Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 46 (2015) 93 – 100

Procedia Computer Science

International Conference on Information and Communication Technologies (ICICT 2014)

Summarizing User Opinions: A Method for Labeled-Data Scarce Product Domains

Neethu Kurian[a]*, Shimmi Asokan[a]

[a]*Department of Computer Science and Engineering, Rajagiri School of Engineering and Technology,Kochi,India*

Abstract

Product reviews contain valuable information that can influence the online purchases. Extracting relevant opinions regarding the product by merely reading all the reviews is a herculean task. An automatic method for mining and summarizing opinions in these reviews is necessary for this purpose. Existing methods for opinion summarization requires pre-labeled data from the target domain or other sophisticated lexical resources. We solve the problems of existing methods by using cross-domain sentiment classification coupled with distributional similarity of opinion words to classify and summarize product reviews. Experimental analysis shows that using cross-domain sentiment classification for opinion summarization gives encouraging results.

Peer-review under responsibility of organizing committee of the International Conference on Information and Communication Technologies (ICICT 2014)

Keywords: Opinion Summarization; Cross-Domain Sentiment Classification

1. Introduction

Sentiment analysis and related streams are found to be the most promising research areas in computer science. With the technological advancements in internet, WWW and devices to access them, sharing of information has become easier. This information explosion has also led to generation of huge volumes of sentiment-laden information, which contains opinions regarding topics, products or services. This information is shared largely over blogs, forums and as user reviews. Sentiment analysis and its related technologies help to extract meaningful

* Corresponding author. Tel.:+919447049532.
*E-mail address:*neethukurian20@gmail.com

Peer-review under responsibility of organizing committee of the International Conference on Information and Communication Technologies (ICICT 2014)
doi:10.1016/j.procs.2015.01.062

information from this vast amount of opinion containing information. Product reviews are the key area that benefits from sentiment analysis.

With e-commerce gaining popularity, people are getting more and more used to online shopping. One of the obvious problems of online purchases is that, customer need to buy product without directly seeing it. It is commonplace that, online customers rely on reviews posted by other customers while making purchase decision. In a way, product reviews are indirect experiences on the product. However, there are usually thousands of reviews on a single product. Reading all these reviews to understand the pros and cons of a product and its aspects is not feasible. Creating easily understandable summary of opinions towards product and its aspects is therefore necessary.

Various opinion summarization methods have been proposed in literatures. These techniques require either an in-domain prelabeled set of reviews or some sophisticated lexical resources or sentiment lexicons like SentiWordNet. Even though these techniques give acceptable performance, they have certain disadvantages. Creating labeled data for each new product domain is labor-intensive and time consuming. Using precompiled lexicons do not consider the context in which an opinion word is used. To solve these problems, we have proposed a method that does not require in-domain data and that considers the context in which the opinion word is used. In this paper, we used a cross-domain sentiment classification based method combined with distributional similarity of words for opinion summarization. We also analyze the performance of the proposed method for summarizing reviews on various products.

2. Related Works

2.1. Opinion Summarization

Opinion summarization is a subtask in sentiment analysis. Opinion summarization aims at creating concise summaries from large number of reviews on a product. This enables a customer to better understand the positive and negative aspects of a product, by which, he can make wiser purchase decisions. Opinion summarization in itself consists of several steps such as,

i. Identifying the important aspects of the product
ii. Identifying the opinion words
iii. Classifying the sentences containing aspects according to their polarity
iv. Generating the actual summary.

Existing review summarization methods can be broadly classified into abstractive summarization methods and extractive summarization methods. Abstractive summarization requires deep understanding of concepts expressed in the source text. These methods attempt to understand the source texts and then express them in condensed form in natural language. It usually requires paraphrasing of source text. Abstractive methods initially examine and interpret the source text using various natural language processing techniques, to identify important concepts. It then expresses the identified concepts using new condensed text, which captures the most important information from the source text[24]. On the other hand, extractive summarization methods generate concise summaries by selecting important sentences from the source text and concatenating them together. An opinion summarization system consists of three basic steps[25]: Aspect identification, Sentiment prediction and Summary generation.

