CONNECTING NEAR BY DOCTORS BY IDENTIFYING & DIAGNOSING DISEASES

1 Lakshmi S V S S , 2Ninisha G, 3 Pravallika AY , 4 Akash KP

1,2,3,4Department of CSE, Anil Neerukonda Institute Of Technology And Sciences(A), Visakhapatnam-531162, India

Abstract: In this we are going to implement a system which is related to smart health prediction inorder to reduce the time of a user. The main objective of developing this project is to provide a proper medical guidance to the patient for their health issues by providing accurate results. Here, the system concentrates on the symptoms of patient’s disease and based on the symptoms, the data is classified from the dataset and finally the disease name is predicted. We have 3 modules, they are patient, doctor, and admin modules. The primary step is, the user(patient/doctor) needs to register if he is new to the application or else he can directly login with his credentials. The admin authenticates the credentials and allows the user to access. There is a facility that user can upload records and there is a feedback system available to the user which is directly visible by the admin, so, the doctor can study the patient and give a proper treatment.

Index Terms: Disease prediction, Doctor, Patient, Symptoms, Feedback, Appointment, Admin.

1. Introduction

Human beings are getting affected by various diseases now a days and this has become a common criteria. They are getting affected by different diseases and neglecting the cause because of enlarged time taken to know the type of disease in their busy lives and this is becoming a critical situation when it goes to a chronic stage. This system reduces the tedious task of an applicant who apply manually by waiting for hours in the queue. So, in order to overcome the above mentioned problems, we are developing this proposed system to provide an ease of access to the victims of the disease through online by saving their time. The two domains which we are using for this system is Data mining and machine learning. Data Mining refers to the extraction of raw data from data warehouses into a useful one. Machine Learning is the ability of making the machine to learn.
Machine Learning consists of two types of Learning. They are:

- **Supervised Learning**
- **Unsupervised learning**

**Supervised Learning**

Learning something with the help of some guide comes under Supervised Learning. This learning process is completely dependent. Here, the input vector is presented to the network which will produce an output vector. The main aim is, the output vector and the target vector i.e, the desired output must be equal or else the adjustment of weights will continue till the outputs match with each other.

**Unsupervised Learning**

Learning is performed without the help of a guide. Here, the network receives n number of input patterns and organise it into a form of clusters. The network itself has the ability to learn by itself, discover the patterns & features from input data. This type of learning is also called as Self-Organised Learning.

2. Literature Survey

In this system we are using naïve bayes theorem for classification of data and disease prediction. Naïve Bayes classifiers are defined as a collection of classification algorithms based on **Bayes’ Theorem**. The main principle is each instance is independent of each other. Naïve Bayes Algorithm is an example of supervised Learning i.e. it is completely trained by examples in the dataset. The most probable classification of new instance is obtained by combining the predictions of hypothesis. There are two events A and B where P (B) is true. P (A) is the **priori** of A (the prior probability, i.e. Probability of event before evidence is seen). The probability of event after evidence is seen as,

\[ P(A/B) = \frac{P(B/A) \times P(A)}{P(B)} \]  

(1)
1) **Naïve Bayes Algorithm**

Naïve Bayes Algorithm is a classification algorithm based on Bayes theorem’s use in predictive modelling and this algorithm uses Bayesian techniques. This Algorithm is less computationally intense then other algorithms.

Requirements for naïve bayes models:

* **Single key column**: Each model must contain unique column which uniquely identifies each record.

* **Input columns**: In Naive Bayes model, it is important to ensure that the input attributes are independent of each other.

* **At least one predictable column**: The values of the predictable column can be treated as inputs.

2) **Decision Tree Algorithm**

It comes under one of the type of supervised learning algorithm which is especially used in classification problems. It works on both continuous as well as categorical input and output variables. This comprises of a single root node with similar set of different nodes.

Decision tree is classified into 2 types. They are:

1) **Categorical Variable Decision Tree**: If a tree consists of a categorical target value then it is termed as Categorical variable decision tree.

2) **Continuous Variable Decision Tree**: If a tree consists of a continuous target variable then it is termed as Continuous Variable Decision tree.

* **Working of a Decision Tree**:  

A decision tree is a graphical representation of all possible solutions to a decision based on certain conditions. Each path of the tree is from root to leaf.

Decision tree learning characteristics:

1) Instances are represented by attribute pair values. Attributes takes on a small number of disjoint possible values.

2) The target function has some discrete output values. It even assigns a Boolean classification (Yes or No/0 or 1) to each example.

3) Disjunctive descriptions may be required.

4) Decision tree learning methods are robust to errors.

5) The training data may contain missing attribute values.
3) **Random forest Algorithm:**

Random forest algorithm is the term for an ensemble classifier which consists of many decision trees and therefore outputs the class that is the mode of the classes output by individual trees. Random forests are set of trees, all slightly different. It randomizes the algorithm, not the training data and generally improves the decisions of decision trees. Every decision tree will be constructed by using a Random subset of the training data from the dataset. To improve the predictive accuracy and control over-fitting, random forest acts as an estimator that fits various samples of dataset. Random Forests are especially used for predicting continuous variables. These are also used for estimating the probability that a particular outcome occurs, Outcomes can be either “yes/no” events i.e. True or false. It could have many possible outcomes but typically multi-class problems have 8 or fewer outcomes.

Every tree generates its own specific predictions and acts independent for each record in the database. For predicting we would average the predictions made by our trees. By yes vs no percentage we can know the prediction.

**Drawbacks of other algorithms:**

One more algorithm is also there for implementing this framework namely Hill climbing algorithm but it has some drawbacks when compared to Naïve bayes algorithm. The following are the disadvantages of hill climbing algorithm.

