

**Mining Opinion Targets and Opinion Words Using Word Alignment Model  
with Positive, Negative Reviews**Ms. Sunita Patil<sup>1</sup>, Mrs. N. S. Patil<sup>2</sup><sup>1</sup>Department of Computer Engineering, D. Y. Patil College of Engineering, Akurdi<sup>2</sup>Department of Computer Engineering, D. Y. Patil College of Engineering, Akurdi

---

**Abstract** — *Opinion mining also called as sentiment analysis, has enticed a great deal of attention recently due to many practical applications and challenging research problems. The important and challenging task in opinion mining is to mine opinion targets and opinion words from online reviews. The key component of mining opinion targets and opinion words is to detecting the relations among the targets and words. To extract opinion targets, opinion words and identifying the relations between them as an alignment process partially-supervised word alignment model (PSWAM) is used. Then, a graph-based algorithm is used to estimate the confidence of each candidate and the candidates with higher confidence will be extracted as the opinion targets or opinion words. This model captures opinion relations more precisely, especially for long span relations as compared to previous methods based on the nearest-neighbor rules. When dealing with informal online texts, the word alignment model effectively solve the problem of parsing errors. Because of the usage of partial supervision the proposed model obtained better result as compared to unsupervised alignment model. To decrease the probability of error generation graph-based co- ranking algorithm is used when estimating candidate confidence. Sentiment analysis is used to get positive negative and neutral reviews. The manufactures can get the feedback from product reviews to improve the quality of their products in a timely fashion.*

---

**Keywords**—*Opinion Mining, Opinion Target Extraction, Opinion Words Extraction.*

**I. INTRODUCTION**

Data mining is defined as the process of extracting information from large sets of data. Text mining is one application of data mining. Text mining is the process of deriving high- quality information from text. To get valuable structured information, text data mining can be search carefully as well as systematically through a text document or resource. There are mainly two categories of textual information one is facts and another is opinions. Facts are nothing but the objective statements about entities and events in the world. Opinions are nothing but the subjective statements. People's sentiments or perceptions about the entities and events are reflects from subjective statements. Opinion mining is a type of natural language processing. Tracking the mood of the public about a particular product is the opinion mining. Opinion mining involves building a system to collect and categorize opinions about a product.

Given a set of evaluative text documents  $D$  that contain opinions (or sentiments) about an object, opinion mining intent to extract attributes and components of the object that have been commented on in each document  $d \in D$  and to determine whether the comments are positive, negative or neutral. There are three basic components of an opinion. Such as first is opinion holder: is the person or organization which can holds the opinion about an object. Second one is object: on which an opinion is expressed and last one is opinion: a view, attitude, or appraisal on an object from an opinion holder [1].

Everyday a large number of product reviews are exploding on the Web. These opinion reviews are very useful and has attracted a lot of attentions from many researchers. Customers use these reviews and obtain initial appraisal about product information and direct control of their purchase actions. Con- currently, manufacturers can obtain immediate feedback and opportunities to improve the quality of their products in a timely fashion.

To extract and analyze opinions from on-line reviews, it is disappointing to simply obtain the overall sentiment about a product. Mostly customers expect to find fine-grained sentiments about an aspect or feature of a product that is reviewed. For example: "This phone has a colorful and big screen, but its LCD resolution is very disappointing."

The reviewer gives positive opinion about the phone's screen and a negative opinion about screen's resolution. Readers expect these things, not only the reviewer's overall sentiment. So to complete this goal, it is important to identify opinion target as well as opinion words.

Opinion target is an object on which opinions are express by the users. Opinion targets are usually nouns/noun phrases. There are two opinion targets in the above example such as "screen and "LCD resolution. In previous

methods from online product reviews an opinion target list is generated. As a result, opinion targets usually are product features or attributes so called as product feature extraction [2], [3]. The words that are used to express user’s opinions are the Opinion words. There are three opinion words in the above example such as “colorful”, “big” and disappointing.

