Chapter 6
A Perspective on Data Mining Integration with Business Intelligence

Ana Azevedo
CEISE/STI, ISCAP/IPP, Portugal

Manuel Filipe Santos
University of Minho, Portugal

ABSTRACT

Business Intelligence (BI) is an emergent area of the Decision Support Systems (DSS) discipline. Over the past years, the evolution in this area has been considerable. Similarly, in the last years, there has been a huge growth and consolidation of the Data Mining (DM) field. DM is being used with success in BI systems, but a truly DM integration with BI is lacking. The purpose of this chapter is to discuss the relevance of DM integration with BI, and its importance to business users. From the literature review, it was observed that the definition of an underlying structure for BI is missing, and therefore a framework is presented. It was also observed that some efforts are being done that seek the establishment of standards in the DM field, both by academics and by people in the industry. Supported by those findings, this chapter introduces an architecture that can conduct to an effective usage of DM in BI. This architecture includes a DM language that is iterative and interactive in nature. This chapter suggests that the effective usage of DM in BI can be achieved by making DM models accessible to business users, through the use of the presented DM language.

INTRODUCTION

Business Intelligence (BI) can be presented as an architecture, tool, technology or system that gathers and stores data, analyzes it using analytical tools, and delivers information and/or knowledge, facilitating reporting, querying, and, ultimately, allows organizations to improve decision making (Clark, Jones, & Armstrong, 2007; Kudyba & Hoptroff, 2001; Michalewicz, Schmidt, Michalewicz, & Chiriac, 2007; Moss & Shaku, 2003; Negash, 2004; Raisinghani, 2004; Thierauf, 2001; Turban, Sharda, Aroson, & King, 2008). To put it shortly, Business Intelligence can be defined as the process that transforms data into information and then into knowledge (Golfarelli, Rizzi, & Cella, 2004).

DOI: 10.4018/978-1-60960-067-9.ch006

Being rooted in the Decision Support Systems (DSS) discipline, BI has suffered a considerable evolution over the last years and is, nowadays, an area of DSS that attracts a great deal of interest from both the industry and researchers (Arnott & Pervan, 2008; Clark et al., 2007; Hannula & Pirttimäki, 2003; Hoffman, 2009; Negash, 2004; Richardson, Schlegel, Hostmann, & McMurchy, 2008; Richardson, Schlegel, & Hostmann, 2009).

Data Mining (DM) is being applied with success in BI and several examples of applications can be found (Linoff, 2008; Turban et al., 2008; Vercellis, 2009). Despite that, DM has not yet reached to non specialized users. The authors consider that the real issue is related with the fact that the Knowledge Discovery in Databases (KDD) process, as presented by Fayyad et al. (1996), is not fully integrated in BI. Consequently, its full potential could be not completely explored by decision makers using the systems. Currently, DM systems are functioning as separate isles, but the authors consider that only the full integration of the KDD process on BI can conduct to an effective usage of DM in BI.

The authors have point out three main reasons for DM to be not completely integrated with BI. Firstly, the models/patterns obtained from DM are complex and there is the need of an analysis from a DM specialist. This fact can lead to a non-effective adoption of DM in BI, being that DM is not really integrated on most of the implemented BI systems, nowadays. Secondly, the problem with DM is that there is not a user-friendly tool that can be used by decision makers to analyze DM models. Usually, BI systems have user-friendly analytical tools that help decision makers in order to obtain insights on the available data and allow them to take better decisions. Examples of such tools are On-Line Analytical Processing (OLAP) tools, which are widely used (Negash, 2004; Turban et al., 2008). Powerful analytical tools, such as DM, remain too complex and sophisticated for the average consumer. Finally, but extremely important, it has not been given sufficient emphasis to the development of solutions that allow the specification of DM problems through business oriented languages, and that are also oriented for BI activities. With the expansion that has occurred in the application of DM solutions in BI, this is, currently, of increasing importance.

Most of the BI systems are built on top of relational databases. As a consequence, DM integration with relational databases is an important issue to consider when studying DM integration with BI. Codd´s relational model for database systems is long ago adopted in organizations. One of the reasons for the great success of relational databases is related with the existence of a standard language – SQL (Structured Query Language). SQL allows business users to obtain quick answers to ad-hoc questions, through queries on the data stored in databases. SQL is nowadays included in all the Relational Database Management Systems (RDBMS). SQL serves as the core above which are constructed the various Graphical User Interfaces (GUI's) and user friendly languages, such as Query By Example (QBE's), included in RDBMS. It is also necessary to define a standard language for data mining, which can operate likewise for data mining. Some efforts are being made in order to overcome this issue. Efforts involve the definition of standards for DM that arises both by academics and by people in the industry. It is the authors' belief that the effective integration of DM with BI systems must involve final business users' access to DM models. This access is crucial in order to business users to develop an understanding of the models, to help them in decision making. With this in mind, the authors present a high-level architecture that pretends to conduct to an effective usage of DM with BI. This architecture includes a DM language, named as Query Models By Example (QMBE), which is iterative and interactive in nature, thus allowing final business users to access and manipulate DM models.

Related Content

Creation and Integration of Reference Ontologies for Efficient LOD Management

Mariana Damova, Atanas Kiryakov, Maurice Grinberg, Michael K. Bergman, Frédérick Giassonand Kiril Simov (2012). *Semi-Automatic Ontology Development: Processes and Resources (pp. 162-199).*

www.irma-international.org/chapter/creation-integration-reference-ontologies-efficient/63902

Changing the Grant Culture of a College

James E. McLeanand Alanna Rochelle King Dail (2012). *Cases on Institutional Research Systems (pp. 22-38).*

www.irma-international.org/chapter/changing-grant-culture-college/60838

Cross-language Ontology Learning

Hans Hjelmand Martin Volk (2011). *Ontology Learning and Knowledge Discovery Using the Web: Challenges and Recent Advances (pp. 272-297).*

www.irma-international.org/chapter/cross-language-ontology-learning/53891

Fuzzy Miner: Extracting Fuzzy Rules from Numerical Patterns

Nikos Pelekis, Babis Theodoulidis, Ioannis Kopanakisand Yannis Theodoridis (2008). *Mathematical Methods for Knowledge Discovery and Data Mining (pp. 299-321).*

www.irma-international.org/chapter/fuzzy-miner-extracting-fuzzy-rules/26147

Learning SKOS Relations for Terminological Ontologies from Text

Wei Wang, Payam M. Barnaghiand Andrzej Bargiela (2011). *Ontology Learning and Knowledge Discovery Using the Web: Challenges and Recent Advances (pp. 129-152).*

www.irma-international.org/chapter/learning-skos-relations-terminological-ontologies/53884