Chapter 68
A Survey on Aspect Extraction Approaches for Sentiment Analysis

Vrps Sastry Yadavilli
National Institute of Technology, Tadepalligudem, India

Karthick Seshadri
https://orcid.org/0000-0002-5658-141X
National Institute of Technology, Tadepalligudem, India

ABSTRACT

Aspect-level sentiment analysis gives a detailed view of user opinions expressed towards each feature of a product. Aspect extraction is a challenging task in aspect-level sentiment analysis. Hence, several researchers worked on the problem of aspect extraction during the past decade. The authors begin this chapter with a brief introduction to aspect-level sentimental analysis, which covers the definition of key terms used in this chapter, and the authors also illustrate various subtasks of aspect-level sentiment analysis. The introductory section is followed by an explanation of the various feature learning methods like supervised, unsupervised, semi-supervised, etc. with a discussion regarding their merits and demerits. The authors compare the aspect extraction methods performance with respect to metrics and a detailed discussion on the merits and demerits of the approaches. They conclude the chapter with pointers to the unexplored problems in aspect-level sentiment analysis that may be beneficial to the researchers who wish to pursue work in this challenging and mature domain.

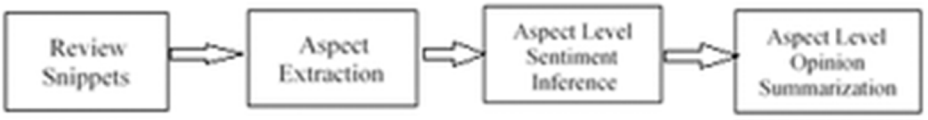
INTRODUCTION

With tremendous growth of world wide web and internet, many users are interested to post reviews about a product in social blogs, merchant web sites and social networking sites. Analyzing and identifying emotions and opinions in this review text can give better insights about a product or service to manufacturers or buyers. Sentiment analysis identifies hidden emotion or opinions in the review text.

DOI: 10.4018/978-1-6684-6303-1.ch068

Sentiment analysis could be applied to each review document (document level sentiment analysis), or to each review sentence (sentence level sentiment analysis) or to each feature phrase (aspect level sentiment analysis). Sentiment analysis at document level deals with mining hidden sentiment in a document about a product. Sentence Level Classification identifies polarity labels such as positive, negative, neutral in each sentence. Aspect level sentiment analysis (ALSA) performs inference of opinions at a fine-grained level i.e., it extracts opinions about a feature of a product. Aspect level sentiment analysis mainly deals with three phases as shown in figure 1. (i) Aspect Extraction, (ii) Aspect Level Sentiment Inference and (iii) Opinion Summarization of extracted aspects.

Figure 1. A generic model for ALSA



The aspect level sentiment inference refers to identification of opinion words e.g., adjectives, adverbs expressed towards each aspect. Aspect Level Opinion summarization is the process of summarizing opinions expressed towards each aspect as a statistical or textual summary. The following example outlines the phases of aspect level sentiment analysis. From the review snippet "This camera is cheap", extracting "price" is an aspect extraction task; inferring the opinion orientation of the aspect "price" as positive is referred to as the aspect level sentimental inference task; summarizing number of such opinions expressed towards the aspect "price" across all the review snippets is considered as the aspect level opinion summarization task. Out of the three tasks, aspect extraction is important and challenging since aspects can either be explicitly mentioned (explicit aspects) in the text or hidden in the text (implicit aspects). Aspect extraction can be thought of as an imbalanced classification problem because the distribution of product aspects over all sentences is not uniform, the users tend to specify few aspects very frequently, leaving some infrequent but popular aspects in review sentences. It is difficult to prepare enough training data to train a classifier to extract implicit aspects which requires extensive domain knowledge as they are implicitly stated in the text. So, extraction of such aspects can try to address the imbalanced data classification issue. Various approaches for aspect extraction are predominantly classified as follows.

- Unsupervised methods
- Supervised methods
- Semi-supervised methods
- Reinforcement Learning methods
- Soft Computing methods
- Hybrid methods

The unsupervised methods extract aspects based on manually curated rules or heuristics or using some statistical techniques. Hu et al. (2004) adopted association rule mining (Agarwal et al.,1994) to find frequent features in product reviews. Quan et.al. (2014) used a combination of PMI (Bouma et al.