

APPLICATIONS OF DATA MINING IN HEALTH PROTECTION AND HOSPITAL MANAGEMENT SYSTEM

[1]. N. Lakshmi Suvarna Latha, Assistant Professor, GTMW Proddatur,

[2]. B. Sree Harini, Assistant Professor, GTMW Proddatur.

ABSTRACT:

A hospital cannot be managed without the information system, where almost all the medical information is stored as multimedia databases. In this paper, we applied temporal data mining and exploratory data analysis techniques to manage the hospital data. We also propose a novel approach to analyze the health datasets collected by electronic health record (EHR) systems, insurance claims, health surveys, and other sources, using data mining techniques is very complex and is faced with very specific challenges, including data quality and privacy issues. This paper also propose a new approach to analyze the relationships between the wait time of outpatients and treatment processes they received based on the temporal pattern mining algorithm proposed by Batal et al.

Keywords—data mining , health data analysis,

INTRODUCTION:

Electronic healthcare records (EHR) have become essential tools to efficiently and effectively provide health protection services in hospitals. The information stored therein is primarily used for treatment of the patients; however, secondary use of the information now receives much attention as it provides a new way of understanding diseases, supporting medical decisions, as well as evaluating and improving hospital services. In this work, we attempt to analyze the relationships between the stay time of outpatients and treatment processes applied to these patients, as a framework to provide them with the prediction of time required for hospital visits. A treatment process is a sequence of clinical actions such as consultation and examinations. In order to capture the

relationships between a treatment process and stay time, we have to take care the temporal orders of actions and existence of overlaps. For example, if necessary examinations can be scheduled prior to consultation, a patient may be able to finish his/her visit earlier compared to a case where examinations have to be carried out during the consultation. In this paper, we report the results of mining frequent patterns of treatment process that have association to the stay time. We utilized the temporal mining algorithm proposed by Batal et al. as it can deal with temporal intervals of clinical events.

DATA MINING:

In the middle of 1990s, data mining came into existence as a strong tool to extract useful information from large datasets and find the relationship between the attributes of the data. Data mining originally came from statistics and machine learning as an interdisciplinary field [4], but then it was grown a lot that in 2001 it was considered as one of the top 10 leading technologies which will change the world.

A. Main Techniques

Data mining techniques are divided into two main categories: descriptive (or unsupervised learning) and predictive (or supervised learning) .

Descriptive data mining is an exploratory analysis that attempts to measure the similarity of records, and discover the patterns and relationships. The most important techniques in descriptive data mining are clustering and association. On the other hand, predictive data mining tries to generate predictive rules as a model to classify the

records based on a specific target (or label). Classification is the most widely used technique in predictive data mining. As there is no standard framework for a data mining process [1], it is important that the analyst himself holds good amount of skills in this area and understands the techniques very well to design an appropriate framework and achieve high quality and reliable outcomes. The main techniques have been introduced briefly in continue.

1) Classification

This technique is used when the data is required to be classified into different groups based on a target attribute – e.g. patient cost. Therefore, the classifiers predict the target label for each record using the input attributes. Some of the famous classification techniques are: decision trees, neural networks, K-nearest neighbors, support vector machines, Bayesian methods. According to a survey, decision tree algorithms are the most popular ones among all other classification techniques in the applications of data mining in healthcare. Classification techniques are widely used in health data analyses, including: analyzing microarray data , diagnosing skin diseases, performance of different classifiers on cancer dataset], predicting cost of healthcare services, identifying significant factors in healthcare coverage and predicting the status.

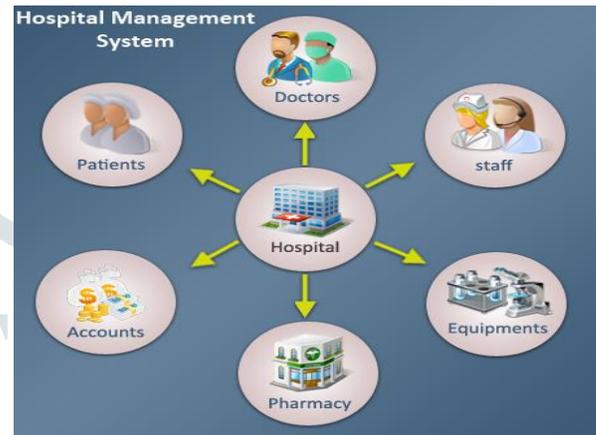
2) Clustering

This technique is used when we do not have much information about the different types of data objects involved in a population. As it is an unsupervised learning, it tries to find the cluster of data objects that have similarities to each other without considering any specific target label. Therefore, there are no predefined classes in contrast to classification. Different clustering techniques are: partitioned clustering, hierarchical clustering, and density based clustering.

3) Association

This technique is used when the relationship of attributes in a dataset needs to be identified – e.g. the association among the purchased items of a customer’s basket. This technique is mainly used in healthcare to detect the relationship

between diseases. In addition, association techniques can also be combined with classification techniques to increase the capability of analysis. For instance, the rules in a database or relationship of attributes in a dataset are detected, and then an efficient classifier is built by just considering the identified rules and including just the main attributes.



METHOD:

Frequent temporal pattern mining Temporal pattern mining method proposed by Batal et al. and is capable of finding frequent patterns from datasets containing multidimensional, numerical or categorical data, taking into account the temporal relationships between events by the use of abstraction and temporal logic.

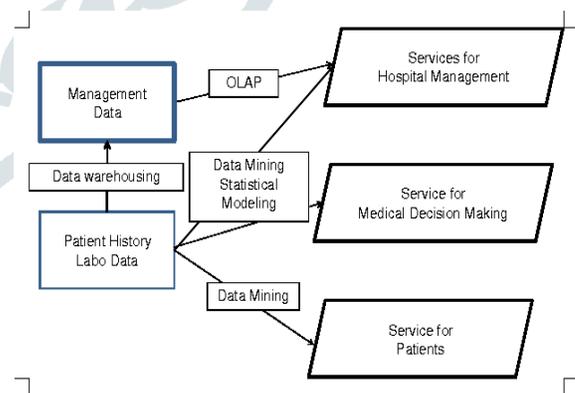


Fig: Service Oriented Hospital management

CONCLUSION:

In this paper, we applied several exploratory data analysis techniques to data extracted from hospital information systems. The results show several interesting results, which suggests that the reuse of stored data will give a powerful

tool to support a long-period management of a university hospital and we have also presented a method for mining rules that predict the stay time of outpatients from treatment processes based on the temporal mining method proposed by Batal et al.

REFERENCES:

[1] S. Tsumoto, “Knowledge discovery in clinical databases and evaluation of discovered knowledge in outpatient clinic,” *Information Sciences*, no. 124, pp. 125–137, 2000.

[2] —, “G5: Data mining in medicine,” in *Handbook of Data Mining and Knowledge Discovery*, W. Kloesgen and J. Zytkow, Eds. Oxford: Oxford University Press, 2001, pp. 798–807.

[3] Jensen PB, Jensen LJ, Brunak S.: Mining electronic health records: towards better research applications and clinical care. *Nat Rev Genet*.13(6):395–405 (2012).

[4] Batal I, Valizadegan H, Cooper GF, Hauskrecht M.; A Temporal Pattern Mining Approach for Classifying Electronic Health Record Data. *ACM Trans Intell Syst Technol*. 4(4) (2013).

