Using Polarity Scores of Words for Sentence-level Opinion Extraction

Lun-Wei Ku, Yong-Sheng Lo and Hsin-Hsi Chen
Department of Computer Science and Information Engineering
National Taiwan University
Taipei, Taiwan
{lwku, yslo}@nlg.csie.ntu.edu.tw; hhchen@csie.ntu.edu.tw

Abstract

The opinion analysis task is a pilot study task in NTCIR-6. It contains the challenges of opinion sentence extraction, opinion polarity judgment, opinion holder extraction and relevance sentence extraction. The three former are new tasks, and the latter is proven to be tough in TREC. In this paper, we introduce our system for analyzing opinionated information. Several formulae are proposed to decide the opinion polarities and strengths of words from composed characters and then further to process opinion sentences. The negation operators are also taken into consideration in opinion polarity judgment, and the opinion operators are used as clues to find the locations of opinion holders. The performance of the opinion extraction and polarity judgment achieves the f-measure 0.383 under the lenient metric and 0.180 under the strict metric, which is the second best of all participants.

Keywords: *Opinion Extraction, Sentiment Mining*

1 Introduction

The processing of opinion information has been widely discussed these days. People are concerned about opinions, and this makes the techniques of opinion information processing practical. Generally speaking, opinions are divided into three categories: positive, neutral and negative. Opinions of different polarities in documents are useful references or feedbacks for governments or companies helping them improve their services or products [2].

Opinions are usually about a theme, and are viewed after grouping by the target which opinions toward to, the opinion holders or the opinion polarities. Therefore, for applications, in spite of the opinion sentence extraction and polarity judgment, the opinion holder identification and the relevance judgment are also important. To extract the relevant opinion sentences, techniques of relevant sentence retrieval are vital. One of the three major conferences, TREC, tried to survey these techniques by having the novelty track. [10] However, this task is proven to be tough because of the lack of information in only one sentence. Moreover, extracting opinion holders is beyond extracting named entities. All named entities, pronouns, and job titles are candidates for opinion holders. Even if all these entities can be extracted, we still need to decide which of them are holders of opinions. In order to group opinions of the same holders, techniques for anaphor and coreference resolution must be applied. These issues raise the degree of difficulty of opinion information processing.

Many researchers have started the study of opinion information processing. Generally speaking, the unit for opinion information can be one document, one sentence, or a single word. Wiebe, Wilson and Bell [9] and Pang, Lee, and Vaithyanathan [6] processed opinion documents and their sentiment or opinion polarities. Researches of extracting opinions in documents of a specific genre, reviews, also use one document as their judging unit. Dave's and Hu's researches both focused on extracting opinions of 3C product reviews [2][3], while Bai, Padman and Airoldi [11] use movie reviews as experimental materials. As for sentences, they are the basic unit for a person to express a complete idea. Riloff and Wiebe distinguished subjective sentences [7], while Kim and Hovy proposed a sentiment classifier for English words and sentences [4]. Of course, the composed opinion words must be recognized first to process opinion documents and sentences. Riloff, Wiebe and Wilson [12] learned opinion nouns from patterns, and Takamura, Inui and Okumura [8] adopt a physical model to decide opinion polarities of words.

Many techniques of NLP were also studied for opinion information processing. Machine learning approaches such as Naive Bayes, maximum entropy classification, and support vector machines have been investigated [6]. Both information retrieval [2] and information extraction [1] technologies have also been explored. However, various metrics and testing beds are employed, which leads to incomparable results. Building common testing sets and evaluation

metrics are always important, and these are what NTCIR provides. With the equivalent testing documents under the same evaluation metrics, it is possible to find the pros and cons of each technique, and also the way to enhance the performance.

We proposed our method of opinion information processing for NTCIR pilot task in this paper. A Chinese opinion extraction system is introduced, and the components in this system are used to deal with subtasks of the pilot task. Frequency-based formulae are adopted in the system kernel to calculate the opinion scores of words and sentences, which tell whether sentences are opinionated and if so, their opinion polarities. Evaluation results are shown and compared with the other participants. At last, a discussion of the performance is also included.

