ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 3, Issue 5, May 2016

Extraction of Hidden Opinion Based On Sentiment Analysis Using Word Alignment Model

[1]Jayshri Vilas Borole, [2] Nilesh S. Vani
[1]M.E. Computer Engineering, [2] Assistant Professor in Computer Science Department

Abstract: **---- In opinion mining, extracting opinion mining from online reviews is quite important and tedious job. Extraction of opinion target which proposes the novel approach by using partially-supervised word alignment model.Firstly partially-supervised word alignment model is a unique scenario in sentences and estimates the relations between words for mining opinion relations. Then to increase the confidence in each candidate graph-based algorithm can be implement and for more confidence will be extracted as the opinion targets and On higher degree vertices in our graph-based algorithm, to decrease the possibility of random walk running into the unrelated region in the graph which makes penalties. To avoid parsing error during handling the informal sentences by using partially-supervised word alignment model in online reviews as compared with existing syntax-based method. On the other hand, to capture opinion relation more efficiently over partial supervision from partial alignment links when compare with existing syntax-based method. These results, that error can be avoided.**

The online market is going up day by day and new products are launching daily so based on word alignment model we are extracting the hidden sentiments in the online reviews. There are 2 parts in this paper opinion targets and opinion words. For example: The dress is good but not beautiful. Here dress is opinion target and good and beautiful are opinion words. Here we are extracting the hidden patterns based on this strategy.

***Keywords*— Opinion mining, Opinion target, Opinion words, Partially Supervised Word Alignment Model**

I. INTRODUCTION

Many researchers have been attracted towards mining opinion and analysing sentiment in online reviews[1][2][3][14]. Only the one basic problem is to extract opinion target, which are expressed by users on their opinion, typically as nouns or noun phrases. Generally user are not satisfied with just the overall sentiment condition of a product, but expect to find even minute sentiment about an aspect and feature of the product which are mentioned in the review.To fulfil this task, existing studies usually regarded opinion words as strong indicators Generally changes in opinion target opinion relation and connection between them are possible by using the strategy which is based on the observation that opinion words are used[4][6][13][15].

For example, in reviews of mobile phones we observed that there are words like "awesome" and "fantastic", so it gives good impact. If "awesome" and "fantastic" had been known to be opinion words, "design" is likely to be an opinion target in this domain.

To expand more number of opinion words we can used for extracted opinion targets. It is a mutual reinforcement procedure. For opinion target extraction, mining opinion relation in sentences and estimating relations between opinion words and opinion target are keys. At this end some heuristic pattern depend on syntactic parsing [4][5][16] are used for several method. So, the paring be prone to generate mistakes in online reviews which usually have informal writing style including grammar mistakes, typos, improper punctuation etc.

It resulted that syntax based methods are heavily depended on parsing performance which would suffer from parsing error and even don't work. Formulae identifying opinion relations between words as an alignment process to solve this problem [3]. An opinion target can find its corresponding modifier through monolingual word alignment model (WAM) without using parsing, so that the noises from Opinion Target Extraction Using Partially-Supervised Word Alignment Model [6] parsing errors can be effectively avoided. Experimental results have reported that their method have better performance than syntax-based methods, especially for large corpora. Nevertheless, we notice that WAM used in Liu's method are trained in a completely unsupervised manner, which makes the alignment quality still unsatisfactory. Even by using the supervised framework we can improve the alignment performance. Manually labeling full alignment for sentences is still time-consuming and impractical. However, in many situations, we can easily obtain a portion of links of the full alignment in a sentence. Partially supervised alignment problem can be used to constrain the alignment process. For improving the alignment performance we argue that it would be beneficial.

