Opinion Analysis based on Lexical Clues and their Expansion

Youngho Kim Sung-Hyon Myaeng
Information and Communications University
119, Moonji-ro, Yuseong-gu, Daejeon, 305-714, South Korea
{yhkim, myaeng}@icu.ac.kr

Abstract

The challenge of an automatic opinion analysis has been the focus of attention in recent years in many domains such as online product review. Especially, in online news articles opinion analysis has good prospects, since newspaper is the most powerful media to disseminate people's opinions. We introduce a lexical information based approach to this task by exploiting lexical information, based on the quantitative analysis of opinions in the news articles. The method comprises semi-supervised subjectivity classification, gloss based sentiment classification, and rule based opinion holder finder. The method we present is remarkable since numbers of lexical clues we discovered were effective to this task. The experimental results show that our system achieves 45% of performance to extract opinionated sentences and 35% of performance to identify opinion holders.

Keywords: *Opinion Extraction, Opinion Holder, Relevance, Polarity, Opinion-based Application, NTCIR, Sentiment Classification, Opinion Analysis, Opinion Mining.*

1 Introduction

Opinion mining is a recent sub-discipline of information retrieval which is concerned not with the topic a document is about, but with the opinion it expresses [5]. Opinionated content management has several killer applications, such as collecting critics' opinions about a product through the classification of online product reviews or tracking the public attitudes toward a political candidate through mining online forums. Since newspaper is a mixture of subjective feature (i.e., opinions) and objective feature (i.e., facts), analyzing news article is the first step toward opinion mining.

Early attempts for opinion mining [2] included opinion extraction (i.e., determining the subjectivity of sentences or expressions in a document) and sentiment classification (i.e., determining the polarity of the subjective expressions and the strength of its polarity). Despite the successful attempts in sentiment classification in the past [15, 13, 22], further detailed analysis such as finding opinion holder (i.e., the one who maintains opinions) and topic related opinions still remain as an active research area. As such, this paper addresses the problem of identifying not only opinions but also holders and topic-relevance of opinions from news articles.

Opinion mining in news articles is more challengeable than in other areas. Unlike a product review whose content is mostly opinions about the product, opinions and facts are mingled in a news article. Thus, mining opinions in news may need more elaborated process. Furthermore, news articles include different opinions which come from very different opinion holders (e.g., people, organizations, or government offices). So, identifying the opinion holders in opinionated news text is more challengeable.

We participated in the Opinion Analysis Pilot Task at NTCIR-6 [16] is such a detailed analysis which includes:

1. *determining the subjectivity at the sentence level*, as in deciding whether a given sentence has a factual nature or expresses an opinion on its subject matter. This is a binary classification task under classes Objective and Subjective;

2. *determining the polarity (or semantic orientation) of the subjective sentence*, as in deciding whether the extracted subjective sentence (i.e., opinion) expresses a positive, negative or neutral sentiment on its subject matter (i.e., topic);

3. *finding opinion holders*, as in searching opinion holders (e.g., people, organizations) who have a positive, neutral or negative attitude to the given topic.

Since we analyze at the sentence level, the task of recognizing the topic-related opinions is determining the relevance of a sentence to the topic. To identify opinion holders indicated by a demonstrative pronoun, anaphora resolution for opinions is subsumed by Task 3. Functionality to the above tasks is the identification of opinions and those holders present in news articles, such as extracting a negative opinion, "They accused certain media companies" which is expressed by "The Japanese Society for History Textbook Reform" (i.e., "They" is resolved as "The Japanese Society for History Textbook Reform" in this case).

In this paper, we present a lexical information based methodology for opinion analysis. It relies on the application of a semi-supervised learning method to the task of classifying sentences as Subjective or Objective (Task 1). The essence of our method is a balance between a rule-based approach and a machine learning-based method. Despite the fact that heuristic rules are concrete, rule-based systems are weak in generalizability and its coverage. On the other hand, the approach based on statistical learning such as Support Vector Machine (SVM) are powerful [15] and applicable to a large domain, but feature selection based on the context of the task and the high cost of designing training data are quite difficult problems. Thus, we propose a semi-supervised learning which overcomes some of the drawbacks. In order to test our hypothesis that an opinionated sentence has its own lexical clues not present in a f actual sentence, we adopted SVM by using presence of clue words and their part-of-speech (POS) information as features. We first extracted relatively complete seed rules to determine a sentence's subjectivity, and then our SVM trained by the seed sentences extracted by the seed rules performed the classification of the subjectivity at the sentence level.

