

Disease Predicting Online Web Portal using Decision Tree Algorithm and React

Neeshit Bundele¹, Manasi Hodavdekar², Archita Patil³

¹Student, Dept. of Computer Engineering, Rajiv Gandhi Institute of Technology, Maharashtra, Mumbai

²Student, Dept. of Computer Engineering, Rajiv Gandhi Institute of Technology, Maharashtra, Mumbai

Abstract : Every year, nearly 5 million deaths occur in India due to medical negligence, which includes errors like not receiving proper treatments in time or lack of information about the patient's medical history. For instance, if a person allergic to anesthesia is involved in a life-threatening accident, he might lose his life if the doctors fail to realize his allergy. But, if he/she has his Medicare on him, the doctors can access his medical history via the card or and prevent his death. This project aims to minimize such medical errors and in turn prevent these mortalities by generating a unique identity medical card that can be accessed by their unique 16-digit identification number and thereby making their medical history accessible in a dire situation. MedCare will provide the user with not only a physical card containing medical details, but also a digital access to the records through a portal. This portal will contain user's relevant medical information like basic patient details, medical history and documents that may be life-saving in certain situations. In addition to these functionalities, the MedCare portal will allow the user to have access to a disease prediction chatbot.

Key Words: Machine Learning, Chatbot, Medicare, REST API, Medical records, Online portal.

1.INTRODUCTION

In India, there has always been a dearth of technology aiding medical science. There are various incidents where people lose their lives or do not get proper resources for their medical needs. Additionally, India has a manual paper-based record maintenance of medical history of all the citizens. A lack of digital revolution in this field has led to a hard copy based physical files record keeping. Additionally, there is no facility for a doctor to view his previous diagnosis for a patient. Whenever a patient visits a doctor, the doctor needs to diagnose the patient and provide a proper medicine for his illness. Now, there might be situations wherein a doctor might want to check the previous medication provided to the patient when he/she was suffering from the same illness in the past. To achieve this, doctors maintain a log book or a register and manually write their important observations. Again, this is inefficient as it requires you to physically write down everything and maintain the records properly and also it leads to wastage of paper. These are the major setbacks in the medical field in India and it needs to be changed for the well-being of the masses. On the same line of

thoughts, the proposed system addresses these problems by aiding medical science by using concepts of databases, machine learning & data analytics. Additionally, if a disease prediction system can be made for patients it will save a lot of time. Generally, the patient is confused as a large amount of medical information on different mediums are available. The idea behind this system is to adapt to cope with the special requirements of the health domain related with users.

The project aims to minimize medical errors and in turn mitigate mortalities by generating a unique identity medical card that can be accessed by their unique 16-digit identification number and thereby making their medical history accessible in a dire situation. To register for MedCard, the user will have to input all his medical details after which, each user will be assigned a 16-digit unique identification number which will be the key for the user's medical information (i.e., basic details, medical history etc.). On completion of the registration process, the user will be able to access his medical records via an app or website using his 16-digit unique pin. In addition to digital access, the user will also be provided with a physical card. The records will be stored in our online repository and will be maintained using database technologies. The user will be able to log into the portal where he can view his/her medical history and access a disease prediction chatbot. The user will be able to access a chatbot for all his medical queries and will be provided with all the necessary precautions he/she needs to undertake based on his/her symptoms.

The scope of this project is to provide a framework to the user which includes such features that are much required for him/her to seamlessly deliver his past medical history to concerned doctors along with some additional features. The set of features provided will be: 1. Generate a physical card with a unique 16-digit code. 2. Search doctors by various filters. 3. Chatbot for disease prediction. 4. Platform where users can maintain past medical history and reports. The future scope of this project includes the addition of more features to make things much more convenient for the user. The features that can be added is analysis of users' medical information. Another feature that can be added is based on the analysis, the system can predict the prognosis or can suggest the doctor specialized in a particular field.

