**Aspect Ranking: Identifying Important Product Aspects from Online**

**Consumer Reviews**

Abstract

The important aspects are identified according to two observations: (a) the important aspects of a product are usually commented by a large number of consumers; and (b) consumers’ opinions on the important aspects greatly influence their overall opinions on the product. In particular, given consumer reviews of a product, we first identify the product aspects by a shallow dependency parser and determine consumers’ opinions on these aspects via a sentiment classifier. We then develop an aspect ranking algorithm to identify the important aspects by simultaneously considering the aspect frequency and the influence of consumers’ opinions given to each aspect on their overall opinions.

1 Introduction



Our assumption is that the important aspects of a product should be the aspects that are frequently commented by consumers, and consumers’ opinions on the important aspects greatly influence their overall opinions on the product. Given the online consumer reviews of a specific product, we first identify the aspects in the reviews using a shallow dependency parser (Wu et al., 2009), and determine consumers’ opinions on these aspects via a sentiment classifier. We then design an aspect ranking algorithm to identify the important aspects by simultaneously taking into account the aspect frequency and the influence of consumers’ opinions given to each aspect on their overall opinions.

The main contributions of this paper include,

1) We dedicate to the topic of aspect ranking, which aims to automatically identify important aspects of a product from consumer reviews.

2) We develop an aspect ranking algorithm to identify the important aspects by simultaneously considering the aspect frequency and the influence of consumers’ opinions given to each aspect on their overall opinions.

3) We apply aspect ranking results to the application of document-level sentiment classification, and improve the performance significantly.

2 Aspect Ranking Framework

2.1 Notations and Problem Formulation

We define *ork* as the opinion on aspect *ak* in review *r*.



Our task is to derive the important weights of aspects, and identify the important aspects.

2.2 Aspect Identification

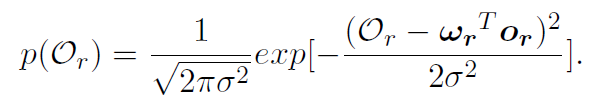
To identify the aspects in free text reviews, we first parse each review using the Stanford parser and extract the noun phrases (*NP*) from the parsing tree as aspect candidates. While the obtained aspects may contain some synonym terms, such as “*earphone*” and “*headphone*,” we further perform synonym clustering to get unique aspects. Specifically, we first expand each aspect term with its synonym terms obtained from the synonym terms Web site 2, and then cluster the terms to obtain unique aspects based on unigram feature.

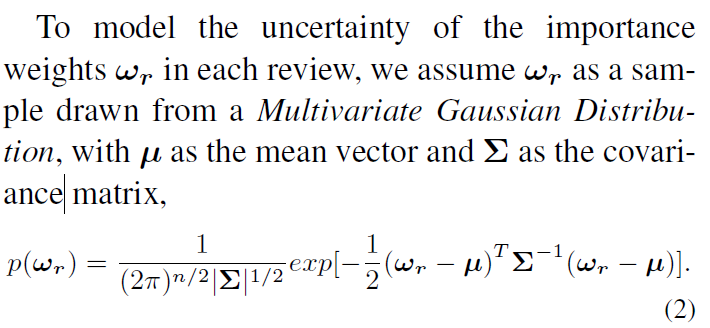
2.3 Aspect Sentiment Classification

Given a free text review, since it may cover various opinions on multiple aspects, we first locate the opinionated expression modifying each aspect, and determine the opinion on the aspect using the learned *SVM* classifier. In particular, since the opinionated expression on each aspect tends to contain sentiment terms and appear closely to the aspect (Hu and Liu, 2004), we select the expressions which contain sentiment terms and are at the distance of less than 5 from the aspect *NP* in the parsing tree.

2.4 Aspect Ranking

Generally, consumer’s opinion on each specific aspect in the review influences his/her overall opinion on the product.





3 Evaluations

3.1 Data and Experimental Setting

To examine the performance on aspect identification and sentiment classification, we employed *F*1-measure, which was the combination of *precision* and *recall*, as the evaluation metric.

3.2 Evaluations on Aspect Identification

3.3 Evaluations on Sentiment Classification

1) Unsupervised method.

2) Supervised method.

3.4 Evaluations on Aspect Ranking

In this section, we compared our aspect ranking algorithm against the following three methods.

1) Frequency-based method. The method ranks the aspects based on aspect frequency.

2) Correlation-based method. This method measures the correlation between the opinions on specific aspects and the overall opinion. It counts the number of the cases when such two kinds of opinions are consistent, and ranks the aspects based on the number of the consistent cases.

3) Hybrid method. This method captures both the aspect frequency and correlation by a linear combination, as *λ· Frequency-based Ranking* + (1 *− λ*)*·* *Correlation-based Ranking*, where *λ* is set to 0.5.

4 Applications

The identification of important aspects can support a wide range of applications. For example, we can provide product comparison on the important aspects to users, so that users can make wise purchase decisions conveniently.

5 Related Work

6 Conclusions and Future Works