Aspect identification is the task of identifying important topics, within the text to be summarized. In product reviews, this includes the words related to products, their parts and attributes. Syntax tree parsing and association rule mining are two basic methods for aspect identification. Nouns and noun phrases are commonly regarded as aspect words in product reviews. The opinion summarization technique proposed by Lu et al[3] used shallow parsing for aspect identification. The method assumed that opinions are expressed as phrases containing aspect and its associated opinion word. Popescu and Etzioni[4] also used syntactic parsing to extract aspects from reviews. They extracted noun phrases from the parsed reviews and classified them into parts and properties with the help of WordNet. Hu and Liu[5] used supervised association rule mining for aspect identification. Rules for identifying feature words are initially generated from the labeled reviews based on POS-tags. These rules were then used for extracting features from input reviews. Ku et al[6] relied on word frequencies to extract aspect words. They considered both paragraph-level word frequencies and document-level word frequencies to identify important aspects. Zha et al[7] used an aspect ranking approach to identify the most important aspects of products. In this paper, we used POS tag based parsing for feature extraction.

The aspect identification is often followed by sentiment prediction. Sentiment prediction aims at identifying the sentiment orientation of opinions on the extracted aspects. Hu and Wu[8] used a simple Sentence Weight classifier to calculate the orientation of a sentence. They initially calculated word weights based on their frequency statistics across training data and their linguistic type. With scored words as features, they trained Sentence Weight classifier to calculate the orientation of a sentence. Hu and Liu[5] used a method based on WordNet for identifying sentiment orientation. The orientations of opinion words were determined by tracing their synonym/antonym relation with some seed words in the WordNet. Ku et al[6] used an automatically created sentiment dictionary to predict sentiment orientation of words. The words in the dictionary and their strengths were learned from multiple sentiment lexicons and thesauri. SentiWordNet is also considered as an efficient method for identifying the sentiment orientation in opinion summarization systems[2]. Yang et al[9] used sentiment dictionaries, automatically constructed based on the rating information of the reviews for sentiment classification. In this paper, we used a cross-domain sentiment classification based method for predicting sentiment orientation of opinion words.

Summary generation, which is the final step, aims at creating an understandable summary of opinions, based on the previous steps. Statistical summaries[10] and textual summaries are the general formats in which summary is organized. While statistical summaries are based on the number of positive and negative opinions associated with each aspect, textual summary aims at selectively displaying important opinion sentences/phrases. Graphical representation of opinion distribution is one form of statistical summaries[11]. Yang et al[9] expressed the summarization of review by showing the score of each product feature. Popescu and Etzioni[4] selected the strongest opinion word associated with each aspect as the summary. Lu et al[3] used clustering to choose representative opinion phrases, which can serve as short summary. Mei et al[12] scored each opinion sentence associated with each topic using TSM model and have chosen the top ranked sentence in each category as the most representative sentence. In another approach, a feature-orientation table[2] is used for creating the summaries. This method generates a feature-orientation table (FO table) that records the aspects and their associated positive and negative opinion descriptors. Some authors[8] classify and summarize all the customer reviews of a product to a list of pros and cons phrases. They generate the compound phrase in the summary by selecting opinion words with high Chi-square correlation with that of aspect words. Zhu et al[13] proposed a graph-based sentence selection method for opinion summarization. In this paper, our aim is to create statistical summaries of aspects.

2.2. Cross-Domain Sentiment Classification

Cross-domain sentiment classification is the task of classifying opinionated text in a domain (source domain) into subjective classes (positive/negative), by utilizing sentiment information obtained from another domain (target domain). It aims at bridging the information gap between the two domains. Cross-domain sentiment classification helps in adapting a classifier trained on one domain into a different domain. Generally, methods for sentiment analysis require labelled data from the same product domain. Manually annotating the text for each new product domain is not feasible, as it is time-consuming and expensive. These difficulties led to the idea of cross-domain sentiment classification. Cross-domain sentiment classification is useful in the case of domains where no labelled data or few labelled data is available. A good number of approaches to cross-domain sentiment classification are based on sentiment transfer across source and target domains.

Aue and Gamon[14] conducted a study on different ways to adapt a sentiment classification system to a new domain. They experimented with different cases, such as, single classifier trained using equal amounts of training data from each of the domains, classifier trained with features that appear in target domain, using an ensemble of classifiers, and using both labelled and unlabeled data.

