2015 International Conference on Technologies for Sustainable Development (ICTSD-2015), Feb. 04 – 06, 2015, Mumbai, India

Electronic Recording System (Heart Disease Prediction System)

Mrs. Sana Shaikh[1], Amit Sawant[2], Shreerang Paradkar[3], Kedar Patil[4]
Department of computer engineering,
Don Bosco Institute of Technology, Mumbai, India
[1]sana.siraj.shaikh@gmail.com, [2]sawant.amit39@gmail.com, [3]shreerang.scp@gmail.com, [4]falkan07@gmail.com

***Abstract*— In medical field there is a growing need for development of software to be used for practical applications rather than only in research field. Doctors need not only to know the main symptoms in a patient but also the other factors in life and surrounding that lead to disease in the patient. For collecting all the necessary information about the patient the doctors needs to spend a lot of time with the patient asking necessary questions to get the important details. The doctor needs to document all the data for further use and needs to keep track of all the patient details and remedies. The proposed system is intended to develop an Intelligent System using data mining modeling technique, namely, Naive Bayes. It is implemented as Java application in which user answers the predefined questions. It retrieves hidden data from stored database and compares the user values with trained data set. It can answer complex queries for diagnosing heart disease and thus assist healthcare practitioners to make intelligent clinical decisions which traditional decision support systems cannot.**

Keywords—Data Mining, Naïve Bayes, Heart disease, Prediction, Remedy.

I. INTRODUCTION

Nowadays to live luxurious life people work like machine in order to earn lot of money hence they forget to take care of their health. Because of this, there is change in the food which they consume, their lifestyle changes. They are more tensed and in very much pressure to earn money so this leads to blood pressure, diabetes and various other diseases at young age. [1]

All these reasons lead to negligence of their health which increases the chances of heart disease. Heart is the most essential organ of the human body and if it gets affected then it also affects the other major organs of the body.

The aim of proposed system is to design an automated system which would manage complete clinical details, patient's history and their appointment details in a single database. Doctors will use this system to keep track of the patient consulting to them. The intentions of the proposed system are to reduce over-time pay and increase the number of patients that can be treated accurately.

It is a prototype Heart Disease Prediction System (HDPS) using data mining techniques, namely, Decision Trees, Naïve Bayes. After successfully predicting whether the patient is prone to heart disease or not, the proposed system will provide with the set of questions which the patient will have to answer. Depending upon the answers given by the patient the pattern matching algorithm will generate the corresponding remedy.

II. LITERATURE REVIEW

A. Existing system

The Heart Disease Prediction System (HDPS) can discover and extract hidden knowledge associated with diseases from a heart disease database. It can answer complex queries for diagnosing disease and thus assist healthcare practitioners to make intelligent clinical decisions which traditional decision support systems cannot. By providing effective treatments, it also helps to reduce treatment costs. To enhance visualization and ease of interpretation, it displays the results in tabular forms. The system uses various data mining techniques [1].

The system very well reduces its first drawback as the paper work is replaced by database and it can retrieve the values any time.

B. Proposed system

The proposed system is developed in such a way that it is very easy to enter the data. The data collection is done using simple questions and wherever necessary the options are provided. The collected data is stored in the database. The data collection is done using Java application thus the patient can fill in the details according to his/her convenience. After the data is entered the Data mining technique is used to find out the probability whether the patient is prone to heart disease. If yes then the patient will have to answer various questions and enter his symptoms. The symptoms are checked in the medicine database by various pattern matching algorithm [9] and the corresponding remedy is generated. The proposed system consists of 3 main modules.

978-1-4799-8187-8/15/$31.00

The following will be the main modules of the proposed system:

A. Registration and Data Collection

- The patient needs to have his/her account. For that the patient needs to enter his personal details and create an account.
- After registration the patient will be provided with a unique username.
- The patient will have to fill the forms containing questionnaires.
- After the completion of forms, the proposed system database will be updated with the patient records.The doctor can also view the patient's health records.

B. Finding the Probability

- After the collection of the data entered by the patient, the proposed system will generate the probability whether the patient is in risk of a heart disease or not[2].

