Sentiment Classification Using Graph Based Word Sense Disambigution

Abbas Jalilvand and Naomie Salim

Faculty of Computer Science and Information Systems
Universiti Teknologi Malaysia, Johor Bahru
jabbas2@live.utm.my, naomie@utm.my

Abstract. Sentiment classification is the most active field in opinion mining that aims to determine whether an opinionated text expresses a positive, negative or neutral opinion. Existing lexicon based sentiment classification methods are unable to deal with context or domain-specific words. To solve this problem, Word Senses Disambiguation (WSD) is useful to identify the most related meaning (sense) of a word in a sentence. In this paper, a sense level sentiment classification method is proposed that determine the sentiment polarity of words using graph based WSD algorithm and a multiple meaning (sense) sentiment lexicon. To evaluate the impact of WSD on sentiment classification, the proposed method compared against a baseline method using two subjectivity lexicons, namely the MPQA and SentiWordNet. Experimental results using a benchmark dataset show that the WSD is effective for sentiment classification.

Keywords: opinion mining, sentiment classification, word sense disambiguation, context dependent word.

1 Introduction

In recent years, with the rapid growth of social media, such as forums, blog, discussion boards and social networks, people can freely express and respond to opinion on variety of topics. Thus, a very large amount of user-generated content has been available on the Web. However, the high volume of reviews makes it difficult for individuals and organizations to read and understand all of them. To solve this problem, a hot research area has recently emerged, which is called opinion mining and sentiment analysis. Sentiment classification is the most active field in opinion mining that aims to determine whether an opinionated text expresses a positive, negative or neutral opinion. Sometimes, subjectivity classification is used as an input data pre-processing step for sentiment classification. Sentiment classification is applied at word-level, sentence-level, document-level and feature/aspect-level using different methods ranging from unsupervised to supervised approaches that can be categorized into two main methodologies: semantic-orientation and machine learning approaches[1-4]. In the semantic-orientation approach, a text is classified based on average polarity of words/phrases containing positive or negative sentiment using a sentiment lexicon

A. Ell Hassanien et al. (Eds.): AMLTA 2012, CCIS 322, pp. 351–358, 2012.

and linguistic rules. This lexicon consists of a list of opinion words and their polarity (positive, negative or neutral). In fact, this approach works at word-level to classify at sentence or document-level. One of the shortcomings of these methods is that they are unable to deal with context or domain-specific words. For example, the word "fight" expresses a positive sentiment in "the football team was full of fight" while is negative in "the people are always fighting". Words have different meaning (sense) in different context, and their sentiment polarities are different. Thus, to have better result in classification, it should be applied at sense-level rather than word-level. In this paper, a sense-level sentiment classification method is proposed. In this method, first, the most related sense of word is found according to its context and then the sentiment of word is determined using the multiple meanings sentiment lexicon. To find the related sense of words, WSD method is useful. WSD is an active field of natural languages processing to identify more related meaning of a word in a sentence. The graph based WSD algorithm builds a graph corresponding to a word sequence from WordNet [5] and then finds the stronger link as a related sense [6]. In this research, this algorithm is used for sentiment classification due to its high accuracy of performance. By using this strategy the problem of context dependent will be solved in sentiment classification. The remainder of the paper is organized as follows: Section 2 provides a review of related work on sentiment classification methods. Section 3 provides the research design. Section 4 shows the experimental results. Finally, Section 5 outlines conclusions.

2 Related Works

In the semantic-orientation approach, a text is classified based on average polarity of words/phrases containing positive or negative sentiment using a sentiment lexicon incorporated to syntactic rules. In recent years, two kinds of approaches have been proposed to build a sentiment lexicon: thesaurus and corpus based. Thesaurus based approach aims to grow a small set of seed opinion words using their synonyms, antonyms, and hierarchies in a thesaurus, e.g., WordNet [5] to generate a sentiment lexicon based on a bootstrapping process [7-10].

WordNet is an online lexical reference system whose design is inspired by psycholinguistic theories of human lexical memory. The WordNet lexicon contains nouns, verbs, adjectives, and adverbs. Lexical information is organized in terms of word meanings, rather than word forms. Senses in the WordNet database are represented relationally by synonym sets (synsets) that sets of all words sharing a common sense.

