A Cross-Domain Sentiment Classification Method Based on Extraction of Key Sentiment Sentence

Shaowu Zhang[1,2], Huali Liu[1], Liang Yang[1], and Hongfei Lin[1(✉)]

[1] School of Computer Science and Technology, Dalian University of Technology, Dalian 116024, China
{zhangsw,hflin}@dlut.edu.cn, {liuhuali,yangliang}@mail.dlut.edu.cn
[2] School of Computer Science and Engineering, Xinjinag University of Finance and Economics, Urumqi 830012, China

Abstract. Cross-domain sentiment analysis focuses on these problems where the source domain and the target domain are from different domains. However, traditional sentiment classification approaches usually perform poorly to address cross-domain problems. So, this paper proposed a cross-domain sentiment classification method based on extraction of key sentiment sentence. Firstly, based on the observation that not every part of the document is equally informative for inferring the sentiment orientation of the whole document, the concept of key sentiment sentence was defined. Secondly, taking advantage of three properties: sentiment purity, keyword property and position property, we construct heuristic rules, and combine with machine learning to extract key sentiment sentence. Then, data is divided into key and detail views. Integrating two views effectively can improve performance. Finally, experimental results show the superiority of our proposed method.

Keywords: Cross-domain · Key sentiment sentence · Multi-view ensemble

1 Introduction

With the rapid development of the Internet and the rise of social networking platforms, more and more people choose to express their views and opinions on the network. Subsequently, a large amount of text data with potential value emerges. Mining sentiment orientation of these data is of importance for scientific significance and commercial value. Under this situation, sentiment analysis has become one of the most active research areas in natural language processing and attracted a lot of research interest in recent years.

Sentiment analysis is the process of automatically detecting whether a text expresses a positive or negative semantic orientation. On this issue of vital importance, traditional supervised machine learning approaches have been shown promising and effectiveness [1–3]. However, these successful studies are based on

J. Li et al. (Eds.): NLPCC 2015, LNAI 9362, pp. 90–101, 2015.
DOI: 10.1007/978-3-319-25207-0_8

an assumption, namely training data and test data are independent and identically distributed. But, in practice, data drawn from different domains is difficult to meet the independent and identically distributed condition. For example, in book reviews, "*obscure*", "*relevant*" show strong sentiment intensity, while rarely appear in hotel reviews. And equally, "*tatty*", "*comfortable*" are usually used to comment on hotel and indicate strong sentiment orientation, but hardly occur in book reviews. Thus, how to solve the problem that the performance of classifier reduced caused by the difference of the distributions of training data and testing data, is a core task of cross-domain sentiment analysis.

In cross-domain sentiment analysis field, for supervised learning techniques, most classifiers are domain-specific. So, it becomes challenging to adapt a classifier trained on one domain to another. To address this problem, Blitzer et al. [4] proposed structural correspondence learning (SCL). *SCL*, selected a set of pivot features to model the relationship between pivot features and non-pivot features and to link the source and target domains. Also, Bollegala et al. [5] proposed a feature-level Sentiment Sensitive Thesaurus(SST) method. Firstly, they created a sentiment sensitive thesaurus. And then used the thesaurus to expand feature vectors for classification. Finally experiment proved the created sentiment sensitive thesaurus can accurately capture words that express similar sentiments. In addition, Li et al. [6] performed active learning for cross-domain sentiment analysis. They selected informative samples based on *QBC* strategy by the source and target classifiers. And combination-based to make decision. At last results showed comparable performances to in-domain method.

Follow the ideas of mentioned paper above, we proposed a cross-domain sentiment classification method based on extraction of *key sentiment sentence*. We divide a document into *key sentiment sentences* and *detailed sentences*. *Key sentiment sentence* is brief, discriminative and usually the most representative sentence in a document for the overall sentiment orientation and opinion. While *detailed sentence* is often more details, trivial, expressed complexly and ambiguously. Then, *key view* and *detail view* are constructed by *key sentiment sentences* and *detailed sentences* respectively. Through effectively integrating two views, we can not only acquire knowledge from domain-independent part to a large extent, but also bridge the distribution gap from *detail view* by domain-dependent part to a certain extent. Experimental results show that our proposed method achieve a very impressive improvement across three domains on 6 tasks.