To determine the sentiment of the subjective sentences (Task 2), we postulate that sentences with the same polarities may contain similar clue words. Therefore, we first extract frequent words in the subjective sentences already labeled by Task 1, and then determine the sentiment of the extracted terms through those glosses based on sentiment seed set expanded by Esuli's method [4]. More precisely, we obtain terms' glosses from WordNet and label a term for Negative, if its gloss contains negative seed terms. If the gloss does not contain any seed terms, we identify it as neutral. Furthermore, we find opinion holders (Task 3) with a simple anaphor resolution method. We hypothesize that only people (e.g., President Noh, the author) and organizations (e.g., Microsoft, Ministry of Information and Communications) are able to express their opinions. Therefore, we adopt Named Entity Recognizer (NER) to identify certain entities. We test our hypotheses on the online news articles gathered from NTCIR-6.

This paper is organized as follows: Section 2 introduces previous work for opinion analysis. Section 3 describes our method that extracts opinionated sentences, determines the sentiment of opinions, and searches opinion holders. Section 4 reports the experimental results on online news articles with discussions. Finally, Section 5 contains a conclusion.

2 Related Works

For opinion analysis, different names have been used such as opinion mining [2], sentiment classification [15], sentiment analysis [13], opinion extraction [12], affective classification [1] and affective rating [3, 14]. It has emerged in the last few years as a research area, largely driven by interests in developing applications such as mining opinions in online corpora, or customer relationship management (e.g., customer's review analysis).

Previously, numerous research activities have focused on the determination of the semantic orientation at the word level, since seed words play an important role as sentiment clues in indicating the sentiment of a sentence or a document as sentiment clues. Hatzivassiloglou and McKeown [6] have attempted to predict semantic orientation of adjectives by analyzing pairs of adjectives (i.e., adjective pair is adjectives conjoined by and, or, but, either-or, neither-nor) extracted from a large unlabelled document set. Turney [18] has obtained remarkable results on the sentiment classification of terms by considering the algebraic sum of the orientations of terms as representative of the orientation of the document they belong to. Furthermore, Turney and Littman [19] have bootstrapped from a seed set[1] , and determined semantic orientation according to PMI[2] method . Kamps et al [10] have focused on the use of lexical relations defined in WordNet. They defined a graph on the adjectives contained in the intersection between the Turney's seed set and WordNet, adding a link between two adjectives whenever WordNet indicate the presence of a synonymy relation between them. Esuli and Sebastiani [4] proposed semi-supervised learning method started from expanding an initial seed set based on Turney and Littman's seed set [19], by using WordNet. They determined the expanded seed term's semantic orientation thru gloss classification by statistical technique. Wilson et al [21] have attempted to classify the strength of opinions and the subjectivity of deeply nested clasuses. Wiebe et al [20] described a sentence-level Naive Bayes classifier using as features the presence or absence of particular syntactic classes[3] , punctuation, and sentence position. Subsequently, Hatzivassiloglou and Wiebe [7] showed that automatically detected gradable adjectives are a useful feature for opinion classification. Pang et al [15] adopted a statistical technique-based approach, using supervised machine learning with words and n-grams as features to predict orientation at the document level.

Ku et al [12] suggested an opinion extraction, sum-

[1] Seed set contains 7 positive words and 7 negative words as good, nice, excellent, positive, fortunate, correct, superior in positive set and bad, nasty, poor, negative, unfortunate, wrong, inferior in negative set.

[2] Pointwise Mutual Information

[3] pronouns, adjectives, cardinal number, modal verbs, adverbs

marization and tracking system which can summarize topic-relevant opinions. Also, Kim et al [11] proposed a Semantic Role Labeling-based method to identify an opinion, firstly and then find its holder and topic. They utilized FrameNet to label semantic roles related to the opinions.