In this essence, Kim and Hovy [7] adopted the method which is based on synonym and antonym lists originated from WordNet to calculate the probability of sentiment polarity of words. Kamps et al. [11] used a hypothesis that synonyms have similar sentiments. They used related synonyms from the thesaurus in order to construct a lexical network; consequently, the sentiment of the word can be measured with the distance from seed words (''good'' and ''bad''). Furthermore, Hu and Liu [8] improved the method of Kamps et al. by using both synonyms and antonyms in order to

construct lexical network. After that, Esuli and Sebastiani [12]found the sentiment of words according to the glosses of subjective terms.

MPQA [13] and SentiWordNet [9, 10] are also two popular sentiment lexicon. MPQA lexicon consists of over 8,000 subjective single-word clues. SentiWordNet is Built upon WordNet. In SentiWordNet each synset of WordNet is automatically annotated with three sentiment scores regarding their positivity, negativity, and objectivity. Many researchers have used SentiWordNet as the sentiment lexicon.

Corpus-based approaches expand a seed list of opinion words using syntactic or co-occurrence patterns in a large corpus[1, 14]. Since the corpus-based approach is inherently context based, it can handle domain and context-specific opinion words; unlike the thesaurus based approach. On the other hand, thesaurus-based approach is more effective than corpus-based, because it is difficult to prepare a huge corpus.

Recently some researchers have addressed the context dependent problem for words such as "high, small, low, and etc." Ding et al. [3] proposed a holistic lexicon-based approach that deals with context dependent opinion words using three rules based on contextual information in other reviews and sentences in same domain. Shortcoming of this approach is that, it is limited to a specific domain. Wanton et al. [15] used WSD-based method in a corpus of newspaper quotations. They worked on two-word phrases as ambiguity terms, and used the sources of SentiWordNet [10] and General Inquirer (GI) [16] to calculate the sentiment polarity. The weaknesses of this method are that it is unable to disambiguate words in sentence and used sources are not annotated based on sense polarities.

Wu and Wen [17] focused on 14 adjectives such as "large, small, many, few, high, low, and etc.", as Dynamic sentiment ambiguous adjectives (DSAAs). They defined the semantic expectation for nouns that shows the tendency of polarity. They predicted semantic expectation for nouns using search engine, and determined the sentiment of word based on semantic expectation and DSAAs. The performance of this method is depending on search engine.

3 Research Design

The proposed method aims to determine the sentiment of words based on the most related sense of words. In order to develop this method, a multiple meanings sentiment lexicon is needed. To construct this lexicon, a bootstrapping technique is applied that start with a list of seed words and uses the synonyms and antonyms of words to enlarge this list from WordNet. In this paper, SentiWordNet is used as a multiple meanings sentiment lexicon.

Firstly, in this method the most related sense of word is found according to its context and then, the sentiment of the word is determined using the multiple meanings sentiment lexicon. To find the related sense of words, graph based WSD algorithm [6] is used.

This algorithm builds a graph corresponding to words sequence from WordNet and then finds the stronger link as a related meaning (sense) of target word. Figure 1 illustrates the framework of proposed approach.

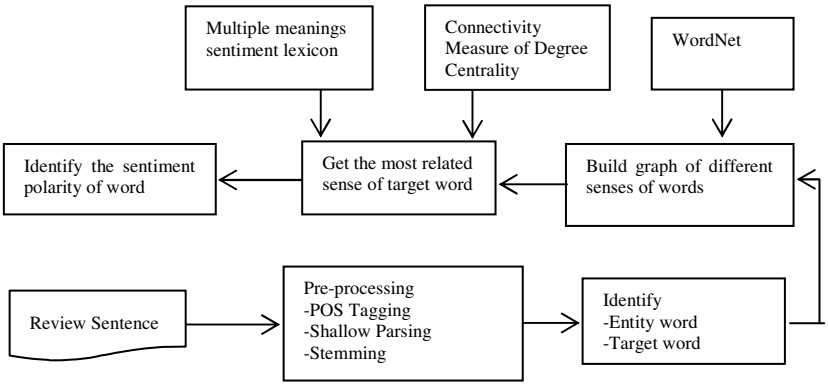

Fig. 1. Framework of WSD Based Sentiment Classification

For example, consider the sentence "people are always fighting". In the first step, preprocessing (POS Tagging, Shallow Parsing and Stemming) will be applied. After preprocessing, "people" and "fight" as entity and target words are identified respectively. Then, all different senses of "people" and "fight" words are extracted from WordNet which are four and nine respectively. To build the graph of words, four different senses of "people" in WordNet and the important words in their glosses are connected to "people". Then these words are followed leading to "fight" using depth-first search (DFS). To assign the most related senses of "fight", the shortest path between the two words is found using connectivity measure of degree centrality[6]. After finding the appropriate sense for "fight", its sentiment polarity is determined using multiple meanings sentiment lexicon[10].