The remainder of this paper is organized as follows. Section 2 introduces the problem setting and related concepts. Section 3 describes our proposed method in detail. Section 4 presents experimental results and discusses related parameters. At last, we conclude the paper and outline the future work in Section 5.

2 Problem Setting and Related Concepts

In this section, we introduce the definition of the problem that will be discussed below and several concepts related to our proposed method.

2.1 Problem Setting

Source Domain: $S = \{(x_i, y_i)\}_{i=1}^{n_s}$ refers to a set of labeled instances from a certain domain. Here, x_i is the i_{th} labeled instances. Particularly, it denotes a piece of product review in this paper. y_i denotes the sentiment label which is assigned for the i_{th} instances, and $y_i \in \{+1, -1\}$, here the sentiment labels +1 and -1, respectively, denote positive and negative. In addition, n_S indicates the number of labeled instances in source domain.

Target Domain: $T = \{(x_i)\}_{i=1}^{n_T}$ refers to a set of unlabeled instances from a certain domain, which is different from source domain but is related to source domain. Here, x_i is the i_{th} unlabeled instances. n_T denotes the number of unlabeled instances in target domain.

Cross-Domain Sentiment Classification: cross-domain sentiment classification seeks to generalize a model trained on source domain and uses it to infer the sentiment label of unlabeled instances in target domain.

2.2 Related Concepts

Domain-Independent Feature: If some features appear frequently in both source and target domains, and their sentiment orientations are coherent across different domains, they are *domain-independent*. Such as "*like*", "*hate*".

Domain-Dependent Feature: Contrary to *domain-independent feature*, if some features usually occur in a certain domain, whereas hardly occur in other domains, or their sentiment orientations are not coherent even opposite between source and target domains, they are *domain-dependent*. Such as "*obscure*" in book reviews, "*short battery life*" in computer reviews.

Sentiment Contribution: *Sentiment contribution* is the importance of subsentence in a document for predicting the overall sentiment label.

Key Sentiment Sentence: *Key sentence* for short, refers to the sentence that can substantially express the sentiment orientation of a document. Generally, the *sentiment contribution* of *key sentence* is larger than the other sentences.

Detailed Sentence: *Detailed sentence*, mainly refers to the sentence that is about what happened or attribute of products and so on. Usually, the *sentiment contribution* of *detailed sentence* is relative smaller compared with *key sentence*.

3 Cross-Domian Sentiment Classification Method Based on Extraction of Key Sentiment Sentence

In this section, we describe how to extract *key sentence*, and then present the process of integrating each view to promote the performance of base classifier for cross-domain sentiment classification.

3.1 Key Sentiment Sentence Extraction Algorithm

As a special issue of text classification, sentiment classification has received the considerable attention recently. However, due to the personalization and arbitrariness of human expression, sentiment classification, aiming at discriminating sentiment orientation of text, is more complicated than the ordinary text classification. Take a *NB* review for example, as shown in Tab. 1.

Table 1. *NB* reviews with mixture sentiments

Overall sentiment label	Sub-sentence content	Sub-sentence label
+1	<s1>Inadequate: USB interface is too few,and is left!	-1
	<s2>Installing system is a little trouble, trouble...	-1
	<s3>But,overall it is pretty good price.	+1

In this review, although there are some shortcomings(such as "*few*", "*trouble*), the overall sentiment orientation towards *NB* is positive, which can be seen from <s3>. In addition, compared with the other sentences, <s3> is shorter and unambiguous, contains "*overall*" and at the end of this review, intuitively it is much more important for inferring the whole sentiment label, namely the *sentiment contribution* of <s3> is larger. Thus, on the basis of the different *sentiment contribution* for each sub-sentence, the whole sentiment label of this review above is positive. However, those negative words may cause the classifier to discriminate the whole sentiment orientation negative, which leads to misclassification unfortunately.

To solve this specific problem, some scholars have found that when predicting sentiment orientation of a document, different sentences in the same document have different sentiment contribution [7]. Although the sentiment orientations of different sentences may be different, seeking out those having larger sentiment contribution as *key sentences*, then making a distinction between *key sentences* and *detailed sentences* will be helpful for inferring the overall sentiment orientation. Inspired by above, this paper proposed a two-stage extraction algorithm combining heuristic rules with machine learning. In the first stage, we make use of heuristic rules to construct the initial training sets for *key view* and *detail view.* In the second stage, we utilize the initial training sets to extract *key sentences* for remaining documents based on machine learning.

