Knowledge-Based Systems 61 (2014) 29–47

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

Incorporating appraisal expression patterns into topic modeling for aspect and sentiment word identification

Xiaolin Zheng [a,b,*], Zhen Lin [a,c], Xiaowei Wang [a], Kwei-Jay Lin [d], Meina Song [e]

[a] College of Computer Science & Technology, Zhejiang University, Hangzhou 310027, China
[b] Stanford University, Stanford, CA 94305, USA
[c] Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[d] Dept. Electrical Engineering & Computer Science, University of California, Irvine, CA 92697, USA
[e] School of Computer Science, Beijing University of Posts and Telecommunications, Beijing 100876, China

ARTICLE INFO

Article history:
Received 12 March 2013
Received in revised form 1 February 2014
Accepted 2 February 2014
Available online 11 February 2014

Keywords:
Opinion mining
Appraisal expression pattern
Dependency analysis
Topic modeling
Aspect and sentiment analysis

ABSTRACT

With the considerable growth of user-generated content, online reviews are becoming extremely valuable sources for mining customers' opinions on products and services. However, most of the traditional opinion mining methods are coarse-grained and cannot understand natural languages. Thus, aspect-based opinion mining and summarization are of great interest in academic and industrial research. In this paper, we study an approach to extract product and service aspect words, as well as sentiment words, automatically from reviews. An unsupervised dependency analysis-based approach is presented to extract Appraisal Expression Patterns (AEPs) from reviews, which represent the manner in which people express opinions regarding products or services and can be regarded as a condensed representation of the syntactic relationship between aspect and sentiment words. AEPs are high-level, domain-independent types of information, and have excellent domain adaptability. An AEP-based Latent Dirichlet Allocation (AEP-LDA) model is also proposed. This is a sentence-level, probabilistic generative model which assumes that all words in a sentence are drawn from one topic – a generally true assumption, based on our observation. The model also assumes that every review corpus is composed of several mutually corresponding aspect and sentiment topics, as well as a background word topic. The AEP information is incorporated into the AEP-LDA model for mining aspect and sentiment words simultaneously. The experimental results on reviews of restaurants, hotels, MP3 players, and cameras show that the AEP-LDA model outperforms other approaches in identifying aspect and sentiment words.

1. Introduction

The development of Web 2.0 has enabled people to express their opinions on various products or services easily via blogs, forums, microblogs, etc. Therefore, online customer reviews are extremely important to both merchants and customers. However, the number of online reviews is often too large for people to make decisions by reading all useful reviews. To address this problem, researchers have proposed various methods for extracting and summarizing opinions expressed in reviews, such as sentiment classification and subjectivity analysis [1–9].

Nevertheless, merely evaluating opinions at document level is insufficient. Reviewers may be satisfied with the products or services in general, but not with their every aspect, so a reviewer's positive opinion on one aspect does not mean that he/she favorably views all aspects of the product, and vice versa. For instance, this is a review for Maize Mexican Grill, a Mexican restaurant, written by a customer on Yelp[1] (We have modified some words, for better adaptability to our paper):

> *Best Mexican food in Chambana (and anywhere else that I've eaten)! Food at this restaurant is delicious! The steak, fish, pork and chicken are great too! That is all that needs to be said about the food.*
> *The space is tiny with very limited seating and limited parking.*

(*continued on next page*)

* Corresponding author. Address: Zhejiang University, Hangzhou 310027, China. Tel.: +86 571 87951453.

E-mail addresses: xlzheng@zju.edu.cn (X. Zheng), nblin@zju.edu.cn (Z. Lin), alexwang@zju.edu.cn (X. Wang), klin@ece.uci.edu (K.-J. Lin), mnsong@bupt.edu.cn (M. Song).

[1] http://www.yelp.com/.

http://dx.doi.org/10.1016/j.knosys.2014.02.003
0950-7051/

If you go during lunch or dinner rush, your best bet is to get food to go or even call ahead and order for pickup.
I recommend the asada burrito, but you really can't go wrong with anything on the menu.

As shown in the sample review, the first paragraph clearly indicates that the customer is quite satisfied with the food in the restaurant. Here, "food" is an aspect in the domain "restaurant". In the second paragraph, the customer shows a little discontent with the space, including seating and parking. Here, "space" is another aspect in the domain "restaurant". Therefore, reviewers' opinions should be evaluated at aspect level instead of review level.

In the last decade, researchers have proposed several aspect-based opinion mining methods [10–26] for extracting reviewers' opinions on different aspects of products or services. In these methods, a set of product or service reviews are taken as input, and then aspects and their corresponding opinions are extracted. More specifically, summarizing reviews via an aspect-based opinion mining framework includes identifying the aspects of a product or service and extracting opinions for each aspect.