Bradley and Lang [1] tried psychological studies which have found measurable associations between words and human emotions. They arranged the ANEW list which is a set of normative emotional ratings for a large number of English words. Those words were evaluated in three dimensions: pleasure, arousal, and dominance. With utilizing this list, Owsley et al [14] had a domain specific affective classification approach used ANEW corpus's valence to evaluate affective adjective words that are identified by the part-of-speech tagging.

3 Methods

3.1 System architecture

The system has three components corresponding to the three steps applied to a given document. The first step is to determine the subjectivity of each sentence. In the next step, our system classifies subjective sentences into one of the three categories: Positive, Negative or Neutral. The final step is to identify the opinion holders for the subjective sentences. Figure 1 shows the flow of the steps with sub-components. We adopted MALLET NER Tagger [4] and Stanford Part-Of-Speech Tagger [17]. To extract topic-related sentences, we implemented the relevance checker which determines the topic-relevance as true if a sentence contains any of the relevant words such as the relevant concepts, topic-description, and headline key-words, provided as background information (see Section 4.1). In Task 1, we adopt SVMlight developed by Joachim [9].

After finishing all the steps, each sentence is associated with a quadruple

$$Relevance, Subjectivity, Sentiment, Holders$$

where Relevance and Subjectivity expresses the degree to which the sentence is relevant to the topic and contains an opinion,respectively. The Sentiment element is one of the three values: Positive, Negative, or Neural, whereas the Holders element contains a string that represents the holders of the opinion.

3.2 Determining the subjectivity and its polarity

As addressed in Section 1, our method for determining the subjectivity of a sentence uses semi-supervised learning. Thus, we analyzed first a set

[4]MALLET Project (http://mallet.cs.umass.edu/)

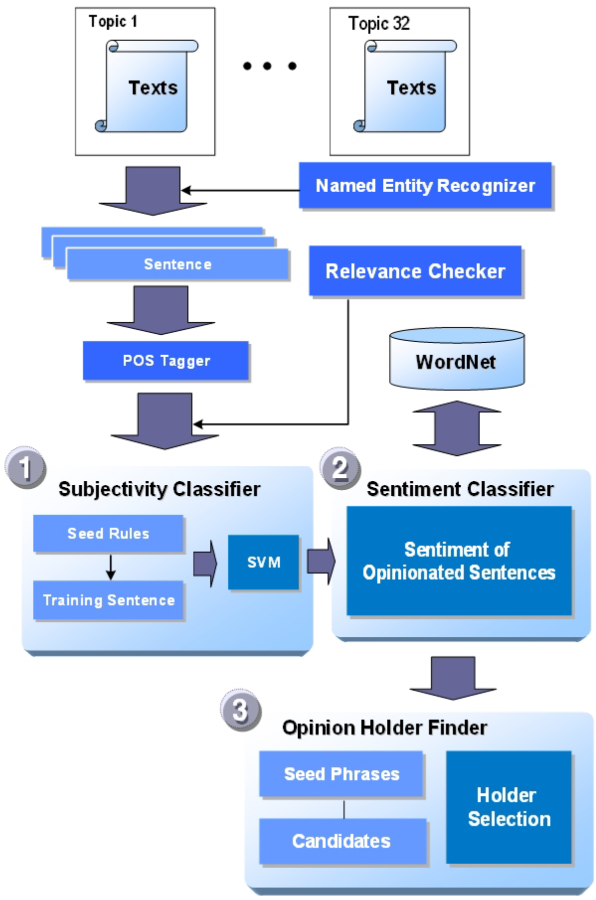

Figure 1. System framework with information flows

of training documents to extract several seed rules by which a sentence's subjectivity can be determined (See Section 4.1). The rules are based on some lexical clues such as "insist" that signal an opinionated sentence. We found some verbs and auxiliary verbs such as "would" are strong indicators for a non-factual statement. The six seed rules predict opinion sentences with a high precision (85% in a sample of the collection). In essence, the rules are used to classify sentences. If any of the rules can be applied, the sentence is considered having an opinion. Otherwise, we can not guarantee the sentence opinionated.

