

DATA MINING: A PREVIEW

Sarjon Defit, Mohd Noor Md Sap
Faculty of Computer Science and Information System
University Technology of Malaysia, KB. 791
80990 Johor Bahru, Malaysia
Telp: (07)-5576160, Fax: (07) 5566155
{sarjon_d@hotmail.com},{mohdnoor@fksm.utm.my}

Abstract

Today's databases contain large amount of data. It is almost impossible to manually analyze valuable decision data from large databases. There are many techniques that can be used to automatically extract the knowledge from data stored in the database. One of the techniques is data mining. This paper discusses the importance of data mining, the process of knowledge discovery in database, the kind of data to be mined, the data mining task, data mining prototype and discussion about data mining today and the next future.

Keywords: Predictive Data Mining, Knowledge Discovery in Database, Data Mining Task, Data Mining Prototype

1. Introduction

With the development of high capacity storage technology, a large amount of data can be stored. The data stored in databases consist of simple information, text document, or complex information such as multimedia, spatial databases, and hypertext documents (Zaiane, 1999). The stored data carry many valuable knowledge that can be extracted using many available tools. One of the tools is data mining. The result of the extracted knowledge can be applied to information management, query processing, decision process, process control and many other applications (Yongjian, 1996).

Data mining can be defined as searching for valuable information in large volumes of data. It is the process of extraction of implicit and potentially useful information such as knowledge, rules, constraints, and regularities from data stored in repositories. Data mining uses pattern recognition technologies as well as statistical and mathematical techniques (Jun and Keng, 1999).

Data mining is used for knowledge discovery, predictive modeling and forensic analysis process (William & Chapple, 1999). In terms of knowledge discovery, data mining is used to find interesting knowledge, patterns and then visualize the knowledge to the users. In predictive modeling, the extracted patterns or knowledge is used to predict the future behavior and in