2 An Chinese Opinion Extraction System: CopeOpi

The **C**hinese **op**inion **e**xtraction system for **op**inionated **i**nformation (CopeOpi) is a web-based system developed from news documents. This system works on a large set of documents. It can extract sentiment words, sentences and documents. Moreover, opinion summarization is also one of its functions. Based on opinion summaries generated everyday, it tracks opinions toward a specific topic and generate a tracking plot for visualization. The tracking topic is in the format of a query in this system, so the user can easily find opinions they concern. The detail framework of this system is introduced in [13], and an example of the tracking plots it outputs is shown in Figure 1.

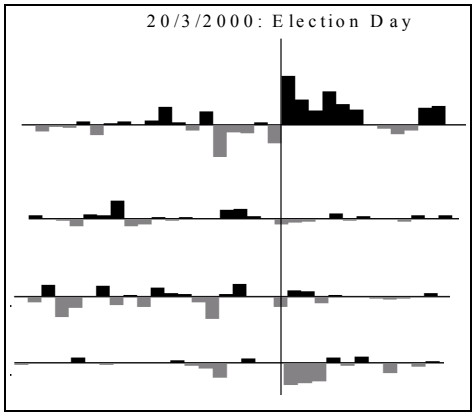
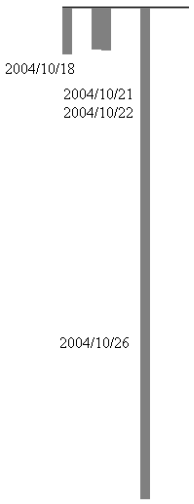


Figure 1. Opinions towards four persons in presidential election

Each bar shows the summarized opinion score of one day, and the x-axis is the timeline. Black bars show the positive opinion scores, and the gray bars show the negative opinion scores. The first person is the president elect. From these plots, we can observe the reputation of four candidates before and after the election. This tracking system also tracks opinions according to different information sources, including news agencies and the Web. Therefore, the results of opinion analysis can be applied as a feature to find the position of each news agency.



This is an example of querying with "typhoon". One of the advantages of arranging opinions with time sequence is easily to see the corresponding event bursts. Obviously there are three typhoons in October 2004. The first is No. 0418 typhoon "Aere", the second is No. 0423 typhoon "Tokage", and the third is No. 0424 typhoon "NOCK-TEN". The second one last longer, while the third one caused greater damage. This figure shows the opinions toward typhoons are all negative, which is the same as we have expected.

Figure 2. Opinions towards typhoons

The components for extracting opinion words and sentences, and then decide their opinion polarities, are essential in this system. Therefore, the documents of NTCIR opinion analysis task are fed into this system and processed by these components, and then the extracted opinion information is reported as the experiment result.

3 Opinion Extraction

Opinions are extracted from sentences, which are the unit defined by NTCIR opinion task. Four factors are considered when extracting opinion passages and determining their tendency: the sentiment words, the opinion operators, the opinion holders and the negation operators.

We postulate that the opinion of the whole is a function of the opinions of the parts. That is, the opinion degree of a sentence, which decides if this sentence is opinionated and its polarity, is a function of sentiment words, negation words, opinion operators, and opinion holders. For negation and opinion operators, word lists are collected. For recognizing sentiment words, their opinion scores are calculated. The definition of the opinion scores of words is introduced in the next section.

3.1 Opinion Score of Words

Sentiment words are employed to compute the tendency of a sentence. Intuitively, a Chinese sentiment dictionary is indispensable. We adopt a Chinese opinion dictionary NTUSD [14]. NTUSD consists of 2,812 positive and 8,276 negative opinion words.

However, looking up in dictionaries may suffer from the problem of coverage. In our system, a method to learn sentiment words and their strengths based on this dictionary is developed. Scores here indicate the strengths.

