ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 4, Special Issue 4, March 2017

A Modified Word Alignment model for extracting opinion targets and opinion words from online reviews

Anette Regina.I
Research Scholar
Ph.D Research Scholar [Part Time],
Department of Computer Science,
Muthurangam Government Arts College
Vellore – 2

Dr. Sengottuvelan. P
Research Supervisor
Associate Professor
Department of Computer Science
Periyar University P.G Extension Centre
Dharmapuri-5

***Abstract*— Data mining - an analytical process designed to explore data in which the opinion mining deals with the computational treatment of opinion, sentiment and subjective in text. The main application of opinion mining is collecting the online reviews about the product, social networks informal text. The research problem is extracting the opinion targets and the opinion words and detecting the opinion relations among the words. A novel approach based on the partially supervised alignment model for identifying the opinion relations as an alignment process have been proposed to satisfy the long span relations. To precisely mine the opinion relations among words, the Word Alignment Model (WAM) is used and to progress the error propagation, the graph based co-ranking algorithm is motivated. By comparing with the syntax based method, the word alignment model effectively reduces the parsing errors and the co ranking algorithm decreases the error probability. The datasets CRD, COAE 2008 and Large are used in various methods.**

I. INTRODUCTION (*HEADING 1*)

Data mining is the process of collecting, searching through, and analyzing a large amount of data in a database, as to discover patterns or relationships. A series of challenges have emerged in data mining and in that one of the major challenges is opinion mining. Opinion mining is the field of study that analyses the people opinions, sentiments, appraisals and emotion towards the entities such as products, services. The main objective is to gather the opinion about the products from the online review websites. The emergence of user-generated content via social media had an undeniable impact on the commercial environment. In fact, social media has shifted the content publishing from business towards the customer. With the explosive growth of social media for like microblogs, amazon, flipkart. On the web, individuals and organizations are increasingly using the content in these media for decision making. Each site typically contains a huge volume of opinion text.

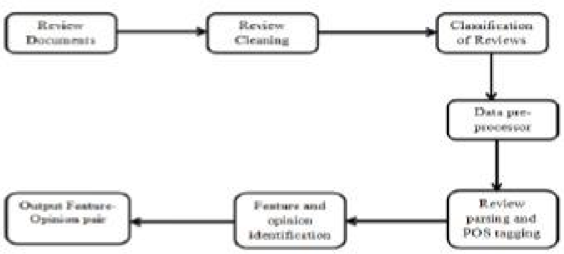
The average human reader will have difficulty in identifying the relevant sites and extracting and summarizing the opinions in them. So automated sentiment analysis systems are needed. In general, sentiment analysis has been classified at three levels. First level is document level, classifies whether a whole opinion document expresses a positive or negative opinion about the product. Second level is sentence level, classifies whether each sentence express a positive, Negative or neutral opinion. Third level is aspect level, performs a fine grained classification of an opinion about the product. In opinion mining, the fundamental subtasks are extracting or noun hrases defined as the object about which user express their opinions. Opinion word is a verb or adjectives used to express users' opinion about the object. For example: "This phone has an amazing and big screen "Here, the customers are expect to know whether this review express the positive opinion or negative opinion about the phone. To achieve this aim, the extraction of opinion word and opinion target should be detected. After that, an opinion target list and an opinion word list should be extracted. In above example, the "screen" is the opinion target and the "amazing", "big" are opinion words for that particular review [1].After the extraction, the next step is to provide the relation among those words [1]. For this process, the graph co ranking algorithm [13] is used and the opinion relation graph is constructed to provide the relations among them.

Recently, a number of online shopping customers have dramatically increased due to the rapid growth of e-commerce,

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 4, Special Issue 4, March 2017

and the increase of online merchants. To enhance the customer satisfaction, merchants and product manufacturers allow customers to review or express their opinions on the products or services. The customers can now post a review of products at merchant sites, e.g., amazon.com, cnet.com, and epinions.com. These online customer reviews, thereafter, become a cognitive source of information which is very useful for both potential customers and product manufacturers. Customers have utilized this piece of this information to support their decision on whether to purchase the product. For product manufacturer perspective, understanding the preferences of customers is highly valuable for product development, marketing and consumer relationship management. Since customer feedbacks influence other customer's decision, the review documents have become an important source of information for business organizations to take it development plans.



