

Article

Sentence Level Domain Independent Opinion and Targets Identification in Unstructured Reviews

Kahirullah Khan ¹ and Wahab Khan ^{2,*}

¹ Department of Computer Science, University of Science & Technology, 28100 Bannu, Pakistan; khairust@gmail.com

² Department of Computer Science and Software Engineering, International Islamic University, Islamabad 44000, Pakistan

* Correspondence: wahab.phdcs72@iiu.edu.pk

Received: 5 November 2018; Accepted: 4 December 2018; Published: 11 December 2018



Abstract: User reviews, blogs, and social media data are widely used for various types of decision-making. In this connection, Machine Learning and Natural Language Processing techniques are employed to automate the process of opinion extraction and summarization. We have studied different techniques of opinion mining and found that the extraction of opinion target and opinion words and the relation identification between them are the main tasks of state-of-the-art techniques. Furthermore, domain-independent features extraction is still a challenging task, since it is costly to manually create an extensive list of features for every domain. In this study, we tested different syntactic patterns and semantic rules for the identification of evaluative expressions containing relevant target features and opinion. We have proposed a domain-independent framework that consists of two phases. First, we extract Best Fit Examples (BFE) consisting of short sentences and candidate phrases and in the second phase, pruning is employed to filter the candidate opinion targets and opinion words. The results of the proposed model are significant.

Keywords: sentiment analysis; advanced analytics; opinion targets

1. Introduction

The increasing trend of social media usage on the internet and mobiles has amazingly diverted the ways of business, marketing, and other day to day activities. Every moment, huge data is floated on the web and is most popularly used in tremendous ways for various applications. In fact, social media data is used by customers, businesses, manufacturers, and administrative machinery for decision-making and e-governance. Thus, opinion mining from customer reviews has become an interesting topic of research. The data sciences techniques and data-driven predictive models are most popularly employed to automate this process [1]. Opinion extraction can be carried out in two ways, i.e., one way is to obtain the overall summary of sentiment about a product and the other way is to produce a feature-based opinion summary about a product referred to as fine-grained analysis. The fine-grained analysis is also referred to as aspect-based opinion analysis. The fine-grained analysis or the aspect-based opinion analysis is in fact the most accurate and popular approach as the users can expect to get reviews about every aspect of the product [2]. It is necessary to identify opinion targets and opinion words to carry out a fine-grained opinion mining process. The primary focus of the current research work is to improve the recognition of opinion targets, opinion words, and relationships between the two from unstructured text.

To demonstrate the efficiency of the proposed model, we tested our model on real datasets of online reviews from different domains. In the current study we made comparison of our proposed

approach with state-of-the-art OM techniques. The experimental result figures confirms that the proposed approach ameliorates the performance over existing methods.

The paper is organized as follows: Section 2 details the research questions that surround aspect-based opinion analysis, Section 3 presents related work, Section 4 describes the methodology, Section 5 presents the results, and Section 6 concludes the paper.

2. Domain Independent Opinion Target Identification

Most commonly, the opinion target is delimited as the expression of the product around which opinion is stated and normally it consists of nouns or noun phrases. However, it is important that all nouns and noun phrases are not opinion targets. Furthermore, a single sentence carries multiple nouns, out of which some of the nouns may be opinion targets and some may not. Similarly, a single sentence can carry multiple opinion targets with different opinions. Hence, the identification of opinion targets is an important element of the opinion mining problem. The second subtask of opinion mining is opinion words or expressions identification, which is defined as the words that express opinion. It is also important to detect the relation between aspect and opinion. The relation represents links between opinion words and opinion targets.

Existing approaches employ a co-occurrence-based mechanism, which is described as opinion words and opinion targets that normally co-occur at a certain distance in a sentence [3]. Hence, various methods have been adopted to extract the two opinion components. A recently reported technique is a word alignment model that employs the co-ranking of opinion targets and opinion words [4].

In order to further explain the opinion mining (OM) problem, let us look into a review taken from amazon datasets about a DVD player (see Figure 1).

“I think, the apex is the best DVD player you can get for the price. Some of my friends had well-branded DVD-players but they keep complaining that sometimes it stucks and sucks! most of them are already switched to the apex and they are happy. If you are skeptical about trying apex take my words and give it a try. Hopefully, you wo n't regret. I bought this to replace an expensive (\$300+) Onkyo DVD player that quit after only 3 years. This has so many capabilities that you 'll always have room for new technologies as they evolve. I use this with a home theater system and it's amazing how it sounds, the picture clarity is unmatched!”

Figure 1. Example of a raw review taken from an Amazon dataset.

In the first sentence, the reviewer writes that apex is the best DVD player. In this sentence, the evaluative expression is the phrase “apex is the best DVD player”, which contains the target as “apex” and opinion as “best”.

The next sentence is rather a complex one. This sentence states that another DVD player has problems of stucks and sucks. In other words, the reviewer means that the apex does not have the problems of stuck and sucks. In this sentence, the phrase “there are other well-branded DVD players, but they keep complaining that some time it stuck and sucks” is an evaluative expression, with the target “other branded DVD player” and opinion “complaining of stuck and sucks”. This sentence implicitly compares the features of another player with apex player. Hence, the explicit target is that other than the apex DVD player and the implicit target is the apex DVD player. Further, the sentence states that most of the reviewer's friends have “switched to apex player and are happy”. This phrase is again an evaluative expression with the target “apex player” and opinion “happy”.