It is postulated that the meaning of a Chinese sentiment word is a function of the composite Chinese characters. This is exactly how people read ideogram when they come to a new word. A sentiment score is then defined for a Chinese word by the following formula. This formula, not only tells us the possible opinion tendency of an unknown word, but also indicates their strength. We start the discussion from the definition of the formulas of Chinese characters.

$$P_{c_i} = \frac{fp_{c_i}}{fp_{c_i} + fn_{c_i}} \quad (2)$$

$$N_{c_i} = \frac{fn_{c_i}}{fp_{c_i} + fn_{c_i}} \quad (3)$$

Where fp_{ci} and fn_{ci} denote the frequencies of a character c_i in the positive and negative words, respectively; n and m denote total number of unique characters in positive and negative words, respectively.

Formulas (2) and (3) utilize the probability of a character in positive/negative words to show its sentiment tendency. However, there are more negative words than positive ones in NTUSD. Hence, the frequency of a character in a positive word may tend to be smaller than that in a negative word. That causes bias for learning, so formulas (2) and (3) are normalized into formulae (4) and (5).

$$P_{c_i} = \frac{fp_{c_i} / \sum_{j=1}^{n} fp_{c_j}}{fp_{c_i} / \sum_{j=1}^{n} fp_{c_j} + fn_{c_i} / \sum_{j=1}^{m} fn_{c_j}} \quad (4)$$

$$N_{c_i} = \frac{fn_{c_i} / \sum_{j=1}^{m} fn_{c_j}}{fp_{c_i} / \sum_{j=1}^{n} fp_{c_j} + fn_{c_i} / \sum_{j=1}^{m} fn_{c_j}} \quad (5)$$

Where P_{ci} and N_{ci} denote the weights of c_i as positive and negative characters, respectively. Formulae (4) and (5) calculate the possibility of one character to carry a positive and negative meaning, respectively. The difference of P_{ci} and N_{ci}, i.e., P_{ci} - N_{ci} in Formula (6), determines the sentiment tendency of character c_i. If it is a positive value, then this character occurs more often in positive Chinese words than negative ones, and vice versa. A value close to 0 means that it is not a sentiment character or it is a neutral sentiment character.

$$S_{c_i} = (P_{c_i} - N_{c_i}) \quad (6)$$

Formula (7) defines that a sentiment tendency of a Chinese word w is the average of the sentiment scores of the composing characters $c_1, c_2, \ldots, c_p$.

$$S_w = \frac{1}{p} \times \sum_{j=1}^{p} S_{c_j} \quad (7)$$

According to these formula, a character will be given at most the score 1 and at least the score –1, and so is a word. Take the word "好人" (good people) and "壞人" (bad people) as examples. There are 7,898 characters in positive opinion words and 24,212 characters in negative opinion words in total. The character "人" (people) appears in positive opinion words 79 times and negative opinion words 265 times. Therefore, the opinion score of "人" is –0.04, which is very neutral. The character "好" (good) appears in positive opinion words 68 times and in negative opinion words 52 times, and it is scored 0.60. Similarly, the character "壞" (bad) appears in positive opinion words 0 times and in negative opinion words 83 times, and it is scored -1. At last, we find the opinion score of "好人" (good people) 0.28, while the opinion score of "壞人" (bad people) –0.52. In spite of polarity information, the opinion score provides strength information. For example, the Chinese word "富貴" means wealth. Its sentiment score 0.61 is computed from the sum of "富" (rich, 0.75) and "貴" (expensive, 0.48). To determine the context polarity, "富貴" (wealth, 0.61) is stronger than"有錢" (have money, 0.33), which is another Chinese word describing rich in a subtler degree. The strength information of sentiment words help when finding the dominate sentiment words in one sentence.

The magnitude of the opinion score of an unknown word is also the indication of whether it should be counted. In our system, if a word does not appear in the dictionary, that is, it is unknown, only the word whose opinion score is above 0.4 or below –0.4 is taken into consideration, i.e. treated as a sentiment word.