Among the 2 main types of textual information - facts and opinions, a major portion of current information processes methods such as web search and text mining work with the former. Opinion Mining refers to the broad area of natural language processing, computational linguistics and text mining involving the computational study of opinions, sentiments and emotions expressed in text. A thought, view, or attitude based on emotion instead of reason is often referred to as a sentiment. Hence, an alternate term for Opinion Mining, namely Sentiment Analysis. This field ends critical use in areas where organizations or individuals wish to know the general sentiment associated to a particular entity - be it a product, person, public policy, movie or even an institution. Opinion mining has many application domains including science and technology, entertainment, education, politics, marketing, accounting, law, research and development. In earlier days, with limited access to user generated opinions, research in this field was minimal. But with the tremendous growth of the World Wide Web, huge volumes of opinionated texts in the form of blogs, reviews, discussion groups and forums are available for analysis making the World Wide Web the fastest, most comprehensive and easily accessible medium for sentiment analysis. However, finding opinion sources and monitoring them over the Web can be a formidable task because a large number of diverse sources exist on the Web and each source also contains a huge volume of information. From a human's perspective, it is both difficult and tiresome to find relevant sources, extract pertinent sentences, read them, summarize them and organize them into usable form. An automated and faster opinion mining and summarizing system is thus needed.

II. RELATED WORK

M. Hu and B. Liu (2007) have proposed a sentiment based classification. The main objective is identifying the opinion sentence from reviews and deciding whether each opinion sentence is positive or negative and summarizing the results [2]. This method extracts the opinion sentences from review.F. Li, S. J. Pan, O. Jin, Q. Yang, and X. Zhu (2012) have proposed a Relational Adaptive bootstrapping (RAP) algorithm [3]. The objective is extracting the sentiment word from the text and generating the seed. This model precisely generates only the seed word (opinion target).

L. Zhang, B. Liu and S. H. Lim (2010) have proposed the Syntax based method to capturing the relation and ranking the product [4]. This method is effectively provides the relations among words for formal text.K. Liu, L. Xu, and J. Zhao (2012) have proposed the Word based translation model (WTM). The main objective is extracting opinion targets in document level from the reviews [5]. This method is precisely mine only the opinion targets.

Z. Liu, X. Chen, and M. Sun (2011) have introduced a Word trigger method (WTM) to suggest tags according to the text description of a resource [6]. By considering both the description and tags of a given resource as summaries. This method provides the WTM model for summarizing The tags and description of the text. Q. Gao, N. Bach, and S. Vogel (2011) have proposed a Constrained hill-climbing algorithm [7].

The main objective is extracting the opinion targets and providing high precision and low recall. They used precision and recall is used as an evaluation metrics.B. Wang and H. Wang (2008) have proposed an Iterative Learning Method [8]. The task of identifying product features with opinion words and learning opinion words through features alternately and iteratively. This model extracts only the opinion words.

T. Ma and X. Wan (2010) have used a Focused Concepts model. The main purpose is extracting explicit and implicit opinion targets from news comments [9]. It extracts the implicit and explicit opinion targets. F. Li, C. Han, M. Huang, X. Zhu, Y. Xia, S. Zhang, and H. Yu (2010) have described a Structure Aware Model Conditional Random Fields [10]. The process of summarizing the review based on document level extraction and extracts positive opinions, negative opinions and object features for review sentences. This model based on document level extraction. A.-M. Popescu and O. Etzioni (2007) have proposed a Word Semantic Orientation [11]. The main objective is identifying product features and determines

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 4, Special Issue 4, March 2017

the polarity of opinions. The datasets CRD and Large are used. Even though, several methods are proposed for the extraction of opinion word and opinion target from online reviews have some problems. In order to improve the precision and recall evaluation metric, the Word alignment Model (WAM) and Graph Co-Ranking algorithms are suggested with some other features.