The heuristic rules in the first stage mainly extract three factors that would impact on the sentiment contribution of sentences, namely, *sentiment purity*, *keyword property* and *position property.*

Sentiment Purity. As a representation of the overall sentiment orientation of a document, *key sentence* should contain quite pure sentiment orientation. Therefore, we define *sentiment purity* as a factor to measure the *sentiment contribution* of a sentence. Higher *sentiment purity* means greater *sentiment contribution* and more likely to be a *key sentence*. The formula of *sentiment purity* score function is shown as below:

$$sentiPurity(s_i) = e^{\frac{|\sum_{w \in s_i} polarity(w)|}{|s_i|}}. \quad (1)$$

Here, $|s_i|$ is the number of words in a sentence. $polarity(w)$ denotes the sentiment label of the word w, if the label of w is positive, then $polarity(w) = 1$, label is negative, $polarity(w) = -1$, otherwise, $polarity(w) = 0$.

In this paper, we utilize the sentiment vocabulary ontology published by Dalian University of Technology [8] to judge whether a word is sentiment vocabulary or not. From formula (1), we can see that *sentiment purity* is inversely proportional to the length of a sentence, indicating that the more brief, the more likely a sentence is a *key sentence*. Besides, the more simple the sentiments are, $|\sum polarity(w)|$ is relatively larger, the higher the *sentiment purity* score is. It illustrates that a quite simple sentence is more likely to be *key sentence*. Based on the habits of human expression, when individuals tend to describe the details, sentiments usually are complex and mixed, whereas sentiments are unambiguity and concise when presenting overall sentiment orientation.

Keyword Property. Inspired by human language habits, *key sentence*, laying the sentiment tone of the whole document, should contain summary keywords generally, such as "*overall*", "*in my opinion*" and so on. Therefore, summary keywords are also a critical factor for extraction of *key sentence*. We selected the top 15 terms as the keywords by term frequency in the first and last sentences. The formula of *keyword property* score function is shown as below:

$$keyWordNum(s_i) = \sum_{w \in s_i} I_K(w). \quad (2)$$

Here, K is the set of keywords used in this paper. $I_K(w)$ is the indicator function, when $w \in K$, $I_K(w) = 1$, otherwise $I_K(w) = 0$.

Position Property. It can be found that people often present their opinions at the beginning of a document, and make a summary in the end generally. So, *position property* is a key factor that cannot be ignored. The formula of *position property* score function is shown as below:

$$position(s_i) = i^2 - n \times i + 100. \quad (3)$$

Here, n presents the number of sub-sentence in a document. i denotes i_{th} sub-sentence, and $i \in [1, n]$. Constant term "100" guarantees the *position property* score is a positive number. Formula (3) is a parabola with opening upwards, the symmetry axis at the very center of the document($n/2$), which ensures that the beginning and end sentences have larger advantage, namely relatively high *position property* score. On the contrary, the *position property* score of these sentences in the middle of a document is quite low.

In summary, we normalize scores of three factors above, and work out the weighted sum of these normalized scores. Then we get the *sentiment contribution* score of each sentence based on the formula (4).

$$sentiContri(s_i) = w_1 \times sentiPurity(s_i) + w_2 \times keyWordNum(s_i) + w_3 \times position(s_i). \quad (4)$$

Here, w_1, w_2 and w_3 deonte the weights of the corresponding factors respectively. We get the value of each weight by *Analytic Hierarchy Process* [9]. First, construct judgement matrix based on ratio scale(Tab. 2). Then, obtain the weight value by finding the eigenvector of the judgement matrix. Finally, $[w_1, w_2, w_3] = [0.17, 0.53, 0.30]$.

Table 2. The judgement matrix

three factors	sentiment Purity	keyword Property	position Property
sentiment Purity	1	1/3	1/2
keyword Property	3	1	2
position Property	2	1/2	1

For thresholds δ_1 and δ_2 , if $sentiContri(s_i) > \delta_1$, s_i is divided into *key view*, whereas if $sentiContri(s_i) < \delta_2$, s_i is divided into *detail view*. Thus, the initial training sets of two views can be drawn.