3.2 Possible Sentiment Words

Opinion scores are not calculated for all words. Since the sentences are segmented, the part of speech information is used to extract possible sentiment words. From observations, Chinese words are composed mostly by more than one character, and one character itself usually cannot express a complete concept. Here words with part of speech A (adjective), V (verb), Na (proper noun), D (adverb) and Cbb (conjunction) and of length more than one are selected for further calculations of opinion scores.

3.3 Negation Operator

Negation operators are words such as "不" (no), "沒有" (not), "從不" (never), "也不" (neither), "不可能" (impossible), etc.. These words reverse the meanings of sentences. Moreover, if they modify sentiment words, the opinion polarities of these sentiment words will be reversed, too.

In CopeOpi, 41 negation operators are collected. For each sentence, after assuring sentiment words by the formula in section 3.1, each negation operator will negate the opinion polarity of the closest sentiment word, that is, change the opinion score of that word from S to –S. The effect of a negation operator will not cross commas, periods, question marks, semicolons, and exclamation marks. This sentence segments separated by these punctuation marks are referred to as "sentence fragments". Negation operators themselves can also express negative attitudes. Therefore, if there are no sentiment words in one sentence fragment, the scores of the negation operators within are counted.

3.4 Opinion Operator and Opinion Holder

Opinion operators are hints for extracting opinions. Words like "說" (say), "認為" (think), "相信" (believe) are actions of expressing thoughts. However, not all sentences containing opinion operators are opinionated. For example, "The central weather bureau says the highest temperature today is 32 Celsius degree" is considered a weather report, while the sentence "John thinks today is hot to death" is without question an opinion. In the experience of developing our system, we found that using opinion operators as the only cues for opinion extraction achieves the f-measure around 0.55 under the lenient metric and 0.35 under the strict metric.

Generally, opinion operators do not tell the overall opinion polarities. The polarities depend on the content of opinions. For example, in sentences "Mary told me that her teacher is a good person" and "Mary told me that her teacher is not good at teaching", the opinion polarities have nothing to do with the opinion operator "told". However, some opinion operators do express the attitudes of the holders and should be considered together with the content of opinions. For example, the opinion operators "hope" and "support" show the positive attitudes towards the following opinions, while "criticize" and "blame" show the negative attitudes. In the current system, the opinion scores of the opinion operators are counted when deciding the opinion polarities.

Another important function of opinion operators is to indicate the opinion holders. Since the opinion operators are the actions of expressing opinions, the subjects prior to opinion operators are likely to be the holders of the corresponding opinions. A word prior to an opinion operator is considered an opinion holder of an opinion sentence by our system if either one of the following two criteria is met:

1. The part of speech is person name (Nb_PERSON), organization name (Nb_ORGANIZATION) or personal (Nh). For example, "金大中" (Kim Dae-Jung) and "我們" (we) could be possible opinion holders.
2. The word is in class A (human), type Ae (job) of Cilin. (tong2yi4ci2ci2lin2, Mei et al., 1982). For example, "教授" (professor) and "學生" (student) could be possible opinion holders.

3.5 Algorithm

Because opinion polarities and opinion holders are information in opinion sentences, our system extracts opinion sentences first. Once the opinion sentences are found, their corresponding polarities and holders are reported. The algorithm of the opinion extraction is shown in Figure 3.

Algorithm: *Opinion Sentence Extraction*

1. **For** every sentence p
2. **For** every word in p, decide whether it's a sentiment word.
3. **For** every negation operator in p
4. **Find** the nearest sentiment word, and reverse its opinion score from S to -S.
5. **Extract** the candidate of the opinion holder if there is any opinion operator.
6. **Decide** the opinion polarity of p by the function of sentiment words and the opinion holder as follows.

$$S_p = S_{opinion-holder} \times \sum_{j=1}^{n} S_{w_j}$$

Where S_p, $S_{opinion\text{-}holder}$, and S_{wj} are the opinion score of sentence p, the weight of *opinion holder*, and the opinion score of sentiment word w_j, respectively, and n is the total number of sentiment words in p.