We propose a novel approach to come-out opinion target byusing partially-supervised word alignment model. To capture partial opinion relation (partial alignment links) in sentences we use some high-precise-low-recall syntactic pattern. Although existing syntactic parsing algorithms cannot obtain the precise whole syntactic tree Opinion Target Extraction Using Partially-Supervised Word Alignment Model of the informal sentences, we believe some short or direct dependency relations between words can be still obtained precisely. Then these extracted partial alignment links would be regarded as ground truths. And a constrained EM algorithm based on hill-climbing is performed to determine all alignments in sentences, where the model will be consistent with these links as far as possible. In this way, more correct opinion relations can be mined. Our model can not only inherit the advantages of word alignment model: in global process (word co-occurrence frequencies, word position etc.) considering multiple factors, noises from syntactic parsing error can be effectively avoiding when dealing with the informal text like online review. But by using partial supervision can improve the mining performance. Use PSWAM for better performance than traditional methods is more reasonable.

In graph based framework on the mined association, opinion target can be extract where noun phrases are regarded as opinion target candidates. A bipartite graph is constructed to model the opinion relations between words. We assume that two candidates are modified by similar opinion words, they are likely to belong to the similar category. If we have known one of them is an opinion target, the other one has high probability to be an opinion target. Thus, the opinion target confidence can propagate among vertices. A random walk algorithm can be applied to estimate the confidence of each candidate, and the candidates with higher confidence will be extracted as the opinion targets. So we can observe that other vertices put more impact prone by the higher degree vertices to collect more information. These words usually are general words and may introduce noises. For example, the opinion word “awesome”, may be used to modify multiple objects like “awesome design”, “awesome feeling” and “awesome things”. The degree of “awesome” will be high in the graph. If we have known that the “design” has higher confidence to be an opinion target, its confidence will be propagated to “feeling” and “thing” through “awesome”. As a result, “feeling” and “thing” will probably to be given higher confidence as opinion targets. It’s unreasonable. In this way, errors can be effectively avoided by Opinion Target Extraction Using Partially-Supervised Word Alignment Model. In Opinion target extraction using Partially-Supervised Word Alignment Model errors can be avoided effectively ,To resolve this problem, we make penalty on the higher-degree vertices to weaken the impacts of them and decrease the probability of the random walk running into the unrelated regions in the graph.

II. RELATED WORK

To track the mood of the public for specific product or topic, sentiment analysis is a type of natural language processing. Opinion mining is called as sentiment analysis includes building a system to collect and examine opinion about the product made in blog posts, comments, reviews or tweets. Opinion mining, which is also called sentiment analysis, involves building a system to collect and categorize opinions about a product. Automated opinion mining often uses machine learning, a type of artificial intelligence (AI), to mine text for sentiment. The success of marketers an ad campaign or new product launch, determine which version of product or service are popular and identify which demographics like or dislike particular product features evaluates in several ways.

For example, a review on a website might be broadly positive about a digital camera, but be specifically negative about how heavy it is. Being able to identify this kind of information in a systematic way gives the vendor a much clearer picture of public opinion than surveys or focus groups do, because the data is created by the customer. Versions of a product or service are popular and even identify which demographics like or dislike particular features helps in judging the success of an ad campaign or new product launch in marketing.

There are several challenges in Sentiment analysis. The first challenge, opinion word is considered to be positive in one case and might be negative in other case. For example if someone told that battery life is long it would be positive opinion. If the customer said that the laptop's start-up time was long, however, that would be is a negative opinion. In this way the opinion type can be gathering.

Other second challenge is that, People are expressing the sentiment about movies are like,” the picture was good" and other is "Picture is not good" both are oppose in their I.e. Positive and negative comments Which is somewhat manageable by analyzing sentences one at a time. So to combine different opinion in the same sentence which easy for human but difficult for computer to parse in more informal medium like twitter or blogs.

For example "That movie was as good as its last movie" this opinion is expressing the opinion thoughts of the previous model due to the lack of context peoples are difficult in understanding what someone thoughts. In tradition text mining concentrate on facts and sentiment analysis concentrate on attitude. There are few main fields of research predominate in Sentiment analysis: Entire document can be classify according to the opinion towards certain object for sentiment classification. Opinion classification is different than the traditional classification because in opinion classification is depend on the feature of the product are mined on which the customers have expressed their opinion. Opinion summarization does not summarize the reviews by selecting a subset or rewrite some of the original sentences from the reviews to capture the main points as in the classic text summarization. Word alignment is the natural language processing task of identifying translation relationships among the words (or more rarely multiword units) in a bitext, resulting in a bipartite graph between the two sides of the bitext, with an arc between two words if and only if they are translations of one another. When sentence alignment has already identified pairs of sentences that are translation of one another after that only the word alignment is typically done which is best fits a statistical machine translation model. In an instance of the expectation-maximization algorithm are resulted by circular application of these ideas.

