International Journal of Smart Home
Vol. 9, No. 3 (2015), pp. 23-34
http://dx.doi.org/10.14257/ijsh.2015.9.3.03

Effect Research of Aspects Extraction for Chinese Hotel Reviews Based on Machine Learning Method

Yan Zhao[1], Suyu Dong[2] and Jing Yang[3]*

[1,2]*College of Management, Inner Mongolia University of Technology, Huhhot, China*
[3]*College of Mechanics, Inner Mongolia University of Technology, Huhhot, China*
[1]*zhaoyanimut@126.com*, [2]*dongsuyu@126.com*, [3]*yangj81@126.com*

Abstract

With the development of Web, there are more and more online reviews. These reviews have become the important information for people. Therefore, automatic analysis of online reviews has become the active research area in natural language processing and management sciences. Apart from the researches about document sentiment classification and sentence sentiment classification, there are an increasing number researches about aspect-based sentiment analysis. This paper tries to use machine learning to extract aspect from Chinese hotel reviews. Through a lot of experiments, we find that machine learning methods are suitable for the aspect extraction of Chinese reviews. This paper adopts different dimensions of features, feature representation methods and classifiers to analyze the integral effect of aspect extraction. The experiment result shows that ME is the best classifiers and presence is most suitable feature representation method for aspect extraction.

Keywords: *Machine learning, Chinese hotel reviews, Aspect extraction*

1. Introduction

With the rapid development of Internet, there is a great deal of online reviews. Many websites allow people to express their opinion. At the same time, more and more hotels supply the same service for their customers to express their views about the hotel. These reviews make an effect on customers and hotel. The reviews can help customers to make wise decision. For hotels, reviews can provide detail information for managers to find shortcoming of hotel and make improvement. However, confronting with a large number of reviews which update at all times, it is difficult for people to complete the analysis by themselves. So, automatic analysis of online reviews has become an urgent work.

Up to now, sentiment analysis has become an active area for researchers [1]. Many researchers focus on document sentiment classification [2-3], sentence sentiment classification [4-5] and aspect-based sentiment classification [6-9]. Comparing with the researches about document sentiment classification and sentence sentiment classification, aspect-based sentiment analysis is more important. The real reason is that even though the targets for document and sentence sentiment classification are assumed as a single entity, a positive/negative opinion document/sentience does not mean the author has positive/negative opinion about all aspects of the entity. Aspect-based sentiment analysis can discover the sentiment on each aspect.

Aspect-based sentiment analysis, which also called feature-based opinion mining [10], two core tasks are aspect extraction and aspect sentiment classification [11]. Aspect extraction is

* Corresponding author

ISSN: 1975-4094 IJSH

the crucial work for aspect-based sentiment analysis. It because that the veracity of aspect extraction will have an impact on the aspect sentiment classification.

There are four main methods for aspect extraction, they are finding frequent nouns and noun phrases [8, 12-14], using opinion and target relations [15-17], using supervised learning [18-21] and utilizing topic models [22-24]. This paper will extract aspect using machine learning. We seek to exploit the effectiveness of using machine learning to extract aspect from Chinese hotel reviews.

2. Related Works

So far, aspect-based sentiment analysis has become an active area, researchers have proposed four main methods for this problem.

The method of finding frequent nouns and noun phrases is that frequent nouns and noun phrase are the aspect which would be extracted. Using this method is principally used for explicit aspect extraction. Hu and Liu [10] adopted finding frequent nouns and noun phrases to extract explicit aspects of the product. They used association miner which based on Apriori algorithm to extract all frequent nouns and noun phrases. In their work, 1% was the threshold and used for selecting candidate aspects. But association mining did not consider the position of item. So, some aspects which are extracted by association mining were not as the real aspects. They utilized compactness pruning to solve the problem. Otherwise, they use p-support (pure support) to get rid of redundant aspects. At last, they got the real aspects of product. Propescu and Etzioni adopted the method of computing the PMI (pointwise mutual information) of aspect to improve the precision of finding frequent nouns and noun phrases.

Zhuang, Jing and Zhu [15] used dependency parser to identify dependency relation between sentiment and aspect. This relation is used for aspect extraction from movie reviews. They firstly set up a keywords lexicon for the document which they would analyze. And then they applied dependency grammar graph to mine dependency relation between features and sentiment words. At last, when it identified feature-opinion pairs, they firstly made use of keywords lexicon to seek feature/opinion which from every sentence of reviews. After that, they utilized existing dependency relation to examine the path between features and opinion words. For the implicit feature-opinion expression, they handled it through two methods. For short sentence which has not more than three words, they thought movie/film was the target of sentiment words; for other expressions of implicit feature-opinion, given mappings which is between opinion words and aspect were used for feature extraction. The result of [15] show that the method has high efficiency.

