International Journal of Research, Science, Technology & Management ISSN Online: 2455-2240, Vol I Issue II, November 2015

Mining Risk Factor: A Rule Based Approach in Hospitals

J. Lala, Professor, MCA,
Poornima Engineering College, Chennai
JUSTIN B JOY PG Scholar, MCA,
Poornima Engineering College, Chennai

Abstract- **Organizations in new era grow larger and complex to provide advanced services due to the varieties of social demands .They are highly efficient for routine work processes but during unexpected situations known to be not robust. Over the past years the risk management has been noticed in organizations in various observations. It is expected that data mining methods will find interesting patterns from databases of such stored data and will be important for medical research and practice because human beings cannot deal with such a huge amount of data. This paper provides a practical introduction to knowledge discovery and data mining in medical databases, especially focusing on the following points: (1) the kind of problems medical people want to solve, (2) characteristics of medical data, (3) problems with medical data mining, especially the importance of preprocessing, and (4) an overview of existing research. We applied this technique to the following three medical domains: risk aversion of nurse incidents, infection control and hospital management. The results show that data mining methods were effective to detection of risk factors.**

I. INTRODUCTION

Since 1980's storage of data's includes from accounting information to laboratory data and even patient records are now stared to be accumulated: in other words, a hospital cannot function without the information system, where almost all the pieces of medical information are stored as multimedia databases. The development of computer systems has contributed to both medical research and practiceand their contribution is now entering into a new phase. Since the early 1980s, there has been a rapid growth in hospital information systems leading to a large proportion of laboratory examinations being stored as a huge database. Other types of data, including medical images, will be stored in hospital information systems within the twenty-first century. Human data analysis is characterized by a deep and short- range investigation based on their experienced "cases", whereas one of the most distinguished features of computer-based data analysis is to enable us to understand from the different viewpoints by using "cross-sectional" search. It is expected that the intelligent reuse of data in the hospital information system provides us to grasp the all the characteristics of university hospital and to acquire objective knowledge about how the hospital management should be and what kind of medical care should be served in the university hospital.

Medical data mining is still in its early days and many problems are still to be solved, even with existing data mining techniques. This suggests that this field should be a hot research topic in medical informatics in the twenty-first century and is awaiting further contributions. Especially, if the implementation of electronic patient records is progressed into the improvement on the efficiency of information retrieval, it may not be a dream for each patient to benefit from the personal database with all the healthcare information, "from cradle to tomb".

The analysis gives interesting results: (1) malignant neoplasm is the first major category which determines the profitability of Hospital, which is stable for twenty years. (2) In a global view, the length of stay is the principle factor for the revenue of the hospital, whose distribution follows the log-normal distribution. (3) Treatment method may be a secondary factor to determine the distribution of the length of stay for each disease, which may be correlated with the property that the length of stay follows log- normal distribution for each minor division in total. (4) Treatment without a surgical operation should be more examined by additional information, which is also important to evaluate the profitability of the university hospital.

II.SCOPE

This paper focuses on application of data mining to medical risk management. To err is human. However, medical practice should avoid as many errors as possible to achieve Safe medicine. Thus, it is a very critical issue in clinical environment how we can avoid the near misses and achieve the medical safety. Errors can be classified into the following three types of errors. First one is systematic errors, which occur due to problems of system and workflow. Second one is personal errors, which occur due to lack of expertise of medical staff. Finally, the third one is random error. The important point is to detect systematic errors and

personal errors, which may be prevented by suitable actions, and data mining is expected as a tool for analysis of those errors.

III. RELATED WORK

In existing System focuses on application of data mining to medical risk management. Medical accidents include not only careless mistakes of doctors or nurses, but also prescription errors, intra hospital infections or drug side-effects. The cause for such accidents may not be well investigated and it is unknown whether such accidents can be classified into systematic errors or random errors. Since the occurrence of severe accidents is very low, case studies are used for their analysis. However, in such investigations, personal errors tend to be the cause of the accidents.

IV.PROPOSED SYSTEM

This paper proposes risk mining where data including risk information is analyzed by using data mining methods and mining results are used for risk prevention. We assume that risk mining consists of three major processes: risk detection, risk clarification and risk utilization.

IV. MODULES DESCRIPTION

RECEPTION

The receptionist enters the patient details. The patient ID will be generated by the system. Receptionist provides them with a unique Patient ID to the Patients. Finally they handover the Patient Details to corresponding Doctors.

