Sentiment Analysis Based on Chinese Thinking Modes

Liang Yang, Hongfei Lin, and Yuan Lin

Department of Computer Science and Engineering,
Dalian University of Technology, Dalian 116023, China
{yangliang,yuanlin}@mail.dlut.edu.cn, hflin@dlut.edu.cn

Abstract. Sentiment analysis is an important research domain for NLP, and currently it mainly focuses on text context. While our research concentrates on the thinking modes, which influence the formation of language. “Spiral graphic mode”, “concreteness” and “scattered view”, are taken into consideration to assist sentiment analysis and classification in this paper. According to these explicit Chinese modes, a Chinese sentiment expression model (CSE) is proposed, which can effectively improve the accuracy of emotion classification. In order to solve the implicit Chinese sentiment expression, Latent Semantic Analysis (LSA) is applied when the CSE model could not classify the implicit emotions accurately. By comparing with two traditional sentiment analysis methods, experimental results show that the performance of sentiment analysis included the Chinese thinking mode factors is significantly better than which not included.

Keywords: Chinese thinking mode, Chinese sentiment expressing model, LSA.

1 Introduction

Sentiment analysis is a popular research topic in NLP in recent years. It aims to assist computers to recognize the human emotions, which is more widely used in industry and academia. For example, whenever we need to make a decision, we often seek out the opinions of others. Qiu et al [1] constructed the double propagation between opinion word and target aspect to get the sentiment orientation of the opinion to the target aspect. Wang et al [2] proposed a rating regression approach to analyze the rating text review data, which could generate summaries about aspects for consumers. Another example is the word-of-mouth on social media, such as Blog and Micro-Blog, and mining the tendency of a group of peoples’ opinions is badly needed. There has been some work by researchers. Such as Wang et al [3] presented a graph-based hashtag sentiment classification approach to analyze the sentiment tendency in Twitter.

In China, researchers concentrate on phrase level and sentence level sentiment classification recently, such as COAE (Chinese Opinion Analysis Evaluation) [4]. Less work has been done on passage level comparing with the former two levels, for the structures and patterns of emotion expressions are more complex. Turney [5] proposed a lexicon based method to identify the passage’s emotion by using the average score of emotion word and phrase. Tan et al [6] constructed a group of domain lexicons to guarantee the accuracy. Although the existing lexicon based method gets a

M. Zhou et al. (Eds.): NLPCC 2012, CCIS 333, pp. 46–57, 2012.

good result, its flexibility and human annotation accuracy are still a potential problem. Xu et al [7] proposed a method based on semantic resources, and Condition Random Field (CRF) is applied to label emotion sentence by sentence; after the emotion chain formed, the emotion of the passage is determined.

In this paper, we analyze relationships between thinking modes and language, and three Chinese thinking modes are taken into consideration to assist sentiment analysis and emotion classification. Then a Chinese sentiment expression model (CSE) is proposed based on the thinking modes, and DUTIR affective ontology [8] is also combined to identify the explicit emotions. Finally Latent Semantic Analysis (LSA) is applied when the CSE model could not classify the implicit emotions accurately.

The paper is organized as follows: Section 2 introduces the Chinese and Western Thinking Modes and their features. Then Section 3 describes the quantification of Chinese thinking modes and the Chinese sentiment expression model (CSE). In section 4, implicit emotion mining method based on Latent Semantic Analysis (LSA) is explained. Section 5 presents experiments and results. Finally, we conclude this work and point out some directions for future research in Section 6.

2 Thinking Modes

In Wikipedia, thought (or thinking) [9] generally refers to any mental or intellectual activity involving an individual's subjective consciousness. It can refer either to the act of thinking or the resulting ideas or arrangements of ideas. Different civilizations cradle different thinking mode, and thinking modes determine the expression way of languages. From the concept of thinking, we can find that thinking and language are inseparable. Language is the expression way of thinking, and it carries the abstraction of the reality. When we express our thinking, to a large degree, the features of thinking play a decisive role. Every nation has its own thinking mode, so researches on thinking mode are essential to sentiment analysis. Later, we will introduce Western and Chinese thinking modes respectively.

