View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Open Research Online

Open Research Online

The Open University's repository of research publications
and other research outputs

Exploring English lexicon knowledge for Chinese sentiment analysis

Conference or Workshop Item

How to cite:

He, Yulan; Alani, Harith and Zhou, Deyu (2010). Exploring English lexicon knowledge for Chinese sentiment analysis. In: CIPS-SIGHAN Joint Conference on Chinese Language Processing, 28-29 Aug 2010, Beijing, China.

For guidance on citations see FAQs.

© 2010 The Authors

Version: Version of Record

Link(s) to article on publisher's website:
http://wing.comp.nus.edu.sg/ antho/sighan.html

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data policy on reuse of materials please consult the policies page.

oro.open.ac.uk

Exploring English Lexicon Knowledge for Chinese Sentiment Analysis

Yulan He Harith Alani
Knowledge Media Institute
The Open University
Milton Keynes MK6 6AA, UK
{y.he, h.alani}@open.ac.uk

Deyu Zhou
School of Computer Science and Engineering
Southeast University
Nanjing, China
d.zhou@seu.edu.cn

Abstract

This paper presents a weakly-supervised method for Chinese sentiment analysis by incorporating lexical prior knowledge obtained from English sentiment lexicons through machine translation. A mechanism is introduced to incorporate the prior information about polarity-bearing words obtained from existing sentiment lexicons into latent Dirichlet allocation (LDA) where sentiment labels are considered as topics. Experiments on Chinese product reviews on mobile phones, digital cameras, MP3 players, and monitors demonstrate the feasibility and effectiveness of the proposed approach and show that the weakly supervised LDA model performs as well as supervised classifiers such as Naive Bayes and Support vector Machines with an average of 83% accuracy achieved over a total of 5484 review documents. Moreover, the LDA model is able to extract highly domain-salient polarity words from text.

1 Introduction

Sentiment analysis aims to understand subjective information such as opinions, attitudes, and feelings expressed in text. It has become a hot topic in recent years because of the explosion in availability of people's attitudes and opinions expressed in social media including blogs, discussion forums, tweets, etc. Research in sentiment analysis has mainly focused on the English language. There have been few studies in sentiment analysis in other languages due to the lack of resources, such as subjectivity lexicons consisting of a list of words marked with their respective polarity (positive, negative or neutral) and manually labeled subjectivity corpora with documents labeled with their polarity.

Pilot studies on cross-lingual sentiment analysis utilize machine translation to perform sentiment analysis on the English translation of foreign language text (Banea et al., 2008; Bautin et al., 2008; Wan, 2009). The major problem is that they cannot be generalized well when there is a domain mismatch between the source and target languages. There have also been increasing interests in exploiting bootstrapping-style approaches for weakly-supervised sentiment classification in languages other than English (Zagibalov and Carroll, 2008b; Zagibalov and Carroll, 2008a; Qiu et al., 2009). Other approaches use ensemble techniques by either combining lexicon-based and corpus-based algorithms (Tan et al., 2008) or combining sentiment classification outputs from different experimental settings (Wan, 2008). Nevertheless, all these approaches are either complex or require careful tuning of domain and data specific parameters.

This paper proposes a weakly-supervised approach for Chinese sentiment classification by incorporating language-specific lexical knowledge obtained from available English sentiment lexicons through machine translation. Unlike other cross-lingual sentiment classification methods which often require labeled corpora for training and therefore hinder their applicability for cross-domain sentiment analysis, the proposed approach does not require labeled documents. Moreover, as opposed to existing weakly-supervised sentiment classification approaches which are rather complex, slow, and require careful parameter tuning, the proposed approach is simple and computationally efficient; rendering more suitable for online and real-time sentiment classification from the Web.

Our experimental results on the Chinese reviews of four different product types show that the LDA model performs as well as the supervised classifiers such as Naive Bayes and Support Vector Machines trained from labeled corpora. Although this paper primarily studies sentiment analysis in Chinese, the proposed approach is applicable to any other language so long as a machine translation engine is available between the selected language and English.

The remainder of the paper is organized as follows. Related work on cross-lingual sentiment classification and weakly-supervised sentiment classification in languages other than English are discussed in Section 2. The proposed mechanism of incorporating prior word polarity knowledge into the LDA model is introduced in Section 3. The experimental setup and results of sentiment classification on the Chinese reviews of four different products are presented in Section 4 and 5 respectively. Finally, Section 6 concludes the paper.

