

Teaching a Data Mining Course to MBA Students

Sathasivam Mathiyalakan

Winston-Salem State University, USA

George Heilman

Winston-Salem State University, USA

Sharon White

Winston-Salem State University, USA

INTRODUCTION

The growth in use of sophisticated Information Technology continues to generate huge amounts of data. The quantity of data has become so enormous that it is now often referred to simply as “big data.” Conventional wisdom says “big data” has become so big that many organizations are “data rich” but lack the tools and expertise to deal effectively and efficiently with the data, and so end up “information poor.” So how can a business filter through all those data to extract important and meaningful information? Three commonly used terms that describe various collections of methods, tools and techniques used to tame “big data” include Business Intelligence, Business Analytics and Data Mining.

Business Intelligence (BI) is a broad category of processes, methods, measurements, technologies and software applications that gathers, stores, provides access to, and analyzes an enterprise’s raw data. The intent is to help business users cut costs, identify inefficient business processes, identify new business opportunities and make better decisions. BI includes applications and activities like querying and reporting, statistical analysis, forecasting, decision support systems, online analytical processing (OLAP), Business Analytics and Data Mining (Rouse, 2006; McCabe, 2010; Mulcahey, 2013).

Business Analytics (BA) emphasizes the use of statistical analysis to methodically and iteratively

investigate organizational data. The intent is to gain insights from the data that will improve decision making (Rouse, 2010). While BA is generally considered to be a subset of BI, sharing many of its methods, tools and outcomes, Sandeep Raut (2011), Director of BI and Analytics at Syntel, indicates a distinction by pointing to BA’s greater emphasis of statistical tools and techniques.

Data Mining focuses on sorting and extracting useful information from large data sets in order to identify patterns and establish relationships (Han, Kamber, & Pei, 2011; Rouse, 2008). It aims to improve organizational decision making and strategic advantage by providing a better understanding of customers, suppliers, employees and other stakeholders. Data Mining is by nature an interdisciplinary field of study, drawing from knowledge domains like statistics, business, economics and IT, and may incorporate a wide variety of tools, techniques and methods including association, classification, clustering, decision trees, forecasting, genetic algorithms, neural networks, predictive analytics, rule induction, and sequence or path analysis (Alexander, 1997; Han 2001; Rouse, 20008).

There is growing evidence that Business Intelligence, Business Analytics and Data Mining will pay off for firms willing to make a commitment and investment. For example, a review of prior research conducted by Evans and Linder (2012) revealed a strong relationship between the use of analytics and a firm’s performance as measured

DOI: 10.4018/978-1-4666-5202-6.ch220

in terms of profitability, revenue and shareholder return. The attractiveness of investment in analytics is also supported by other research such as Nucleus Research's report (2011) which says that each \$1 invested in Business Analytics returns \$10.66, and the Information Management article (2002) reporting on an IDC study that found an average 431% ROI on analytics projects.

The expanding use of analytics in business and industry has created a growing demand for qualified analytic personnel. In response, many universities have developed programs or courses in Business Intelligence, Business Analytics or Data Mining. Evidence of academia's response can be seen at the Web site of KDnuggets, an online community for Data Mining practitioners and students. KDnuggets maintains a page entitled "Education in Data Mining, Analytics, and Data Science in USA/Canada" which lists colleges and universities providing degrees in Data Mining and other forms of analytics. The page currently lists about 50 universities with analytics-related degree granting programs (KDnuggets, 2013).

The growth of BI, BA and Data Mining offerings in academia has been aided by technology advances in hardware and software that have enabled courses dealing with the analysis of "big data" to be offered in a desktop environment. In this paper we report on our experience incorporating desktop technology into an elective graduate course in data analysis named Data Mining for Business Intelligence. We also discuss lessons learned and provide suggestions for other instructors interested in developing a course in Data Mining.

BACKGROUND

To support the effort of implementing Data Mining in academia, the Association for Computing Machinery (ACM) Special Interest Group for Knowledge Discovery of Data (ACM SIGKDD) has established a Curriculum Committee. In 2006, the Intensive Working Group of ACM SIGKDD Curriculum Committee released Version 1.0

of a proposal for a Data Mining Curriculum (Chakrabarti, Ester, Fayyad, Gehrke, Han, Morishita, Piatetsky-Shapiro, & Wang, 2006). The committee considers Data Mining's core endeavor to be extracting knowledge from data and capturing that knowledge in a human-understandable form. To impart the principles and concepts needed to discover structure in data, they recommended a curriculum that includes the following components: 1) database and data management issues, 2) data preprocessing, 3) choice of model and statistical inference considerations, 4) interestingness metrics, 5) algorithmic complexity considerations, 6) post-processing of discovery structure, 7) visualization and understandability, and 8) maintenance, updates, and model life cycle considerations.

ACM has also teamed with the Association for Information Systems (AIS) to develop a curriculum guideline for undergraduate degree programs in Information Systems (IS) (Topi, Valacich, Wright, Kaiser, Nunamaker, Sipior, & de Vreede, 2010). The authors recommend that, even at the undergraduate level, IS students study the concepts of Business Intelligence, Data Warehousing, Data Mining and analytics in at least 2 courses.

MAIN FOCUS

Course Design

The Data Mining for Business Intelligence course is offered to students enrolled in an evening MBA program. The course has an applied focus, with an emphasis on using automated statistical techniques to analyze business scenarios and solve business problems. The course has seven objectives measured by the following outcomes:

1. Understand the value of Data Mining and its use as a strategic resource,
2. Approach business problems with an analytical mind,

5 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/teaching-a-data-mining-course-to-mba-students/107428

Related Content

Pattern Retrieval through Classification from Pattern Warehouse: Issues and Challenges

Ramjeevan Singh Thakur and Vivek Tiwari (2014). *International Journal of Business Intelligence Research* (pp. 1-10).

www.irma-international.org/article/pattern-retrieval-through-classification-from-pattern-warehouse/122448

Peer-to-Peer Lending as an Alternative Investment Asset Class in the EU

Manuel Gutierrez (2018). *Disruptive Technologies for Business Development and Strategic Advantage* (pp. 56-98).

www.irma-international.org/chapter/peer-to-peer-lending-as-an-alternative-investment-asset-class-in-the-eu/206831

Agile Methodologies for Business Intelligence

Deanne Larson (2012). *Business Intelligence and Agile Methodologies for Knowledge-Based Organizations: Cross-Disciplinary Applications* (pp. 101-119).

www.irma-international.org/chapter/agile-methodologies-business-intelligence/58567

Inventory Models for Deteriorating Items

Vinod Kumar Mishra (2014). *Encyclopedia of Business Analytics and Optimization* (pp. 1224-1233).

www.irma-international.org/chapter/inventory-models-for-deteriorating-items/107321

Integration Platform for De-Centralized Investment Projects Appraisal

M. Stanojevic (2007). *Adaptive Technologies and Business Integration: Social, Managerial and Organizational Dimensions* (pp. 329-349).

www.irma-international.org/chapter/integration-platform-centralized-investment-projects/4242