C. Generating the Remedy

- Depending upon the probability of the heart disease of the patient, the further questions will appear.
- The symptoms will be entered by the patients and depending upon the probability and the symptoms, the remedy will be generated and will be displayed [9].

III. ARCHITECTURE OF THE PROPOSED SYSTEM

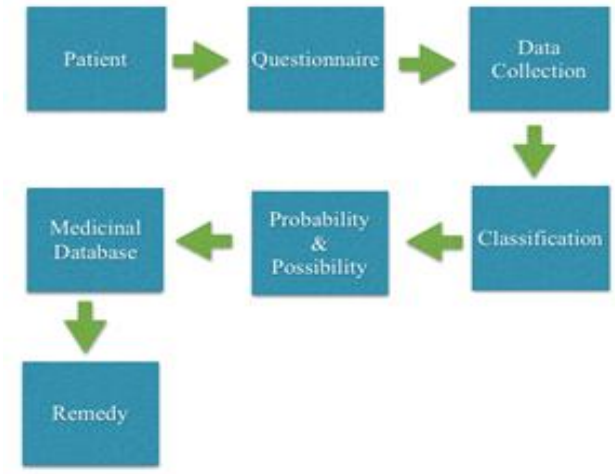

Figure 1.1 - Architecture of the proposed system

The proposed system is done as Java Application. The patients need to answer various questions and the data is collected and then stored in the database. The data is classified (Figure-1.1) based on the different possibilities. The patient can check whetherthe patient is prone to heart disease or not. The data entered by the patient is considered and the probability is generated whether the patient is in risk of getting the heart attack or not. Remedy is generated depending on the probability and it is check in medicinal database. The doctors can view the patient details as well as the medicine database.

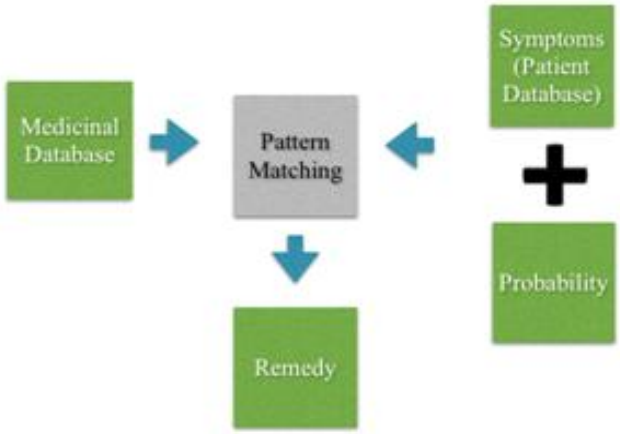

Figure 1.2 - Remedy generation

The symptoms are taken from the patient database and segregated into categories. These symptoms categories are matched with the symptom file created from the medicine database. After pattern matching the necessary remedies with their probability is displayed in Figure – 1.2.

The proposed system is divided into three parts

- Data collection
- Probability generation
- Remedy Finder

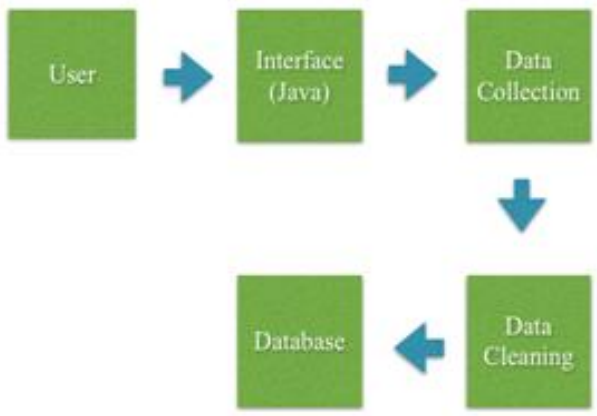

Figure 1.3 - Data Collection

Data collection in Figure – 1.3is meant to be used by the patients and the probability generation to be used by doctors or patients. In Data collection the main functionality of the system is carried out. The system would direct the patients to answer a few questions in order to get necessary diagnosis.