Then, we take advantage of the initial training sets of two views, and employ machine learning algorithm to extract *key sentences*. Each iteration selects these sentences with high-confidence to add to corresponding view. Repeat iteration until all reviews have been extracted *key sentences*. Then, the remaining sentences are considered as *detail sentence*s. The procedure of our algorithm is shown in Tab. 3.

3.2 Multi-view Ensemble

Ensemble learning has demonstrated impressive capacities to improve the performance of base learning algorithm [11,12]. Ensemble classifier refers that combining the classifiers that have independent decision-making ability together. It has been proved that the ensemble classifier is more predictive than the single classifier normally.

After dividing data into two independent and complementary views, *key view* and *detail view*, we train two base classifiers f_{key} and f_{detail} on these two views separately. Effectively integrating the two base classifiers, can learn the knowledge that a single base classifier cannot learn. In this paper, two well-known simple fixed rules [13] for integrating the set of base classifiers into a combining classifier will now be described as follows.

Table 3. The procedure of *key sentence* extraction

Input:The training data(in the source domain)
Output:The *key view* and the *detail view*. Procedure: 1: Segment reviews into sentences. And segment Sentence into words via NLPIR [10]. 2: Apply heuristic rules to construct the initial training sets for *key/detail view*. 3: Train a classifier f with the initial training sets for the two views. 4: Employ f to classify the remaining sentences. And put top-confident sentences into corresponding view. 5: Repeat 4, 5 until all the reviews have been extracted *key sentence*. 6: Return *key view* and *detail view*.

Product Rule. *Product rule* is to combine base classifiers by multiplying the posterior possibilities and using the multiplied possibility for decision.

$$p_i = \prod_{f=1}^{n} p_f(c_i|x) \qquad y = arg \max_i p_i \tag{5}$$

Sum Rule. *Sum rule* is to combine base classifiers by summing the weighted posterior possibilities and using the accumulative possibility for decision.

$$p_i = \sum_{f=1}^{n} w_f p_f(c_i|x) \qquad y = arg \max_i p_i \tag{6}$$

4 Experiments

In this section, we will present experiments on a real-world dataset to prove the effectiveness of the method we proposed. Besides, we also discuss the sensitivity of related parameter.

4.1 Dataset Description

The dataset used in our experiments is Chinese sentiment corpus published by TanSongbo [14]. It contains a collection reviews of three products: *book*, *hotel*, and *NB*. So we can construct 6 cross-domain sentiment classification tasks: *book→hotel*, *book→NB*, *hotel→book*, *hotel→NB*, *NB→book*, *NB→hotel*, where the domain before an arrow is source domain and the domain after an arrow is target domain. For the sake of convenience, we use *ChnSentiCorp* to denote this dataset. The detail of the dataset is shown in Tab. 4.

Fig. 1 demonstrates the sentence distribution of *key view* and *detail view*. As can be seen from Fig. 1, the number of sentences in *detail view* is more than that in *key view* in three domains. Moreover, in *book* and *hotel*, the number of sentences in *detail view* is twice as many as that in *key view*, while the number

Table 4. Scale of the datasets

Domain	Positive	Neative	AveLength	AveSentence	Features
book	2000	2000	73	3.7	17517
hotel	2000	2000	72	4.4	13835
NB	2000	2000	30	2.3	6565

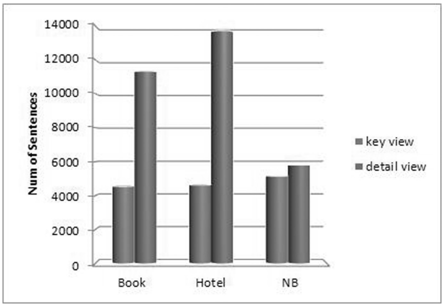

Fig. 1. Sentence number in each view

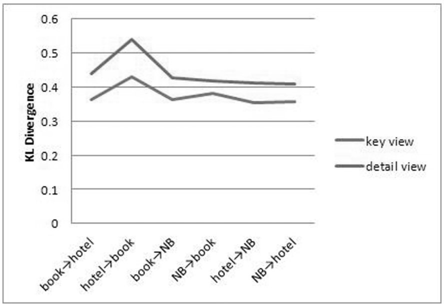

Fig. 2. KL divergence across domains

of sentences in *key view* is not much different from that in *detail view* in *NB*. Through analysis and comparison, there are a lot of sentences about storylines in *book* and details of what happened in *hotel*. However, in *NB*, there is mainly simple description about attributes of *NB*. Therefore, in *NB*, the number of sentences in two views is relatively balance than that in *book* and *hotel*. Besides, we can find that the average sentence number of *NB* is 2.3 from Tab. 4, which proved that in *NB*, the sentences in two views is more balance again.