Word alignment quality will be unsatisfactory which are trained in a completely unsupervised manner used in Liu's method although we can improve the alignment performance by using the supervised manually labeling full alignment for sentences is still time-consuming and impractical. However, in many situations, we can easily obtain a portion of links of the full alignment in a sentence. To improve the alignment performance, they can be used to constrain the alignment process, which is partially supervised alignment problem [6].

It observed that previous studies focused on opinion target extraction, such as [4][7][8][10][14], can be divided into two main categories: supervised and unsupervised methods.

In supervised approaches, the opinion target extraction task was usually regarded as a sequence labeling task [8][9][10][11]. The main limitation of these methods is that labeling training data for each domain is time consuming and impracticable.

In unsupervised methods, similar to ours, most approaches regarded opinion words as the important indicators for opinion targets. [13] exploited an association rule mining algorithm and frequency information to extract frequent explicit product features in a bootstrapping process.

[4] designed some syntactic patterns to extract opinion targets. [5] proposed a Double Propagation method to expand sentiment words and opinion targets iteratively, where they also exploited syntactic relations between words. The main limitation of Qiu's method is that the patterns based on dependency parsing tree may introduce many noises for the large corpora. [2] extended Qiu's method. Besides the patterns used in Qiu's method, they adopted some other special designed patterns to increase recall.

In addition they used the HITS [12] algorithm to compute opinion target confidences to improve the precision. [3] is similar to our method; they use a completely unsupervised WTM to capture opinion relations in sentences. Then the opinion targets were extracted in a standard random walk framework where two factors were considered: opinion relevance and target importance.

Opinion Relevance [3] reflects the degree that a candidate is associated to opinion words. If an adjective has higher confidence to be an opinion word, the noun/noun phrase it modifies will have higher confidence to be an opinion target.Similarly, if a noun/noun phrase has higher confidence to be an opinion target, the adjective which modifies it will be highly possible to be an opinion word.Candidate Importance [3] reflects the salience of a candidate in the corpus.

To model these two factors, a bipartite graph [3] is constructed, the vertices of which include all nouns/noun phrases and adjectives.An edge between a noun/noun phrase and an adjective represents that there is an opinion relation between them. The weight on the edges represents the association between them, which are estimated by using word translation model.

III. PROPOSED SYSTEM ARCHITECTURE

Natural language processing (NLP), Information Retrieval (IR), Structured and unstructured data mining are used for sentiment analysis; opinion mining and subjectivity analysis are interrelated areas of research.

Unstructured text data, speech, audio and video poses important research challenges are handling by traditional method like i.e. NLP information retrieval and information came into existence.We are using word alignment model (WAM) which can capture more complex relationships also it is robust and does not need to parse informal texts. Opinion target and opinion word extraction are divided into 2 parts:

Sentence level extraction and corpus level extraction

In sentence-level extraction, the task of opinion target/ word extraction is to identify the opinion target mentions or opinion expressions in sentences. In addition, much research focused on corpus-level extraction. They did not identify the opinion target/word mentions in sentences, but aimed to extract a list of opinion targets or generate a sentiment word lexicon from texts. The Basic System Architecture Shown below:-

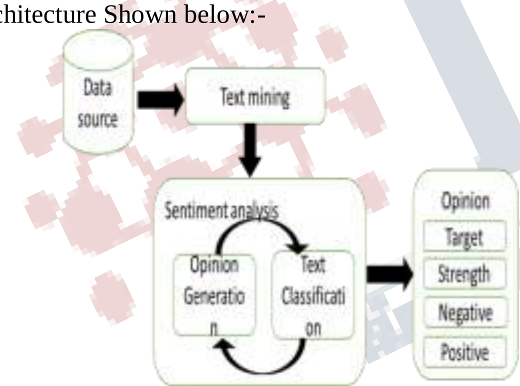

Fig. 1 Proposed System Architecture

For data collection, we can use sources like Face book, twitter, or Google. Data can be analyze text using Word Net for classifying, extracting and stemming text; can classify using extracting positive and negative opinion. We plan to consider additional types of relations between words, such as topical relations, in Opinion Relation Graph.

