

Context-Sensitive Opinion Mining using Polarity Patterns

Saeedeh Sadat Sadidpour

Department of Computer Science,
Faculty of ICT
Malek-Ashtar University of
Technology
Tehran, Iran

Hossein Shirazi

Department of Computer Science,
Faculty of ICT
Malek-Ashtar University of
Technology
Tehran, Iran

Nurfadhlina Mohd Sharef

Department of Computer Science,
Faculty of Computer Science and
Information Technology
Universiti Putra Malaysia
Malaysia

Behrouz Minaei-Bidgoli

Department of Computer Engineering
Iran University of Science and Technology
Tehran, Iran

Mohammad Ebrahim Sanjaghi

Department of Management and Soft Technology
Malek-Ashtar University of Technology
Tehran, Iran

Abstract—The growing of Web 2.0 has led to huge information is available. The analysis of this information can be very useful in various fields. In this regards, opinion mining and sentiment analysis are one of the most interesting task that many researchers have paid attention for two last decades. However, this task involves to some challenges that a very important challenge is the different polarity of words in various domain and context. Word polarity is an important feature in the determination of review polarity through sentiment analysis. Existing studies have proposed n-gram technique as a solution which allows the matching of the selected words to the lexicon. However, identification of word polarity using the standard n-gram method poses limitation as it ignores the word placement and its effect according to the contextual domain. Therefore, this study proposes a linguistic-based model to extract the word adjacency patterns to determine the review polarity. The results reflect the superiority of the proposed model compared to other benchmarking approaches.

Keywords—Opinion mining; Polarity patterns; Pattern matching; Context-sensitive; Politics domain

I. INTRODUCTION

In the past, people tried very to acquire data and knowledge. Whereas the appearance of web, and especially Web 2.0, brings on huge information is generated by users. Although, this is not desire because too much information leads to confused. Therefore, analysis, summarization and other related tasks are very useful and applicable results present for users and researchers. In this regard, opinion mining and sentiment analysis are one of the most important and helpful task that has been defined from two last decades.

Opinion mining is a field that its results can be used for different researches of various fields. In this regard, researchers have been attracted because of its application from the sociological and psychological analysis to the extraction of users' opinion in business field such as about products and services, or political discussions.

The first time, Wiebe was presented the most widely definition for subjectivity and opinion mining based on a linguist's idea in 1994. Also, she defined as the linguistically expression of opinions, sentiment, emotions, evaluations, beliefs, and speculations [1]. Hence, the goal of subjectivity analysis is determining the subjective or objective sentences [2] that Liu considered subjective and objective sentences against each other. In this respect, objective and subjective sentences define, respectively, as a fact information indicative about world and a person's emotions or beliefs indicative [3].

Next and in lower level, Zhang and Liu define opinion mining or sentiment analysis as "the computational study of opinions, appraisals, attitudes, and emotions about entities such as products, services, organizations, persons, events, and their different aspects" [4]. Also, they focus on recognizing the orientation and strength of polarity in different level using various methods and settings.

In this respect, various definitions in these years have caused to not be specific different between opinion mining and sentiment analysis and are used instead of each other. Of course, it is notable that some ones consider different levels for opinion and sentiment.

Regards, researchers investigated the various sides of opinion mining. For this purpose, some issues are defined such as how polarity is expressed, which level polarity is measured, or the polarity of words is fix or not.

Hence, primary researches on opinion mining determined the polarity as positive and negative [5]. However, Koppel and Schler proposed neural is also considered to be improved results [7 ,6]. In addition, opinion mining is defined as orientation (such as positive and negative) and strength (such as weak positive, middle positive, and strong positive) [8].

Meanwhile, researches have shown polarity is not fix and even changes in a domain [10 ,9]. Also, Ding expressed the polarity is fix only rather than a context [11].

scenario Give example to illustrate your idea.

What do you mean by 'replacement of words is not considered for its statistical calculation'? do you mean the vector space model/bag of words?