As a result of the analysis of sample documents (See Section 4.1), we developed the six rules from 86 lexical clues including 34 verbs, 28 adjectives, 13 nouns, and 11 other patterns. For example, the pattern "It is certain that" is an indicator for the author's opinion. Among those candidates, we chose top 14 high precision clues (i.e., *insist, claim, criticize, think, believe, would, could, should, might, will, may, in fact, unfortunately, consequently*) based on the sample data and designed seed rules as shown in Table 1.

The rest of the initial candidates were used to detect

Table 1. Seed rules for subjectivity classification

No	Description
1	If a sentence has “insist” as the main verb
2	If a sentence has “claim” as the main verb
3	If a sentence has “criticize” as the main verb
4	If a sentence has “think” as the main verb
5	If a sentence has “believe” as the main verb
6	If a sentence contains “would”, “could”, “should”, “might”, “will”, or “may” as an auxiliary verb and concurrently has a phrase such as “in fact”, “unfortunately”, or “consequently”

the lack of subjectivity (i.e., a sentence has no opinion). If a sentence satisfies any one of the conditions of the seed rules, the sentence is labeled as True for Subjectivity (i.e., positive example for the subjectivity classifier). A sentence that does not contain any of the 86 lexical clues, however, is tagged as False for Subjectivity (i.e., negative example for the subjectivity classifier).

Using the seed rules, we obtained a training set for our SVM-based classifier, where sentences are tagged with True or False for subjectivity. All the verbs, nouns, adjectives, and adverbs in the positive sentences tagged with True were extracted as features in the training documents. Also included to our feature set were the 13 strong clues that often appear in positive sentences.

Having selected opinionated sentences, the second step is to classify them into three classes: Positive, Negative, and Neutral. This process of judging the sentiment of opinions is quite challenging for news article, due to the unique aspects of new articles. First of all, some previous research results on sentiment analysis are not transferable to the news articles. For example, Turney [19], Kamps [10], and Hatzivassiloglo [7]developed a set of seed terms to determine sentiment of sentences but those seeds rarely occur in news articles, making them hardly useful. Second, the sentiment judgment of news articles needs abundant prior knowledge. For instance, a sentence reporting on a merge of two companies in Japanese newspaper article should be judged to have negative sentiment whereas the same kind of activities in the US would be a positive event. Possessing such knowledge would be difficult for automated text processing systems. This is the reason why we wanted to focus on sentiment analysis based on lexical clues: simplicity and applicability.

Our approach to the sentiment analysis phase is based on the hypothesis that polarity can be determined not only by seed terms but also those semantically related to them. This hypothesis, if proven to be true, would alleviate the first problem mentioned above, i.e. difficulty to find the seed terms in newspaper articles. In order to obtain the terms that are semantically related to the seed terms, we employed Esuli's seed term expansion method.

First, we sorted the frequent terms in the opinionated sentences obtained from the first phase. Next, we classify the frequent terms through their glosses as Esuli's seed term classification [4]. That is,we gathered the glosses of such sorted terms from WordNet because sentiment seed terms do not often appear on newspaper text. In order to classify the glosses as positive, negative, or neutral, we started with Turney's [19] minimal seed set consisting of 7 positive terms and 7 negative terms[5], and expanded it with Esuli's seed term expansion method (Figure 2).

The function, *ExpandSimple* is invoked to produce an expanded seed set utilizing lexical relations defined in WordNet, a union of synonymy and indirect antonymy, according to Esuli's suggestion. The expansion procedure runs iteratively by using SVM. That is, at the first step, the new seed set is obtained from the old seed by the function. Next, for the glosses (from WordNet) of the terms in the new seed, SVM which is trained from the glosses of the old seed terms with the weighting scheme of cosine-normalized *tfidf* performs the classification. Empirically, over 4 repetitions guarantee relatively high performance. As a result, we collected 522 negative terms and 415 positive terms (Sentiment Seed Set).

Based on the 937 term set, we classify the glosses of the frequent terms. If the gloss contains any of sentiment seed terms, the frequent term which has the gloss is polarized as the detected sentiment seed term. For the collision of the sentiment seed terms (e.g., Term A's gloss has both of positive terms and negative terms), the judgment is based on the first seed term's sentiment. Moreover, in case of the gloss containing none of seed terms, we determined the term as neutral.