Extensive research has been done on sentiment analysis of review text and subjectivity analysis (determining whether a sentence is subjective or objective). Another related area is feature/topic-based sentiment analysis, in which opinions on particular attributes of a product are determined. Most of this work concentrates on finding the sentiment associated with a sentence (and in some cases, the entire review). There has also been some research on automatically extracting product features from review text. Though there has been some work in review summarization, and assigning summary scores to products based on customer reviews, there has been relatively little work on ranking products using customer reviews.

A. *EXISTING SYSTEM*

Existing Systems on feature-based opinion mining have applied various methods for feature extraction and refinement, including NLP and statistical methods. However, these analyses revealed two main problems. First, most systems select the feature from a sentence by considering only information about the term itself, for example, term frequency, not bothering to consider the relationship between the term and the related opinion phrases in the sentence. As a result, there is a high probability that the wrong terms will be chosen as features. Second, words like ‘photo,’ ‘picture,’ and ‘image’ that have the same or similar meanings are treated as different features since most methods only employ surface or grammatical analysis for feature differentiation. This results in the extraction of too many features from the review data, often causing incorrect opinion analysis and providing an inappropriate summary of the review analysis.

Level of Opinion Mining

The opinion mining tasks at hand can be broadly classified based on the level at which it is done with the various levels being namely,

- The document level,
- The sentence level and
- The feature level.

At the document level, sentiment classification of documents into positive, negative, and neutral polarities is done with the assumption made that each document focuses on a single object O (although this is not necessarily the case in many realistic situations such as discussion forum posts) and contains opinion from a single opinion holder. At the sentence level, identification of subjective or opinionated sentences amongst the corpus is done by classifying data into objective (Lack of opinion) and subjective or opinionated text. Subsequently, sentiment classification of the aforementioned sentences is done moving each sentence into positive, negative and neutral .

At the feature level, the various tasks that are looked at are:

- Task1: Identifying and extracting object features that have been commented on in each review/text.
- Task 2: Determining whether the opinions on the features are positive, negative or neutral.
- Task 3: Grouping feature synonyms and producing a feature-based opinion summary of multiple reviews/text.

When both F (the set of features) and W (synonym of each feature) are unknown, all three tasks need to be performed. If F is known but W is unknown, all three tasks are needed, but Task 3 is easier. It narrows down to the problem of matching discovered features with the set of given features F. When both W and F are known, only task 2 is needed.

WORD ALIGNMENT MODEL (WAM)

WAM method is based on the monolingual model, which precisely mine the opinion relations among the words.“This phone has an amazing and colorful screen “ based on WAM, the opinion word and opinion target was extracted. In the above example, “amazing” and “colorful” is the opinion target and the “screen” is an opinion word [1]. When compare to previous method syntactic patterns [3], the WAM precisely mine the words and target.

The previous nearest-neighbour [5] method precisely mines the relation for short span sentences. But WAM method precisely mine relation for both short span and long span relations. The WAM method has some following constraints [1]:

- Nouns/noun phrases should be aligned with adjectives/verbs/a null word.
- other unrelated words, such as prepositions conjunctions and adverbs should be aligned only with themselves.

Then the hill-climbing algorithm is used to perform local optimizations. For calculating the associations among the words are estimated by

$$P(w_t \backslash w_o) = \frac{\text{Count}(w_t, w_o)}{\text{Count}(w_o)}$$

Where, w t means the opinion target and wo means the opinion word, and then P (w t \w o) means the problem

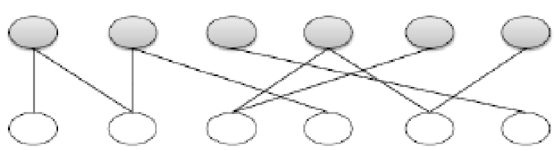
ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 4, Special Issue 4, March 2017

between these two words. The above formula was referred from [1].

GRAPH CO-RANKING ALGORITHM

After extracting the opinion word and the opinion target, the relations has been constructed by the opinion relation graph [1] was shown in fig 1. Graph co-ranking method is estimated by candidate confidence of each opinion word and opinion target and this can be constructed on the graph. The word which has higher problem will be extracted as opinion word or opinion target.