The change of polarity leads to context-sensitive opinion mining. For this purpose, some different solutions such as ontology are proposed for the context-sensitive problem. Although, the context-sensitive issue poses limitation to existing solution. For an example, the methods often consider the polarity of words as uni-gram. While words get together, they affect on the polarity of each other.

In contrast, n-gram is not a suitable method because the replacement of words is not considered for its statistical calculation; As a result, n-gram leads to the sparseness in repeating adjacent same words and lack of generalizability. For an instance, "supporting of terrorism" and "supporting of spies" are two different n-gram that their value of language model is very low. While the generality of them in pattern "supporting of (Neg. Exp.)" occurs more.

Thus, this paper proposes opinion mining is done using polarity pattern extraction based on language model helps to be possible words, which are synonym or have similar polarity, are replaced.

The polarity patterns categorize expressions which their polarity is same. For example, "supporting of terrorism" is negative; however, "supporting" is positive and "terrorism" is negative, and bag of word method often recognize the term as neutral. While the polarity pattern of "supporting of (Neg. Exp.)" leads to a negative result. Consequently, the results of context-sensitive opinion mining using polarity pattern matching express the significant improvement of accuracy rather than other methods.

In continue, section 2 presents a review on context-sensitive opinion mining. Then, context-sensitive opinion mining using polarity pattern matching and evaluation are discussed, respectively. Finally, conclusion is expressed.

II. A REVIEW ON CONTEXT-SENSITIVE OPINION MINING

Although many researchers have paid attention to opinion mining and sentiment analysis, the field poses several room of improvements to solve issues in data mining, web mining, and information retrieval. Therefore, different challenges are observed in the field such as how the polarity is calculated for word groups. So, the polarity is calculated using different methods [12] such as deep sentiment analysis [13], Markov logic [14], Fuzzy [16,15], LDA [17], and language model [18].

The primary opinion mining researches prepared opinion lexicons and calculated the polarity based on the lexicons. However, growing researches indicated that opinion mining faces to a big challenge. In this regard, the challenge is the polarity of words sometimes changes in different texts. Thus, some topics were considered such as domain-specific opinion mining [20,19,10,9], and context-sensitive opinion mining [21].

Contextual opinion mining expresses the polarity of words can be changed in a text, and words may have different polarities even in a domain [11]. In the meanwhile, co-

occurrence and conjunctions [22], top word collocations [23], Emoticons [24], opposite pairs [25], an approach based on holistic lexicon prepared using WordNet [11], and using features such as the characteristics of before and next words, modifiers, and dependency tree [26] are solutions for changing polarity in a text.

In addition, lexicon adapting is another strategy that finds context-sensitive polarity using contextual semantics and updates orientation and strength using a primary polarity lexicon and rule-base adapting [27].

However, the effectiveness of adjacent words is a significant criterion that has not been explored by any existing methods. Although language model can find adjacent words, it is typically rigid. Even co-occurrence and synonym words are not considered in statistical calculations. As a result, this paper proposes a method which is based on n-grams to categorize expressions that have similar polarity. Then, some polarity patterns are extracted that indicate for polarity language model. Finally, the polarity patterns are used to context-sensitive opinion mining and sentiment is calculated based on them.

III. CONTEXT-SENSITIVE OPINION MINING USING POLARITY PATTERN MATCHING

Different studies have used patterns for determining polarity. Pattern recognition is a helpful step for opinion mining tasks that detects text relations. In this regard, Riloff and Wiebe [28] extracted patterns to detect subjective and objective sentences. The patterns were extracted based on some templates and using corpus investigating.

Furthermore, the patterns for getting together words and language model are two another effective methods. Wiebe and others [2] reached to subjectivity language indicates adjacent words and language model accompanied with part of speech (PoS) patterns. Thereof, the aim of statistical language model is learning the probability function of word sequence. However, this problem needs many data that leads to sparseness. Therefore, words which operate same are considered as similar tokens and replaced to each other [30,29]. In the other hand, Tang and others [31] extracted possible bi-grams to calculate polarity. The determination of N in n-grams is another problem in opinion mining.

