IOSR Journal of Engineering (IOSRJEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
Vol. 06, Issue 10 (Oct. 2016), ||V1|| PP 16-21

Applications of Data Mining Approaches In Healthcare

Najwa shehadeh Ahmad altabeiri
The Ministry of Educationmaster degree of vurricula and methods of teaching

Abstract:- We'll talk in this search for Data Mining, one of the interesting topics of his research, which has become increasingly popular in the WHO because there is a wealth of data within the health care systems and the lack of effective tools for thatData mining plays an important role for the detection of new trends in the health institution which in turn is beneficial to all parties associated with this field. This survey will discover the benefitsofdifferentdataminingtechniques.In this paper, we present a brief overview of these technologies and the most important advantages and disadvantages. We will highlight the most important applications and future challenges and issues of data mining in the area of health care.

Keywords*: Healthcare, health data, data mining, Classification*

I. INTRODUCTION

We will speak in this search for Data Mining, which has become one of the topics interesting, which aims to find useful information from large data sets At the present time is to extract the most popular data in the field of health care, because he needs to be an effective methodology for the detection of non-information is known in the health data .In the field of health is a mining and one data of the most important reasons for the detection of fraud in the health insurance and the provision of medical solutions to patients at a lower cost and detect the causes of disease and how to get to the style of treatment at the lowest cost as it also helps the researcher in the field of health care to make special for them in the field of welfare policies health and the purpose of data mining is to extract useful information from large databases and use these applications on commercial and scientific aspects in the field of health care (1)We will discuss in this study methods to extract data in a scientific aspect and style data mining distinguishes itself by the nature of the databases be quite different from the extraction of traditional data and here in this work we will conduct a comprehensive and detailed survey on the extraction of data in the health care sector and the types of data used, as well as details the extracted data to ensure a very useful data for use in healthcare .Here in this research data mining will provide a new approach toward the user to identify new patterns in the extraction of data that can be used by health care officials in order to improve the type of service provided they have also acquired knowledge of doctors to reduce the cost of treatment.In data mining is used information from previous data for the analysis of the results of a specific problem or a case and are working to extract data to analyze the data stored in the data warehouse this data may come from all over the world and is also used data in marketing strategies for their products and is also used the data to compare the competitors that can be used at the moment to increase sales and to promote the new product and delete products that do not represent the value of the company.Data Mining and Knowledge Discovery in Databases (KDD) are related terms and are used interchangeably but many researchers assume that both terms are different as Data Mining is one of the most important stages of the KDD process [2, 3].And organizes knowledge of the different stages of the data discovery process which is about the first stage is the selection of data, where at this stage the data from various different sources assembling In the second stage which is the analysis and data processing and Phase III stage is a conversion of data into the appropriate form or shape required to increase and the fourth phase of treatment is a data mining where here in this stage are suitable for data mining the data to extract information valuable and useful in recent fun as in Figure 1 technology application[2, 4].

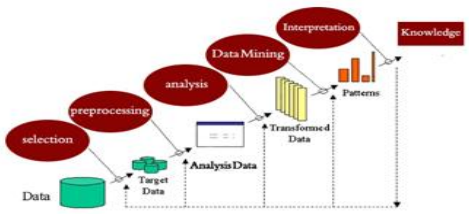

Figure (1) : Stages of Knowledge Discovery Process

Data Mining

Data mining or knowledge discovery process in databases is the process by which to extract implicit information that is not insignificant and can be useful data and also include groups such as technical curriculum and summarize data and classification[4].The stage data development has evolved significantly in recent years, as in Figure 2, which began in 1960 data mining system seemed a data mining to processing file and followed the next stage DBMS to be in 1970 and until 1980 was the data modeling tools have evolved from database management system in three categoriesIt must be resolved

1 - Advanced data that is evaluated in 1980 the rules of systems to provide data and applications in process models
2 - Work data storage and extract data from the late 1980's until now
3 - Database systems on the Internet and this began in 1990 and the systems existing databases to xml and this created a new process called the new generation of information system and this started in 2000

