Special Issue of International Journal of Computer Applications (0975 – 8887)
on Issues and Challenges in Networking, Intelligence and Computing Technologies – ICNICT 2012, November 2012

Evidence based Guideline System

Dharmendra
CDAC, Noida (INDIA)

Teena Verma
KIIT college of Engineering, Gurgaon (INDIA)

Puneet Kumar Aggarwal
BIT, Meerut (INDIA)

Parita Jain
KIET, Ghaziabad (INDIA)

ABSTRACT

Evidence-based Guideline System is providing an effective process to treating a patient. it totally work on the "Clinical Evidences" which are created by applying data mining techniques on collected data about particular disease and the rule generated by data mining process is further analyze by various medical field experts. These verified rules become the "Clinical Evidence" then these evidences are used to predict patient disease and provide Clinical guideline related to particular "Clinical Evidence" to patient if Patient medical parameters of disease matching with medical parameter of "Clinical Evidence". In order to obtain the best evidence for a given disease, external clinical expertise as well as internal clinical experience must necessary. In this research Data warehousing and data mining support the creation of the Evidence-based rules by providing a platform and tools for knowledge discovery. Large amounts of data can be analyzed to confirm known or discover unknown trends and correlations in data.

In this Dissertation, We will predict Thyroid Disease and giving appropriate clinical guidelines to Doctor to diagnose their patient according to Thyroid "Clinical Evidences". This dissertation is intended to provide a roadmap for achieving sustainable healthcare decision support system based on data warehouses and data mining, facilitating evidence-based medicine that which are used to diagnose patients.

Keywords— Clinical Evidence, EBM, EBGS, Thyroid, K-mean and K-medoid**.**

1. INTRODUCTION

Health organizations are facing big problem in providing good quality of service that treat the patients at inexpensible costs. Good quality of services is used to treat the patient and give positive results and is procured from poor clinical decisions. The decisions making in Healthcare that give positive result are totally dependent on doctor's perception and knowledgeable information present in clinical databases. Poor clinical decisions can lead to disastrous consequences which are therefore unacceptable. Most of the hospitals use their clinical data to diagnose patients and producing positive results on patients. The hidden knowledge present in clinical databases procures us to make a poor decision in patient diagnosis. Clinical decision support with computer-based patient records could reduce medical errors, enhance patient safety, decrease unwanted practice variation, and improve patient outcome. This hidden information may be utilized to diagnose a patient who is suffering from particular diseases e.g. Heart Diseases and also predict the disease on the basis of symptoms. Any doctor can use hidden information to treat his patient. Finding of evidence from knowledge extract from clinical databases is great job in front of medical persons.

EBM using data mining technologies makes it possible to automatically analyze huge clinical databases and to discover patterns behind them [6]. The integration of evidence-based medicine rules into clinical decision-support systems would both improve quality and reduce costs of care, by recommending only the most efficient treatments and medications. Practicing evidence-based medicine often relies on published clinical evidence in form of books, magazines, journals, healthcare protocols, clinical trials and best practice guidelines.

In order to obtain the best evidence for a given disease, external clinical expertise, as well as internal clinical experience, must be available to the healthcare practitioners at the right time and in the right manner. Consolidation of heterogeneous medical information sources and building of a huge medical data repository are the prerequisites for practice of evidence-based medicine.

There are two major requirements for the information technology in order to improve the quality of care:

1. to enable the inter-operability across heterogeneous medical information systems,
2. to integrate evidence-based knowledge into medical decision making process.

Data warehousing and Data Mining offers a comprehensive support for gathering, analyzing and presenting medical data [1]. EBGS can be used for building sustainable healthcare systems which contribute to more effective healthcare. The benefits of such system would be as follows:

A. QUALITY OF CARE:

EBGM relies entirely on proven medical experience. Incorporating EBM into the clinical decision-making process ensures that only impartial and scientifically verified knowledge will be offered to physicians.

B. INTEROPERABILITY:

In order to give the best possible care to the patient, physicians need insight into the patient's complete health record. During their lifetime, patients receive healthcare from diverse medical institutions that keep the records of the individual treatments for a mandatory period of time and creates a unique patient's health record which is then made available to the decision makers in a user-friendly manner.

C. DECENTRALIZED DATA STORAGE:

An electronic health record (EHR) is not necessarily stored as a single physical entity in a centralized system. Instead, it can be aggregated into a single coherent record from data stored at various geographical locations, when required.

