

Domain Independent Feature Extraction using Rule Based Approach

Sint Sint Aung^{*1}, Myat Su Wai²

¹Department of Academic Affairs, University of Computer Studies, Mandalay, +95, Myanmar

²Web Mining Lab, University of Computer Studies, Mandalay, +95, Myanmar

ARTICLE INFO

Article history:

Received: 16 November, 2017

Accepted: 09 January, 2018

Online: 30 January, 2018

Keywords:

Sentiment Analysis

Features Extraction

Opinions

ABSTRACT

Sentiment analysis is one of the most popular information extraction tasks both from business and research prospective. From the standpoint of research, sentiment analysis relies on the methods developed for natural language processing and information extraction. One of the key aspects of it is the opinion word lexicon. Product's feature from online reviews is an important and challenging task in opinion mining. Opinion Mining or Sentiment Analysis is a Natural Language Processing and Information Extraction task that identifies the user's views or opinions. In this paper, we developed an approach to extract domain independent product features and opinions without using training examples i.e, lexicon-based approach. Noun phrases are extracted using not only dependency rules but also textblob noun phrase extraction tool. Dependency rules are predefined according to dependency patterns of words in the sentences. StanfordCoreNlp Dependency parser is used to identify the relations between words. The orientation of words is classified by using lexicon-based approach. According to the experimental results the system gets good performance in six different domains.

1. Introduction

Sentiment analysis is one of the most popular information extraction tasks both from business and research prospective. It has numerous business applications, such as evaluation of a product or company perception in social media [1]. From the standpoint of research, sentiment analysis relies on the methods developed for natural language processing and information extraction. One of the key aspects of it is the opinion word lexicon. Opinion words are such words that carry opinion. Positive words refer to some desired state, while negative words to some undesired one. For example, "good" and "beautiful" are positive opinion words, "bad" and "evil" are negative [2].

Opinion phrases and idioms exist as well. Many opinion words depend on context, like the word "large". Some opinion phrases are comparative rather than opinionated, for example "better than". Auxiliary words like negation can change sentiment orientation of a word[3].

Opinion words are used in a number of sentiment analysis tasks. They include document and sentence sentiment classification, product features extraction, subjectivity detection etc. [4]. Opinion words are used as features in sentiment classification. Sentiment orientation of a product feature is usually computed based on the sentiment orientation of opinion words nearby [5]. Product features can be extracted with the help of phrase or dependency patterns that include opinion words and placeholders for product features themselves. Subjectivity detection highly relies on opinion word lists as well, because many opinionated phrases are subjective in [6]. Thus, opinion lexicon generation is an important sentiment analysis task. Detection of opinion word sentiment orientation is an accompanying task.

Opinion lexicon generation task can be solved in several ways. The authors of [7] point out three approaches: manual, dictionary-based and corpus-based. The manual approach is precise but time-consuming. The dictionary based approach relies on dictionaries such as WordNet. One starts from a small collection of opinion words and looks for their synonyms and antonyms in a dictionary [8]. The drawback of this approach is that the dictionary coverage

^{*}Corresponding Author: Sint Sint Aung, Email: ssaung@gmail.com

is limited and it is hard to create a domain-specific opinion word list. Corpus-based approaches rely on mining a review corpus and use methods employed in information extraction. The approach proposed in [9] is based on a seed list of opinion words. These words are used together with some linguistic constraints like “AND” or “OR” to mine additional opinion words.

Clustering is performed to label the mined words in the list as positive and negative. Part of speech patterns are used to populate the opinion word dictionary in and Internet search statistics is used to detect semantic orientation of a word. In [10], the authors extend the mentioned approaches and introduces a method for extraction of context-based opinion words together with their orientation. Classification techniques are used in [11] to filter out opinion words from text. The approaches described were applied in English. There are some works that deal with Russian. For example, in [12], the author proposed to use classification. Various features, such as word frequency, weirdness, and TF-IDF are used there.

Most of the research done in the field of sentiment analysis relies on the presence of annotated resources for a given language. However, there are methods which automatically generate resources for a target language, given that there are tools and resources available in the source language. Different approaches to multilingual subjectivity analysis are studied and are summarized in [13].

In one of them, subjectivity lexicon in the source language is translated with the use of a dictionary and employed for subjectivity classification. This approach delivers mediocre precision due to the use of the first translation option and due to word lemmatization. Another approach suggests translating the corpus. This can be done in three different ways: translating an annotated corpus in the source language and projecting its labels; automatic annotation of the corpus, translating it and projecting the labels; translating the corpus in the target language, automatic annotation of it and projecting the labels. Language Weaver 1 machine translation was used on English-Roman and English-Spanish data. Classification experiments with the produced corpora showed similar results. They are close to the case when test data is translated and annotated automatically. This shows that machine translation systems are good enough for translating opinionated datasets. In [14], the authors also confirmed when they used Google Translate 2, Microsoft Bing Translator 3 and Moses 4.