The other two sentences, consisting of “If you are skeptical about trying apex take my words and give it a try. Hopefully you won't regret”, state that if you are doubtful about trying apex, then you

should take the reviewer's surety that you will not be unhappy with it. Here, again, the target is "apex DVD player" and "will not regret" is the opinion.

The sentence "I bought this to replace an expensive (\$300+) Onkyo DVD player that quit after only 3 years" has again both an implicit and explicit nature. Implicitly, it shows a good opinion about the apex DVD player and explicitly, it shows that the Onkyo DVD player quit after only three years. It again compares apex with the Onkyo DVD player.

These examples clearly show that it is not simple to obtain opinion terms and features relation. The OM process needs to consider the semantic and context nature of the sentences. Although there have been many approaches for opinion target identification, we have observed that there are still some limitations, as explained below.

Adjectives and nouns have been potentially exploited for opinions and targets, respectively, while the relationship between them is detected by collocation [5,6] or syntactic patterns [7,8] in a limited window. A sentence may have different nouns and adjectives and there can be long-span modified relations and diverse evaluative expressions [4,9]. Similarly, reviewers may express an opinion about different products or topics in the same or a single review. Hence, simply instance-based nouns or adjectives and their nearness might not provide a particular result. In order to address this problem, several methods have exploited lexico-sequential relations that contain opinions and words. It can otherwise be stated as part of speech dependency relations. To this effect, several regular expressions have been designed to identify mutual relationships between targets and opinions [6]. However, the online reviews are usually written in an informal manner, having typographical, punctuation, and grammatical errors. Hence, machine learning tools trained on standard text corpuses might produce deficient results.

Another important issue in iterative procedures is error propagation. The erroneous extraction in preceding iteration is difficult to eliminate in subsequent iterations. For example, selecting all base noun phrases (BNPs) and adjectives may have numerous irrelevant entities and may lead to extracting irrelevant opinion targets. Therefore, it might be better to filter out redundant features in the first step.

Furthermore, supervised learning algorithms have the limitation of domain dependency and are difficult to prepare training sets for multi domain operations.

To address these challenges, we have proposed the BFE algorithm. BFE extracts the relation between opinion targets and opinion words based on grammatical and contextual information. Furthermore, BFE extracts only those sentences that have no grammatical errors. The BFE provides a list of seed target features that are employed to collectively mine opinion targets and opinion words from a candidate list. Since the BFE model is the only exactly relevant target in the first phase, error propagation is avoided. It may also be noted that BFE is domain independent.

3. Related Work

OM has been an interesting area of research and significant efforts have been taken [1,2,5,6,8]. Therefore, the research community has contributed a number of lexical dictionaries, corpuses, and tools for opinion mining. The appearance of social media on the internet has further attracted researcher attention in this field due to the availability of a huge collection of online and real-time data. The machine learning algorithms are employed for the opinion mining process in two different modes, i.e., extraction of opinion from sentences and documents. The sentence mode of the process extracts the opinion targets and opinions at the sentence level, while the document mode sees opinion targets and opinion words at the document full-text level.

Earlier works concentrated on the document level extraction of opinion words and opinion targets in pipelines [10–13], without regarding the relation between them, while in recent works, sentence level techniques are reported that are based on the relationship between opinion words and opinion targets [1,2].

In this work, we have focused on the sentence level domain independent extraction of opinion words and opinion targets based on the grammatical structure of the sentences. The contextual

and semantic relations among the synset of WordNet provide the best clues for the identification of opinions and targets [14]. A number of unsupervised learning approaches have potentially exploited syntactic and contextual patterns [12,13,15–20]. We regarded syntactic patterns for the identification of relationships between opinion targets and opinion words.

We employed regarded opinion lexicon for the detection of opinion words, while opinion targets are extracted through syntactic relations with opinion words. A number of studies have exploited basic opinion words to key out opinion aspects. Some common words, for example, good, well, right, bad, badly, awful, like, dislike, jolly, beautiful, pretty, wonderful, fantastic, impressive, amazing, and excellent, have been used as words to generate opinion lexicon for opinion detection through machine learning techniques [14,21–28]. Kobayashi et al. [19] presented an unsupervised approach for the extraction of evaluative expressions that contain target and opinion pairs. Popescu and Etzioni [29] introduced a machine learning tool referred to as OPINE. This system is established on the un-supervised machine learning paradigm to dig out the most relevant product features from the given reviews. OPINE applies syntactical patterns' semantic preference of words for the recognition of phrases containing opinion words and their polarity. Carenini, Ng et al. [25] presented an ameliorated un-supervised approach for the task of feature descent that exercises the taxonomy of the product features. Holzinger et al. [30] have proposed domain ontologies based on a more eminent set up of knowledge for the identification of product features. L. Zhuang, Jing, & Zhu [31] proposed a multi-knowledge-based approach exploiting WordNet (WordNet is a lexical dictionary), syntax rules, and the keyword list for mining feature-opinion pairs. Bloom et al. [32] efficaciously made use of the adjectival assessment aspects for target recognition. Ben-David et al. [33] presented a structural correspondence learning (SCL) algorithm for domain classification to predict new domain features based on training domain features. Ferreira, Jakob et al. [13] presented a modified Log Likelihood Ratio Test (LRT) for opinion target identification. Blitzer et al. [34] presented a novel structural correspondence learning algorithm, which take advantage of the untagged training data from the target domain to take out several relevant features that might have the capability to bring down the difference. Kessler, Eckert et al. [35] presented an annotated corpus with different linguistic features, such as mentions, co-reference, meronym, sentiment expressions, and modifiers of evaluative expressions, i.e., neutralizers, negators, and intensifiers. The authors of [15] presented a semi-supervised approach to bunch up features through a bootstrap process using lexical characteristics of terms. Lin & Chao [36] presented feature-based opinion mining model particularly pertained to hotel reviews. The authors presented model is based on supervised learning. The main objective of their work is to train classifier for touristy associated opinion mining. Goujon [10], Qu et al. [11] and Nigam [37] have exploited linguistic patterns and bag of word models for target identification and sentiment extraction.