In order to verify the distribution gap in *key view* is much smaller, we calculate the Kullback-Leibler (KL) [15] divergence of the same view in different domains. The closer two distributions are, the smaller KL divergence is. Fig. 2 shows the KL divergence curves between different domains. It can be seen that the KL divergence of *detail view* is much larger than that of *key view* overall, which reveals that the distribution among *key view*s is much closer. Besides, it coincides with our intuition that features in *key view* is mainly *domain-independent* while *domain-dependent* features usually appear in *detail view*.

4.2 Baselines Setting

In order to evaluate the effectiveness of extracting *key sentence* for cross-domain sentiment classification, we will compare our proposed method with several other methods in this sub-section.

No Transf. *No Transf* denotes that we train a classifier using source domain and directly apply the classifier on target domain. Just because there is not any transfer learning, this method can be considered as a lower bound.

SCL. SCL, short for structural correspondence learning, is proposed by Blitzer et al. [4]. We follow the details described in Blitzers thesis to implement *SCL*.

GraphOA. GraphOA applied the graph-ranking algorithm using the accurate labels of source domain as well as the pseudo labels of target domain for cross-domain sentiment classification.

Product Rule. In order to verify the robustness and effectiveness of our proposed view mining strategy, we employ *product rule* that is different from *sum rule* to integrate multi-view for cross-domain sentiment analysis.

For *No Transf*, *SCL*, *product rule* and *sum rule*, we use Support Vector Machine (SVM) [16] as the basic classifier. We choose linear kernel function for experiments. For all experiments on the 6 tasks, we randomly split each domain data into a training set of 3200 instances and a test set of 800 instances. We use information gain to select features. Accuracy is used to evaluate the cross-domain sentiment classification result.

4.3 Overall Comparision Results

The performances of different methods are shown in Tab. 5.

Table 5. Accuracy comparison of different methods

methods	*book→hotel*	*hotel→book*	*book→NB*	*NB→book*	*hotel→NB*	*NB→hotel*
No Transf	54.88%	60.00%	56.25%	54.88%	72.25%	72.50%
SCL	69.30%	80.42%	69.85%	65.38%	80.55%	73.80%
GraphOA	71.45%	76.08%	74.43%	79.30%	78.68%	79.80%
product rule	84.57%	79.97%	79.90%	72.90%	84.20%	86.08%
sum rule	85.98%	80.95%	80.98%	74.18%	84.88%	86.43%

Fig. 3 and Tab. 5 show the results of different methods together. Compared with *No Transf*, *SCL*, *GraphOA* and *product rule*, our method *sum rule* get significant improvements in all tasks. It is not surprising to get this result. The most intuitional reason is that capturing common information between domains can find a reasonable representation for cross-domain sentiment classification, and it will reduce the gap of distributions. It is coincident with our analysis that *No Transf* can be seen as a lower bound. Besides, *sum rule* performs better than other methods including *SCL*, which is well-known and classic method. On the one hand, the performance of *SCL* highly relies on pivot features, but heuristically selecting method might not guarantee the best performance always. On the other hand, it reveals that extraction of *key sentence* is significative for cross-domain sentiment classification. For *GraphOA*, except for the task *NB→book*, the accuracy of *sum rule* is higher than that of *GraphOA*. Moreover, average accuracy increased 5.6% as a whole. The reason might be that *GraphOA* is based on the instance similarity. However, in instance-level sentiment classification, the dataset is relatively sparse. In this case, it is hard to guarantee each test instance and all instances which have higher similarity have the same sentiment label. So, the performance of *GraphOA* is not so superior compared with the other methods. But, *GraphOA* presents its advantage on *NB→book* and *NB→hotel* tasks. Especially on the *NB→hotel* task, the performance is the best among all methods. The analysis indicates that *NB* review is quite short, and thus it becomes easy to detect the similarity with those instances in other domains accurately.