The information collected from the patient includes personal details, associated complaints and reactions experienced. The symptom file is generated from the medicine database. The symptoms from the segregated sections are pattern matched with symptoms present in the symptom file. The matched symptoms are then used to generate remedies from the medicine database and the probability of the remedies. The symptoms are segregated depending upon the attributes(Figure – 1.4).

Figure 1.4 - Segregation of Symptoms

IV. CLASSIFICATON ELEMENTS

Questionnaires have advantages over some other types of medical symptoms that they are cheap, do not require as much effort from the questioner as verbal or telephone surveys, and often have standardized answers that make it simple to compile data. However, such standardized answers may frustrate users. Questionnaires are also sharply limited by the fact that respondents must be able to read the questions and respond to them. Here our questionnaire is based on the attribute given in the data set, so the questionnaire contains various input attributes: sex, chest pain type, fasting blood sugar, ECG, exang, slope, blood pressure, cholestrol, thalach, depression

V. DESIGN OF THE PROPOSED SYSTEM

Finding the Probability using Naïve Bayes

- Bayes' theorem relates the conditional and marginal probabilities of two random events. It is often used to compute posteriorprobabilities given observations.
- For example, a patient may be observed to have certain symptoms. Bayes' theorem can be used to compute the probability of a particular disease the patient is suffering from.

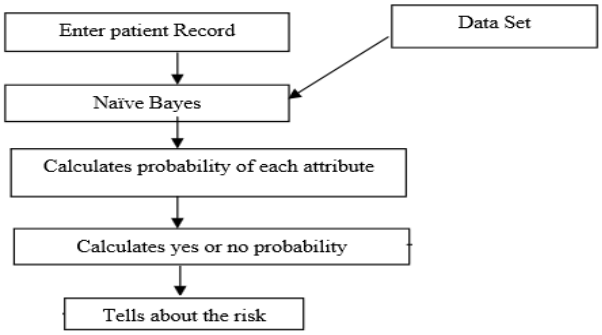

Figure 1.5 - Finding the probability

A. Bayes Rule

- A conditional probability is the likelihood of some conclusion, A, given some evidence/observation, *B*, where a dependence relationship exists between *A*and *B*.
- This probability is denoted as P*(A|B)* where A=High B=Low
- P(High | Low) = $\frac{P(Low \mid High)\, P(High)}{P(Low)}$ ...(I)

B. For proposition High and evidence Low

- P(High), the prior, is the initial degree of belief in A.
- P(High| Low), the posterior, is the degree of belief having accounted for B.
- P(Low | High) / P(Low) represents the support B provides for A.

C. Remedy Finder

- After the probability is found, the patient will come to know about the whether he is prone to heart disease or not.
- If yes, then various set of question will appear where the patient will have to enter the symptoms. After the symptoms are entered, the remedy will be generated using Pattern Matching algorithms.

D. Boyer Moore

- The algorithm scans the characters of the pattern from right to left beginning with the rightmost one.
- In case of a mismatch (or a complete match of the whole pattern) it uses two pre-computed functions to shift the window to the right.

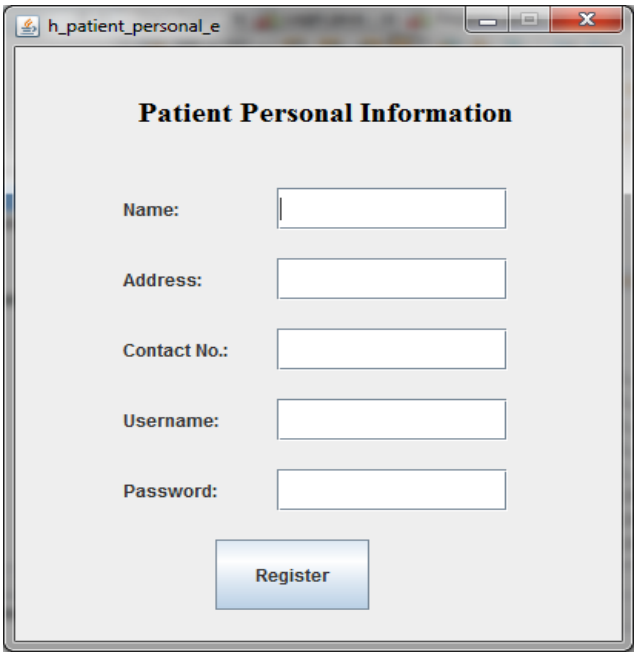

Figure 1.6 - Patient Registration Form

If the patient has not registered then the patient will have to register using the above form (Figure 1.6).