2 Related Work

Pilot studies on cross-lingual sentiment analysis rely on English corpora for subjectivity classification in other languages. For example, Mihalcea et al. (2007) make use of a bilingual lexicon and a manually translated parallel text to generate the resources to build subjectivity classifiers based on Support Vector Machines (SVMs) and Naive Bayes (NB) in a new language; Banea et al. (2008) use machine translation to produce a corpus in a new language and train SVMs and NB for subjectivity classification in the new language. Bautin et al. (2008) also utilize machine translation to perform sentiment analysis on the English translation of a foreign language text.

More recently, Wan (2009) proposed a co-training approach to tackle the problem of cross-lingual sentiment classification by leveraging an available English corpus for Chinese sentiment classification. Similar to the approach proposed in (Banea et al., 2008), Wan's method also uses machine translation to produced a labeled Chinese review corpus from the available labeled English review data. However, in order to alleviate the language gap problem that the underlying distributions between the source and target language are different, Wan builds two SVM classifiers, one based on English features and the other based on Chinese features, and uses a bootstrapping method based on co-training to iteratively improve classifiers until convergence.

The major problem of the aforementioned cross-lingual sentiment analysis algorithms is that they all utilize supervised learning to train sentiment classifiers from annotated English corpora (or the translated target language corpora generated by machine translation). As such, they cannot be generalized well when there is a domain mismatch between the source and target language. For example, For example, the word 'compact' might express positive polarity when used to describe a digital camera, but it could have negative orientation if it is used to describe a hotel room. Thus, classifiers trained on one domain often fail to produce satisfactory results when shifting to another domain.

Recent efforts have also been made for weakly-supervised sentiment classification in Chinese. Zagibalov and Carroll (2008b) starts with a one-word sentiment seed vocabulary and use iterative retraining to gradually enlarge the seed vocabulary by adding more sentiment-bearing lexical items based on their relative frequency in both the positive and negative parts of the current training data. Sentiment direction of a document is then determined by the sum of sentiment scores of all the sentiment-bearing lexical items found in the document. The problem with this approach is that there is no principal way to set the optimal number of iterations. They then suggested an iteration control method in (Zagibalov and Carroll, 2008a) where iterative training stops when there is no change to the classification of any document over the previous two iterations. However, this does not necessarily correlate to the best classification accuracy.

Similar to (Zagibalov and Carroll, 2008b), Qiu et al. (2009) also uses a lexicon-based iterative process as the first phase to iteratively enlarge an initial sentiment dictionary. But instead of using a one-word seed dictionary as in (Zagibalov and Carroll, 2008b), they started with a much larger HowNet Chinese sentiment dictionary[1] as the initial lexicon. Documents classified by the first phase are taken as the training set to train the SVMs which are subsequently used to revise the results produced by the first phase.

Other researchers investigated ensemble techniques for weakly-supervised sentiment classification. Tan et al. (2008) proposed a combination of lexicon-based and corpus-based approaches that first labels some examples from a give domain using a sentiment lexicon and then trains a supervised classifier based on the labeled ones from the first stage. Wan (2008) combined sentiment scores calculated from Chinese product reviews using the Chinese HowNet sentiment dictionary and from the English translation of Chinese reviews using the English MPQA subjectivity lexicon[2]. Various weighting strategies were explored to combine sentiment classification outputs from different experimental settings in order to improve classification accuracy.

Nevertheless, all these weakly-supervised sentiment classification approaches are rather complex and require either iterative training or careful tuning of domain and data specific parameters, and hence unsuitable for online and real-time sentiment analysis in practical applications.

3 Incorporating Prior Word Polarity Knowledge into LDA

Unlike existing approaches, we view sentiment classification as a generative problem that when an author writes a review document, he/she first decides on the overall sentiment or polarity (positive, negative, or neutral) of a document, then for each sentiment, decides on the words to be used. We use LDA to model a mixture of only three topics or sentiment labels, i.e. positive, negative and neutral.