The main contribution of author is to detect opinion relations among words. Based on monolingual word alignment model (WAM), author has proposed a method to mine opinion relations among words. Through the word alignment, an opinion target can find its corresponding modifier, for example in figure 1, “colorful” and “big” are opinion words aligned with the target word “screen”. The WAM does not constrain identifying modified relations as compared to previous nearest-neighbor rules [5], [6], [7] to a limited window; so more complex relations can be capture, such as long-span modified relations.

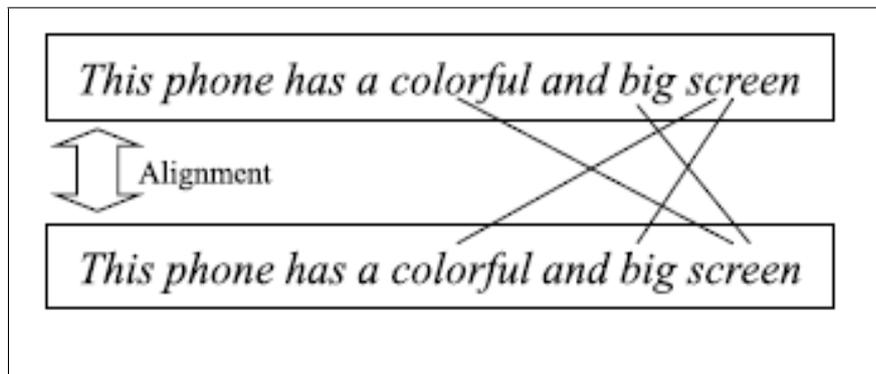


Figure1: Mining opinion relations between words using the word alignment model [1].

The WAM has does not need to parse informal texts as compared to syntactic patterns so WAM has more robust. Also word co-occurrence frequencies and word positions are intuitive factors integrating by the WAM, into a unified model for indicating the opinion relations among words. More precise results are expecting on opinion relation identification. For opinion target extraction, the alignment model used in [8] has proved to be effective. But there is still no straightforward evidence to determine the WAM’s effectiveness for opinion word extraction.