4. Proposed Framework

Although there are a lot of OM techniques, we noticed that there are few recent methods that are particularly similar to our proposed mechanism [1,2,33]. Similar to our proposed BFE approach, Ben-David et al. [33] presented a structural correspondence learning (SCL) algorithm for domain classification to predict new domain features based on training domain features. However, their approach needs pre-knowledge of the domain and our approach exploits syntax rules for the identification of BFE. [1,2] presented an automatic extraction technique for opinion words and opinion targets based on word alignment relations. Their technique is semi-supervised topic modeling for a generalized cross domains opinion target extraction. However, it is unlike our proposed model, which is fully heuristic and does not need well-written expert training examples. Similar to this, our proposed model aligns the best fit example patterns to sentences for the prediction of relations between opinion targets and opinion words. Furthermore, we derived the method from the existing work [14,21–24] to exploit opinion lexicon and the nearest neighbor rule for candidate selection from the sentences where proposed BFE rules are not applicable.

This proceeding section reports and describes the paces of the total process of the proposed method for the identification of opinion targets, opinion words, and the relations between them in sentence-based opinion extraction.

Overview of the Proposed Framework

As explained above, we have to extract opinion targets, opinion words, and the target-opinion relations from each sentence of the document. We noted that the existing work has commonly adopted nouns/noun phrases for opinion targets and adjectives for opinion words. Additionally, for relation extraction, diverse approaches have been adopted, as explained in related work. The most crucial and challenging issue is the relation identification. Here, in this work, we considered the nouns/noun phrases for target identification and adjectives for opinion words. For relation identification, we regarded key factors that can be employed to predict target-opinion relations: contextual dependency, location information, and opinion lexicon. Considering all the three factors, we can get better results, as shown through experiments. The proposed framework employs the following analogy for the prediction of the three opinion components.

If an adjective in a sentence is likely to be an opinion word, then a noun/noun phrase modified by the adjective would be opinion target and vice versa. As explained in related work [10–13], adjectives are used for opinions and nouns for opinion targets in opinionated sentences. In this study, our proposed approach considers the adjectives and noun phrases for the subject task.

Therefore, our approach first detects opinion words in sentences and then searches for opinion targets, as explained in Algorithm 1.

Algorithm 1. Extract BFE and Candidate Opinion-Target.

```

1.  Input: Reviews
2.  Output: BFE-LIST, COT-LIST
3.  Begin
4.  Initialize: OP, BFE, BFE-LIST, COP-LIST
5.  For Each Review Ri
6.  For Each Sentence Si in Ri
7.  Find Opinion Word: OP in Si
8.  If Si Contains OP then
9.  Find BFE Pattern
10. If BFE Exits then
11. Add BFE to BFE-LIST
12. else
13. Search Candidate Pattern: COT
14. IF found COT then
15. Add COT to COT-LIST
16. End if
17. End if
18. End if
19. Next For
20. Next For
21. End

```

The steps of the proposed framework's process are given below.

1. Pre-Process Documents
2. Extract Best Fit Examples (BFEs)
3. Extract Candidate Features (CFs)
4. Prune BFEs and CFs

5. Extract Opinion-Targets

Pre-Processing

The initial step is to pre-process the input review to prepare it from the start for additional actioning. This step necessitates noise removal, to accomplish part of speech (POS) tagging task and sentence breaking. In POS tagging, a unique correct syntactic category or the POS tag is assigned to each word of a phrase, which is mandatory for pattern generation, e.g., the extraction of noun phrases, subjective expressions, etc. In this step, noise removal is also performed, which is used to remove incomplete sentences and unidentified words.

Extract Best Fit Examples

In the first step, we extract strong candidate sentences through context-dependent sequential patterns, which is referred to as BFE. The BFE looks for sentences which are grammatically correct and are opinionated, having both target and opinion words linked by a linking phrase. The BFE consists of OW, a linking phrase (LP), and an opinion target (OT), and the main task of BFE is to produce a grammatically correct pattern by the execution of rules shown in Figures 3 and 4. The opinion targets extracted through BFE are used as a seed list to extract opinions and targets from sentences exhibiting no BFE patterns with the help of semantic similarity.

In addition to this, BFE provides a seed list that is employed to extract targets and opinions from the remaining sentences of the document. A regular BFE pattern is explained below.

Rule 1. OT-{LP}-OW: This rule searches for a target and opinion linked by a linking verb, where OT means an opinion target, LP means a linking phrase, and OW means an opinion word. The example of OTLP OW is given in Figure 2.



Figure 2. Rule 1: Opinion Target Linking Phrase-Opinion Word.