As a result, this paper proposes a method for opinion mining using polarity patterns which are the replacement of n-gram and solve its problems.

A. Polarity Pattern Extraction

Polarity pattern extraction proposed is based on n-gram method and creates a polarity language model for opinion mining. In according to the investigations of corpus, if X is the noun phrase of affect, three types of phrases, that can be shown as sub-sentences too, are effective on polarities:

- X + A: for example, "دشمنی تروریسم" means "hostility of terrorism" is negative when a negative affect is done by a negative effective.
- X + Conjunction + B: for example, "دشمنی با دین" means "hostility with religion" is negative when a negative affect is done rather than a positive affected.

- X + A + Conjunction + B: for example, “دشمنی تروریسم با دین” means “hostility of terrorism with religion” is negative when a negative affect is done by a negative effective rather than a positive affected.

X, A and B are sub-sections that can be positive or negative and led various phrases with different polarity. These three type of phrases are templates for finding polarity patterns. TABLE I is two examples of polarity of phrases.

TABLE I. THE EXAMPLES OF POLARITY PATTERNS

X		A	Conjunction	B	Polarity of phrase
Noun phrase of affect	translate				
دشمنی	Supporting		از of		
+		+		+	+
+		+		-	-
+		-		+	-
+		-		-	-
دشمنی	Hostility of		با with		
-		+		+	-
-		+		-	+
-		-		+	-
-		-		-	+

For this purpose, a rule-based method is used and polarity patterns are extracted based on pre-defined templates which were mentioned before. The steps of proposed method are shown in “Fig. 1”.

Firstly, natural language processing is done on corpus and tokens and their PoS are determined. Then, the corpus is parsed and its dependency trees are extracted. As follow, the dependency trees are adapted to templates; for each phrase group, if y_p and Y_q are the p th phrase group and q th desired phrase group, respectively, that will be obtained by (1).

$$Y_q = y_{p|DT_p=T_i} \quad (1)$$

that DT_p is the dependency tree of p th phrase group and T_i is different templates with $i = 1, 2, 3$. Thus, the desired phrase groups will be selected if the dependency tree of p th phrase group is matched with one of templates T_i . Also, the type of template, or i , which has been matched is determined.

In the next step, the polarity of sub-sections is determined. The sub-section X is a noun which is done an affect but A and B can be from a noun to a noun phrase which are an effective and an affected, respectively. For every Y_q , an adapted template AT_q is built based on matching among the desired phrase group, the polarity of sub-sections, and template T_{iq} in (2).

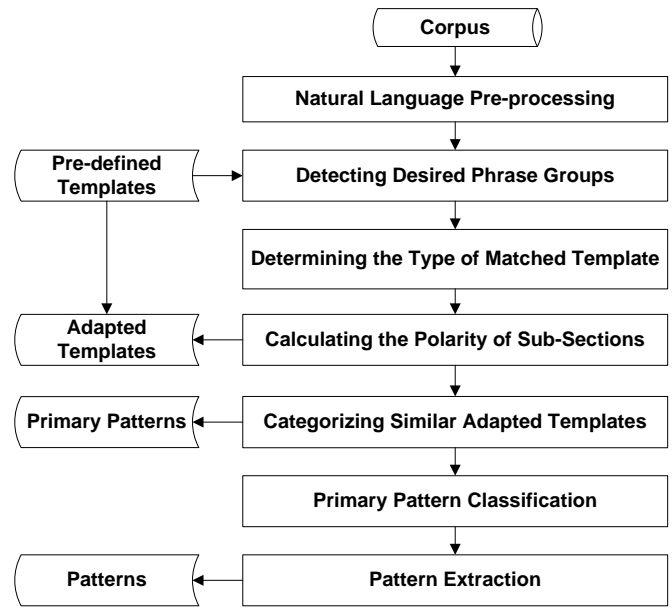


Fig. 1. The steps of proposed method for polarity pattern extraction

$$AT_q = Y_q \begin{cases} AT_{aq} = \text{The polarity of sub-section } A_q \\ \text{if } A_q \in T_{iq} \\ AT_{bq} = \text{The polarity of sub-section } B_q \\ \text{if } B_q \in T_{iq} \end{cases} \quad (2)$$

that q index indicates q th phrase group. AT_{aq} , AT_{bq} and T_{iq} are the polarity of sub-sections A and B, and i th template, which correspond with Y_q .

