Multimed Tools Appl (2018) 77:21265–21280
https://doi.org/10.1007/s11042-017-5529-5

Generate domain-specific sentiment lexicon for review sentiment analysis

Hongyu Han [1] **· Jianpei Zhang** [1] **· Jing Yang** [1] **·**
Yiran Shen [1] **· Yongshi Zhang** [1]

Received: 9 January 2017 / Revised: 31 October 2017 / Accepted: 11 December 2017 /
Published online: 7 January 2018

Abstract Lexicon-based approaches for review sentiment analysis have attracted significant attention in recent years. Lots of sentiment lexicon generation methods have been proposed. However, the generation of domain-specific lexicon with unlabeled data has not been effectively addressed. In this paper, we propose a new domain-specific sentiment lexicon generation method, mutual information is introduced to assign terms with Part-Of-Speech (POS) tags in the lexicon, the training data are selected from unlabeled corpus according to their sentiment scores which are evaluated by the SentiWordNet (SWN) based sentiment classifier. Then we propose a completed lexicon-based sentiment analysis framework which uses the domain-specific sentiment lexicon generated by the proposed domain-specific sentiment lexicon generation method. The experiment is carried out on publically available datasets. Results show that the proposed lexicon-based sentiment analysis framework using domain-specific lexicons generated by the proposed method gets a good performance.

Keywords Lexicon-based approach · Sentiment analysis · SentiWordNet · Mutual information

1 Introduction

In recent years, along with the network popularization and the vigorous development of e-commerce, more and more people are shopping online. There are a large number of users comment on what they purchase online each day. The reviews provide important guidance for both customers and businesses. Customers make their decisions based on existing reviews, and business companies discover the problems of their products or services from the attitudes of users in the reviews [11]. Sentiment analysis is the computational study of people's opinions, sentiments, emotions, and attitudes [10].

✉ Hongyu Han
hanhongyu@hrbeu.edu.cn

1 College of Computer Science and Technology, Harbin Engineering University, No.145 Nantong Street, Nangang District, Harbin 150001, China

Sentiment analysis methods can be generally divided into two categories, machine learning and lexicon-based methods. The former uses machine learning techniques for sentiment polarity classification. These kinds of methods usually need a lot of labeled training data. However, collecting sufficient labeled data is a challenge in itself. Lexicon based methods utilize sentiment lexicons to compute sentiment scores of given reviews. Then they group the scored reviews into positive or negative categories by the sentiment scores. Sentiment lexicon is the foundation of lexicon-based method, therefore, how to construct a sentiment lexicon has attracted lots of attention naturally. The sentiment generation methods are mainly divided into two categories: dictionary-based and corpus-based approaches. The study [14] indicates that the domain-specific sentiment lexicon leads to better sentiment analysis compared to a general sentiment lexicon. For example, *long* is a positive word when it is used to describe the phone's standby time in a review that comments on a phone, but it is a negative word when it is used to describe the time that a printer prints a paper. Despite a significant amount of research, there is still no an effective method for discovering and determining domain-dependent sentiment lexicon [10]. The existing methods have some deficiencies: 1. these methods only apply to some specific domains in which emoticons are used frequently [18]; 2. they need human-annotated data [14]; 3. the generated sentiment lexicon contains more positive or negative words [7].

In this paper, we propose a domain-specific sentiment lexicon generation method and a sentiment analysis framework based on the generated domain-specific sentiment lexicon. The proposed method exploits the unlabeled reviews to generate the domain-specific sentiment lexicon. Specifically, we first label the training reviews with 'Positive' or 'Negative' tags using the SWN-based sentiment analysis method. Then part of the labeled reviews are selected according to the ranking of their sentiment scores. Second, we extract the terms with special POS tags from the selected training data as the entries of the generated sentiment lexicon, then the mutual information of each entry can be obtained as the sentiment strength by statistical calculation. Finally, we achieve the sentiment classification on the testing data by weighing the positivity and negativity of the terms in the generated domain-specific lexicon. The proposed domain-specific sentiment lexicon generation method and the proposed sentiment analysis framework can be widely used in more domains for they have almost no requirement for the corpus.

The remainder of this paper is structured as follows. Section 2 discusses the related works. The proposed method is presented in Section 3. Section 4 presents the experiments and discussions. Finally, we conclude the paper in Section 5.

2 Related works

The lexicon-based approach calculates the final sentiment tendency of a review by rating the sentiment tendency of each word or phrase in a given review [16]. There are several well-known sentiment lexicons such as SentiWordNet [2], MPQA [20] etc. A large amount of research has been studied on these lexicons. The study [8] makes use of SentiWordNet and treats it as the labeled corpus (For each entry in SentiWordNet, if its positivity is greater than negativity, its gloss is labeled as a positive document; if its positivity is less than negativity, its gloss is labeled as a negative document; otherwise, its gloss is labeled as a neural document.) for training, builds upon the mutual information calculated from these terms, generates a general sentiment dictionary, then proposes a completed lexicon-based sentiment analysis framework. The study [1] deals with the lexicon-based approach in document-level and sentence e-level sentiment analysis (SA) in Arabic. It experimented four different lexicons: a translation of Harvard IV-4 Dictionary (HarvardA), translation of the MPQA subjectivity lexicon developed by Pittsburgh University (HRMA) and two different implementations of MPQA. The study [17] presents a lexicon-based method for sentiment analysis via making use of adjectives and extending the Semantic Orientation CALculator to other parts of speech.