7. **If** the absolute value of S_p exceeds the specific threshold, report this sentence as an opinion. Report its polarity according to the sign of S_p, and its opinion holder.

Figure 3. Algorithm of Opinion Sentence Extraction

4 Experiments and Discussion

The experiment results are shown in Table 1, 2, 3 and 4. For opinion sentence extraction under the lenient metric, there are two groups of performance. One group is of f-measure around 0.6, and the other is around f-measure 0.7. Our system (NTU) is in the group of f-measure 0.7. The f-measures of all runs in this task are close to each other. In this group, our system has the best precision 0.664. If we consider performance of the opinion extraction together with the polarity judgment (field OpAndPolarity), our system achieves the f-measure 0.383, which is the second best. For the performance under the strict metric, we are still the second best. However, we also find that the system with higher precision will achieve better performance under the strict metric. Therefore, the difference between our system and the system of CHUK becomes larger under the strict metric. We believe that the most important work is to improve the precision in the future.

For sentence extraction, if there is any sentiment word in one sentence, it will be extracted. Therefore, missing one sentiment word will not influence the performance much. However, for the polarity judgment task, every sentiment word is important. As mentioned, to avoid noise, single character word will not be considered a sentiment word in our system. However, there are several opinion words which consists only one character, and misses occur. In addition, the negation operator will negate the nearest sentiment word. Therefore, the influence of missing a sentiment word will propagate if the missing word happens to be the target for negation.

The algorithm of dealing with negation operators is also very important in the polarity judgment. Our system negates the closest sentiment words. However, we found that the prior and the later sentiment words are both possible targets for negation, and the distance may not be the most important factor in choosing the correct one. Also if the target is wrong, the result is usually wrong. Besides, sometimes the negation operator in fact negates a non-sentiment noun, instead of a sentiment word. Since the noun is non-sentiment, our algorithm will ignore it and find the closest sentiment word to negate. To solve this problem, we may need a shallow parser to find the exact targets of negation operators.

Similar to the negation operators, some verbs have the abilities to negate sentiment words. However, this kind of words is not considered as a negation operator by our system now. For example, in the sentence fragment "結束戰爭的恐怖" (end the terror of wars), the verb "end" reverses the sentiment of terror, therefore "end the terror" is actually something good. However, in our system, "結束" (end), "戰爭" (war) and "恐怖" (terror) are all negative. This makes this sentence fragment very negative and that is wrong.

For the opinion holder extraction task, we achieve a relatively high precision with a low recall. The loss of the precision is mostly due to wrong segmentations. However, the loss of the recall may cause by the limitation that the opinion holders must appear prior to opinion operators. Since the opinion operators are collected manually and suffer from the coverage problem, many opinion holders are not extracted. Also if the prior word is not of part of speech Nb_PERSON, Nb_ORGANIZATION, Nh, or a job name in Cilin, nothing will be reported, even though the opinion operators are found. To achieve a better recall, a double-check mechanism may be needed to re-exam the location of the opinion holders when either opinion operators or potential holders are detected.

Group	Opinionated			Relevance			OpAndPolarity		
	P	R	F	P	R	F	P	R	F
CHUK	0.818	0.519	0.635	0.797	0.828	0.812	0.522	0.331	0.405
ISCAS	0.590	0.664	0.625	---	---	---	0.232	0.261	0.246
Gate-1	0.643	0.933	0.762	---	---	---	---	---	---
Gate-2	0.746	0.591	0.659	---	---	---	---	---	---
UMCP-1	0.645	0.974	0.776	0.683	0.516	0.588	0.292	0.441	0.351
UMCP-2	0.630	0.984	0.768	0.644	0.936	0.763	0.286	0.446	0.348
NTU	0.664	0.890	0.761	0.636	1.000	0.778	0.335	0.448	0.383