Rule 2. RT-{LP}-OW-OT: This rule is slightly different, having Root Target (RT) linked by Linking Phrase (LP) with Opinion Word (OW), followed by Opinion Target (OT). In this case, the RT is the main topic and OT is the feature. The RTLPOWOT is explained in Figure 3.



Figure 3. Rule 2: Root Target Linking Phrase Opinion Word Opinion Target.

Obviously, Rule 1 and 2 extract short sentences from the text. It is also worth mentioning that these types of sentences are commonly used and have a fair chance of grammatical errors.

Extract Candidate Targets

In this step, we extract candidate target features and opinion words depending upon location and semantic context from sentences other than BFE. As discussed in related work, noun phrases and adjectives have been widely used for opinion and target extraction. However, simply selecting all BNPs and adjectives cannot provide a materialized result and has a high false positive rate.

Hence, it is important to select BNPs with semantic and contextual dependency. In this regard, various studies have presented different sequences or patterns to identify the target and opinion. We derived a combination of dependency patterns (cBNP) from [38] and employed random walk through an algorithm to extract candidate opinion-targets.

The proposed patterns that are employed for candidate extraction have the following description:

1. Entity-to-Entity and Entity-to-Feature are interlinked through the preposition “of/In”. For example, in the sentence “physical appearance of the apex compared to one previous (ad1100w) is most appreciated.”, the opinion target is the “physical appearance of the Apex”.
2. This pattern considers noun phrases with subjective adjectives, for example, “good/JJ color/NN setting/NN, funny/JJ pictures/NN”.

We experimentally evaluated the patterns as shown in the scattered graph Figure 4 and Table 1, where the specificity and sensitivity of the patterns based on the precision and recall are presented. From this graph, it is clear that the Precision and Recall of the cBNP are highly balanced compared to the other patterns.

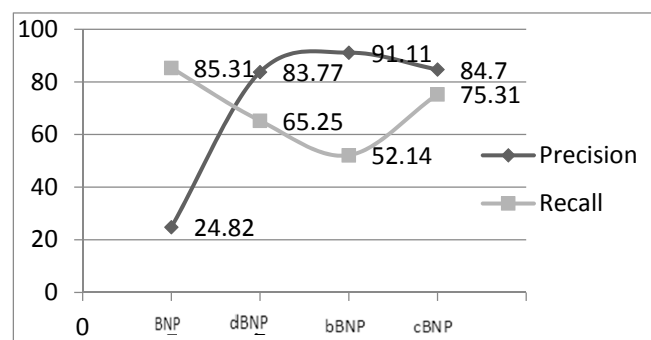


Figure 4. Sensitivity and specificity of the patterns.

Table 1. Observed Sensitivity and specificity of the various patterns.

Observed			
Method	Precision	Recall	F-Score
cBNP	84.7	75.31	79.7
dBNP	83.77	65.25	73.35
bBNP	91.11	52.14	66.16
BNP	24.82	85.31	38.45

Targets Features Pruning

In order to refine the target features from impurities, we propose two steps of pruning, i.e., BFE pruning and CF pruning. BFE pruning pertains to refining target features extracted through a best-fit examples algorithm, while CF pruning pertains to refining the candidate features. Both pruning steps depend on semantic-based relevance scoring. The relevance scoring is calculated by using WordNet Based Semantic Similarity. The WordNet dictionary offers conceptual-semantic and lexical relations. It is important to mention why the IS-A relation is proposed for semantic similarity in this work. Since this relation does not cross the part of speech boundaries, the similarity measures are limited to making judgments between noun pairs (e.g., cat and dog) and verb pairs (e.g., run and walk) [39]. As we only consider noun phrases and opinion targets, the IS-A relation is proposed. Hence, for pruning, we use the WordNet-based path length semantic similarity, as given in “1”.

$$Sim(T_i, T_j) = \frac{1}{Distance(T_i, T_j)} \quad \forall T_i \text{ and } T_j \in BNPs \quad (1)$$

where $Sim(T_i, T_j)$ is the semantic similarity, $Distance(T_i, T_j)$ is the path length from T_i to T_j , and T_i and T_j are terms.

BFE Pruning

Although BFE provides compact results, there is still a chance of impurities, like the reviewer may discuss different topics. e.g., reviews on the Amazon website sometimes discuss product features and the services of Amazon. Actually, we are concerned with product features. Thus, to exactly obtain the target and opinion about the product, there is a need to refine the BFE results before applying relevance scoring with candidate feature so that error propagation may be avoided. In this pruning step, we divide targets into groups based on WordNet-based path length similarity. Hence, a target group is generated by finding semantically similar features where the similarity score might be greater than a threshold value. Mathematically, the group function is presented in Equation (2).

$$BT = U_{bt \in BNP} \forall Sim(bt_i, bt_j) > \alpha \exists BNP \in BFE \quad (2)$$

where BT is a base term, $U_{bt \in BNP}$ represents the union of base terms, $Sim(bt_i, bt_j)$ represents the similarity between base term bt_i and bt_j , and α is a threshold value.