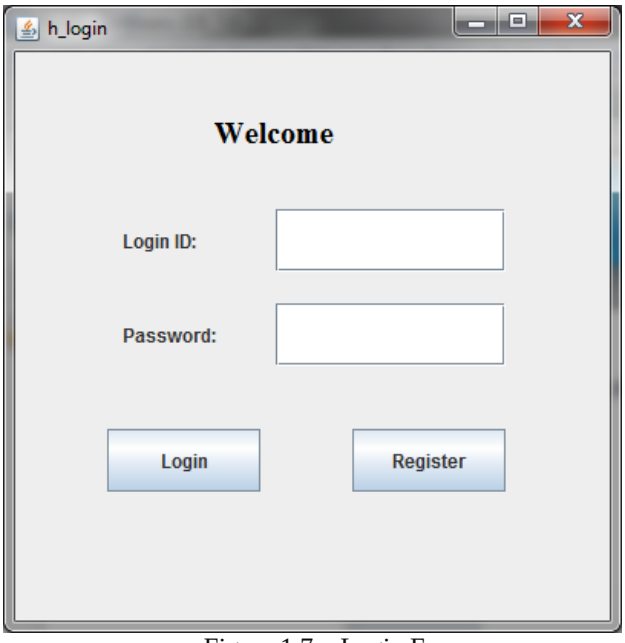

Figure 1.7 – Login Form

The above figure – 1.7 represents the Login form. After the patient has registered he/she can login using his/her Login ID and password.

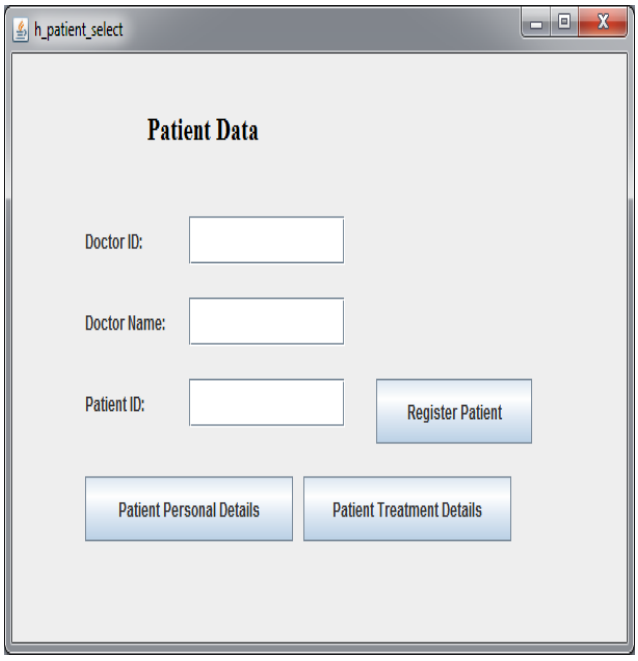

Figure 1.8 – Doctor Form

The above figure 1.8 represents the Doctor login. After Doctor has logged in, then the Doctor can view the details of the patients and assign the patient to the particular Doctor.

Figure – 1.9 Patient's account

The above figure – 1.9 represents the account details of the patient. Once the patient has logged into his account he has the options to select from the above shown two options i.e Details and Checkup. If the patient selects Details then he can see his previous records. If he wants to do a medical checkup regarding the probability of heart disease then he can check with the help of checkup option.


