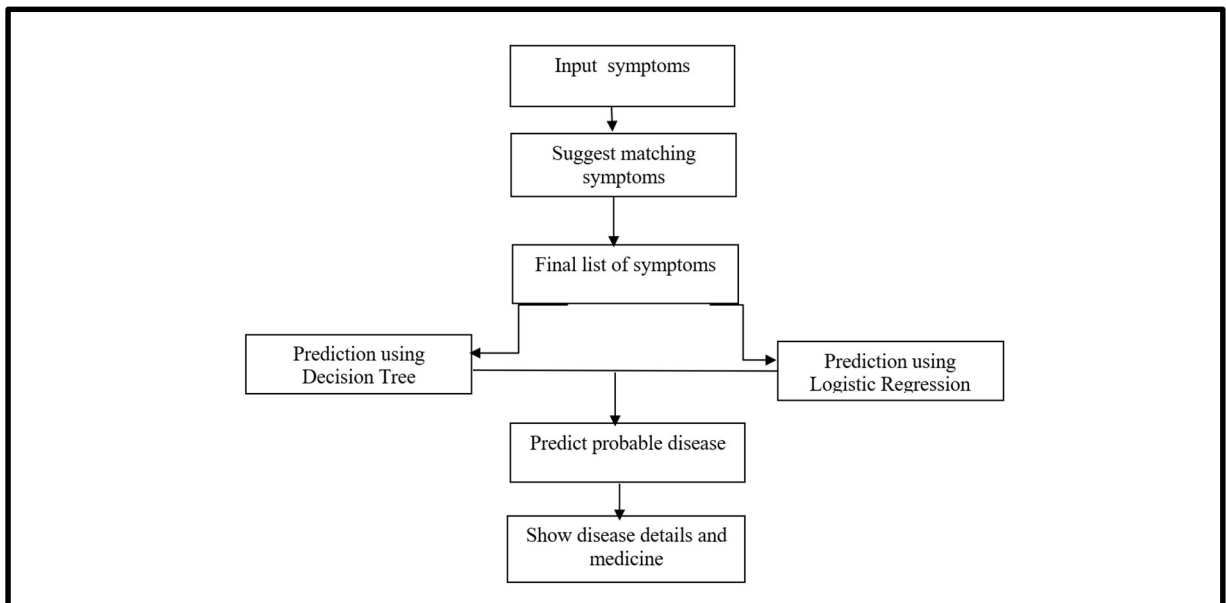
ADVANTAGES

1. Logistic regression is simpler to implement, interpret, and really economical to train.
2. It makes no assumptions approximately distributions of training in function space.
3. It will simply reach multiple categories and a natural probabilistic read of sophistication predictions.
4. It could be very speedy at classifying unknown data

