

Chapter 7

OpAL: A System for Mining Opinion from Text for Business Applications

Alexandra Balahur
University of Alicante, Spain

Ester Boldrini
University of Alicante, Spain

Andrés Montoyo
University of Alicante, Spain

Patricio Martínez-Barco
University of Alicante, Spain

ABSTRACT

The past years have marked the birth and development of the Social Web, where people freely express and search for opinions on all possible topics. This phenomenon has been proven to have a great impact on many business sectors globally. Given the proven importance of the subjective data on the Web, but bearing in mind the difficulties inherent to their textual peculiarities and large volume, efficient techniques must be employed to process this data, so that it can be fully exploited to the benefit of potential users and companies. We present the OpAL system, which implements an efficient approach to mine, classify and statistically summarize opinions, grounded on the feature-based Opinion Mining paradigm. In this approach, all components are studied, implemented and optimized using different NLP techniques. Results of different in-house and competition evaluations show that the system components have a good performance and that the techniques considered are efficient. We finally complete the proposed approach by presenting a method for opinion retrieval, which is robust and multilingual. Thus, we offer an integrated solution to build a system that is able to fully respond to user needs, from the querying to the summarized output stage. Implemented at a large scale, such systems can benefit the business environment and its customers everywhere.

DOI: 10.4018/978-1-61350-038-5.ch007

INTRODUCTION

Humans are social beings. They cannot reach the level of what we call “human” unless they develop in organized societies, where they are taught norms, rules and laws governing the existence and co-existence of people. Although most of the times unconsciously, we continuously shape our behavior and attitudes on the basis of these social conventions, of public and private opinions and events of the world surrounding us. We give and accept advice as part of our every-day lives, as part of a ritual to knowing, better understanding and integrating into our surrounding reality.

In a globalized world, however, the whole idea of context changes. Supported by the fast development of the Internet and the Web 2.0 technologies, with the predominant presence of social networks, forums, “blogging” and reviewing as world-wide phenomena, giving and receiving advice has become a global phenomenon. One that we give into more and more every day, as decisions to buy products or contract services, for example, are nowadays preceded by an internet search for opinions in many of the cases [Pang and Lee, 2008]. People express and search for such opinions on blogs, forums, in reviews and comments - a phenomenon which led to the creation of extensive quantities of subjective data that cannot be manually processed, although the knowledge they enclose is crucial to the business and social environments.

At the economic level, the globalization of markets combined with the fact that people can freely express their opinion on any product or company on forums, blogs or e-commerce sites led to a change in the companies’ marketing strategies, in the rise of awareness for client needs and complaints, and a special attention for brand trust and reputation. Specialists in market analysis, but also IT fields such as Natural Language Processing, demonstrated that in the context of the newly created opinion phenomena, decisions for economic action are not only given by factual information, but are highly affected by rumors

and negative opinions. Studies showed that financial information presented in news articles have a high correlation to social phenomena, on which opinions are expressed in blogs, forums or reviews. Investigations carried out in market analysis, as for example the Technorati survey series¹, assesses that opinions found in blogs and news correlate with the subsequent changes in sales. This is also exemplified by Mishne and Glance(2005), where the authors demonstrate that references to movies in blogs correlate well with the previous and subsequent success rate of a movie. Koppel and Shtrimberg (2004) investigate the influence of news of positive and negative polarity on rises and falls in stock market. Lexical-based opinion analysis models have shown up to 70% accuracy in predicting the corresponding actual price change and other research [Devitt and Ahmad, 2007] suggests that markets react to the same extent to affect-related parts of text as to the informative sections.

On the other hand, many tasks that involved extensive efforts from the companies’ marketing departments are easier to perform. An example is related to market research for business intelligence and competitive vigilance. New forms of expression on the web made it easier to collect information of interest, which can help to detect changes in the market attitude, discover new technologies, machines, markets where products are needed and detect threats. Moreover, using the opinion information, companies can spot the market segments their products are best associated with and can enhance their knowledge on the clients they are addressing and on competitors. The analysis of the data flow on the web can lead to the spotting of differences between the companies’ products and the necessities expressed by clients and between the companies’ capacities and those of the competitors. Last, but not least, the interpretation of the large amounts of data and their associated opinions can give companies the capacity to support decision through the detection of new ideas and new solutions to their technological or economic problems. The advantage and, at the same

29 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/opal-system-mining-opinion-text/58415

Related Content

Measuring the Effect of EWOM Readability and Sentiment on Sales: Online Cellphone Reviews

Aakash Aakashand Anu Gupta Aggarwal (2020). *International Journal of Business Analytics* (pp. 24-42).

www.irma-international.org/article/measuring-the-effect-of-ewom-readability-and-sentiment-on-sales/264261

AI for IoT Application: An Systematic Review

Ganesh Shivaji Piseand Sachin D. Babar (2023). *Handbook of Research on AI and Knowledge Engineering for Real-Time Business Intelligence* (pp. 131-142).

www.irma-international.org/chapter/ai-for-iot-application/321490

Credit Rating Classification Using Self-Organizing Maps

Roger P.G.H. Tan, Jan van den Bergand Willem-Max van den Bergh (2002). *Neural Networks in Business: Techniques and Applications* (pp. 140-153).

www.irma-international.org/chapter/credit-rating-classification-using-self/27264

Voice Engagement Leading to Business Intelligence: A Systematic Review and Agenda for Future Research

Praveen Kumar Sattarapu, Deepti Waderaand Jaspreet Kaur (2021). *International Journal of Business Intelligence Research* (pp. 1-23).

www.irma-international.org/article/voice-engagement-leading-to-business-intelligence/294568

Biologically Inspired Techniques for Data Mining: A Brief Overview of Particle Swarm Optimization for KDD

Shafiq Alam, Gillian Dobbie, Yun Sing Kohand Saeed ur Rehman (2016). *Business Intelligence: Concepts, Methodologies, Tools, and Applications* (pp. 2275-2284).

www.irma-international.org/chapter/biologically-inspired-techniques-for-data-mining/142727