CF Pruning

As a matter of fact, candidate features are greater in number and have impurities. In the free text, there is no boundary for the reviewers. They can touch upon different topics in a single review. However, our proposed patterns for candidate features selection only predict the subjectivity of the phrase or sentence. Hence, the candidate features algorithm extracts all those BNPs that are modified by adjectives. Because of this, there is a greater chance of including irrelevant target features. In order to remove the irrelevant features, we propose CF pruning. This pruning step employs BT obtained in the previous step. Thus, in this step, we extract all the features from CF that have a semantic relation with BFE.

$$CT = U_{ct \in CF} \exists Sim(bt, ct) > \alpha \quad (3)$$

where CT represents candidate terms, $U_{ct \in CF}$ is the union of candidate terms, $Sim(bt, ct)$ is the similarity between the base term and candidate term, and α is a threshold value.

5. Experiments

5.1. Datasets

We tested our proposed framework on state-of-the-art datasets. The dataset used in study comprises of product review of five products. These datasets have been reported in numerous works for opinion mining and target identification. The contents of these datasets are collected from Amazon review sites and its annotation process was manually carried by [12] and are freely available for research purposes from the authors' website (<http://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html>). The description and summary of the dataset is given in Table 2.

Table 2. Description and summary of the dataset Manually Tagged Opinion Targets by [12].

Description	Product				
	Apex	Cannon	Creative	Nikon	Nokia
#Reviews	99	45	95	34	41
#Sentences	739	597	1716	346	546
#Sentences with OT and OW	344 (46%)	238 (39%)	720 (41%)	159 (45%)	265 (48%)
#Distinct BNP	779	811	1641	556	724
#OT	347	257	736	185	310
Average($\frac{\#BNP}{\#OT}$)	3.15	1.36	2.22	3.00	2.34
#Distinct OT	100	161	180	74	109
($\frac{\#Distinct OT}{\#OT}$)	0.38	0.27	0.245	0.40	0.35
#Distinct OW	164	175	238	182	261

5.2. Tools and Implementation

For experiments, we employed state-of-the-art software. For part-of-speech tagging, we used an the Stanford part of speech tagger [40] open source software available on the website (<http://nlp.stanford.edu/software/tagger.shtml>). We incorporated an open source WordNet.Net Library for similarity scoring developed by Troy Simpson, which is available from the author’s website (<http://www.ebswift.com>). We created a database of noun phrases containing similarity scores from all the datasets. Our algorithm extracts similarity scores to identify the relatedness between BFE and candidate features. For opinion word identification, we used the opinion lexicon available from the website (<https://web.stanford.edu/~jjurafsky>).

5.3. Qualitative Analysis

We first present the feasibility of our proposed method to extract opinion targets. Tables 3 and 4 show the targets annotated by a human, retrieved through BFE. In this table, the sentence no.1 is according to the OT-{LP}-OW rule, where camera is OT and perfect is the opinion word. Sentences No. 2 and 3 are also according to this rule, where “it” points to camera and thus represents the target, and small and light enough represents the opinion. Similarly, the same is true for sentence No. 4, where feature is the target and easy is the opinion word. In sentence No. 5, RT-{LP}-OW-OT has been observed, where camera is RT, very good is OW, and picture is OT.

Table 3. Sentences from manually labeled datasets.

1.	camera[+2]##this camera are perfect for an enthusiastic amateur photographer.
2.	size[+2][u]##it is small enough to fit easily in a coat pocket or purse.
3.	weight[+1][u]##it is light enough to carry around all day without bother.
4.	feature[+2][u]##operating its many features is easy and often obvious—I’m no annie lebovitz, but I was able to figure out most of its abilities just messing around with it at a camera store.
5.	camera[+2]##this camera is ideal for people who want more power, but don’t want to spend 1000s dollars on a camera.
6.	picture[+3], delay[+1]##otherwise, it takes very good pictures: shutter delay isn’t so bad either.
7.	battery life[+1]##the battery life is ok.
8.	##the store clerk concluded that the blurriness may be caused by shaking the camera slightly when i press the button, which is understandable since this camera is lighter than the other two models.
9.	picture[+3]##after nearly 800 pictures I have found that this Nikon takes incredible pictures.

Table 4. Extracted BFE.

1. Camera/NN IS/VBZ Perfect/JJ
2. IT/PRP is/VBZ Small/JJ Enough/RB
3. IT/PRP IS/VBZ LIGHT/JJ ENOUGH/RB
4. FEATURES/NNS IS/VBZ EASY/JJ AND/CC OFTEN/RB OBVIOUS/JJ
5. Camera/NN is/VBZ IDEAL/JJ for/IN PEOPLE/NNS
6. IT/PRP Takes/VBZ VERY/RB GOOD/JJ Pictures/NNS ELAY/NN Is/VBZ N'T/RB so/RB BAD/JJ
7. MODE/NN iss/VBZ also/RB WORKING/VBG GREAT/JJ
8. BATTERY-LIFE/NN IS/VBZ OK/RB
9. CAMERA/NN IS/VBZ LIGHTER/JJR THAN/IN THE/DT OTHER/JJ
10. NIKON/NN TAKES/VBZ INCREDIBLE/JJ PICTURES/NNS

5.4. Quantitative Analysis

As explained in Section 2, lexical patterns have been potentially exploited for opinion target extraction. In this study, we tested different patterns and categorized them into two different classes. The patterns showing the highest recall are proposed for extracting best fit examples of opinion targets, while the other patterns are considered as candidates. This section explains the analysis and results of the patterns and semantic terms. We carried out the experiment on the datasets given in Table 2. Table 5 represents the error rate of the proposed BFE, while Table 6 shows the influence of the semantic-based relevancy of candidate features with the BFE. Table 7 shows the precision (p), recall (R), and F-score (F) for all instances of opinion and target pairs.

