Expert Systems with Applications 36 (2009) 10760–10773

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

A survey on sentiment detection of reviews

Huifeng Tang, Songbo Tan *, Xueqi Cheng

Information Security Center, Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100080, PR China

ARTICLE INFO

Keywords:
Sentiment detection
Opinion extraction
Sentiment classification

ABSTRACT

The sentiment detection of texts has been witnessed a booming interest in recent years, due to the increased availability of online reviews in digital form and the ensuing need to organize them. Till to now, there are mainly four different problems predominating in this research community, namely, subjectivity classification, word sentiment classification, document sentiment classification and opinion extraction. In fact, there are inherent relations between them. Subjectivity classification can prevent the sentiment classifier from considering irrelevant or even potentially misleading text. Document sentiment classification and opinion extraction have often involved word sentiment classification techniques. This survey discusses related issues and main approaches to these problems.

1. Introduction

Today, very large amount of reviews are available on the web, as well as the weblogs are fast-growing in blogsphere. Product reviews exist in a variety of forms on the web: sites dedicated to a specific type of product (such as *digital camera*), sites for newspapers and magazines that may feature reviews (like *Rolling Stone* or *Consumer Reports*), sites that couple reviews with commerce (like *Amazon*), and sites that specialize in collecting professional or user reviews in a variety of areas (like *Rottentomates.com*). Less formal reviews are available on discussion boards and mailing list archives, as well as in Usenet via Google Groups. Users also comment on products in their personal web sites and blogs, which are then aggregated by sites such as *Blogstreet.com*, *AllConsuming.net*, and *onfocus.com*.

The information mentioned above is a rich and useful source for marketing intelligence, social psychologists, and others interested in extracting and mining opinions, views, moods, and attitudes. For example, whether a product review is positive or negative; what are the moods among Bloggers at that time; how the public reflect towards this political affair, etc.

To achieve this goal, a core and essential job is to detect subjective information contained in texts, include viewpoint, fancy, attitude, sensibility etc. This is so-called *sentiment detection*.

A challenging aspect of this task seems to distinguish it from traditional topic-based detection (classification) is that while topics are often identifiable by keywords alone, sentiment can be expressed in a much subtle manner. For example, the sentence "What a bad picture quality that digital camera has! ... Oh, this new type camera has a good picture, long battery life and beautiful appearance!" compares a negative experience of one product with a positive experience of another product. It is difficult to separate out the core assessment that should actually be correlated with the document. Thus, sentiment seems to require more understanding than the usual topic-based classification.

Sentiment detection dates back to the late 1990s (Argamon, Koppel, & Avneri, 1998; Kessler, Nunberg, & SchÄutze, 1997; Spertus, 1997), but only in the early 2000s did it become a major subfield of the information management discipline (Chaovalit & Zhou, 2005; Dimitrova, Finn, Kushmerick, & Smyth, 2002; Durbin, Neal Richter, & Warner, 2003; Efron, 2004; Gamon, 2004; Glance, Hurst, & Tomokiyo, 2004; Grefenstette, Qu, Shanahan, & Evans, 2004; Hillard, Ostendorf, & Shriberg, 2003; Inkpen, Feiguina, & Hirst, 2004; Kobayashi, Inui, & Inui, 2001; Liu, Lieberman, & Selker, 2003; Raubern & Muller-Kogler, 2001; Riloff and Wiebe, 2003; Subasic & Huettner, 2001; Tong, 2001; Vegnaduzzo, 2004; Wiebe & Riloff, 2005; Wilson, Wiebe, & Hoffmann, 2005). Until the early 2000s, the two main popular approaches to sentiment detection, especially in the real-world applications, were based on machine learning techniques and based on semantic analysis techniques. After that, the shallow nature language processing techniques were widely used in this area, especially in the document sentiment detection. Current-day sentiment detection is thus a discipline at the crossroads of NLP and IR, and as such it shares a number of characteristics with other tasks such as information extraction and text-mining.

Although several international conferences have devoted special issues to this topic, such as ACL, AAAI, WWW, EMNLP, CIKM etc., there are no systematic treatments of the subject: there are neither textbooks nor journals entirely devoted to sentiment detection yet.

* Corresponding author.
E-mail addresses: tanghuifeng@software.ict.ac.cn (H. Tang), tansongbo@software.ict.ac.cn (S. Tan), cxq@ict.ac.cn (X. Cheng).

0957-4174/$ - see front matter © 2009 Published by Elsevier Ltd.
doi:10.1016/j.eswa.2009.02.063

This paper first introduces the definitions of several problems that pertain to sentiment detection. Then we present some applications of sentiment detection. Section 4 discusses the subjectivity classification problem. Section 5 introduces semantic orientation method. The sixth section examines the effectiveness of applying machine learning techniques to document sentiment classification. The seventh section discusses opinion extraction problem. The eighth part talks about evaluation of sentiment detection. Last section concludes with challenges and discussion of future work.