2.1 "Spiral Graphic Mode" and "Straight Line Mode"

"Spiral graphic mode" is one of Chinese thinking modes, and it commonly reflects in the passage organization aspect. Because of the implicit characteristic, Chinese normally introduces the topic in an indirect way like a spiral graphic, so the topic of the passage is discussed after examples. On the other hand, "Straight Line Mode" is the feature of English passage organization, for Western thinking is influenced by the Roman philosophy, so they focus on deduction and thinking in a straight line way. On passage organization aspect, they tend to state their views directly and frankly. Take these sentences below as examples:

(1) Chinese: "他被眼前的一幕震惊了。"
English: "He was shocked by what he saw."
(2) Chinese: "经过反复的思考，我终于得到了完美的答案。"
English: "I got a perfect answer after deeply thinking."

Form the above mentioned examples, we can find that the key emotion words "shocked" and "perfect" appear in the front part of the sentence, but the Chinese emotion word "震惊" and "完美" locate at the end part of the sentences.

2.2 "Concreteness" and "Abstractness"

The second Chinese thinking mode is called "concreteness", for Chinese is a hieroglyphic language, which means that quantities of specific words, shapes, sounds and description are used in Chinese to illustrate abstract things, and some character components can reflect this characteristic. But the English lay emphasize on "abstractness", which is another western thinking mode. They tend to implement general vocabularies and their variants to express abstract feelings or opinions, such as "-ion", "-ance" and "-ness", but concrete things are seldom applied to explain.

(3) Chinese: "土崩瓦解。" English: "Disintegration."
(4) Chinese: "有志者，事竟成。" English: "Where there is a will, there is a way."

From example (3), we found that concrete things, soil, tile, appeared in Chinese idiom, while only the variant of disintegrate is used to explain these meaning in English. And in example (4), we could get an overview that verb has advantages in Chinese, but noun is more frequent used flexibly in English.

2.3 "Scatter View" and "Focus View"

"Scatter view" [10] is the third Chinese thinking modes. From the angle of ontology, Chinese tend to emphasize unified whole, which means think from more to one. When Chinese express their feelings or comments, they are usually inclined to use various words to enhance their emotion. "Scatter View" is usually applied in the Chinese expression, for example, we can frequently find that more than one verb is used in one Chinese sentence, but the sentence is still fluent. While English pay more attention to logical reasoning or deduction, they tend to express their feelings or emotions briefly, and it is a kind of focusing thinking, which is called "Focus View". "Focus view" can be reflected by the only one verb, which normally is the core of one sentence. "Focus view" is widely adopted in English. Here are two examples about "Scatter view" and "Focus view" below:

(5)Chinese: "他拿着课本走进了教室。"
English: "He walked into the classroom with a textbook in hands."
(6)Chinese: "他们俩青梅竹马，两小无猜。"
English: "The boy and the girl were playmates in their childhood."

From the examples (5), we can clearly find that more than one word are implemented in Chinese sentence, but each English sentence only contain one word, and it is the core of the sentence. In example (6), "Scatter View" is embodied in Chinese expression, which is reflected by more words used to express the same meaning.

To sum up the above mentioned Western and Chinese thinking modes [11], we find that "Scatter view" and "Focus View" is the feature of sentence structure;

"concreteness" and "abstractness" are the characteristic of vocabulary usage; "Spiral graphic mode" and "Straight line mode" are reflections on passage expression model. "Scatter view" and "Spiral graphic mode" are explicit in Chinese articles, such as vocabulary variants, while "concreteness" mostly appears in an implicit way. These thinking modes provide us a new angle to analyze text sentiment orientation.