Table 5. Least Square Error.

Product	MLF	BFE	Other(x)	$x-\mu$	$(3-\mu)^2$
Apex	110	25	2	−0.2	0.04
Cannon	100	20	1	−1.2	1.44
Creative	180	50	3	0.8	0.64
Nikon	74	16	3	0.8	0.64
Nokia	109	22	2	−0.2	0.04
Mean		26.6	2.2		2.8

Table 6. Influence of semantic-based feature scoring of candidate features and BFE.

Dataset	MLF	BFE	Candidate	Pruned Candidates	Total Predicted
Apex	110	25	200	82	107
Cannon	100	20	180	78	98
Creative	180	50	250	122	172
Nikon	74	16	190	53	69
Nokia	109	22	220	82	104

Table 7. Confusion Matrix.

Product	NP	Actual TF	Predicted TF	TP	FP	TN	FN	P	R	F
Apex	779	347	340	330	10	422	17	0.97	0.95	0.96
Cannon	811	257	279	240	39	515	17	0.86	0.93	0.90
Creative	1641	736	775	700	75	830	36	0.90	0.95	0.93
Nikon	556	185	200	140	60	311	45	0.70	0.76	0.73
Nokia	724	310	326	280	46	368	30	0.86	0.90	0.88

5.5. Comparative Analysis

In order to verify the performance of our proposed method, we select the following state-of-the-art relevant methods for comparison:

- HU [12] has contributed a method which extracts opinion targets and then opinion words by a nearest neighbor technique.
- HL [41] has proposed a hybrid likelihood ratio test method to detect opinion targets and opinion words using syntactic patterns and semantic relation.
- WAM [2] has used a semi-supervised model with a word alignment model based on syntactic patterns and then employed a co-ranking graph.
- OursBFE: Best Fit Example is our proposed method, where we first extract short sentences based on syntactic patterns with the highest accuracy rate of opinion targets and opinion words, which we refer to as the BFE. Then, we extract candidate opinion targets and opinion words and filter them through semantic relations with BFE. We apply pruning to improve recall. Hence, the OursBFE experiment is carried before and after pruning.

When referring to the results shown in the Table 8, it is clear that both WAM and the proposed BFE model outperform the other techniques. However, it is pertinent to note that WAM is a partially supervised technique and needs manual work, while the BFE model does not need any training materials. Significance is tested using a paired t-test with $p < 0.05$. Table 9 shows that our methods are highly statistically significant when compared with the baseline approaches. However, the proposed methods are marginally significant against the WAM method.

Table 8. Opinion targets extraction using [12] an annotation scheme.

Product	HU			HL			WAM			Ours BFE before Pruning			Ours BFE after Pruning		
	P	R	F	P	R	F	P	R	F	P	R	F	P	R	F
Apex	0.74	0.8	0.78	0.91	0.74	0.82	0.92	0.88	0.90	0.91	0.84	0.87	0.97	0.95	0.89
Cannon	0.75	0.8	0.78	0.85	0.65	0.74	0.87	0.84	0.85	0.87	0.79	0.83	0.86	0.93	0.85
Creative	0.69	0.8	0.75	0.87	0.68	0.76	0.82	0.83	0.82	0.86	0.78	0.82	0.90	0.95	0.83
Nikon	0.71	0.8	0.75	0.87	0.68	0.76	0.89	0.84	0.86	0.88	0.85	0.86	0.74	0.79	0.87
Nokia	0.72	0.8	0.74	0.89	0.74	0.81	0.9	0.9	0.90	0.89	0.84	0.86	0.86	0.90	0.89
Average			0.76			0.78			0.87			0.85			0.87

Table 9. Significance of our methods in terms of T-Pair test.

Baseline	Proposed	$P(T \leq t)$ Two-Tail	$P(T \leq t)$ One-Tail	Remarks
HU	BFE before Pruning	0.002349	0.001175	Highly Significant
	BFE after Pruning	0.001792	0.000896	Highly Significant
HL	BFE before Pruning	0.002619	0.00131	Highly Significant
	BFE after Pruning	0.000658	0.000329	Highly Significant
WAM	BFE before Pruning	0.087645	0.043823	Highly Significant
	BFE after Pruning	1	0.5	Marginal Significant

6. Conclusions

User reviews, blogs, and social media data are widely used for various types of decision-making. For effective decision-making, a suitable, smart, and reliable feature set is mandatory. However, domain-independent features extraction is a challenging task, since it is costly to manually create an extensive list of features for every domain. In this work, we have focused on the sentence level domain independent extraction of opinion words and opinion targets based on the grammatical structure of the sentences. Our goal is to propose a novel detection approach that permits the retrieval of domain independent opinions and targets from sentences. Our presented approach identifies opinions and targets without any prior training sets. The effectiveness of the proposed approach was demonstrated on a benchmark dataset by comparing it with four baseline approaches. The experiments provide evidence that the proposed approach significantly outperforms baseline

approaches. The significance of the proposed approach over the baseline approaches is proved using the standard T-Pair test. The outcomes of the test confirm that the proposed approach is highly significant over the baseline approaches, except in the case of BFE after pruning with the WAM baseline, where it is marginally significant. Based on the initial performance of the proposed approach, there are several possible extensions and improvements for it. One potential extension would be to find opinion targets in different languages, not only in English. Among these methods, we can extend the approach for implicit features extraction.