SCALABILITY: A federated DWH utilizes a component-based architecture. Each new data source can be easily

included into the federation without redesign of the existing system.

2. RELATED WORK

A lot of work has been done on Evidence Based Medicine. These researches tell about how evidence based medicine can help to treat patients and what strategies they have used to predict disease on the basis of disease parameters. At what level they exactly classify the disease and up to what extends they are unable to identify disease.

In Ref. [3], Dynamic Evidence Based Medicine (DEBM) is defined as the process of finding evidence about the care of individual patients automatically and dynamically in the case that we cannot rely on any literatures or guidelines. In this paper, A framework for DEBM using data mining technologies that make it possible to automatically analyze huge clinical databases and to discover patterns behind them .

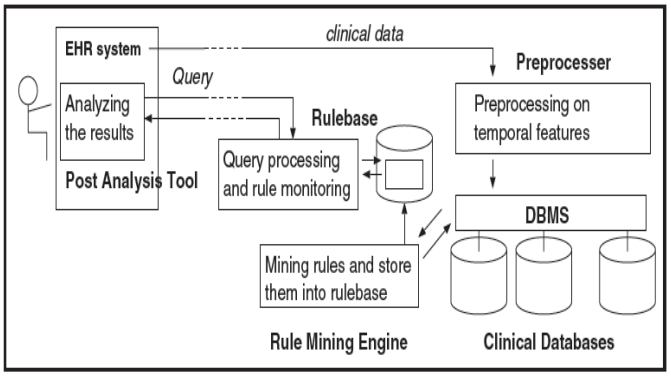

Figure 1. A Framework for DEBM

In this paper, there are some requirements of data mining system for DEBM are defined. These are two function for this system.

1. Supporting decision regarding healthcare.
2. Search some hidden patterns or rules.

For clinical decision making, rule discovery methods such as association rule mining are applied to this framework. Discovered rules are collected into a rule base for further analysis. By executing queries to the rule base, users can obtain the keys to evidence for making decisions about the clinical care. Here a prototype of rule base and post analysis tool is developed based on this framework. This tool can assist users to analyze discovered rules.

In Ref. [6], this paper presents a review of current research on the integration of HIT into clinical guidelines, a methodology for updating and altering the clinical guideline development process to include exploration of HIT and an updated version of the Patient Journey Modeling Architecture to include a technology layer to assist in visually depicting how HIT can benefit healthcare.

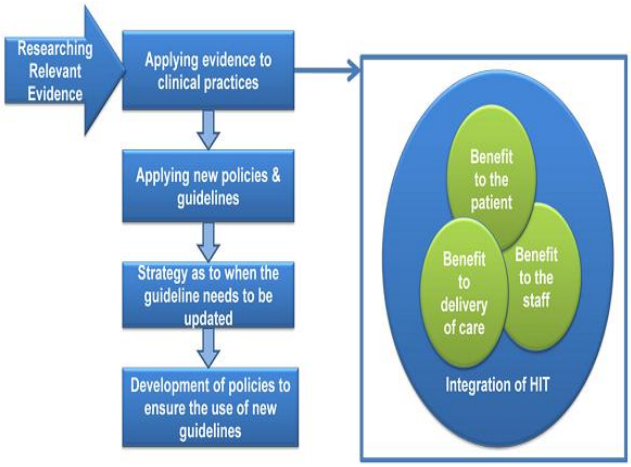

Figure 2. HIT

In this paper a new method is developed generating clinical guidelines. In above figure shows how can we apply evidence to clinical practices. A more defined process of collecting research based HIT evidence needs to be included within clinical guideline development to ensure that HIT is explored further. Understanding the benefits that HIT can bring to the patient, staff and to the delivery of care is an important part of guideline development. Once the benefits of HIT have been weighed, the guideline developers must then apply the use of HIT to clinical practice. Polices need to be reviewed and updated to ensure that the HIT is fully incorporated and followed by the healthcare professionals.

In Ref. [7], this paper has presented a DSS based on OLAP with data mining. The system is powerful because (1) it discovers hidden patterns in the data, (2) it enhances real-time indicators (3) it improves information visualization. The healthcare industry collects huge amounts of data which, unfortunately, are not turned into useful information for effective decision making. Decision support systems (DSS) can now use advanced technologies such as On-Line Analytical Processing (OLAP) and data mining to deliver advanced capabilities. This paper presents a model for clinical decision support system which combines both OLAP and data mining. It provide a way to extract knowledge from clinical Database that is impossible to achieve individually.