The candidate confidence can be estimated by random walking method. Here the confidence of an opinion target candidates and opinion word candidates in the iterations, then the higher confidence than the threshold are obtained as an opinion word or opinion target. The previous bootstrapping method has the error propagation problem. The graph based co-ranking algorithm effectively decreases the error problem [1].

The following features are used to represent the candidates [1]:

• Salience feature: This feature indicates the salience degree of the candidates.

• Domain relevance feature: The opinion targets are domain specific and the difference between them has different domains.

B. PROPOSED SYSTEM

In this, we can present a feature-based product ranking technique that mines various customer reviews. We first identify product features and analyze their frequencies. We model the relationships among products by using the information obtained from customer reviews, by constructing a weighted and directed graph. We mine this graph to determine relative quality of products. On the other hand, customer reviews, i.e. text describing features of the product, their comparisons and experiences of particular product provide a rich source amount of information to compare products and to make the good purchasing decisions, online retailers like Amazon.com, and flipcart.com allow us customers to add reviews of products that they have purchased. These reviews become diverse to aid the other customers. Traditionally, many customers have used expert rankings. Moreover, the product usually has multiple product features, their advantages and some drawbacks, which plays a vital role in different manner. Different customers may be interested in different features of a product, and their preferences may vary accordingly.

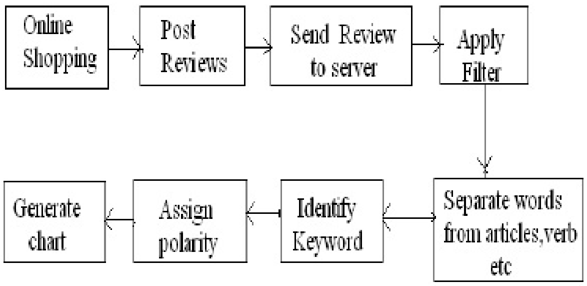
III. SYSTEM ARCHITECTURE

We select real online reviews from different domains and languages as the evaluation datasets. We compare our method to several state-of-the-art methods on opinion target/word extraction. We present the main framework of our method. As mentioned, we regard extracting opinion targets/words as a co-ranking process. We assume that all nouns/noun phrases in sentences are opinion target candidates, and all adjectives/verbs are regarded as potential opinion words, which are widely adopted by previous methods. Each candidate will be assigned a confidence, and candidates with higher confidence than a threshold are extracted as the opinion targets or opinion words. To assign a confidence to each candidate, our basic motivation is as follows. “If a word is likely to be an opinion word, the nouns/ noun phrases with which that word has a modified relation will have higher confidence as opinion target. If a noun/noun phrase is an opinion target ,the word that modifies it will be highly likely to be an opinion word ". We can see that the confidence of a candidate (opinion target or opinion word) is collectively determined by its neighbors according to the opinion associations among them. Simultaneously, each candidate may influence its neighbors. This is an iterative reinforcement process.

The fig. 1.1 says that when a particular customer does online shopping, after that according to that particular product he or she should post reviews i.e. feedback of customer about product. Those reviews may be either positive or negative. After sending the reviews, system will send reviews to the server. Server will apply filter for those review. Filter is applied to separate positive or negative review So that extraction of positive reviews and negative reviews will be done. As well as separation of words those are meaningful will be extracted. For this separation Hill climbing algorithm is used. To identify the keyword partially supervised algorithm is used and the polarity of the sentence is distinguished.

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 4, Special Issue 4, March 2017



Based on our Opinion Relation Graph, we propose a graph-based co-ranking algorithm to estimate the confidence of each candidate. Briefly, there are two important problems: 1) how to capture the opinion relations and calculate the opinion associations between opinion targets and opinion words 2) how to estimate the confidence of each candidate with graph co-ranking. For the first problem, we adopt a monolingual word alignment model to capture opinion relations in sentences A noun/noun phrase can find its modifier through word

Alignment. We additionally employ a partially-supervised word alignment model, which performs word alignment in a partially supervised framework. After that, we obtain a large number of word pairs, each of which is composed of a noun/noun phrase and its modifier. We then calculate associations between opinion target candidates and opinion word candidates as the weights on the edges. For the second problem, we exploit a random walking with restart algorithm to propagate confidence among candidates and estimate the confidence of each candidate on Opinion Relation Graph. More specifically, we penalize the high-degree vertices according to the vertices' entropies and incorporate the candidates' prior knowledge. In this way, extraction precision can be improved.