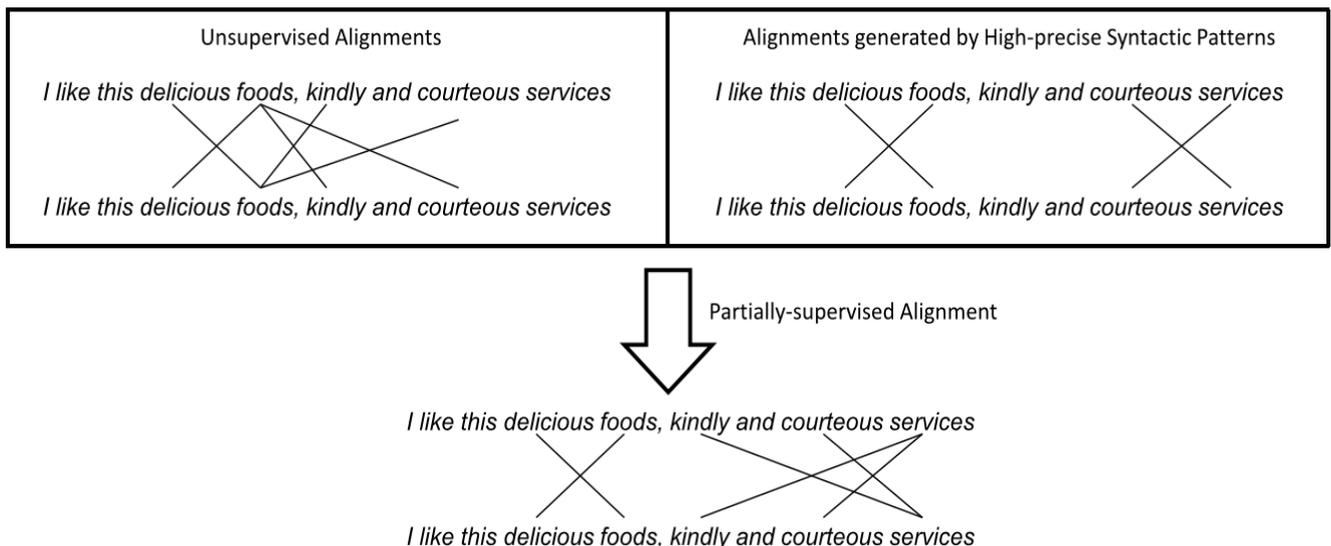


Figure2: Mining opinion relations between words using the word alignment model. [1]

The standard word alignment model is completely unsupervised. The results in alignment quality for word alignment model may be unsatisfactory. The alignment quality was improving by supervision [9]. But to manually label full alignments in sentences it is time consuming and also impractical. So to remove this problem author has exploited a partially-supervised word alignment model (PSWAM). To obtain partial alignments, author resort to syntactic parsing. To determine all of the alignments in sentences a constrained EM algorithm based on hill-climbing is used. Here the model will be consistent with as much as possible links. In this way, some errors of WAMs will be corrected. For example, in figure 2, “kindly” and “courteous” are not correctly identified as modifiers for “foods” if the WAM is performed in a wholly unsupervised manner. However, it can fact that “courteous” should be aligned to “services” by

using some syntactic patterns. By using the PSWAM, “kindly” and “courteous” are correctly linked to “services”. This model inherits the advantages of the word alignment model for opinion relation identification. To solve the problem of error propagation used graph co-ranking. Extracting opinion targets/ words is regarded as a co-ranking process. To model all opinion target/word candidates and the opinion relations among them, the Opinion Relation Graph is constructed.

For opinion word extraction and opinion target extraction there are many methods are described in related work in section II. System architecture can be described in section III. Algorithm and mathematical model are in system analysis in section IV.

## II. REVIEW OF LITERATURE

According to extraction aims, Opinion target and opinion word extraction are divided into two categories: sentence level extraction and corpus-level extraction. The task of opinion target/ word extraction is to identify the opinion target mentions or opinion expressions in sentences is comes under the sentence-level extraction. Thus, these tasks are usually regarded as sequence-labeling problems. Probably, to indicate opinion targets/words in sentences contextual words are selected as the features. To build the extractor classical sequence labeling models are used, such as CRFs [10] and HMM [11]. To perform opinion mining Jin and Huang [11] proposed a lexicalized HMM model . To extract opinion targets from reviews both [10] and [12] used CRFs. But these methods always need the labeled data to train the model. To facilitate cross domain extraction of opinion targets/words, author proposed a method [13] based on transfer learning. But this method also needed the labeled data from out-domains and the performance of extraction is heavily depended on the relevance between in-domain and out-domain.

In corpus-level extraction, aimed to extract a list of opinion targets or generate a sentiment word lexicon from texts but cannot identify the opinion target/word mentions in sentences. Most previous approaches are based on unsupervised extraction framework. Detecting opinion relations and calculating opinion associations among words are the key component of this type of method. To indicate opinion associations Wang and Wang [6] adopted the co-occurrence frequency of opinion targets and opinion words. To identify opinion relations among words Hu and Liu [5] exploited nearest-neighbor rules . But not get the precise result for the detection of opinion relations among words. Better results are obtained in [14] in which syntax information was exploited to extract opinion targets, and designed some syntactic patterns to capture the opinion relations among words. To expand sentiment words and opinion targets iteratively double Propagation method [16] and [15] exploited syntactic relations among words. But in this method, the patterns based on the dependency parsing tree opinion relations among words. To expand sentiment words and opinion targets iteratively double Propagation method [16] and [15] exploited syntactic relations among words. But in this method, the pattern based on the dependency parsing tree could not cover all opinion relations. This limitation can overcome by Zhang et al. [17]. The patterns used in [15], Zhang et al. further designed specific patterns to increase recall. HITS [18] algorithm is used to compute opinion target confidences to improve precision. Liu et al. [8] focused on opinion target extraction which is based on the WAM. They used a completely unsupervised WAM to capture opinion relations in sentences. By using standard random walk frame- work opinion targets were extracted. Lius experimental results showed that the WAM was effective for extracting opinion targets. But on opinion word extraction they cannot give any evidence to demonstrate the effectiveness of the WAM.

