Chapter 1.32
A Comparison and Scenario Analysis of Leading Data Mining Software

John Wang
Montclair State University, USA

Xiaohua Hu
Drexel University, USA

Kimberly Hollister
Montclair State University, USA

Dan Zhu
Iowa State University, USA

ABSTRACT

Finding the right software is often hindered by different criteria as well as by technology changes. We performed an analytic hierarchy process (AHP) analysis using Expert Choice to determine which data mining package was best suitable for us. Deliberating a dozen alternatives and objectives led us to a series of pair-wise comparisons. When further synthesizing the results, Expert Choice helped us provide a clear rationale for the decision. The issue is that data mining technology is changing very rapidly. Our article focused only on the major suppliers typically available in the market place. The method and the process that we have used can be easily applied to analyze and compare other data mining software or knowledge management initiatives.

INTRODUCTION

Based on the *knowledge life cycle* model, four stages of knowledge creation, knowledge storage/retrieval, knowledge transfer, and knowledge application have been proposed by Alavi and Leidner

(2001) and confirmed by Jennex (2006). "To be effective knowledge management systems, KMS, must support the various knowledge management functions of knowledge capture, storage, search, retrieval, and use" (Jennex, 2006, p.3). Knowledge discovery is generally one of the important stages or phases of KM. And while this incorporates identifying critical knowledge (this may also be what this stage is called), using data mining to aid in knowledge discovery is appropriate as being a useful KM tool.

Data mining is a promising tool that assists companies to uncover patterns hidden in their data. These patterns may be further used to forecast customer behavior, products and processes. It is important that managers who understand the business, the data, and the general nature of the analytical methods are involved. Realistic expectation can yield rewarding results across a wide range of applications, from improving revenues to reducing costs (Davenport & Harris, 2007; Porter & Miller, 2001). It is crucial to properly collect and prepare the data, and to check the models against the real figures. The best model is often found after managers build models of several different types or by trying different technologies or algorithms. This alone demonstrates the active role managers play in the data mining or other knowledge management processes.

Selecting software is a practical yet very important problem for a company (James, Hakim, Chandras, King, & Variar, 2004). However, not enough attention is given to this critical task. Current literature is quite limited because selecting software is such a complex problem, due to many criteria and frequent technology changes (Elder IV & Abbott, 1998; Giraud-Carrier & Povel, 2003). Haughton, Deichmann, Eshghi, Sayek, Teebagy, and Topi (2003) generally reviewed several computer *software packages* for *data mining, including* SPSS Clementine, XLMiner, Quadstone, GhostMiner, and SAS Enterprise Miner. Corral, Griffin, and Jennex (2005) exam-

ined the potential of knowledge management in data warehousing from an expert's perspective. Jennex (2006) introduced technologies in support of knowledge management systems.

Firstly, this article will take a brief look at data mining today, through describing some of the opportunities, applications and available technologies. We will then discuss and analyze several of the most powerful data mining software tools available on the market today. Ultimately, we will also attempt to provide an analytical analysis and comparison among the brands we have selected. Our selection is based, in part, on our own experience using data mining software as well as writing data mining code, SQL code and our work as relational database administrators. For our analytical comparison we will be using *Expert Choice* (Version 11) advanced decision support software.

DATA MINING SOFTWARE

Data mining software analyzes- based on open-ended user queries- relationships and patterns that are stored in transaction data. Available are several types of analytical software: statistical, machine learning and neural networks, decision trees, Naive-Bayes, K-Nearest Neighbor, rule induction, clustering, rules based, linear and logistical regression time sequence, and so forth. Along the lines of Mena (1998) and Martin (2005), the basic steps of data mining for knowledge discoveries are:

1.   Define business problem
2.   Build data mining data base
3.   Explore data
4.   Prepare data for modeling
5.   Build model
6.   Evaluate model
7.   Deploy model
8.   Results

Related Content

A Novel Approach to Parkinson's Disease Progression Evaluation Using Convolutional Neural Networks
Mhamed Zineddine (2023). *International Journal of Software Innovation (pp. 1-26).*
www.irma-international.org/article/a-novel-approach-to-parkinsons-disease-progression-evaluation-using-convolutional-neural-networks/315655

Modular Game Engine Design
Aaron Boudreauxand Brandon Primeaux (2014). *Software Design and Development: Concepts, Methodologies, Tools, and Applications (pp. 1179-1199).*
www.irma-international.org/chapter/modular-game-engine-design/77753

Applications of Filter Banks to Communications
Cihan Tepedelenliogluand Gerogios B. Giannakis (2002). *Multirate Systems: Design and Applications (pp. 225-256).*
www.irma-international.org/chapter/applications-filter-banks-communications/27229

Development of a Master of Software Assurance Reference Curriculum
Nancy R. Mead, Julia H. Allen, Mark Ardis, Thomas B. Hilburn, Andrew J. Kornecki, Rick Lingerand James McDonald (2010). *International Journal of Secure Software Engineering (pp. 18-34).*
www.irma-international.org/article/development-master-software-assurance-reference/48215

Hardware/Software Implementation for Wireless Sensor Network Applications
Mohamed Wassim Jmal, Oussema Ghorbel, Olfa Gaddourand Mohamed Abid (2013). *Embedded Computing Systems: Applications, Optimization, and Advanced Design (pp. 306-322).*
www.irma-international.org/chapter/hardware-software-implementation-wireless-sensor/76962