Structural correspondence learning (SCL) proposed by Blitzer et al[15] is an algorithm for domain adaptation[16] in sentiment classification. SCL chooses pivot features based on their common frequency across source and target domains, and point wise mutual information. The relationship between pivot features and non-pivot features are then modelled by constructing a set of related tasks between them. Pan et al[17] proposed spectral feature alignment that exploits domain independent features to construct a bipartite graph. They used feature clusters to co-align domain independent and domain specific features.

Tan et al[18] proposed MIEA method in which sentiment scores are iteratively calculated to determine the sentiment polarity of documents. Sentiment scores are assigned to documents by considering all possible relationships between documents and words, in both source and target domains. The cross-domain sentiment classification method proposed by Wu and Tan[19] consisted of two stages. In the first stage, a relationship is established between source domain and target domain and the most confidently labelled documents are identified. In the second stage, they employ the manifold-ranking algorithm[20] to compute the ranking score for every unlabeled document, based on the ranks of the identified seed documents. An active learning approach is proposed in[21], which combines the active learning strategy called Query by Committee[20] and the label propagation (LP) algorithm[21] to make the classification decision.

Bollegala et al[1] proposed a semi-supervised approach that uses multiple source domains for document level sentiment classification. The method uses a sentiment sensitive thesaurus coupled with feature expansion. In this paper, we adopted cross-domain sentiment classification method proposed by Bollegala et al[1] for identifying the sentiment orientation of opinion words. We adapted their method to word level sentiment classification, making it suitable for opinion summarization task.

3. Proposed Method

In this paper, a method for opinion summarization for product domains with no prelabeled data is proposed. A cross-domain sentiment classification based method is used for this purpose. In this method, the sentiment information from target domain is leveraged by establishing similarity relationship between target domain reviews and another domain (source domain) where labeled data is available. The steps involved in the process are preprocessing, seed lexicon creation, aspect identification, sentiment classification and summary generation**.** Fig. 1 shows the general architecture of the proposed system.

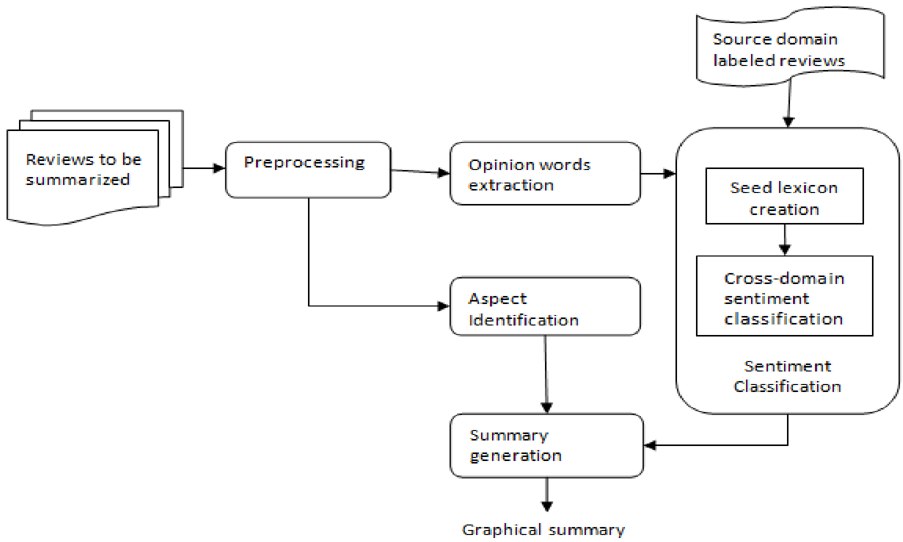

Fig. 1. General Architecture of the proposed system.

3.1. Preprocessing

Preprocessing step consists of sentence boundary detection, word tokenization, and POS tagging. Each review is initially broken down to component sentences. Sentences are then converted to a list of component word tokens. Redundant characters and white spaces appearing in reviews are removed. The words are then tagged with their Part-of-Speech using a POS-Tagger. After preprocessing, each review will be a list of POS tagged words.