Table 1. Chinese opinion analysis lenient results

Group	Opinionated			Relevance			OpAndPolarity		
	P	R	F	P	R	F	P	R	F
CHUK	0.341	0.575	0.428	0.468	0.900	0.616	0.197	0.596	0.296
ISCAS	0.221	0.662	0.331	---	---	---	0.059	0.314	0.099
Gate-1	0.253	0.979	0.402	---	---	---	---	---	---
Gate-2	0.330	0.696	0.448	---	---	---	---	---	---

UMCP-1	0.245	0.986	0.393	0.404	0.565	0.471	0.085	0.615	0.150
UMCP-2	0.239	0.993	0.768	0.354	0.953	0.516	0.081	0.604	0.143
NTU	0.258	0.921	0.404	0.343	1.000	0.511	0.104	0.662	0.180

Table 2. Chinese opinion analysis strict results

Group	Lenient			Strict		
	P	R	F	P	R	F
CHUK	0.647	0.754	0.697	0.707	0.785	0.744
ISCAS	0.458	0.405	0.430	0.470	0.406	0.436
Gate-1	0.427	0.154	0.227	0.419	0.156	0.227
Gate-2	0.373	0.046	0.082	0.368	0.052	0.091
UMCP-1	0.241	0.410	0.303	0.293	0.438	0.351
UMCP-2	0.221	0.376	0.278	0.274	0.410	0.329
NTU	0.652	0.172	0.272	0.661	0.177	0.279

Table 3. Chinese opinion holders analysis: sentence based

Group	Lenient			Strict		
	P	R	F	P	R	F
CHUK	0.742	0.932	0.826	0.794	0.806	0.800
ISCAS	0.516	0.445	0.478	0.527	0.456	0.489
Gate-1	0.525	0.171	0.258	0.517	0.175	0.262
Gate-2	0.398	0.042	0.076	0.397	0.048	0.086
UMCP-1	0.297	0.429	0.351	0.357	0.453	0.400
UMCP-2	0.272	0.393	0.321	0.333	0.423	0.373
NTU	0.745	0.169	0.276	0.760	0.175	0.284

Table 4. Chinese opinion holders analysis: holder based

5 Conclusion and Future Work

This paper introduces a Chinese opinion extraction system. In this system, opinion scores are used to show the opinion polarities and strengths of words. This system adopts bottom up formulae which calculate the opinion scores of potential sentiment words in sentences from characters. Together with the negation operators and opinion operators, the polarities and opinion holders can be decided for all sentences. The experimental results are satisfactory.

The proposed formulae work well in general cases. However, they are not good enough in some cases. First, the polarities of some opinion words are context dependent. For example, "增加" (increase) is positive when its object is "薪水" (salary), while it is negative when its object is "稅賦" (tax). Moreover, if we consider the multi-perspective issue, "增加稅賦" (increase tax) may benefit the budget deficit, so it is positive for the government. Therefore, the opinion polarities of these words depend on the roles they play in the sentences or documents. Moreover, there are perspective issues to be studied in the future. Second, this method depends a lot on the part of speech of words. Some words of part of speech noun (Na) are the noun form of adjectives, but some are not. Analyzing the components in those general nouns, which are not the noun form of adjectives or adverbs, is meaningless. However, we cannot distinguish one kind from the other in our system, and this results in false alarms.

The negation issue is also important in the polarity judgment. To find the exact target word for negation, a shallow parser is necessary. In addition, some other words have the same effects as negation operators. For example, words expressing "not to do" like "防止" (prevent) or "勸阻" (discourage). Even some opinion operators can negate the opinions, too. In the future, words of the concept "not to do" should be able to be extracted automatically. The combinations of the attitudes of the opinion operators and their corresponding opinions should be considered together.

To find opinion holders are important in applications of extracting opinions. With opinion holders, the public not only can find opinions of a specific person, but also people having the same attitudes toward a public issue can be grouped. The possible entities of persons or organizations may not be necessarily opinion holders in opinion sentences. Sometimes they are the targets criticized. Therefore,

we extract opinion holders with the hints from opinion operators. However, it seems too strict. Not only the coverage of opinion holders is limited, but also the opinion holders do not always appear together with the opinion operators. We found that sometimes the opinion holders appear with the possessives. For example, sentence fragments like "the opinions/attitudes of A are …" or "B's thoughts on … are …" contain opinion holders obviously but do not accompany opinion operators. To solve this problem, we need to learn more patterns in the future.