After that, the similar adapted templates are categorized using the polarity of each existing A or B sub-sections to be constructed primary polarity patterns. In meanwhile, every primary polarity pattern is an index for the set of adapted templates that their “A” and “B” have been replaced by polarity and their X are a set includes the noun phrase of affect. According to (3),

$$PP_k = AT_h, \forall h, l \begin{cases} AT_{ah} = AT_{al} & \text{if } A \in T_{iq} \\ AT_{bh} = AT_{bl} & \text{if } B \in T_{iq} \\ AT_{ah} = AT_{al}, AT_{bh} = AT_{bl} & \text{if } A, B \in T_{iq} \end{cases} \quad (3)$$

that PP_k is the k th primary polarity pattern. Also, $\forall h, l$ expresses only an adapted template is considered as PP for replacing all adapted templates which the polarity of their sub-sections except X is same.

In the next step, the set of X of every primary polarity pattern is classified to positive and negative classes. The classification is done rather than the polarity of “X”s. Then, the classified primary polarity patterns and the final polarity of phrases are investigated to be extracted polarity patterns. For this purpose, final polarity patterns are provided based on the different states of “A” and “B” in similar primary polarity patterns which their polarity of “X” is same. If P_m is the m th extracted polarity pattern, that will be presented based on (4)

$$P_m = \begin{cases} \{X^+\} + (PP_k - \{X\}) & \text{if } \{X\} \text{ is } + \\ \{X^-\} + (PP_k - \{X\}) & \text{if } \{X\} \text{ is } - \end{cases} \quad (4)$$

that $\{X\}$, $\{X^+\}$ and $\{X^-\}$ are the set of “X”s, and classified “X”s, respectively. According to the detected phrase groups in corpus, the corresponding polarity of every state is determined by experts, and as a result, the polarity patterns are prepared for opinion mining.

For generalization of polarity patterns, the sets of “X” are extended using semantic similarity strategies, because similar words can be replaced each other. For this purpose, the synonyms, hypernyms, hyponyms, co-occurrences, and co-locations of all sub-sets of “X”s are extracted. Then, every word would be added to its set of “X” if there was not. Finally, the extracted polarity patterns and the prepared sets of “X” are used for context-sensitive opinion mining.

B. Opinion Mining Using Polarity Pattern Matching

The goal of proposed method in this paper is to analyze the opinion of texts using polarity pattern matching. The main stages of context-sensitive opinion mining are indicated in “Fig. 2”.

The first stage is the pre-processing for context-sensitive opinion mining, which involves normalization, tokenizing, sentence detection, and Part-Of-Speech (POS) tagging. The dependency parser [32] is used for syntactic relations identification besides words and their POS. Following, the syntactic relations are used for polarity pattern matching.

In the next stage, which is main stage for the proposed method, the polarity identification of sections, which are matched with the polarity patterns, is begun. Firstly, the words of sentence are compared with the sets of “X”. If each word is matched, its related phrases as sub-sections are extracted based on its corresponded polarity pattern and the dependency tree. Then, the polarity of sub-sections is measured as direct or recursive using polarity patterns. In meanwhile, the direct polarity of words are calculated using SentiFarsNet. Thus, the polarity of phrase is considered as the result of polarity pattern.

Finally, opinion mining measures the polarity of every sentence. After that the polarity of all sections of sentence was calculated, the polarity of sentence would be determined dependent on the polarities of sections.

IV. EVALUATION

This paper implements the context-sensitive opinion mining using polarity pattern matching. For this purpose, the polarity patterns were extracted for politics domain. Also, the

SentiFarsNet that is the translation of SentiWordNet based on Farsnet is applied to find prior opinion.