Multilingual opinion lexicon generation is considered in [that presents a semi-automatic approach with the use of triangulation. The authors use high-quality lexicons in two different languages and then translate them automatically into a third language with Google Translate. The words that are found in both translations are supposed to have good precision. It was proven for several languages including Russian with the manual check of the resulting lists. The same authors collect and examine entity-centered sentiment annotated parallel corpora [15].

The process of automatic extraction of knowledge by means of opinion of others about some particular product, topic or problem. Opinion mining is also called sentiment analysis due to large volume of opinion which is rich in web resources available online. Analyzing customer review is most important, it tend to rate the

product and provide opinions for it which is been a challenging problem today. Opinion feature extraction is a sub problem of opinion mining, with the vast majority of existing work done in the product review domain. Main fields of research in sentiment analysis are Subjectivity Detection, Sentiment Prediction, Aspect based Sentiment Summarization, Text summarization for opinions, Contractive viewpoint, Summarization, Product Feature Extraction, Detecting opinion spam [16].

2. Related Works

Many researchers have addressed the problem of constructing subjective lexicon for different languages in recent years. In [17] to compile a subjective lexicon, the author investigated three main approaches and they are outlined in this section.

knowledge to extract the domain-specific sentiment lexicon based on constrained label propagation. According to [18], the authors had divided the whole strategy into six steps. Firstly, detected and extracted domain-specific sentiment terms by combining the chunk dependency parsing knowledge and prior generic sentiment lexicon. To refine the sentiment terms some filtering and pruning operations were carried out by others. Then they selected domain-independent sentiment seeds from the semi-structured domain reviews which had been designated manually or directly borrowed from other domains. As the third step, calculated the semantic associations between sentiment terms based on their distribution contexts in the domain corpus. For this calculation, the point-wise mutual information (PMI) was utilized which is commonly used in semantic linkage in information theory. Then, they defined and extracted some pair wise contextual and morphological constraints between sentiment terms to enhance the associations. The conjunctions like “and” and “as well as” were considered as the direct contextual constraints whereas “but” was referred to as a reverse contextual constraint. The above constraints propagated though out the entire collection of candidate sentiment terms. Finally, the propagated constraints were incorporated into label propagation for the construction of domain-specific sentiment lexicon. In [19], the authors proposed approach showed an accuracy increment of approximately 3% over the baseline methods. Opinion analysis has been studied by many researchers in recent years. Two main research directions are sentiment classification and feature-based opinion mining. Sentiment classification investigates ways to classify each review document as positive, negative, or neutral. Representative works on classification at the document level include. These works are different from ours as we are interested in opinions expressed on each product feature rather than the whole review.

In [20], sentence level subjectivity classification is studied, which determines whether a sentence is a subjective sentence (but may not express a positive or negative opinion) or a factual one. Sentence level sentiment or opinion classification is studied in. Our work is different from the sentence level analysis as we identify opinions on each feature. A review sentence can contain multiple features, and the orientations of opinions expressed on the features can also be different, e.g., “the voice quality of this phone is great and so is the reception, but the battery life is short.” “voice quality”, “reception” and “battery life” are features. The opinion on “voice quality”, “reception” are positive, and the opinion on

“battery life” is negative. Other related works at both the document and sentence levels include those in [21].

Most sentence level and even document level classification methods are based on identification of opinion words or phrases. There are basically two types of approaches: (1) corpus-based approaches, and (2) dictionary-based. approaches. Corpus-based approaches find co-occurrence patterns of words to determine the sentiments of words or phrases, e.g., the works in [22].

In [23], the authors proposed the idea of opinion mining and summarization. It uses a lexicon-based method to determine whether the opinion expressed on a product feature is positive or negative. In [24] and [25]. these methods are improved by a more sophisticated method based on relaxation labeling. We will show in Section 5 that the proposed technique performs much better than both these methods. In [26], a system is reported for analyzing movie reviews in the same framework. However, the system is domain specific. Other recent work related to sentiment analysis includes in. In [27], the authors studied the extraction of comparative sentences and relations, which is different from this work as we do not deal with comparative sentences in this research.

