Analysis of Datamining Prediction Techniques in Healthcare Management System

K. Rajalakshmi	Dr. S. S. Dhenakaran
M.Phil Research Scholar	Assistant Professor
Computer Science &Engineering,	Computer Science &Engineering,
Alagappa University, Tamil Nadu, India	Alagappa University, Tamil Nadu, India

Abstract— One of the rapid growing fields is health care industries. The medical industries have great amount of data set collections about diagnosis, patient details and medications. To turns these data is into useful pattern and to predicting forthcoming trends data mining approaches are used in health care industries. The medical industries come across with new treatments and medicine every day. The healthcare industries should provide better diagnosis and therapy to the patients to attaining good quality of service. This paper explores different data mining techniques which are used in medicine field for good decision making.

Keywords— data mining, prediction techniques, decision making.

I. INTRODUCTION

Data mining is the methodology for finding hidden values from enormous amount of data. As the patients population increases the medical databases also growing every day. The transactions and analysis of these medical data is complex without the computer based analysis system. The computer based analysis system indicates the automated medical diagnosis system. This automated diagnosis system support the medical practitioner to make good decision in treatment and disease. Data mining is the massive areas for the doctors to handling the huge amount of patient's data sets in many ways such as make sense of complex diagnostic tests, interpreting previous results, and combining the different data together. Traditionally Infirmaries decision is shaped by the medical practitioner's observations and fore knowledge rather than the knowledge which obtain from the large amount of data. This automated diagnosis system leads to increases the quality of service provided to the patients and decreases the medical expenditure.

II. DATA MINING

Data mining is the process of combining the different data source and derives the new pattern from that data collection. The following diagram represents different stages of data mining process:
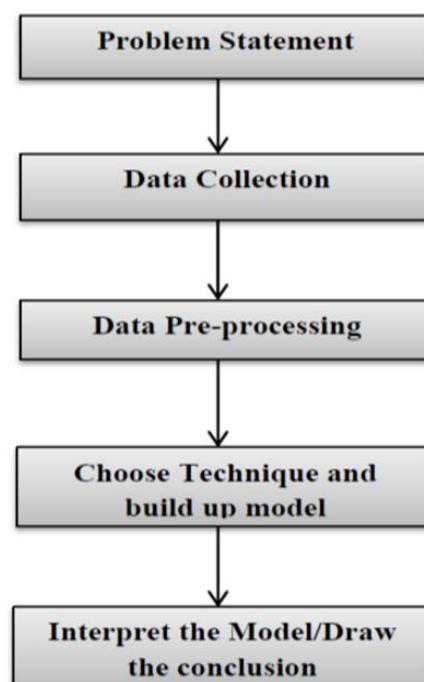

Fig. 1 Data mining process

III. HEALTHCARE DATA MINING

Healthcare data mining is the growing research area in data mining technology. Data mining holds great promising for healthcare management to allow health system to systematically use data and analysis to improve the care and reduce the cost concurrently could apply to as much as 30% of overall healthcare spending. In the healthcare management data mining prediction are playing active role. Some of the prediction based data mining techniques are as follows:

1. Decision tree
2. Bayesian Classifiers
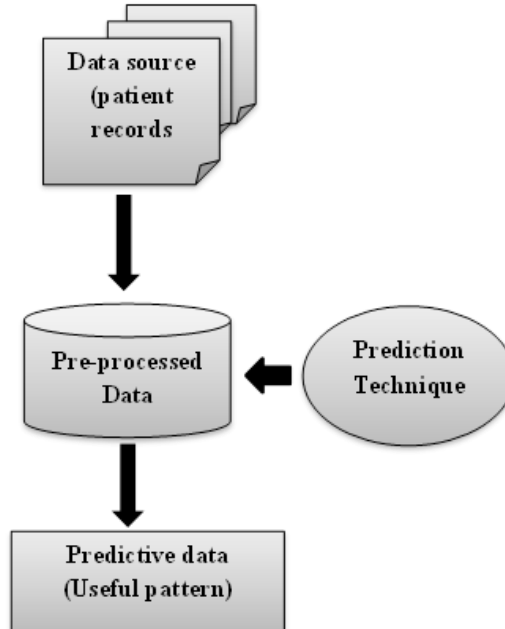3. Neural network
4. Support Vector Machine

Fig. 2 Healthcare Prediction

IV. PREDICTION TECHNIQUES

1. Decision Tree

The decision tree is the model that consists of root node, branch and leaf node. The root node is the top most nodes in the tree structure, each internal node specifies the test on attributes, the class label is hold by the leaf node, and the branch node is used to hold the test results.