The limitation of WAM method can be overcome by the method PSWAM. PSWAM can be describe in the section of system architecture.

## III. SYSTEM ARCHITECTURE / SYSTEM OVERVIEW

Figure 3 shows that it first take the reviews of people from any source such face book, twitter, or any internet websites. Then it extracts these reviews as opinion words and opinion target. From opinion words and opinion target identify the relations between them. After that calculate the estimation of confidence of each candidate. Extract the candidate among them which have higher confidence.

### 3.1 Word alignment model:

Opinion relation identification is defined as a word alignment process. To carry out monolingual word alignment author used the word-based alignment model. Replicated the every sentence to achieve a parallel corpus. For the monolingual scenario a bilingual word alignment algorithm is used. Because of this scenario it align a noun/noun phrase (potential opinion targets) with its modifiers (potential opinion words) in sentences.

The word alignment  $A$  is define as  $A = \{(i, a_j) \mid [1, n], a_j \in [1, n]\}$  There are several word alignment models for usage, such as IBM-1, IBM-2 and IBM-3 [19]. IBM-3 model have perform better than other models [8]. Thus,

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

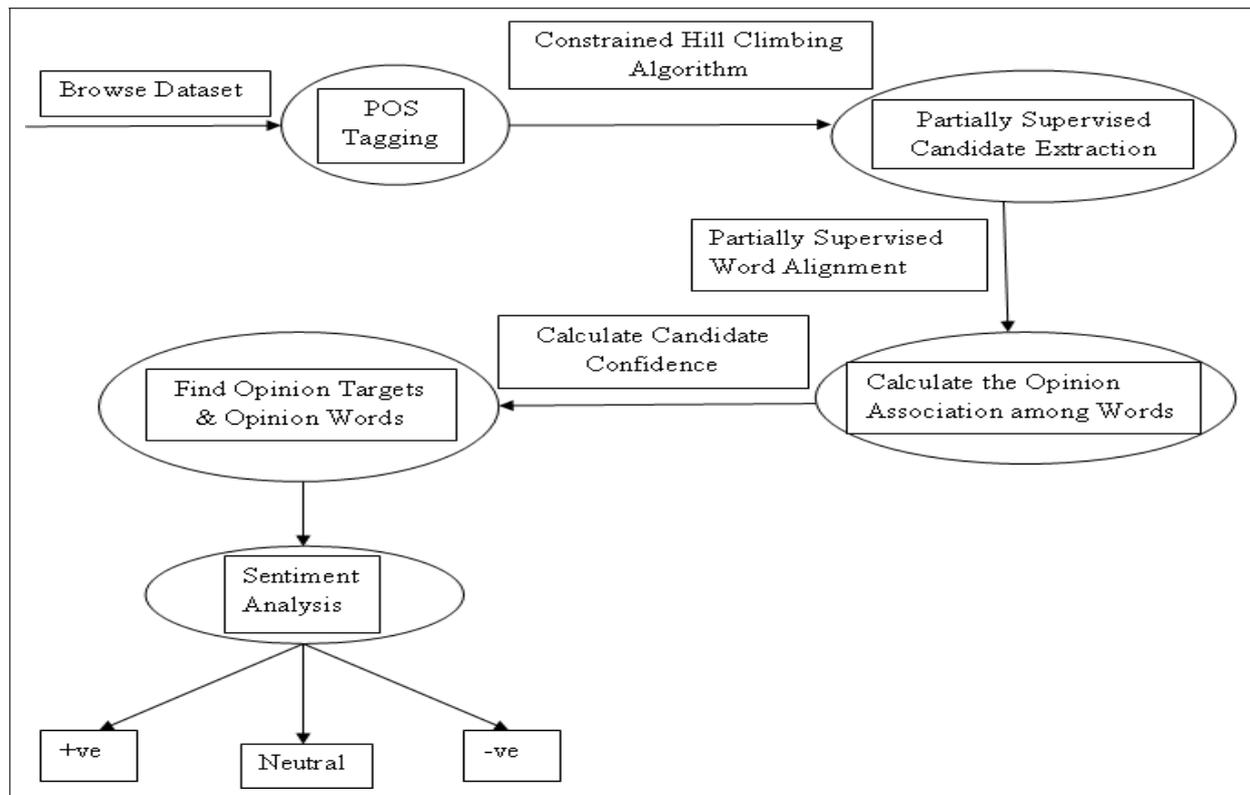


Figure 3: System Architecture.

There are mainly three factors in above model such as  $t(w_j|w_{aj}, d(j|a_j, n))$  and  $n(\phi_i|w_i)$ . To indicate the opinion relations among words these factors are used.

$t(w_j|w_{aj})$  can give concurrence of information of two words in corpora. When word is frequently modifies a noun or noun phrase, then it will have a higher value of  $t(w_j|w_{aj})$ . For example, in reviews of cell phone, big often co-occurs with “phone’s size”; therefore, “big” has high association with “phone’s size”.

$d(j|a_j, n)$  is indicate the word position information. It describes the probability that a word a word in position is aligned with a word in position j.

$n(\phi_i|w_i)$  describes the ability of a word for “one to many” relation, which means that a word can modify (or be modified by) several words.  $\phi$  specifies the number of words that are aligned with  $w_i$ . For example,

“Iphone4 has an amazing screen and software”.