3.2. Aspect identification

Aspects of a product usually appear as nouns and noun phrases in product reviews. Most important aspects are the ones that are frequently mentioned by reviewers. The nouns and noun phrases are extracted from the parsed reviews as candidate aspects. From the candidate aspects, those with frequency above a threshold are selected as important aspects.

3.3. Seed lexicon creation

A seed lexicon is created from the labeled reviews of source domain. Opinion words are extracted from the source reviews to create seed lexicons. Adjectives and adverbs are generally considered as opinion words in sentiment analysis literature. The POS tagged reviews are parsed using Regular Expression based parser. From the parsed sentences adjectives, adverbs and adverbial phrases are extracted as opinion words. The sentiment class of opinion words is assumed to be the sentiment class of the labeled review from which they are extracted. Thus, a positive seed lexicon is created from positive source reviews and a negative seed lexicon is created from negative source reviews.

3.4. Sentiment classification

The cross-domain sentiment classification method proposed by Bollegala et al[1] is used for identifying sentiment orientation of opinion words. In the first step, each opinion word extracted from parsed source and target reviews are represented by feature vector. The feature vector of a word contains point wise mutual information values between the word and every other opinion word that co-occur with it in reviews. Every other opinion word that co-occurs with a word contributes a feature to the word's feature vector. Feature vector represents how a word is distributed among the set of all opinion words.

In the next step, we find out the similarity between opinion words. For each opinion word, a similarity matrix is created. The similarity matrix contains the relatedness values[1] of the word with all other opinion words extracted from both source labeled reviews and target unlabeled reviews. Bollegala et al[1] describes the relatedness value between two words as a measure of features shared between the two word vectors. The distributional similarity between words is used here. Based on the sentiment class of the neighboring words in the similarity matrix, the sentiment class of an opinion word is determined. If majority of the neighbors of a word belongs to positive seed lexicon, the word is positive. If majority of neighbors belongs to negative seed lexicon, the word is negative. The seed lexicons are updated with opinion words from target reviews.

3.5. Summary generation

The opinion words that co-occur with an aspect in parsed review sentences are found. The extracted opinion words are then classified into positive words and negative words using the sentiment lexicons. Similar aspects are grouped together. The lists of positive and negative words are identified for each aspect in this way. The positive and negative opinions about an aspect are identified as the number of sentences in which the aspect and the corresponding positive or negative opinion words co-occur. This information is used to create aspect wise graphical summary of positive and negative opinions on the product.

4. Results and Discussion

The proposed system is experimented with user reviews on various products, collected from Amazon.com. Each set of product reviews are summarized using proposed method and the performance is evaluated. We used reviews of various product domains from Blitzer et al.'s Dataset as source domain for cross-domain sentiment classification. Initial experiments are conducted with Books as the source domain and Electronic devices as target domain. Further experiments are conducted by considering other product domains as source. In addition to Books domain, we evaluated the system with DVDs and Kitchen Appliances as source domains.

The proposed system is implemented using Python 2.7. The preprocessing of reviews is done using the Natural Language Toolkit. The Penn Tree Bank tagset is used for POS tagging. Review sentences are parsed with the help of regular expression parser. Aspects and opinion words are extracted from the parsed reviews. In our experiments, we considered nouns and noun phrases with a frequency greater than 5 as important aspects. Nouns and noun phrases that refer to same aspect are identified and are manually grouped into single set. Adjectives and adverbs extracted from sentences containing aspect words are selected as opinion words. Opinion words associated with each aspect are classified into positive/negative using cross-domain sentiment classification. Graphical summary of positive and negative opinions on each aspect is created based on this.

To evaluate the performance of our method, we compared it with the method proposed in[2] which makes use of SentiWordNet, a lexical database with polarity scores, for sentiment classification. Performance of the proposed system was evaluated for precision, recall and F-measure (Table 1). The results obtained for proposed system were comparable with that of SentiWordNet based approach[2].

Table 1. Evaluation of opinion sentence identification.

