Chapter 14  

Data Mining in Nonprofit Organizations, Government Agencies, and Other Institutions

Zhongxian Wang  
Montclair State University, USA

Ruiliang Yan  
Indiana University Northwest, USA

Qiyang Chen  
Montclair State University, USA

Ruben Xing  
Montclair State University, USA

ABSTRACT

Data mining involves searching through databases for potentially useful information, such as knowledge rules, patterns, regularities, and other trends hidden in the data. Today, data mining is more widely used than ever before, not only by businesses who seek profits but also by nonprofit organizations, government agencies, private groups and other institutions in the public sector. In this paper, the authors summarize and classify the applications of data mining in the public sector into the following possible categories: improving service or performance; helping customer relations management; analyzing scientific and research information; managing human resources; improving emergency management; detecting fraud, waste, and abuse; detecting criminal activities; and detecting terrorist activities.

INTRODUCTION

Data mining involves searching through databases for potentially useful information such as: knowledge rules, patterns, regularities, and other trends hidden in the data. In order to complete these tasks the contemporary data mining packages offer techniques such as neural networks, inductive learning decision trees, cluster analysis, link analysis, genetic algorithms, visualization etc. (Hand, Mannila, & Smyth, 2001; McPhail, 2008; Ranjan, 2009). In general, data mining is a data analytical technique that assists businesses in learning and understanding their customers so that
decisions and strategies can be implemented most accurately and effectively to maximize profitability. Data mining is not general data analysis, but a comprehensive technique that requires analytical skills, information construction, and professional knowledge.

Businesses are now facing globalized competition, and are being forced to deal with an enormous amount of data. The vast amounts of data and the increasing technological ability to store it also facilitated data mining. In order to gain a certain level of competitive advantage, data mining is now commonly adopted among businesses. Nowadays, data mining is more widely used than ever before; not only by businesses who seek profits, but also by nonprofit organizations, government agencies, private groups and other institutions in the public sector. Organizations use data mining as a tool to forecast customer behavior, reduce fraud and waste, and assist in medical research.

BACKGROUND

Data mining uses statistical analysis, artificial intelligence, and machine learning technologies to identify patterns that could not be found by manual analysis alone. The primary function of data mining has already amazed many people and is now considered one of the most critical issues towards a business’s success. However, data mining was not born all of a sudden. The earliest usage of data mining can be traced back in the World War II years. Data analytical methods such as model prediction, database segmentation, link analysis, and deviation detection, were used for military affairs and demographic purposes by the U.S. government, but data mining had not been seriously promoted until the 1990’s (Meletiou & Katsirikou, 2009).

Gramatikov (2006) compared statistical methods to data mining, differentiating them by the ultimate focus of these two tools. Statistical methods use data which is collected with a pre-defined set of questions. Statisticians are looking either for describing parameters of data or making inferences through statistics within intervals. With data mining, knowledge is generated from hidden relations, rules, trends and patterns which emerge as the data are mined.

The reason that data mining has been developed enormously again in the last few years is that huge amount of information was demanded by modern enterprises due to globalization. Important information regarding the markets, customers, competitors, and future opportunities were collected in the form of data to the database and needed data mining to unearth useful information and knowledge. Otherwise, a huge, overloaded, and unstructured database could just make it very difficult for companies to utilize and mislead the database users.

Public administration is, broadly speaking, the study and implementation of policy. The term may apply to government, private sector organizations and groups, and individuals. The adjective ‘public’ often denotes government at federal, state, and local levels, although it increasingly encompasses nonprofit organizations such as those of civil society or any not specifically acting in self-interest. Then, a long list exists: colleges and universities, health care organizations, charities, as well as postal offices, libraries, prisons, etc.

In the public sector, data mining initially were used as a means to detect fraud and waste, but have since grown into the use for purposes such as measuring and improving program performance. Data mining has been increasingly cited as an important tool for homeland security efforts, crime prevention, medical and educational application to increase efficiency, reduce costs, and enhance research.