The sentiment lexicon generation methods are mainly divided into two categories: dictionary-based and corpus-based approach.

Dictionary-based approaches make use of WordNet or another dictionary to propagate sentiment from seed words (with known positive or negative sentiment orientation) [4, 6]. While the main disadvantage of dictionary-based approaches is that the sentiment orientations of words collected by this way are general or domain and context independent [10].

Corpus-based approaches discover sentiment words and their orientations from a domain corpus. These approaches are more appropriate to induce domain-specific lexicons. The study [14] extracts a sentiment lexicon from a domain-specific corpus by annotating an intelligently selected subset of documents in the corpus, the subset is selected by an active learner. Although this method could generate excellent sentiment lexicon, the training data acquisition is a laborious process. The study [18] builds large-scale sentiment lexicon from Twitter with a representation learning approach. On one hand, it casts sentiment lexicon learning as a phrase-level sentiment classification task. On the other hand, it develops a dedicated neural architecture and integrates the sentiment information of text into its hybrid loss function for learning sentiment-specific phrase embedding (SSPE). The neural network is trained from massive tweets collected with positive and negative emoticons, without any manual annotation. However, this method only applies to the field where emoticons are widely used. The study [7] combines domain-specific word embeddings with a sentiment propagation framework (SenProp) to induce accurate domain-specific sentiment lexicons using small sets of seed words. However, domain-specific lexicon generated by the sentiment propagation method may induce more positive or negative words.

The method proposed in this paper is different from these methods mentioned above, we use unlabeled reviews as the training data, then acquire the sentiment polarity tag for each review of the training data by utilizing the SWN-based sentiment classifier, select the reviews with high absolute value of the sentiment score as the sentiment lexicon training set, and finally generate the domain-specific sentiment lexicon. A completed sentiment analysis framework is constructed to evaluate the generated domain-specific sentiment lexicon.

3 The proposed sentiment analysis method

In this section, we will show the content of our proposed method. First, we demonstrate the generation of the domain-special sentiment lexicon, and then describe the usage of the newly generated sentiment lexicon for review sentiment classification. Figure 1 shows the process of generating the domain-specific sentiment lexicon.

3.1 Domain-specific dataset acquisition

The domain-specific sentiment lexicon is extracted from a specific domain, so the first step is to acquire a domain-specific review dataset. In this paper, we mainly consider how to induce the sentiment lexicon, so we use the domain-specific dataset collected by other researchers.

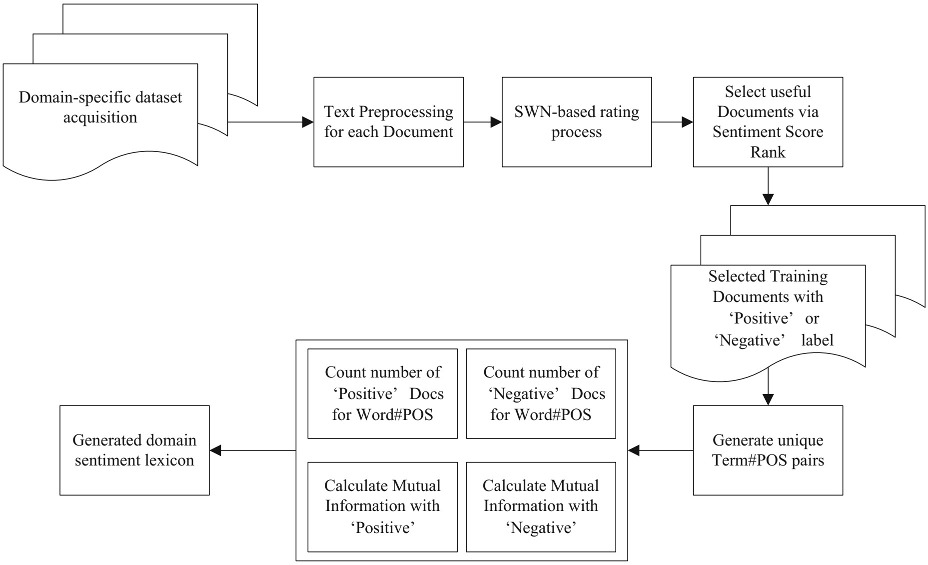


Fig. 1 Construction process of domain-special sentiment lexicon

3.2 Text preprocessing

Text preprocess is mandatory for sentiment analysis, in order to cope with texts published by users on social media platforms [15]. The text preprocessing in the proposed method consists of four parts as follows:

3.2.1 Remove URLs and special symbols

URLs in user reviews do not convey any emotional tendencies, therefore we remove them. Some special symbols (such as () @ # $%, etc.,) are often used in the reviews, while these symbols do nothing but interfere with the sentiment classification. In order to avoid these disturbances, special symbols are removed.

