

AN EFFICIENT PREDICTION OF MEDICAL DISEASES USING PATTERN MINING IN DATA EXPLORATION

V.Geetha,

Assistant Professor,
Department of Computer Science Engineering,
SCSVMV University,
India.

T.Jayanthi,

Assistant Professor,
Department of Computer Science Engineering,
SCSVMV University,
India.

S. Indhumathi,

UG Scholar,
Department of Computer Science Engineering,
SCSVMV University,
India.

Dr.C.K.Gomathy,

Assistant Professor,
Department of Computer Science Engineering,
SCSVMV University,
India.

R. Jayashree,

UG Scholar,
Department of Computer Science Engineering,
SCSVMV University,
India.

E. Avinash,

UG Scholar,
Department of Computer Science Engineering,
SCSVMV University,
India.

Abstract: Interactive Data Exploration is a key ingredient of a diverse set of discovery-oriented application. Data discovery is a highly ad hoc interactive process where it is difficult to retrieve relevant data from large, unorganized pools. In this project, the diseases can be predicted based on the analysis from their symptoms and the report is generated from the systematic analysis of particular disease. Early detection and prevention of diseases plays a very important role in reducing the mortality rate caused by those diseases. It is a multilayered method which uses significant pattern mining using iterative search techniques to build a risk prediction system which predicts various disease. It is user friendly, time and cost saving. This research uses data mining technology such as classification and prediction to identify potential treatments for patients according to their diseases. The gathered data is preprocessed, fed into the database and classified to yield significant patterns using pattern mining algorithm. Finally a prediction system is developed to analyze risk levels which help in prognosis. This helps in detection of a person's predisposition for any common diseases before going for clinical and lab tests which is cost and time consuming.

Keywords: Data exploration, Data sampling, Mortality rate, Symptoms

I.INTRODUCTION

Traditional data management systems assume that when users pose a query they have good knowledge of the schema, meaning and contents of the database and they are certain that this particular query is the one they wanted to pose. In short, traditional DBMSs are designed for applications in which the users know what they are looking for. However, as data are being collected and stored at an unprecedented rate, we are building more dynamic data-driven. Interactive data exploration (IDE) is one such example.

The healthcare environment is generally perceived as being 'information rich' yet 'knowledge poor'. There is a wealth of data available within the healthcare systems. However, there is a lack of effective analysis tools to discover hidden relationships and trends in data. Knowledge discovery and data mining have found numerous applications in business and scientific domain. Valuable knowledge can be discovered from application of data mining techniques in healthcare system.

Health examination has played an important role for maintaining people's health since it can not only help people understand their own health conditions clearly but also avoid missing the best timing of disease treatment. However, in current health examination systems, people get only a basic report from single health examination and no advanced health risk analysis is provided. In this paper, we proposed an effective mechanism for chronic disease risk prediction by mining the data containing historical health records and personal life style information. Value change trends of the data are important for disease status prediction, and we defined significant ones as health risk patterns in our mechanism.

The main objective of this study is to create a fast, easy and an efficient algorithm for disease prediction, with less error rate and can apply with even large data sets and show reasonable patterns with dependent variables. For disease identification and prediction in data mining a new hybrid algorithm was constructed. The Disease Identification and Prediction(DIP) algorithm, which is the combination of decision tree and association rule is used to predict the chances of some of the disease

Hypertension is a major disease, being one of the top ten causes of death in Taiwan. The exploration of three-dimensional (3-D) anthropometry scanning data along with other existing subject medical profiles using data mining techniques becomes an important research issue for medical decision support. This research attempts to construct a prediction model for hypertension using anthropometric body surface scanning data.

II.RELATED SYSTEM

In the existing system, data exploration was limited only to data mechanisms which are difficult to learn. In existing approach, it doesn't provide complete coverage of operations needed by the users. In previous work, it was manual based data exploration. An automated system is crucial for deriving insights from complex datasets found in many big data applications such as scientific and health care applications.

III.METHODOLOGIES

The main objective of this paper is to predict diseases from the given set of symptoms and the generation of report from the systematic analysis of particular disease that is found using iterative search through the symptoms. The collected data is pre-processed and stored in the knowledge base to build the model. The pattern mining is done using iterative search which is build using the classification rules, the significant frequent pattern and its corresponding weightage. Finally the report is generated.

MODULES:

- User Page
- Doctor Page
- Admin Page

MODULE DESCRIPTION:

User Module:

After careful analysis the system has been identified to have the following modules:

- Registration
- Login
- Disease Prediction
- Report Generation
- Appointment Fixation

Registration:

The Registration is the process of making the user to register on our sites. If the user gets registered in our sites then the prominent user will allow to proceed the other process. Once they choose the Symptoms from the list of symptoms then they will be followed with that plan. Only registered users can access the site. The user will get the notification when they logout the session.