Our holistic lexicon-based approach to identifying the orientations of context dependent opinion words is closely related to works that identify domain opinion words . In [28], the authors used conjunction rules to find such words from large domain corpora. In [29], the conjunction rule basically states that when two opinion words are linked by “and” in a sentence, their opinion orientations are the same. For example, in the sentence, “this room is beautiful and spacious”, both “beautiful” and “spacious” are positive opinion words. Based on this rule or language convention, if we do not know whether “spacious” is positive or negative, but know that “beautiful” is positive, we can infer that “spacious” is also positive. Although our approach will also use this linguistic rule or convention, our method is different in two aspects. First, we argue that finding domain opinion words is still problematic because in the same domain the same word may indicate different opinions depending on what features it is applied to. For example, in the following review sentences in the camera domain, “the battery life is very long” and “it takes a long time to focus”, “long” is positive in the first sentence, but negative in the second. Thus, we need to consider both the feature and the opinion word rather than only the opinion word as in [30].

Opinion target and opinion word extraction are not new tasks in opinion mining. There is significant effort focused on these tasks. They can be divided into two categories: sentence-level extraction and corpus-level extraction according to their extraction aims. In sentence-level extraction, the task of opinion target/word extraction is to identify the opinion target mentions or opinion expressions in sentences. Thus, these tasks are usually regarded as sequence-labeling problems. Intuitively, contextual words are selected as the features to indicate opinion targets/words in sentences. Additionally, classical sequence labeling models are used to build the extractor, such as CRFs and HMM. In [31], the authors proposed a lexicalized HMM model to perform opinion mining. In [32], the authors used CRFs to extract opinion targets from reviews. However, these methods always need the labeled data to train the model. If the labeled training data are insufficient or come from the different domains than the current texts, they

would have unsatisfied extraction performance. Although in [33], the authors proposed a method based on transfer learning to facilitate cross- domain extraction of opinion targets/words, their method still needed the labeled data from out-domains and the extraction performance heavily depended on the relevance between in-domain and out-domain.

Although many target extraction methods exist, we are not aware of any attempt to solve the proposed problem. According to [34], although in supervised target extraction, one can annotate entities and aspects with different labels, supervised methods need manually labeled training data, which is time-consuming and labor-intensive to produce. Note that relaxation labeling was used for sentiment classification in, but not for target classification.

3. Proposed method

This section presents the detailed of step by step process about the system.

3.1. Preprocessing the Input sentences

Input sentences with xml file are prepared before parsing to the StanfordCoreNLP parser. Xml tag are removed. ASCII code characters are replaced with Unicode characters because StanfordCoreNLP cannot process non-Unicode characters.

3.2. Rules for Features and Opinions Extraction

In this section, we describe how to extract opinion and product features using extraction rules. They are the most important tasks for text sentiment analysis, which has attracted much attention from many researchers. Based on the relations between features and opinions, there are four main rules in the double propagation;

1. extracting features using opinion words
2. extracting features using the extracted features
3. extracting opinion words using the extracted features
4. extracting opinion words using both the given and the extracted opinion words

In the following extraction rules, O is opinion word, H is the third word, {O} is a set of seed opinion lexicon, F is product feature, and O-Dep is part-of-speech information and dependency relations. {JJ}, {VB} and {NN} are sets of POS tags of potential opinion words and features, respectively. And {DR} contains dependency relations between features and opinions such as mod, pnmmod, subj, s, obj, obj2, conj. We used rule 1 and 2 to extract features, and rule 3 and 4 use to extract opinion words. Moreover, we also used some additional patterns to extract features and opinions.

R11: If a word F whose POS is NN is directly depended by an opinion word O through one of the dependency relations mod, pnmmod, subj, s, obj and obj2, then F is a feature. It can be defined as follows;

$$O \rightarrow O\text{-Dep} \rightarrow F$$

$$F \rightarrow F\text{-Dep} \rightarrow O$$

such that $O \in \{O\}$, $O\text{-Dep}$ and $F\text{-Dep} \in \{DR\}$, where $\{DR\} = \{\text{mod, pnm, subj, s, obj, obj2, desc}\}$ and $POS(O) \in \{NN\}$.

For eg. Overall a sweet machine.

R12: If an opinion word O and a word F , whose POS is NN , directly depend on a third word H through dependency relations except conj , then F is a feature. It can be expressed as follows;

$$O \rightarrow O\text{-Dep} \rightarrow H \leftarrow F\text{-Dep} \leftarrow F$$

such that $O \in \{O\}$, $O\text{-Dep}$ and $F\text{-Dep} \in \{DR\}$, $POS(F) \in \{NN\}$. For eg. Canon is the great product.

