NEUOM: Identifying Opinionated Sentences in Chinese and English Text

Chunliang Zhang, Ke Wang, Muhua Zhu, Tong Xiao, Jingbo Zhu
Natural Language Processing Lab, Northeastern University
{zhangcl, xiaotong, zhujingbo}@mail.neu.edu.cn
wangke@ics.neu.edu.cn, zhumuhua@gmail.com

Abstract

This paper introduces our NEUOM system which participates in the opinionated sentence detection task, one of evaluation tasks in Multilingual Opinion Analysis Task (MOAT) of NTCIR-7. NEUOM system adopts a sentiment lexicon-based(SLB) approach to identifying opinionated sentences in a Chinese text and English text. For English task, a machine learning algorithm, naïve Bayesian classification model, is also tried with the use of the English training corpora, such as MPQA and NTCIR-6 data set. Experimental results show that in the English task SLB method achieved better F1 performance than Naïve Bayesian model.

Keywords: *Opinionated Sentence detection, sentiment lexicon, Naïve Bayesian classification model*

1. Introduction

Opinion analysis has received lots of concerns in text mining research in recent years. With the advance in Internet technologies, BBS or personal blogs are becoming a popular medium for expressing various public opinions such as product reviews, stock market predictions and social issue discussions. An automatic approach to opinion analysis is quite desirable since it can extract online reviews and thus help make instant response for some commercial applications, like product review analysis, public opinion survey generation, opinion summarization and question answering.

The Multilingual Opinion Analysis Task(MOAT), which consists of 6 subtasks—opinionated sentence detection, opinionated unit detection, opinion polarity classification, opinion holder detection, opinion target detection and topic relevance detection, provides a platform to evaluate various opinion analysis techniques. We participate in opinionated sentence detection subtask, and submit 2 runs for this subtask in English and 1 run for Chinese(Simplified). This paper reports our methods and their results of each run.

The rest of the paper is structured as follows. Section 2 presents our approach to identifying opinionated sentences in a natural text. Then we present the results of our system in two languages in Section 3. At last, we address conclusions in Section 4 with discussion of future work.

2. Methods

In this section, we describe our system briefly. In our system two sentiment lexicon based approaches and a naïve Bayesian classification model are adopted respectively.

2.1 Sentiment Lexicon-based Approach(SLB)

The motivation behind the sentiment lexicon-based approach is that an opinionated sentence can be identified if it contains an opinionated word or an opinionated fragment. Hatzivassiloglou and Wiebe(2000) reported that adjectives are good indicators for subjective sentences. In the work of Ellen Riloff *et al.* (2003), subjective nouns are viewed as important subjectivity features for opinionated sentence detection. Actually adjectives, adverbs, verbs and nouns can all be used to feature a subjective expression, as discussed in the following example.

1) It's ***unacceptable*** for Japan's leader to visit Yasukuni Shrine where Class-A war criminals are enshrined," Tang was quoted as ***telling*** secretaries general of three ruling coalition parties in Beijing.

2) 对此，马利基***说***：“这***绝对***不是***报复***行为。”

In sentence 1), both the adjective “unacceptable” and the verb “telling” indicate the subjectivity of the English sentence. In sentence 2), the verb “说”, the adverb “绝对” and the noun “报复” can convey that this Chinese sentence is opinionated.

To implement SLB method, a sentiment dictionary should be constructed in advance, which is comprised of a large number of sentiment words such as adjective “unacceptable”. A well-established sentiment dictionary is of great help to identify opinionated sentence.

2.1.1 Acquisition of Sentiment Lexicons in English

OpinionFinder[1] is an open source system for opinion analysis task, which can automatically analyze documents and extract subjective sentences. The system source of OpinionFinder provides feature files which contain words and fragments called subjective clues. We extract the words from the feature files as our sentiment lexicons for our lexicon-based method. Altogether 8,203 words are extracted to build our English sentiment dictionary. We use these feature words to fulfill the opinionated sentence detection task.

[1] http://www.cs.pitt.edu/mpqa/opinionfinderrelease/

Table1. The format of sentiment dictionary (English)

word	value
abhorrent	1
……	…
amuse	1
……	…

The left column of the table above shows sentiment words used in our system, and the right column corresponds with a weight of each feature word. Our system initially sets the weight of each feature word as 1, and other non-sentiment words not listed in the table as 0.