3 Description of Chinese Sentiment Expression Model

3.1 External Resource

So far, there is no standard in emotional classification, the paper uses DUTIR Emotion Ontology [8] as the external resource. Because human emotion is complicated and changeable, and people have insufficient cognition about it, so the emotion is divided into 4, 6, 8, 10 and 20categories etc. While in DUTIR Emotion Ontology, the emotion is classified into 7 categories and 20 subcategories, and it can be applied widely in sentiment analysis and emotion classification.

To recognize the affective lexicon, we calculate mutual information between the lexicon and ontology in the resource, and then we combine some affective lexicon rules, such as part-of-speech rules, co-occurrence rules, and context rules at el. Machine learning method is also used to automatically expand the emotion ontology. Conditional Random Fields [12, 13] (CRFs) is adopted as the automatic acquisition method, the formula is defined as follows:

$$P_\theta(y \mid x) = \exp(\sum_{e \in E,k} \lambda_k f_k(e, y|_e, x) + \sum_{v \in V,k} u_k g_k(v, y|_v, x) \tag{1}$$

Where x is a data sequence and y a label sequence, $y|v$ is the set of components of y associated with the vertices in sub graph S. The features f_k and g_k are given and fixed.

While in Chinese Opinion Analysis Evaluation (COAE) [4], emotion is classified into 4 categories, which are "happy", "angry", "sad" and "fear", and opinion evaluation consists of positive and negative. So we made some adjustments on DUTIR Emotion Ontology to complement the standard of COAE.

3.2 Quantification of Similarities between Thinking Modes

Although there are differences between Western and Chinese thinking modes, there are still some similarities. For example, no matter what Chinese or English articles are, the negative words locate in front of emotion or opinion words to express contrary feelings or attitudes; adverb of degree is placed before adjective to enhance or decrease the its intensity; if an adversative occurs, the emotion of the sentence is determined by the part after the adversative.

Sentence is the foundation of constructing paragraph and passage, and it is the minimum level to meet sentiment phenomenon on the above mentioned part. So sentence level is chosen as minimum research object level. Modifier window strategy is adopted to analyze the emotion of a sentence and its score. In this strategy, negative words, adverb of degree and adversative are detected if they exist in the modifier window.

3.2.1 Quantification of "Spiral Graphic Mode"

"Spiral graphic mode" is a Chinese sentence or passage organization characteristic, the topic of article is introduced in an indirect way. Demonstration comes first, and the theme usually locates the tail of the sentence or passage. On the contrary to the above mentioned, the nearer emotion words locate the tail of the sentence, the more important they are to determine the emotion of the sentence. For example:

(7)Chinese: "这个酒店什么都好，就是服务让人失望。"
English: "Every aspect about the hotel is ok except the disappointing service."

Although the former part presents a positive opinion, the word "失望" (disappointing) in later part determines the emotion of the whole sentence. So the example sentence shows a negative opinion. Based on the statistic data of our corpus, we find that the emotion-determining words mostly locate the end part of Chinese sentences. And this conclusion is still useful to paragraphs and passages [14].

In order to simulate the "spiral graphic mode", equation (2) is applied:

$$score(A) = \sum_i \left(1 + position(a_i) / count(a \mid A)\right) \times score(a_i) \tag{2}$$

Where A can be a passage, paragraph or sentence, a can be a paragraph, sentence or a word. $score(A)$ is the emotion score of A, and $position(a_i)$ is the position of the emotion unit a_i .

From the equation (2) we can find that if the closer a_i locates the tail, the larger the $score(a_i)$ will be. Take example (7), the word "失望" (disappointing) locates near the tail of the sentence, it can get a larger score than the word "好" (good) after computing by equation (2), so different emotion words regain new emotion weights, so this sentence tends to have a negative tendency.