Assuming that we have a total number of S sentiment labels; a corpus with a collection of D documents is denoted by $C = \{d_1, d_2, ..., d_D\}$; each document in the corpus is a sequence of N_d words denoted by $d = (w_1, w_2, ..., w_{N_d})$, and each word in the document is an item from a vocabulary index with V distinct terms denoted by $\{1, 2, ..., V\}$. The generative process is as follows:

- Choose distributions $\varphi \sim Dir(\beta)$.
- For each document $d \in [1, D]$, choose distributions $\pi_d \sim Dir(\gamma)$.
- For each of the N_d word position w_t, choose a sentiment label $l_t \sim Multinomial(\pi_d)$, and then choose a word $w_t \sim Multinomial(\varphi_{l_t})$.

The joint probability of words and sentiment label assignment in LDA can be factored into two terms:

$$P(\mathbf{w}, \mathbf{l}) = P(\mathbf{w}|\mathbf{l})P(\mathbf{l}|d). \quad (1)$$

Letting the superscript $-t$ denote a quantity that excludes data from the t^{th} position, the conditional posterior for l_t by marginalizing out the random variables φ and π is

$$P(l_t = k|\mathbf{w}, \mathbf{l^{-t}}, \beta, \boldsymbol{\gamma}) \propto \frac{N^{-t}_{w_{t,k}} + \beta}{N^{-t}_k + V\beta} \times \frac{N^{-t}_{k,d} + \gamma_k}{N^{-t}_d + \sum_k \gamma_k}, \quad (2)$$

where $N_{w_t,k}$ is the number of times word w_t has associated with sentiment label k; N_k is the the number of times words in the corpus assigned to sentiment label k; $N_{k,d}$ is the number of times sentiment label k has been assigned to some word tokens in document d; N_d is the total number of words in the document collection.

Each words in documents can either bear positive polarity ($l_t = 1$), or negative polarity ($l_t = 2$), or is neutral ($l_t = 0$). We now show how to incorporate polarized words in sentiment lexicons as prior information in the Gibbs sampling process. Let

$$Q_{t,k} = \frac{N^{-t}_{w_{t,k}} + \beta}{N^{-t}_k + V\beta} \times \frac{N^{-t}_{k,d} + \gamma_k}{N^{-t}_d + \sum_k \gamma_k} \quad (3)$$

[1] http://www.keenage.com/download/sentiment.rar

[2] http://www.cs.pitt.edu/mpqa/

We can then modify the Gibbs sampling equation as follows:

$$P(l_t = k|\mathbf{w}, \mathbf{l^{-t}}, \beta, \boldsymbol{\gamma}) \propto \begin{cases} \mathbb{I}(k = S(w_t)) \times Q_{t,k} & \text{if } S(w_t) \text{ is defined} \\ Q_{t,k} & \text{otherwise} \end{cases} \quad (4)$$

where the function $S(w_t)$ returns the prior sentiment label of w_t in a sentiment lexicon and it is defined if word w_t is found in the sentiment lexicon. $\mathbb{I}(k = S(w_t))$ is an indicator function that takes on value 1 if $k = S(w_t)$ and 0 otherwise.

Equation 4 in fact applies a hard constraint that when a word is found in a sentiment lexicon, its sampled sentiment label is restricted to be the same as its prior sentiment label defined in the lexicon. This constraint can be relaxed by introducing a parameter to control the strength of the constraint such that when word w_t is found in the sentiment lexicon, Equation 4 becomes

$$P(l_t = k|\mathbf{w}, \mathbf{l^{-t}}, \beta, \boldsymbol{\gamma}) \propto (1 - \lambda) \times Q_{t,k} + \lambda \times \mathbb{I}(k = S(w_t)) \times Q_{t,k} \quad (5)$$

where $0 \leq \lambda \leq 1$. When $\lambda = 1$, the hard constraint will be applied; when $\lambda = 0$, Equation 5 is reduced to the original unconstrained Gibbs sampling as defined in Equation 2.

While sentiment prior information is incorporated by modifying conditional probabilities used in Gibbs sampling here, it is also possible to explore other mechanisms to define expectation or posterior constraints, for example, using the generalized expectation criteria (McCallum et al., 2007) to express preferences on expectations of sentiment labels of those lexicon words. We leave the exploitation of other mechanisms of incorporating prior knowledge into model training as future work.

The document sentiment is classified based on $P(\mathbf{l}|d)$, the probability of sentiment label given document, which can be directly obtained from the document-sentiment distribution. We define that a document d is classified as positive if $P(\mathbf{l_{pos}}|d) > P(\mathbf{l_{neg}}|d)$, and vice versa.

Table 2: Data statistics of the four Chinese product reviews corpora.