After the frequent terms' classification, for each opinionated sentence,we judge its sentiment in accordance with the sentiment of the most frequent term which belongs to the target sentence. For example, the sentence contains term A and term B. Term A (negative) is more frequent than Term B (positive) in overall opinionated sentences. And then, the sentiment of the sentence is negative with the hypothesis that negative sentence has higher chance to have the frequent negative term A.

[5] Positive set is { good, nice, excellent, positive, fortunate, correct, superior } and Negative set is { bad, nasty, poor, negative, unfortunate, wrong, inferior}

function $ExpandSimple$

Input :

(S_p, S_n) : seed set for the Positive and Negative categories

G_{rel} : graph defined on terms by the lexical relation rel

S_{rel} : boolean flag specifying if the relation expresses similarity or opposition of orientation

Output :

(S'_p, S'_n) : expanded seed set

Body :

1. $S'_p \leftarrow S_p$; $S'_n \leftarrow S_n$;
2. foreach $term$ in S_p do
   - $Temp \leftarrow$ set of all terms directly connected to $term$ in G_{rel};
   - if S_{rel} then
     - $S'_p \leftarrow S'_p \cup Temp$;
   - else
     - $S'_n \leftarrow S'_n \cup Temp$;

   foreach $term$ in S_n do
   - $Temp \leftarrow$ set of all terms directly connected to $term$ in G_{rel};
   - if S_{rel} then
     - $S'_n \leftarrow S'_n \cup Temp$;
   - else
     - $S'_p \leftarrow S'_p \cup Temp$;
3. $S'_p \leftarrow S'_p - S_n$; $S'_n \leftarrow S'_n - S_p$;
4. $Dup \leftarrow S'_p \cap S'_n$; $S'_p \leftarrow S'_p - Dup$; $S'_n \leftarrow S'_n - Dup$;

Figure 2. Esuli's expansion function

3.3 Finding opinion holders

After finishing the sentiment classification, we run the Opinion Holder Finder which looks for the holders of opinionated sentences. We mostly focus on a person and an organization identified by the NER Tagger as an opinion holder. We extracted lexical clues such as “According to President Roh” or “President Roh saying” which are simple but critical. Hypothetically, we assumed that there are not significantly many various patterns to express an opinion holder. Therefore, we developed a set of rules, according to the extracted lexical clues, to identify holders as in Table 2. More precisely, opinion holder can be extracted if sentence satisfies any of the rules. Otherwise, we need another process to extract the holder since all opinionated sentences must have their holders in this phase. So, we consider the nearest person or organization from the verb as the holder of the opinion which can not be observable by the rules.

Rule 2 looks for the holder preceding “say” or following “say” (i.e., inversion). For instance, “said Choi Sang Yong” is an inversion case. Rule 3 is case-sensitive since the sentence starting with “By” and the proper entity following the “By” are strong clues. Rule 4 and 5 are used to capture the holder as ”the author”. Practically, Rule 6 is useful to identify the holder.

No matter whether the extracted holder is correct, the anaphora resolution for opinion holder is still problem. Since pronoun can denote opinion holder as an anaphoric clue, we detected the anaphoricity of the holder as pronoun is considered in the rules. More specifically, the identified holder which contains pronoun is anaphoric. For example, ”they” indicates a holder from somewhere in previous sentences as an anaphoric clue. Thus, the pronoun is linked to the nearest antecedent in the list of only people or organization entities from the previous sentences.

Table 2. Rules for the search of holders

No	Description
1	If “according to” occurs in a sentence, the holder is the next people, organizations, or pronoun.
2	If “say” exists in a sentence, holder is the nearest people, organizations, or pronoun.
3	If If a sentence starts with “By” and people, organizations, or pronouns follow it, holder is the following entities.
4	If a sentence starts with “I” and concurrently has one of “think, criticize, claim, believe agree, insist, express, announce, talk, tell, note, deliver” or any auxiliary verbs, the holder is the author
5	If a sentence contains one of clue phrases such as “I am of the opinion that, I know that, It is certain that, Seems to me”, the holder is the author.
6	If a sentence includes one of “think, criticize, claim, believe agree, insist, express, announce, talk, tell, note, deliver” and people, organizations, or pronouns are in front of such verb, the holder is one of them.