3.2.2 Part-of-speech tagging

Part-Of-Speech Tagger (POS Tagger) is used to assign part of speech to each word in the text (and other tokens), such as noun, verb, adjective, etc. [19]. We can assign out the useful words which may convey opinions of users according to their POS tags. It is very useful to the review sentiment analysis. As this process makes use of Penn POS Tags, and SentiWordNet with only four POS Tags (a, r, v, and n) is used, therefore, SWN POS Tags are used as the standard POS Tags. The relationship between Penn and SWN POS Tags are shown in Table 1.

3.2.3 Filtering

Traverse the reviews to retain the terms with POS Tags in Table 1 and to remove all the terms with other POS Tags. Removing terms with other POS Tags can increase the speed of the program without affecting the accuracy of the review sentiment polarity classification.

Table 1 Look-up table for POS [8]

Penn POS tag	Description	Equivalent SWN POS tag
JJ	Adjective	a
JJR	Adjective, comparative	a
JJS	Adjective, superlative	a
NN	Noun, singular or mass	n
NNS	Noun, plural	n
NNP	Proper noun, singular	n
NNPS	Proper noun, plural	n
RB	Adverb	r
RBR	Adverb, comparative	r
RBS	Adverb, superlative	r
VB	Verb, base form	v
VBD	Verb, past tense	v
VBG	Verb, gerund or present participle	v
VBN	Verb, past participle	v
VBP	Verb, non-3rd person singular present	v
VBZ	Verb, 3rd person singular present	v

3.2.4 Lemmatization

Lemmatization is used to map word-forms to a standardized lexicon entry [9]. SentiWordNet is used as the standardized lexicon in this process. For example, "enjoying" is converted to "enjoy". Lemmatization procedure in NLTK (Natural Language Toolkit) [13] is utilized to perform this process.

3.3 SWN-based sentiment classifier

SentiWordNet 3.0 is a general sentiment lexicon. Table 2 shows a piece of SentiWordNet 3.0.

The pair (POS, ID) uniquely identifies a WordNet [5] synset. *PosScore* and *NegScore* (range from 0 to 1) represent the positivity and negativity of the corresponding synset. SynsetTerms column reports the terms, with sense number, belonging to the synset. The sentiment score of the corresponding synset *SynsetScore* is calculated as follows:

$$SynsetScore = PosScore–NegScore \tag{1}$$

Table 2 A piece of SentiWordNet 3.0

POS	ID	PosScore	NegScore	SynsetTerms	Gloss
a	643,598	0.5	0	notional#3 fanciful#1	indulging in or influenced by fancy; "a fanciful mind"; "all the notional vagaries of childhood"
n	8,388,871	0.375	0	knighthood#1	aristocrats holding the rank of knight
r	370,046	0.25	0	incisively#1	in an incisive manner; "he was incisively critical"
v	1,922,895	0.125	0.125	mountaineer#1	climb mountains for pleasure as a sport

For a term with a particular POS tag, if it appears in n synsets of the SentiWordNet 3.0, the final sentiment score of the term can be calculated as follow by weighing the synsets according to their sense numbers:

$$TermScore = \left(\sum_{r=1}^{n} SynsetScore(r)/r\right) / \left(\sum_{r=1}^{n} 1/r\right) \tag{2}$$

Where r is the sense number of the synset contained the term.

For a given review, preprocess it at first, TermScore for each term with its corresponding POS tag in the review can be obtained via formula (1) and (2), and then the TermScore of each term is added to get the final sentiment score for the review. If the sentiment score of the review is greater than 0, the review is classified as positive, otherwise, the review is classified as negative.

3.4 Training data selection

A SWN-based sentiment classifier is utilized to score reviews in the domain-specific dataset, each review can be classified as a positive or negative review according to its sentiment score (that is, if the sentiment score of a review is greater than 0, it is classified as a positive review, otherwise, it is classified as a negative review). We rank the positive and negative reviews according to their absolute value of sentiment score separately. The results show that the review classification precision of sentiment polarity classification is increasing as the absolute value of sentiment score grows. Thus we select the top ω_1 (the value of ω_1 is from 0 to 100%) of both positive and negative reviews as the lexicon training data. The effect of ω_1 on sentiment polarity classification will be discussed in Section 4. The selected lexicon training data is utilized to generate the domain-specific sentiment lexicon.

3.5 Domain-special sentiment lexicon generation

In this paper, we extract sentiment lexicon from the above newly selected lexicon training data by introducing mutual information to evaluate the sentiment tendencies of the terms.

3.5.1 Mutual information

The basic idea of mutual information is: the greater the mutual information, the greater the degree of co-occurrence of the pair (Term, POS) t_i and category (Positive, Neutral and Negative) C_j. The mutual information of t_i and C_j can be calculated as:

$$I\left(t_i, C_j\right) = \log \frac{P\left(t_i, C_j\right)}{P\left(C_j\right)P\left(t_i\right)} \cong \log \frac{A \times N}{(A+B) \times (A+C)}, \tag{3}$$

Table 3 Contingency table

	t_i occurs	t_i does not occur
C_j occurs	A	B
C_j does not occur	C	D

The meanings of the representatives of A, B and C are shown in Table 3. N is the total number of reviews. If feature t_i is not relevant to category C_j, then $I(t_i,C_j) = 0$.