2.1.2 Acquisition of Sentiment Lexicons in Chinese(Simplified)

Hownet[2] is an online commonsense knowledge base which has been widely used for Chinese language processing applications, such as word sense disambiguation, information extraction and topic analysis. Recently, Hownet provides a version of Chinese/English Vocabulary for Sentiment Analysis(VSA, Beta version)[3] purpose, in which sentiment words are categorized into six classes such as "Plus Feeling", "Minus Feeling", "Plus Sentiment", "Minus Sentiment", "opinion", and "degree".

We first build our Chinese sentiment dictionary by extracting sentiment words from Hownet belonging to the above six classes, and removing the single-character Chinese lexicons out of Chinese VSA, since the single Chinese characters are unreliable to indicate whether the sentence is opinionated. After that, 7,757 words are extracted to generate our Chinese sentiment dictionary. The format of the Chinese feature words is the same as the English one. Here we give an example to explain the problem of the single character Chinese word as follows:

3) 她的皮肤好*白*啊！

4) 屋子的墙壁被刷上了一层*白*漆。

For the same single character word "白", in sentence 3) it can be taken as a subjective feature referring to a positive sentiment toward "皮肤", while in sentence 4) it is just an adjective to describe the color. Therefore, it is quite uncertain for such a lexicon to express opinions.

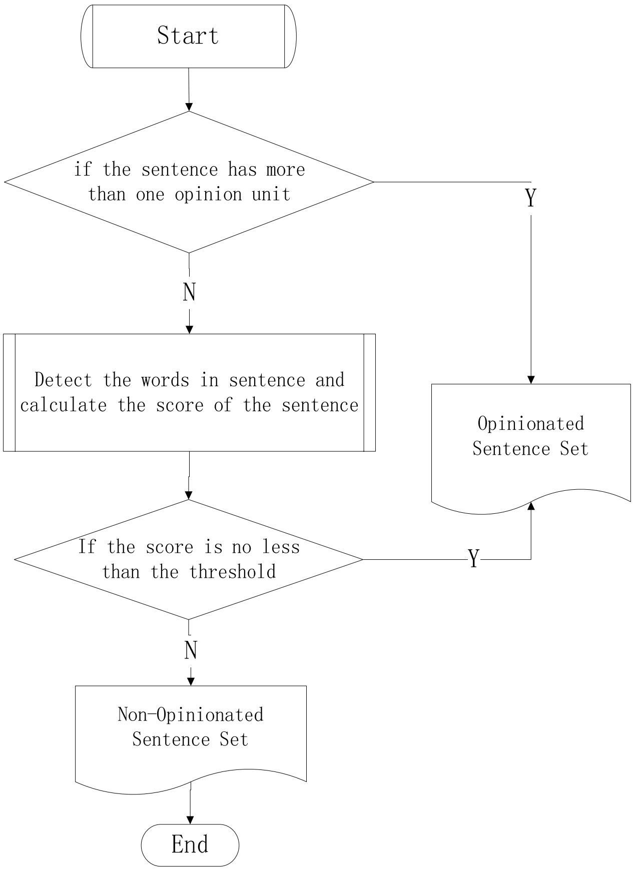
Table2. The format of sentiment lexicon list (Chinese Simplified)

Word	value
百孔千疮	1
……	…
百听不厌	1
……	…

2.1.3 Opinionated Sentence Identification

Opinion units in each sentence have been marked out in the test data provided by NTCIR, and the official NTCIR standards tell that the sentence is opinionated if it contains more than one opinion unit. As default, our system labels the sentences containing more than one opinion unit as "Y" indicating the sentence is opinionated, and then outputs these sentences with Y label into the result set. The procedure of the opinionated sentence identification algorithm used in our system is summarized as follows.

