

DISEASE PREDICTION MODEL USING MACHINE Learning

¹Sahil Talreja, ¹Rohit Talwar, ¹Khushboo Bhatia, ¹Sagar Raheja ²Anjali Yeole

¹Day Scholar, ²Assistant Professor

Department of Computer Science Engineering, Vivekanand Education society's Institute of Technology, Chembur, Maharashtra, India.

Abstract—The objective of the system is to identify symptoms of patients and then prediction of the diseases. Identification of similar patients cases is useful for improving patient outcome, for treatment or disease recommendation to all patients, prediction of very efficient outcome, research on those cases. The world is moving at a fast speed and in order to keep up with the whole world we tend to ignore the symptoms of disease which can affect our health to a large extent. The correct prediction of the disease now-a-days is the most challenging task. We proposed disease prediction using symptoms. For the disease prediction, we use Naive Bayes machine learning algorithm for accurate prediction of disease. In our System disease prediction required disease symptoms dataset. The accuracy of general disease prediction by using Naive Bayes is 88%. After general disease prediction, this system is able to search for the nearest doctors and users can book an appointment. Once your appointment is booked then it will show the nearest medical shops. You can also be able to purchase medicines from it.

Index Terms: Naive Bayes, Machine learning, Disease Prediction.

I. INTRODUCTION

Artificial Intelligence can enable the computer to think and make computers more intelligent. AI study considers machine learning as a subfield in numerous research work. Numerous analysts feel that without machine learning, awareness can't be created. There are many kinds of Machine Learning Techniques like Unsupervised, Semi Supervised, Supervised, Reinforcement, Evolutionary Learning and Deep Learning. These techniques are used to classify a lot of huge data very fastly. So we use Naive Bayes machine learning algorithm for classification of big data and accurate prediction of disease.

Because medical data now-a-days is increasing day by day. So usage of that medical data for predicting correct disease is an important task but processing data is very important data mining plays a very important role and classification of large dataset using machine learning becomes very easy.

One can possibly use Machine Learning many times a day. Today Machine Learning is present everywhere. Some machine learning algorithms only work on structured data and time required for computation is high. Also they are lazy because they store entire data as a training dataset and use complex methods for calculation.

The section I explains the Introduction of general disease prediction using classification methods such as Naive Bayes. Section II presents the literature review of system and Section III presents proposed system architecture and state diagram Section IV presents results and discussion of the system. Section V concludes the system. While at the end references paper are presented

II. LITERATURE REVIEW

The paper "Predicting Disease by Using Data Mining Based on Healthcare Information System" [1] applies the information mining process to predict high blood pressure from patient medical records with eight alternative diseases. Under-sampling technique has been applied to come up with coaching knowledge sets, and data processing tool wood hen has been wont to generate the Naive Bayesian and J48 classifiers created to improve the prediction performance, and rough set tools were wont to scale back the ensemble supported the concept of second-order approximation. But the choice trees generated by J-48 is typically lacking within the leveling therefore the overall improvement of victimization ensemble approach is a smaller amount.

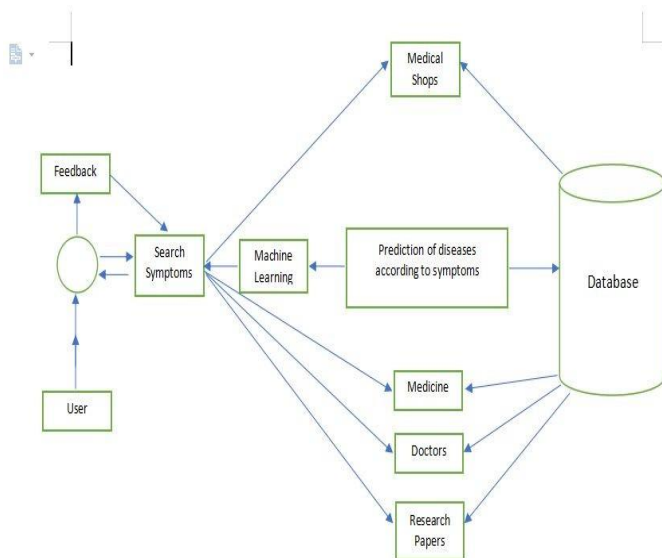
Author[2] has proposed a best clinical decision-making system which predicts the disease on the basis of historical data of patients. In this predicted multiple diseases and unseen pattern of patient condition. Designed a best clinical decision-making system used for the accurate disease prediction on the historical data. In that also determined multiple diseases concept and unseen

pattern. For the visualization purpose in this used 2D/3D graph and pie Charts.And 2D/3D graph and pie charts designed for visualization purpose.