4 Experimental Results

4.1 Data

The test collection from NTCIR-6 [16] consists of 28 topics (439 documents, 8,379 sentences) in English. The sample data containing 19 documents (786 sentences) relevant to the topic “Economic influence of the European monetary union” are provided prior to actual running of our system. In addition, background-information which contains topic relevant concepts, titles, and descriptions is given to each topic. In order to design our methods, we analyzed the sample data in advance.

4.2 Results

We present the results from testing of our methods with the test collection. Prior to presenting the results, we introduce how the standard (i.e., gold-standard) was established to evaluate the experimental results. The gold-standard is based on the annotations of three assessors. *Strict standard* is the case where the annotations of all three assessors are the same, and *lenient standard* is the case for the agreements of two or more of all assessors are the same. Since the gold-standard was constructed by NTCIR-6, further detailed information is available in [16].

We could be confident through the results from the preceding experiment on the sample data (i.e., 19 documents) but we found several flaws through the sample results. Our method in Task 1 is competitive but weak for sentences which have no lexical clues (e.g., "The main difference is that Japanese firms usually do not act on their desires to merge."). The sample results show the difficulties of the sentiment task (i.e., lower performance). In Task 3, there are many various patterns not covered by the rules (See Section 3.3). One of them is passive sentence (e.g., "which has been criticized by some Asian countries"), since detecting the past perfect form of verb is slightly hard task. The number of holders is also the problem in the anaphora resolution (e.g., "they" can indicate not only "Japanese firms" but also "AOL"). That is, several people or organizations can express the same opinion.

Table 3. Evaluation on sample data

Task	Precision	Recall	F-Measure
Subjectivity	0.701	0.645	0.672
Relevance	0.737	0.448	0.557
Positive	0.500	0.421	0.457
Negative	0.480	0.648	0.551
Neutral	0.321	0.663	0.433
None	0.785	0.774	0.779
All	0.246	0.462	0.321
Holder	0.621	0.368	0.462

We obtained our system's performance as shown on Table 4, 5, and 6. Since our approach was driven by the sample data, the performance on the testing collection was slightly lower than that on the sample documents. In the subjectivity task (Task 1), the difference of precision results between the strict and lenient standard denotes the diverse subjectivity of the assessors (i.e., each assessor has very different perspective with respect to the subjectivity). Also, certain diversity led lower recall in lenient evaluation (i.e., Much more opinions were annotated in lenient standard than those in strict standard). In the sentiment task (Task 2), the sentiment of whole sentence is not revealed directly by lexical information (i.e., the context such as topic or previous sentence has more influence than the sentence itself). However, negativity was recognized more accurately (i.e., negative sentiment of the sentence is expressed by negative seed terms, influentially) than those of the other classes. Although the precision of opinion holder in the sample data is 0.621, that in the testing set is low because of the limitation of rule based extraction (i.e., rules are effective in the sample data, but not as much effective in the testing set).

Table 4. Strict evaluation on testing collection

Task	Precision	Recall	F-Measure
Subjectivity	0.102	0.616	0.175
Relevance	0.177	0.266	0.213
Positive	0.035	0.578	0.066
Negative	0.090	0.198	0.123
Neutral	0.016	0.489	0.031
None	0.980	0.702	0.818
All	0.034	0.301	0.061

Table 5. Lenient evaluation on testing collection

Task	Precision	Recall	F-Measure
Subjectivity	0.396	0.524	0.451
Relevance	0.409	0.263	0.320
Positive	0.154	0.385	0.221
Negative	0.303	0.176	0.223
Neutral	0.101	0.341	0.156
None	0.881	0.740	0.804
All	0.151	0.264	0.192

4.3 Discussion

Even though we recognized several drawbacks of our methodology through the analysis of the sample experiment, more difficulties were identified from the analysis of the testing results. First of all, less objectivity in the standard caused the significantly low precision in strict evaluation. That is, there are many contrary assessments in the gold-standard. Thus, the

Table 6. Opinion holder evaluation on testing collection

Standard	Precision	Recall	F-Measure
Strict	0.085	0.515	0.146
Lenient	0.303	0.404	0.346

performance of the system is so flexible depending on annotator's subjectivity (e.g., tendency such as some assessor regards any sentences as opinions if the sentences are said by only people). It is much more important to capture certain subjectivity to obtain better performance. Ideally, gathering all public's perspectives is the best, but it is not feasible. So, an analysis of annotators may bring us practical breakthrough.