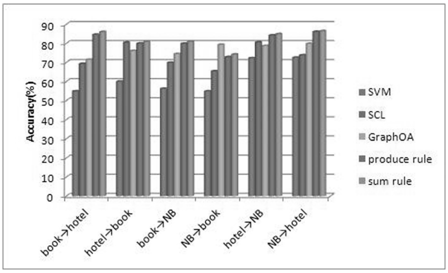

Fig. 3. Sum rule vs. other algorithms

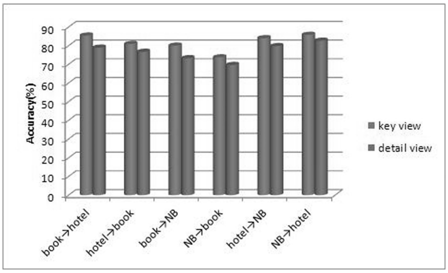

Fig. 4. Performance of *key/detail view*

As for *product rule*, *sum rule* has about 1% increase. From the formula (5) and (6), we can observe that *sum rule* allocates different weights for each base classifier respectively, whereas *product rule* does not. To reduce the gap between different domains is the core aim for cross-domain sentiment classification. As proved before, distribution gap of *key view* is much smaller than that of *detail view*. Therefore, in theory *key view* should be allocated much larger weight. Above all, it is reasonable that *sum rule* is more suitable for our proposed method.

Fig. 4 shows the performance of base classifiers trained on *key view* and *detail view*. Clearly, the classifier trained on *key view* is better than that on *detail view*. Compared with *detail view*, *key view* improves about 4.8%. This coincides with the analysis that the distribution gap of *key view* is much smaller than that of *detail view*.

4.4 Parameter Sensitivity Analysis

In this part, we discuss the effect of parameter w_{key} which is the weight of *key view* on the performance of ensemble classifier.

In Fig. 5, we present the weight-dependent accuracy curves on 4 tasks. From these figures, we can see that although each set of results obtained the best integrated performance at different w_{key}, w_{key} is more than 0.5 at the optimal performance. For example, for the task *NB→hotel*, when the value of w_{key} is nearly 0.65, accuracy is the best, the value of w_{key} is about 0.75 on the task *hotel→NB* as well. Thereby this verified the conclusion that allocating much larger weight for *key view* than *detail view* will be of benefit to cross-domain sentiment classification. Clearly, *sum rule* is much more suitable for our proposed

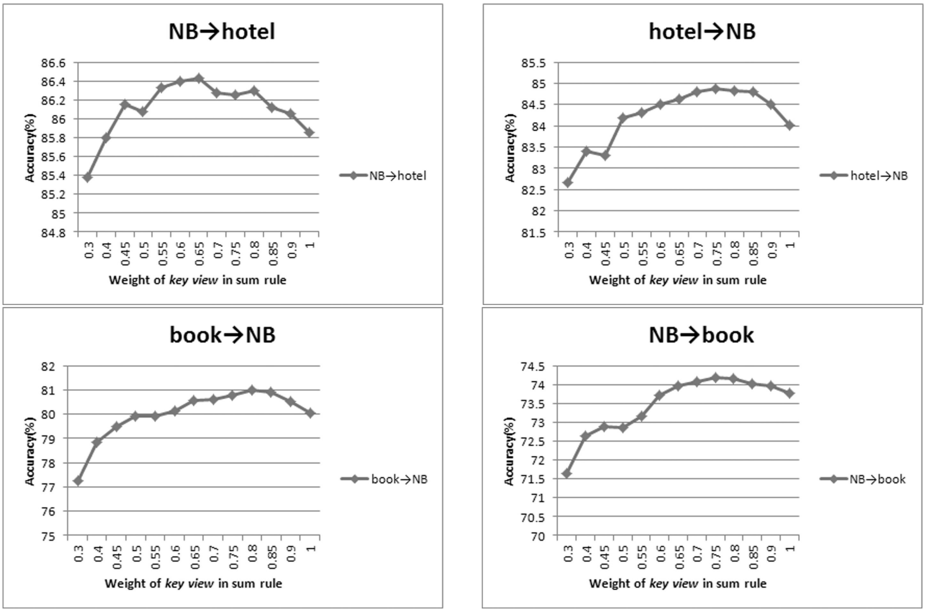

Fig. 5. Parameter sensitivity on different tasks

method. Consequently, our proposed view mining strategy based on extraction of *key sentence* is valid for cross-domain sentiment classification.