Yu, *et al.,* [20] adopted supervised learning to extract aspects from product reviews. The products came from eleven classes of four areas. They used different methods for aspect extraction of different reviews. The reason is that the reviews which they dealt had different forms. For the prons and crons, they used finding frequent nouns and noun phrases to extract aspect. For free text reviews, they used supervised learning. When they extracted aspects from free text reviews, they firstly adopted grammar to identify nouns and these nouns were regarded as candidate aspects. And then it use the aspects which is identified by frequent nouns to train one-class SVM classifier to recognize real aspects. For attained aspects, they clustered synonym and got the unique binary aspects.

Topic models, in these years, an unsupervised learning method are widely used for information retrieve. Titov and McDonald [22] proposed Multi-grain Topic Models which is based on extended LDA [25] and PLSA [26] models. Because of LDA and PLSA model all use bag-of-words to represent text. They only can study document level problems. In addition, LDA and PLSA use diverse distribution and co-occurrence of text topics to identify targets. However, all reviews for given items have same topics. If using LDA and PLSA, the extracted aspects will have same topics. This deviates from the purpose of aspect extraction. So, based on LDA and PLSA, Titov and McDonald [22] use local topics to extract aspects and global topics to obtain entity.

This paper uses machine learning to extract aspects of hotel, aiming to prove the feasibility of using machine learning to extract aspects from Chinese reviews.

3. Theory Model

In this paper, each document is represented as a vector with feature weights. Let $\{t_1,t_2,\cdots,t_m\}$ be a predefined set of m features that can appear in a document. Let w_i is the feature weight in a document. Each document d is represented by the document vector $d=\{w_1,w_2,\cdots,w_m\}$.

3.1 Feature Weighting

The ability of every feature to distinguish document is different, and this ability can be measured by feature weighting. Feature weighting gets from the statistical information of documents. This paper will compare three feature weighting methods: Presence, TF and TF-IDF.

Presence is based on the feature whether or not appears in the text. If the feature appears in the document, the value is 1. Otherwise the value is 0. Even though presence can not represent the effect of features for the document, in different applications, Boolean is better than other feature weighting methods.

TF is widely used for text analysis. It uses the times of feature appearance in the text to represent the documents. TF maybe ignore some low-frequency features. But, sometimes many low-frequency features perhaps have the greater ability to distinguish the document. The low-frequency features will have an effect on document analysis.

TF-IDF is the most widely used feature weight calculation method for the text classification. It is based on the idea: if one feature has high-frequency, and rarely appears in other texts, then the feature has a good ability to distinguish.

3.2 Machine Learning

3.2.1 Naive Bayes: Naive Bayes classifier is widely used in the text classification, it uses the Bayes formula to calculate the probability of document d belongs to C_i, the equation is $P(C_i|d)=\frac{P(d|C_i)*P(C_i)}{P(d)}$. $P(C_i)$ denotes the probability of a document belonging to C_i. On the basis of the assumption of independence conditions, NB uses the joint probability between features and categories to estimate the probability of categories given a document, namely that $P_{NB}(C_i|d)=\frac{P(C_i)(\prod_{t_i\in V}P(t_i|C_i)^{W(t_i,d)})}{\sum_j[P(C_j)\prod_{t_i\in V}P(t_i|C_j)^{W(t_i,d)}]}$. Thereinto, feature t_i is independent of document d, $W(t_i,d)$ indicates the weights of feature t_i in document d. $P(t_i|C_i)$ indicates the Laplacean probability estimation value of conditional probability of documents belonging to C_i if it contains feature t_i. $P(t_i|C_i)$ is calculated by the following equation: $P(t_i|C_i)=\frac{1+W(t_i,C_i)}{|V|+\sum_j W(t_j,C_i)}$. $W(t_i,C_i)$ indicates the number of documents containing features t_i and belonging to C_i . $|V|$ is the size of $\{t_1,t_2,\cdots,t_m\}$, which are all features coming from all documents.

Although the assumption is harsh, NB performs well and is efficient in the text categorization.

3.2.2 Maximum Entropy Classifier: Maximum entropy classifier (ME) is based on the maximum entropy model. Its basic idea is that it does not make any hypothesis and remain maximum entropy for the unknown information. This is an advantage for maximum entropy compared with Naive Bayes. Maximum entropy model must satisfy the constraint of known information and the principle of maximum entropy. Hence, maximum entropy model is got through solving an optimization problem with constraints. The classical algorithm to solve this problem is Lagrange multiplier method. In this paper, we give the conclusion directly. The result is following:

$$p^*(C_i|t_i)=\frac{1}{\sum_{C_i}\exp\left(\sum_i \lambda_i f(t_i,C_i)\right)}\exp\left(\sum_i \lambda_i f(t_i,C_i)\right)$$

P^* indicates a predictive model for classification; V indicates the feature vectors; C_i indicates the type which the document belongs to. λ_i indicates the feature weight of feature vectors containing many feature t_i. $f(t_i,C_i)$is an indicator function.

3.2.3 SVM: Unlike Naive Bayes and maximum entropy, support vector machine (SVM) classifier is got by solving the optimal hyperplane represented by vector $\vec{w}$. The hyperplane is used to accomplish classification which can ensure maximum separation between a certain amount of data from the training set and hyperplane. Solving the maximum margin hyperplane eventually is converted into solving a convex quadratic programming problem.