2) REVIEW OF LITERATURE AND ANALYSIS

2.1) Domain Explanation

2.1.1) Web Development

A) MERN Stack:

MERN Stack is a JavaScript Stack that is used for easier and faster deployment of full-stack web applications. MERN Stack comprises 4 technologies namely: MongoDB, Express, React and Node.js. It is designed to make the development process smoother and easier. Each of these 4 powerful technologies provides an end-to-end framework for the developers to work in and each of these technologies play a big part in the development of web applications.

1) React: Front-End Library: React is a JavaScript library that is used for building user interfaces. React is used for the development of single-page applications and mobile applications because of its ability to handle rapidly changing data. React allows users to code in JavaScript and create UI components.

2) Node.js: JS Runtime Environment Node.js provides a JavaScript Environment which allows the user to run their code on the server (outside the browser). Node pack manager i.e., npm allows the user to choose from thousands of free packages (node modules) to download. 3) Express: Back-End Framework: Express is a Node.js framework. Rather than writing the code using Node.js and creating loads of Node modules, Express makes it simpler and easier to write the back-end code. Express helps in designing great web applications and APIs. Express supports many middleware which makes the code shorter and easier to write.

4) MongoDB: Cross-platform Document-Oriented Database MongoDB is a NoSQL database where each record is a document consisting of key-value pairs that are similar to JSON (JavaScript Object Notation) objects. MongoDB is flexible and allows its users to create schema, databases, tables, etc. Documents that are identifiable by a primary key make up the basic unit of MongoDB. Once MongoDB is installed, users can make use of the Mongo shell as well. Mongo shell provides a JavaScript interface through which the users can interact and carry out operations (e.g.: querying, updating records, deleting records).

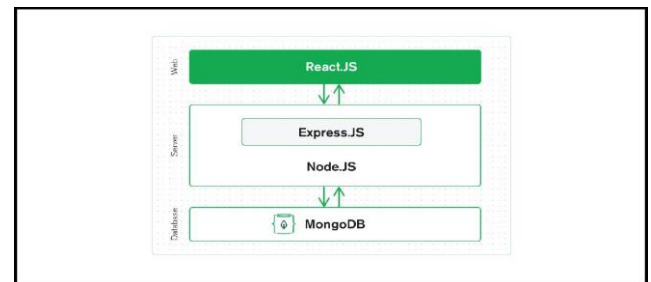


Fig 2.1 Block Diagram of MERN Stack

B) Flask: Flask is a web application framework written in Python. It was developed by Armin Ronacher, who led a team of international Python enthusiasts called Pocco. Flask is based on the Werkzeug WSGI toolkit and the Jinja2 template engine. Both are Pocco projects. Unlike the Django framework, Flask is very Pythonic. It's easy to get started with Flask, because it doesn't have a huge learning curve. On top of that it's very explicit, which increases readability. To create the "Hello World" app, you only need a few lines of code.

```
from flask import Flask
app = Flask(__name__)

@app.route('/')
def hello_world():
    return 'Hello World!'

if __name__ == '__main__':
    app.run()
```

Fig 2.2 Flask sample code

2.1.2) Machine Learning

Machine Learning plays a significant role among the areas of Artificial Intelligence (AI). During recent years, Machine Learning (ML) has been attracting many researchers, and it has been successfully applied in many fields such as medical, education, forecasting etc., Right now, the diagnosis of diseases is mostly from expert's decision. Diagnosis is a major task in clinical science as it is crucial in determining if a patient is having the disease or not. This in turn decides the suitable path of treatment for disease diagnosis. Applying machine learning techniques for disease diagnosis using intelligent algorithms has been a hot research area of computer science. This project throws a light on the comprehensive survey on the machine learning applications in the medical disease prognosis during the past decades.

2.1.2.1) Decision Tree algorithm

Decision Tree algorithm belongs to the family of supervised learning algorithms. Unlike other supervised learning algorithms, the decision tree algorithm can be used for solving regression and classification problems too. The goal

of using a Decision Tree is to create a training model that can use to predict the class or value of the target variable by learning simple decision rules inferred from prior data (training data).

In Decision Trees, for predicting a class label for a record we start from the root of the tree. We compare the values of the root attribute with the record's attribute. On the basis of comparison, we follow the branch corresponding to that value and jump to the next node.