	No. of Reviews		*Vocab*
Corpus	*positive*	*Negative*	*Size*
Mobile	1159	1158	8945
DigiCam	853	852	5716
MP3	390	389	4324
Monitor	341	342	4712

4 Experimental Setup

We conducted experiments on the four corpora[3] which were derived from product reviews harvested from the website IT168[4] with each corresponding to different types of product reviews including mobile phones, digital cameras, MP3 players, and monitors. All the reviews were tagged by their authors as either positive or negative overall. The statistics of the four corpora are shown in Table 2.

We explored three widely used English sentiment lexicons in our experiments, namely the MPQA subjectivity lexicon, the appraisal lexicon[5], and the SentiWordNet[6] (Esuli and Sebastiani, 2006). For all these lexicons, we only extracted words bearing positive or negative polarities and discarded words bearing neutral polarity. For SentiWordNet, as it consists of words marked with positive and negative orientation scores ranging from 0 to 1, we extracted a subset of 8,780 opinionated words, by selecting those whose orientation strength is above a threshold of 0.6.

We used Google translator toolkit[7] to translate these three English lexicons into Chinese. After translation, duplicate entries, words that failed to translate, and words with contradictory polarities were removed. For comparison, we also tested a Chinese sentiment lexicon, NTU Sentiment Dictionary (NTUSD)[8] (Ku and Chen, 2007) which

[3] `http://www.informatics.sussex.ac.uk/users/tz21/dataZH.tar.gz`

[4] `http://product.it168.com`

[5] `http://lingcog.iit.edu/arc/appraisal_lexicon_2007b.tar.gz`

[6] `http://sentiwordnet.isti.cnr.it/`

[7] `http://translate.google.com`

[8] `http://nlg18.csie.ntu.edu.tw:`

Table 1: Matched polarity words statistics (positive/negative).

Lexicon	*Chinese*				*English*			
	Mobile	DigiCam	MP3	Monitors	Mobile	DigiCam	MP3	Monitors
(a)MPQA	261/253	183/174	162/135	169/147	293/331	220/241	201/153	210/174
(b)Appraisal	279/165	206/127	180/104	198/105	392/271	330/206	304/153	324/157
(c)SentiWN	304/365	222/276	202/213	222/236	394/497	306/397	276/310	313/331
(d)NTUSD	338/319	263/242	239/167	277/241	–			
(a)+(c)	425/465	307/337	274/268	296/289	516/607	400/468	356/345	396/381
(a)+(b)+(c)	495/481	364/353	312/280	344/302	624/634	496/482	447/356	494/389
(a)+(c)+(d)	586/608	429/452	382/336	421/410	–			

was automatically generated by enlarging an initial manually created seed vocabulary by consulting two thesauri, tong2yi4ci2ci2lin2 and the Academia Sinica Bilingual Ontological WordNet 3.

Chinese word segmentation was performed on the four corpora using the conditional random fields based Chinese Word Segmenter[9]. The total numbers of matched polarity words in each corpus using different lexicon are shown in Table 1 with the left half showing the statistics against the Chinese lexicons (the original English lexicons have been translated into Chinese) and the right half listing the statistics against the English lexicons. We did not translate the Chinese lexicon NTUSD into English since we focused on Chinese sentiment classification here. It can be easily seen from the table that in general the matched positive words outnumbered the matched negative words using any single lexicon except SentiWordNet. But the combination of the lexicons results in more matched polarity words and thus gives more balanced number of positive and negative words. We also observed the increasing number of the matched polarity words on the translated English corpora compared to their original Chinese corpora. However, as will be discussed in Section 5.2 that the increasing number of the matched polarity words does not necessarily lead to the improvement of the sentiment classification accuracy.

We modified GibbsLDA++ package[10] for the model implementation and only used hard constraints as defined in Equation 4 in our experiments. The word prior polarity information was also utilized during the initialization stage that if a word can be found in a sentiment lexicon, the word token is assigned with its corresponding sentiment label. Otherwise, a sentiment label is randomly sampled for the word. Symmetric Dirichlet prior β was used for sentiment-word distribution and was set to 0.01, while asymmetric Dirichlet prior γ was used for document-sentiment distribution and was set to 0.01 for positive and neutral sentiment labels, and 0.05 for negative sentiment label.

5 Experimental Results

This section presents the experimental results obtained under two different settings: LDA model with translated English lexicons tested on the original Chinese product review corpora; and LDA model with original English lexicons tested on the translated product review corpora.