Generally, it translates the above problem into the constrained optimization problem of dual variables through Lagrange Duality. The solution can be written as: $\vec{w}=\sum_{i=1}^{n}\alpha_i C_i \vec{d_i}$. C_i is the correct category for document $\vec{d_i}$. α_i are support vector and greater than zero.

What's more, for linear inseparable problems, kernel function can be used for SVM to convert low dimensional space nonlinear problem to a high dimension space linear problem. Mapping of kernel function can be a good control of the computational complexity of nonlinear expansion and can avoid the curse of dimensionality. There are many kernel functions: linear kernel, Gaussian kernel function, radial basis function and so on. In this paper, we used linear kernel function and optimize the parameter of SVM model, which will be used for following experiments.

4. Data Collection

Through a crawler, this paper acquired 3000 reviews from some hotel websites. Observing the download reviews, we found that different customers used various ways to express same content. Thus, we discussed with experts and determined that sum up different expressions to six aspects for the experiment of this paper. The six aspects are service, traffic, facilities, network, sanitation, price performance ratio.

The annotation of reviews adopts same rules. We labeled service, traffic, facilities, network, sanitation and price performance ratio use SE, TR, FA, NT, SA and PP respectively. If one aspect is contained in a review, it will be labeled by 1. On the contrary, one aspect is not contained in a review; it will be labeled by 0. For example, if “service” is contained in a review, the review will be labeled by SE-1; if not, the review will be labeled by SE-0.

5. Experiment

We adopt our own implementation for text pre-processing, NLPIR toolkit is used for Chinese text segmentation, McCallum’s Mallet toolkit [27] implementation of naive Bayes classifier and maximum entropy classifier and Chang’s LIBSVM [28] implementation of a Support Vector Machine classifier are used for aspect extraction. We ran each classifier with various feature representations and different number of features to experiment.

This paper used accuracy to measure the result of aspect extraction. For the effect of all aspects extraction, we used the mean value of accuracy to measure the result.The accuracy can be calculated according to Table 1. The calculation formula is following.

Table 1. Results of Experiments

	Actual Y	Actual N
Labeled Y	a	b
Labeled N	c	d

$$\text{Accuracy} = \frac{a+d}{a+b+c+d}.$$

5.1 The Experiment Results of ME

Figure 1-3 show the experiment results of ME with three feature representation methods and different number of features, the figures also display the mean value of accuracies of all aspects extraction. Figure 1 demonstrates the result of aspect extraction which uses ME and presence. Apart from little accuracy below 80%, most accuracies of every aspect extraction are higher. Accuracies of traffic and price performance ratio have achieved 95%, and the accuracies of network and service are about 90%. The statistical data of Figure 2 indicates the result of method using TF as the feature weighting method. Observing the figure, we found that the accuracy of traffic is also about 95%, other aspect extractions all about 90%. The result of ME with TF-IDF is worse than presence and TF. When using ME and TF-IDF, the best accuracy achieved by service extraction, other accuracies are most about 80%. For all results, comparing with other aspects, the accuracies of sanitation and facilities are lower.