Some aspect-based opinion mining approaches [16,17,19,21,22] impose constraints on frequently appearing noun phrases to identify aspects. However, these methods extract an excessive number of non-aspects, and occasionally miss appearing aspects. In addition, these approaches cannot group semantically similar aspects, but require manual parameter tuning. Several topic model-based techniques [27–33] have also been proposed, which can automatically group similar semantic aspects, and explain the degree of a word's relevance to a topic. However, most of these studies have identified aspect and sentiment words separately, thereby neglecting the relationship between them. To overcome this problem, many scholars have explored and summarized linguistic knowledge for opinion mining [10,18,22,34–37]. Linguistic knowledge-based methods generally utilize syntactic knowledge. Some of these methods can even extract relationships between aspect and sentiment words at the language level, rather than domain level, by syntactic analysis. Therefore, linguistic knowledge-based methods exhibit better domain adaptability than other methods.

In order to simultaneously mine aspect and sentiment words, we focused on two areas: topic modeling and linguistic knowledge. Even though people use different aspect and sentiment words to express their opinions on a product or service, the Shortest Dependency Paths (SDPs) between aspect and sentiment words can overlap, and so can be extracted from a dependency graph of sentences. Through a series of SDP processing, such as statistics and generalization, the Appraisal Expression Patterns (AEPs) can be obtained.

We extend the Latent Dirichlet Allocation (LDA) model [38] by incorporating AEP information into topic modeling, constructing what we call the AEP-based LDA (AEP-LDA) model. AEP-LDA is a sentence-level, probabilistic generative model. It assumes that:

1. All the words in a single sentence are drawn from one topic-which, based on our observation in this study and several previous works [27,39,40], is often true.
2. T aspects are expressed in reviews, and each aspect includes several semantically similar words.
3. T sentiment word topics exist in reviews, and each sentiment word topic corresponds to an aspect.
4. Reviews include a background word topic.

As far as is known, this study is the first to combine dependency analysis with topic modeling for simultaneously mining aspect and sentiment words. Unlike the existing topic models, the AEP-LDA model uses linguistic knowledge-based AEP information to improve the precision of aspect and sentiment words' identification. In addition, it identifies aspect and sentiment words jointly, which is essential for determining the polarity of the sentiment word in different domains. For example, in the expressions "low LCD resolution" and "low price", the sentiment word "low" represents a negative opinion for "LCD resolution" but positive for "price". If the influence of an opinion's aspect is disregarded, "low" might be treated as a negative or a positive word. Therefore, the polarity of the sentiment word cannot be determined without its corresponding aspect.

The word distribution in reviews varies in different domains, making it difficult to design a robust, domain-independent opinion mining system. One solution is to construct a single model for each domain, but this is time-consuming. The AEP information is highly consistent among different domains (see Sections 6.2 and 6.6), and has excellent domain similarity and adaptability.

Various evaluations on the model were conducted. Experiments show that our model performs well in the following evaluations:

Aspect word identification: The AEP-LDA model can identify concept-coherent aspect words more accurately than other models, such as the standard LDA [38], Local LDA [27], and Sentence LDA (SLDA) [39]. AEP-LDA obtains higher F-score than baseline in various domains.

Sentiment word identification: The AEP-LDA model can capture substantive sentiment words for each aspect. The precision of the top n sentiment words extracted by our model is also higher than the baseline values.

Domain adaptation: The problem of domain adaptation of the AEP information is also considered when a system is trained on reviews from one source domain but is deployed on another. The proposed AEP-LDA model is used to verify the excellent domain adaptability of the AEP information.

The rest of the paper is organized as follows: Section 2 provides an extensive review on related work, and Section 3 introduces the problem statement. Section 4 presents the definitions of SDP and AEP, and discusses the methods for building and processing the dependency path collection. Two related probabilistic graphical models, standard LDA and AEP-LDA, are presented in Section 5. Section 6 details the results of the experimental evaluation, and Section 7 concludes the paper with a summary and discussion of future work.

2. Related work

In this section, we conduct a literature review on related research in aspect-based opinion mining. We divide previous research into three parts: frequency-based approaches, topic model-based techniques, and linguistic knowledge-based methods. These are comprehensively compared with the proposed AEP-LDA model.

2.1. Opinion mining based on aspect frequency

Aspect-based opinion mining has been studied since the early 2000s, beginning with the work of Hu and Liu [16]. The two major tasks involved are aspect identification and opinion identification. A widely used approach in aspect identification involves calculating the frequency of noun phrases, treating high-frequency noun phrases as aspect candidates, and filtering out irrelevant ones on the basis of heuristic rules. Hu and Liu [16] applied associative rule mining to extract aspects of electronic products and treat the nearest adjectives as sentiment words, in order to predict whether each sentence was positive or negative. However, aspect and sentiment words are not always closely related. Although high-frequency aspects are popular among consumers, some other aspects are also discussed, making it necessary to identify such low-frequency aspects.