Chapter 6
Interaction Mining:
Making Business Sense of Customers Conversations through Semantic and Pragmatic Analysis

Vincenzo Pallotta
Webster University, Switzerland

Lammert Vrieling
Webster University, Switzerland

Rodolfo Delmonte
Università "Ca Foscari," Italy

ABSTRACT

In this chapter we present the major challenges of a new trend in business analytics, namely Interaction Mining. With the proliferation of unstructured data as the result of people interacting with each other using digital networked devices, classical methods in text business analytics are no longer effective. We identified the causes of their failure as being related to the inadequacy of dealing with conversational data. We propose then to move from Text Mining towards Interaction Mining, and we make several cases for this transition in areas such as marketing research, social media analytics, and customer relationship management. We also propose a roadmap for the future development of Interaction Mining by challenging the current practices in business intelligence and information visualization.

1 INTRODUCTION

Via the Web a wealth of information for business research is ready at our fingertips. Analyzing this—unstructured—information, however, can be very difficult. *Analytics* has become the business buzzword distinguishing traditional competitors from 'analytics competitors' who have dramatically boosted their revenues. The latter competitors distinguish themselves through "expert use of statistics and modeling to improve a wide variety of functions" (Davenport, 2006, p. 105). However, not all information lends itself to statistics and

DOI: 10.4018/978-1-61350-038-5.ch006

models. Actually, most information on the Web is made for, and by, people communicating through 'rich' language. This richness of our language is typically missed or not adequately accounted for in (statistical) analytics (e.g. Text-mining)—and so is its real meaning—because it is hidden in *semantics* rather than form (e.g. syntax). In our efforts of turning unstructured data into structured data, important information—and our ability to distinguish ourselves from competitors—gets lost.

Search engines (Büttcher, Clarke, & Cormack, 2010) have exploited statistical (frequency-based, TF-IDF[2]) methods to its extreme, but building indexes of Web content with keywords is not enough for understanding beyond keyword-based search. The use of semantics in search would be a great improvement and new generation search engines (Grimes, 2010) are starting to address this. Semantic search can be approached from several perspectives. The most common one is to go beyond word forms and consider concepts with their semantic relationships. Concepts can be extracted implicitly or explicitly. In the first case, the contexts of words in a document base determine the concept (Landauer, Foltz, & Laham, 1998). In the second case, concepts are assigned to word forms through a semantic lexicon or ontology such as WordNet[3].

However, semantics is not only necessary for search, it is necessary for any processing of information from the Web. After all, semantics simply means "making sense" and we would like to argue that sophisticated semantic analysis of content is a necessary tool for quality business research of Web data.

For instance, any business analyst in *fast moving consumer goods* (FMCG) is looking for so much more than just text when analyzing online focus group interview data. A FMCG analyst would analyze interaction, which would reveal shared language, beliefs and myths, argumentative reasoning, justifications, and changes of opinion or (re)interpretation of experiences (Catterall & Maclaran, 1997).

Focus group interview data is both qualitative and interactive. It is not just text; it is conversational data as people are responding to one another. As such traditional (manual) analysis of focus group data is labor intensive, complex, analyst dependent, inconsistent and subjective[4].

The key problem is that good analysis of unstructured data is costly, complex and time-consuming

However, the power of current state-of-the-art NLU[5] systems makes automated analysis of these—and other—types of unstructured data feasible and possible (Delmonte, Bristot, & Pallotta, 2010).

This chapter will put to rest the myth that computers cannot extract rich information from unstructured data[6] even from conversational data. We will put forward a new generation of "Interaction Mining" technology that is analyst independent, consistent regardless of the quantity of data, 'Machine-like' precision in its analysis in multiple languages and—compared to manual analysis—a quantum leap faster.

First, we will conduct first a survey of current technology for Interaction Mining by assessing the strengths, weaknesses and limits of current approaches such as Text Mining. Additionally, we will present a study in eliciting business requirements for Interaction Mining in different domains. We will present a new approach, which exploits information extracted from automatic analysis of conversational data, which solves some of the challenges highlighted in the requirements section. We will conclude the chapter by outlining a roadmap for research in Interaction Mining.

2 INTERACTION MINING

In this section we review some current approaches to Business Analytics such as Data Mining and Text Mining by assessing their benefit, strengths, weaknesses and limits to deal with conversational data. We then propose a novel paradigm that we