A Framework designed for the Medicines classification and its data extraction in data mining

Aishwarya Datalkar

[1]Student, Information Technology, Atharva College of Engineering, Mumbai, India.
Email: aishwaryadatalkar.1996@gmail.com

ABSTRACT

The purpose of the paper is to provide the information about the medicines not being used currently through data mining. It takes the data regarding various medicines into consideration so as to demonstrate the processing of the suggestions to the user based on the searches of the previous users. The system forms the various schemas to fetch the data at ease. While the previously Organizations used the single database to retrieve and fetch the data, making the process slower. Thus this paper put forth the system of the fetching the data at the faster rate and so as to suggest the medicines based upon previous searches to the users using rule based and apriori algorithm.

Keywords : Data Mining, Database System, Medicines and healthcare data.

1 INTRODUCTION

The purpose of data mining, whether it's being used in healthcare or business, is to identify useful and understandable patterns by analyzing large sets of data. These data patterns help predict industry or information trends, and then determine what to do about them. In healthcare, data mining has proven effective in areas such as predictive medicine, customer relationship management, detection of fraud and abuse, management of healthcare and measuring the effectiveness of certain treatments. Electronic health records (EHR) are common among healthcare facilities. With increased access to a large amount of patient data, healthcare providers are now focused on optimizing the efficiency and quality of their organizations use of data mining. Businesses have used data mining for things like credit scoring and fraud detection. Today, healthcare organizations are seeking similar benefits from data mining and predictive analytics. However, inspite of these there instances where people are still unaware about the medicines that are not used anymore for the various treatments. Inorder to make these data easily avaliable for the users and spread the awareness about the same the system is developed.

The system includes the data for various medicines and then retrived as and when searched for by the user. The data will be bifurcated into various forming categories and so as storing the data for different medicines in each category. On a search of a particular medicine the data from the database will be provided to the user. However, it will provide the data about the disease for which the medicine is being used or currently it is banned due to some side effects. This will help people get the updates about the various medicines and avoid future risks.

2 PROPOSED SYSTEM

The Proposed System shows storage of data efficiently. Ideally a data could be stored in a single database. And the corresponding tables of the Database could be normalized. But the proposed System stores the data at Schema level so that the stored data is fetched efficiently. The following are the features that will be provided by the system:

- . The medicines would be segregated according to the categories and located at individual schemas which nullifies the interdependency.
- Medicines Data could be fetched at a comparatively faster speed which indicates improvised performance of the system.
- New diseases categories can be added to the system in runtime without any disturbances because of decreased dependencies.
- The searches of the user will be recorded for further analysis of the data.
- Based on the searches of the medicines ranks are allotted to the medicines which is used while suggesting the most relevant medicine to the user.
- Related Medicines are suggested to the User based on the mostly searches.

As shown in fig1, we can form the product data with the help of various categories. These categories can help form the various sche-

mas for the different products, making the process of fetching the data faster.

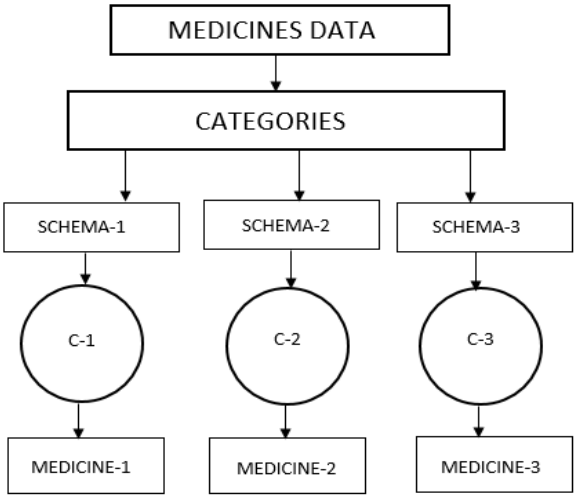

Fig 1: Proposed System

3 METHODOLOGY

3.1 Admin Module

This module takes into consideration the users data and how the data should be managed into the various schemas so as to fetch the data in less amount of time. It looks after the system is working as per the requirement or not. Admin is responsible for the proper working of the website and so as database system.

3.2 Client Module

Client Modules manages client data. It has mainly Client Details. Client details sub module is for storing clients personal details. The details of clients are displayed in the client details page, with an edit button to edit their details as well as search button to find client.

3.3 Information Module

Information Modules manages the medicines information with sub modules such as their current usage. The users can search for the medicine by entering the name of the medicine and hence its details will be displayed. As the medicines are placed according to the categories they can retrieve at faster rate.

3.4 User Module

User modules takes into consideration the login, wherein the users will ensure the security of their actions over the system is maintained properly. There the authentication and authorization of the users is maintained and users data is not miss used.

4 DESIGN DETAILS

Data mining involves exploring and analyzing large blocks of information to glean meaningful patterns and trends. It can be used in a variety of ways, such as database marketing, fraud detection, spam Email filtering, or even to discern the sentiment or opinion of users.

The data mining process breaks down into six steps: First, organizations collect data and load it into their database systems. Next, they store and manage the data, either on in-house servers or the cloud and select data. Business analysts, management teams and information technology professionals access the data and determine how they want to organize it. Then, application software sorts the data based on the user's results, and finally, the end-user presents the data in an easy-to-share format.

