Chapter 8.7
Symbiotic Data Miner

Kuriakose Athappilly
Western Michigan University, USA

Alan Rea
Western Michigan University, USA

INTRODUCTION

Symbiotic data mining is an evolutionary approach to how organizations analyze, interpret, and create new knowledge from large pools of data. Symbiotic data miners are trained business and technical professionals skilled in applying complex data-mining techniques and business intelligence tools to challenges in a dynamic business environment.

BACKGROUND

Most experts agree (Piatetsky-Shapiro, 2000; Thearling, 2007) that data mining began in the 1960s with the advent of computers that could store and process large amounts of data. In the 1980s, data mining became more common and widespread with the distribution of relational databases and SQL. In the 1990s, business saw a boom in data mining as desktop computers and powerful server-class computers became affordable and powerful enough to process large amounts of data in data warehouses (Havenstein, 2007) as well as real-time data via online analytical processing (OLAP). Today we see an increasing use of advanced processing of data with the help of artificial intelligence technology tools such as fuzzy logic, decision trees, neural networks, and genetic algorithms (Brachman, et al., 1996; Gargano & Raggad, 1999). Moreover, current trends are moving organizations to reclassify data mining as business intelligence using such tools as Cognos (Cognos, 2007).

We also see three distinct theoretical approaches to data mining: statistical (classical), Artificial Intelligence (heuristics), and machine learning (blended AI and statistics). The three approaches do not adhere to the historical boundaries applied to data mining; rather, they are embarkation points for data-mining practitioners (Thuraisingham, 1999; Kudyba & Hoptroff, 2001; Kepner & Kim, 2003; Padmanabhan & Tuzhilin, 2003). It is not the intent of this discussion to argue which approach best informs data mining. Instead, we note that many software platforms adhere to one or more methods for solving problems via data-mining tools.

Most organizations agree that sifting through data to create business intelligence, which they can use to gain a competitive edge, is an essential business component (Lee & Siau, 2001; Brown, 2004; MacInnis, 2004; Burns, 2005). Whether it is to gain customers, increase productivity, or improve business processes, data mining can provide valuable information if it is done correctly. In most cases, a triad of business manager, information technology technician, and statistician is needed to even begin the data-mining process. Although this combination can prove useful if a symbiotic relationship is fostered, typically the participants cannot effectively work with one another because they do not speak the same language. The manager is concerned with the business process, the technician with software and hardware performance, and the statistician with analyses of data and interpretations of newfound knowledge. While this may be an overgeneralization, it is not far from the truth (O'Hara, 2007).

What is needed, then, is an individual who can pull all three components together: a symbiotic data miner trained in business, technology, and statistics.

MAIN FOCUS

In this chapter we will discuss how an individual trained not only in business but also in technology and statistics can add value to any data mining and business intelligence effort by assisting an organization to choose the right data-mining techniques and software, as well as interpret the results within an informed business context.

Data Mining in Contemporary Organizations

Data mining is the "semi-automatic discovery of patterns, associations, changes, anomalies, rules, and statistically significant structures and events in data" (Dhond, et al., 2000, p. 480). Analyzed data is many times larger than the sum of its parts. In other words, data mining can find new knowledge from observing relationships among the attributes in the form of predictions, clustering, or associations that many experts might miss. The new knowledge in a continuously changing environment is the most potent weapon for organizations to become and remain competitive (Inmon, 1996; Amato-McCoy, 2006; Corbitt, 2006).

In today's business organizations intelligence is necessary to anticipate economic trends, predict potential revenue streams, and create processes to maximize profits and efficiency. This is especially true for strategic and other mid-level managers (Athappilly, 2003). In the past, many decisions were made using corporate experience and knowledge experts. This is still true today. However, with the increased influx of data—some experts argue that the amount of information in the world doubles every 20 months (Dhond, et al., 2000; Tallon & Scannell, 2007)—many high-level managers now turn to data-mining software in order to more effectively interpret trends and relationships among variables of interest (Deal, 2004).

Related Content

Social Acceptability of Open Source Software by Example of the Ubuntu Operating System
Mateusz Szotysik (2014). *Frameworks of IT Prosumption for Business Development (pp. 158-176).*
www.irma-international.org/chapter/social-acceptability-of-open-source-software-by-example-of-the-ubuntu-operating-system/78773

Clinical Decision Support Systems in Nursing
Dawn Dowding, Rebecca Randell, Natasha Mitchell, Rebecca Foster, Valerie Lattimerand Carl Thompson (2010). *Business Information Systems: Concepts, Methodologies, Tools and Applications (pp. 1056-1070).*
www.irma-international.org/chapter/clinical-decision-support-systems-nursing/44122

Evaluating Conceptual Modeling Practices: Composites, Things, Properties
Graeme Shanks, Jasmina Nurediniand Ron Weber (2005). *Business Systems Analysis with Ontologies (pp. 28-55).*
www.irma-international.org/chapter/evaluating-conceptual-modeling-practices/6118

Ontological Analysis of Reference Models
Peter Fettkeand Peter Loos (2005). *Business Systems Analysis with Ontologies (pp. 56-81).*
www.irma-international.org/chapter/ontological-analysis-reference-models/6119

Towards a Middleware Based on SOA for Ubiquitous Care of Non-Communicable Diseases
Henrique Damasceno Vianna, Fábio Pittoli, Emerson Butzen Marquesand Jorge Luis Victoria Barbosa (2017). *Exploring Enterprise Service Bus in the Service-Oriented Architecture Paradigm (pp. 305-326).*
www.irma-international.org/chapter/towards-a-middleware-based-on-soa-for-ubiquitous-care-of-non-communicable-diseases/178077