



Chapter 8

Data Mining for Web-Enabled Electronic Business Applications

Richi Nayak

Queensland University of Technology, Australia

ABSTRACT

Web-enabled electronic business is generating massive amounts of data on customer purchases, browsing patterns, usage times, and preferences at an increasing rate. Data mining techniques can be applied to all the data being collected for obtaining useful information. This chapter attempts to present issues associated with data mining for Web-enabled electronic business.

INTRODUCTION

Web-enabled electronic business (e-business) is generating massive amounts of data such as customer purchases, browsing patterns, usage times, and preferences at an increasing rate. What can be done to utilize this large volume of Web data with rich description? One possible solution is the processing of all the data being collected to obtain some useful information. For instance, mining of such Web-enabled e-business data can provide valuable information on consumer buying behaviour, which is buried deep within the data otherwise, resulting in an improved quality of business strategies.

As corporations look toward the next phase of e-business (i.e., Web-enabled), one thing is clear—it will be hard to continue to capture customers in the future without the help of data mining. Examples of data mining in Web-enabled e-business applications are generation of user profiles, enabling customer relationship management, and targeting Web advertising based on user access patterns that can be extracted from the Web data. E-business companies can improve product quality or sales by anticipating problems before they occur with the use of data mining techniques. Data mining, in general, is the task of extracting implicit, previously unknown, valid and potentially useful information from data (Fayyad, Piatetsky, Shapiro, & Smyth, 1995).

Data mining in Web-enabled e-business domain is currently a "hot" research area. The

objective of this chapter is to present and discuss issues associated with data mining for Web-enabled e-business applications. This chapter starts with brief description of basic concepts and techniques of data mining. This chapter then extends these basic concepts for the Web-enabled e-business domain. This chapter also discusses challenges for data mining techniques when faced with e-business data, and strategies that should be implemented for better use of Web-enabled electronic business.

WHAT IS DATA MINING?

A typical data mining process starts with identifying a data mining problem depending on the goals and interest of a data analyst. Next, all sources of information are identified and a subset of data is generated from the accumulated data for the data mining application. To ensure quality, the data set is preprocessed by removing noise, handling missing information, and transforming to an appropriate format. A data mining technique or a combination of techniques appropriate for the type of knowledge to be discovered is then applied to the derived data set. The discovered knowledge is then evaluated and interpreted, typically involving some visualization techniques. Finally, the information is presented to the user to incorporate into the company's business strategies.

A data mining task can be decomposed into many sub-tasks when dealing with Web-enabled e-business data. Figure 1 illustrates a typical data mining process for Web documents. The process starts with locating and then retrieving intended Web documents or Web access logs. The next and most important task is analysis of data obtained from Web document(s) or logs. This includes preprocessing, actual mining process, and knowledge assimilation. In the end, the discovered knowledge is presented to user in a format that is appropriate to its goal. The analysis may indicate how a Web site is useful to a user in making decision or not. Information for a company to improve its Web site can be concluded from this analysis. The analysis may indicate business strategies to acquire new customers and retaining the existing one.

Various Data Mining Tasks and Techniques

Depending on the goals and interests of an end-user, a data mining process can have three possible tasks—predictive modelling, clustering, and link analysis.¹

Predictive Modelling The goal of predictive modelling is to make predictions based on essential characteristics about the data (Berry & Linoff, 2000). These goals are achieved by classification and regression tasks of data mining. The classification task of data mining builds a model to map (or classify) a data item into one of several predefined classes. The regression task of data mining builds a model to map a data item to a real-valued prediction variable. Both the tasks have same basic objective—to make a prediction about variable(s) of interest. The difference lies in the nature of the variable(s) being predicted - categorical variable(s) for the classification data mining task and continuous variable(s) for the regression data mining task.

Any supervised machine-learning algorithm that learns a model on previous or existing data can be used to perform predictive modelling on the data set. The model is given some already known facts with correct answers, from which the model learns to make accurate predictions. Mainly three techniques—neural induction, tree induction and bayesian classifiers—are used for classification data mining tasks (Lim & Loh, 2000). Some other classification methods are K-nearest neighbour classifiers, case-based reasoning, genetic algorithms, rough set, and fuzzy set approaches (Berry & Linoff, 2000; Han & Kamber, 2001).

9 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/data-mining-web-enabled-electronic/5195

Related Content

Privacy Factors for Successful Ubiquitous Computing

Linda Little and Pam Briggs (2009). *International Journal of E-Business Research* (pp. 1-20).

www.irma-international.org/article/privacy-factors-successful-ubiquitous-computing/3922

Using Social Networks Sites in the Purchasing Decision Process

Francisco Javier Miranda, Sergio Rubio, Antonio Chamorro and Sandra M. C. Loureiro (2014). *International Journal of E-Business Research* (pp. 18-35).

www.irma-international.org/article/using-social-networks-sites-in-the-purchasing-decision-process/116623

IT-Driven Business Model Innovation: Sources and Ripple Effects

Sune Müller and Mads Hundahl (2018). *International Journal of E-Business Research* (pp. 14-38).

www.irma-international.org/article/it-driven-business-model-innovation/201880

Mistakes in Remittance to Account

Masao Yanaga (2008). *Cyberlaw for Global E-business: Finance, Payments and Dispute Resolution* (pp. 182-193).

www.irma-international.org/chapter/mistakes-remittance-account-masao-yanaga/7496

Application of Web Services in the Context of E-Procurement: An SME Foci

Stanley Oliver and Kiran Maringanti (2009). *Electronic Business: Concepts, Methodologies, Tools, and Applications* (pp. 1064-1088).

www.irma-international.org/chapter/application-web-services-context-procurement/9337