Author Contributions: Writing—original draft, K.K; Writing—review & editing, W.K.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Othman, R.; Belkaroui, R.; Faiz, R. Extracting product features for opinion mining using public conversations in twitter. *Procedia Comput. Sci.* **2017**, *112*, 927–935. [\[CrossRef\]](#)
- Liu, K.; Xu, L.; Zhao, J. Co-extracting opinion targets and opinion words from online reviews based on the word alignment model. *IEEE Trans. Knowl. Data Eng.* **2015**, *27*, 636–650. [\[CrossRef\]](#)
- Zhang, L.; Liu, B.; Lim, S.H.; O'Brien-Strain, E. Extracting and ranking product features in opinion documents. In Proceedings of the 23rd International Conference on Computational Linguistics: Posters, Beijing, China, 23–27 August 2010; Association for Computational Linguistics: Stroudsburg, PA, USA, 2010.
- Liu, K.; Xu, L.; Liu, Y.; Zhao, J. Opinion target extraction using partially-supervised word alignment model. In Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI 2013), Beijing, China, 3–9 August 2013.
- Liu, K.; Xu, L.; Zhao, J. Opinion target extraction using word-based translation model. In Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, Jeju Island, Korea, 12–14 July 2012; Association for Computational Linguistics: Stroudsburg, PA, USA, 2012.
- Khan, K.; Baharudin, B.; Khan, A.; Ullah, A. Mining opinion components from unstructured reviews: A review. *J. King Saud Univ. Comput. Inf. Sci.* **2014**, *26*, 258–275. [\[CrossRef\]](#)
- Wu, Y.; Zhang, Q.; Huang, X.; Wu, L. Phrase dependency parsing for opinion mining. In Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing, Singapore, 6–7 August 2009; Association for Computational Linguistics: Stroudsburg, PA, USA, 2009; Volume 3.
- Qiu, G.; Liu, B.; Bu, J.; Chen, C. Opinion word expansion and target extraction through double propagation. *Comput. Linguist.* **2011**, *37*, 9–27. [\[CrossRef\]](#)
- Gao, Q.; Bach, N.; Vogel, S. A semi-supervised word alignment algorithm with partial manual alignments. In Proceedings of the Joint Fifth Workshop on Statistical Machine Translation and MetricsMATR, Uppsala, Sweden, 15–16 July 2010; Association for Computational Linguistics: Stroudsburg, PA, USA, 2010.
- Goujon, B. Text Mining for Opinion Target Detection. In Proceedings of the 2011 European Intelligence and Security Informatics Conference (EISIC), Athens, Greece, 12–14 September 2011.
- Qu, L.; Ifrim, G.; Weikum, G. The bag-of-opinions method for review rating prediction from sparse text patterns. In Proceedings of the 23rd International Conference on Computational Linguistics, Beijing, China, 23–27 August 2010; Association for Computational Linguistics: Stroudsburg, PA, USA, 2010; pp. 913–921.
- Hu, M.; Liu, B. Mining and summarizing customer reviews. In Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Seattle, WA, USA, 22–25 August 2004; ACM: New York, NY, USA, 2004; pp. 168–177.
- Ferreira, L.; Jakob, N.; Gurevych, I. A comparative study of feature extraction algorithms in customer reviews. In Proceedings of the IEEE International Conference on Semantic Computing, Santa Clara, CA, USA, 4–7 August 2008.
- Pang, B.; Lee, L. Opinion mining and sentiment analysis. *Found. Trends Inf. Retr.* **2008**, *2*, 135. [\[CrossRef\]](#)
- Zhai, Z.; Liu, B.; Xu, H.; Jia, P. Clustering product features for opinion mining. In Proceedings of the 4th ACM International Conference on Web Search and Data Mining, Hong Kong, China, 9–12 February 2011; ACM: New York, NY, USA, 2011; pp. 347–354.

