ABSTRACT

of the Thesis

APPLICATIONS AND SOCIAL IMPACT OF DATA MINING TECHNIQUES IN HEALTH CARE MANAGEMENT

Many important disciplines such as databases, statistics, machine learning, information retrieval and artificial intelligence have influenced the development of data mining. The discovery of relational data model was of significant importance in the evolution of database systems. With the tremendous growth of data, users are expecting more relevant and sophisticated information which may be lying hidden in the data. Data mining is often described as a discipline to find hidden information in a database. It involves different techniques and algorithms to discover useful knowledge lying hidden in the data. Data mining and the term Knowledge Discovery in Databases (KDD) are often used interchangeably. It is commonly agreed that KDD is the process of finding useful information and pattern in data where as data mining may be described as the use of algorithms to extract useful information and patterns derived by the KDD process.

Several books such as Data Mining concepts and techniques by Fayyad *et al.* (1996), Berry and Linoff (1997), Han and Kamber (2001), Hand *et al.* (2001), Groth (1998), and Dunham (2003), examine data mining algorithms and other important concepts relating to data mining. Knowledge discovery by exploring new phenomena and gaining experiences is a daily and natural task for human being. It is also one of the desirable functions that artificial intelligence and expert systems need to perform. However, due to the

limited insight on the internal mechanism of the human reasoning process and incomparable computation architectures, this task is still far beyond what modern computer technology can reach. Some alternative approaches, such as neural network paradigms and inductive learning programming, have been proposed to deal with this task. These methods all have advantages and problems. Understanding how the knowledge patterns are derived and which phenomena they are based upon is critical for several application fields. In the medical domain, where the outcomes are affected by a large number of factors, one can try to understand how the outcomes are inferred with analogous cases and this can help the health care providers to arrange for the best treatments and for preventing adverse outcomes. To make explored knowledge patterns clinically useful, the combination of medical expertise and evidences from previous cases is the best practical approach.

It is well known that in Information Technology (IT) driven society, knowledge is one of the most significant assets of any organization. The role of IT in health care is well established. Knowledge Management in health care offers many challenges in creation, dissemination and preservation of health care knowledge using advanced technologies. Pragmatic use of database systems, data warehousing and knowledge management technologies can contribute a lot to decision support systems in health care.

Data mining technology provides a user oriented approach to novel and hidden patterns in the data. The discovered knowledge can be used by the health care administrators to improve the quality of service. The discovered knowledge can also be used by the medical practitioners to reduce the number of adverse drug effect, to suggest less expensive therapeutically equivalent

alternatives.

This thesis proposes use of Knowledge Discovery in Medical Databases, where medical experts can express their concerns and preferences to guide knowledge exploration from the data sets. When applying the derived knowledge patterns in medical work, the domain experts can further justify the decision support information and then refine the scope of the knowledge.

We propose to apply the principles of Knowledge Discovery in Databases (KDD) to a Medical databases. We have considered the creation of medical information system with the intention of mining. We have considered rule mining with reference to medical datasets. The health care providers can directly help in knowledge exploration process with their expertise and evidence from medical data sets.