# Applications of Data Mining in Healthcare

**Chris Sosa Jacob, Swethamol V S, Jisha J Nair**

Department of Computer Applications

Saintgits College of Applied Sciences, Kottayam, India

-------------------------------------------------------------------***-------------------------------------------------------------------

**Abstract -** Nowadays many organizations are using data mining technologies in various fields. We all know that healthcare is very important for everyone. Many technologies are evolving for the examination of physical conditions and details about many diseases. There is a large quantity of data about the physical conditions of a person and about many diseases. There are some diseases that can be prevented before getting affected. If the person is aware about the symptoms then he will be able to take care and thereby he can overcome it. As we know there is a large set of data related to the medical field an effective method is required to collect data from several databases. Here we need modern data mining technologies. This paper is a study about the various data mining applications in the medical care. These applications in data mining are being relevant and specific for the medical needs as they are successfully used in the healthcare sector for its high accuracy in prediction and results.

*Key Words:* data mining, medical care, technologies

# 1.INTRODUCTION

Data mining means the obtaining of hidden data relationships, patterns and other knowledge from a large data which is difficult to choose by using the traditional methods. From any organization data warehouse we can extract data by using the data mining technology to find the most modern patterns in the medical industry. There are many data mining tools. We can say that data mining is a quantitative approach and is user-friendly also. It is easy to read and correct the errors. Data mining has been extensively used in many healthcare organizations because it is being essential nowadays. In our India, we have many healthcare organizations in the public sector and private sector. Public sector healthcare organizations are undertaken by the government which gives a low-cost medical treatment for the poor people. From both these sectors we have huge data associated with patients, medicines, diseases etc. to manage this largedata we need new technologies and new techniques. Here we give a review about applications in healthcare.

# 2.DATA MINING

We can say that data mining tools will help us to predict what will happen in future with the existing data or by processing the data. Data mining has many frameworks such as path analysis classification clustering and forecasting. Path analysis is asearching of patterns with the help of data related to one another. Classification parameters search for new patterns. Clustering parameter groups the facts that are unknown. Forecasting parameter is known as predictive analysis.

## 2.1DATA MINING TASKS

There are different tasks for the execution of a model, for example clustering, association rule, correlation analysis. These are some of the methods used for descriptive models. Although classification, regression, categorization are utilized for predictive models.

- Classification
- Regression
- Categorization
- Anomaly Detection

## 2.2 DATA MINING METHODS

The anomaly recognitions are namely, standard support vector data description, densitybasedsupport, vector data description, gaussianmixture.Vector standardization way aremostly important in clustering ,more over in classification the widely used methods are statistical, classifiers, support vector.

# 3.DATA MINING APPLICATION

Nowadays mining applications are broadly integrated in many fields. It indicates prediction from the large data. Few possibilities of data mining used in different areas are:

- Business
- Healthcare
- Telecommunication
- Marketing and sales

## 3.1 DATA EXPLORATION TECHNIQUES IN HEALTHCARE

As we know that the medical industry produces a huge set of data about the patients, doctors, hospitals, diseases, medicines etc. This technique is mostly used for predicting diseases and aid healthcare professionals in necessary decisions.It serves patient's life by predicting about the medicines. Data mining helps in managing healthcare at different levels. Categorization of information probing strategy are as follows;
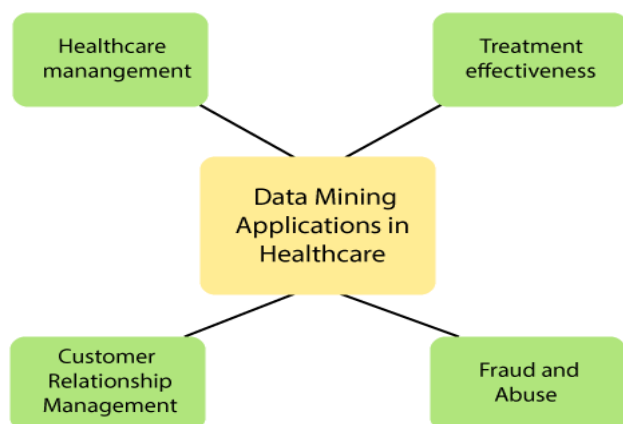
**Fig -1**: Figure

### 3.1.1 TREATMENT EFFECTIVENESS

Data mining operation can be used to evaluate effective treatment for various diseases.These are done by evaluating the indications of the diseases at the early stage. Thesevariables differentiate various treatment methods and from that study the adequatetreatment is identified. Other applications include identifying sideeffects,symptoms,effective drug compounds etc.

### 3.1.2 HEALTHCARE MANAGEMENT

Healthcare management means the managing or administrating hospital networks and alsoa wide area consisting of healthcare informatics, medical device industry, pharmaceuticalindustry, system biology and many more.

➢ **Medical Device Industry**

Medical devices are very important in medical area for best communication.

➢ **Pharmaceutical Industry**

Determining technology is adapted in the pharmaceutical industry for developingadvancedmedicines.For organizational decision making utilize large fact in the pharmaceutical sector.