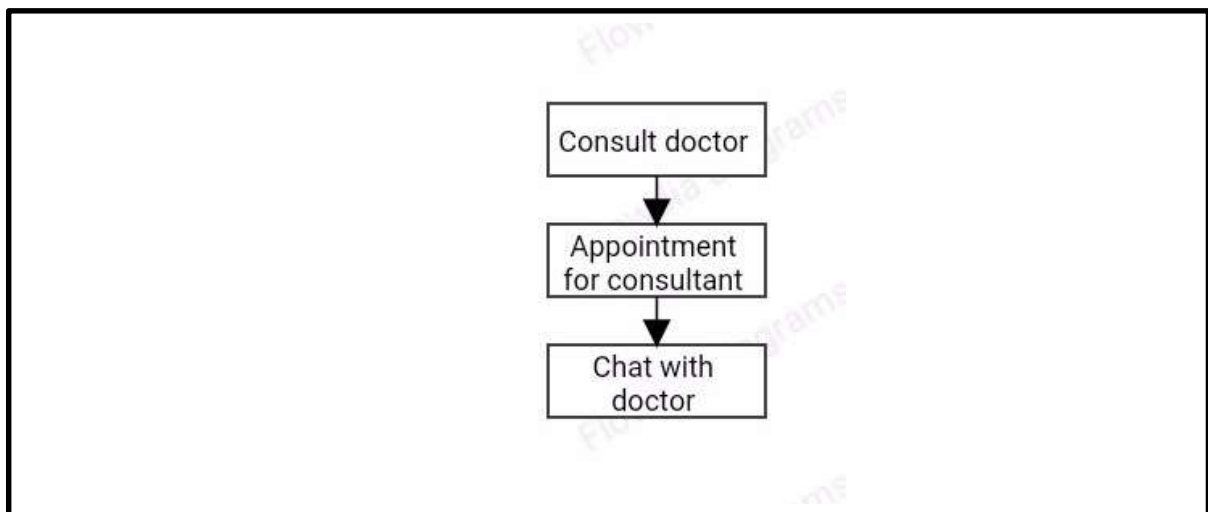
3. PROPOSED WORK

- We aim at building an android application that processes the inputs, predicts the different diseases and also helps in providing disease details before an emergency or catastrophe arises
- Creating a android application for digitalized communication between doctor and patient and to provide immediate remedies for their health issues. Our application aims to make Associate in Nursing atmosphere wherever numerous patients needing doctor facilitate at their home will consult doctor chat with doctors, inform then their problems and talk remedies.

WorkFlow:



Disease Prediction System Architecture



Doctor Consult Architecture

4. EXPERIMENTAL RESULTS

The Application is designed using Python and Java. Datasets consists of disease and symptoms. Diseases as target labels and symptoms as attributes. Datasets are collected from national health portal and Wikipedia. By cross validation accuracy logistic regression(LR) gives 89.19% accuracy and Decision Tree gives 83.57% accuracy so we decided to use logistic regression for our application. To connect android application and python we created flask server and deployed in Heroku and connected to android application.

Disease prediction

1) Accuracy Output of used algorithms in disease prediction

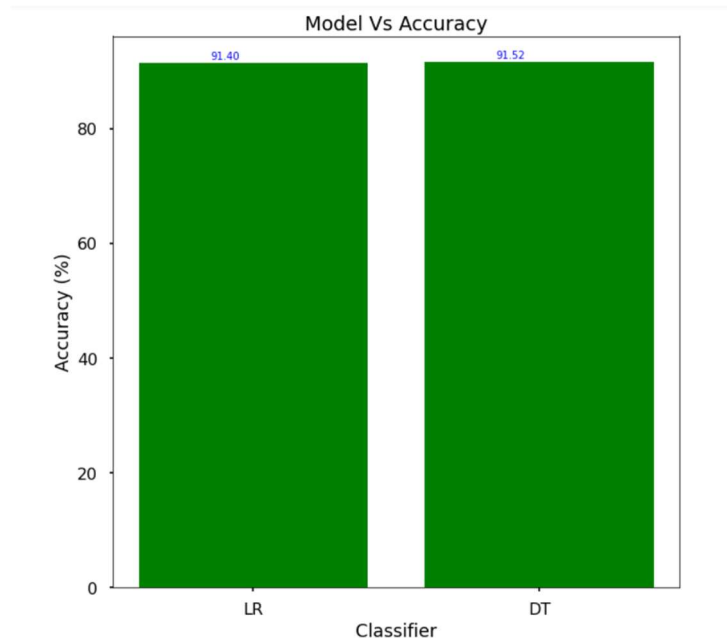


Figure 1

The figure 2 shows the comparison of Accuracy of algorithms used.