In this paper, a framework is design for making decision regarding patient care and appropriate action is taken for treatment. Here data warehousing and Data Mining concept is use extracting different rules from clinical database. Data Warehousing concept play a vital role for providing source in the form of Clinical Database and different operation of Data Warehousing are applied on this data for taking different view for analysis.

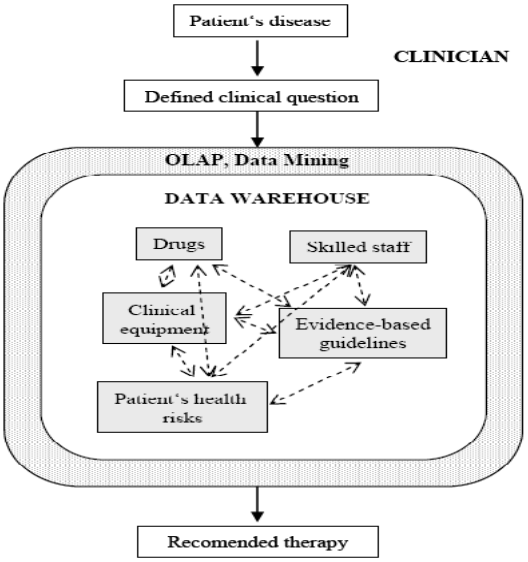

Figure 3. Data warehouse at the point of care [2]

The goal of the use of data warehouse and data mining to support evidence-based medicine in the domain of clinical pathways is to improve the coordination of care between different clinical departments and so to improve the quality of care and to reduce the length of patient's hospitalisation. Data mining techniques can be deployed for discovering the patterns of clinical pathways. Based on the patient record data, administrative data, clinical log data and evidence based rules, mining process is applied. With their usage, we can detect the structure of clinical paths and the sequence among activities, which human beings could hardly find [2].

3. METHODOLOGY

Evidence-based guidelines use methodical reviews to inform specific clinical conditions. These guidelines just give what type of treatments should apply on patient to treat them. These are not guiding a doctor to how to apply these guidelines on patients. Evidence-based guidelines are also beneficial for patient in getting idea about particular treatment. Evidence-based guidelines can improve the quality of care of patients by supporting involvements of proven benefit [3]. However, guidelines can also be confusing and cause harm, particularly if recommendations are based on an incomplete or erroneous clinical data available in clinical databases. Evidence-based guidelines aim to minimize these potential harms and provide effective treatments. For generating medical guidelines clinical expert opinion alone is not sufficient. There is need of science and methodology recommending medical guidelines. Evidence-based guidelines use the principle Evidence-based Medicine for producing medical guidelines.

Evidence-based guidelines are designed to review the evidence and address a specific question regarding a medical condition. Defining the question is a critical first step, and involves clearly defining the patient population, intervention, comparison and outcome. Guidelines can also increase the consistency of care, such that patients with identical clinical problems receive consistent and appropriate treatment, irrespective of where and by whom they are treated. Health care source can also benefit from guidelines since they can improve the quality of clinical decisions. Guidelines offer recommendations for clinicians who may be uncertain how to best diagnose or treat a specific condition, and can educate clinicians by providing a summary of the supporting evidence behind the recommendations.

Here I have design this system to provide clinical guidelines for Thyroid disease. All Calculation is done on Thyroid Disease Dataset for identify Disease Clusters. After applying clustering techniques on Thyroid Disease Dataset four clusters are formed Negative cases, Primary cases, compensated cases and Secondary cases.