Figure1. The procedure of our lexicon-based method


Start
if the sentence has more than one opinion unit
Y
N
Detect the words in sentence and calculate the score of the sentence
Opinionated Sentence Set
If the score is no less than the threshold
Y
N
Non-Opinionated Sentence Set
End

In the lexicon-based method, the score of an unlabeled sentence is calculated by the following formula:

$$Score(S_i) = \sum value(w_j) \qquad (1)$$

The w_j stands for j-th word in the sentence S_i. A word of value (i.e. weight) zero means that it is not defined in our sentiment dictionary. Otherwise, the weight of each word can be retrieved from our sentiment dictionary. The score of a sentence can be calculated as sum over the scores of all words appearing in the sentence. A sentence with a larger score than the predefined threshold is taken as opinionated. The predefined threshold is set as 1.0 in our experiments.

[2] http://www.keenage.com/

2.2 Naïve Bayesian classification approach(NB)

The opinionated sentence identification aims to detect whether a sentence in the natural text is opinionated or not. Therefore, the task can be taken as a problem of binary classification.

In the English task, we take Multi-Perspective Question Answering(MPQA)[3] corpus and NTCIR-6 English corpus as our training data. Due to lack of training data in Chinese (Simplified), we do not use the framework of classification in Chinese.

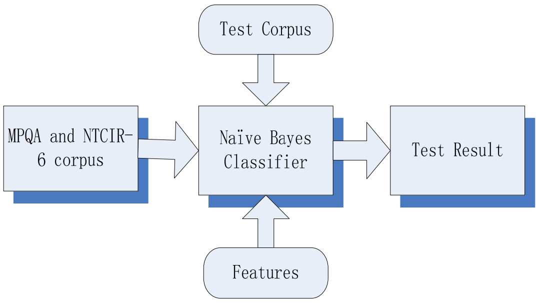
In this work, Naïve Bayes(NB) classification model is utilized to design the classifier for opinionated sentence task, which has been wildly used for text classification tasks. The formula used is as follows:

$$C^* \equiv \arg\max p(C_j)\prod_{t=1}^{|v|} p(w_t \mid C_j)^{n(w_t)} \qquad (2)$$

where $p(C_j)$ is the class priori probabilities, $|V|$ is the size of vocabulary, w_t is the t^{th} word in the vocabulary, $n(w_t)$ is the frequency of a specific word w_t in the given sentence, and $P(w_t \mid C_j)$ represents the associated probability of w_t under the class C_j.

In our system, we first do feature selection using the Information Gain(IG) method from the training data, and extract the words in both the IG feature list and sentiment lexicons list as the features in the classification. The whole procedure is shown as follows:

Figure 2. The classification system



As shown in the figure, a naïve Bayesian text classification framework is adopted to solve the problem.

3. Results and Discussion

As the NTCIR evaluation principle, evaluation will be reported against two gold standards: the strict and lenient standards. Precision, Recall, and F-measure will be reported over each sub-tasks.

3.1 The result with Chinese(Simplified)

We use SLB approach in Chinese(Simplified), achieving the following result.

Table 3. Results of Opinionated Sentence Identification with Chinese(simplified)

	Chinese	Precision	Recall	F1
SLB	Lenient	0.4721	0.7116	0.5676
	Strict	0.4358	0.7339	0.5469

As shown in table 3, the precision of our approach is relatively low, while the recall is high, which indicates that some sentiment words in our dictionary are noise. These noise sentiment words would cause negative effect on precision performance. It is worth studying how to assign a proper sentiment weight to each word in our dictionary which indicates the ability of reflecting its sentiment polarity. The sentiment weight of each word can be calculated from a pre-given training corpus. However, constructing such a large-scale sentiment corpus is time-consuming.

In real world applications, many sentiment words have multiple polarities such as positive, negative and neutral. For example, a sentiment word “下降”(decrease) shows positive polarity with the collocation of a context word “成本”(cost), and negative with a context word “利润”(profit). Therefore, the second critical problem of a dictionary-based method is how to determine the appropriate polarity of a sentiment word in the opinionated sentence detection tasks, for a sentiment word of negative or positive polarity is a strong indicator for an opinionated sentence. Similarly, a sentiment word is not an opinion indicator if it shows neural polarity, though included in the sentiment dictionary. The sentiment word “下降”(decrease) will be neutral if it is collocated with the word “趋势”(trend). Hereby we think such collocations are very helpful for opinionated sentence detection, rather than single sentiment words. We will study this issue of how to use collocations for improve opinionated sentence identification.