2) Cross validation accuracy of used algorithms in Disease prediction

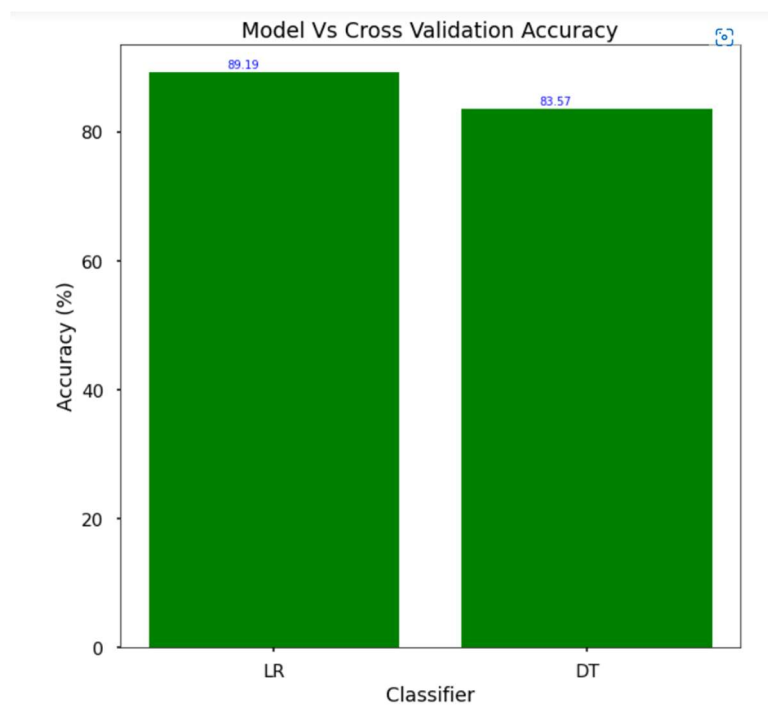


Figure 2

The figure 2 shows the comparison of cross validation accuracy of algorithms used.

3) The predicted output

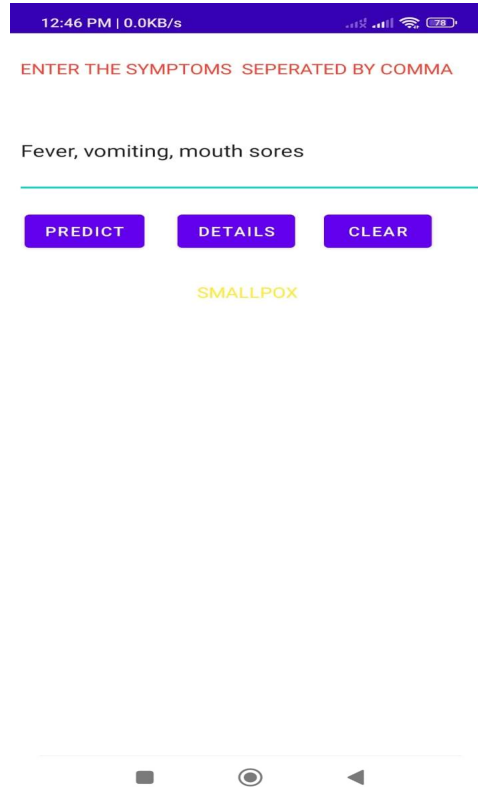


Figure 3

Figure 3 shows the predicted disease based on Symptoms entered by patient in Android application.

