Potential Power and Problems in Sentiment Mining of Social Media

John Wang, Department of Information & Operations Management, Montclair State University, Montclair, NJ, USA

Qiannong Gu, Department of Information Systems and Operations Management, Miller College of Business, Ball State University, Muncie, IN, USA

Gang Wang, Department of Global Management Studies, Nathan Weiss Graduate College, Kean University, Union, NJ, USA

ABSTRACT

Sentiment mining research has experienced an explosive growth in awareness and demand as Web 2.0 technologies have paved the way for a surge of social media platforms that have significantly and rapidly increased the availability of user-generated opinioned text. The power of opinions has long been known and is beginning to be tapped to a fuller potential through sentiment mining research. Social media sites have become a paradise for sentiment providing endless streams of opinioned text encompassing an infinite array of topics. With the potential to predict outcomes with a relative degree of accuracy, sentiment mining has become a hot topic not only to researchers, but to corporations as well. As the social media user base continues to expand and as researchers compete to fulfill the demand for sentiment analytic tools to sift through the endless stream of user-generated content, the growth of sentiment mining of social media will continue well into the future with an emphasis on improved reliability, accuracy, and automation.

Keywords: Ontology-based Analysis, Opinion Mining, Opinion Spamming, Sentiment Mining, Social Media, Web 2.0 Technologies

INTRODUCTION

Sentiment mining (SM), also called opinion mining or sentiment analysis, has evolved over the last decade from text mining and natural language processing, but aims to determine the attitudes of individuals/groups with respect to some specific topics. More recently, SM has greatly assisted decision makers in extracting opinions from unstructured human-authored documents. SM is a computational process which is profoundly rooted in the field of computer science. SM utilizes analytical tools including natural language processing, computational linguistics, and text analysis to isolate and extract subjective information from factual information from a source of text (Pang & Lee,
2008; Anderson, 2011). Once extracted, SM seeks to analyze and determine the opinions, feelings, emotions, attitudes, and sentiments embedded in the text by aggregating and categorizing text as positive, negative, or neutral.

Montoyo, Martínez-Barco, and Balahur (2012) consider SM to deal with the detection of “private states” (opinions, emotions, sentiments, beliefs, speculations). SM can detect, extract and classify opinions and sentiments concerning different topics, as expressed in textual input. The process of SM is significant as it provides a basis for the determination of distinctive insights into the thoughts and opinions of individuals based on written text. This information is of tremendous value to a diverse population of researchers particularly those concentrated in the marketing spectrum. The demand for information regarding the opinions and sentiments of individuals and groups of individuals has grown tremendously. Opinions are a crucial aspect in the decision making process both on a personal and business level (Liu, 2010; Gopal, Marsden, & Vanthienen, 2011). Opinions ultimately influence our decisions and behaviors which ultimately define an outcome. It is this influential relationship as well as the outcome(s) of the opinions and behaviors that have heightened the interest of SM to decision making entities including individuals, researchers, corporations, politicians, and even governments.

Individuals’ opinions are not only figuratively valuable, but also literally valuable. Corporations spend millions of dollars each year to gain insights into customers’ thoughts, attitudes, and opinions. Traditionally, researchers have used conventional methods such as opinion polls or questionnaires amassed of intricately crafted questions designed to produce heavily opinioned responses to gauge the sentiment of a population. While these methods may be effective in determining the sentiment of the given population, the information obtained more than likely does not represent an accurate reading of true sentiment due to population size limitations and questionnaire bias (Wang et al., 2012).

While the concept and value of SM has existed for some time, its potential has been severely stifled by the absence and availability of opinionated text. However, with the relatively recent advent of Web 2.0 and its resulting social media revolution, many have reawakened to the potential and power of SM. Blogs, Micro-Blogs, Forums, wikis, social networking sites, namely Facebook and Twitter, and user review/comment portals have significantly accelerated the production of user generated content immensely increasing the volume and availability of opinioned text which is both free and easily accessible. With such a tremendous surge of the publication of user generated opinioned text, the World Wide Web has become a valuable trove of opinion laden information waiting to be mined.

With such an explosion of social media, current SM research is almost exclusively focused on social media as user generated content from social media has provided the fuel for SM fire. However, although Web 2.0 has negated the lack of opinionated text which had hindered the advancement of SM, there will continue to be inheriting challenges in the field of SM as the complexities of natural language have not yet been overcome by artificial intelligence making SM an imperfect science. However, while SM may remain an imperfect science for the time being, it is by no means a futile science. Current SM research has shown to be between 70 and 80% accurate (Cai et al., 2010; Bai, 2011). This statistic may spawn a community of skeptics, it is crucial to realize the field of SM is in its infancy and as additional research is conducted and the field continues to generate interest, advancements in SM tools and techniques are taking shape leading to more accurate results and to the development of more complex mining tools and techniques that have the potential to forever change how research is
Related Content

A Decision Model on Corporate Social Responsibility and Business Strategies
[www.igi-global.com/chapter/a-decision-model-on-corporate-social-responsibility-and-business-strategies/176789?camid=4v1a](www.igi-global.com/chapter/a-decision-model-on-corporate-social-responsibility-and-business-strategies/176789?camid=4v1a)

The Evaluation of Library Services Methods: Cost Per Use and Users’ Satisfaction
[www.igi-global.com/chapter/evaluation-library-services-methods/66727?camid=4v1a](www.igi-global.com/chapter/evaluation-library-services-methods/66727?camid=4v1a)
Investment Climate Factors with Reference to Firm Performance in Bangladesh: A Prospective Cohort Study

Application of Conjoint Analysis in Improving the Value of New Product Development: A Hotel Case Study Analysis