In this sentence, “amazing” is used to modify two words: “screen and “software. Thus,  $\phi$  equals to 2 for “amazing. To train the model EM-based algorithm is used by which obtained the optimal alignments in sentences. To carry out sums over all possible alignments EM-based algorithm is used. So it overcome the drawback of iterations. The IBM-1, IBM-2 and HMM are used sequentially trained as the initial alignments for the training of IBM-3. After that to find a local optimal alignment the hill-climbing algorithm is used.

### 3.2 PSWAM

The standard word alignment model is usually trained in a completely unsupervised manner. Precise alignment results may not obtained by WAM. So to improve alignment performance, perform a partial supervision on the statistic model and then used a partially-supervised alignment model. To integrate partial alignment links into the alignment process partially-supervised alignment model is used. The partial alignment links are regarded as constraints for the trained alignment model. From the neighbor alignments optimal alignment are obtained. Then the neighbors of the current optimal alignment are searched. As far as new optimal alignment is found the algorithm does not end. In each step of hill-climbing algorithm, find the alignment, and update the current center alignment.

### 3.3 Partially Supervised Candidate Extraction and Word Alignment

The coarser techniques, such as frequency statistics and phrase detection are adopted to detect the proper opinion targets/words. Obtained a set of word pairs after getting the results obtained from alignment. Each pairs was composed of a noun/noun phrase which are nothing but the opinion target candidate and corresponding of its modified word i.e. opinion word candidate. After that calculate the alignment probabilities between a potential opinion target  $w_t$  and a potential opinion word  $w_o$ . Then aligned probability obtained  $P(w_o/w_t)$  by changing the alignment direction in the alignment process. After getting alignment probability, calculate the opinion association  $OA(w_t/w_o)$  between  $w_t$  and  $w_o$ .

### 3.4 Calculate Candidate Confidence

Confidence of each candidate is estimated. After calculating the opinion association, extracted the candidates with higher confidence than a threshold. The confidence of each candidate is estimated in a global process with graph co-ranking rather than that of step by step the opinion targets/words are extracted which is drawback of bootstrapping strategy. The error propagation problem is solved.