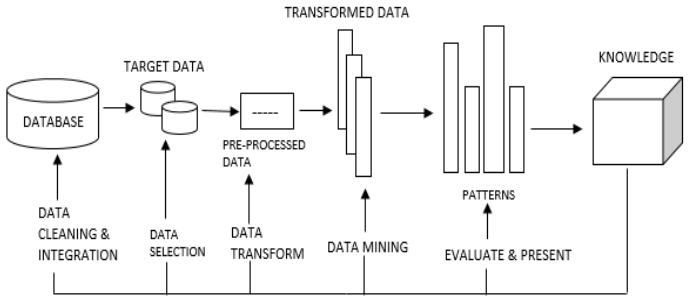

Fig 2: Functionalities of data mining

The system is designed as below shown in the fig 3, where the web application contacts the backend where the mining is done on the database system and this backend is done in the java programing language. The web application provides the different users data regarding the transactions and thus the data is stored in the databases and hence timely scheduled so to keep updating the database. The rule based algorithm is used in order to provide the suggestions to the users based on previous searches. The data is stored in the form of the different schemas so as to fetch the data easily and avoid long processing time.

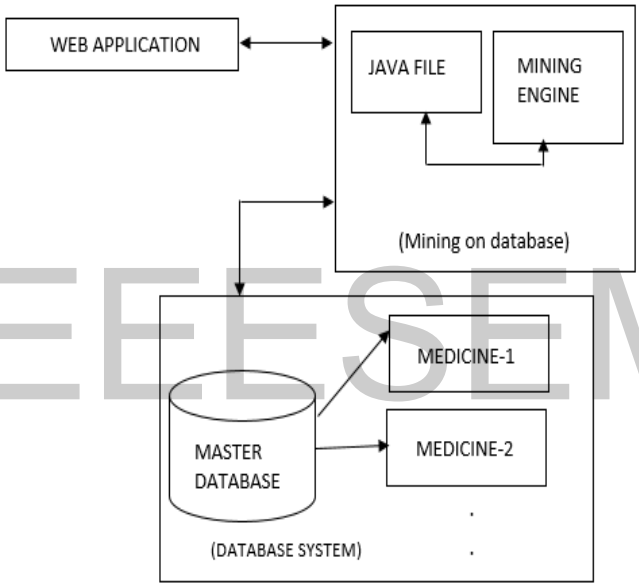

Fig 3: Designed System

5 ALGORITHM

5.1 Rule-Based Algorithm:

A rule-based prediction alogorithm program consists of a collection of rules for each task. Each rule has the structure of an if-then statement.

Then if part checks for some condition on the features of an item.

When the condition is true, the then part provides something to be used in the final output. In the programs we consider, the rules will be collected in an ordered list format, so that rules earlier in the list take precedence over rules later in the list if they produce conflicting outputs.

We can express a rule in the following from − IF condition THEN conclusion

Consider a rule R1,

R1: IF age = youth AND student = yes THEN buy_computer = yes

5.2 Apriori Algorithm:

Apriori is an algorithm for frequent item set mining and association rule over transactional databases.

The algorithm proceeds by identifying the frequent individual items in the database and extending them to larger and larger item sets as long as those item sets appear sufficiently often in the database.

The frequent item sets determined by the algorithm can be used to determine association rules which highlight general trends in the database. This has applications in domains such as analysis.

It uses a bottom up approach, where frequent subsets are extended one item at a time. The algorithm terminates when no further successful extensions are found.

In this system the searches are taken into consideration and then the medicines are allotted particular id which is considered and compared with medicines that occur frequent in searches and hence suggestions are provided to the user.

6 CONCLUSION

A Healthcare application has been designed with modules for Medicines, Categories, Account. The proposed model can be used by any medical organization wherein their own application that works efficiently in all browsers and devices. Healthcare application plays a key role in every organization and can be used to manage its entities and modules easily.

The application will make Independent Schemas for every disease category which will make the creation and deletion of entities more convenient.

The developed system is cost effective and can be scaled at economical costs. If we can develop an application which can customize itself based on user needs, then that healthcare application can be easily consumed and installed by the organization itself. In future we can see integrating virtual reality and the Internet of Things in their frameworks.

REFERENCES

[1] HianChyeKoh and Gerald Tan, ―Data Mining Applications in Healthcare‖, journal of Healthcare Information Management – Vol 19, No 2.

[2] JayanthiRanjan, ―Applications of data mining techniques in pharmaceutical industry‖, Journal of Theoretical and Applied Technology, (2007).

[3] RubanD.Canlas Jr., MSIT., MBA , ― Data mining in Healthcare: Current applications and issues‖.

[4] K. Srinivas , B. Kavitha Rani and Dr. A. Govrdhan, ―Applications of Data Mining Techniques in Healthcare and Prediction of Heart Attacks‖ International Journal on Computer Science and Engineering (2010).

[5] ShwetaKharya, ―Using Data Mining Techniques ForDiagnosis And Prognosis Of Cancer Disease‖, International Journal of Computer Science, Engineering and Information Technology (IJCSEIT), Vol.2, No.2, April 2012.

[6] EliasLemuye, ―Hiv Status Predictive Modeling Using Data Mining Technology.