Types of Decision Trees: Types of decision trees are based on the type of target variable we have. It can be of two types:

- i) Categorical Variable Decision Tree: Decision Tree which has a categorical target variable then it is called a Categorical variable decision tree.
- ii) Continuous Variable Decision Tree: Decision Tree has a continuous target variable then it is called Continuous Variable Decision Tree.

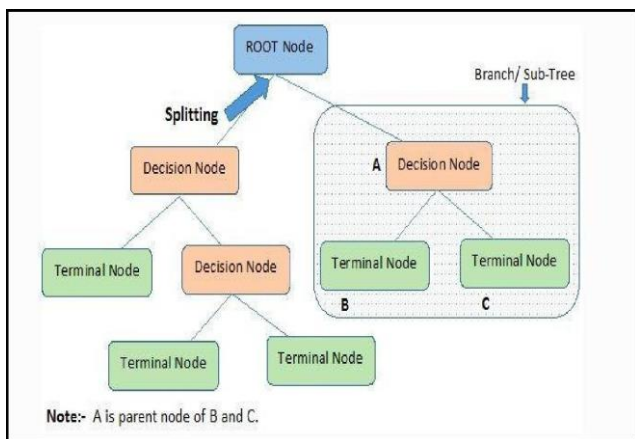


Fig 2.3. Decision Tree Architecture

2.1.2.2) Support Vector Machine Algorithm

Support vector machines (SVMs) are a set of supervised learning methods used for classification, regression and outlier's detection. The advantages of support vector machines are:

- i) Effective in high dimensional spaces.
- ii) Still effective in cases where the number of dimensions is greater than the number of samples.
- iii) Uses a subset of training points in the decision function (called support vectors), so it is also memory efficient.
- iv) Versatile: different Kernel functions can be specified for the decision function. Common kernels are provided, but it is also possible to specify custom kernels.

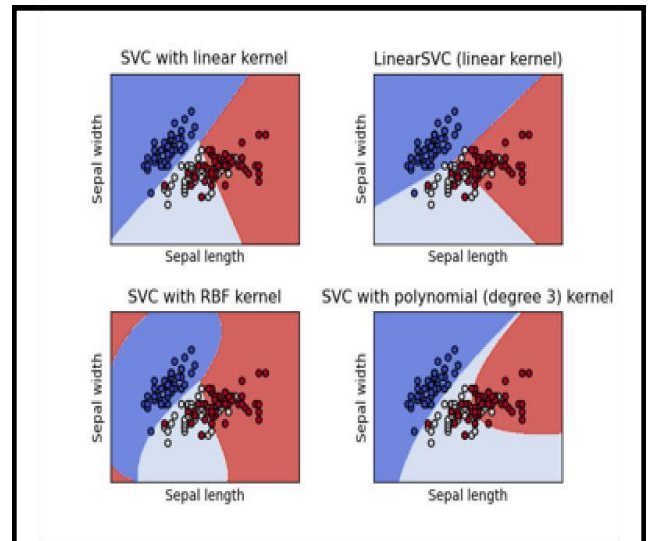


Fig 2.4 SVC Classifier

2.2) Existing Solution

In the previous system, medical records of all Indian citizens dispersed at various sites like dispensaries and hospitals were made available onto the doctor's terminal at a given point of time. The primary goal of the existing system was to provide digital records and amalgamate the records at one site whenever a patient visits a doctor. In addition to this, our system will provide additional features such as disease prediction with the help of a chatbot and we are also aiming to include statistical analysis of a patient's medical report.

2.3) Proposed System

In the proposed system, the entire medical record of all Indian citizens registered through the portal will be made available. The system revolves around an entity- physical card which provides a unique identification of every citizen. Using this entity, medical records can be analyzed and used for better medical diagnosis and assessment. The primary goal of the proposed system is to provide digital records, help patients identify a disease by inputting symptoms he/she is facing using the chatbot and amalgamate the records at one site whenever a patient visits a doctor. In addition to this, search for a doctor based on specialization and degree.