### 3.5 Opinion Target and Opinion Words

In existing system, they co-extracts opinion targets and opinion words for quick understanding. In well form it extract. But user cannot get positive opinions and negative opinions from these opinion targets and opinion words. We proposed sentiment analysis to solve this problem from this co-extracting opinion targets and opinion words. It classifies each opinion words to positive, negative and neutral. Then provide opinion results to each opinion targets.

#### A. Software Requirement Specification

##### 1) Hardware requirements

- Processor Type : Pentium IV
- Speed : 2.4 GHZ
- RAM : 256 MB
- Hard disk : 20 GB

##### 2) Software requirements

- Operating System : Windows 7
- Programming Package : Net Beans IDE 7.3.1
- Coding Language : JDK 1.7
- Database : MySql

## IV. SYSTEM ANALYSIS

### 4.1 Algorithm

The optimal alignments from neighbor alignments are got. After that next search is started in the neighbors of the current optimal alignment. Greedy algorithm is used here. Until no new optimal alignment is found, algorithm does not end. In each step of greedy algorithm, find the alignment, and update the current center alignment.

#### 4.1.1 Hill-Climbing Algorithm

In hill-climbing algorithm S and M are the swap operator and move operator respectively. i and j are the links. a is the alignment.  $f_i$  and  $f_j$  are the arbitrary set of words. If m [i,j] are the set in which changes  $a_j := i$  i.e. arbitrary set word  $f_j$  in source sentence to align to word  $f_i$  in target sentence. nba is neighbor alignment set of current center alignment a as nb(a).

---

**Algorithm 1: Constrained Hill-Climbing Algorithm**

---

**Input:** Review sentences  $S_i = \{w_1, w_2, \dots, w_n\}$

**Output:** The calculated alignment  $\hat{a}$  for sentences.

- 1) Initialization: Calculate the seed alignment  $a_0$  orderly using simple model.
- 2) Step 1: Optimize toward the constraints
- 3) while  $N_{ill}(\hat{a}) > 0$  do
- 4) if  $a : N_{ill}(a) < N_{ill}(\hat{a}) =$  then
- 5) break
- 6)  $\{\hat{a} \text{ argmax}_a \text{ enb}(\hat{a})\} \text{Pro}(f|e, a)$
- 7) end
- 8) Step 2: Toward the optimal alignment under the constraint
- 9) for  $i < N$  and  $j < N$  do
- 10)  $M_{ij} = -1, \text{ if } (i, j) \notin \hat{A};$
- 11) end
- 12) while  $M_{i_1, j_1} > 1$  or  $S_{j_1, j_2} > 1$  do
- 13) If  $(j_1, a_{j_2}) \notin \hat{A}$  or  $S_{j_2, a_{j_1}} \notin \hat{A}$  then
- 14)  $S_{j_1, j_2} = -1$
- 15) end
- 16)  $M_{i_1, j_1} = \text{argmax} M_{i, j}$
- 17)  $S_{j_1, j_2} = \text{argmax} S_{i, j}$
- 18) If  $M_{i_1, j_1} > S_{j_1, j_2}$  then
- 19) Update  $M_{i_1, *}, M_{j_1, *}, M_{*, i_1}, M_{*, j_1}$
- 20) Update  $S_{i_1, *}, S_{j_1, *}, S_{*, i_1}, S_{*, j_1}$
- 21) set  $\hat{a} := M_{i_1, j_1}(a)$
- 22) end
- 23) else
- 24) Update  $M_{i_1, *}, M_{j_2, *}, M_{*, i_1}, M_{*, j_2}$
- 25) Update  $S_{j_2, *}, S_{j_1, *}, S_{*, j_2}, S_{*, j_1}$
- 26) set  $\hat{a} := S_{i_1, j_1}(a)$
- 27) end
- 28) end
- 29) return  $\hat{a};$

**end**

#### 4.1.2 Algorithm for capturing Topical Relation

For capturing topical relation following algorithm is used. In the algorithm the topical relation is in between  $k_1$  and  $l_1$ , is optimal alignments found by hill climbing algorithm, is the threshold value for consideration of topical relation. In the algorithm  $f_{k_1, l_1}$  is the arbitrary set of words.  $\hat{A}$  is partial alignment.  $n$  is the number of words.  $w$  represents the word.