DOCTOR

Doctor's view the Patient Details and give the prescription to Patients. The Prescription details will be mailed to Nurses and give them the order to handle the patients.

NURSES

Nurses give the tablets, Injection, Trips etc to patients according to Doctor Prescription. Nurses send the Medicines Details to Pharmacy which are unavailable.

PHARMACY

The pharmacy in charge will get the patient ID and provide them with the necessary medicines asked from the nurse department The patients can also view patient details with help of the patient ID.

V.RULE INDUCTION METHOD

Rule induction methods have been proposed in order to discover knowledge automatically from databases. However, conventional approaches do not focus on the implementation of induced results into an expert system. This paper focuses not only on rule induction but also on its evaluation and presents a systematic approach from the former to the latter as follows. First, a rule induction system based on rough sets and attribute-oriented generalization is introduced and was applied to a database of congenital malformation to extract diagnostic rules. Then, by the use of the induced knowledge, an expert system which makes a differential diagnosis on congenital disorders is developed. Finally, this expert system was evaluated in an outpatient clinic, the results of which show not only that the system performs as well as a medical expert, but also that the system is very useful for instruction to medical residents.

V..RISK DETECTION

Patterns or information unexpected to domain experts may be important to detect the possibility of large-scale accidents. So, first, mining patterns or other types of information, which are unexpected to domain experts, is one of the important processes in risk mining. We call this process risk detection, where acquired knowledge is referred to as detected risk information.

The dataset includes the types of the near misses, the patients' factors, the medical staff's factors and the shift (early-night, late- night, and daytime) and the number of items of incidents collected was 245.

RULE INDUCTION. We obtained the followin interesting rules.
(Medication error):
If late-night and lack of checking, Then medication errors occur: probability (53.3%, 8/15). (Injection error):
If daytime and lack of checking, then injection incidents occur: probability (53.6%, 15/28). (injection error):
If early-night, lack of checking, And error of injection rate, Then injection incidents occur:
probability (50%, 2/4)

Those rules show that the time shift of nurse and lack of checking were the principal actors for medication and injection errors. Interestingly, lack of expertise (personal errors) was not selected. Thus, time shift and lack of checking could be viewed as risk factor for these errors. Since the conventional format of incident reports did not include future information about workflow, we had decided to ask nurses' to fill out new report form for each incident.

V.2.RISK CLARIFICATION

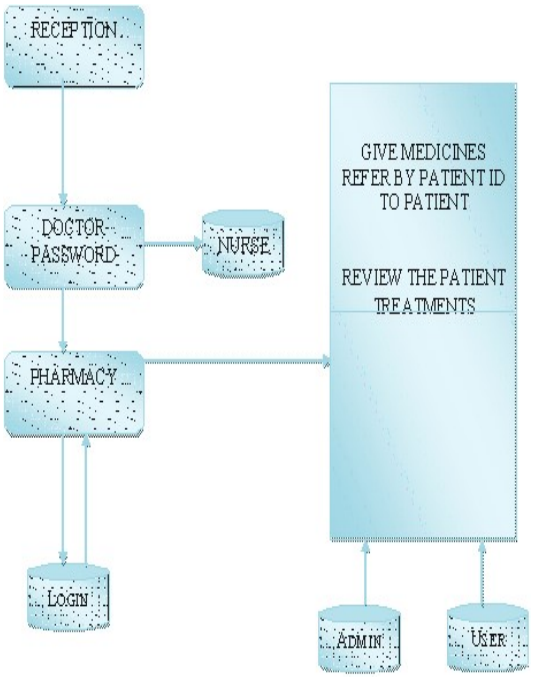
Focusing on detected risk information, domain experts and data miners can focus on clarification of modeling the hidden mechanism of risk. If domain experts need more information with finer granularity, we should collect more data with detailed information, and apply data mining to newly collected data. We call this process risk clarification, where acquired knowledge is referred to as clarified risk information.

RULE INDUCTION

The following rules were obtained: (medication error):
If the number of disturbing patients is one or more,
then medication errors occur:
probability (90%, 18/20). (medication error):
If nurses' work interrupted, then medication errors occur:
probability (80%, 4/5).

By addition of "the environmental factors", these high probability rules of medication errors were extracted.