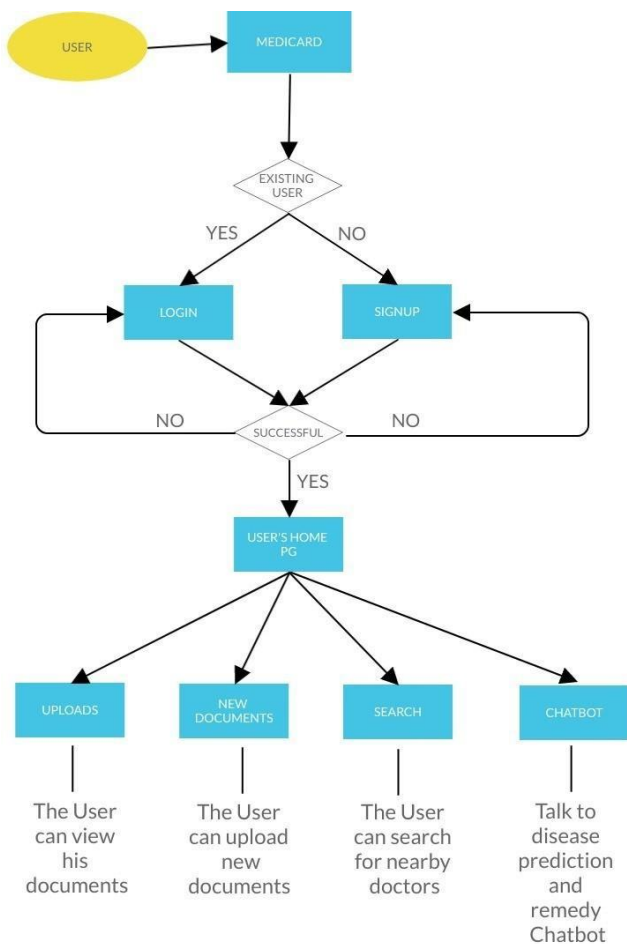


Fig 3.1 Architecture Diagram

3) DESIGN AND IMPLEMENTATION

3.1) Design Consideration for Designing our Software Model, we have incorporated MVC Design Architecture to plan our application and how the interaction between different models and routes occurs. In order to make our application reachable from different front-end views (Mobile and PCs) and enable cross-platform features, we have considered building our backend server as a RESTful API.

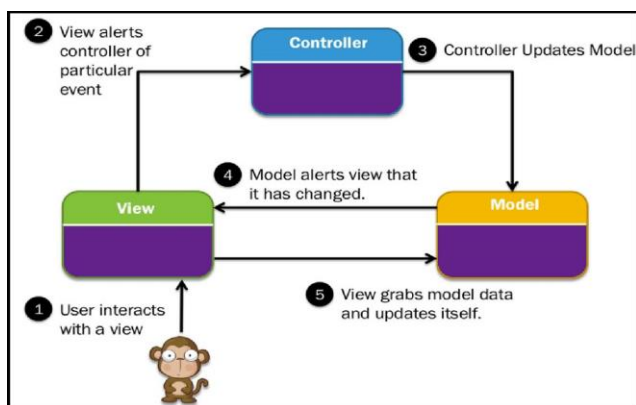


Fig 3.1 Block Diagram of MVC Pattern

The Model View Controller (MVC) design pattern specifies that an application consist of a data model, presentation information, and control information. The pattern requires that each of these be separated into different objects. MVC is more of an architectural pattern, but not for complete application. MVC mostly relates to the UI / interaction layer of an application. You're still going to need a business logic layer, maybe some service layer and data access layer. i) The Model contains only the pure application data; it contains no logic describing how to present the data to a user. ii) The View presents the model's data to the user. The view knows how to access the model's data, but it does not know what this data means or what the user can do to manipulate it. iii) The Controller exists between the view and the model. It listens to events triggered by the view (or another external source) and executes the appropriate reaction to these events. Since the view and the model are connected through a notification mechanism, the result of this action is then automatically reflected in the view.