3.5.2 Sentiment lexicon extraction

First, we traverse each term with its corresponding POS tag in each review in the training set, if it has not existed in the proposed domain-special sentiment lexicon, we add it to the lexicon. Then we traverse the whole training set, and count the numbers of times each term with its corresponding POS tag in the domain-special sentiment lexicon appears in the positive reviews and negative reviews separately. According to the formula (3), the mutual information of positive and negative of corresponding POS tagged term can be obtained. The format of the domain-special sentiment lexicon shown in Table 4.

3.6 The proposed lexicon-based sentiment analysis framework

The flow chart of the proposed sentiment classifier which utilizes the domain-specific sentiment lexicon generated by out proposed method is shown in Fig. 2.

For a given review, we preprocess it at first, and then extract the sentiment score from our proposed sentiment lexicon for each term in the review, if a term is not found in the dictionary, we assign a score of 0 to it. A weighting factor α is introduced to ensure the accuracy of the sentiment classification, it ranges from 0 to 1. The final score for each term with its POS tag is calculated as follows:

$$TS(t_i) = \alpha \times P(t_i) + (1-\alpha) \times (-N(t_i)), \tag{4}$$

$TS(t_i)$ is the final sentiment score of the term with its POS tag, $P(t_i)$ and $N(t_i)$ are the values of the positivity and negativity of the corresponding term with its POS tag.

The determination process of α can be achieved automatically. First, the SWN-based approach is used to evaluate and classify the testing dataset, rank the reviews according to their positivity and negativity separately, and then select the top ω_2 (range from 0 to 100%) of the reviews with the highest absolute scores as the training data for the determination of α. The effect of ω_2 on sentiment polarity classification will be discussed in the experiment. The parameter α is taken from 0 to 1, and a value is taken every 0.001. For each value, the sentiment tendencies are re-evaluated using the previously generated sentiment lexicon for the reviews of the training data, if the score is greater than 0, the review is classified as positive, otherwise it is classified as negative. The sentiment classification results of the training dataset corresponding to each α value are compared with those of the SWN-based sentiment classification. Select the value which makes the classification results of the greatest degree of similarity as the final value of α, If there are multiple values corresponding to the highest similarity at the same time, select their average as the final value of α.

Table 4 A piece of the proposed domain-special lexicon

Term#POS	PosCnt	NegCnt	PosMI	NegMI
enjoy#a	59	26	1.21022	0.0280145
hate#v	123	174	0.465161	0.96559
sad#r	21	32	0.401363	1.00905