After introducing the triple check system, the total number of the medication errors during the last 6 months decreased to 24 cases. It was considered that the nurses' education work was improved by the triple check system during the last 6 months.



V.3.RISK UTILIZATION

We have to evaluate clarified risk information in a real world environment to prevent risk events. If risk information is not enough to prevention, then more analysis is required. Thus, additional data collection is evoked for a new cycle of risk mining process. We call this process risk utilization. Where acquired knowledge is referred to as clarified risk information.

Therefore, it was decided that two nurses who had finished their shifts would prepare medicines for the next shift, and one nurse in charge of the medication would check the dose and identification of medicines alone (triple check by a total of 3 nurses). (However, heated discussions among domain experts (nurses) needed for this decision, as shown in Section 5.) Improvement was applied to the check system as a result of their discussion. During the last 6 months (April 2002 to October 2002), incident reports were collected.

VI.ARCHITECTURE DESIGN

VI.CONCLUSION

All the clinical information have been stored electronically as a hospital information system and expected that mining such combined data will give a new insight to medical accidents.

We applied risk-mining process to analysis of nurses' incident data. First, data collected in 6 months were analyzed by rule induction methods, which detect several important factors for incidents (risk detection). Since data do not include precise information about these factors, we recollect incident data for 6 months to collect precise information about incidents. Then, rule induction is applied to new data. Domain experts discussed all the results obtained and found several important systematic errors in workflow (risk clarification). Finally, nurses changed workflow to prevent incidents and data were recollected for 6 months.The frequency of medication errors has been reduced to one-tenth
(risk utilization).

REFERENCES

[1] J. Quinlan. C4.5 - Programs for Machine Learning. Morgan Kaufmann, Palo Alto, 1993.

[2] S. Tsumoto. Knowledge discovery in clinical databases and evaluation of discovered knowledge in outpatient clinic. Information Sciences, (124):125–137, 2000.

[3] S. Tsumoto. G5: Data mining in medicine. In W. Kloesgen and J. Zytkow, editors, Handbook of Data Mining and Knowledge Discovery, pages 798–807. Oxford University Press, Oxford, 2001.

[4] R. Bayardo, R. Agrawal, and D. Gunopulos, Constraint-based rule mining in large, dense database, *D*ata Mining and Knowledge Discovery Journal 2/3(2000) 217-240.

[5] J. Roberto J. Bayardo and R. Agrawal, Mining the most interesting rules, in:U. Fayyad, S. Chaudhuri and D. Madigan eds., *P*roceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining(ACM, New York, 1999) 145-154.

[6] G. I.Webb, Efficient search for association rules, in: R. Ramakrishnan, S. Stolfo, R. Bayardo and I. Parsa eds., *P*roceedinmgs of the 6th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (ACM, New York, 2000) 99-107.

[7] G. I.Webb and S. Zhang, K-optimal rule discovery, *D*ata Mining and Knowledge Discovery Journal 1(2005) 39-79.

[8] Y. Cheung and A. Fu, Mining association rules without support threshold: with and without item constraints, *I*EEE Transactions on Knowledge and Data Engineering 9(2004) 1052-1069.

[9] H. Li, J. Li, L. Wong, M. Feng, and Y.-P. Tan, Relative risk and odds ratio: a data mining perspective, in: C. Li, ed., *P*roceedings of the Twenty-fourth ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems (ACM, New York, 2005) 368-377.

[10] M. Ohsaki, S. Kitaguchi, K. Okamoto, H. Yokoi, and T. Yamaguchi, Evaluation of rule interestingness measures with a clinical dataset on hepatitis, in: J.-F. Boulicaut, F. Esposito, F. Giannotti, and D. Pedreschi, eds., *P*roceedings of the 8th European Conference on Principles and Practice of Knowledge Discovery in Databases (Springer-Verlag, New York, 2004) 362-373.

[11] M. M. Triola and M. F. Triola, *B*iostatistics for the Biological and Health Sciences (2nd ed.) (Addison-Wesley, Boston, 2005).

[12] J. Wang, J. Han, and J. Pei, Closet+: searching for the best strategies for mining frequent closed itemsets, in: L. Getoor, T. E. Senator, P. Domingos, and C. Faloutsos, eds., *P*roceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (ACM, New York, 2003) 236-245.[27] A. Asuncion and D. J. Newman, UCI repository of machine learning databases