Recent research on big data analytics has developed and identified major analytics problems, such as sentiment analysis. While the sentiment analysis on product reviews and political debates on the Web is not a new research problem, social networking services (e.g., Twitter and Facebook) and blogospheres are now producing enormous amount of datasets to be analyzed from the sentiment analysis. The difference between the amount of datasets from the old-fashioned Web and modern social media is just one facet of the increasing difficulties in the sentiment analysis. For example, as more Internet users produce documents with a certain sentiment, more sentiment words are introduced to represent a specific sentiment.

Furthermore, the new sentiment words are either (1) new words; (2) words with completely different meanings, but that are used to show a certain sentiment in a specific domain; or (3) traditional sentiment words that have been collected and analyzed by researchers. Arising of new sentiment words is particularly rapid in the big data era, so we need a new approach to resolve this problem. While we limit the sentiment analysis to the sentiment identification area, capturing new sentiment words is one of the keys to improving the result of the analysis. Either supervised classification or unsupervised clustering approaches utilize a set of sentiment keywords, or sentiment lexicon.

One of the most popular sentiment lexicons is Senti Word Net. The challenge of Senti Word Net is its slow adaptability to the constantly piling documents and their embedded sentiment words. Moreover, Senti Word Net points out which words show which sentiment without regarding the context of the words. These domain specific sentiment words turn out to be a key determinant in identifying a sentiment from the past experiments. A solution to the inadaptability and context awareness is a dynamic generation of sentiment lexicons based upon a specific domain. This dynamic generation fundamentally relies on how to capture new sentiment words.

Based on the above discussions, the proposed co-extraction algorithm can be concluded in Figure 2. Firstly, collected data are analyzed and classified using natural language processing. After classifying text into opinion target and opinion word we extract target and word and also stemming them by word net. Then by using text classification we can generate hidden opinion. Finally, negative opinion and positive opinion generated. We can find the strength of that opinion for easily understand the sentiment of user.

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 3, Issue 5, May 2016

IV.PROBLEM DESCRIPTION AND IMPLEMENTATION

A. Data Collection:-Database Contents:- As of November 2012 Word Net's latest Online-version is 3.1. The database contains:

1. 155,287 words organized in 117,659 synsets for a
2. Total of 206,941 word-sense pairs;
3. The size of database in compressed form; it is about 12 megabytes in size.

Word Net includes the lexical categories nouns, verbs, adjectives and adverbs but ignores prepositions, determiners and other function words. Words from the same lexical category that are roughly synonymous are grouped into synsets. Synsets include simplex words as well as collocations like "eat out" and "car pool." The different senses of a polysemous word form are assigned to different synsets. The meaning of a synset is further clarified with a short defining gloss and one or more usage examples.

