Survey on Sentiment Analysis and Aspect Extraction of Reviews

Kazuaki Hiraga (1510754)

School of Information Science, JAIST, `kazu.hiraga@jaist.ac.jp`

Extended Abstract

Online reviews are becoming popular and primary factor in a customer's decision to purchase products or afford services. Opinion mining, which is a technique to reveal how much people satisfy or dissatisfy a given product or service, has a great impact on not only manufactures and retailers but also customers. In spite of the fact that many people make use of reviews to obtain others' opinions, there are potential concerns that some people have created false reviews to artificially promote or demote products and services. Furthermore, some reviews that are irrelevant to products and services are intentionally posted to increase the number of reviews. Those reviews are called fake reviews or spam reviews, which are written by groups of people who are employed by manufactures or retailers. Obviously, it is useful for customers to automatically detect fake and spam reviews.

The current research of opinion mining mainly focuses on polarity classification that is able to identify whether reviewers have positive or negative opinions on products or services. There exists various targets to be classified into polarity classes: a document, sentence, phrase and aspect. Recently, sentiment analysis on an aspect of a product is paid much attention. There is a wide variety of studies on sentiment analysis. For example, many researchers have tried to explore effective features for machine learning in the task of aspect-based sentiment analysis. As for the polarity classification of an aspect, there exists several subtasks: extraction of aspect to be classified into its polarity class, development of sentiment lexicon and labeled corpus, and so on. As for the aspect extraction, many methods are proposed: employing part-of-speech tagger, syntactic parser, supervised/unsupervised machine learning and so on. Therefore, it is rather hard to get a clear perspective on the vast amount of studies about sentiment analysis.

Considering the above situation, the goal of this study is to explore a trend of studies that related to sentiment analysis. They are categorized into five categories: (1) polarity classification, (2) aspect extraction, (3) construction of sentiment lexicon, (4) detection of fake/spam reviews, and (5) identification of product-aspect relation. For each category, current research trends are investigated by comprehensive survey of the papers. It is meaningful to understand the progress of research to reveal current unsolved issues and what solutions we can propose in future.

This study is carried out as follows. Search engines and database of technical papers are used to retrieve papers related to polarity classification. Google

Scholar and SemanticScholar are used as the search engines, while ACL anthology and the proceedings of The Association for Natural Language Processing are used as the database. First, 308 papers are retrieved by entering several keywords about polarity classification. Among them, the papers are chosen when they can be categorized in one of the five groups and the number of citation to them is greater than 100. Since the number of papers in the category (1) "polarity classification" is quite large, the two most cited papers per year are chosen. Finally, 40 papers are chosen for precise survey.

The major outcome obtained by this study for each category is summarized as follows.

(1) It is found that the level of polarity classification has been changed from coarse to fine grained piece of a text, i.e. from a document-level, sentence-level, phrase-level to aspect-level. Regardless of the level of polarity classification, however, the features such as part-of-speech, word n-gram, results of syntactic parsing, polarity words and term frequency are used consistently. On the other hand, over the past a few years, novel approaches which utilize deep learning-based models are proposed, which intend not to use conventional features. Despite the fact that deep-learning based models are getting popular and many researchers are involved in developing new techniques, they have not significantly outperformed the traditional approaches, such as Bayesian classifier, Support Vector Machine etc., in the task of polarity classification. At this moment, it is unclear that the most researches on polarity classification are shifted to the deep learning methods.

(2) As for aspect extraction, many methods based on new topic models as well as rule based methods that rely on part-of-speech and syntactic dependency are proposed. In recent years, deep learning based methods are proposed, where the features such as part-of-speech are not explicitly given. They often outperform the conventional methods significantly in terms of F-measure, but both precision and recall are not improved by applying deep learning. Therefore, it is uncertain that the deep learning can be the major approach for aspect extraction.

(3) As for construction of a sentiment lexicon, several method based on bootstrapping are proposed. Given an initial lexicon consisting of a few sentiment words, new sentiment words are automatically extracted from corpora or other language resources, the polarity of them is identified, and finally the new sentiment words with their polarity are added to the lexicon. Part-of-speech and syntactic dependency are used in all papers in this survey. Since it is known that the polarity of words is dependent to domains, it is very useful to automatically construct domain dependent sentiment lexicon.

(4) Most of researches on spam detection define the task as binary classification where a given text is judged as a spam or not. However, it is rather difficult to detect a spam using only the information derived from a text. Therefore, several studies propose methods to use the information of reviewers' behavior such as pattern of reviewers' rating. Machine learning based methods using reviewers' behavior as features are also proposed. However, it is problematic that the information of reviewers' behavior is not always available.

(5) We try to find the research that aims at identifying whether the automatically extracted aspects are ones of a target product of which a user want to know a reputation. However, such a study was not found in this survey. One related work is Jindal and Lin's method. Although they try to identify whether a given document is related to a target product, but not for an aspect. A technique to judge if an aspect is related to a target is useful to filter out ineffectual comments. Therefore, we believe that it is one of the important future work.