Thyroid Disease Clinical Guidelines Generation Steps

1. Filling Missing values of Thyroid Disease Dataset with appropriate Values (by taking mean according to majority).
2. Compute Different Thyroid Disease Clusters.
3. Calculating Distance between any Disease Vector and Cluster Centroid after Normalization.
4. Prepare Clinical Evidence Database for Thyroid Disease
5. Checking Thyroid Disease and Provide Clinical Guideline to Patient

Cluster Analysis

Cluster is a collection of data objects. Grouping a set of data objects into clusters such that the data objects are similar to one another within the same cluster dissimilar to the objects in other clusters. The quality of a clustering result depends on both the similarity measure used and on the clustering method employed. Clustering is unsupervised classification: no predefined classes and no training examples. Similarity and Dissimilarity Between Objects:

1. Distances are normally used to measure the similarity or dissimilarity between two data objects

2. Some popular ones include: *Minkowski distance*:

$$d(i,j)=\sqrt[q]{(|x_{i_1}-x_{j_1}|^q+|x_{i_2}-x_{j_2}|^q+...+|x_{i_p}-x_{j_p}|^q)}$$

Where i = (x_{i1}, x_{i2}, …, x_{ip}) and j = (x_{j1}, x_{j2}, …, x_{jp}) are two p-dimensional data objects, and q is a positive integer

– If $q = 1$, d is Manhattan distance

$$d(i,j)=|x_{i_1}-x_{j_1}|+|x_{i_2}-x_{j_2}|+...+|x_{i_p}-x_{j_p}|$$

– *If $q = 2$*, d is Euclidean distance:

$$d(i,j)=\sqrt{(|x_{i_1}-x_{j_1}|^2+|x_{i_2}-x_{j_2}|^2+...+|x_{i_p}-x_{j_p}|^2)}$$

Properties

– $d(i,j) \geq 0$

– $d(i,i) = 0$

– $d(i,j) = d(j,i)$, *symmetry*

– $d(i,j) \leq d(i,k) + d(k,j)$, *triangular inequality*

Now we have unlearned data about Thyroid disease with 29 medical parameters. So, four clusters are formed Negative cases, Primary cases, compensated cases and Secondary cases. Let we assume four Thyroid disease vectors all four cases and

cluster whole Thyroid dataset into these four groups by using Data mining algorithm **K-mean** clustering. After Clustering four clusters are formed are below:

- 52.45308 F f f f f f f f f f f f f f f t 1.9404491 t 1.8719923 t 119.42871 t 1.0201257 t 116.78532 f
- 53.192593 F f f f f f f f f f f f f f f t 6.6424665 t 0.6780247 t 43.582718 t 0.43856782 t 40.54815 f
- 32.1 F f f f f f f f f f f f f f f t 8722.301 t 83.0 t 3132.7998 t 91.439995 t 3007.7 f
- 43.32219 F f f f f f f f f f f f f f f t 0.014027355 t 0.011246201 t 0.21276596 f 0.002705167 f 0.16413374 f

4. RESULT ANALYSIS

When we apply clustering technique of Data Mining on Hypothyroid disease then we got four cluster centroids which show level of disease. We traverse whole Dataset Hypothyroid disease with these cluster centroids disease case move towards their cluster centroid. Initially Disease Cases in Dataset are given in below table which will basically use for comparison.

Negative case : 3480

Compensated case : 194

Primary Case : 95

Secondary Case : 2

These are those disease cases which are exactly identify by clustering technique.

Negative case : 3095

Compensated case : 153

Primary Case : 75

Secondary Case : 2

Here is accuracy percentage that how much cases are exactly classified and how much over all accuracy of cases

Negative case : 88.93%

Compensated case : 78.86%

Primary Case : 78.94%

Secondary Case : 100%

Overall Accuracy of clustering: **88.17%**

5. EXAMPLE

In Below form you can insert disease reading of hypothyroid disease and press on checking and guideline generate button then if disease is present then result show in text area of Result and guideline generated for this reading in guideline text area.