3.2.2 Quantification of "Concreteness"

Based on the analysis of the differences between "abstractness" and "concreteness", we find that the characteristic or property of a certain word is helpful to sentiment analysis. In all nations, adjective tends to have a better performance to express feeling or emotion than the other part of speech word. While in Chinese sentiment expression, verb also plays an important role, such as "溃败" (be defeated) and "脸红" (blush). The former tries to express a "fear" feeling and the latter means shy or embarrassed. In order to simulate this thinking and language characteristic, different part of speech word and sentence structure are adopted different tactics. The details can be seen as follows in equation (3).

$$score(a) = \sum_i^n W_i^{adj} + \sum_j^m W_j^{verb} + \sum_k^l W_k \tag{3}$$

Where a is a sentence, and *score(a)* is the score of sentence a. W_i^{adj} is the i th adjective, W_j^{verb} is the j th adjective, and W_k is other part of speech word in sentence a.

In equation (3), it does not simply add all these part of speech words' weights together, but give the highest priority to adjective, then the higher priority is given to the verb, finally other words are processed. "Concreteness" is implemented by the priority to the verb, and which meets the Chinese thinking explicitly.

3.2.3 Quantification of "Scatter View"

In Chinese writing, authors are likely to use a number of similar expressions to express the same sentiment or emotion. And a great number of nouns or adjectives are used as predicate and this has been discussed above. In order to simulate the Chinese thinking mode "scatter view", view-window is adopted, which means that in specific position different sizes of windows are applied to reflect this characteristic. Based on our experiment, when the size of window is fixed at 6, we can get the best result. If the size of view-window is too small, it cannot capture the words which significantly assist the sentiment analysis; on the contrary, too much noise will be introduced, which will also mislead the analysis.

3.3 Chinese Sentiment Expression Model

After analyzing the features of Western and Chinese thinking modes and their specific quantification methods, in this part, we will introduce a Chinese sentiment expression model, which is the combination of Chinese thinking modes. When we try to get the sentiment or emotion tendency of the text content, all of the Chinese thinking modes are taken into consideration.

"Scatter view" and "Spiral graphic mode" are conclusions on Chinese expression, and they are both explicit features of Chinese thinking modes. "Scatter view" is the feature of sentence structure, and "Spiral graphic mode" is a characteristic of passage organization. While in English expression, only one verb is contained in one sentence, so the emotion of this sentence is easy to determine after analyzing the key verb. But due to the "Scatter view" thinking mode, Chinese tend to use more verb and their variants to express emotions.

In order to enrich the vocabulary, DUTIR Emotion Ontology [8] is adopted. In Chinese Opinion Analysis Evaluation (COAE) [4], the emotions are classified into four group, which are "happy", "angry", "sad", and "fear", and opinion-bearing words are classified into "positive" and "negative". For this coarse granularity partition can avoid the intersection among different emotions, which can significantly reduce the classification error rate and improve the accuracy. 22773 emotion words and 6 groups are chosen manually to testify the influence of "Scatter view" on sentiment analysis. Therefore the number of emotion words can explain the characteristic "Scatter", but the emotion groups can maintain that every passage has an explicit topic.

Due to the complexity of the sentence structure, sometimes sentiment analysis or emotion classification would be misled by negative words and degree words. For example, the negative words can change the emotion tendency, and the degree words can enlarge or lessen the emotion intensity. So in this article, degree words, negative words and adversative are all taken into consideration.

After integrating the two Chinese thinking modes, “Scatter view” and “Spiral graphic mode”, the Chinese sentiment expression model is proposed. DUTIR Ontology [8] and classification principle are applied to implement the characteristic of “Scatter view”, and “spiral graphic mode” is simulated based on the position-influenced strategy.