- It got failed to find a better solution.
- Either algorithm may terminate not by finding a goal state but by getting to a state from which no better state can be generated.
- Not an efficient method- It gets stuck at all peaks.

**3. Methodology:**

Features Available in the System are,

- Patient Registration & Login.
- View Details.
- Diseases Prediction.
- Admin Login

---

**Fig 2.1:** Decision Tree sample diagram
• Notification.
• Feedback
• Automatic Location Tracker

In this system there are three modules. They are:

MODULES:

MODULE 1: ADMIN MODULE

• Admin Login: Admin can login into the system by using Id and password.
• Add Disease: Admin has the ability to add disease details along with symptoms and its type.
• View Disease: Admin can view details stored in database of different diseases.
• View Patient: Admin can view patients details along with their diseases.
• View Feedback: Admin can view the feedback of the user at the end.
• Add Doctor: Admin can add the Details of the doctors.
• View Doctor: Admin can view doctor details.

MODULE 2: PATIENT MODULE

• Patient login: Patient can use his ID and password to login to the system.
• Patient registration: If patient is a new user then he has to enter his personal details or else he can directly give his credentials to login.
• My details: Personal data can be viewed by the patient.
• Edit patient record: Personal details of the patient can be edited.
• Disease prediction: System generates certain random questions regarding his illness and it predicts the disease based on the symptoms using naive bayes specified by the patient. The system will also suggest the specialized doctors based on the disease.
• Search Doctor with known disease: Patient can search for a particular doctor by specifying Disease name/type.
• My Stuff: Patient has the facility to manage by uploading their reports.
• Appointment: Patient can proceed for appointment after selecting their doctor from list.
• Feedback: Patient will post the feedback and this will be reported to the admin.
MODULE 3: DOCTOR MODULE

- **Doctor login**: By using id and password doctor can login to the system.
- **Doctor registration**: If Doctor is a new user he will enter his personal details and he will have user id and password through which he can login to the system.
- **Appointment dates/slots**: Doctor have to provide his available dates and slots for treatment.
- **Connected patient details**: Doctor can view his connected patient details after completion of appointment.

![Fig 3.1: Architecture Of System](image)

3.1.1. DATASET DETAILS

In this module the dataset is gathered which is a collection of diseases and their corresponding symptoms. We have created these datasets which consists of diseases dataset and random questions dataset.
3.1.2. PERFORMANCE MEASURES

The performance for Diagnosing a disease and connecting nearby doctor is measured by time and accuracy.

**Performance measure with respect to Time:** The performance with respect to time is measured by plotting the graph by considering total time and people.

- X-axis represents number of people and Y-axis represents Time in minutes.

**Table 3.1.2(a):** The Time taken by 5 people in real world approach is as follows:

<table>
<thead>
<tr>
<th>People</th>
<th>Time taken to Appointment</th>
<th>Time taken to Connect doctor</th>
<th>Time taken to know Disease</th>
<th>Total Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>120mins</td>
<td>300mins</td>
<td>200mins</td>
<td>620mins</td>
</tr>
<tr>
<td>2</td>
<td>240mins</td>
<td>350mins</td>
<td>300mins</td>
<td>890mins</td>
</tr>
<tr>
<td>3</td>
<td>300mins</td>
<td>420mins</td>
<td>320mins</td>
<td>1040mins</td>
</tr>
<tr>
<td>4</td>
<td>240mins</td>
<td>300mins</td>
<td>350mins</td>
<td>890mins</td>
</tr>
<tr>
<td>5</td>
<td>360mins</td>
<td>450mins</td>
<td>400mins</td>
<td>1210mins</td>
</tr>
</tbody>
</table>

**Table 3.1.2(b):** The Time taken by the DDCND to compute is as follows:

<table>
<thead>
<tr>
<th>Users</th>
<th>Time taken to predict Disease</th>
<th>Time taken to get list Of doctors</th>
<th>Time taken to fix an appointment</th>
<th>Total Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3mins</td>
<td>50sec</td>
<td>10sec</td>
<td>4mins</td>
</tr>
<tr>
<td>2</td>
<td>2mins</td>
<td>50sec</td>
<td>10sec</td>
<td>3mins</td>
</tr>
<tr>
<td>3</td>
<td>3mins</td>
<td>50sec</td>
<td>10sec</td>
<td>4mins</td>
</tr>
<tr>
<td>4</td>
<td>4mins</td>
<td>50sec</td>
<td>10sec</td>
<td>5mins</td>
</tr>
<tr>
<td>5</td>
<td>2mins</td>
<td>50sec</td>
<td>10sec</td>
<td>3mins</td>
</tr>
</tbody>
</table>
3.1.3. Performance measure with respect to accuracy:

The performance with respect to accuracy is measured by plotting the graph disease given by doctors to corresponding patients.

Table 3.1.3: The diseases given by different doctors for the same patient are as follows:

<table>
<thead>
<tr>
<th>Patient</th>
<th>Disease given by doctor1</th>
<th>Disease given by doctor2</th>
<th>Disease given by doctor3</th>
<th>Disease given by doctor4</th>
<th>Disease given by doctor5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Viral fever</td>
<td>Dengue</td>
<td>Malaria</td>
<td>Dengue</td>
<td>Viral fever</td>
</tr>
</tbody>
</table>

Fig 3.1.3: Pie diagram related to the above table

4. Conclusion

In this proposed system, we develop a system to mine or extract the data which is very important because we are getting valuable information which is not available easily and all this information are real time information. Finally, we want to tell you that this is a fully developed unique system which will be helpful for us. Hope this system will be very demandable in future. This system involves fundamental parts, for example, basic login, enter symptoms in the system and recommend medications, proposes an adjacent specialist. It takes the contribution of different manifestations from the patient, does the examination of the entered symptoms and gives fitting sickness expectation.

References