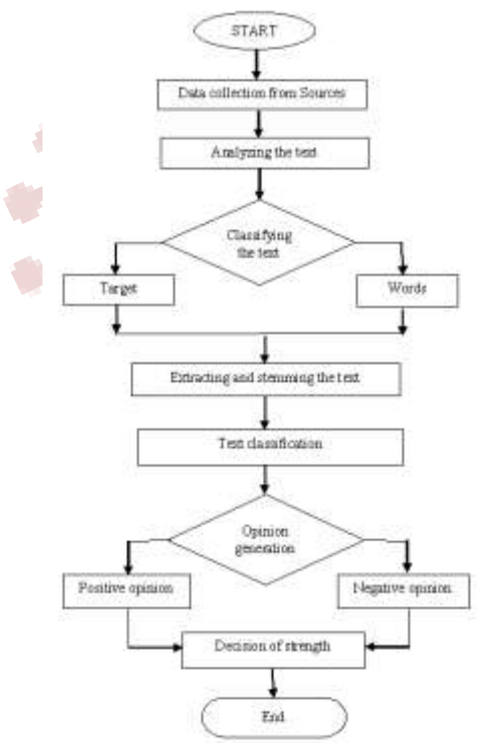

Fig. 2 Proposed Co-extracting Algorithm

An example adjective synset is: good, right, and ripe – (most suitable or right for a particular purpose; "a good time to plant tomatoes"; "the right time to act"; "thetime is ripe for great sociological changes") All synsets are connected to other synsets by means of semantic relations.

In our system, database will store all the information related to registration, authentication and modification. It is an important part of the architecture as the first process that is login starts with the database. All the web services related to the SQL Server will be included in our implementation of the project. Services are like developing of web page. User registration information stores the information in user table. Comments are stored in comment database table. Products are stored in product table and finally categories are stored in category table.

B. Proposed Work Procedure:-

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 classes. At this level as well, I make the assumption that a sentence contains only one opinion which as in our previous levels is not true in many cases. An optional task is to consider clauses.

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.

C. Sentiment Analysis:-

In this algorithm we are checking sentiments and also taking score of it for checking confidence of the customer. Sentiment Analysis or Opinion Mining is the computational study of people's opinions, attitudes and emotions toward an entity. The entity can represent individuals, events or topics. These topics are most likely to be covered by reviews. The two expressions sentiment analysis or opinion mining are interchangeable. They express a mutual meaning. However, some researchers stated that opinion mining and sentiment analysis have slightly different notions. Opinion Mining extracts and analyzes people's opinion about an entity while Sentiment Analysis identifies the sentiment expressed in a text then analyzes it. Therefore, the target of sentiment analysis is to find opinions, identify the sentiments they express and then classify their polarity as shown in Fig.

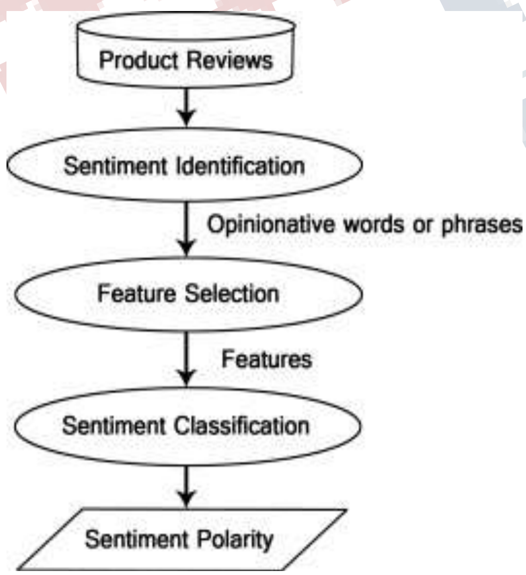

Fig.3Sentiment analysis process on product reviews.

Sentiment Analysis can be considered a classification process as illustrated in Fig. There are three main classification levels in SA: document-level, sentence-level, and aspect-level SA. Document-level SA aims to classify an opinion document as expressing a positive or negative opinion or sentiment. It considers the whole document a basic information unit (talking about one topic). Sentence-level SA aims to classify sentiment expressed in each sentence. The first step is to identify whether the sentence is subjective or objective. If the sentence is subjective, Sentence-level SA will determine whether the sentence expresses positive or negative opinions. Wilson et al. have pointed out that sentiment expressions are not necessarily subjective in nature. However, there is no fundamental difference between document and sentence level classifications because sentences are just short documents. Classifying text at the document level or at the sentence level does not provide the necessary detail needed opinions on all aspects of the entity which is needed in many applications, to obtain these details; we need to go to the aspect level. Aspect-level SA aims to classify the sentiment with respect to the specific aspects of entities. The first step is to identify the entities and their aspects. The second step is to identify or extract the features by feature identification techniques. The third step is to classify the features by sentiment classification techniques and last and fourth step is to calculate the sentiment polarity.