A. Dataset

For this paper, we gathered data by a crawler from the political news. The news are collected from Mar. 2013 to Dec. 2015 on three news sites. Also, the crawled data contains 208,000 political news. After that, we selected our corpus from gathered data. The corpus includes about 14,000 news between 2KB and 10KB. Then, pre-processes were implemented on the corpus and dependency trees were obtained for opinion mining.

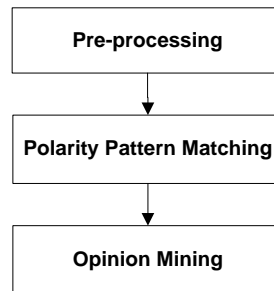


Fig. 2. The stages of context-sensitive opinion mining

B. Results

According to the proposed method of this paper, the polarity patterns are extracted that TABLE II is shown some samples. The investigation of corpus indicates 38% group nouns belong to the polarity patterns.

TABLE II. THE EXAMPLES OF POLARITY PATTERNS

X		A	Conjunction		B	Polarity of phrase
Noun phrase of affect	translate					
پشتیبانی	Supporting		از	of		
+		+			+	+
+		+			-	-
+		-			+	-
+		-			-	-
Other words of this polarity pattern						
اعتماد	Trust		به	of		
اطمینان	Confidence		به	of		
مقاومت	Resistance		در مقابل	against		

The first evaluation is calculating the exactness and completeness of polarity patterns. For this purpose, precision is defined as the fraction of obtained polarity patterns that are relevant to the content. In the other hand, recall is the fraction of the polarity patterns are relevant to the content that are successfully retrieved in documents. Finally, F-measure is an average of precision and recall that “Fig. 3” presents the quality of the extracted polarity patterns. Whereas the coverage of required polarity patterns for opinion mining is the most measurement, the value of precision, which is more than 90%, demonstrates the performance of the proposed method.

Of course, the precision of n-gram is more and better than the proposed method but in contrast, the recall of proposed method is better. As a result, the polarity pattern method needs less calculations and has more f-measure than n-gram.

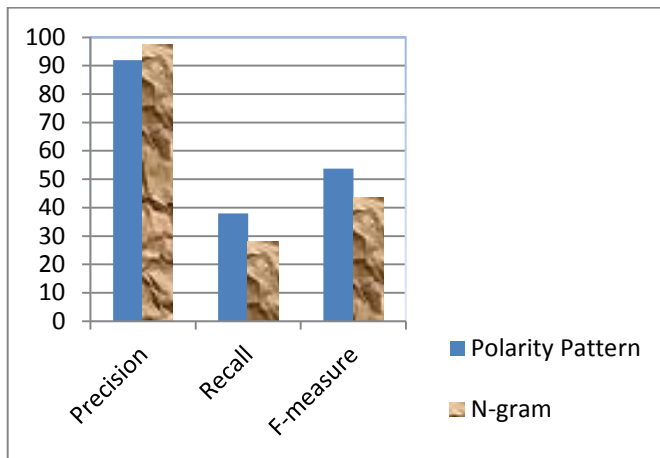


Fig. 3. The quality of the extracted polarity patterns and n-grams

Also, the accuracy of proposed method is compared with a base method that this paper uses bag of words and their polarity. The results are observed in “Fig. 4”. As the results show, the determination of polarity based on the polarity patterns leads to increase accuracy in opinion mining.