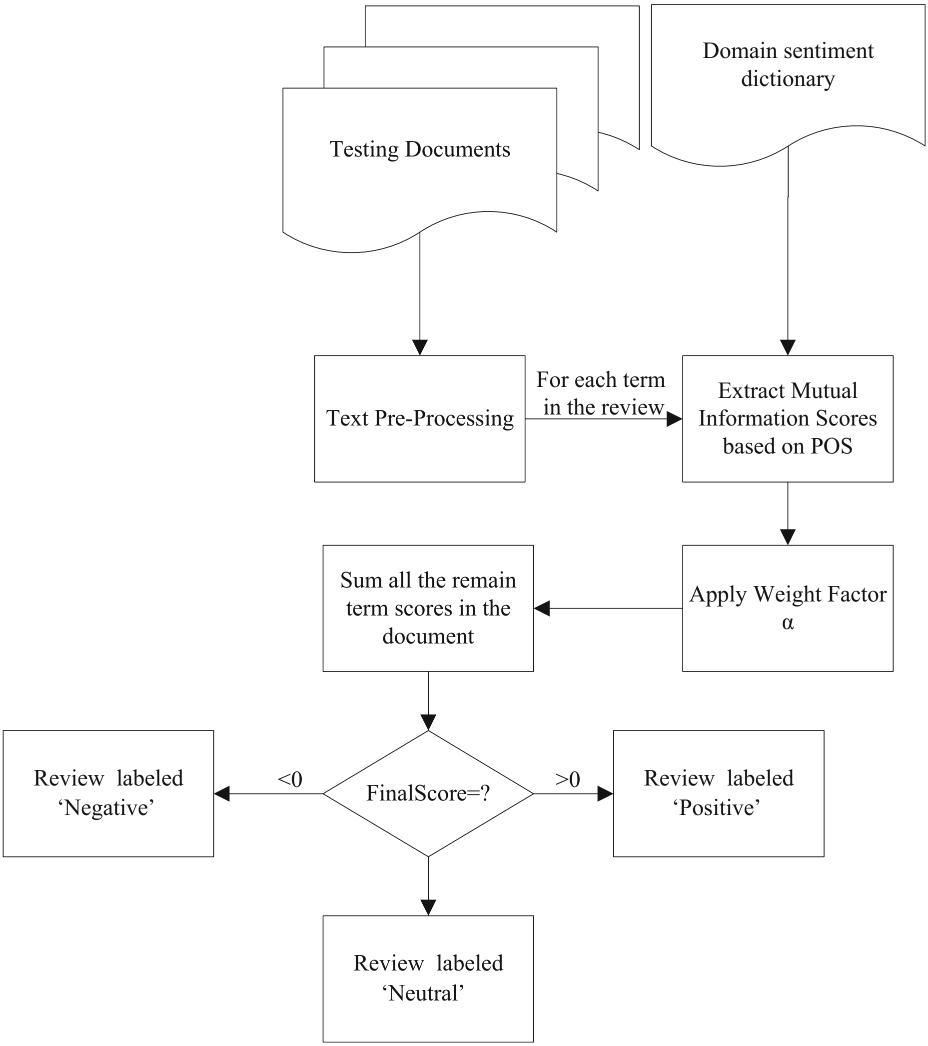


Fig. 2 The proposed sentiment analysis framework

Finally, the score of each term is added to get the final sentiment score of a given review. If the sentiment score is larger than 0, the review is classified as 'Positive'; if the sentiment score is less than 0, the review is classified as 'Negative'; otherwise, the review is classified as 'Neutral'.

4 Results and discussion

4.1 Experimental dataset introduction

In this paper, the Large Movie Review Dataset [12] and the Amazon Product Review Dataset [3] are used as the experimental datasets.

The Large Movie Review Dataset is collected from IMDb, the core dataset contains 50,000 reviews split evenly into 25,000 training and 25,000 testing sets. The overall distribution of labels is balanced(25,000 positive and 25,000 negative reviews) [12]. 10,000 positive and 10,000 negative reviews in the train set are selected as the experimental dataset, the average length of the review is 153. Five experimental datasets are generated by using 5-fold cross-validation. For each experimental dataset, 2000 positive and 2000 negative reviews are taken as the testing set, and the remaining 8000 positive and 8000 negative reviews are used as the training set to generate the domain-special sentiment lexicon proposed in this paper.

Amazon Product Review Dataset: it is a benchmark dataset which is widely used in sentiment analysis. Review datasets for several products are included, collected from Amazon.com. There are 1000 positive and negative reviews in each product review dataset, and large unlabeled reviews. Four product (Books, DVD, Electronics, and Kitchen) review datasets are used in our experiments. An overview of the four datasets is shown in Table 5. The unlabeled reviews are used as the training set and the labeled (positive and negative) reviews are used as the testing set.

4.2 Parameter evaluation

The following experiments show the influence of different values of ω_1 and ω_2 on sentiment classification. Select one of the five data sets generated above as the experimental set.

The influence of ω_1 on the sentiment classification is analyzed, ω_2 is assigned 50%. Figures 3 and 4 show the effect of different values of ω_1 on the sentiment classification.

It can be seen from the experimental results, when the value of ω_1 ranges from 20% to 80%, the method proposed in this paper has achieved better and more stable classification results.

When ω_1 is assigned 10%, 90% and 100%, more reviews are classified as negative ones and we can draw the following observations:

1. In Fig. 3, less positive reviews are detected, which leads to the low recall rate; and less negative reviews are wrongly classified as positive reviews, that is the reason for the raising precision rate.
2. In Fig. 4, more negative reviews are detected, which leads to the high recall rate; but a high proportion of positive reviews are wrongly classified as negative reviews, that is the reason for the dropping precision rate.

When ω_1 is assigned 10% only 800 reviews are selected as the lexicon training data, the less amount of reviews lead to poorly generated sentiment lexicon, therefore, the performance is dropped. When ω_1 is assigned 90% or 100% too many wrongly classified reviews are selected as the lexicon training data, therefore, poorly generated sentiment lexicon get poor performance when it is used in the lexicon-based sentiment classification framework.

Table 5 Four amazon product review datasets

	Unlabeled	Labeled positive	Labeled negative	Average length
Books	975,193	1000	1000	169
DVD	122,438	1000	1000	192
Electronics	21,009	1000	1000	115
Kitchen	17,856	1000	1000	97

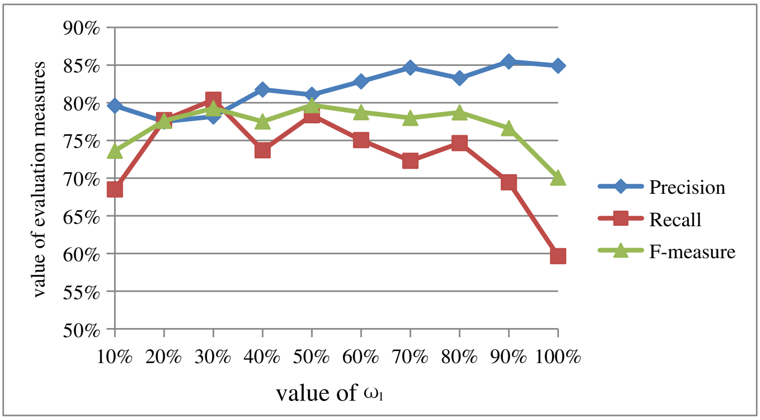


Fig. 3 Effect of ω_1 on the positive review detection

The influence of ω_2 on the sentiment classification is analyzed bellow, ω_2 is assigned 50%. Figures 5 and 6 show the effect of different values of ω_2 on the sentiment classification.