5 Conclusion

In this paper, we proposed a cross-domain sentiment classification based on extraction of *key sentiment sentence*. Firstly, we construct heuristic rules, and use them to initialize the training sets for two views, *key view* and *detail view* respectively. Secondly, machine learning is adopted to extract *key sentiment sentence*s for all reviews. Thus data is divided into *key view* and *detail view*. Finally, two views are integrated effectively for cross-domain sentiment classification.

In the future, we are planning to detect deep properties of sentence for extraction of *key sentiment sentence*. In addition, multi-view ensemble strategy *product rule* does not distinguish the importance of integrated views, how to improve this rule to make it have a distinguish ability for different views is also future research content.

Acknowledgments. This work is partially supported by grant from the Natural Science Foundation of China (No. 61277370, 61402075), Natural Science Foundation of Liaoning Province, China (No.201202031, 2014020003), State Education Ministry and The Research Fund for the Doctoral Program of Higher Education (No.20090041110002), the Fundamental Research Funds for the Central Universities.

References

1. Pang, B., Lee, L., Vaithyanathan, S.: Thumbs up? sentiment classification using machine learning techniques. In: Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing, pp. 79–86 (2002)
2. Singh, V.K., Piryani, R., Uddin, A., Waila, P., et al.: Sentiment analysis of textual reviews; Evaluating machine learning, unsupervised and SentiWordNet approaches. In: 5th International Conference on Knowledge and Smart Technology (KST), pp. 122–127. IEEE Press (2013)
3. Mohammad, S.M., Kiritchenko, S., Zhu, X.: NRC-Canada: building the state-of-the-art in sentiment analysis of tweets. In: Proceedings of the Second Joint Conference on Lexical and Computational Semantics (SEMSTAR 2013) (2013)
4. Blitzer, J., McDonald, R., Pereira, F.: Domain adaptation with structural correspondence learning. In: Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing, pp. 120–128 (2006)
5. Bollegala, D., Weir, D., Carroll, J.: Cross-domain sentiment classification using a sentiment sensitive thesaurus. IEEE Transactions on Knowledge and Data Engineering **25**(8), 1719–1731 (2013)
6. Li, S., Xue, Y., Wang, Z., et al.: Active learning for cross-domain sentiment classification. In: Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence. AAAI Press, pp. 2127–2133 (2013)
7. Lin, Z., Tan, S., Cheng, X.: Sentiment classification analysis based on extraction of sentiment key sentence. Journal of Computer Research and Development **49**, 2376–2382 (2012)
8. Xu, L., Lin, H., Zhao, J.: Construction and analysis of emotional corpus. Journal of Chinese Information Processing **22**, 116–122 (2008)
9. Saaty, T.L.: Decision making with the analytic hierarchy process. International Journal of Services Sciences **1**(1), 83–98 (2008)
10. NLPIR. http://ictclas.nlpir.org
11. Chalothom, T., Ellman, J.: Simple Approaches of Sentiment Analysis via Ensemble Learning, pp. 631–639. Springer, Heidelberg (2015)
12. Xia, R., Zong, C.: A Pos-based Ensemble Model for Cross-domain Sentiment, pp. 195–197. Citeseer (1981)
13. Alexandre, L.A., Campilho, A.C., Kamel, M.: On combining classifiers using sum and product rules. Pattern Recognition Letters **22**, 1283–1289 (2001)
14. Chinese Sentiment Corpus. http://www.searchforum.org.cn/tansongbo/corpusenti.htm
15. Kullback, S., Leibler, R.A.: On information and sufficiency. The Annals of Mathematical Statistics, 79–86 (1951)
16. Chang, C., Lin, C.: LIBSVM: a library for support vector machines. ACM Transactions on Intelligent Systems and Technology (TIST) **2**, 27 (2011)