2. Sentiment detection

2.1. Subjectivity classification

Subjectivity in natural language refers to aspects of language used to express opinions and evaluations (Wiebe, 1994). Subjectivity classification is stated as follows: Let $S = \{s_1, \ldots, s_n\}$ be a set of sentences in document D. The problem of subjectivity classification is to distinguish sentences used to present opinions and other forms of subjectivity (subjective sentences set S_s) from sentences used to objectively present factual information (objective sentences set S_o), where $S_s \cup S_o = S$. This task is especially relevant for news reporting and Internet forums, in which opinions of various agents are expressed.

2.2. Sentiment classification

Sentiment classification includes two kinds of classification forms, i.e., binary sentiment classification and multi-class sentiment classification. Given a document set $D = \{d_1, \ldots, d_n\}$, and a pre-defined categories set $C = \{\text{positive}, \text{negative}\}$, binary sentiment classification is to classify each d_i in D, with a label expressed in C. If we set $C^* = \{\text{strong positive}, \text{positive}, \text{neutral}, \text{negative}, \text{strong negative}\}$ and classify each d_i in D with a label in C^*, the problem changes to multi-class sentiment classification.

Most prior work on learning to identify sentiment has focused on the binary distinction of positive vs. negative. But it is often helpful to have more information than this binary distinction provides, especially if one is ranking items by recommendation or comparing several reviewers' opinions. Koppel and Schler (2005a, 2005b) show that it is crucial to use neutral examples in learning polarity for a variety of reasons. Learning from negative and positive examples alone will not permit accurate classification of neutral examples. Moreover, the use of neutral training examples in learning facilitates better distinction between positive and negative examples.

3. Applications of sentiment detection

In this section, we will expound some rising applications of sentiment detection.

3.1. Products comparison

It is a common practice for online merchants to ask their customers to review the products that they have purchased. With more and more people using the Web to express opinions, the number of reviews that a product receives grows rapidly. Most of the researches about these reviews were focused on automatically classifying the products into "recommended" or "not recommended" (Pang, Lee, & Vaithyanathan, 2002; Ranjan Das & Chen, 2001; Terveen, Hill, Amento, McDonald, & Creter, 1997). But every product has several features, in which maybe only part of them people are interested. Moreover, a product has shortcomings in one aspect, probably has merits in another place (Morinaga, Yamanishi, Tateishi, & Fukushima, 2002; Taboada, Gillies, & McFetridge, 2006).

To analysis the online reviews and bring forward a visual manner to compare consumers' opinions of different products, i.e., merely with a single glance the user can clearly see the advantages and weaknesses of each product in the minds of consumers. For a potential customer, he/she can see a visual side-by-side and feature-by-feature comparison of consumer opinions on these products, which helps him/her to decide which product to buy. For a product manufacturer, the comparison enables it to easily gather marketing intelligence and product benchmarking information.

Liu, Hu, and Cheng (2005) proposed a novel framework for analyzing and comparing consumer opinions of competing products. A prototype system called *Opinion Observer* is implemented. To enable the visualization, two tasks were performed: (1) Identifying product features that customers have expressed their opinions on, based on language pattern mining techniques. Such features form the basis for the comparison. (2) For each feature, identifying whether the opinion from each reviewer is positive or negative, if any.

Different users can visualize and compare opinions of different products using a user interface. The user simply chooses the products that he/she wishes to compare and the system then retrieves the analyzed results of these products and displays them in the interface.

3.2. Opinion summarization

The number of online reviews that a product receives grows rapidly, especially for some popular products. Furthermore, many reviews are long and have only a few sentences containing opinions on the product. This makes it hard for a potential customer to read them to make an informed decision on whether to purchase the product. The large number of reviews also makes it hard for product manufacturers to keep track of customer opinions of their products because many merchant sites may sell their products, and the manufacturer may produce many kinds of products.

Opinion summarization (Ku, Lee, Wu, & Chen, 2005; Philip et al., 2004) summarizes opinions of articles by telling sentiment polarities, degree and the correlated events. With opinion summarization, a customer can easily see how the existing customers feel about a product, and the product manufacturer can get the reason why different stands people like it or what they complain about.

Hu and Liu (2004a, 2004b) conduct a work like that: Given a set of customer reviews of a particular product, the task involves three subtasks: (1) identifying features of the product that customers have expressed their opinions on (called product features); (2) for each feature, identifying review sentences that give positive or negative opinions; and (3) producing a summary using the discovered information.

Ku, Liang, and Chen (2006) investigated both news and web blog articles. In their research, TREC, NTCIR and articles collected from web blogs serve as the information sources for opinion extraction. Documents related to the issue of animal cloning are selected as the experimental materials. Algorithms for opinion extraction at word, sentence and document level are proposed. The issue of relevant sentence selection is discussed, and then topical and opinionated information are summarized. Opinion summarizations are visualized by representative sentences. Finally, an opinionated curve showing supportive and non-supportive degree along the timeline is illustrated by an opinion tracking system.

3.3. Opinion reason mining

In opinion analysis area, finding the polarity of opinions or aggregating and quantifying degree assessment of opinions