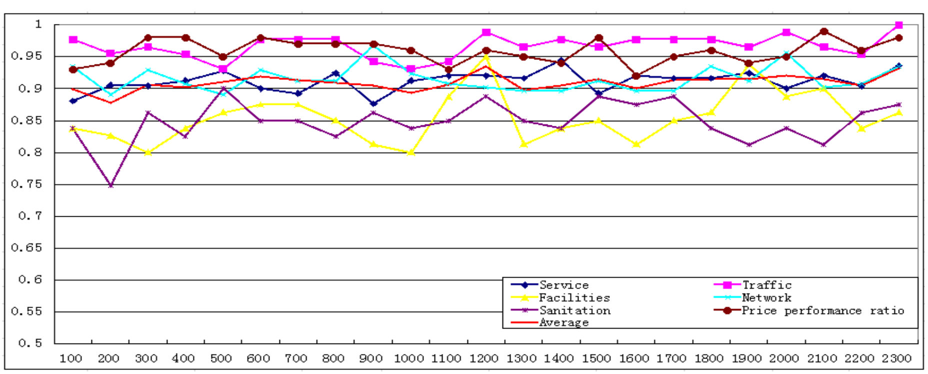

Figure 1. The Experiment Result of ME with Presence

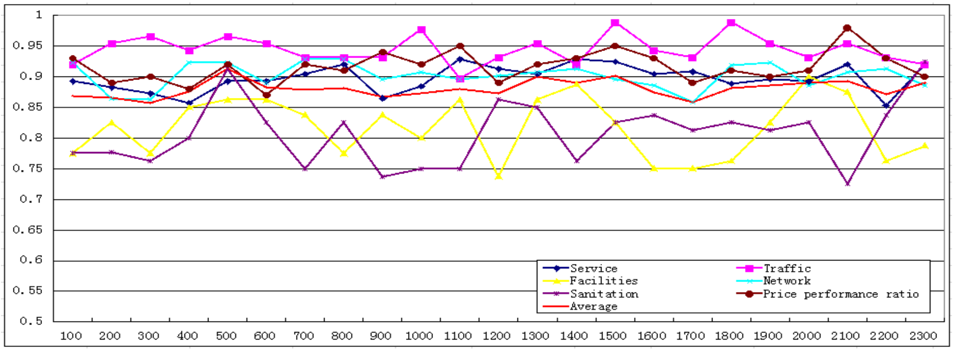

Figure 2. The Experiment Result of ME with TF

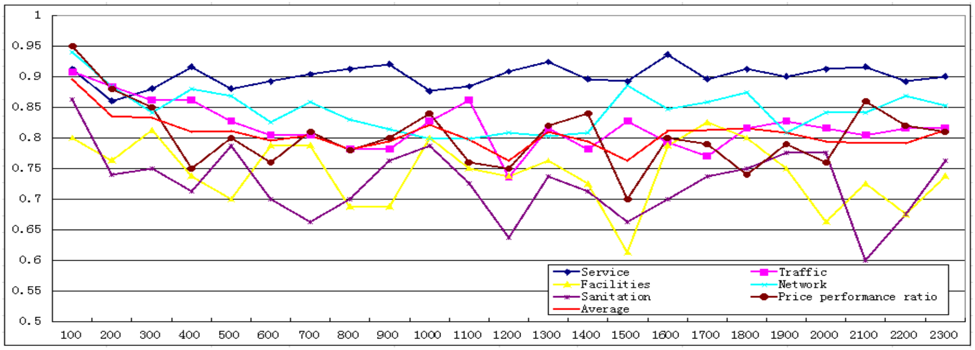

Figure 3. The Experiment Result of ME with TF-IDF

5.2 The Experiment Results of NB

Figure 4-6 show the experiment results of NB with three feature representation methods and different number of features, the Figures also display the mean value of accuracies of all aspects extraction. Overall, when using presence and TF as the feature representation methods, traffic extraction is better than other aspects, and the highest accuracy is about 96.6%. In addition, the accuracies of facilities and sanitation are lower, but more than 95% accuracies are higher than 75%. When using presence, the most accuracy is approximately 85%. Using TF as the feature representation method, almost all of accuracies of every aspect are about 80%. In the case of ME with TF-TDF, the experiment result is worse, the accuracies of all aspects are lower, and the accuracies of most aspects are about 75%. Network extraction attains the higher accuracy.