Evaluation metrics	Method in[2]	Proposed Method
Precision	0.916	0.921
Recall	0.782	0.807
F-measure	0.844	0.860

Classification accuracy of the system is also evaluated by considering different product domains as source. Fig. 2 shows the results obtained for three different product domains: Books, DVDs and Kitchen Appliances on summarizing reviews from electronics domain. Among the three source domains, Kitchen appliances showed the highest accuracy. The results show that the choice of source domain can affect the cross-domain sentiment classification performance. The high accuracy with kitchen domain is due to the similarity of vocabulary in kitchen domain and electronics domain.

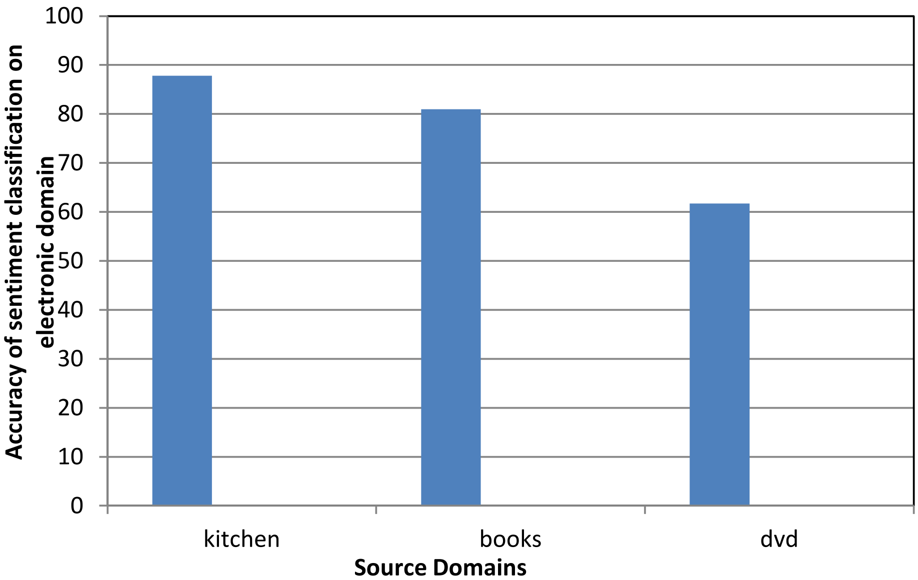

Fig. 2. Sentiment classification performance with different source domains

5. Conclusion

Opinion summarization aims at mining and organizing the opinions within large volumes of opinionated text, such as, product reviews, into an easily understandable form. Opinion summaries help online customers to make a quick analysis of the general opinion towards products, while making an online purchase. Sentiment classification is a crucial step in opinion summarization. Generally, opinion summarization techniques make use of either labeled data or similar sentiment information from the product domain to accomplish this task. Requirement of labeled data makes opinion summarization nearly impossible for product domains in which labeled data is scarce. In this paper, we propose a method for opinion summarization in product domains for which prelabeled data is not available. The proposed method makes use of cross-domain sentiment classification for creating aspect wise graphical summaries. Cross-domain sentiment classification utilizes the sentiment information obtained from another product domain to predict the sentiment class of opinions in target domain. The experiments conducted with product reviews from Amazon.com shows encouraging results. The cross-domain sentiment classification method achieved accuracies comparable to that of SentiWordNet based approach. Experiments conducted with different source domains showed that considering source domains more similar to target domains improves performance. Further researches could investigate the possibility of utilizing domain similarity factor in cross-domain sentiment classification.