### 3.1.3 CUSTOMER RELATIONSHIP MANAGEMENT

CRM permits organizing the communication between commercial organizations. It has less importance in medical industry. Customer interactions are mainly done through call centers, building sections, inpatient settings and ambulatory care settings. Mainly types of CRM, first one is for healthcare organizations to connect with their patients and second one is for a healthcare organization to connect with the organizations. The reasons for these two types are because healthcare organizations develop new patients in two ways. The first way relays on the patient's choice. People may seek help for a doctor when they are ill or they may solve their problems themselves. Healthcare organizations use CRM to administer relationship with customers. Data Mining of patient surveys are made for getting an idea about the patients view on healthcare. Pharmaceutical companies use CRM and healthcare information for tracking doctors, which drugs they prescribe to the patient and they can decide whom to target.The healthcare CRM system coordinates data from laboratory, billing sector, drug supply and other sources. A generic CRM system looks for manual entry of tasks and opportunities on the part of healthcare workers.

### 3.1.4 FRAUDULENCE IN HEALTHCARE

Irrelevant payments by insurance organizations may cause this. The usage of traditional method will not help to identify this medical abuse. Health care fraud means double dealing for obtaining illegal benefits. Cheating behaviour is considered a crime. It is examined that 10 percentage of medical expenses are wasted because of fraud activities. Data collection can highlight wrong medical data have been submitted for medical insurance etc. A rule has been issued by the health department and human services on May 17,2013 the rule is that" Static Medicaid Fraud Control Units; Data mining "it states that MFCU has to submit the data collection methods to the office of inspector general for the approval. The major healthcare fraud was the prescription fraud a claim for patients who do not exist in the records and another one is upcoding the claim for medical treatments which are not performed by the hospital.

## 4.DATAMINING ALGORITHMS IN HEALTHCARE

Healthcare comprises the procedures of diagnosis such as curing and prevention of physical and mental problems in human beings. Further data collection in medical industry is adapted to predict diseases to help doctors in various purposes. In majority of the nations, it moves in a steady manner, according to the region the healthcare industries will be rich in data while they built abundance of records including electronic medical records, benchmark findings etc. might be utilized. In medical field industry the discussions on different methods are:

### 4.1 ANOMALY DETECTION

It detects the most changes in data. Bo Lie et al used tree types of anomaly detection methods, this is analyzed using AUC accuracy.

### 4.2CLUSTERING

The descriptive task where a person can recognize set of clusters to describe the data. The forecasting of readmission in a huge medicine used vector quantization method, this use algorithm like k-means, k-methods and x-means. The facts for making this study were collected from patientmedical process and medical results.

### 4.3. CLASSIFICATION

Classification illustrates a model that describes data class and deploy it to understand the classes of object where we know the label. The classification works are:

### 4.3.1 STATISTICAL

MTS algorithm is mainly used here. It makes a statistical prediction from one form in another also MS reflex the abnormality degree of monitoring from the known recommendation group. The class variance issues are common in health data.

### 4.3.2 DISCRIMINENT ANALYSIS

Linear Discriminant Analysis(LDA) is mostly utilized to forecast the class depend on the calculation on latest nameless examination. The analysis depends on probability density functions of the futurist tracks a normal distribution based upon the class value.

### 4.3.3 DECISION TREE

Predictions are made by testing the data features and constructing the tree and its protocol. The type of data set used in this investigation is quite balanced dataset. Decision tree can't be taken in decision process to resolve unbalanced problems as the decision tree periodically split observation to branches to make a tree.

## 5. DATAMINING LIMITATION

A major limitation is the obtainability of data. Owing to existence of data in various sectors such as laboratory,clinic, administration and more. Before starting data mining we must collect all these data and combine them.A solution is the deployment of a data warehouse. Since highly priced and time-taking. Another limitation is the problem related to data. Database may include corrupted missing data. There may be problems associated to the controlling of the records and privacy issues. Third limitation is the insufficient knowledge about mining technologies. A data collection team should have accurate knowledge about the domain.Finally healthcare organization should have an investment in time effort and money. It requires accurate planning, good preparation before beginning the data mining.

## 6. ADVANTAGES OF HEALTHCARE DATA GATHERING TOOLS

Computer technologies have made the medical care to make the patient records digitized.This includes all the information related to that patient whenever they visit the hospital.It will help to simplify the workflow of the management. We know that health care records are private information so before starting the mining process healthcare organizations must develop a privacy policy.

## 5.CONCLUSION

Data mining has a major place in the health area. Information gained by using data mining is applied for the better decision making of organizations to enhance their organizations working and health of their patients. Information can be gathered and kept in warehouse and can implement whenever we need it. Data mining provides a customer focus towards these data.The application of data mining using healthcare organizations can predict future aspirations of the patients. We can say that by using the emerging information technology data mining will get a potential in extracting information from hidden medical data.

## REFERENCES

1.Rakhi Ray–Advances in Data Mining: Healthcare Applications, 2018
2.Manaswini Pradhan–Data Mining& Health Care: Techniques of Application,2012
3.System Manish Pandey– Applications of Data Mining in Healthcare,2011