Author has concluded that machine learning algorithms such as Naive Bayes and Apriori [3] are highly useful for disease diagnosis on the given data set. Here small volume data used for prediction like symptoms or previous knowledge obtained from the physical diagnosis.

Nowadays medical data is growing so needs to classify that and classification of that data is challenging.

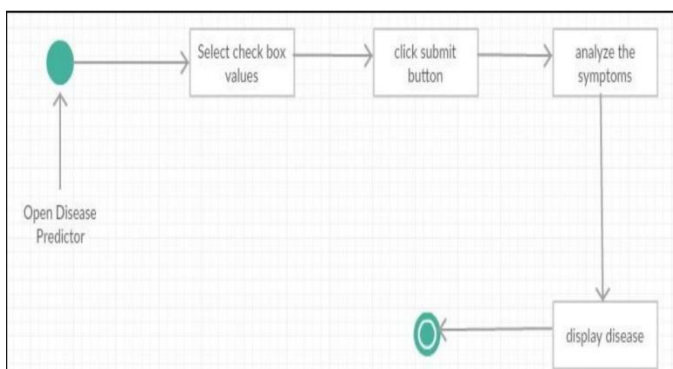
III. SYSTEM ARCHITECTURE



A. Architecture Overview

Fig 1. System Architecture

Fig 1 shows the system users will enter the symptoms and then there will be the prediction of diseases. Based on machine learning we can predict accurate disease. Users will also be able to find the nearest doctors. They will also find the nearest pharmacist store. The data of medicines and pharmacy is stored



in the database. Fig 2. State Diagram

Fig 2 shows a state diagram that explains how the disease prediction will work. As the user opens the login page and will click the diagnosis button user will click the symptoms he/she

is having. Then the diagnosis will take place according to naive bayes algorithm and then the system will analyze the symptoms and the most probable disease will be displayed.

B. Algorithm Used

The alternative algorithms are Decision making, Random Forest and Naive Bayes. We have selected Naive Bayes because it is easy to implement ,it requires less training data,it can make probabilistic predictions and easy to update on arrival of new data.

	Naive Bayes	Random Forest	Decision Tree
Unigram	44.01%	43.53%	32.74%
bigram	34.82%	33.75%	28.4%
trigram	24.66%	23.47%	24.1%
-/bi-/trigra m	45.22%	43.93%	34.48%

Fig 3.Average Classification Accuracy

Fig 3. Illustrates that the average values of classification accuracy of Naive Bayes, Random Forest and Decision Tree are similar and Naive Bayes has achieved 1-2 % higher average classification accuracy results in comparison to Random Forest and Decision tree.

Naive Bayes

Naive Bayes algorithm is an algorithm that checks the probability of an object with some features which belong to a particular group. For example, if somebody is attempting to spot a vegetable on its color, shape and taste, then a red colored, spherical, round, and sweet vegetable will be most likely a tomato. All these properties contribute to the probability that the fruit is an orange and that’s why it is known as “naive”. As for the “Bayes” part, it refers to statistician and philosopher, Thomas Bayes and the theorem named after him , Bayes’ theorem , which is the base for Naïve Bayes Algorithm. More formally, Bayes ‘Theorem is stated as following equation:

$$P(A/B) = (P(B/A)*P(A)) / P(B)$$

IV. RESULT AND DISCUSSIONS



Fig 4 Home Page

Fig 4 shows the home page of the system. Users first have to sign up and if the user has signed up, the user can directly login to the system.



Fig 5 User type

As we can see in Fig 5 the user type page for the system. They are Customer, Doctor and pharmacy.