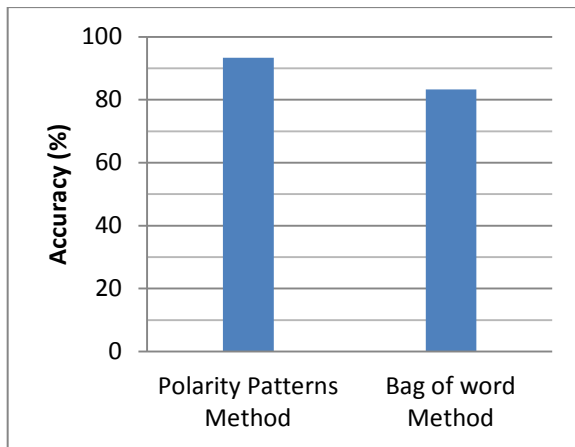


Fig. 4. The comparison of polarity patterns and bag of words usage

V. CONCLUSION

The growing of Web 2.0 has lead to extend interesting to opinion mining in research society. Whereas the aim of opinion mining is to recognize the positive or negative opinion of documents and opinionated sentences [33], the polarity of words and expressions is the most important issue. In the meanwhile, researches have shown polarity is not fix and even changes in a domain. Also, Ding expressed the polarity is fix only rather than a context [11]. Therefore, context-sensitive opinion mining is presented and indicates the primary polarity does not determine correct polarity.

Most methods often consider the polarity of words as uni-gram, but how words are placed adjacent to each other is very useful for opinion mining. Also, n-gram is used only as features for machine learning algorithms.

When words get together, they affect on the polarity of each other. In the other hand, n-gram is not a suitable method because the replacement of words is not considered for its

statistical calculation; As a result, n-gram leads to sparseness and lack of generalizability. Thus, this paper proposes opinion mining is done using polarity pattern extraction based on language model helps to be possible words, which are synonym or have similar polarity, are replaced.

The polarity patterns categorize expressions which their polarity is same. For example, “supporting of terrorism” is negative because “terrorism” is negative and the polarity pattern of “supporting of (Neg. Exp.)” leads to a negative result. Consequently, the results of context-sensitive opinion mining using polarity pattern matching show the significant improvement of accuracy the determination of polarity based on the polarity patterns rather than other methods.

REFERENCES

- [1] J. Wiebe, "Tracking point of view in narrative," *Computational Linguistics*, vol. 20, pp. 233-287, 1994.
- [2] J. Wiebe, et al., "Learning subjective language," *Computational linguistics*, vol. 30, pp. 277-308, 2004.
- [3] B. Liu, "Sentiment analysis and subjectivity," *Handbook of natural language processing*, vol. 2, pp. 627-666, 2010.
- [4] L. Zhang and B. Liu, "Aspect and entity extraction for opinion mining," in *Data mining and knowledge discovery for big data*, ed: Springer, 2014, pp. 1-40.
- [5] B. Pang, et al., "Thumbs up?: sentiment classification using machine learning techniques," in *Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10*, 2002, pp. 79-86.
- [6] M. Koppel and J. Schler, "Using neutral examples for learning polarity," in *INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE*, 2005, p. 1616.
- [7] M. Koppel and J. Schler, "The importance of neutral examples for learning sentiment," *Computational Intelligence*, vol. 22, pp. 100-109, 2006.
- [8] A. Esuli and F. Sebastiani, "Determining Term Subjectivity and Term Orientation for Opinion Mining," in *EACL*, 2006, p. 2006.
- [9] A. Aue and M. Gamon, "Customizing sentiment classifiers to new domains: A case study," in *Proceedings of recent advances in natural language processing (RANLP)*, 2005, p. 2.1.
- [10] S. Mahalakshmi and E. Sivasankar, "Cross Domain Sentiment Analysis Using Different Machine Learning Techniques," in *Proceedings of the Fifth International Conference on Fuzzy and Neuro Computing (FANCCO-2015)*, 2015, pp. 77-87.
- [11] X. Ding, et al., "A holistic lexicon-based approach to opinion mining," in *Proceedings of the 2008 International Conference on Web Search and Data Mining*, 2008, pp. 231-240.
- [12] A. Ko, et al., "Ontology Supported Policy Modeling in Opinion Mining Process," in *On the Move to Meaningful Internet Systems: OTM 2012 Workshops*, 2012, pp. 252-261.
- [13] H. Kanayama and T. Nasukawa, "Fully automatic lexicon expansion for domain-oriented sentiment analysis," in *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing*, 2006, pp. 355-363.
- [14] Z. Chen, et al., "Joint model for subsentence-level sentiment analysis with Markov logic," *Journal of the Association for Information Science and Technology*, 2015.
- [15] R. Y. Lau, et al., "Social analytics: Learning fuzzy product ontologies for aspect-oriented sentiment analysis," *Decision Support Systems*, vol. 65, pp. 80-94, 2014.
- [16] L. Liu, et al., "Toward a fuzzy domain sentiment ontology tree for sentiment analysis," in *Image and Signal Processing (CISP)*, 2012 5th International Congress on, 2012, pp. 1620-1624.
- [17] Y. Jo and A. H. Oh, "Aspect and sentiment unification model for online review analysis," in *Proceedings of the fourth ACM international conference on Web search and data mining*, 2011, pp. 815-824.