4) Disease Details

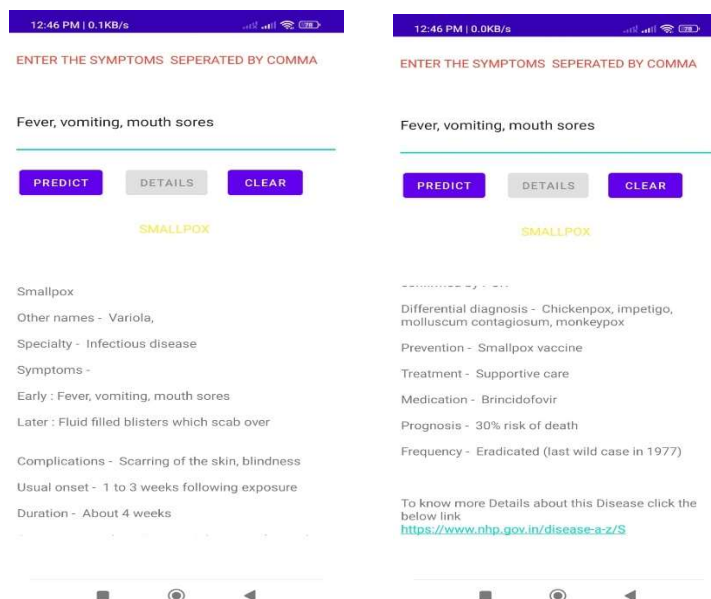


Figure 4

Figure 4 shows Providing details of fetched Disease

Doctor Consult

1) Patient taking appointment for consultation

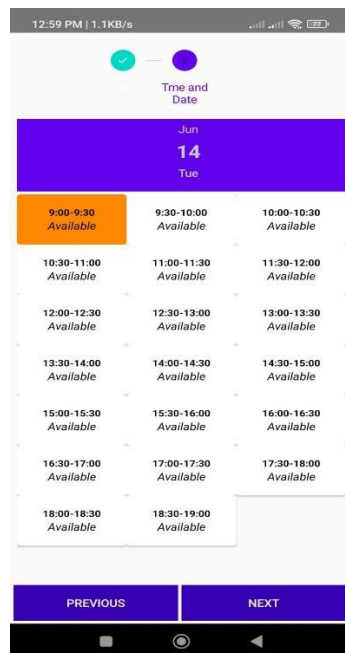


Figure 1

Figure 1 shows the patient taking appointment for consultation.

2) Doctor accepting the appointment



Figure 2.

The figure 2 shows the doctor accepting the appointment of patient.

3) Chat with Doctor

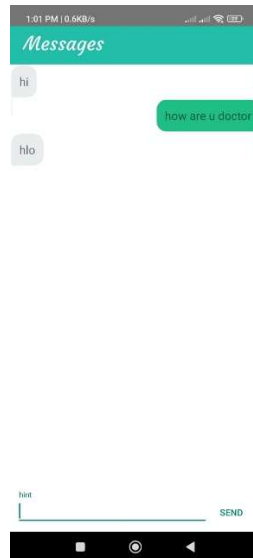


Figure 3.

The figure 3 shows the patient chatting with doctor.

5. CONCLUSION AND FUTURE SCOPE.

The predicting a disease by using symptoms and proper implementation of algorithms have proved that we can predict disease using symptoms. From above work I conclude that for disease prediction Logistic Regression is good with cross validation accuracy 89.19% compare to Decision Tree (83.57%).

6. ACKNOWLEDGEMENT.

I would love to explicit my unique thank you of gratitude to my guide “Mrs. Naganandini”, who gave me the possibility to try to this excellent analysis work on the “AUTOMATED DISEASE PREDICTION AND MEDICINE RECOMMENDATION SYSTEM” Project. which additionally helped me in doing plenty of studies and I got here to recognise many stuff. I am definitely grateful to my Guide.

7. REFERENCES

1. Dhiraj Dahiwade, Prof. Gajanan Patle, Prof. Ektaa Meshram , Designing Disease Prediction Model Using Machine Learning Approach 2019
2. Divya Krishnani, Anjali Kumari, Akash Dewangan, Aditya Singh, Nenavath Srinivas Naik, Prediction of Coronary Heart Disease Using Supervised Machine Learning Algorithms 2019
3. Sneha Grampurohit, Chetan Sagarnal , Disease Prediction using Machine Learning Algorithms 2020
4. Dr. P.Hamsagayathri , Mr .S. Vigneshwaran , Symptoms Based Disease Prediction Using Machine Learning Techniques 2021
5. S.Vijayalakshmi, Anshul Saini, Atulya Srinivasan, Nikhil Kumar Singh, Disease Prediction Over Big Data from Healthcare Institutions 2021
6. Rahatara,Ferdousi,M.Anwae,Hossain,Abdulmotaleb El Saddik,Early Stage Risk Prediction of Non Communicable Disease Using Machine Learning in Health CPS 2021