4 Implicit Chinese Sentiment Expression Mining Based on LSA

The above mentioned Chinese sentiment expressions are extracted based on the explicit Chinese thinking mode “Scatter view” and “spiral graphic mode”. But the Chinese thinking mode “concreteness” mostly appears in an indirect way. If the emotion of the sentence cannot be analyzed by using explicit features, the thinking mode “concreteness” is helpful. Chinese express their emotions or feelings influenced by “concreteness” thinking mode, and they tend to make use of familiar and similar object. So it is an implicit feature for implicit emotion mining. In this paper LSA is used to determine the emotion of the implicit emotions, and it has been widely applied to calculate the similarity between word and word, word and passage, or passage and passage. In section 4.1, we discuss what kind of texts is labeled as implicit emotion text. Then we introduce the method to determine the emotion of implicit emotion text.

4.1 The Criterion of Implicit Emotion Sample

When people express their emotions or feelings, some articles can be identified just by analyzing the tendency of emotion words, but others are not. The articles, which cannot be determined by emotion words, are implicit emotion articles. Because the implicit emotion articles mostly have low scores comparing with the explicit ones, so the threshold is needed to classify the explicit emotion sample and the implicit ones.

In this paper, a group of samples are chosen to determine the threshold, and we call them threshold sample. The threshold samples are the sentences which only contain one explicit emotion word, and its intensity is 9 in DUTIR Emotion Ontology [8]. The strong intensity of the emotion ontology is applied to guarantee the explicitness of sentiment expression. The scores are implemented to determine the threshold.

4.2 Implicit Emotion Classification Based on LSA

LSA (Latent Semantic Analysis) is proposed by Dumais et al in the year 1988, which is used in statistic method to analyze mass texts, and get the latent semantic relationships among words. The main object of LSA is to map the higher dimension Vector Space Model (VSM) to lower dimension latent semantic space.

The main reason why the implicit samples exist is that the no emotion words are indexed in emotion lexicon. For the implicit emotion samples express emotion in an indirect way, so we need to further analysis, and the samples which scores are larger than the threshold are chosen as seed samples [15]. Then LSA is implemented to mine the latent semantic relationships between the implicit samples and the seed samples.

After the above steps, we can obtain the emotion tendency of the implicit samples. Equation (4) is used to compute the score of the implicit sample, and then identify the emotion of the sample：

$$score_j\left(A_i\right)=\sum_k sim\left(A_i,s_{jk}\right)\times score\left(s_{jk}\right) \tag{4}$$

Where $Score(A_i)$ denotes the score of implicit sample A_i in Emotion j; S_{jk} is the k th sample in Emotion j; sim(A_i, S_{jk}) denotes the similarity of implicit sample A_i and sample S_{jk}; $score(s_{jk})$ is the score of sample S_{jk} in Chinese sentiment expression model.

5 Experiment Setting and Evaluation

5.1 Experiments of Chinese Thinking Modes in Different Domains

In order to test and verify whether the Chinese thinking modes can assist sentiment analysis, three group experiments have been done on Evaluation data set provided by Tan [16]. The data set contains three domains, which are 4000 hotel reviews, 1608 electronics reviews and 1047 stock reviews, and each domain has its positive and negative subset. The reviews in the data set cover sentence level, paragraph level, and passage level, and this can avoid the particularity of sentence structure. The baseline is the method proposed by Turney [5], and three thinking modes are integrated to see whether they are helpful to assist sentiment analysis. The results of thinking modes method and method proposed by Turney are shown in figure 1, and the columns in it from 1 to 6 respectively represent the subsets: elec-neg, elec-pos, hotel-neg, hotel-pos, stock-neg and stock-pos.

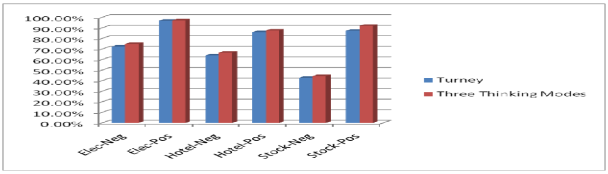

Fig. 1. Precision Results between Turney and Chinese thinking modes

From the results in figure 1, we can get a conclusion that no matter positive or negative review subset are, there is still an increase in all domains. It proves that the Chinese thinking modes can assist the sentiment analysis indeed, and they have independence to different fields. Other three groups of experiments have been done to analyze the three Chinese thinking modes in different domains.