The algorithm for relevant sentence retrieval is not integrated yet in this system, because the CopeOpi originally cooperate with an IR system. Since all documents are relevant to the selected opinion topics, sentences in these testing documents are all treated as relevant to achieve the baseline performance and we focus on the opinion related tasks this year. In the future, for selecting topical words and further retrieving relevant sentences, the existing algorithm, which works well on TREC materials [5], can be applied to improve the performance.

We have developed a Chinese opinion extraction system. And the large-scale experiments are done on the materials from the NTCIR opinion pilot task. From the evaluation results, we find the directions to improve our techniques on opinion sentence extraction, opinion polarity judgment, opinion holder extraction and relevant sentence retrieval. The future goal is to enhance our system with improved techniques and apply this system in real applications.

References

[1] C. Cardie, J. Wiebe, T. Wilson and D. Litman. Combining low-level and summary representations of opinions for multi-perspective question answering. *Proceedings of AAAI Spring Symposium Workshop*, pages 20-27. 2004.

[2] K. Dave, S. Lawrence and D.M. Pennock. Mining the peanut gallery: opinion extraction and semantic classification of product reviews. *Proceedings of 12th International Conference on World Wide Web*, pages 519-528. 2003.

[3] Minqing Hu and Bing Liu. Mining and Summarizing Customer Reviews. *SIGKDD 2004*, pages 168-177. 2004.

[4] Soo-Min Kim and Eduard Hovy. Determining the Sentiment of Opinions. *Coling*, pages 1367-1373. 2004.

[5] L.-W. Ku, L.-Y. Li, T.-H. Wu and H.-H. Chen. Major topic detection and its application to opinion summarization. *SIGIR 2005*, pages. 627-628. 2005.

[6] B. Pang, L. Lee and S. Vaithyanathan. Thumbs up? Sentiment classification using machine learning techniques. *Proceedings of the 2002 Conference on EMNLP*, pages 79-86. 2002.

[7] E. Riloff and J. Wiebe. Learning extraction patterns for subjective expressions. *Proceedings of the 2003 Conference on EMNLP*, pages 105-112. 2003.

[8] H. Takamura, T. Inui and M. Okumura. Extracting Semantic Orientations of Words Using Spin Model. *ACL 2005*, pages 133-140. 2005.

[9] J. Wiebe, T. Wilson and M. Bell. Identify collocations for recognizing opinions. *Proceedings of ACL/EACL2001 Workshop on Collocation*. 2001.

[10] Soboroff, I. and Harman, D. Overview of the TREC 2003 novelty track. *Proceedings of the Twelfth Text REtrieval Conference,* National Institute of Standards and Technology, pages 38-53. 2003.

[11] Bai, X., Padman, R. and Airoldi, E. On learning parsimonious models for extracting consumer opinions. *Proceedings of the 38th Annual Hawaii International Conference on System Sciences*, Track 3, Volume 03, page 75.2. 2005.

[12] Riloff, E., Wiebe, J. and Wilson, T. Learning subjective nouns using extraction pattern bootstrapping. *Proceedings of the Seventh Conference on Natural Language Learning*, pages 25-32. 2003.

[13] Ku, L.-W., Wu, T.-H., Lee, L.-Y. and Chen., H.-H. (2005). Construction of an evaluation corpus for opinion extraction. *Proceedings of the 5th NTCIR Workshop Meeting*, pages 513-520.

[14] Ku, L.-W., Liang, Y.-T. and Chen, H.-H. Opinion extraction, summarization and tracking in news and blog Corpora. *Proceedings of AAAI-2006 Spring Symposium on Computational Approaches to Analyzing Weblogs, AAAI Technical Report*. Pages 100-107. 2006.