Extracting Opinions and Facts for Business Intelligence

Horacio Saggion*, Adam Funk*

*Department of Computer Science
University of Sheffield
Regent Court
211 Portobello Street
Sheffield - S1 5DP
{H.Saggion,A.Funk}@dcs.shef.ac.uk

Abstract. Finding information about companies on multiple sources on the Web has become increasingly important for business analysts. In particular, since the emergence of the Web 2.0, opinions about companies and their services or products need to be found and distilled in order to create an accurate picture of a business entity. Without appropriate text mining tools, company analysts would have to read hundreds of textual reports, newspaper articles, forums' postings and manually dig out factual as well as subjective information. This paper describes a series of experiments to assess the value of a number of lexical, morpho-syntactic, and sentiment-based features derived from linguistic processing and from an existing lexical database for the classification of evaluative texts. The paper describes experiments carried out with two different web sources: one source contains positive and negative opinions while the other contains fine grain classifications in a 5-point qualitative scale. The results obtain are positive and in line with current research in the area. Our aim is to use the result of classification in a practical application that will combine factual and opinionated information in order to create the reputation of a business entity.

1   Introduction

It is accepted that incidents which damage a company's reputation for honesty or safety may cause serious damage to finances. For example, in 1999 Coca Cola lost $60 million (by its own estimate) after schoolchildren reported suffering from symptoms like headaches, nausea and shivering after drinking its products. The reputation of a company not only depends on the quality of its products but also on the behaviours of their employees and in particular their managers. Many businesses have public relations departments dedicated to managing their reputation. The public relations industry is growing due to the demand for companies to build corporate credibility and hence reputation. One important component of the reputation of a company is the opinion that stake-holders have about it, its products, and its services. The Web has become an increasingly important source of information in all areas of society in particular, in the field of business intelligence, business analysts are turning their eyes on the web in order to obtain factual as well as more subtle and subjective information (opinions) on companies.