The first group experiments are adopted on electronics negative and positive reviews. The reviews in elec-neg data set are relatively longer and rich in verbs. From figure 2, we can find that after adding each Chinese thinking mode the precision has increased. The precision of elec-pos reviews is not increasing obviously. That is because too much nouns are in the text context and less emotion words are used to express the only one emotion.

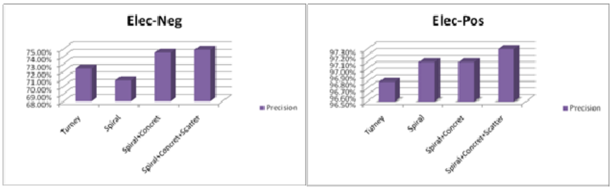

Fig. 2. Precisions on Elec-Neg data set and Elec-Pos data set

The second group experiments are implemented on negative and positive hotel reviews. Adding the "spiral graphic mode", the results on hotel-neg and hotel-positive data sets are basically remained the same level. But after the "concreteness" is taken into consideration, the results are much higher than before. It verifies that "concreteness" is much helpful in Chinese sentiment analysis. The size of view-window is decided by the specific data set to get the best result. The experiment results of hotel reviews are shown in figure 3 as follows.

Stock reviews are used to do the third group of experiments results. The result of stock-neg data set is not good in figure 4. After we analyze the errors, we find the reason that a great number of specialized words exist and part of them does not appear in DUTIR emotion ontology [8]. That is also the limitation of lexicon based method. From figure 4, there is an improvement after adding these three Chinese thinking modes. Due to most stock reviews are passage-level passages, the length of the review is relatively long than the former two electronic and hotel data sets. And it can explicitly reflect the Chinese thinking modes – spiral graphic mode and its advantage.

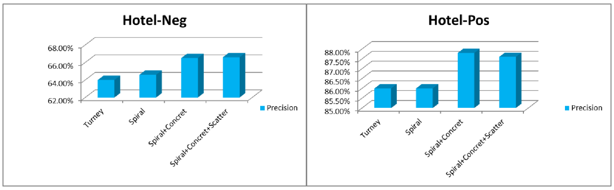

Fig. 3. Precisions on Hotel-Neg data set and Hotel-Pos data set

5.2 Experiment of Chinese sentiment Expression Model and LSA

In this experiment part, the experiment data set is ChnSentiCorp [17], and it is about hotel evaluation, which consists of hotel service quality, food quality, surrounding

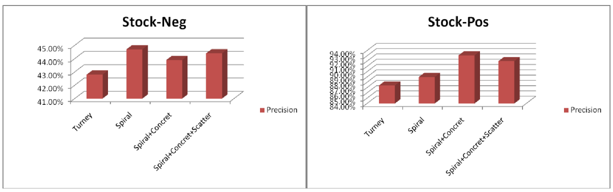

Fig. 4. Precisions on Stock-Neg data set and Stock-Pos data set

and so on. There are 2000 positive passages and 2000 negative passages in it. After checking the label result, the noisy texts are removed and parts of the passages are relabeled. Then we get 1828 positive passages and 2163 negative passages.

For comparison, we implement two baseline methods: one is lexicon based method proposed by Turney [5], which decides the passage polarity only by its emotion words. If they are positive-orientated, then the passage is positive-orientated. The same goes for a negative-orientated passage; the other method is based on semantic resource proposed by Xu [7], which takes semantic resource into consideration, such as negation and transition. The result is listed in table 1.

