



Chapter XI

Transforming Textual Patterns into Knowledge

Hércules Antonio do Prado, Brazilian Enterprise for Agricultural
Research, Brazil and Catholic University of Brasília, Brazil

José Palazzo Moreira de Oliveira, Federal University of
Rio Grande do Sul, Brazil

Edilson Fernela, Catholic University of Brasília, Brazil

Leandro Krug Wives, Federal University of Rio Grande do Sul,
Brazil

Edilberto Magalhães Silva, Brazilian Public News Agency, Brazil

Stanley Loh, Catholic University of Pelotas and
Lutheran University of Brazil, Brazil

ABSTRACT

Business Intelligence (BI) can benefit greatly from the bulk of knowledge that stays hidden in the large amount of textual information existing in the organizational environment. Text Mining (TM) is a technology that provides the support to extract patterns from texts. After interpreting these patterns, a business analyst can reach useful insights to improve the organizational knowledge. Although text represents the largest part of the available information in a company, just a small part of all Knowledge Discovery (KD) applications are in TM. By means of a case study, this

chapter shows an alternative to how TM can contribute to BI. Also, a discussion on future trends and some conclusions are presented that support the effectiveness of TM as source of relevant knowledge.

INTRODUCTION

In order to help companies to remain competitive, Business Intelligence (BI) requires adequate tools to transform the large amount of information existing in the organizational environment into models that accurately represent the reality. By analyzing and interpreting these patterns, a business analyst can reach interesting insights that lead to a better understanding of the business, to improve internal processes, and to enable more accurate decision making. In this environment, Knowledge Discovery in Databases (KDD) has been shown to be an important approach, providing the methodologies to extract models from large data sets by means of relatively user-friendly tools.

This chapter discusses the application of Text Mining (TM) in the BI realm. TM is part of the broader field of KDD that, departing from a data set relevant to the solution of a specific problem, looks for interesting and previously unknown patterns. After human-specialized interpretation of these patterns, important knowledge that leads to the solution of that problem can be discovered. KDD comprises the preliminary tasks of data collection, cleansing, and transformation, the core task of data mining, and the post-processing pattern interpretation activity. Data Mining (DM) is concerned with the extraction of patterns from structured data. In our point of view, TM is situated at the same level of DM, but is addressed to extract patterns from textual information. The patterns issued by a TM application can provide feedback to the business specialist, triggering the mental process that leads to insights on the business problems.

The next section draws the relations between TM and BI, making clear at which point TM can interfere proactively in the BI process. It also discusses the weak use of TM in an environment in which the majority of the available information is represented in textual form. Next, the methodology applied to guide the case study is presented. After that, a description of some procedures applied to identify clusters in a set of texts is shown. The subsequent section is devoted to discuss the case study carried out in a Brazilian public news agency. Future trends on TM and some conclusions are presented in the last two sections of the chapter.

19 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/transforming-textual-patterns-into-knowledge/6072

Related Content

Chinese Cultural Values and Knowledge Sharing

Alan K.M. Au and Matthew C.H. Yeung (2016). *Business Intelligence: Concepts, Methodologies, Tools, and Applications* (pp. 1482-1490).

www.irma-international.org/chapter/chinese-cultural-values-and-knowledge-sharing/142684

Information Retrieval (IR) and Extracting Associative Rules

Asmae Dami, Mohamed Fakir and Belaid Bouikhalene (2016). *Business Intelligence: Concepts, Methodologies, Tools, and Applications* (pp. 713-732).

www.irma-international.org/chapter/information-retrieval-ir-and-extracting-associative-rules/142647

Some New Alternative Formulations of Adaptive Kalman Filter for Market Risk Beta Estimation

Atanu Das (2021). *International Journal of Business Analytics* (pp. 17-37).

www.irma-international.org/article/some-new-alternative-formulations-of-adaptive-kalman-filter-for-market-risk-beta-estimation/276444

The Performance Mining Method: Extracting Performance Knowledge from Software Operation Data

Stella Pachidi and Marco Spruit (2015). *International Journal of Business Intelligence Research* (pp. 11-29).

www.irma-international.org/article/the-performance-mining-method/132821

Business Intelligence in the Bayou: Recovering Costs in the Wake of Hurricane Katrina

Gregory Smith, Thilini Ariyachandra and Mark Frolick (2010). *International Journal of Business Intelligence Research* (pp. 21-29).

www.irma-international.org/article/business-intelligence-bayou/43679