Regardless of the assessment problem, determining the sentiment of opinions implied much more complexity. In other words, many problems such as prior knowledge about topic (See Section 3) belong to the task. A partial negativity is one of such problems. That is, partially negative expressions can not guarantee the whole sentence's negativity. For example, the topic "Japanese text book distortion" is intuitively negative, and many negative expressions such as "distorted" are identified. However, there exist positive sentences which even have negative expressions concurrently. To resolve certain problem, we need to know semantic relationships between topic words and sentiment seed words (i.e., the function to ignore the partial negativity is required).

We discovered some unexpected problems involved in Task 3. One of the problems is hidden opinion holders. That is, many of opinion holders were not revealed explicitly in their sentences without any anaphoricity clues (e.g., pronouns). To illustrate this, "The absence of a regulatory framework was supposed to enable firms to thrive in cyberspace." is obviously opinion. However, identifying the author as an opinion holder is not clear even for human without reading the previous sentences (i.e., context knowledge). Even though the anaphoricity clue is explicitly detected, it is difficult to extract accurate holders. For example, "One says" should be rephrased as "A member of Japanese Society for History Textbook Reform." Since there are many organizations and people between the correct antecedent (i.e., the answer for the anaphor) and the anaphor, determining an exact antecedent among candidates is an overnice task. Actually, this is one of the challengeable problems in Anaphora Resolution [8].

5 Conclusions and Future Works

This paper presents a lexical information based methodology for the opinion analysis from Opinion Analysis Pilot Task at NTCIR-6. To identify the subjectivity at the sentence level, we propose semi-supervised learning method based on highly precise seed rules. The main thrust of this method is the enhanced combination of rule-based algorithms and machine learning techniques. Next, we determine the sentiment of the opinionated sentences based on the lexical information of the sentiment seed terms derived from the sentiment term expansion function. Since opinion holder identification is still challengeable because of the anaphoric opinion and its abbreviation, we present competitive rules including the detection of the anaphoricity and its resolution. In order to verify our methodology, we tested it on online news articles gathered from the NTCIR-6 opinion corpus. The evaluation shows the difficulties of opinion analysis, but our method is promising since its potential enhancement. We have concentrated on extracting lexical clues of only opinions. Considering the nature of factual sentences, however, will improve the performance. Also, adding more clues such as exaggeration (e.g., "dream combination") will contribute the enhancement. To improve the performance of finding holders, we need more accurate detection of subject and verb in much complicated sentences such as compound sentences. Using a parser to identify certain components will guarantee much better performance. As discussed in Section 4.3, recognizing the sentiment of opinions without prior knowledge of the topic and the context information is so complicated. Topic analysis such as capturing the intuitive tendency of the given topic (e.g., several Asian countries demand Ministry of Education in Japan to re-examine "Japanese text book distortion") can promise more detailed determination of the sentiment.

References

[1] M. M. Bradley and P. J. Lang. Affective norms for english words (anew): Stimuli, instruction manual and affective ratings. Technical Report Technical Report C-1, The Center for research in Psychophysiology, University of Florida, Gainesville, Florida, US, 1999.

[2] K. Dave, S. Lawrence, and D. M. Pennock. Mining the peanut gallery: Opinion extraction and semantic classification of product reviews. In *Proceedings of the 12th International Conference on the World Wide Web*, Budapest, Hungary, 2003. ACM Press, New York, US.

[3] S. D. Durbin, J. N. Richter, and D. Warner. A system for affective rating of texts. In *Proceedings of the 3rd Workshop on Operational Text Classification*, Washington, US, 2003.