R13: If a word F whose POS is NN is indirectly depended by an opinion word O through another word H through two dependency relations dobj and amod or nmod:poss , then F is a feature. It can also be expressed as follows;

$$O \rightarrow O\text{-Dep} \rightarrow H \rightarrow F\text{-Dep} \rightarrow F$$

$$O \leftarrow O\text{-Dep} \leftarrow H \leftarrow F\text{-Dep} \leftarrow F$$

such that $O \in \{O\}$, $O\text{-Dep} \in \{DR\}$, $F\text{-Dep} \in \{DR\}$, $POS(F) \in \{NN\}$. For eg. I like the computer's battery.

R21: If a word F_j , whose POS is NN , directly depends on a feature F_i through conj , then F_j is a feature. It can also be expressed as follows;

$$F_i \rightarrow F_i\text{-Dep} \rightarrow F_j$$

such that $F_i \in \{F\}$, $F_j\text{-Dep} \in \{CONJ\}$, $POS(F_j) \in \{NN\}$. For eg. Overall, I like the system features and performance.

R22: If a word F_j , whose POS is NN , and a feature F_i , directly depend on a third word H through the same dependency relation, then F_j is a feature. It can also be expressed as follows;

$$F_i \rightarrow F_i\text{-Dep} \rightarrow H \leftarrow F_j\text{-Dep} \leftarrow F_j$$

such that $F_i \in \{F\}$, $F_i\text{-Dep} \in \{DR\}$, $F_j\text{-Dep} \in \{DR\}$, $POS(F_j) \in \{NN\}$. For eg. Canon has done an excellent job.

R31: If a word O whose POS is JJ or VB directly depends on a feature F through one of the dependency relations mod , pnm , subj , s , obj , obj2 and desc , then O is an opinion word. It can also be expressed as follows;

$$O \rightarrow O\text{-Dep} \rightarrow F$$

$$F \rightarrow F\text{-Dep} \rightarrow O$$

such that $F \in \{F\}$, $O\text{-Dep}$ and $F\text{-Dep} \in \{DR\}$, $POS(O) \in \{JJ, VB\}$. For eg. Overall a sweet machine.

R32: If a word O whose POS is JJ or VB and a feature F directly depend on a third word H through dependency relations except conj , then O is an opinion word. It can also be expressed as follows;

$$O \rightarrow O\text{-Dep} \rightarrow H \leftarrow F\text{-Dep} \leftarrow F$$

such that $F \in \{F\}$, $O\text{-Dep}$ and $F\text{-Dep} \in \{DR\}$, $POS(O) \in \{JJ, VB\}$. For eg. Canon is the great product.

R33: If a word O whose POS is JJ or VB indirectly depends on a feature F through another word H through dependency relations

dobj , amod and nmod:poss , then O is an opinion word. It can also be expressed as follows;

$$O \rightarrow O\text{-Dep} \rightarrow H \rightarrow F\text{-Dep} \rightarrow F$$

$$O \leftarrow O\text{-Dep} \leftarrow H \leftarrow F\text{-Dep} \leftarrow F$$

such that $F \in \{F\}$, $O\text{-Dep}$ and $F\text{-Dep} \in \{DR\}$, $POS(O) \in \{JJ\}$. For eg. I like the computer's battery.

R41: If a word O_j , whose POS is JJ or VB , directly depends on an opinion word O_i through dependency relation conj , then O_j is an opinion word. It can also be expressed as follows;

$$O_i \rightarrow O_i\text{-Dep} \rightarrow O_j$$

such that $O_i \in \{O\}$, $O_i\text{-Dep} \in \{CONJ\}$, $POS(O_j) \in \{JJ, VB\}$. For eg. Nice and compact.

R42: If a word O_j , whose POS is JJ or VB , and an opinion word O_i , directly depend on a third word H through the same dependency relation, then O_j is an opinion word. It can also be expressed as follows;

$$O_i \rightarrow O_i\text{-Dep} \rightarrow H \leftarrow O_j\text{-Dep} \leftarrow O_j$$

such that $O_i \in \{O\}$, $O_i\text{-Dep} = O_j\text{-Dep}$, $POS(O_j) \in \{JJ, VB\}$. For eg. The screen size and screen quality is amazing.

3.3. Features and Opinion Extraction

In this section This system takes raw data as input and xml tag and ascii code characters are removed. After that, word tokenization, part-of speech tagging and dependency identification between words are done by using StanfordCoreNLP dependency parser. We used the algorithm also from [17]. Table 1 shows some examples of English stop word list.