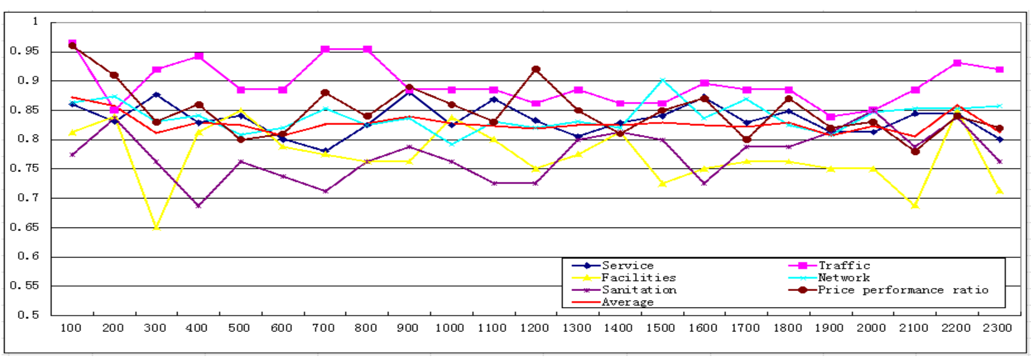

Figure 4. The Experiment Result of NB with Presence

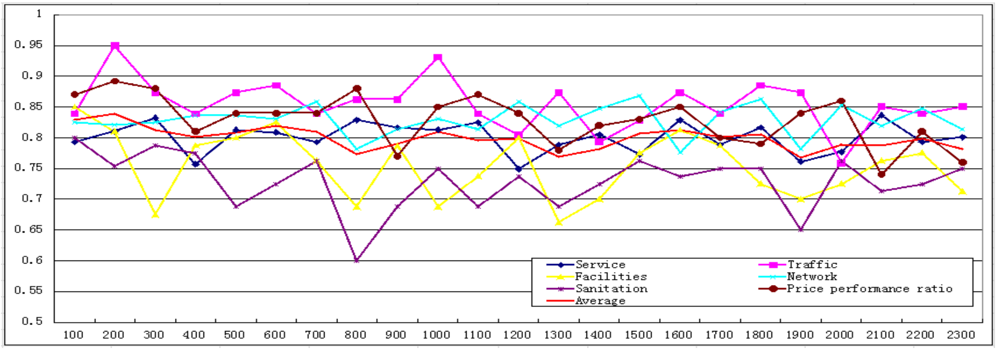

Figure 5. The Experiment Result of NB with TF

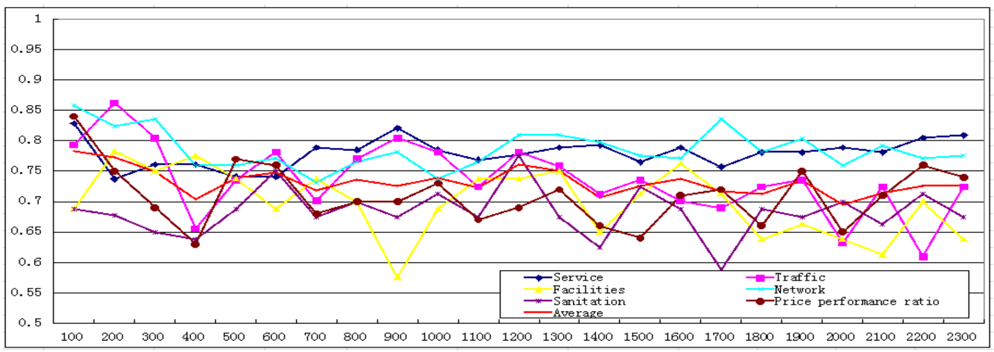

Figure 6. The Experiment Result of NB with TF-IDF

5.3 The Experiment Result of SVM

Figure 7-9 show the experiment results of NB with three feature representation methods and different number of features, the figures also display the mean value of accuracies of all aspects extraction. With three feature weighting methods, the accuracies of traffic are higher, and price performance ratio also attains the better accuracies. The accuracies of facilities and service are lower than other aspects. When we use presence as the feature representation method, Figure 7 displays that apart from service and facilities, the accuracies of other aspects are all greater than 80%, the accuracies of facilities are about 70%-75% and the most accuracies of service are more than 80%. Figure 8 shows the experiment result of SVM with TF, the most accuracy of other aspects are more than 75% besides the accuracies of facilities. In the case of using TF-IDF, apart from facilities, the statistical data indicate that almost accuracy is more than 70%.

