Chapter 7

Web Mining for the Integration of Data Mining with Business Intelligence in Web–Based Decision Support Systems

Marcos Aurélio Domingues
University of São Paulo, Brazil

Alípio Mário Jorge
University of Porto, Portugal

Carlos Soares
University of Porto, Portugal

Solange Oliveira Rezende
University of São Paulo, Brazil

ABSTRACT

Web mining can be defined as the use of data mining techniques to automatically discover and extract information from web documents and services. A decision support system is a computer-based information sy Analysis stem that supports business or organizational decision-making activities. Data mining and business intelligence techniques can be integrated in order to develop more advanced decision support systems. In this chapter, the authors propose to use web mining as a process to develop advanced decision support systems in order to support the management activities of a website. They describe the Web mining process as a sequence of steps for the development of advanced decision support systems. By following such a sequence, the authors can develop advanced decision support systems, which integrate data mining with business intelligence, for websites.

INTRODUCTION

The management of web sites imposes a constant demand for new information and timely updates due to the increase of services and content that site owners wish to make available to their users, which in turn is motivated by the complexity and diversity of needs and behaviors of the users. Such constant labor intensive effort implies very high financial and personnel costs.

A decision support system is a computer-based information system that supports business and organizational decision-making activities. The integration of data mining with business intelligence techniques is important for the development of advanced decision support systems. Data mining is the application of specific algorithms for extracting patterns from data (Fayyad, Piatetsky-Shapiro, & Smyth, 1996). Business intelligence can be presented as a technology to collect and store data, analyze it using analytical tools, and deliver information and/or knowledge, facilitating reporting, querying, and allowing organizations to improve decision making (Azevedo, & Santos, 2011). An advanced decision support system, which integrates data mining with business intelligence, can support several management activities of a web site, reducing the effort to manage it.

Web mining is usually defined as the use of data mining techniques to automatically discover and extract information from web documents and services (Etzioni, 1996). According to (Kosala, & Blockeel, 2000), web mining is commonly categorized into three areas: web content mining that describes the discovery of useful information from content (Chakrabarti, 2000), web structure mining that analyzes the topology of web sites (Chakrabarti, Dom, Kumar, Raghavan, Rajago-palan, Tomkins, Gibson, & Kleinberg, 1999), and web usage mining that tries to make sense of the data generated by the navigation behavior and user profile (Srivastava, Cooley, Deshpande, & Tan, 2000). These areas naturally overlap and complement one another.

In this chapter, we propose to use web mining as a process to develop advanced decision support systems in order to support the management activities of a web site. We describe the web mining process as a sequence of steps for the development of advanced decision support systems. By following such a sequence, we can develop advanced decision support systems which integrate data mining with business intelligence. We have applied our proposed web mining process in the development of intelligent monitoring/management systems to guarantee the quality of web sites. Examples of monitoring activities include:

- **Usage:** Keep track of the paths users take during their accesses, the efficiency of pages/hyperlinks in guiding the users to accomplish their goals;
- **Users:** How users are grouped taking into account their browsing behavior, how groups change with time, how groups of users relate with the success of the site;
- **Data quality:** How adequate the content and meta-data of a web site are;
- **Automation:** The effect of personalization actions. For instance, if users are following the recommendations of products and pages or not.

This chapter is organized as follows: We start by presenting some work related to web mining. Then, we present our proposal. We show the main data sources that are used in the web mining process. We pay special attention to the pre-processing and storage of web data for use in web site management. Then, we present a case study where web mining is used as a process to develop an advanced decision support system that monitors the quality of the meta-data describing content in an e-news web portal, supporting decision making (Domingues, Soares, & Jorge, 2013). Finally, we present future research directions and conclusion.

21 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/web-mining-for-the-integration-of-data-mining-with-business-intelligence-in-web-based-decision-support-systems/116812

Related Content

Mining Association Rules in Data Warehouses
Haorianto Cokrowijoyo Tjioeand David Taniar (2005). *International Journal of Data Warehousing and Mining (pp. 28-62).*
www.irma-international.org/article/mining-association-rules-data-warehouses/1755

Automatic Ontology Construction: Ontology From Plain Text Using Conceptualization and Semantic Roles
Amita Arora (2020). *Critical Approaches to Information Retrieval Research (pp. 109-149).*
www.irma-international.org/chapter/automatic-ontology-construction/237643

Algebraic Reconstruction Technique in Image Reconstruction Based on Data Mining
Zhong Qu (2006). *International Journal of Data Warehousing and Mining (pp. 1-15).*
www.irma-international.org/article/algebraic-reconstruction-technique-image-reconstruction/1767

An Improvement of K-Medoids Clustering Algorithm Based on Fixed Point Iteration
Xiaodi Huang, Minglun Renand Zhongfeng Hu (2020). *International Journal of Data Warehousing and Mining (pp. 84-94).*
www.irma-international.org/article/an-improvement-of-k-medoids-clustering-algorithm-based-on-fixed-point-iteration/265258

A Review on Spatial Big Data Analytics and Visualization
Bangaru Kamatchi Seethapathyand Parvathi R (2018). *Modern Technologies for Big Data Classification and Clustering (pp. 179-197).*
www.irma-international.org/chapter/a-review-on-spatial-big-data-analytics-and-visualization/185983