Table 1. Some English Stopwords List

to	Of	I	Me	My
Mine	You	At	They	In
Which	With	On	Under	Below
Above	Thing	Things	Some	Someone
Sometime	Something	Somebody	No one	nobody

To start the extraction process, a seed opinion lexicon, a list of general words, review data and extraction rules are input to the proposed algorithm. The extraction process uses a rule-based approach using the relations defined in above. The system assumed opinion words to be adjectives, adverbs and verbs in some cases. And product features are nouns or noun phrases and also verbs in some cases.

Its primary idea is that opinion words are usually associated with product features in some ways. Thus, opinion words can be recognized by identified features, and features can be identified by known opinion words. So, the extracted opinion words and product features are used to identify new opinion words and new product features. The extraction process ends when no more opinion words or product features can be found.

Table 2. Some Unigram Features

Price	System	Computer	Laptop	Reviews
Customer	Battery	Camera	Power	Screen
Staff	Food	Wine	Window	Notebook
Work	Use	Download	Connection	Cost
D-link	Device	Feature	Model	quality

Moreover, textblob is used to extract ngram noun phrase words from the sentences that are not covered with extraction rules. It can increase the performance of the system in term of precision, recall, and f1-score. In order to increase the accuracy, stop words and some general words are removed during the extraction time. Table 2, 3 and 4 describe some extracted of unigrams, bigrams, trigrams and n-gram words from the system.

Table 3. Some Bigram Features

Picture quality	Battery life	Tech support	Screen quality
Power supply	Printer sharing	Return policy	Set up
Security setting	Service tech	Setup software	Mac support
Guest feature	8GB RAM	web cam	charger unit
connect quality	cooling system	cordless mouse	data rate
Window 7	Cd drive	Hard drive	Service tech

The system constructs n-gram dictionary to refine the noun phrase extraction. If the phrase contains in this dictionary, the system extracts it as a feature. Otherwise, remove it.

Table 4. Sme N-gram Features

D-Link support crew	customer service agents
cover for the DVD drive	Dell's customer disservice
D-Link support crew	extended life battery
Garmin GPS software	sound quality via USB
customer service center	design based programs
direct Electrical connectivity	built it web cam
fingerprint reader driver	technical service for dell

4. Classifying Polarity Orientation

Opinions are classified by using Vader lexicon and qualitative analysis techniques developed by C.J. Hutto and Eric Gilbert, 2014. This deep qualitative analysis resulted in isolating five generalizable heuristics based on grammatical and syntactical cues to convey changes to sentiment intensity. They incorporate word-order sensitive relationships between terms:

Punctuation, namely the exclamation point (!), increases the magnitude of the intensity without modifying the semantic orientation.

Capitalization, specifically using ALL-CAPS to emphasize a sentiment relevant word in the presence of other non-capitalized words, increases the magnitude of the sentiment intensity without affecting the semantic orientation

Degree modifiers (also called intensifiers, booster words, or degree adverbs) impact sentiment intensity by either increasing or decreasing the intensity.

The **contrastive conjunction** “but” signals a shift in sentiment polarity with the sentiment.

By examining the **tri-gram** to deeply analyze the intensity of sentiment orientation.

Table 5. shows some example of polarity classification. Polarity score included negative sign (-) indicates negative opinion. And, score with positive sign (+) indicates positive opinion.

Table 5. Polarity Classification of the System

Words	Polarity Score	Polarity Label
Set up	0.4404	positive
excellent	0.5719	positive
work	0.7264	positive
installation disk	-0.296	negative
problem	-0.4019	negative

5. Experimental Results

For experiment, we use core i7 processor, 4GB RAM and 64-bit Ubuntu OS, And, we implement the proposed system with python programming language (PyCharm 2016.3 IDE for python).

Table 6. Dataset Used in the System

Dataset	no of sentences	no of features
ABSA15 restaurant	1083	1193
ABSA15 Hotel	266	212
Router	245	304
Speaker	291	435
Computer	239	346
iPod	161	293
Linksys Router	192	375
Nokia 6000	363	633
Norton	210	302
Diaper Champ	212	239

In this paper, 10 product review datasets are collected and evaluated in term of precision, recall and f1-score. Table 6 shows the domains according to their names, the number of sentences

and the number of features. For performance evaluation on product feature extraction, the comparative results between the proposed approach and Qiu's approach are also analyzed. Precision, recall and f1-score of both are described in figure 1, 2 and 3.

From figure 1, we can see that the proposed approach has higher precision in 6 datasets, dropped in 3 datasets and nearly the same in 1 dataset over Qiu's approach. The proposed approach relies only on the review data itself and no external information is needed.