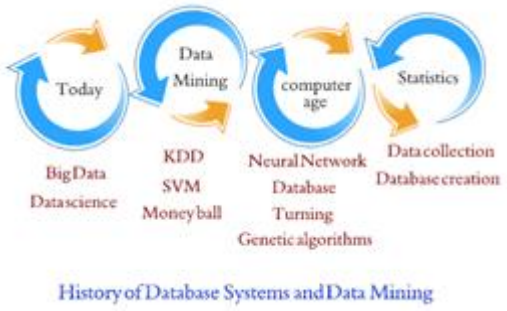

Figure (2) : History of Database Systems and Data Mining

But now, in the current era in February 2015 became the data extracted on a large scale in all the fields of trade, science, engineering, medicine and extraction of credit card transactions and the movement of the money market and national security has become the big data is now a common data collection has become cheaper and became the devices capable of collecting data spread fullylassification of data mining : Classification is one of the most domain names that are used in the extraction of data in health institutions and using different methods such as decision tree and SVM and many other approaches to data analysis [5].Used classification techniques to predict the cost of treatment in health care services and increase with the rapid growth every year and has become a major concern for all use to predict health care costs by using all three data years collected from health insurance companies to implement the experiment half-life tree approach is to use the data for two years and was before used for only one year to compare the expected results The following will discuss the classification algorithms used in the field of healthcare:[6].

1 - K-Nearest Neighbors (KNN) :

K-Nearest Neighbors (KNN) Seed is one of the simplest seed who is trying to discover the identity of an unknown point by using the knowledge of the data points (nearest neighbor) and points disaggregated data in accordance with the KNN are classified as more data points using one of the nearest neighbors.{ 7 } KNN has many applications in various fields, for example health data bases and fields of images and cluster analysis and pattern recognition and online marketing and etc. are used KNN to analyze distinct for LDA for the classification of chronic diseases in order to generate early warning system and is also used to analyze the relationship between the heart and the arteries, high blood pressure and factors of disease the risk for various chronic diseases in order to build a system based on early warning to reduce the incidence of complications of these diseases, as shown in Figure (3) is used for the analysis of patients with cardiovascular disease {8} .

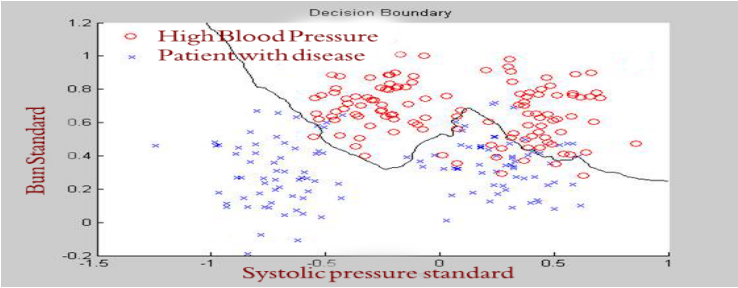

Figure (3) : K-NN Classifier for Chronic Disease

2 - Decision Tree (DT):

Decision tree is a graph which shows the structure where shows each test node on the sublime value of every branch represents the test results and tree leaves represent categories or a distribution of the decision tree class is a model predictive often used for classification and decision tree working on space division contribution in the cells where each cell belongs to one category represents this division as a series of tests each internal node in the decision tree represents the value of some variable input and called the branches of the node with the possibility of the potential of the test results. The node is a paper representing the cells and determine the category of return if the paper has been reached and thus the classification of homosexual types of inputs by starting at the root node and depending on the results of the tests until there is a node [9].Decision tree is represented in figure 4

.

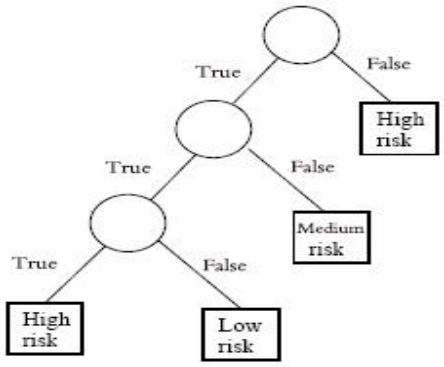

Figure (4) : Decision tree

3 - Artificial Neural Network

Artificial neural network has become a powerful tool in pattern recognition and how to report the problem, or the application of bonds tasks and is considered one of the technology that is used in signal processing ANN is an adaptation because learns to perform the function of the data and considers this phase adjustment is a change during the training process and after the process is completed Training is reform and if there is a lot of data or unintelligible bad problems is the use of the model ANN flour and there is also a linear properties of ANN provide a lot of flexibility to achieve the input map production and the artificial neural network to provide the capabilities of the user to select the network.(10)

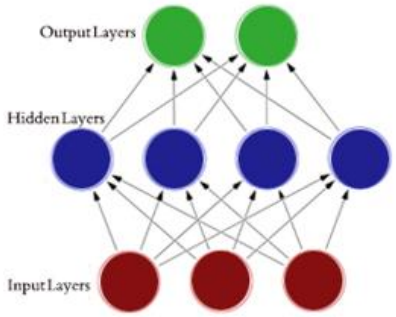

Figure (5) : Artificial Neural Network

4 - Support Vector Machine (SVM) .