It can be seen from the experimental results, when the value of ω_2 range from 10% to 90%, the method proposed in this paper has achieved better and more stable classification results. When ω_2 is assigned 100% the review polarity detection performs worse.

When ω_2 is assigned 100%, more reviews are classified as positive ones and we can draw the following observations:

1. In Fig. 5, more positive reviews are detected, which leads to the high recall rate; but a high proportion of negative reviews are wrongly classified as positive reviews, that is the reason for the dropping precision rate.
2. In Fig. 6, less negative reviews are detected, which leads to the low recall rate; and less positive reviews are wrongly classified as negative reviews, that is the reason for the raising precision rate.

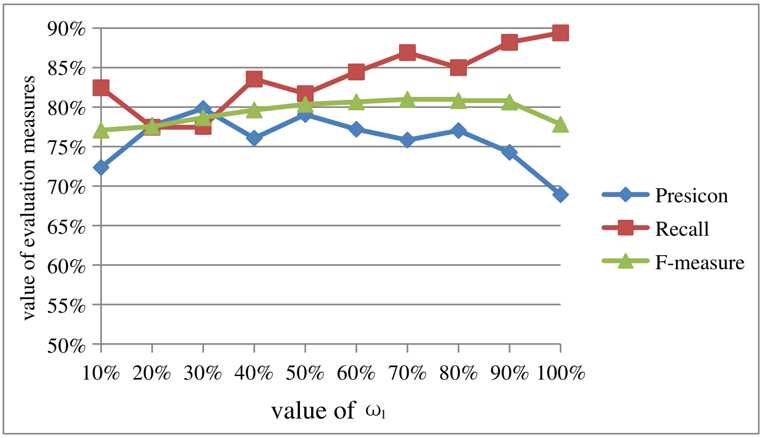


Fig. 4 Effect of ω_1 on the negative review detection

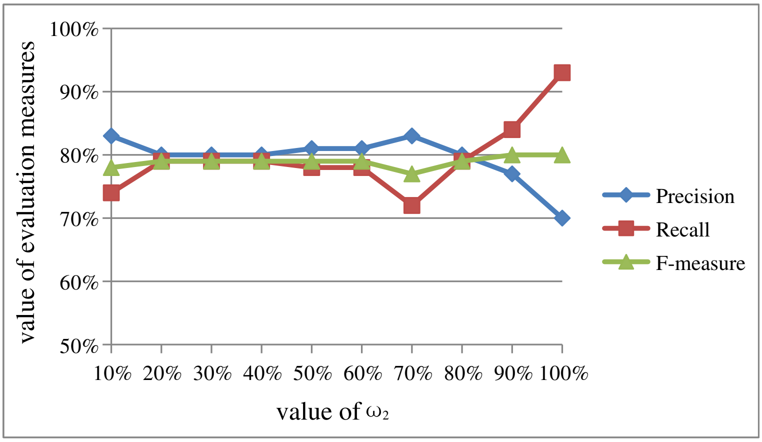


Fig. 5 Effect of ω_2 on the positive review detection

When $\omega_2 = 100\%$, too many wrongly classified reviews are selected as the parameter training data and more reviews are classified as positive reviews, parameter α will get a value closer to 1, lead to testing reviews get more positive than negative sentiment score, therefore more testing reviews are classified as positive reviews, the positive review detection get a high recall rate and negative review detection get a low recall rate. Therefore the polarity detection gets worse performance when $\omega_2 = 100\%$.

4.3 Performance evaluation

The lexicon-based sentiment classifiers which use SentiMI [8] and domain-specific lexicon generated by SenProp [7] are used as the benchmark methods to compare with the method proposed in this paper.

According to the observation of the above experiments about ω_1 and ω_2, we use $\omega_1 = \omega_2 = 50\%$ to test the sentiment analysis method proposed in this paper on the Large Movie Review Dataset, the sentiment classification results are shown in Table 6.