Word Alignment Model

As mentioned in the above section, we formulate opinion relation identification as a word alignment process. We employ the word-based alignment model [23] to perform monolingual word alignment, which has been widely used in many tasks such as collocation extraction [24] and tag suggestion [25]. In practice, every sentence is replicated to generate a parallel corpus. A bilingual word alignment algorithm is applied to the monolingual scenario to align a noun/noun phase (potential opinion targets) with its modifiers (potential opinion words) in sentences.

Formally, given a sentence S with n words (w1; w2; . . . ; wn), the word alignment

$$A = \{(i, a_i) \,|\, i \in [1, n], a_i \in$$

$$A^* = \underset{A}{argmax}\, P(A\,|\,S),$$

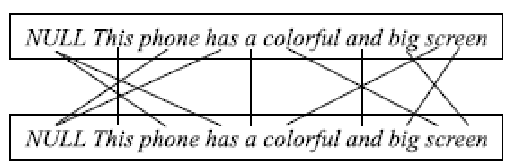
Where (i; ai) means that a noun/noun phrase at position i is aligned with its modifier at position ai. There are several word alignment models for usage, such as IBM-1, IBM-2 and IBM-3 [23]. We select IBM-3 model in our task, which has been proven to perform better than other models for our task [4]. Thus, we have

Notably, if we are to directly apply the standard alignment model to our task, an opinion target candidate (noun/noun phrase) may align with the irrelevant words rather than potential opinion words (adjectives/verbs), such as prepositions and conjunctions. Thus, we introduce some constraints in the alignment model as follows:

1) Nouns/noun phrases (adjectives/verbs) must be aligned with adjectives/verbs (nouns/noun phrases) or a null word. Aligning to a null word means that this word either has no modifier or modifies nothing;

2) Other unrelated words, such as prepositions, conjunctions and adverbs, can only align with themselves. According to these constraints, for the sentence in Fig. 1, we obtain the following alignment results shown in Fig. 4, where "NULL" means the null word. From this example,

We can see that unrelated words, such as "This", "a" and "and", are aligned with themselves. There are no opinion words to modify "Phone" and "has" modifies nothing; therefore, these two words may align with "NULL". To obtain the optimal alignments in sentences, we adopt an EM-based algorithm [23] to train the model. Specifically, for training the IBM-3 model, the simpler models (IBM-1,IBM-2 and HMM) are sequentially trained as the initial alignments for the subsequent model. Next, the hill-climbing algorithm, a greedy algorithm, is used to find a local optimal alignment.



ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 4, Special Issue 4, March 2017

$$P_{ibm3}(A \mid S) \propto \prod_{i=1}^{n} n(\phi_i \mid w_i) \prod_{j=1}^{n} t(w_j \mid w_{a_j}) d(j \mid a_j, n)$$

Input: Review sentences $S_i = \{w_1, w_2, \ldots, w_n\}$
Output: The calculated alignment $\hat{a}$ for sentences

Initialization: Calculate the seed alignment a_0 orderly using simple model (IBM-1, IBM-2, HMM)
Step 1: Optimize toward the constraints
while $N_{ill}(\hat{a}) > 0$ **do**
  if $\{a: N_{ill}(a) < N_{ill}(\hat{a})\} = \varnothing$ **then**
  break
  $\hat{a} = argmax_{a \in nb(\hat{a})} Pro(f \mid e, a)$
end
Step 2: Toward the optimal alignment under the constraint
for $i < N$ *and* $j < N$ **do**
  $M_{i,j} = -1$, if $(i, j) \notin \hat{A}$;
end
while $M_{i_1 j_1} > 1$ *or* $S_{j_1 j_2} > 1$ **do**
  If $(j_1, a_{j_2}) \notin \hat{A}$ *or* $(j_2, a_{j_1}) \notin \hat{A}$ **then**
    $S_{j_1, j_2} = -1$
  end
  $M_{i_1, j_1} = \arg\max M_{i,j}$
  $S_{j_1, j_2} = \arg\max S_{i,j}$
  If $M_{i_1, j_1} > S_{j_1, j_2}$ **then**
    Update $M_{i_1,*}, M_{j_1,*}, M_{*,i_1}, M_{*,j_1}$
    Update $S_{i_1,*}, S_{j_1,*}, S_{*,i_1}, S_{*,j_1}$
    set $\hat{a} := M_{i_1, j_1}(a)$
  end
  else
    Update $M_{i_1,*}, M_{j_2,*}, M_{*,i_1}, M_{*,j_2}$
    Update $S_{j_2,*}, S_{j_1,*}, S_{*,j_2}, S_{*,j_1}$
    set $\hat{a} := S_{j_1, j_2}(a)$
  end