3.2 The result with English

For English, we test two methods in the task: SLB and NB, and the results are shown as below.

Table 4. Results of Opinionated Sentence Identification with English

	English	Precision	Recall	F1
SLB	Lenient	0.352	0.779	0.485
	Strict	0.110	0.820	0.195
NB	Lenient	0.295	0.899	0.444
	Strict	0.088	0.901	0.161

Table 4 depicts the accuracy performances of SLB and NB approaches to English opinionated sentence detection tasks respectively. Seen from Table 4, NB achieves higher recall values than SLB in both lenient and strict standards. However, it surprises us that SLB achieves higher F1 values than NB in both standards. Since NB is one of machine learning algorithms, its performance depends upon the construction of training data. Many machine learning algorithms achieve good accuracy performance in document classification tasks, for example, classifying documents into some predefined categories such as *Sports* or *Education*. Roughly speaking, classes Sport and Education are two different topics. That is to say, machine learning algorithms such as NB are powerful tools to discriminate different topic classes. However, in opinionated sentence detection task, an opinionated sentence indicates a positive or negative polarity, and a non-opinionated sentence indicates neutral polarity. In most cases, both opinionated and non-opinionated sentences in the same document refer to the same topic. We think opinionated sentence identification can not be simply viewed as a document classification task; rather, it should be a semantic-level classification task. In the future work we will study to adopt semantic feature with machine learning techniques for opinionated sentence identification.

4. Conclusions and Future Work

This paper presents some details on two techniques used in our opinionated sentence detection system participating in NTCIR-7, such as SLB method and NB classification. For Chinese, we only apply the SLB method to fulfill the task due to lack of sufficient training data. Both methods achieve good recall performance in Chinese and English opinionated sentence detection. However, the precision performance is not satisfied. In future work, we will focus on the construction of a sentiment collocation dictionary and semantic-level sentiment classification to improve the performance of opinionated sentence identification, particularly for precision performance.

References

[1]Dong Z. D., Dong Q.. "Hownet," http://keenage.com, 2000.

[2]Hatzivassiloglou, V. and McKeown K. 1997. Predicting the Semantic Orientation of Adjectives. 3. In Proceedings of ACL-2007: pages 174–181.

[3]Hatzivassiloglou V. and Wiebe J.M. 2000. Effects of Adjective Orientation and Gradability on Sentence Subjectivity. In Proceedings of ACL 2000: pages 299-305

[4]Hu M.Q. and Liu B. 2004. Mining and Summarizing Customer Reviews. In Proceedings of ACM-KDD 2004: pages 168-177.

[5]Kim S.M.and Hovy E. 2004. Determining the Sentiment of Opinions. In Proceedings of COLING 2004: pages 1367-1373.

[6]Pang B., Lee L., and Vaithyanathan S. 2002. Thumbs up? Sentiment Classification Using Machine Learning Techniques. In Proceedings of EMNLP 2002: pages 79-86.

[7]Riloff E., Wiebe J., and Wilson T. 2003. Learning Subjective Nouns Using Extraction Pattern Bootstrapping. In Proceedings of CoNLL-2003: pages 25-32.

[8]Riloff E. and Wiebe J. 2003. Learning Extraction Patterns for Subjective Expressions. In Proceedings of EMNLP-2003: pages 105-112.

[9]Turney P. D. 2002. Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. In Proceedings of ACL-2002: pages 417–424.

[10]WIEBE, J. Learning Subjective Adjectives From Corpora. 2000. In Proceedings of AAAI-2000: pages 735–740.

[11]WilsonT., Hoffmann P., Somasundaran S., Kessler J., Wiebe J., Choi Y., Cardie C., Riloff E., Patwardhan S. 2005. OpinionFinder: A System For Subjectivity Analysis. In Proceedings of HLT/EMNLP 2005: pages 34–35.

[12]Yu H. and Hatzivassiloglou V. 2003. Toward Answering Opinion Questions: Separating Facts from Opinions and Identifying the Polarity of Opinion Sentences. In Proceedings of EMNLP-2003: pages 129-136.