---

**Algorithm 2: Algorithm for capturing topical Relation**

---

- 1) while  $n_{ft} > 0$  do
- 2) if  $f_{k_1, l_1} \in \hat{A}$  then
- 3) do nothing
- 4) endif
- 5) else
- 6) if  $f_{k_1, l_1} = \varepsilon_w$  then
- 7)  $\hat{A} \cup f_{k_1, l_1}$
- 8) endif
- 9) endelse
- 10)  $n_{ft} = n_{ft} - 1$
- 11) end

## 4.2 Mathematical Model

The mathematical model of the system consist of input(I), processing(P) and the output(O). I= Database of customer reviews.

Processing= {P1, P11, P12, P2, P3, P4, P5}

P1: Word alignment model

$n$ : words in sentence  $S=\{w_1, w_2, \dots, w_n\}$  The word alignment  $A$  is defined as

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

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

where  $(i, a_i)$ : Noun/noun phrase at position  $i$  is aligned with its modifier at position  $a_i$ .

P11: IBM-3 model

The IBM-3 is used to identify opinion relations.

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

P12: Hill climbing algorithm is used to find a local optimal alignment.

P2: PSWAM

The partial alignment links are regarded as constraints for the trained alignment model. The partial alignment links  $(i, a_i) \mid i \in [1, n], a_i \in [1, n]$  the optimal alignment  $A^*$  is rewritten as

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

P3: Partially Supervised candidate Extraction and word alignment

Calculate the alignment probabilities between a potential opinion target  $w_t$  and a potential opinion word  $w_o$  are estimated using

$$P(w_t|w_o) = \frac{\operatorname{Count}(w_t, w)}{\operatorname{Count}(w_o)}$$

where  $P(w_t|w_o)$ : Alignment probability between these two words. The alignment probability  $P(w_o|w_t)$  by changing the alignment direction in the alignment process. Calculate the opinion association  $OA(w_t|w_o)$  between  $w_t$  and  $w_o$  as:

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

P4: Calculate candidate confidence

Calculate confidence of each candidate.

$$C_t^{k+1} = (1 - \mu) M_{to} \times C_o^k + \mu \times I_t$$

$$C_o^{k+1} = (1 - \mu) M_{to}^{to} \times C_t^k + \mu \times I_o$$

Where  $C_t^{k+1}$ : Confidence of opinion target candidate and  $C_o^{k+1}$ : Confidence of opinion word candidate, in the  $k+1$  iteration.

$C_t^k$ : Confidence of opinion target candidate and

$C_o^k$ : Confidence of opinion word candidate, in the  $k$  iteration.

$M_{to}$ : Records opinion associations among candidates.

O: Output Sentiment analysis is used to classify each opinion words to positive, negative and neutral.

## V. RESULT AND ANALYSIS

### 5.1 Dataset

Text file is given as an input to system. In the text file contained the customer reviews dataset. Reviews are about the laptop.