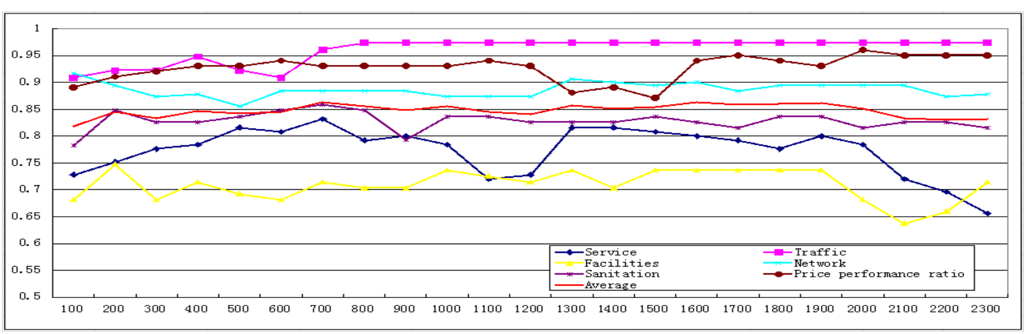

Figure 7. The Experiment Result of SVM with Presence

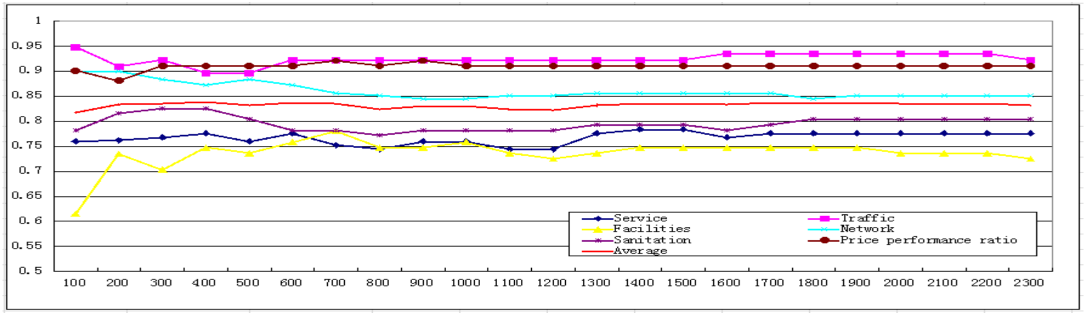

Figure 8. The Experiment Result of SVM with TF

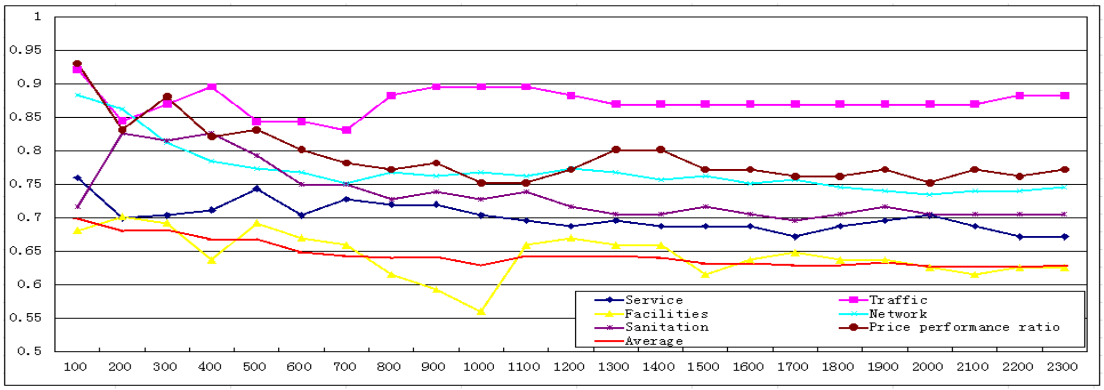

Figure 9. The Experiment Result of SVM with TF-IDF

5.4 Analysis of Experiment Result

5.4.1 Observing Figures 1-9, we found that there is a gap of the highest accuracy and lowest accuracy for different classifiers with three feature representation methods. All the highest and lowest mean values of accuracies for every classifier with different feature weighting methods are displayed by Figure 10. Considering the mean values of the highest accuracies, the ME is the best supervised learning method, as well as the following is NB, and SVM is the worst method. Comparing the highest and lowest mean values of accuracies, three supervised learning methods all with presence achieve their best accuracy and with TF-IDF attain the worst accuracy. The maximum values of highest accuracy of different supervised learning method are 93.47% (ME), 87.82% (NB) and 86.34% (SVM). For the minimum values of three supervised learning methods, the order is ME (76.30%)> NB (69.47%) > SVM (62.72%). Using presence as the feature representation method, the lowest accuracy of ME, NB and SVM are 87.55%, 80.62% and 81.82%. When three supervised methods combine with TF, ME achieves the top of highest and lowest accuracy, they are 91.27% and 85.65%. For NB and SVM, the gap of highest accuracy is small, and the lowest accuracy of SVM (81.79%) is larger than NB (76.77%). When TF-IDF is used as the feature weighting method, the highest accuracy of ME is more 10% and 20% than NB and SVM, respectively.

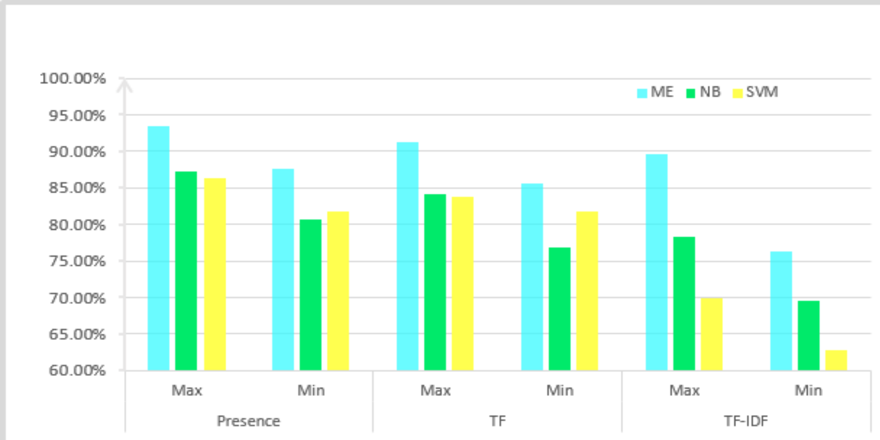

Figure 10. The Highest and Lowest Accuracy of Three Supervised Learning with Different Feature Representation Methods

5.4.2 Figure 10 displays that the minimum value of highest accuracy 69.94% is achieved by SVM with TF-IDF. Observing the whole effect of aspect extraction, apart from the accuracy 69.47% attained by NB with TF-IDF, other accuracies are all more than 69.47%. In addition, besides the accuracy achieved by ME with TF-IDF and NB with TF, other accuracies are all above the highest accuracy 78.25% which reached by NB with TF-IDF. Therefore, the accuracy 69.94% and 78.25% should not be the reference standard for other supervised learning methods. The data of Table 2 show the least value of the number of features, which is used for all supervised learning methods attain the accuracy that is no less than 80.10% (80.10% is the accuracy achieved by NB with TF). The statistical data indicate that all supervised learning methods use 100 features to reach 80.10% or higher accuracy. Moreover, for every classifier, the accuracy achieved by presence is larger than of TF and TF-IDF. For ME, the accuracies of presence and TF-IDF are 89.95% and 89.55%, the gap between them is small. In addition, these accuracies are more than the accuracies achieved by other classifiers with different feature weighting methods and same number of features, and are higher 3% than ME with TF. For SVM, the gap between presence and TF is also small, but they are all lower than ME and NB with presence or TF. The accuracy for NB with presence is larger 4% than NB with TF.