16. Wei, C.-P.; Chen, Y.-M.; Yang, C.-S.; Yang, C.C. Understanding what concerns consumers: A semantic approach to product features extraction from consumer reviews. *Info. Syst. E Bus. Manag.* **2010**, *8*, 149–167. [[CrossRef](#)]
17. Toprak, C.; Jakob, N.; Gurevych, I. Sentence and expression level annotation of opinions in user-generated discourse. In Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics, Uppsala, Sweden, 11–16 July 2010; Association for Computational Linguistics: Stroudsburg, PA, USA, 2010; pp. 575–584.
18. Lu, Y.; Castellanos, M.; Dayal, U.; Zhai, C. Automatic construction of a context-aware sentiment lexicon: An optimization approach. In Proceedings of the 20th International Conference on World Wide Web, Hyderabad, India, 28 March–1 April 2011; ACM: New York, NY, USA, 2011; pp. 347–356.
19. Kobayashi, N.; Inui, K.; Matsumoto, Y.; Tateishi, K.; Fukushima, T. Collecting evaluative expressions for opinion extraction. In Proceedings of the 1st International Joint Conference on Natural Language Processing, Hainan Island, China, 22–24 Marh 2004.
20. Yi, J.; Nasukawa, T.; Bunesco, R.; Niblack, W. Sentiment analyzer: Extracting sentiments about a given topic using natural language processing techniques. In Proceedings of the Third IEEE International Conference on Data Mining (ICDM), Melbourne, FL, USA, 19–22 November 2003.
21. Zhang, C.; Zuo, W.; Peng, T.; He, F. Sentiment classification for chinese reviews using machine learning methods based on string kernel. In Proceedings of the Third International Conference on Convergence and Hybrid Information Technology (ICCIT 2008), Busan, Korea, 11–13 November 2008.
22. Esuli, A. Automatic generation of lexical resources for opinion mining: Models, algorithms and applications. *ACM SIGIR Forum* **2008**, *42*, 105–106. [[CrossRef](#)]
23. Wilson, T.A. Fine-Grained Subjectivity and Sentiment Analysis: Recognizing the Intensity, Polarity, and Attitudes of Private States. Ph.D. Thesis, Faculty of School of Arts and Sciences Intelligent Systems Program, University of Pittsburgh, Pittsburgh, PA, USA, 2008.
24. Somprasertsri, G.; Lalitrojwong, P. Mining feature-opinion in online customer reviews for opinion summarization. *J. Univers. Comput. Sci.* **2010**, *16*, 938–955.
25. Carenini, G.; Ng, R.T.; Zwart, E. Extracting knowledge from evaluative text. In Proceedings of the 3rd International Conference on Knowledge Capture, Banff, AB, Canada, 2–5 October 2005; ACM: New York, NY, USA, 2005; pp. 11–18.
26. Feldman, R.; Fresco, M.; Goldenberg, J.; Netzer, O.; Unger, L. Extracting product comparisons from discussion boards. In Proceedings of the 2007 Seventh IEEE International Conference on Data Mining, Omaha, NE, USA, 28–31 October 2007; IEEE Computer Society: Washington, DC, USA; pp. 469–474.
27. Xu, K.; Liao, S.S.; Li, J.; Song, Y. Mining comparative opinions from customer reviews for Competitive Intelligence. *Decis. Support Syst.* **2011**, *50*, 743–754. [[CrossRef](#)]
28. Jindal, N.; Liu, B. Mining comparative sentences and relations. In Proceedings of the 21st National Conference on Artificial Intelligence, Boston, MA, USA, 16–20 July 2006; AAAI Press: Palo Alto, CA, USA, 2006; Volume 2, pp. 1331–1336.
29. Popescu, A.-M.; Etzioni, O. Extracting product features and opinions from reviews. In Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing, Vancouver, BC, Canada, 6–8 October 2005; Association for Computational Linguistics: Stroudsburg, PA, USA, 2005; pp. 339–346.
30. Holzinger, W.; Krüpl, B.; Herzog, M. Using ontologies for extracting product features from web pages. In Proceedings of the 5th International Semantic Web Conference (ISWC 2006), Athens, GA, USA, 5–9 November 2006.
31. Zhuang, L.; Jing, F.; Zhu, X.-Y. Movie review mining and summarization. In Proceedings of the 15th ACM International Conference on Information and Knowledge Management, Arlington, VA, USA, 6–11 November 2006; ACM: New York, NY, USA, 2006; pp. 43–50.
32. Bloom, K.; Garg, N.; Argamon, S. Extracting appraisal expressions. In Proceedings of the Human Language Technologies 2007: The Conference of the North American Chapter of the Association for Computational Linguistics, Rochester, NY, USA, 22–27 April 2007.
33. Ben-David, S.; Blitzer, J.; Crammer, K.; Pereira, F. Analysis of Representations for Domain Adaptation. In *Advances in Neural Information Processing Systems 2007*; Vancouver, British Columbia, Canada, 3–6 December 2007; ACM: New York, NY, USA, 2007; pp. 137–144.

34. Blitzer, J.; Dredze, M.; Pereira, F. Biographies, Bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification. In Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics, Prague, Czech Republic, 25–27 June 2007.
35. Kessler, J.S.; Eckert, M.; Clark, L.; Nicolov, N. The ICWSM 2010 JDPA sentiment corpus for the automotive domain. In Proceedings of the 4th International AAAI Conference on Weblogs and Social Media Data Workshop Challenge (ICWSMDWC 2010), Washington, DC, USA, 23–26 May 2010.
36. Lin, C.J.; Chao, P.H. Tourism-related opinion detection and tourist-attraction target identification. *Int. J. Comput. Linguist. Chin. Lang. Process.* **2010**, *15*, 37–60.
37. Nigam, K.; Hurst, M. Towards a Robust Metric of Opinion. In *AAAI Spring Symposium on Exploring Attitude and Affect in Text*; American Association for Artificial Intelligence: Menlo Park, CA, USA, 2004.
38. Khan, K.; Ullah, A.; Baharudin, B. Pattern and semantic analysis to improve unsupervised techniques for opinion target identification. *Kuwait J. Sci.* **2016**, *43*, 129–149.
39. Navigli, R. Word sense disambiguation: A survey. *ACM Comput. Surv.* **2009**, *41*, 1–69. [[CrossRef](#)]
40. Toutanova, K.; Klein, D.; Manning, C.D.; Singer, Y. Feature-rich part-of-speech tagging with a cyclic dependency network. In Proceedings of the 2003 Conference of the North American Association for Computational Linguistics (NAACL), Edmonton, AB, Canada, 27 May–1 June 2003; pp. 173–180.
41. Khan, K.; Baharudin, B.; Khan, A. Semantic-based unsupervised hybrid technique for opinion targets extraction from unstructured reviews. *Arabian J. Sci. Eng.* **2014**, *39*, 1–9. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).