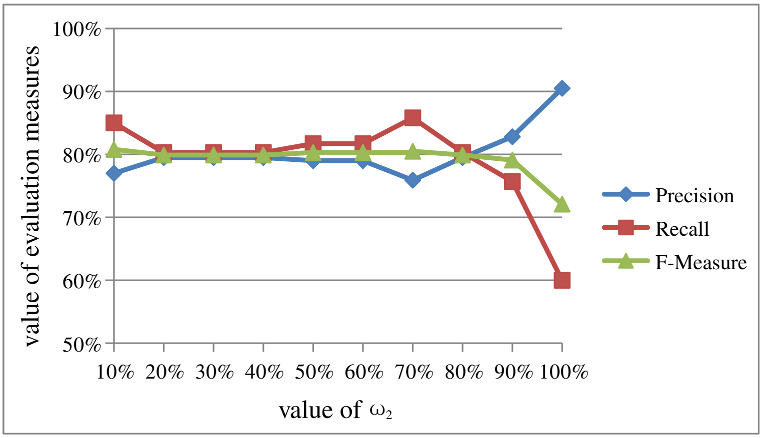


Fig. 6 Effect of ω_2 on the negative review detection

Table 6 5-fold cross validated precision, recall, and F-measure comparison

			Proposed	SentiMI	SentPorp
Positive	Precision	Average	77.50%	71.17%	72.60%
		Improved		6.33%	4.90%
	Recall	Average	75.78%	72.10%	83.45%
		Improved		3.68%	−7.67%
	F-measure	Average	76.52%	71.63%	77.65%
		Improved		4.89%	−1.13%
Negative	Precision	Average	76.41%	71.73%	80.54%
		Improved		4.68%	−4.13
	Recall	Average	77.83%	70.08%	68.50%
		Improved		7.03%	8.73
	F-measure	Average	77.02%	71.26%	74.03%
		Improved		5.76%	2.99%

It can be seen from Table 6 that compare with the SentiMI-based method the proposed one perform well in both positive and negative review detection. F-measure of positive and negative review recognition are increased by 4.89% and 5.76%, respectively. Compare with the SentPorp method, positive review recognition of the proposed method is slightly underperformed by 1.13% of F-measure and negative review recognition is outperformed by 2.99% of F-measure.

Experiments are then carried out on four different Amazon product review datasets. The results are shown in Table 7.

It can be seen from Table 7 that the method proposed in this paper outperforms the SentiMI-based method in the ability of sentiment polarity classification on all four Amazon product review datasets. The negative review recognition ability of the proposed method outperforms the state-of-art method SentProp and the positive review recognition ability of the proposed method and SentProp method is comparable. The domain relevance of the generated sentiment lexicon ensured the performance of our proposed method to review sentiment classification. SentiMI is a sentiment lexicon which extracted from SWN, dictionary glosses are used as the training corpus, the gloss sentiment polarity is consistent with its corresponding entry. It underperforms the proposed domain-specific sentiment lexicon for some words sentiment scores are wrongly assigned and outdated in the specific domain. The proposed method makes use of the unlabeled domain-specific corpus as the training set. So the extracted sentiment lexicon could match the corresponding domain better. Sentiment evaluation

Table 7 Amazon product review polarity classification F-measure comparison

	SentMI		SenProp		Proposed	
	Positive	Negative	Positive	Negative	Positive	Negative
Books	67.78%	62.29%	75.33%	59.41%	74.80%	66.94%
DVD	70.64%	66.49%	75.67%	64.41%	76.28%	68.24%
Electronics	70.58%	67.58%	78.04%	71.39%	76.69%	75.84%
Kitchen	72.58%	68.77%	77.31%	70.63%	77.99%	74.91%

weighting α also determined by the domain-specific corpus, it ensures the proposed sentiment classification could get a good performance. The SentProp method considers the relationships between words by using propagation algorithm, while the induced domain-specific sentiment lexicon tend to be introduced more positive sentiment scores, lead to more words in the lexicon are labeled as positive than negative, therefore the review classification gets a good performance in positive review recognition and worse performance in negative review recognition. The proposed method considers the relationships between words and the sentiment labels directly, therefore the proposed domain-specific lexicon method could induce well sentiment lexicons which could perform well in lexicon-based sentiment polarity classifications. Sentiment evaluation weighting α ensures the domain-specific sentiment lexicon get more balanced performance in the review sentiment polarity classification.

5 Conclusion

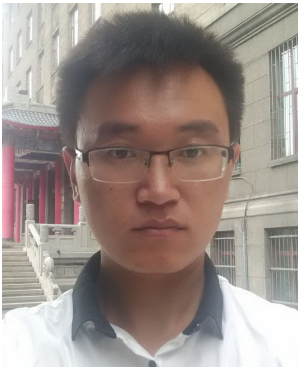
In this paper, we propose a new domain-specific lexicon generation method for review sentiment analysis. It exploits unlabeled domain-specific review datasets to generate domain-specific sentiment lexicons based on the SWN-based sentiment classifier. A completed lexicon-based sentiment classification framework is proposed. Compared with the lexicon-based sentiment classifiers which use SentiMI and domain-specific lexicon generated by SenProp, the method proposed in this paper improves the recognition performance of the negative review greatly, meanwhile gets comparable positive review recognition performance with SentProp method.

Acknowledgements This work is supported by the National Natural Science Foundation of China (No. 61672179, No. 61370083, No. 61402126), the Specialized Research Fund for the Doctoral Program of Higher Education (No.20122304110012), the Heilongjiang Postdoctoral Science Foundation (LBH-Z14071), the Natural Science Foundation for Young Scientists of Heilongjiang Province (QC2016083) and the Natural Science Foundation of Heilongjiang Province (F2015030).