As mentioned in the first section, the standard word alignment model is usually trained in a completely unsupervised manner, which may not obtain precise alignment results. Thus, to improve alignment performance, we perform a partial supervision on the statistic model and employ a partially-supervised alignment model to incorporate partial alignment links into the alignment process. Here, the partial alignment links are regarded as constraints for the trained alignment model. Formally, given the partial alignment links Eq. (1) is rewritten as follows:

$$A^* = \underset{A}{argmax}\, P(A \mid S),$$

Parameter Estimation for the PSWAM Unlike the unsupervised word alignment model, the alignments generated by the PSWAM must be as consistent as possible with the labeled partial alignments. To fulfill this aim, we adopt an EM-based algorithm. For training a simpler alignment model, such as the IBM-1 and IBM-2 models, we easily obtain all possible alignments from the observed data. Those inconsistent alignments with pre-provided partial alignment links (illegal alignments) could be filtered out; therefore, they would not be counted for parameter estimation in subsequent iterations.

Obtaining Partial Alignment Links by Using

High-Precision Syntactic Patterns for training the PSWAM, the other important issue is to obtain the partial alignment links. Naturally, we can resort to manual labeling. However, this strategy is both time consuming and impractical for multiple domains. We need an automatic method for partial alignment generation. To fulfill this aim, we resort to syntactic parsing. As mentioned in the first section, although current syntactic parsing tools cannot obtain the whole correct syntactic tree of informal sentences, some short or direct syntactic relations can be still obtained precisely. Thus, some high-precision low-recall syntactic patterns are designed to capture the opinion relations among words for initially generating the partial alignment links. These initial links are then fed into the alignment model.

Calculating the Opinion Associations among Words

From the alignment results, we obtain a set of word pairs, each of which is composed of a noun/noun phrase (opinion target candidate) and its corresponding modified word (opinion word candidate). Next, the alignment probabilities between a potential opinion target wt and a potential opinion word wo are estimated using

$$P(w_t \mid w_o) = \frac{Count(w_t, w_o)}{Count(w_o)},$$

Where P(wt/wo) means the alignment probability between these two words. Similarly, we obtain the alignment probability by changing the alignment direction in the Alignment process. Next, we use the score function in [24] and [25] to calculate the opinion association between wt and wo as follows:

$$OA(w_t, w_o) = (\alpha * P(w_t \mid w_o) + (1 - \alpha) P(w_o \mid w_t))^{-1},$$

Where a is the harmonic factor used to combine these two alignment probabilities.

Conclusion

We studied a novel method by making use of word alignment model, for co-extraction of opinion targets as well as co-extraction of opinion words. The main goal is to focus on detection of the opinion relations which are present in between opinion targets and opinion words as compared with

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 4, Special Issue 4, March 2017

previous method which is based on nearest neighbor rules and syntactic patterns, Due to the high usage of internet, the extraction of huge volume of reviews about a product from the online websites to clarify the users thought is increasing day by day and to overcome this problem, the extraction of words and targets and providing relation among these words were implemented by WAM and Graph based Co-Ranking algorithm which achieves higher precision when compared to previous methods.

References

[1]M. Hu and B. Liu, "Mining and summarizing customer reviews," in Proc. 10th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, Seattle, WA, USA, 2004, pp. 168–177.