SVM was developed initially Category bilateral, but can be renewed efficiently to the problems of multiple layers and supports seeded vector machine in fragments aircraft or multiple aircraft in fragments in space that use high dimensional, which may be useful for the classification and perform tasks efficiently SVM has many attractive features As a result, gaining international fame and popularity SVM in the performance of tasks is building aircraft in the internal space is used in the functions of the nucleus to draw non-linear maps of training samples into space and different functions such as core polynomial therefore used for this purpose. SVM operates on the principle that classification using excessive aircraft, which increase the separation between the points constructed with the help of support vectors and Figure 6 shows the work of the classification algorithm centers

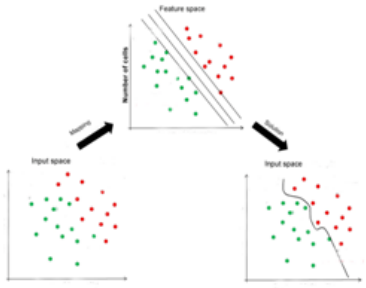

Figure (6) : Support Vector Machine (SVM)

Data Mining Application in healthcare:
The health care sector at the moment provides large amounts of complex data for patients and resources of the hospital and patient records and electronic diagnosis of diseases, as well as medical devices and considers this data is a major supplier of how to process, analyze and extract knowledge that provide support to reduce costs and how decision making can be classified as data mining applications in health care into categories such as :

1 - Treatment effectiveness

Data mining is used for patients and doctors in order to be compared with various treatments technology and provide excellent analysis of the course of the patient, which proves its effectiveness through the symptoms and varied reasons and courses of treatment therapy treatment, as well as Data Mining helps them to identify the side effects of treatment are to take the appropriate decision to reduce risk and develop appropriate to receive methodology treatment .

2 - Improved Healthcare management

Can develop data mining applications for the better can be traced chronic conditions dangerous for patients and data mining and disease reduces admissions number in the hospitals and ask hospitals to help health care management and is collecting a large amount of information with the progress in the electronic health record, as well as the activation of disease data service in a digital format to improve the quality system health care and also extract data helps service providers to determine the current and future patients to enhance the level of patient satisfaction needs

3 - Decrease Insurance Fraud and abuse

The health care system insured long a model for the detection of fraud and abuse in medical claims using the data model extraction techniques is very useful to determine the characteristics of improper third-party organization or phantom and patterns in medical claims submitted by physicians and patients or hospitals to reduce mining, fraud and misuse of applicationsYou can highlight the qualities of non-decent or fraud or fraudulent medical claims

4 - Effective management of Hospital

Allows extraction support data of how to build a model for the management of hospital resources that are important task in the field of health care through data extraction and that for the detection of chronic disease and for patients who have complications priority should be to give them so that they get an effective treatment in a timely and accurate manner and considers Organization non-profit health care provided to patients online access for their medical information on the Internet by filling in a prescription form and allow the exchange of secure e-mail with a health care provider.

5 - Better Customer relationship

Data Mining helps health care institute to understand the needs and patterns and preferences of behavior and educating customers in order to better relationship with customers through data mining is customers are curriculum core to manage the interactions between organizations such as banks, retailers or customers, it is equally important in the field of health care management6 - Management and Ranking of the HospitalIs used method to extract data to analyze various hospital details in order to determine their ranks are chosen the order of the hospital on the basis of their ability to deal with patients with serious cases and be patient treatment with serious cases a top priority for the hospital while the low risk factor and organizations

including modern hospitals are able to generate and collect huge amounts of data and extract data from the application of the data are stored in the hospital system.