References

1. Bollegala, D., Weir, D., and Carroll, J. Cross-domain sentiment classification using a sentiment sensitive thesaurus. *Knowledge and Data Engineering, IEEE Transactions on*, 2012, *25*(8), 1719-1731.
2. Dalal, M. K., and Zaveri, M. A. Semisupervised learning based opinion summarization and classification for online product reviews. *Applied Computational Intelligence and Soft Computing*, 2013, 10.
3. Lu, Y., Zhai, C., and Sundaresan, N. Rated aspect summarization of short comments. In *Proceedings of the 18th international conference on World Wide Web (ACM)*, 2009, 131-140.
4. Popescu, A. M., and Etzioni, O. Extracting product features and opinions from reviews. In *Natural language processing and text mining* , Springer, 2007, 9-28.
5. Hu, M., and Liu, B. Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining (ACM)*, 2004, 168-177.
6. Ku, L. W., Liang, Y. T., and Chen, H. H. Opinion Extraction, Summarization and Tracking in News and Blog Corpora. In *AAAI Spring Symposium: Computational Approaches to Analyzing Weblogs* , 2006, Vol. 100107.
7. Zha, Z., Yu, J., Wang, M., and Chua, T. Product Aspect Ranking and Its Applications. *Knowledge and Data Engineering, IEEE Transactions on*, 2014, vol.26, no.5, 1211-1224.
8. Hu, X., and Wu, B. Classification and summarization of pros and cons for customer reviews. In *Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology,* IEEE, 2009, Volume 03, 73-76.
9. Yang, J. Y., Myung, J., and Lee, S. G. The Method for a Summarization of Product Reviews Using the User's Opinion. In *Information, Process, and Knowledge Management, eKNOW'09. International Conference on*, IEEE, 2009, 84-89.
10. Liu, B., Hu, M., and Cheng, J. Opinion observer: analyzing and comparing opinions on the web. In *Proceedings of the 14th international conference on World Wide Web*, ACM, 2005, 342-351.
11. Hu, M., and Liu, B. Opinion extraction and summarization on the web. In. In *AAAI'06: proceedings of the 21st national conference on Artificial intelligence. AAAI Press,* 2006, Vol. 7, 1621-1624.
12. Mei, Q., Ling, X., Wondra, M., Su, H., and Zhai, C. Topic sentiment mixture: modeling facets and opinions in weblogs. In *Proceedings of the 16th international conference on World Wide Web*, ACM, 2007, 171-180.
13. Zhu, L., Gao, S., Pan, S. J., Li, H., Deng, D., and Shahabi, C. Graph-based informative-sentence selection for opinion summarization. In *Advances in Social Networks Analysis and Mining (ASONAM), 2013 IEEE/ACM International Conference on,* IEEE, 2013, 408-412.
14. Aue, A., and Gamon, M. Customizing sentiment classifiers to new domains: A case study. In *Proceedings of recent advances in natural language processing (RANLP),* 2005, Vol. 1, No. 3.1, 2-1.
15. Blitzer, J., Dredze, M., and Pereira, F. Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification. In *Proc. 45th Ann. Meeting of the Assoc. for Computational Linguistics (ACL '07),* 2007, Vol. 7, 440-447.
16. Blitzer, J., McDonald, R., and Pereira, F. Domain adaptation with structural correspondence learning. In *Proceedings of the 2006 conference on empirical methods in natural language processing (ACL)*, 2006, 120-128.
17. Pan, S. J., Ni, X., Sun, J. T., Yang, Q., and Chen, Z. Cross-domain sentiment classification via spectral feature alignment. In *Proceedings of the 19th international conference on World Wide Web*, ACM, 2010, 751-760.
18. Wu, Q., Tan, S., Cheng, X., and Duan, M. MIEA: a mutual iterative enhancement approach for cross-domain sentiment classification. In *Proceedings of the 23rd International Conference on Computational Linguistics: Posters*, ACL, 2010, 1327-133.
19. Wu, Q., and Tan, S. A two-stage framework for cross-domain sentiment classification. *Expert Systems with Applications*, 2011, *38*(11), 14269-14275.

20. Zhou, D., Weston, J., Gretton, A., Bousquet, O., and Schölkopf, B. Ranking on data manifolds. In *Advances in neural information processing systems*, 2004, *16*, 169-176.
21. Li, S., Ju, S., Zhou, G., and Li, X. Active learning for imbalanced sentiment classification. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, ACL, 2012, 139-148.
22. Freund, Y., Seung, H. S., Shamir, E., and Tishby, N. Information, prediction, and query by committee. In *Advances in neural information processing systems*, 1993, 483-490.
23. Zhu, X., and Ghahramani, Z. *Learning from labeled and unlabeled data with label propagation*. Technical Report CMU-CALD-02-107, Carnegie Mellon University, 2002.
24. Gupta, V., and Lehal, G. S. A survey of text summarization extractive techniques. *Journal of Emerging Technologies in Web Intelligence*, 2010, *2*(3), 258-268.
25. Kim, H. D., Ganesan, K., Sondhi, P., and Zhai, C. Comprehensive review of opinion summarization. 2011.