Login

The users who register their details to the site and accessing it are called login. They have the following permissions. They are

- Can book the appointment.
- Send report to the Suggested Doctors
- Can view the comments from doctor.

Disease Prediction

In this module, the user can check their diseases based on the symptoms which is listed in the symptom checker tab. Based on the symptoms the disease is predicted.

This follows a certain pattern where the disease was accurately predicted.

Report Generation

The report is generated based on the results from the previous page. This generated report is sent to the doctor.

BMI calculator

From the weight and height mentioned by the user, the body mass index is calculated.

The Disease is predicted based on the symptoms .Though these symptoms could be caused by several conditions and these were not appropriate.

Suggested Doctors

For each disease, doctors were suggested .These doctors are the specialist in that particular field. After clicking view button, the user can view the doctor details.

Appointment Fixation

After viewing the doctor details, the user can fix the appointment with that particular doctor. By filling the appointment form, the user can fix the appointment and another option is sending the report to particular doctor.

Doctor's module

After careful analysis the system has been identified to have the following modules:

- Login
- View My Details
- View Appointment
- Add a Comment

Login

The doctors where automatically registered to this site by the administrator. The doctor can log into his/her account by logging into the page. They can access only certain users.

View My Details

In this tab, the doctors can view their own details like hospital ID ,timings and their Contact numbers

View Appointment

In this Section, the doctors can view their appointments fixed by the user. In the appointment table it includes the user name who fixed the appointment with the date and the reason for the fixation. The token generated from the user module is displayed along with the timings of the doctor

Add a comment

This contains the report of the various users who requested the doctors for suggestion. When the user sends the report to particular doctor it will be displayed in this report page where the doctors can view the specific user and can comment regarding the report

Admin Module

After careful analysis the system has been identified to have the following modules:

- Login
- View User Details
- View Doctor Details
- View Feedback

Login

The Administrator has their own login. This login is only for admin page.

View User Details

The Administrator can view the whole user details in a table.

View Doctor Details

The Administrator can view the whole doctor details in a table.

View Feedback:

The feedbacks entered by the user can be viewed by the administrator.

IV.SYSTEM ARCHITECTURE:

Software design sits at the technical kernel of the software engineering process and is applied regardless of the development paradigm and area of application. Design is the first step in the development phase for any engineered product or system. The designer's goal is to produce a model or representation of an entity that will later be built. Beginning, once system requirement have been specified and analyzed, system design is the first of the three technical activities -design, code and test that is required to build and verify software. The importance can be stated with a single word "Quality". Design is the place where quality is fostered in software development. Design provides us with representations of software that can assess for quality. Design is the only way that we can accurately translate a customer's view into a finished software product or system. Software design serves as a foundation for all the software engineering steps that follow. Without a strong design we risk building an unstable system – one that will be difficult to test, one whose quality cannot be assessed until the last stage. During design, progressive refinement of data structure, program structure, and procedural details are developed reviewed and documented. System design can be viewed from either technical or project management perspective. From the technical point of view, design is comprised of four activities – architectural design, data structure design, interface design and procedural design.

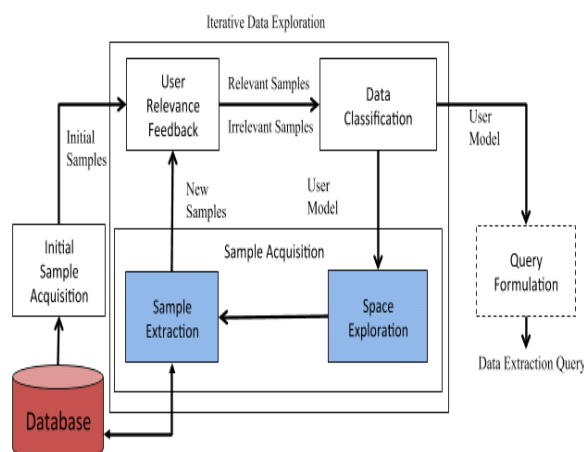


Figure1: Automated Interactive data exploration framework

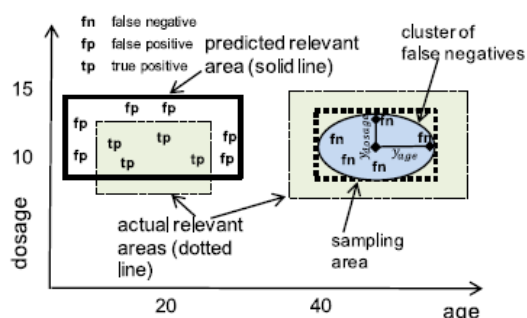


Figure2: Misclassified objects and cluster based sampling

V.ALGORITHM DESCRIPTION:

Pattern mining:

Sequential pattern mining is a topic of data mining concerned with finding statistically relevant patterns between data examples where the values are delivered in a sequence.