Table 2. The Feature Dimensions and Accuracy When the Mean Value of Accuracy is Not Less than 80.10% in First Time

	Presence	TF	TF-IDF
ME	100（89.95%）	100（86.92%）	100（89.55%）
NB	100（87.28%）	100（82.95%）	-
SVM	100（81.82%）	100（81.79%）	-

5.4.3 Observing Figures 1-9, we find that the fluctuation of accuracy is different for different supervised learning with three feature weighting methods. For every classifier with different feature representation methods, they have their stable interval. In addition, the mean value of accuracy is different. In this paper, the stable interval is defined as an interval in which no less than 12 feature dimensions (50% of all feature dimensions), and the gap between 95% accuracy is not more than 3%. Table 3 shows the stable interval of every classifier with different feature representation methods, at the same time, the mean value of accuracy is listed. Through the statistical data, we know that there is no stable interval for ME with TF-IDF and NB with TF, TF-IDF. For ME, using presence has larger stable interval and mean value of accuracy than TF. In the case of NB with presence, the stable interval is similar with ME with presence, but the mean value of accuracy is less 9% than ME. SVM with three feature representation methods all have stable interval. When SVM with presence and TF, the stable interval is larger, the mean value of accuracy is 85.06% and 83.24% respectively, and is better than NB.

Table 3. The Stable Interval and Mean Value of Accuracy of Three Machine Learning Methods

	Presence		TF		TF-IDF	
	Stable interval	The mean value of accuracy	Stable interval	The mean value of accuracy	Stable interval	The mean value of accuracy
ME	[300,2200]	91.01%	[100,1200]	87.68%	-	-
NB	[300,2100]	82.24%	-	-	-	-
SVM	[200,2100]	85.06%	[100,2300]	83.24%	[600,2300]	63.58%

5.4.4 As is shown in Figure 10, we find that three supervised learning all with presence achieve the highest mean value of accuracy. Moreover, Table 1 indicates when we use presence, all supervised learning use same number of features attain the higher accuracy. Therefore, considering aspect extraction from hotel reviews, presence is the best feature weighting method.

5.4.5 The experiment results indicate that ME, NB and SVM are all suitable for aspect extraction. With the increasing of feature dimensions, the most accuracy decline and then keep stable or fluctuate in an interval (only the accuracy achieved by ME and SVM with presence is different). The experiment results are different with sentiment classification [29]. In the experiment of aspect extraction, a small number of features can include the necessary features for aspect extraction. If we increase the number of features, the most increased features are infrequent features and noisy data for aspect extraction. Therefore, with the increasing of feature dimensions, the accuracy will reduce. The primary cause is the texts which are dealt by this paper have large noisy data [30].

6. Conclusions

Up to now, there are more and more researches about sentiment classification. Comparing with document and sentence sentiment classification researches, aspect-based sentiment analysis researches are less. In the work of aspect-based sentiment analysis, aspect extraction is the first and core task. This paper focuses on aspect extraction of Chinese hotel reviews.

For aspect extraction of Chinese hotel reviews, ME is the best machine learning method. Comparing with other machine learning methods with all feature representation methods, the ME achieves the top of highest accuracy and lowest accuracy. In addition, when using the same number of features, the accuracy of ME is highest. At the same time, the stable interval of ME is bigger than other classifiers, as well as the mean value of accuracy in stable interval is largest. For three feature representation methods, comparing the highest and lowest accuracy, the mean value of accuracy using same number of features, and the stable interval, presence is the best feature weighting method. The experiment results also prove that noisy data maybe result in the decline of accuracy. If texts include a large number of noisy data, the accuracy will reduce with the increasing of feature dimension.

Acknowledgements

This work was mainly supported by National Natural Science Foundation of China (71363038), Natural Science Foundation of Inner Mongolia, China(2012MS1008, 2013MS1009), Scientific Research Project of Colleges and universities in Inner Mongolia, China (NJSZ12047), Promotion Plan of Research Center on Modern Management of Inner Mongolia (2013-2015), Scientific Research Project of Inner Mongolia University of Technology (X201302).