From the results in table 1, we find that the precision of lexicon based method is close to that of CSE. After we analyze the error list, the main reason is the implicit emotion samples. Due to the emotions of implicit samples are unclear, the method applied in experiments have a higher probability to be misclassified by them, and that will decrease the precision in the long run. Some statistic jobs on implicit emotion samples have also been done. Later we will illustrate them.

Semantic method focuses on the sentence level, and we can see that there is a precision increase of our proposed method - CSE from the comparing results. In order to identify the opinion tendency of the whole passage, the semantic method proposed by Xu [7] has to count the numbers of positive sentences and negative ones, and the larger number determines the opinion orientation of the passage. So we can see that the CSE model has a better performance in binary opinion classification than Xu [7].

Table 1. Experiment results on ChnSentiCorp

	Lexicon	**Semantic**	**CSE**
Pos	1575/1828(86.16%)	1502/1828(82.17%)	1575/1828(86.16%)
Neg	1727/2163(79.84%)	1811/2163(83.73%)	1827/2163(84.47%)
Total	3302/3991(82.74%)	3313/3991(83.01%)	3402/3991(85.24%)

Implicit emotion sample classification is a difficult problem in sentiment analysis. Based on the classification results of the former experiment in 5.2, and we can find that although the precisions are close, the correct classification number of implicit emotion sample is not good comparing with lexicon based method. So the implicit emotion classification based on LSA is necessary. Some statistic jobs have been done by us, and the details are listed in table 2.

Table 2. Statistic data of implicit samples

	Lexicon	CSE
Pos-Implicit	751/1575(47.68%)	640/1575(40.63%)
Neg-Implicit	836/1827(45.76%)	624/1827(34.15%)
Total	1587/3402(46.65%)	1264/3402(37.15%)

In order to solve the implicit emotion problem, LSA is implemented. Each implicit emotion sample is used as a query to index the relevant seed sample in latent semantic space. Based on the opinion orientation of seed sample, the emotion of the implicit sample is determined. To capture the higher similarities between implicit emotion sample and seed sample, the similarity threshold value is given 0.8 empirically. Figure 5 is the results after secondary classification in positive and negative data sets in ChnSentiCorp [17]. Figure 6 is macro-average precision of the former two data sets.

From figure 5 to 6, we can find that after secondary classification by implementing LSA, there is a significant increase. And the precision of the whole data set reaches higher than 90%.

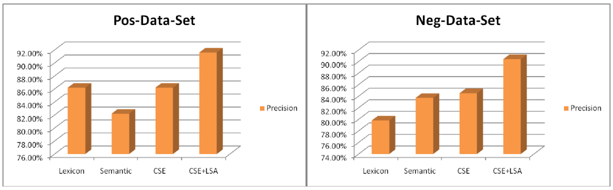

Fig. 5. Precisions on ChnSentiCorp Pos-Data-Set and Neg-Data-Set

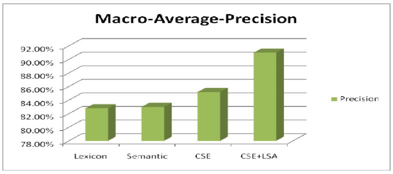

Fig. 6. Macro-Average-Precision of different methods in ChnSentiCorp

Chinese sentiment expression model is applied to get the emotion tendency of the explicit emotion sample from Chinese thinking modes, "spiral graphic mode" and "scatter view", and the results show that it cannot only guarantee the precision, but also make sure the credibility of emotion classification. LSA is adopted in secondary classification to do the mining work between seed sample and implicit sample in latent semantic space based on the "concreteness" thinking mode.

After analyzing the error samples, we find that classification errors in this paper mainly focus on the following two aspects: First, some of the seed samples are misclassified. There are 8 misclassified seed samples in positive data set and 15 of that in negative data set; second, due to the variety of implicit emotion samples, some of the implicit emotion samples cannot index the samples with higher similarities.

