

A Web Service Based Data Mining System

Student: Min-Hong Tsai

Advisor: Dr. Chieh-Yuan Tsai

Institute of Industrial Engineering and Management

Yuan-Ze University

ABSTRACT

With the trend of globalization and the rapidly changing commerce environment, many enterprises introduce knowledge management to increase the efficiency of their organizations and keep their critical competitive advantages. Data mining is a powerful knowledge management tool to extract knowledge from a large number of data. However, several problems have been encountered when implemented in enterprises. First, data exchange among heterogeneous information systems is hard to be achieved. Next, current data mining software are not flexible to provide different mining processes. Third, it is costly for enterprises to maintain or update current data mining software.

In this thesis, a web-based data mining system called “Dynamic Data Mining Web Service System” is introduced to solve the difficulties. In this system, each activity component in a data mining process is viewed as a web service on Internet. An eXtensible Markup Language (XML) is used as a medium to unify the format when data are transferred. Besides, An XML-based language, Business Process Execution Language for Web Service (BPEL4WS), is used to describe the details of data mining processes, including routing and information. The mining results in this system are described by Predictive Model Markup Language (PMML). PMML is an

XML-based language that provides users a formal method to define statistical and data mining models. The proposed system has successfully applied to a simple data mining example and a real-world credit card case. The results of these applications show that Dynamic Data Mining Web Service System can satisfy the requirements of enterprises.

Keywords: Data Mining, eXtensible Markup Language (XML), Web Service, Business Process Execution Language for Web Service (BPEL4WS), Predictive Model Markup Language (PMML)