References

[1] B. Pang and L. Lee, “Opinion mining and sentiment analysis”, Foundations and trends in information retrieval, vol. 2, no. (1-2), **(2008),** pp. 1-135.
[2] B. Pang, L. Lee and S. Vaithyanathan, “Thumbs up?: sentiment classification using machine learning techniques”, Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10. Association for Computational Linguistics, **(2002),** pp. 79-86.
[3] P. D. Turney, “Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews”, Proceedings of the 40th annual meeting on association for computational linguistics. Association for Computational Linguistics, **(2002),** pp. 417-424.
[4] O. Täckström and R. McDonald, “Semi-supervised latent variable models for sentence-level sentiment analysis”, Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers-Volume 2. Association for Computational Linguistics, **(2011),** pp. 569-574.
[5] D. Davidov, O. Tsur and A. Rappoport, “Enhanced sentiment learning using twitter hashtags and smileys”, Proceedings of the 23rd International Conference on Computational Linguistics: Posters. Association for Computational Linguistics, **(2010),** pp. 241-249.
[6] L. Jiang, M. Yu, M. Zhou, *et al.*, “Target-dependent twitter sentiment classification”, Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1. Association for Computational Linguistics, **(2011),** pp. 151-160.
[7] X. Ding, B. Liu and P. S. Yu, “A holistic lexicon-based approach to opinion mining”, Proceedings of the 2008 International Conference on Web Search and Data Mining. ACM, **(2008),** pp. 231-240.
[8] L. Zhuang, F. Jing and X. Y. Zhu, “Movie review mining and summarization”, Proceedings of the 15th ACM international conference on Information and knowledge management. ACM, **(2006),** pp. 43-50.
[9] J. Zhu, H. Wang, M. Zhu, *et al.*, “Aspect-based opinion polling from customer reviews”, Affective Computing, IEEE Transactions on, vol. 2, no. 1, **(2011),** pp. 37-49.
[10] M. Hu and B. Liu, “Mining and summarizing customer reviews”, Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining, ACM, **(2004),** pp. 168-177.
[11] B. Liu, “Sentiment analysis and opinion mining”, Synthesis Lectures on Human Language Technologies, vol. 5, no. 1, **(2012),** pp. 1-167.
[12] S. Blair-Goldensohn, K. Hannan, R. McDonald, *et al.*, “Building a sentiment summarizer for local service reviews”, WWW Workshop on NLP in the Information Explosion Era, **(2008),** pp. 14.
[13] S. Moghaddam and M. Ester, “Opinion digger: an unsupervised opinion miner from unstructured product reviews”, Proceedings of the 19th ACM international conference on Information and knowledge management, ACM, **(2010),** pp. 1825-1828.
[14] A. M. Popescu and O. Etzioni, “Extracting product features and opinions from reviews”, Natural language processing and text mining. Springer London, **(2007),** pp. 9-28.
[15] L. Zhuang, F. Jing and X. Y. Zhu, “Movie review mining and summarization”, Proceedings of the 15th ACM international conference on Information and knowledge management, ACM, **(2006),** pp. 43-50.
[16] N. Kobayashi, R. Iida, K. Inui0, *et al.*, “Opinion mining as extraction of attribute-value relations”, New Frontiers in Artificial Intelligence. Springer Berlin Heidelberg, **(2006),** pp. 470-481.
[17] Y. Wu, Q. Zhang, X. Huang, *et al.*, “Phrase dependency parsing for opinion mining”, Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 3-Volume 3. Association for Computational Linguistics, **(2009),** pp. 1533-1541.

[18] W. Jin, H. H. Ho and R. K. Srihari, “Opinion Miner: a novel machine learning system for web opinion mining and extraction”, Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining, ACM, **(2009),** pp. 1195-1204.
[19] N. Jakob and I. Gurevych, “Extracting opinion targets in a single-and cross-domain setting with conditional random fields”, Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing Association for Computational Linguistics, **(2010),** pp.1035-1045.
[20] J. Yu, Z. J. Zha, M. Wang, *et al.*, “Aspect ranking: identifying important product aspects from online consumer reviews”, Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1. Association for Computational Linguistics, **(2011),** pp. 1496-1505.
[21] G. Somprasertsri and P. Lalitrojwong, “Automatic product feature extraction from online product reviews using maximum entropy with lexical and syntactic features”, Information Reuse and Integration, 2008 IRI 2008, IEEE International Conference on. IEEE, **(2008),** pp. 250-255.
[22] I. Titov and R. McDonald, “Modeling online reviews with multi-grain topic models”, Proceedings of the 17th international conference on World Wide Web. ACM, **(2008),** pp. 111-120.
[23] C. Lin and Y. He, “Joint sentiment/topic model for sentiment analysis”, Proceedings of the 18th ACM conference on Information and knowledge management. ACM, **(2009),** pp. 375-384.
[24] A. Mukherjee and B. Liu, “Aspect extraction through semi-supervised modeling”, Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Long Papers-Volume 1, Association for Computational Linguistics, **(2012),** pp. 339-348.
[25] S. C. Deerwester, S. T. Dumais, T. K. Landauer, G. W. Furnas and R. A. Harshman, Indexing by latent semantic analysis. Journal of the American Society of Information Science, vol. 41, no. 6, **(1990),** pp. 391–407.
[26] T. Hofmann, “Unsupervised Learning by Probabilistic Latent Semantic Analysis”, Machine Learning, vol. 42, no. 1, **(2001),** pp. 177–196.
[27] A. K. McCallum, "MALLET: A Machine Learning for Language Toolkit", **(2012),** http://mallet.cs.umass.edu.
[28] C.-C. Chang and C.-J. Lin, “LIBSVM: a library for support vector machines”, ACM Transactions on Intelligent Systems and Technology, 2:27:1--27:27, **(2011)**, Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.
[29] Z. Zhang, Q. Ye, Z. Zhang, *et al.*, “Sentiment classification of Internet restaurant reviews written in Cantonese”, Expert Systems with Applications, vol. 38, no. 6, **(2011),** pp. 7674-7682.
[30] M. Gamon, “Sentiment classification on customer feedback data: noisy data, large feature vectors, and the role of linguistic analysis”, Proceedings of the 20th international conference on Computational Linguistics, Association for Computational Linguistics, **(2004),** pp. 841.