5.1 Results with Different Sentiment Lexicons

Table 3 gives the classification accuracy results using the LDA model with prior sentiment label information provided by different sentiment lexicons. Since we did not use any labeled information, the accuracies were averaged over 5 runs and on the whole corpora. For comparison purposes, we have also implemented a baseline model which simply assigns a score +1 and -1 to any matched positive and negative word respectively based on a sentiment lexicon. A review document is then classified as either positive or negative according to the aggregated sentiment scores. The baseline results were shown in brackets in Table 3 .

8080/opinion/pub1.html

[9]http://nlp.stanford.edu/software/stanford-chinese-segmenter-2008-05-21.tar.gz

[10]http://gibbslda.sourceforge.net/

Table 3: Sentiment classification accuracy (%) by LDA, numbers in brackets are baseline results.

Lexicon	*Mobile*	*DigiCam*	*MP3*	*Monitors*	*Average*
(a)MPQA	82.00 (63.53)	80.93 (67.59)	78.31 (68.42)	81.41 (64.86)	80.66 (66.10)
(b)Appraisal	71.95 (56.28)	80.46 (60.54)	77.28 (61.36)	80.67 (57.98)	77.59 (59.04)
(c)SentiWN	81.10 (62.45)	78.52 (57.13)	79.08 (64.57)	75.55 (55.34)	78.56 (59.87)
(d)NTUSD	82.61 (71.21)	78.70 (68.23)	78.69 (75.87)	84.63 (74.96)	81.16 (72.57)
(a)+(c)	81.18 (65.95)	78.70 (65.18)	83.83 (67.52)	80.53 (62.08)	81.06 (65.18)
(a)+(b)+(c)	81.48 (62.84)	80.22 (65.88)	80.23 (65.60)	78.62 (61.35)	80.14 (63.92)
(a)+(c)+(d)	82.48 (69.96)	84.33 (69.58)	83.70 (71.12)	82.72 (65.59)	83.31 (69.06)
Naive Bayes	86.52	82.27	82.64	86.21	84.41
SVMs	84.49	82.04	79.43	83.87	82.46

It can be observed from Table 3 that the LDA model performs significantly better than the baseline model. The improvement ranges between 9% and 19% and this roughly corresponds to how much the model learned from the data. We can thus speculate that LDA is indeed able to learn the sentiment-word distributions from data.

Translated English sentiment lexicons perform comparably with the Chinese sentiment lexicon NTUSD. As for the individual lexicon, using MPQA subjectivity lexicon gives the best result among all the English lexicons on all the corpora except the MP3 corpus where MPQA performs slightly worse than SentiWordNet. The combination of MPQA and SentiWordNet performs significantly better than other lexicons on the MP3 corpus, with almost 5% improvement compared to the second best result. We also notice that the combination of all the three English lexicons does not lead to the improvement of classification accuracy which implies that the quality of a sentiment lexicon is indeed important to sentiment classification. The above results suggest that in the absence of any Chinese sentiment lexicon, MPQA subjectivity lexicon appears to be the best candidate to be used to provide sentiment prior information to the LDA model for Chinese sentiment classification.

We also conducted experiments by including the Chinese sentiment lexicon NTUSD and found that the combination of MPQA, SentiWordNet, and NTUSD gives the best overall classification accuracy with 83.31% achieved. For comparison purposes, we list the 10-fold cross validation results obtained using the supervised classifiers, Naive Bayes and SVMs, trained on the labeled corpora as previously reported in (Zagibalov and Carroll, 2008a). It can be observed that using only English lexicons (the combination of MPQA and SentiWordNet), we obtain better results than both NB and SVMs on the MP3 corpus. With an additional inclusion of NTUSD, LDA outperforms NB and SVMs on both DigiCam and MP3. Furthermore, LDA gives a better overall accuracy when compared to SVMs. Thus, we may conclude that the unsupervised LDA model performs as well as the supervised classifiers such as NB and SVMs on the Chinese product review corpora.