1#2 Acute:# 76#3:Acon, 1:Agar, 1:Aloe, 1:Alum, 1:Am-c, 2:Anac, 1:An
1:Carb-v, 1:Cham, 3:Chin, 1:Chin-ar, 2:Cic, 1:Cimic, 2:Cocc, 3:Coff, 2:C
2:Nat-ar, 3:Nat-c, 2:Nat-m, 2:Nat-p, 2:Nux-m, 3:Nux-v, 3:Op, 1:Petr, 2:Ph-
2:Sulph, 3:Tab, 2:Ther, 1:Thuj, 2:Verat, 1:Viol-o, 1:Zing,
2# 3 Morning:# 1#2:Fl-ac,
3# 3 Evening:# 2#1:Coca, 1:Rhod,
4# 4 Bed,in:# 1#1:Kali-c,
5# 4 On falling asleep:# 2#1:Calad, 1:Calc,
6# 3 Night:# 1#1:Atro,
7# 3 Bed,in:# 1#1:Kali-c,
8# 3 Chill,during:# 2#1:Arn, 3:Caps,
9# 3 Cracking in ears,preceded by:# 1#1:Graph,
10# 3 Heat,during:# 8#1:Acon, 1:Bell, 1:Calc, 3:Caps, 3:Con, 1:Ip, 1:L
11# 3 Labor pains,during:# 1#1:Cimic,
12# 3 Menses,during:# 3#2:Hyper, 1:Mag-c, 1:Nux-v,
13# 3 Music:# 0#
14# 4 To:# 15#3:Acon, 1:Aloe, 1:Ambr, 1:Bufo, 2:Cact, 2:Cham, 2:Cof
15# 4 During menses:# 1#2:Nat-c,
16# 4 Organ:# 1#1:Lyc,

Figure 1.10 – Medicine Database Details

The above Figure –1.10represent the database details regarding various symptoms related to heart and their corresponding remedies.

VI. CONCLUSION

Decision Support in Heart Disease Prediction System is a Java Application which is intended to be done using Data mining technique "Naive Bayesian Classification technique". The system extracts hidden knowledge from the database. HDPS can be further enhanced and expanded. It can also incorporate other data mining techniques.

After finding out the probability and depending upon the probability, the corresponding remedies will be generated using Pattern Matching Algorithm.

VII. SCOPE OF THE PROPOSED SYSTEM

[1] Mobile Application

We can develop mobile application on Android and IOS platforms for making the system easily available.

[2] SMS Facility

We can provide a SMS facility that will send a message to the patient regarding his heart condition and how to take effective measures to reduce the danger.

[3] Pacemaker

We can integrate the system with pacemaker. It is a small device that is placed in the chest or abdomen to help control abnormal heart rhythms. With the help of electrical pulses it prompts the heart to beat at a normal rate.

VIII. REFERENCES

[1] Mrs. G. Subbalakshmi, "Decision Support in Heart Disease Prediction System using Naive Bayes", Indian Journal of Computer Science and Engineering.

[2] Fayyad, U: "Data Mining and Knowledge Discovery in Databases: Implications for scientific databases", Proc. of the 9th Int. Conf. on Scientific and Statistical Database Management, Olympia, Washington, USA, 2-11, 1997.

[3] Giudici,P.:"Applied Data Mining: Statistical Methods for Business and Industry", New York: John Wiley, 2003.

[4] Han, J.,Kamber, M.: "Data Mining Concepts and Techniques", Morgan Kaufmann Publishers, 2006.

[5] Intelligent Heart Disease Prediction System Using Data Mining Techniques-SellappanPalaniappan, RafiahAwang 978-1-4244-1968-5/08/ ©2008 IEEE

[6] Obenshain, M.K : "Application of Data Mining Techniques to Healthcare Data", Infection Control and Hospital Epidemiology, 25(8), 690–695, 2004.

[7] Anthony M. Dymond, "Data Extraction Methods in Medium and Large Databases", Dymond and Associates, LLC, Concord, CA.

[8] RajuBhukya, DVLN Somayajulu, "An Index based Forward Backward Multiple Pattern Matching Algorithm", World Academy of Science, Engineering and Technology International Journal of Computer, Information, Systems and Control Engineering Vol:4 No:6, 2010.

[9] 'http://www-igm.univ-mlv.fr/~lecroq/string/node14.html', Boyer Moore, accessed on 25th April,2014

[10] Wu, R., Peters, W., Morgan, M.W.: "The Next Generation Clinical Decision Support: Linking Evidence to Best Practice", Journal Healthcare Information Management. 16(4), 50-55, 2002.