These include building efficient databases and indexes for sequence information, extracting the frequently occurring patterns, comparing sequences for similarity, and recovering missing sequence members.

In general, sequence mining problems can be classified as string mining which is typically based on string processing algorithms and item set mining which is typically based on association rule learning.

$$Sw(i) = \sum (W_i * F_i)$$

VI.RESULTS

More interactive with user. Earlier Prediction of diseases reduces the mortality rate. This prediction system may provide easy and a cost effective way for screening various diseases. It plays a pivotal role in earlier diagnosis process for different types of heart and provides effective preventive strategy. Easy to find the relevant data's.

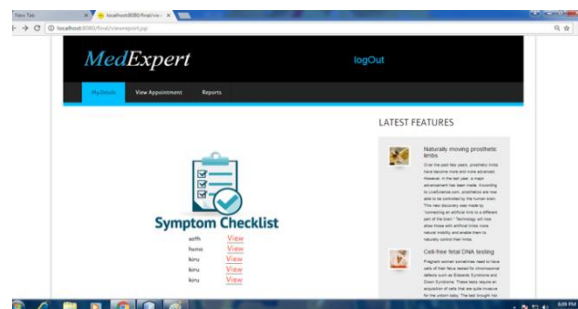
HOME PAGE



REGISTRATION PAGE

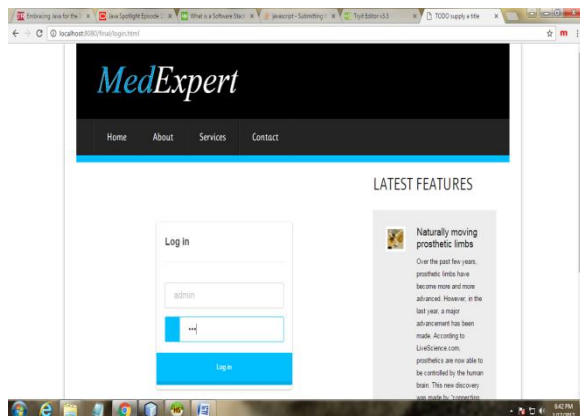


DOCTOR REPORT PAGE



ADMIN LOGIN PAGE

information technology in bio
medicine, vol.11, no.3, may 2007



VII.CONCLUSION:

Interactive Data Exploration is a key ingredient of a diverse set of discovery-oriented application. In these applications, data discovery is a highly ad hoc interactive process where users execute numerous exploration queries using varying predicates aiming to balance the trade-off between collecting all relevant information and reducing the size of returned data. Therefore, there is a strong need to support these human-in-the-loop applications by assisting their navigation in the data space.

In this project, our aim is to detect or predict the diseases based on the symptoms obtained from the user and the report is generated from that predicted analysis. This helps in detection of a person's predisposition for any common diseases before going for clinical and lab tests which is cost and time consuming.

More interactive with user. Earlier Prediction of diseases reduces the mortality rate. This prediction system may provide easy and a cost effective way for screening various diseases. It plays a pivotal role in earlier diagnosis process for different types of heart and provides effective preventive strategy. Easy to find the relevant data's.

VIII.REFERENCES

- [1]. Kyriaki.et.al "AIDE: An Active Learning-Based Approach for Interactive Data Exploration", IEEE transactions on knowledge and data engineering, vol.28, no.11, November 2016.
- [2]. K.Srinivasan,et.al, Associate Professor, Dept. of CSE Principal and Professor of CSE "Applications of Data Mining Techniques in Healthcare and Prediction of Heart Attacks".(IJCSE) International Journal on CSE Vol. 02, No. 02, 2010, 250-255
- [3]. Gomathy, C.K. and S. Rajalakshmi, 2014. Software pattern quality component in service-oriented architectures. Eur. Sci. J., 10: 412-423.
- [4]. Smitha.T et.al, "An Efficient Algorithm for Disease Prediction with Multidimensional Data", International Journal of Computer Applications (0975 – 8887) Volume 63– No.9, February 2013
- [5]. Chiu et.al., "Mining Three-Dimensional Anthropometric Body Surface Scanning Data for Hypertension Detection", IEEE transactions on