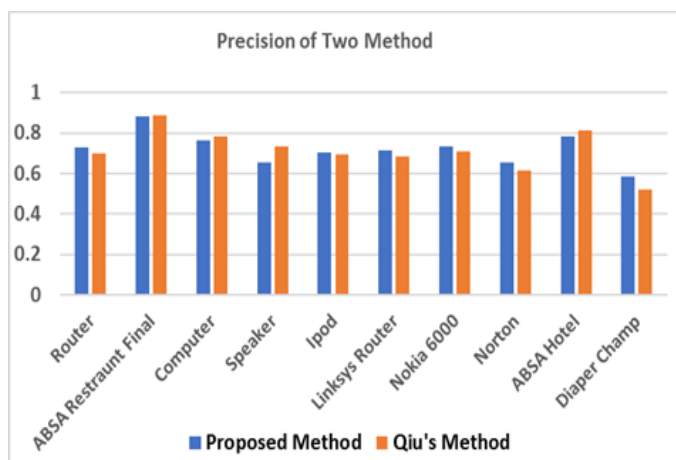


Figure 1: Comparison of Precision on Feature Extraction between Two Approaches

According to the experimental results, the highest precision 0.8819 (88%) are achieved in ABSA Restaurant dataset. In this dataset, the precision, recall and f1-score are nearly the same because there are no verb product features in this dataset.

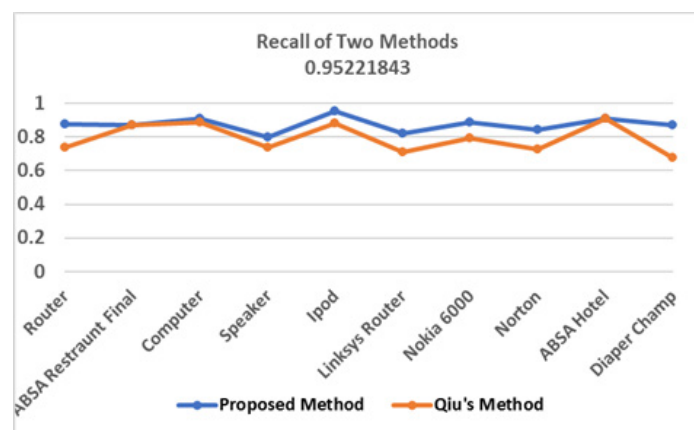


Figure 2: Comparison of Recall on Feature Extraction between Two Approaches

Figure 2 shows that proposed approach outperforms all the other approaches in recall except ABSA Restaurant dataset and ABSA Hotel dataset. The proposed approach has about 14% improvement in Router dataset, about 3% in Computer dataset, about 6% in Speaker dataset, about 7% in iPod dataset, about 10% in Linksys Router dataset, about 9% in Nokia 6000 dataset, about 12% in Norton dataset, about 20% in Diaper Champ dataset. In ABSA Restaurant and ABSA Hotel datasets, the recall of the two

approaches are the same. So, to sum up, the proposed approach outperforms over the Qiu's approach according to the recall.

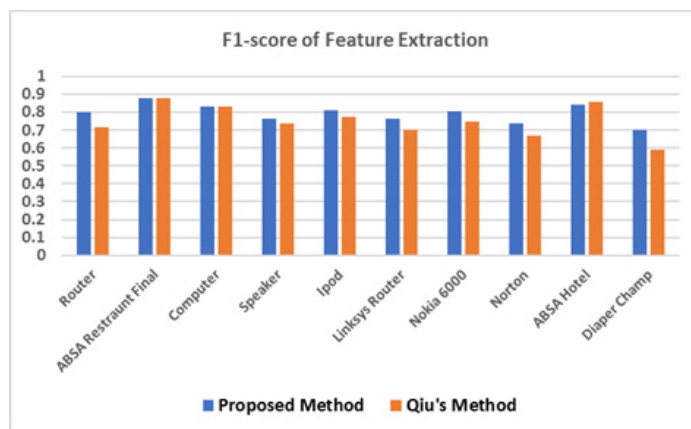


Figure 3: Comparison of F1-score on Feature Extraction between Two Approaches

The comparative results of f1-score in feature extraction are shown in figure 3. According to the experimental results, the proposed approach has higher f1-score (about 7%) in 7 datasets and about 0.003% drop in ABSA Restaurant and Computer datasets and 1% drop in ABSA Hotel dataset