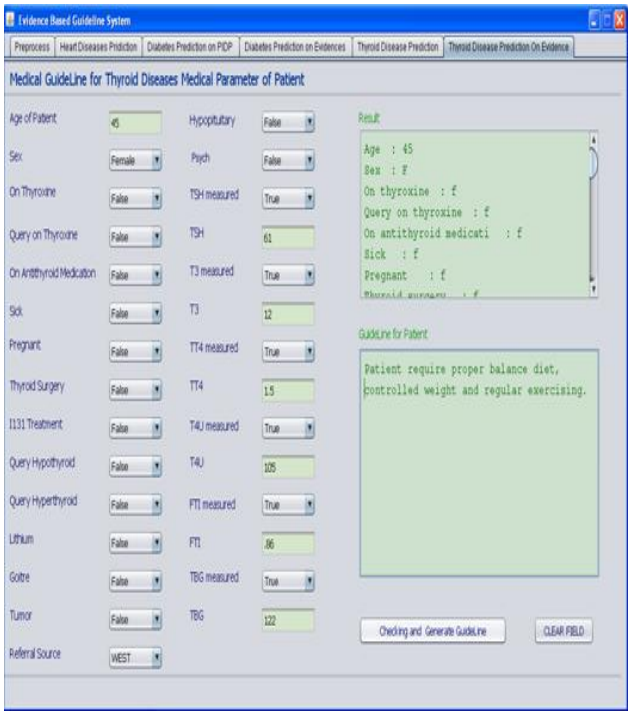

Figure 4. Guideline Generation

6. CONCLUSION

In this Research, Evidence-based guideline System is providing efficient clinical care to patients. EBGS totally works on clinical evidences that are procured/ obtained after very laborious process of clinical research, consultation of medical specialists and positive result with patients. The analysis of diseases is very essential and complex job in health care industries. Exact identification of disease in any patient who is suffering from particular disease in first attempt is impossible up to some extent. The clinical data store in databases and experience of specialist have great opportunity to provide valuable information that can be proved divinity for health care industry. We have a task to identity first valuable information from clinical databases in terms of patterns and these patterns become evidences after long process of analysis, lab research, and specialist verification and apply on patients that give positive result during patient diagnosis.

For identifying valuable pattern from clinical databases, we will apply different Data Mining algorithms like Decision K-mean and K-medoid for checking medical parameter of patient with Clinical evidence. For which Evidence with Disease vector of patient is matched then provide appropriate guideline according to Clinical Evidence is match. This match is not exact matching. This is only near aboutness of disease vector with Clinical evidence. The results that are obtained from clinical databases are used to treat patient and also used to predict the disease by which patient is suffering. And provide appropriate guideline to doctors for treatment of their patient according to clinical evidence.

7. REFERENCES

[1] Wen-Hong Chiu, Yuan-Chieh Chang and Hui-Ru Chi, "Exploring Modulating Effect within Evidence based Medicine Realization Based on Service Innovation Model" Sep,2009 IEEE

[2] Nevena Stolba and A Min Tjoa, "The relevance of data warehousing and data mining in the field of evidence-based medicine to support healthcare decision making" Dec 2005

[3] Gou Masuda, Norihiro Sakamoto and Ryuichi Yamamoto, "A Framework for Dynamic Evidence Based Medicine using Data Mining", Proceedings of the 15 th IEEE Symposium on Computer- Based Medical Systems (CBMS 2002)

[4] A Candelieri, D Conforti, and A Sciacqua, F Perticone, "Knowledge Discovery Approaches for Early Detection of Decompensation Conditions in Heart Failure Patients", Sep 2009 IEEE

[5] Sellappan Palaniappan and Rafiah Awang, "Intelligent Heart Disease Prediction System Using Data Mining Techniques " International Journal of Computer Science and Network Security, VOL. 8 No. 8 , August 2008.

[6] Candice MacDougall, Jennifer Percival and Carolyn McGregor, "Integrating Health Information Technology into Clinical Guidelines" Annual International Confrence of the IEEE EMBS Minneapolis, Minnesota, USA, September 2-6,2009

[7] Sellappan Palaniappan and Chua Sook Ling, "Clinical Decision Support Using OLAP With Data Mining " International Journal of Computer Science and Network Security, VOL. 8 No. 9, September 2008

[8] Latha Parthiban and R.Subramanian, "Intelligent Heart Disease Prediction System using CANFIS and Genetic Algorithm" International Journal of Biological and Life Sciences 3:3 2007

[9] K.Srinivas, B.Kavihta Rani and Dr. A.Govrdhan, "Applications of Data Mining Techniques in Healthcare and Prediction of Heart Attacks" International Journal on Computer Science and engineering Vol. 02, No. 02, 2010, 250-255

[10] Shantakumar B.Patil and Y.S.Kumaraswamy "Intelligent and Effective Heart Attack Prediction System Using Data Mining and Artificial Neural Network" European Journal of Scientific Research ISSN 1450-216X Vol.31 No.4 (2009).

[11] For Dataset :http://repository.seasr.org/Datasets/UCI/arff/ http://www.medicinenet.com/hypothyroidism/article.htm

[13] Jawai Han and Micheline Kamber, "Data Mining: Concepts and Techniques" Morgan Kaufmann, 2006.