REST API: The term REST stands for Representational State Transfer. It is an architectural style that defines a set of rules in order to create Web Services. In a client- server communication, REST suggests to create an object of the data requested by the client and send the values of the object in response to the user. For example, if the user is requesting for a movie in Bangalore at a certain place and time, then you can create an object on the server-side. So, over here, you have an object and you are sending the state of an object. This is why REST is known as Representational State Transfer.²⁵ The architectural style of REST helps in leveraging the lesser use of bandwidth to make an application more suitable for the internet. It is often regarded as the "language of the internet" and is completely based on the resources.

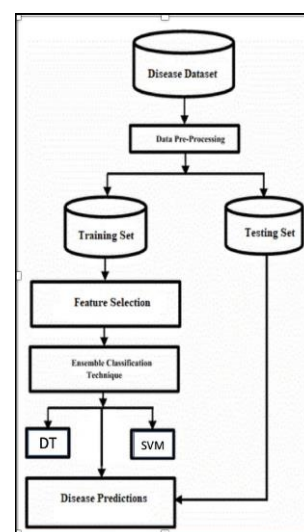


Fig 3.2 Disease Prediction Chatbot Architecture

3.2) Design Details the proposed system is a website application which provides universal access to the users depending on their roles. The website is created to provide

coherent access to all the users irrespective of their system specifications. In addition, the proposed system will also provide extensive insights by providing analytics of data sets as per the user needs. Entire functionality is divided into two primary users of the system as given below:

A) Doctors

The doctors are one of the primary users of the proposed system. Every doctor will have his own user terminal. The doctor simply logs into the website using his login-id and password. Further, whenever any patient approaches the doctor, he/she enters the patient’s code onto his/her terminal. Upon submitting this detail, entire medical

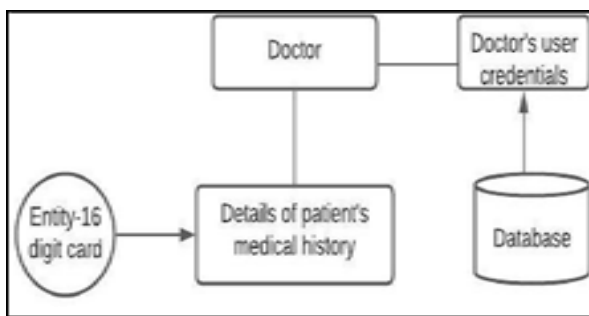


Fig 3.3 Functions of system available for doctors

records that are linked to this code are made available at the doctor’s terminal. However, the data of other doctors are not made available and every doctor gets details of patients persistent to their own diagnosis. The doctor gets essential information like the previous health issues, diseases, allergies, etc. of the patient.

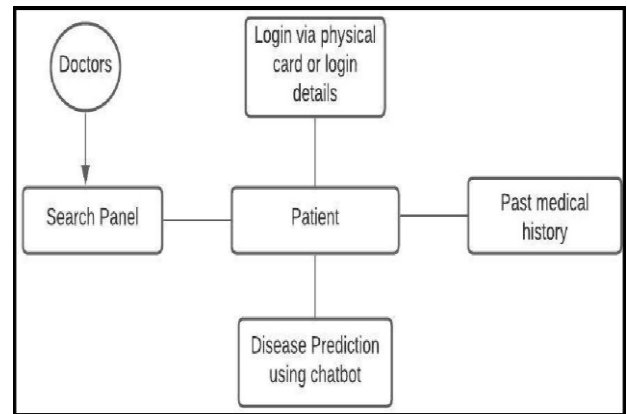
B) Patients

The patients are the other users of the proposed system. Every patient will enter his/her login-id and password to log onto their terminal. Initially, when the patients will register themselves, they will enter their personal details like name, age, sex, etc. as well as their previous allergies. They will also enter their Medicaid policies. The patient’s physical card will be particularly helpful in situations like accidents wherein knowing the previous medications and allergy details of a person can be helpful for the treatment. Additionally, the proposed system also provides a chatbot feature wherein a patient can enter his/her symptoms and find the chances of contracting a particular disease. Further, patients can also search for a doctor based on specialization and degree. This will provide added assistance to the users. Lastly, the entire medical history of every individual patient will be provided to that particular patient. This will help the patients to keep track of their past admits, allergies and medications for providing a proper medical history to doctors in the future. It also provides additional benefit that all the records are digitized. This removes the redundant need of hardcopy papers to maintain the records in the future. It also provides additional benefit that all the records

are digitized. This removes the redundant need of hardcopy papers to maintain the records in a file.