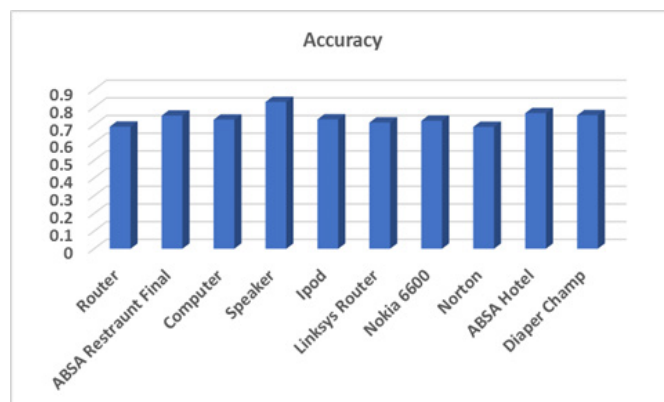


Figure 4 Experimental results of Polarity Classification

The performance analysis of polarity classification of the proposed system is evaluated in all datasets. Figure 4 shows the experimental results of polarity classification in all datasets. The system achieves Highest accuracy 83% in speaker dataset.

6. Conclusion

Opinion mining and sentiment classification are not only technically challenging because of the need for natural language processing. In this work, an effective opinion lexicon expansion and feature extraction approach is proposed. Features and opinions words are extracted simultaneously by using proposed algorithm based on double propagation. So, unlike the existing approach, context dependent opinion words are extracted and domain independence. According to experimental results, the proposed system works well in all datasets and get domain independency without using training examples. As the future extension, we will analyze the performance of the proposed system with more different datasets from SemEval research group.

And we will apply more dependency relations in extraction process.