The challenges and difficulties faced by Data mining in healthcare

We will get rid of this research show that data mining is one of the most important challenges in the field of health care in access to quality medical data relevant because most of the data are accurate, complete and also health data be complex and heterogeneous because they are collected from different sources of medical reports and are these sources either by the patient or by check with your doctor or a discussion with the patient and without the quality of the results there will be no useful results to extract the data successfully and also there are complications in the medical data is one of the big obstacles for the analysis of medical dataThat it is necessary to maintain the quality and accuracy of the data to extract the data for effective decision-making and from another view of health care data is data exchange and be health care institutions are not willing to exchange their data because of patient privacy in dealing with Dr. concerns, as well as most patients do not want to disclose health data, so the preservation of health and the Organization of the health insurance organization type as their data to maintain patient privacy and this is an obstacle in studies to detect fraud in the health insurance is the cost to start a data warehouse is very high and before the application of data mining techniques in the health data to be of collecting and recording data from different sources into a central repository of data which is a very expensive process in time and effort and the defect in the design of the data warehouse to extract the effective data

II. CONCLUSION

In this study, we talked about the application of data mining in the health care sector in order to extract useful information to predict disease using data mining applications is a difficult task, but significantly reduce the human effort and increase the diagnosis accuracy can develop the style of data mining to reduce cost and time constraints in terms of resources and human expertise.You will notice in this research we collected between more data to diagnose extraction techniques or to predict diseases in the health care sector could yield promising results the effective use of data mining in health institutions need to promote and exchange secure health data between the various parties, And some things are appropriate, such as relations between the patient and health care and WHO overcome security issues, and there is also a new approach to building a data warehouse in the recent years due to promote Internet for a huge data are available on the site and also a basic need of data mining techniques are analyzing the data to reveal the hidden information

REFERENCES

[1] H. C. Koh and G. Tan, "Data Mining Application in Healthcare", Journal of Healthcare Information Management, vol. 19, no. 2, **(2005)**. .

[2] U. Fayyad, G. Piatetsky-Shapiro and P. Smyth, "The KDD process of extracting useful knowledge form volumes of data.commun.", ACM, vol. 39, no. 11, **(1996)**, pp. 27-34.

[3] J. Han and M. Kamber, "Data mining: concepts and techniques", 2nd ed. The Morgan Kaufmann Series, **(2006)**.

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

[5] H. Hu, J. Li, A. Plank, H. Wang and G. Daggard, "A Comparative Study of Classification Methods For Microarray Data Analysis", Proc. Fifth Australasian Data Mining Conference (AusDM2006), Sydney, Australia. CRPIT, ACS, vol. 61, **(2006)**, pp. 33-37.

[6] R. Potter, "Comparison of classification algorithms applied to breast cancer diagnosis and prognosis", advances in data mining, 7th Industrial Conference, ICDM 2007, Leipzig, Germany, **(2007)** July, pp. 40-49.

[7] M. Silver, T. Sakara, H. C. Su, C. Herman, S. B. Dolins and M. J. O'shea, "Case study: how to apply data mining techniques in a healthcare data warehouse", Healthc. Inf. Manage, vol. 15, no. 2, (2001), pp. 155-164.

[8] C. H. Jena, C. C. Wang, B. C. Jiangc, Y. H. Chub and M. S. Chen, "Application of classification techniques on development an early-warning systemfor chronic illnesses", Expert Systems with Applications, vol. 39, (2012), pp. 8852-8858.

[9] LiorRokach and OdedMaimon,"Data Mining withDecision Trees: Theory and Applications(Series in

[10] Machine Perception and Artificial Intelligence)", ISBN:981-2771-719, World Scientific Publishing Company, ,2008.

[11] R. Andrews, J. Diederich, A. B. Tickle," A survey andcritique of techniques for extracting rules from trainedartificial neural networks", Knowledge-BasedSystems,vol.- 8,no.-6, pp.-378-389,1995.

[12] C. Hattice and K. Metin, "A Diagnostic Software tool for Skin Diseases with Basic and Weighted K-NN", Innovations in Intelligent Systems and Applications (INISTA), **(2012)**.

[13] W. L. Zuoa, Z. Y. Wanga, T. Liua and H. L. Chenc, "Effective detection of Parkinson's disease using an adaptive fuzzy k-nearest neighbor approach", Biomedical Signal Processing and Control, Elsevier, (2013), pp. 364-373.

[14] 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.

[15] M.Durairaj, K.Meena, ―A Hybrid Prediction System Using Rough Sets and Artificial Neural Networks‖, International Journal Of Innovative Technology & Creative Engineering (ISSN: 2045-8711) VOL.1 NO.7 JULY 2011.