4 Experiments and Result

In this section, experimental results are presented and discussed. To evaluate the impact of WSD method on sentiment classification, the proposed method is compared against a baseline method that determines sentiment of a review, using aggregated sentiment score of its words based on two different sentiment lexicons. Results evaluated using standard evaluation measures accuracy based on same data set.

4.1 Dataset Description

Two publicly available datasets will be used in this research. The movie review (MR)[1] crawled from the IMDB movie archive consists of 1000 positive and 1000 negative movie reviews [2], and the multi-domain sentiment(MDS)[2] used by Blitzer et al.[18] crawled from Amazon.com containing four different types of product reviews (Book, DVD, Electronics and Kitchen). This dataset contains 1000 positive and

[1] http://www.cs.cornell.edu/people/pabo/movie-review-data
[2] http://www.cs.jhu.edu/~mdredze/datasets/sentiment/index2.html

1000 negative examples for each domain. Pre-processing was performed on both of the datasets. Firstly, punctuation, numbers, non-alphabet characters and stop words were removed. Secondly, Porter's stemmer[19], Stanford POS tagger[3] and parser[4] were performed to identify phrases in the form of a pair of head term and modifier. Summary statistics of the datasets before and after preprocessing are shown in Table 1.

Table 1. Dataset in the number of words

Dataset	**MR**	**MDS**			
		Book	DVD	Electronic	Kitchen
Corpus size(before pre-processing)	674,662	214,350	212,413	146,159	129,587
Corpus size(after pre-processing)	450,032	120,553	119,887	74,996	64,443

4.2 Baseline Method

In this method, first, positive and negative words are extracted from review documents and then simply assign a score +1 and -1 to positive and negative words respectively based on sentiment lexicon. A document is classified as positive if the sum of score is above zero and as negative if is below zero. This method was implemented using two subjectivity lexicons, namely the MPQA[13] and SentiWordNet[10]. MPQA subjectivity lexicon contains 2,718 positive and 4,911 negative words. SentiWordNet is a multiple meaning sentiment lexicon. It is automatically annotated from WordNet. In SentiWordNet, each synset has three sentiment scores regarding how positive, negative, and objective. Since, in this method work based on single meaning of words, the first sense of words which is more important sense is considered as opinion words. It should be noted that these lexicons are domain-independent.

4.3 Proposed Method

In contrast to the baseline method, our method considers different sentiment polarities for words based on their different meanings in the context. To find the related meaning (sense) of words, graph based WSD algorithm [6] is used. After finding the appropriate sense, sentiment polarity is determined using multiple meanings sentiment lexicon. In this paper, SentiWordNet is used as a multiple meanings sentiment lexicon. Because each word in SentiWordNet has multiple senses, we calculated the sentiment polarity of each sense from the three polarity scores (positive, negative, and objective) using a sentiment polarity calculation strategy adopted from previous literature [20] (Figure 2).

[3] `http://nlp.stanford.edu/software/tagger.shtml`
[4] `http://nlp.stanford.edu/software/lex-parser.shtml`

```
If (score(sense)_objective > 0.5 or (score(sense)_positive = (score(sense)_negative
    Sentiment polarity (sense) =neutral
Else
    If (score(sense)_positive > (score(sense)_negative
        Sentiment polarity (sense) =positive
    Else
        Sentiment polarity (sense) =negative
```

Fig. 2. Sentiment polarity calculation strategy

4.4 Results

Table 2 gives the experimental results for sentiment classification at document level using the baseline and proposed methods. The standard evaluation measures of accuracy used to evaluate performance of methods. From table 2, we observe that the proposed method yield results better than baseline method in different experiment. This improvement shows that WSD is useful to determine the sentiment polarity of words at sense level using the multiple meanings sentiment lexicon.