References

- [1] Agarwal, B., Poria, S., Mittal, N., Gelbukh, A., and Hussain, A. "Concept-level sentiment analysis with dependency-based semantic parsing: A novel approach". *Cognitive Computation*, pages 1–13, 2015.
- [2] Asghar M, Khan A, Ahmad S, Kundi F, "A Review of Feature Extraction in Sentiment Analysis", *Journal of Basic and Applied Scientific Research*, 4(3): 181–186, 2014.
- [3] Broß, J. "Aspect-Oriented Sentiment Analysis of Customer Reviews Using Distant Supervision Techniques". PhD thesis, Freie Universität Berlin, 2013.
- [4] Chen, L., Wang, W., Nagarajan, M., Wang, S., Sheth, A. P. "Extracting Diverse Sentiment Expressions with Target-Dependent Polarity from Twitter", In the Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media(ICWSM), 50–57, 2012.
- [5] Chinsha, T. and Joseph, S. "A syntactic approach for aspect-based opinion mining". In *Semantic Computing (ICSC)*, 2015 IEEE International Conference on, pages 24–31. IEEE.
- [6] De Marneffe, M.-C. and Manning, C. D. "Stanford typed dependencies manual. Report", Technical report, Stanford University, 2016.
- [7] Fabrizio G, Aker A, Gaizauskas R012) "Summarizing Online Reviews Using Aspect Rating Distributions and Language Modeling", *IEEE Intelligent Systems*, 28(3): 28–37, 2012.
- [8] G. Qiu, B. Liu, J. Bu, and C. Chen, "Opinion word expansion and target extraction through double propagation," *Computational Linguistics* 37, no.1, 2011, pp. 9–27.
- [9] Jiang, M. Yu, M. Zhou, X. Liu, and T. Zhao, "Target-dependent twitter sentiment classification," In Proc. of the 49th Ann. Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1, pp. 151-160. Association for Computational Linguistics, 2011.
- [10] Khan, K., Baharudin, B., Khan, A., and Ullah, A. "Mining opinion components from unstructured reviews: A review". *Journal of King Saud University-Computer and Information Sciences*, 26(3):258–275, 2014.
- [11] Kundi F, Ahmad S, Khan A, Asghar, "Detection and Scoring of Internet Slangs for Sentiment Analysis Using SentiWordNet", *Life Science Journal*, 11(9):66–72, 2014.
- [12] Liu, B. "Web data mining: exploring hyperlinks, contents, and usage data". Springer Science & Business Media, 2011.
- [13] Liu, B. "Sentiment analysis and opinion mining". *Synthesis Lectures on Human Language Technologies*, 5(1):1–167, 2015.
- [14] Liu, B. "Sentiment analysis mining opinions, sentiments, and emotions". 1:1–386, 2015.
- [15] Liu, L. Xu and J. Zhao, "Opinion target extraction using word-based translation model", *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, pp. 1346-1356, 2012.
- [16] Moghaddam, A. S., "Aspect-based opinion mining in online reviews". PhD thesis, Applied Sciences: School of Computing Science, 2014.
- [17] Myat Su Wai and Sint Sint Aung, "Simultaneous Opinion Lexicon Expansion and Product Feature Extraction", *Proceeding of 16th IEEE/ACIS International Conference on Computer and Information Science (ICIS 2017)*, ISBN: 978–1–5090–5506–7.
- [18] Neviarouskaya A, Aono M, "Sentiment Word Relations with Affect, Judgment, and Appreciation", *IEEE Transactions on Affective Computing*, 4(4): 425–438, 2014.
- [19] Pontiki, M., Galanis, D., Papageorgiou, H., Manandhar, S., and Androutsopoulos, I., "Semeval-2015 Task 12: Aspect Based Sentiment Analysis". In *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)*, pages 486–495, Denver, Colorado. Association for Computational Linguistics.
- [20] Pontiki, M., Galanis, D., Pavlopoulos, J., Papageorgiou, H., Androutsopoulos, I., and Manandhar, S., "Semeval-2014 Task 4: Aspect Based Sentiment Analysis". In *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, pages 27–35, Dublin, Ireland. Association for Computational Linguistics.
- [21] S. Poria, B. Agarwal, Gelbukh, A. Hussain, N. Howard, "Dependency-based semantic parsing for concept-level text analysis", In Gelbukh, A. (ed.) *Computational Linguistics and Intelligent Text Proceeding, Part I. LNCS*, vol. 8403, pp. 113–127. Springer, Heidelberg, 2014.
- [22] S. Poria, E. Cambria, G. Winterstein, G. B. Huang, "Sentic patterns: Dependency-based rules for concept-level sentiment analysis,". *Knowledge-Based Systems*, vol.69, pp. 45-63, 2014.
- [23] S. Poria, E. Cambria, L.W. Ku, C. Gui, A. Gelbukh, "A rule-based approach to aspect extraction from product reviews," In *Workshop Proc. of the 25th International Conference on Computational Linguistics, COLING'14*, pp. 28-37, 2014.
- [24] Shariaty, S. and Moghaddam, S., "Fine-grained opinion mining using conditional random fields". In *Data Mining Workshops (ICDMW)*, 2011 IEEE 11th International Conference on, pages 109–114. IEEE.
- [25] Souza M, Vieira R, Busetti D, Chishman R, Alves I, "Construction of a Portuguese Opinion Lexicon from Multiple Resources", in the proceedings of the 8th Brazilian Symposium in Information and Human Language Technology (STIL' 2011), 59–66. Springer, pp.411-448, 2007.
- [26] Xu, T. J. Zhao, D. Q. Zheng, S. Y. Wang, "Product features mining based on Conditional Random Fields model", *Proceedings of the 2010 International Conference on Machine Learning and Cybernetics*, pp. 3353-3357, 2010.
- [27] Xu, X., Cheng, X., Tan, S., Liu, Y., and Shen, H. "Aspect-level opinion mining of online reviews". *China Communications*, 10(3):25–41.
- [28] Y. Wu, Q. Zhang, X. Huang, and L. Wu, "Phrase dependency parsing for opinion mining,". In *EMNLP'09, Volume 3-Volume 3*, pp. 1533- 1541. Association for Computational Linguistics, 2009.
- [29] Y. Zheng, L. Ye, G. Wu and X. Li, "Extracting product features from Chinese customer reviews," *Proceedings of 3rd International Conference on Intelligent System and Knowledge Engineering*, pp. 285-290, 2008.
- [30] Z. Hai, K. Chang, G. Cong, "One seed to find them all: mining opinion features via association", *Proceedings of the 21st ACM international conference on Information and knowledge management*, pp. 255-264, 2012.
- [31] Zhai, Z., Liu, B., Xu, H., and Jia, P., "Clustering product features for opinion mining". In *Proceedings of the fourth ACM international conference on Web search and data mining*, pages 347–354. ACM.
- [32] Zhang D, Dong H, Yi J, Song L, "Opinion summarization of customer reviews", In *proceedings of the International Conference on Automatic Control and Artificial Intelligence (ACAI 2012)*, 1476–1479.
- [33] Zhang, L. and Liu, B. "Aspect and entity extraction for opinion mining". In *Data mining and knowledge discovery for big data*, pages 1–40. Springer.
- [34] Zhang, Y. and Zhu, W. "Extracting implicit features in online customer reviews", In *Proceedings of the 15th ACM international conference on Information and knowledge management*, pages 43–50. ACM.