Feature Selection methods can be divided into lexicon-based methods that need human annotation, and statistical methods which are automatic methods that are more frequently used. Lexicon-based approaches usually begin with a small set of 'seed' words. Then they bootstrap this set through synonym detection or on-line resources to obtain a larger lexicon. This proved to have many difficulties as reported by Whitelaw et al. Statistical approaches, on the other hand, are fully automatic.

Sentiment Classification techniques can be roughly divided into machine learning approach, lexicon based approach and hybrid approach. The Machine Learning Approach applies the famous machine learning algorithms and uses linguistic features. The Lexicon-based Approach relies on a sentiment lexicon, a collection of known and precompiled sentiment terms. It is divided into dictionary-based approach and corpus-based approach which use statistical or semantic methods to find sentiment polarity. The hybrid Approach combines both approaches and is very common with sentiment lexicons playing a key role in the majority of methods.

D. Open NLP:-

In this algorithm we are separating opinions targets and words based on which we are doing further operations. Open NLP is java library to processing the natural language text. It is based on machine learning tools i.e. maximum entropy and perception. This includes pre-built models for some languages and annotated text resources. The Open NLP Library Supported the following NLP tasks:-

- ❖ Tokenization
- ❖ Sentence segmentation
- ❖ named entity extraction
- ❖ chunking
- ❖ parsing
- ❖ coreference resolution (experimental)

The library provides components to approach specific NLP tasks. They are as follows:-

- ❖ The components can be combined to build a NLP processing pipeline
- ❖ Each component interface in general has methods for
- ❖ Execute the NLP processing task on a given input text stream
- ❖ Train a model for the NLP task from examples
- ❖ Evaluate a model on test data
- ❖ The component functionalities can be accessed through a Java API or a command line interface (CLI)
- ❖ Read the model from file
- ❖ Instantiate the model
- ❖ Execute the processing task

❖ ***Tokenizer:-***

The Tokenizer segments an input character sequence into tokens such as words, punctuation, and numbers. OpenNLP has multiple Tokenizer implementations such as,

Whitespace Tokenizer: - non whitespace sequences are identified as Tokens.

Simple Tokenizer: - Sequences of the same character class are Tokens.

Learnable Tokenizer: - A maximum entropy tokenizer; uses a probability model to detect token boundaries.

❖ ***Document categorizer:-***

It performs automatic text classification into a set of predefined classes. It uses a MaxEnt classifier. A specific model must be trained for a given classification task using an annotated corpus. The default format is one document per line, starting with a string representing the category name. The categorizer expects an input segmented into sentences. Training and testing can be performed with CLI commands or by calling the API in a Java program.

❖ ***Chunker:-***

Chunker splits the text into syntactically correlated groups of words such as noun groups, verb groups.

❖ ***Parsing:-***

It uses the Penn Treebank format for parse trees with one sentence per line and it also contains also a Part of Speech tagger.

IV. RESULT AND ANALYSIS

As our contribution we propose a model which overcomes the difficulties in extracting opinion word and opinion target. We use the natural language processing (OpenNLP) to extract the opinion word and opinion target. With the help of this we find the hidden sentiment of user i.e. positive, negative or neutral.By doing this we get know the interest and opinion of user about product and we can make changes in product according to user opinion. And it is also helpful to manufacturer and retailers. According to hidden sentiment everyone has taken her/his decision quickly without wasting of time. By using NLP and Sentiment Analysis getting hidden sentiment becomes easy to admin and user. The method gives the remarkable accuracy and precision.

V. CONCLUSION

We present word alignment model based on sentiment analysis to finding hidden opinion. Both corpus and statement level extraction are used to extract opinion word and opinion target that is adopted for helping to finding the hidden opinion from online reviews. This method can easily find out the hidden sentiment of user. This method has a remarkable accuracy and precision.There are various techniques for extraction of opinion targets and opinion words. An approach for extractive summary generation from opinion targets and opinion words with word alignment model which is based on sentiment analysis this technique is focused on detecting hidden opinion of user. According to hidden sentiment we estimate the score of each candidate. This method captures opinion relations more precisely from all previous methods and therefore is more effective for opinion target and opinion word extraction.