References

1. Awwad H, Alpkocak A, Ieee (2016) Performance comparison of different lexicons for sentiment analysis in arabic. 2016 Third European Network Intelligence Conference (Enic 2016):127–133. https://doi.org/10.1109/enic.2016.25
2. Baccianella S, Esuli A, Sebastiani F (2010) SentiWordNet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. In: LREC, pp 2200–2204
3. Blitzer J, Dredze M, Pereira F (2007) Biographies, Bollywood, Boomboxes and blenders: domain adaptation for sentiment classification. Acl 31(2):187–205
4. Dragut EC, Yu C, Sistla P, Meng W (2010) Construction of a sentimental word dictionary. In: Proceedings of the 19th ACM international conference on Information and knowledge management, 2010. ACM, pp 1761–1764
5. Fellbaum C, Miller G (1998) WordNet: an electronic lexical database. Cognition Brain & Behavior
6. Gatti L, Guerini M (2012) Assessing sentiment strength in words prior polarities. arXiv preprint arXiv: 12124315
7. Hamilton WL, Clark K, Leskovec J, Dan J (2016) Inducing domain-specific sentiment lexicons from unlabeled corpora

8. Khan FH, Qamar U, Bashir S (2016) SentiMI: introducing point-wise mutual information with SentiWordNet to improve sentiment polarity detection. Appl Soft Comput 39:140–153. https://doi.org/10.1016/j.asoc.2015.11.016
9. Leopold, Kindermann, rg (2002) Text categorization with support vector machines. How to represent texts in input space? Mach Learn 46 (1):423–444
10. Liu B (2016) Sentiment analysis: mining opinions, sentiments, and emotions. Comput Linguist 42(3):1–4
11. Lochter JV, Zanetti RF, Reller D, Almeida TA (2016) Short text opinion detection using ensemble of classifiers and semantic indexing. Expert Syst Appl 62:243–249. https://doi.org/10.1016/j.eswa.2016.06.025
12. Maas AL, Daly RE, Pham PT, Huang D, Ng AY, Potts C (2011) Learning word vectors for sentiment analysis. In: The Meeting of the Association for Computational Linguistics: Human Language Technologies, Proceedings of the Conference, 19–24 June, Portland, Oregon, USA, 2011. pp 142–150
13. Natural Language Toolkit. http://nltk.org/
14. Park S, Lee W, Moon IC (2015) Efficient extraction of domain specific sentiment lexicon with active learning. Pattern Recogn Lett 56:38–44. https://doi.org/10.1016/j.patrec.2015.01.004
15. Petz G, Karpowicz M, Fürschuß H, Auinger A, Stříteský V, Holzinger A (2014) Computational approaches for mining user's opinions on the web 2.0. Inf Process Manag 50(6):899–908
16. Saif H, He YL, Fernandez M, Alani H (2016) Contextual semantics for sentiment analysis of twitter. Inf Process Manag 52(1):5–19. https://doi.org/10.1016/j.ipm.2015.01.005
17. Taboada M, Brooke J, Tofiloski M, Voll K, Stede M (2011) Lexicon-based methods for sentiment analysis. Comput Linguist 37(2):267–307
18. Tang D, Wei F, Qin B, Zhou M, Liu T (2014) Building large-scale twitter-specific sentiment lexicon: a representation learning approach. In: COLING, pp 172–182
19. Toutanova K, Klein D, Manning CD, Singer Y (2003) Feature-rich part-of-speech tagging with a cyclic dependency network. In: Conference of the North American chapter of the Association for Computational Linguistics on human language Technology. Association for Computational Linguistics, Edmonton, pp 173–180
20. Wilson T, Wiebe J, Hoffmann P (2005) Recognizing contextual polarity in phrase-level sentiment analysis. Paper presented at the proceedings of the conference on human language technology and empirical methods in natural language processing. Vancouver, British Columbia



Hongyu Han is a PhD candidate at the Department of Computer Science and Technology of Harbin Engineering University (China). His interests include: sentiment analysis and text mining.

Jianpei Zhang PhD in Computer Science, is a professor at the Department of Computer Science and Technology of Harbin Engineering University (China). His interests include: new technology of database, big data analysis and social computing, information security and privacy protection. He has published more than 100 papers in distinguished journals and conferences.

Jing Yang PhD in Computer Science, is a professor at the Department of Computer Science and Technology of Harbin Engineering University (China). Her interests include: big data analysis, privacy protection, social networking, information security and scheduling algorithm. Her has published more than 100 papers in distinguished journals and conferences. She is member of CCF.

Yiran Shen PhD in Computer Science, is an associate professor at the Department of Computer Science and Technology of Harbin Engineering University (China). His interests include: wearable computing, machine learning, security in mobile/cloud computing and crowdsourcing, and underwater sensor networks. He has published 15 papers in distinguished journals and conferences. He is member of IEEE, ACM and CCF.

Yongshi Zhang PhD in Computer Science, is a postdoctoral researcher at the Department of Computer Science and Technology of Harbin Engineering University (China). His interests include: big data analysis and social networking.