Table 2. Experimental results for baseline and proposed methods(accaracy%)

Method	MR	MDS			
		Book	DVD	Electronic	Kitchen
MPQA	65.90	61.95	63.40	60.30	61.85
SentiWordNet	60.60	59.50	59.75	58.25	59.30
Graph based WSD	67.15	63.70	65.60	63.80	64.60

5 Discussion and Conclusion

In this study, a sense level sentiment classification method was presented. However, words have different meanings (senses), but they don't necessarily represent the same sentiment polarity. The existing methods only consider one of the different senses for words; thus, they cannot deal with context dependent problems. The proposed method determines the sentiment polarity of each word based on its most related sense in the context; therefore, has better result in comparison with current lexicon based sentiment classification approach. To find the most related sense of words, a graph based WSD algorithm is applied. The proposed method was compared to a baseline method using two subjectivity lexicons (MPQA and SentiWordNet). The experimental results is shown the proposed method outperform the current method. It can be concluded

WSD can be a useful part of an opinion mining system. We plane to apply WSD in machine learning sentiment classification approach.

Acknowledgment. This work is supported by ministry of higher education and research Management Centre at Universiti Teknologi Malaysia (UTM) under Research University grant Category GUP (VOT Q.J130000.7128.00H72).

References

1. Turney, P.D.: Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In: Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL), Philadelphia, pp. 417–424 (2002)
2. Pang, B., Lee, L., Vaithyanathan, S.: Thumbs up?: sentiment classification using machine learning techniques. In: Proceedings of Conference on Empirical Methods in Natural Language Processing, EMNLP 2002 (2002)
3. Ding, X., Liu, B., Yu, P.: A holistic lexicon-based approach to opinion mining. In: Proceedings of the Conference on Web Search and Web Data Mining, WSDM 2008 (2008)
4. Abbasi, A., France, S., Zhang, Z., Chen, H.: Selecting attributes for sentiment classification using feature relation networks. IEEE Transactions on Knowledge and Data Engineering (99), 1 (2011)
5. Felbaum, C.: Wordnet, an Electronic Lexical Database for English. MIT Press, Cambridge (1998)
6. Navigli, R., Lapata, M.: An experimental study of graph connectivity for unsupervised word sense disambiguation. IEEE Transactions on Pattern Analysis and Machine Intelligence 32(4), 678–692 (2010)
7. Kim, S., Hovy, E.: Determining the sentiment of opinions. In: Proceedings of International Conference on Computational Linguistics, COLING 2004 (2004)
8. Hu, M., Liu, B.: Mining and summarizing customer reviews. In: Proceedings of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD 2004 (2004)
9. Esuli, A., Sebastiani, F.: SentiWordNet: a publicly available lexical resource for opinion mining. In: Proceedings of Language Resources and Evaluation, LREC 2006 (2006)
10. Baccianella, S., Esuli, A., Sebastiani, F.: Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining (2010)
11. Kamps, J., Marx, M., Mokken, R.J., Rijke, M.: Using WordNet to measure semantic orientation of adjectives. In: Proceedings of LREC 2004, 4th International Conference on Language Resources and Evaluation, Lisbon, PT, vol. IV, pp. 1115–1118 (2004)
12. Esuli, A., Sebastiani, F.: Determining the semantic orientation of terms through gloss classification. In: Proceedings of ACM International Conference on Information and Knowledge Management, CIKM 2005 (2005)
13. Wilson, T., Wiebe, J., Hoffmann, P.: Recognizing contextual polarity in phrase-level sentiment analysis. In: Proceedings of the Human Language Technology Conference and the Conference on Empirical Methods in Natural Language Processing, HLT/EMNLP 2005 (2005)
14. Read, J., Carroll, J.: Weakly supervised techniques for domain-independent sentiment classification. ACM (2009)
15. Martın-Wanton, T., et al.: Word sense disambiguation in opinion mining: Pros and cons. Special issue: Natural Language Processing and its Applications, 119 (2010)

16. Stone, P.J., Dunphy, D.C., Smith, M.S.: The General Inquirer: A Computer Approach to Content Analysis (1966)
17. Wu, Y., Wen, M.: Disambiguating dynamic sentiment ambiguous adjectives. Association for Computational Linguistics (2010)
18. Blitzer, J., Dredze, M., Pereira, F.: Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification (2007)
19. Porter, M.F.: An algorithm for suffix stripping. Program (1980)
20. Dang, Y., Zhang, Y., Chen, H.: A lexicon-enhanced method for sentiment classification: An experiment on online product reviews. IEEE Intelligent Systems 25(4), 46–53 (2010)