[4] A. Esuli and F.Sebastiani. Determining the semantic orientation of terms through gloss classification. In *Proceedings of the 14th ACM International Conference on Information and Knowledge Management*, Bremen, Germany, 2005. ACM Press, New York, US.

[5] A. Esuli and F.Sebastiani. Determining term subjectivity and term rientation for opinipon mining. In *Proceedings of the 11th Conference of the European Chapter of the Association for Computational Linguistics*, Trento, Italy, 2006.

[6] V. Hatzivassiloglou and K. R. McKeown. Predicting the semantic orientation of adjectives. In *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics*, Madrid, Spain, 1997. ACL.

[7] V. Hatzivassiloglou and J. M. Wiebe. Effects of adjective orientation and gradability on sentence subjectiv-

ity. In *Proceedings of the 18th International Conference on Computational Linguistics*, Saarbrucken, Germany, 2000.

[8] R. Iida, K. Inui, and Y. Matsumoto. Anaphora resolution by antecedent identification followed by anaphoricity determination. *ACM Transactions on Asian Language Information Processing*, 4(4), 2005.

[9] T. Joachims. *Making large-Scale SVM Learning Practical Advances in Kernel Methods - Support Vector Learning*. MIT Press, 1999.

[10] J. Kamps, M. Marx, R. J. Mokken, and M. D. Rijke. Using wordnet to measure semantic orientation of adjectives. In *Proceedings of the 4th International Conference on Language Resources and Evaluation*, Lisbon, Portugal, 2004.

[11] S. Kim and E. Hovy. Extracting opinions, opinion holders, and topics expressed in online news media text. In *Proceedings of the Workshop on Sentiment and Subjectivity in Text*, Sydney, Australia, 2006. ACL.

[12] L. Ku, Y. Liang, and H. Chen. Opinion extraction, summarization and tracking in news and blog corpora. In *Proceedings of AAAI-2006 Spring Symposium on Computational Approaches to Analyzing Weblogs*, California, US, 2006.

[13] T. Nasukawa and J. Yi. Sentiment analysis: Capturing favorability using natural language processing. In *Proceedings of the 2nd International Conference on Knowledge Capture*, New York, US, 2003. ACM Press.

[14] S. Owsley, S. C. Sood, and K. J. Hammond. Domain specific affective classification of document. In *Proceedings of AAAI-2006 Spring Symposium on Computational Approaches to Analyzing Weblogs*, California, US, 2006.

[15] B. Pang, L. Lee, and S. Vaithyanathan. Thumbs up? sentiment classification using machine learning techniques. In *Proceedings of the 7th Conference on Empirical Methods in Natural Language Processing*, Philadelphia, US, 2002. ACL, Morristown, US.

[16] Y. Seki, D. Evans, L. Ku, H. Chen, N. Kando, and C. Lin. Overview of opinion analysis pilot task at ntcir-6. In *Proceedings of the 6th NTCIR Evaluation Workshop*, Tokyo, Japan, 2007.

[17] K. Toutanova and C. D. Manning. Enriching the knowledge sources used in a maximum entropy part-of-speech tagger. In *Proceedings of the Joing SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora(EMNLP/VLC-2000)*, Hong Kong, 2000.

[18] P. D. Turney. Thumbs up or thumbs down? semantic orientation applied to unsupervised classification of reviews. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, Philadelphia, US, 2002. ACL.

[19] P. D. Turney and M. L. Littman. Measuring praise and criticism: Inference of semantic orientation from association. *ACM Transactions on Information Systems*, 4(21):315–346, 2003.

[20] J. M. Wiebe, R. B. T. Wilson, M. Bell, and M. Martin. Development and use of a gold standard data set for subjectivity classifications. In *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics*, Maryland, US, 1999.

[21] T. Wilson, J. Wiebe, and R. Hwa. Just how mad are you? finding strong and weak opinion clauses. In *Proceedings of the 21st Conference of the American Association for Artificial Intelligence*, San Jose, US, 2004.

[22] H. Yu and V. Hatzivassiloglou. Towards answering opinion question: Separating facts from opinions and identifying the polarity of opinion sentences. In *Proceedings of the 8th Conference on Empirical Methods in Natural Language Processing*, Sapporo, Japan, 2003.