5.2 Results with Translated Corpora

We ran a second set of experiments on the translated Chinese product review corpora using the original English sentiment lexicons. Both the translated corpora and the sentiment lexicons have gone through stopword removal and stemming in order to reduce the vocabulary size and thereby alleviate data sparseness problem. It can be observed from Figure 1 that in general sentiment classification on the original Chinese corpora using the translated English sentiment lexicons gives better results than classifying on the translated review corpora using the original English lexicons on both the Mobile and Digicam corpora. However, reversed results are observed on the Monitor corpus that classifying on the translated review corpus using the English sentiment lexicons outperforms classifying on the

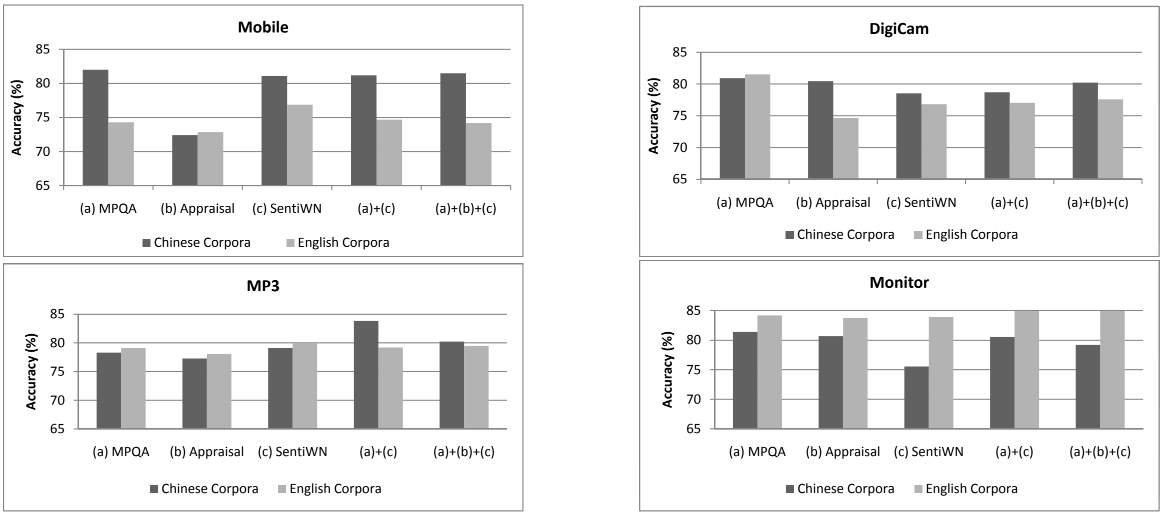

Figure 1: Comparison of the performance on the Chinese corpora and their translated corpora in English.

original Chinese review corpus using the translated sentiment lexicons. In particular, the combination of the MPQA subjectivity lexicon and SentiWordNet gives the best result of 84% on the Monitor corpus. As for the MP3 corpus, classifying on the original Chinese reviews or on the translated reviews does not differ much except that a better result is obtained on the Chinese corpus when using the combination of the MPQA subjectivity lexicon and SentiWordNet. The above results can be partially explained by the ambiguities and changes of meanings introduced in the translation. The Mobile and DigiCam corpora are relatively larger than the MP3 and Monitors corpora and we therefore expect more ambiguities being introduced which might result in the change of document polarities.

5.3 Extracted Polarity-Bearing Words

LDA is able to extract polarity-bearing words. Table 4 lists some of the polarity words identified by the LDA model which are not found in the original sentiment lexicons. We can see that LDA is indeed able to recognize domain-specific positive or negative words, for example, 蓝牙 (bluetooth) for mobile phones, 小巧 (compact) for digital cameras, 金属 (metallic) for MP3, 纯平 (flat screen) and 变形 (deformation) for monitors.

The iterative approach proposed in (Zagibalov and Carroll, 2008a) can also automatically acquire polarity words from data. However, it appears that only positive words were identified by their approach. Our proposed LDA model can extract both positive and negative words and most of them are highly domain-salient as can be seen from Table 4.

6 Conclusions

This paper has proposed a mechanism to incorporate prior information about polarity words from English sentiment lexicons into LDA model learning for weakly-supervised Chinese sentiment classification. Experimental results of sentiment classification on Chinese product reviews show that in the absence of a language-specific sentiment lexicon, the translated English lexicons can still produce satisfactory results with the sentiment classification accuracy of 81% achieved averaging over four different types of product reviews. With the incorporation of the Chinese sentiment lexicon NTUSD, the classification accuracy is further improved to 83%. Compared to the existing approaches to cross-lingual sentiment classification which either rely on labeled corpora for classifier learning or iterative training for performance gains, the proposed approach is simple and readily to

Table 4: Extracted example polarity words by LDA.