Decision tree is easy and fast method since it does not require any domain knowledge. In the decision tree inputs are divided into two or more groups repeat the steps till complete the tree as shown on Fig. 3
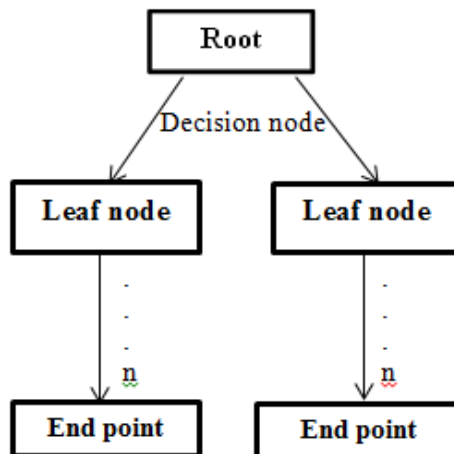

Fig. 3 Decision tree Structure

Some of the decision tree algorithms as follows:

a) ID3 (Iterative Dichotomiser 3)

b) C4.5 (Successor of ID3)

c) CART (Classification & Regression Tree)

d) CHAID (CHI-squared Automatic Interaction Detector)

2. Bayesian Classifier
It is statistical classification approach based on the Bayes theorem

Theorem:
P (B given A) = P (A and B)/P (A) to calculate probability of A given B, the algorithm counts the number of cases where A and B occurs together and divides it by the number of cases where A occurs alone. Let X be a data tuple, In Bayesian terms, X is considered "Evidence". Let H be some hypothesis, such that the data tuple X belongs to class C. P (H|X) is posterior probability, of H conditioned on X. In contract, P (H) is the prior probability of H.

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)}$$

$$Posterior = \frac{Likelihood \times Prior}{Evidence}$$

3. Artificial Neural Network
Neural network is a widely used decision making technique. Since 1959 neural network are proposed for healthcare decision making. In neural network the neurons are started with random weights. Neuron doesn't know anything they have to train.
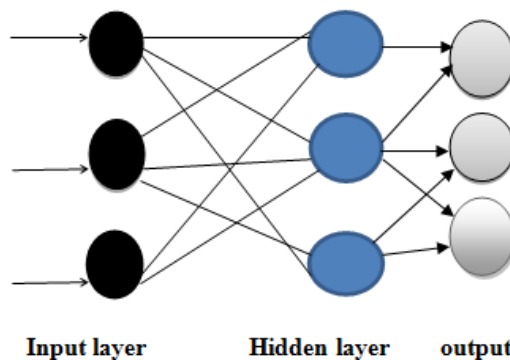

Input layer   **Hidden layer**   **output**
Fig.4 Neural Network

4. Support Vector Machine (SVM)
Normally SVM is the classification technique. Initially it developed for binary type classification later extended to multiple classifications. This SVM creates the hyper plane on the original inputs for effective separation of data points.
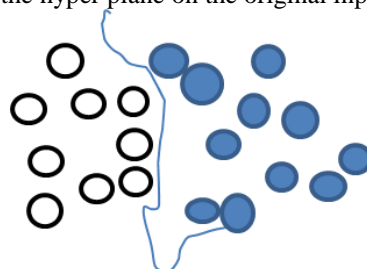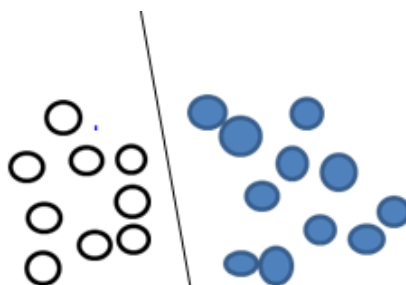