Fig 3.4 Functions of the system available for patients

3.3) Implementation Details



i) View and upload documents This module aims at providing the user with a platform for viewing and uploading his/her medical documents to the system. Users’ data is encrypted with encryption algorithms to ensure security for the data.

ii) Search Doctors this module aims at searching the doctors based on their specialization, Degree and by name. It accepts input from the user and queries the doctor from the database based on the filters.

iii) Disease prediction and remedy using Chatbot

A) Dataset Description The disease prediction dataset that we found is from a study of the University of Columbia. The table in the dataset is a knowledge database of disease-symptom associations generated by an automated method based on information of patients at New York Presbyterian Hospital admitted. The dataset contains 95 symptoms and 41 diseases. We have also used the symptom precaution dataset which advises users to take the necessary measures in order to recover from a particular disease. Apart from these, the symptom Description dataset gives a detailed overview of the disease and the symptom severity dataset provides information about the degree of severity.

B) Data Pre-processing There is a CSV document containing diseases and symptoms, named training.csv, which is utilized to prepare the model. Reads() function is utilized to store the information in the data frame, named df. Utilizing replace() function, prognosis column that are the different diseases, it is replaced by the numbers from 0 to n-1, where n is the number of different diseases present in the .csv record. Head() function is utilized to print the initial five rows of the preparation data frame.

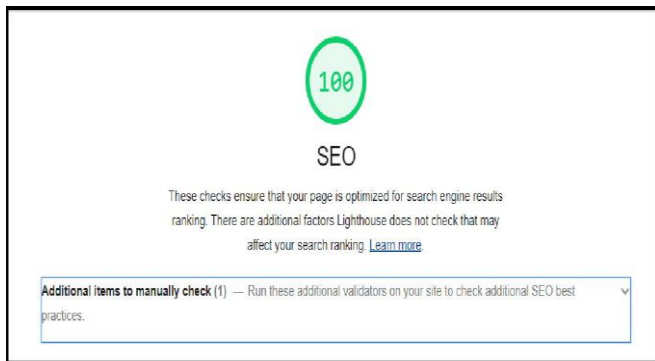


Fig 4.2 SEO report generated by Lighthouse

[3] Dahiwade, D., Patle, G., & Meshram, E. (2019). Designing Disease Predication Model Using Machine Learning Approach. Third International Conference on Computing Methodologies and Communication (pp. 1211-1215). IEEE.

5) CONCLUSIONS

There is a plethora of data related to healthcare in an unorganized form. Moreover, there is a dire need to come up with a remedy which can be universally acceptable. Hence, a solution to these problems is given by our project which includes disease prediction using machine learning. It also provides a solution to revolutionize the medical records maintenance procedure in India by establishing grounds to encompass digital records storage and maintenance. This work presents a system, which can be used as a professional system which can be used to view medical reports, upload latest documents, search for doctors based on their specialty and converse with a chatbot. One of the major hindrances in making a satisfactory healthcare portal available to one and all is the inaccessible nature of a well-developed online portal. This is the gap we have tried to fill.

5.1) Future Scope:

However, our chatbot does not cover all the diseases, so, there is a possibility to give an enhanced version of the chatbot by integrating this idea with a specialized disease prediction system and integrating a skin disease detection which would ease people to identify the skin diseases they are suffering from by merely clicking a picture of the patch that has been infected by parasites or some other disorders. Not just for skin disease, but also for other parts of the body. Additionally, patients can have virtual meetings with their respective doctors for follow ups. In future, with the advancements in technology, one could also get blood reports in place by providing samples to a report generating machine.

REFERENCES

[1] <http://people.dbmi.columbia.edu/~friedma/Projects/DiseaseSymptomKB/index.html>

[2] B., P., & M., M. (2016). Study of Machine Learning Algorithms for Special Disease Prediction Using Principal of Component Analysis. 2016 International Conference on Global Trends in Signal Processing, Information Computing and Communication (pp. 5-10). IEEE.