6 Conclusion

The contribution of this paper is to propose the Chinese sentiment expression model, which focuses on the thinking modes. Three Chinese thinking modes, "spiral graphic mode", "concreteness" and "scattered view", are applied to assist sentiment analysis and emotion classification. According to these explicit Chinese modes, a Chinese sentiment expression model (CSE) is proposed. From the experiment results, it can effectively improve the accuracy of emotion classification. Thinking mode "concreteness" mostly exists in implicit emotion expressions. In order to solve this, Latent Semantic Analysis (LSA) is applied to implement "concreteness" when CSE model could not classify the implicit emotions accurately. Two traditional sentiment analysis methods are used to verify the effectiveness of proposed method, CSE and LSA. Experimental results show that the performance of sentiment analysis included the Chinese Thinking mode factors and LSA mining is better than that not included.

References

1. Guang, Q., Bing, L., Jiajun, B., et al.: Opinion word expansion and target extraction through double propagation. Computational Linguistics 37(1), 9–27 (2011)
2. Hongning, W., Yue, L., Chengxiang, Z.: Latent aspect rating analysis without aspect keyword supervision. In: Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, USA, NY, pp. 618–626 (2011)
3. Xiaolong, W., Furu, W., Xiaohua, L., et al.: Topic sentiment analysis in twitter: A graph-based hashtag sentiment classification approach. In: Proceedings of the 20th ACM Conference on Information and Knowledge Management, USA, NY, pp. 1031–1040 (2011)
4. Hongbo, X., Tianfang, Y., Xuanjing, H., et al.: Overview of Chinese opinion analysis evaluation (EB/OL) (2009), http://ir-china.org.cn/coae2009.html
5. Turney, P.D., Littman, M.L.: Measuring praise and criticism: Inference of semantic orientation from association. ACM Transactions on Information Systems 21(4), 315–346 (2003)
6. Weifu, D., Songbo, T.: Building domain-oriented sentiment lexicon by improved information bottleneck. In: Proceedings of the 18th ACM Conference on Information and Knowledge Management, pp. 1749–1752. ACM Press, New York (2009)
7. Linhong, X., Hongfei, L.: Discourse affective computing based on semantic features and ontology. Journal of Computer Research and Development 44(z2), 356–360 (2007)
8. Linhong, X., Hongfei, L., Yu, P.: Constructing the affective lexicon ontology. Journal of the China Society for Scientific and Technical Information 27(2), 180–185 (2008)
9. Wikipedia. Thought (EB/OL), http://en.wikipedia.org/wiki/Thinking
10. Shuneng, L.: Mdes of thinking: Chinese and western. Foreign Languages and Their Teaching 2, 40–48 (2002)
11. Yan, C.: Thinking difference and modes of English and Chinese expressions. The Northern Forum 5, 111–113 (2001)
12. Jianmei, C., Hongfei, L., Zhihao, Y.: Automatic acquisition of emotional vocabulary based on syntax. Caai Transactions on Intelligent Systems 4(2), 100–106 (2009)
13. Lafferty, J., McCallum, A., Pereira, F.: Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In: Proceedings of the18th International Conference on Machine Learning, Waltham, Massachusetts, pp. 282–289 (2001)
14. Ku, L.W., Liu, I.C., et al.: Sentence-level opinion analysis by CopeOpi in NTCIR-7 (EB/OL), http://research.nii.ac.jp/ntcir/workshop/OnlineProceedings7/index.html
15. Qiong, W., Songbo, T., Gang, Z., et al.: Research on cross-domain opinion analysis. Journal of Chinese Information Processing 1, 77–83 (2010)
16. Qiong, W., Songbo, T., Haijun, Z., et al.: SentiRank: Cross-domain graph ranking for sentiment classification. In: Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology, pp. 309–314. IEEE Computer Society, Washington, DC (2009)
17. ChnSentiCorp (EB/OL), http://www.searchforum.org.cn/tansongbo/corpus-senti.htm