Corpus	*Positive*	*Negative*
Mobile	优点 (advantage), 大 (large), 好用 (easy to use), 快 (fast), 舒服 (comfortable), 蓝牙 (bluetooth), 新 (new), 容易 (easy)	坏 (bad), 差 (poor), 慢 (slow), 没 (no;not), 难 (difficult;hard), 少 (less), 只是 (but), 修 (repair)
DigiCam	优点 (advantage), 小巧 (compact), 强 (strong;strength), 长焦 (telephoto), 动态 (dynamic), 全 (comprehensive), 专业 (professional), 上手 (get started)	后悔 (regret), 坏 (bad), 差 (poor), 慢 (slow), 暗 (dark), 贵 (expensive), 难 (difficult;hard), 耗电 (consume much electricity), 塑料 (plastic), 修 (repair)
MP3	小巧 (compact), 快 (fast), 强 (strong;strength), 更 (even), 质感 (textual), 全 (comprehensive), 金属 (metallic),十分 (very)	不 (no;not), 差 (poor), 坏 (bad), 有点 (rather), 根本 (simply), 次 (substandard), 死机 (crash), 没 (no), 但是 (but)
Monitors	容易 (easy), 新 (new), 纯平 (flat screen), 舒服 (comfortable), 显亮 (looks bright), 锐利 (sharp), 亮 (bright), 自动 (automatic)	变形 (deformation), 偏色 (color cast bad), 坏 (bad), 差 (poor), 没 (no;not), 漏光 (leakage of light), 黑屏 (black screen), 退 (refund;return), 暗 (dark), 抖动 (jitter)

be used for online and real-time sentiment classification from the Web.

One issue relating to the proposed approach is that it still depends on the quality of machine translation and the performance of sentiment classification is thus affected by the language gap between the source and target language. A possible way to alleviate this problem is to construct a language-specific sentiment lexicon automatically from data and use it as the prior information source to be incorporated into the LDA model learning.

References

Banea, C., R. Mihalcea, J. Wiebe, and S. Hassan. 2008. Multilingual subjectivity analysis using machine translation. In *Proceedings of the EMNLP*, pages 127–135.

Bautin, M., L. Vijayarenu, and S. Skiena. 2008. International sentiment analysis for news and blogs. In *Proceedings of the International Conference on Weblogs and Social Media (ICWSM)*.

Esuli, A. and F. Sebastiani. 2006. SentiWordNet: A publicly available lexical resource for opinion mining. In *Proceedings of LREC*, volume 6.

Ku, L.W. and H.H. Chen. 2007. Mining opinions from the Web: Beyond relevance retrieval. *Journal of the American Society for Information Science and Technology*, 58(12):1838–1850.

McCallum, A., G. Mann, and G. Druck. 2007. Generalized expectation criteria. Technical Report 2007-60, University of Massachusetts Amherst.

Mihalcea, R., C. Banea, and J. Wiebe. 2007. Learning multilingual subjective language via crosslingual projections. In *Proceedings of the ACL*, pages 976–983.

Qiu, L., W. Zhang, C. Hu, and K. Zhao. 2009. Selc: a self-supervised model for sentiment classification. In *Proceeding of the CIKM*, pages 929–936.

Tan, S., Y. Wang, and X. Cheng. 2008. Combining learn-based and lexicon-based techniques for sentiment detection without using labeled examples. In *Proceedings of the SIGIR*, pages 743–744.

Tseng, H., P. Chang, G. Andrew, D. Jurafsky, and C. Manning. 2005. A conditional random field word segmenter. In *Fourth SIGHAN Workshop on Chinese Language Processing*, volume 37.

Wan, X. 2008. Using bilingual knowledge and ensemble techniques for unsupervised Chinese sentiment analysis. In *Proceedings of the EMNLP*, pages 553–561.

Wan, X. 2009. Co-training for cross-lingual sentiment classification. In *Proceedings of the ACL*, pages 235–243.

Zagibalov, T. and J. Carroll. 2008a. Automatic seed word selection for unsupervised sentiment classification of Chinese text. In *Proceedings of the COLING*, pages 1073–1080.

Zagibalov, T. and J. Carroll. 2008b. Unsupervised classification of sentiment and objectivity in chinese text. In *Proceedings of the IJCNLP*, pages 304–311.