[2]F. Li, S. J. Pan, O. Jin, Q. Yang, and X. Zhu, "Cross-domain coextraction f sentiment and topic lexicons," in Proc. 50th Annu. Meeting Assoc. Comput. Linguistics, Jeju, Korea, 2012, pp. 410419.

[3]L. Zhang, B. Liu, S. H. Lim, and E. O'Brien-Strain, "Extracting and ranking product features in opinion documents," in Proc. 23th Int. Conf. Comput. Linguistics, Beijing, China, 2010, pp. 1462–1470.

[4] M. Hu and B. Liu, "Mining opinion features in customer reviews," in Proc. 19th Nat. Conf. Artif.Intell., San Jose, CA, USA, 2004, pp. 755–760.

[5] G. Qiu, B. Liu, J. Bu, and C. Che, "Expanding domain sentiment lexicon through double propagation," in Proc. 21st Int. Jont Conf. Artif. Intell., Pasadena, CA, USA, 2009, pp. 1199–1204.

[6] X. Ding, B. Liu, and P. S. Yu, "A holistic lexicon-based approach to opinion mining," in Proc. Conf. Web Search Web Data Mining, 2008, pp. 231–240.

[7] F. Li, C. Han, M. Huang, X. Zhu, Y. Xia, S. Zhang, and H. Yu, "Structure-aware review mining and summarization." inProc.23thInt.Conf.Comput.Linguistics,Beijing,China,2010,pp.653–661.

[8] K. Liu, L. Xu, and J. Zhao, "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, Korea, July 2012, pp. 1346–1356.

[9] M. Hu and B. Liu, "Mining opinion features in customer reviews," in Proceedings of the 19th the National Conference on Artificial Intelligence (AAAI), San Jose, California, USA, 2004, pp. 755–760.

[10] A.-M. Popescu and O. Etzioni, "Extracting product features and opinions from reviews," in Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing, Vancouver, British Columbia, Canada, 2005, pp. 339– 346.

[11] G. Qiu, L. Bing, J. Bu, and C. Chen, "Opinion word expansion and target extraction through double propagation," Computational Linguistics, vol. 37, no. 1, pp. 9–27, 2011.

[12] B. Wang and H. Wang, "Bootstrapping both product features and opinion words from chinese customer reviews with cross-inducing," in Proceedings of the third International Joint Conference on Natural Language Processing, Hyderabad, India, 2008, pp. 289–295.

[13] Kang Liu, Liheng Xu, and Jun Zhao, "Co-Extracting Opinion Targets and Opinion Words from Online Reviews Based on the Word Alignment Model", IEEE Trans . Knowledge and data Engineering, Vol. 27, No. 3, March 2015

[14] M. Hu and B. Liu, "Mining and summarizing customer reviews," in Proc. 10th ACM SIGKDD Int. Conf.Knowledge. Discovery Data Mining, Seattle, WA,USA, 2004

[15] A.Mukherjee and B. Liu, "Modeling review comments," in Proc.50th Annual. Meeting Assoc.comput. linguistics, Jeju, Korea, Jul.2012

[16] K. Liu, L. Xu, and J. Zhao, "Opinion target extraction using word based translation model," In Proc. Joint Conf. Empirical Methods Natural Lang. Process.Comput. Natural Lang. Learn., Jeju, Korea, Jul.2012

[17] Y. Wu, Q. Zhang, X. Huang, and L.Wu, "Phrase dependency parsing for Opinion mining," in Proc. Conf .Empirical methods Natural. Lang. Process, Singapore,2009

[18] M. Hu and B. Liu, "Mining opinion features in customer reviews," in Proc.19th Nat. Conf. Artif. Intell., San Jose, CA, USA, 2004

[19] B. Wang and H. Wang, "Bootstrapping both product features and opinion words from Chinese customer reviews with cross inducing," in Proc.3rd Int. Joint Conf. Natural Lang. Process., Hyderabad, India, 2008

[20] Bo Pang, Lillian Lee, "Opinion Mining and Sentiment Analysis", Foundations and Trends in Information Retrieval Vol. 2, Nos. 1–2 (2008).