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 taught is increasing day by day. To overcome this problem, the extraction of words and targets and providing relation among these words were followed. These processes has implemented by NLP and Sentiment Analysis based on word alignment model and achieves the higher precision when compare to previous methods.

In Future, here we are considering only text comments and in future we can add audio comments and emotions like smileys. Future plans are to recommend the goods based on user likeness so data must be thoroughly researched for the same. Artificial intelligence must be developed for that part and recommendation. Also regarding big data we are suggesting Hadoop and cloud platform.

REFERENCES

1) F. Li, S. J. Pan, O. Jin, Q. Yang, and X. Zhu, “Cross-domain coextraction of sentiment and topic lexicons,” in Proc. 50th Annu. Meeting Assoc. Comput. Linguistics, Jeju, Korea, pp. 410–419, 2012.

2) 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, pp. 1462–1470, 2010.

3) 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, pp. 1346–1356, Jul. 2012.

4) A.-M. Popescu and O. Etzioni, “Extracting product features and opinions from reviews,” in Proc. Conf. Human Lang. Technol. Empirical Methods Natural Lang. Process., Vancouver, BC, Canada, pp. 339–346, 2005.

5) G. Qiu, L. Bing, J. Bu, and C. Chen, “Opinion word expansion and target extraction through double propagation,” Comput. Linguistics, vol. 37, no. 1, pp. 9–27, 2011.

6) R. C. Moore, “A discriminative framework for bilingual word alignment,” in Proc. Conf. Human Lang. Technol. Empirical Methods Natural Lang. Process., Vancouver, BC, Canada, pp. 81–88, 2005.

7) X. Ding, B. Liu, and P. S. Yu, “A holistic lexicon-based approach to opinion mining,” in Proc. Conf. Web Search Web Data Mining, pp. 231–240, 2008.

8) F. Li, C. Han, M. Huang, X. Zhu, Y. Xia, S. Zhang, and H. Yu, “Structure-aware review mining and summarization.” in Proc. 23th Int. Conf. Comput. Linguistics, Beijing, China, pp. 653–661, 2010.

9) T. Ma and X. Wan, “Opinion target extraction in chinese news comments.” in Proc. 23th Int. Conf. Comput. Linguistics, Beijing, China, pp. 782–790, 2010.

10) Q. Zhang, Y. Wu, T. Li, M. Ogihara, J. Johnson, and X. Huang, “Mining product reviews based on shallow dependency parsing,” in Proc. 32nd Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval, Boston, MA, USA, pp. 726–727, 2009.

11) W. Jin and H. H. Huang, “A novel lexicalized HMM-based learning framework for web opinion mining,” in Proc. Int. Conf. Mach. Learn., Montreal, QC, Canada, pp. 465–472, 2009.

12) J. M. Kleinberg, “Authoritative sources in a hyperlinked environment,” J. ACM, vol. 46, no. 5, pp. 604–632, Sep. 1999.

13) Mingqin Hu and Bing Liu, “Mining opinion features in customer reviews”, in Proceedings of Conference on Artificial Intelligence (AAAI), 2004a.

14) Minqing Hu and Bing Liu, “Mining and summarizing customer reviews”, in Proceedings of

the tenth ACM SIGKDD international conference on Knowledge discovery and data mining, KDD ’04, New York, NY, USA, pp. 168–37, 2004b.

15) Liu Bing, Minqing Hu, and Junsheng Cheng. Opinion observer: analyzing and comparing opinions on the web. In Allan Ellis and Tatsuya Hagino, editors, WWW, ACM, pp. 342–351, 2005.

16) GuangQiu, Bing Liu, Jiajun Bu, and Chun Che, “Expanding domain sentiment lexicon through double propagation”, 2009.