Fig. 5 Input

Fig. 6 Output using SVM

V. COMPARITIVE STUDY OF DIFFEEENT PREDICTION IN HEALTHCARE
An ultimate of the data mining in the medical domain is better prediction through the experience and scientific observations. This section explores different data mining prediction applications which are in medical domain.
This section examines data mining applications in medical domain by different research worker given in detail. Various data mining tools are used to predict in different healthcare problems. In this section, the following list of medical issues has been studied and estimated.

a. Heart disease
b. Cancer
c. Eye disease
d. Diabetics

There may be large number of data mining techniques and data mining tools are available for predicting heart disease, various cancers, diabetics, eye disease and dermatological conditions. The following table presents comparison of disease, data mining techniques and the accuracy of the data mining techniques.

Table 1 Comparison of data mining techniques

S.NO	ACCURACY (%)	DATA MINING TECHNIQUE	ALGORITHM	DISEASE
1.	75.21	Decision tree(SPSS)	Chi-Square	Diabetics
2.	91	Decision tree	C4.5	Diabetics
3.	94.3	SVM	SMO	Diabetics
4.	92	Decision tree/Neural network	Back propagation	Eye disease
5.	99	Neural network	Learning algorithm	Urinary system disease
6.	84.14	Classification	Naïve Bayes	Lung cancer
7.	86.7	Decision tree(WEKA)	C4.5	Breast cancer
8.	93.75	Regression Tree	-	Parkinson's disease
9.	88.76	Classification	Naïve Bayes	Heart disease
10.	86	Classification	Laplace Smoothing	Heart disease
11.	98.24	Classification	k- nearest neighbour's algorithm	Heart disease

VI. CONCLUSION

We presented this paper to analyze the various data mining application in the healthcare domain to discover new range of pattern information. There is variety of data mining tools and techniques are available for health care diagnosis systems that are clearly defined in this paper. This data mining based prediction system are reduces the human effects and cost effective one.

REFERENCES

[1] Muhamad Hariz Muhamad Adnan,Wahidah Husain, Nur'Aini Abdul Rashid, *"Data Mining for Medical Systems: A Review"* .
[2] V. Krishnaiah et al," *Diagnosis of Lung Cancer Prediction System Using Data Mining Classification Techniques"*, International Journal of Computer Science and Information Technologies, Vol. 4 (1) , 2013, 39 – 45.
[3] Abdelghani Bellaachia, Erhan Guven, "Predicting Breast Cancer Survivability Using Data Mining Techniques".
[4] Ravi Sanakal, Smt. T Jayakumari*," Prognosis of Diabetes Using Data mining Approach-Fuzzy C Means Clustering and Support Vector Machine"*,International Journal of Computer Trends and Technology (IJCTT) – volume 11 number 2 – May 2014.
[5] L. G. Kabari and E. O. Nwachukwu*," Neural Networks and Decision Trees For Eye Diseases Diagnosis"*.
[6] Qeethara Kadhim Al-Shayea and Itedal S. H. Bahia*,"Urinary System Diseases Diagnosis Using Artificial Neural Networks"*, IJCSNS International Journal of Computer Science and Network security, VOL.10 No.7, July 2010.

[7]   Dhanashree S.Medhekar,Mayur P.Bote,Shruti D.Deshmukh,*"Heart Disease Prediction using Naïve Bayes"*, International Journal Of Enhanced Research In Science Technology & Engineering Vol.2 Issue 3, March 2013.

[8]   Ms.Rupali R.Patil, *"Heart disease prediction system using Naïve Bayes and Jelinek-mercer smoothing"*, International Journal Advanced Research in Computer and Communication Engineering, Vol.3, Issue 5, may 2014.

[9]   A.H. Hadjahmadi, and Taiebeh J. Askari," *A Decision Support System for Parkinson's Disease Diagnosis using Classification and Regression Tree*", The Journal of Mathematics and Computer Science Vol. 4 No.2 (2012) 257 – 263.

[10]   Hian Chye Koh and Gerald Tan." *Data Mining Applications in Healthcare*".

[11]   M. Durairaj, V. Ranjani*," Data Mining Applications In Healthcare Sector: A Study"*, International Journal Of Scientific & Technology Research Volume 2, Issue 10, October 2013.