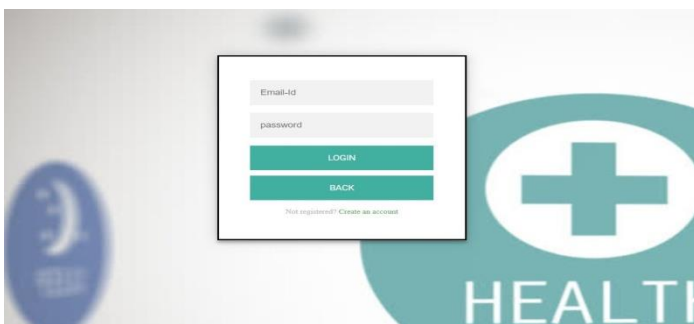


Fig 6 Sign in page

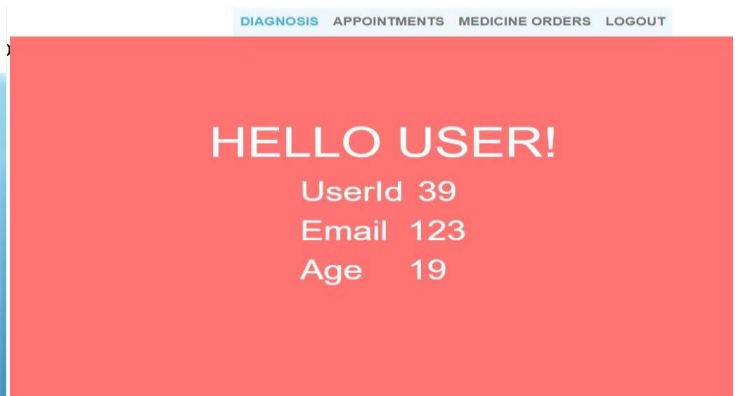


Fig 7 Dashboard for user

Fig 7 shows the dashboard for users where the page that users will have options like diagnosis, appointments, Medicine orders and logout option.

SELECT THE SYMPTOMS FROM THE FOLLOWING SELECT BOXES



Fig 8 Selecting Symptoms

Fig 8 shows selection of symptoms. Diagnosis button will redirect to the page shown in Fig 9.

TOPMOST THREE PREDICTED DISEASES

- 1. COMMON COLD
- 2. FAILURE HEART CONGESTIVE
- 3. ASTHMA

Enter the 4 areas that are most convenient to you for visiting a doctor :

*We will give you suggestions on the pharmacies/chemist stores in these areas as well



Fig 9 Prediction Page

And on that page only users will have to enter his nearest 4 areas that are convenient for him/her to visit a doctor.

MOST PROBABLE DISEASE : COMMON COLD

SPECIALIST(S) REQUIRED TO TREAT THE DISEASE :

- 1. General Physician
- 2. Psychologist
- 3. Internist

SUBMIT

SPECIALITY	DOCTOR NAME	AREA	CONFIRM(Y/N)
General physician	Dr. Subhashini Venkatesh	Anna Nagar	<input type="text"/>
General physician	Dr. Nirmala Jayashankar	Kilpauk	<input type="text"/>
Psychologist	Dr. Hema Karthik	Anna Nagar	<input type="text"/>
Dentist	Dr. Subashini	Anna	<input type="text"/>

Fig 10 Selecting Specialists/Doctors

Fig 10 shows the most probable disease and then after that users will be shown the specialists required to treat that disease and the doctors also. Users have to put Y or N in the Text box to select an appointment with the doctor.

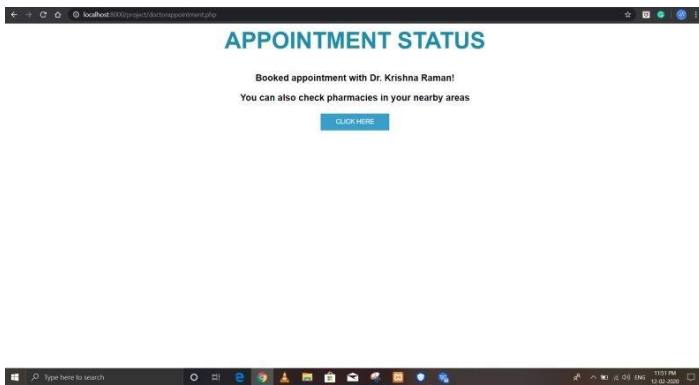


Fig 11 Appointment status

Fig 11 shows the appointment status which is booked and after that user will be able to select the pharmacy store near that doctor.



Fig 12 Selecting Pharmacies

Fig 12 shows users can select the pharmacy store and then submit it. Then again it will redirect you to the home page where you came after login. Users will be able to see whether they have booked a user's appointment or not with the doctor and the user has ordered medicine from the pharmacy store or not

V. CONCLUSION

We proposed a general disease prediction system based on machine learning algorithms. We utilized Naive Bayes algorithm to classify patient data because today medical data is growing very vastly and that needs to process existing data for predicting exact disease based on symptoms. We got accurate general disease risk prediction as output, by giving the input as patients record which help us to understand the level of disease risk prediction.

REFERENCES

- [1] Feixiang Huang, Shengyong Wang, and Chien-Chung Chan, "Predicting Disease By Using Data Mining Based on Healthcare Information System", in IEEE 2012.
- [2] Disease and symptoms Dataset –www.github.com
- [3] .Allen Daniel Sunny1, Sajal Kulshreshtha, Satyam Singh3, Srinabh, Mr. Mohan Ba, Dr. Sarojadevi H "Designing a disease prediction model using machine learning", *International Journal of Innovations in Engineering and Technology (IJJET)* Volume 10 Issue 2 May 2018.