- [18] L. Garcia-Moya, et al., "Retrieving product features and opinions from customer reviews," *IEEE Intelligent Systems*, pp. 19-27, 2013.
- [19] A. Andreevskaia and S. Bergler, "When Specialists and Generalists Work Together: Overcoming Domain Dependence in Sentiment Tagging," in *ACL*, 2008, pp. 290-298.
- [20] P. Sanju and T. Mirnalinee, "Construction of Enhanced Sentiment Sensitive Thesaurus for Cross Domain Sentiment Classification Using Wiktionary," in *Proceedings of the Third International Conference on Soft Computing for Problem Solving*, 2014, pp. 195-206.
- [21] Y. Zhang, et al., "Joint naive bayes and lda for unsupervised sentiment analysis," in *Advances in Knowledge Discovery and Data Mining*, ed: Springer, 2013, pp. 402-413.
- [22] D. R. Rice and C. Zorn, "Corpus-based dictionaries for sentiment analysis of specialized vocabularies," *Proceedings of NDATAD*, 2013.
- [23] S. A. Bahrainian, et al., "Fuzzy Subjective Sentiment Phrases: A Context Sensitive and Self-Maintaining Sentiment Lexicon," in *Proceedings of the 2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)-Volume 01*, 2014, pp. 361-368.
- [24] N. Blenn, et al., "Context-sensitive sentiment classification of short colloquial text," in *NETWORKING 2012*, ed: Springer, 2012, pp. 97-108.
- [25] Z. Zhang, et al., "Context-Dependent Sentiment Classification Using Antonym Pairs and Double Expansion," in *Web-Age Information Management*, ed: Springer, 2014, pp. 711-722.
- [26] T. Wilson, et al., "Recognizing contextual polarity in phrase-level sentiment analysis," in *Proceedings of the conference on human language technology and empirical methods in natural language processing*, 2005, pp. 347-354.
- [27] H. Saif, et al., "Adapting Sentiment Lexicons Using Contextual Semantics for Sentiment Analysis of Twitter," in *The Semantic Web: ESWC 2014 Satellite Events*, ed: Springer, 2014, pp. 54-63.
- [28] E. Riloff and J. Wiebe, "Learning extraction patterns for subjective expressions," in *Proceedings of the 2003 conference on Empirical methods in natural language processing*, 2003, pp. 105-112.
- [29] Y. Bengio, et al., "A neural probabilistic language model," *The Journal of Machine Learning Research*, vol. 3, pp. 1137-1155, 2003.
- [30] I. Maks and P. Vossen, "A lexicon model for deep sentiment analysis and opinion mining applications," *Decision Support Systems*, vol. 53, pp. 680-688, 2012.
- [31] D. Tang, et al., "A joint segmentation and classification framework for sentence level sentiment classification," *Audio, Speech, and Language Processing, IEEE/ACM Transactions on*, vol. 23, pp. 1750-1761, 2015.
- [32] S. Huang, et al., "Automatic construction of domain-specific sentiment lexicon based on constrained label propagation," *Knowledge-Based Systems*, vol. 56, pp. 191-200, 2014.
- [33] B. Pang and L. Lee, "Opinion mining and Sentiment Analysis," *Foundations and Trends in Information Retrieval*, vol. 2, pp. 1-135, 2008.