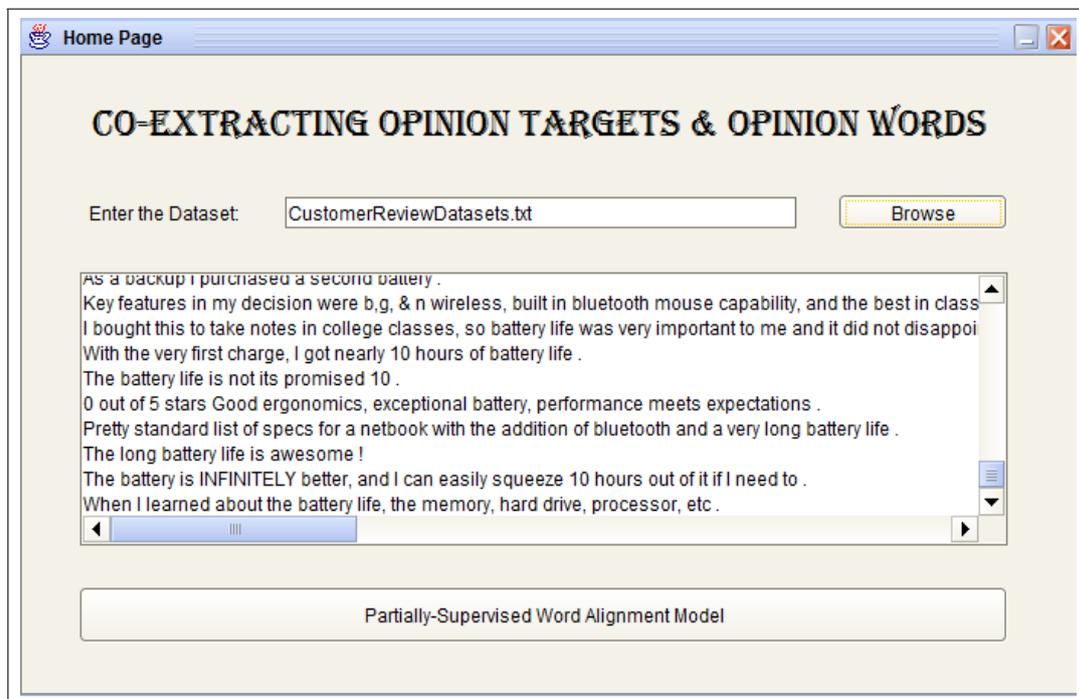


Figure 4: Customer reviews dataset

Fig. 5 shows the calculation of opinion association among the words.

Opinion Target Cand...	Opinion Words Cand...	P(wt wo)	P(wo wt)	OA(wt wo)
addition	ac	0.0625	0.2	2.5312
addition	hot	0.125	0.2	2.5625
addition	charging	0.0238	0.2	2.5119
adapter	ac	0.25	0.1538	3.375
adapter	hot	0.125	0.0385	13.0625
adapter	charging	0.0238	0.0385	13.0119
battery	ac	0.25	0.0071	70.125
battery	hot	0.375	0.0054	93.5208

Figure5: Opinion association among the words

Fig. 6 shows the co-extraction of opinion targets and opinion words by using existing method nearest neighbor rule. This method does not extract exact opinion targets and opinion words.

Opinion Targets	Opinion words
addition	ac
addition	does
adapter	get
adapter	hot
battery	is
battery	charging
Asus	mentions
heat	be
output	alarmed
battery	Realistic
numbers	are
battery	are
battery	using

Figure 6: Nearest neighbor rule

Fig. 7 shows the topical relations among the words.

Opinion Topics	Opinion Words	Opinion Word Confidence
addition	standard	1.465
adapter	snaps	7.5392
battery	protrude	161.432
Asus	mentions	8.1531
user	mentions	1.0225
manual	mentions	0.3837
heat	alarmed	0.4316
numbers	Realistic	0.4316

Figure 7: Topical Relations

Fig. 8 shows the positive, negative and neutral opinions about the customer

Opinion Topics	Opinion Words	Opinion Word Confiden...	Review	Sentiment
battery	protrude	12.0716	The battery does not protrude and is removeable .	Negative
life	extensive	5.5624	The battery life is quite extensive .	Neutral
battery	protrude	12.0716	The battery does not protrude and is removeable .	Negative
life	extensive	5.5624	The battery life is quite extensive .	Neutral
battery	protrude	12.0716	The battery does not protrude and is removeable .	Negative
life	extensive	5.5624	The battery life is quite extensive .	Neutral

Figure 8: Positive, negative and neutral opinion

Graph 9 shows the opinion words with respective their confidence. In this graph X-axis shows opinion words and Y-axis shows confidence of opinion words.

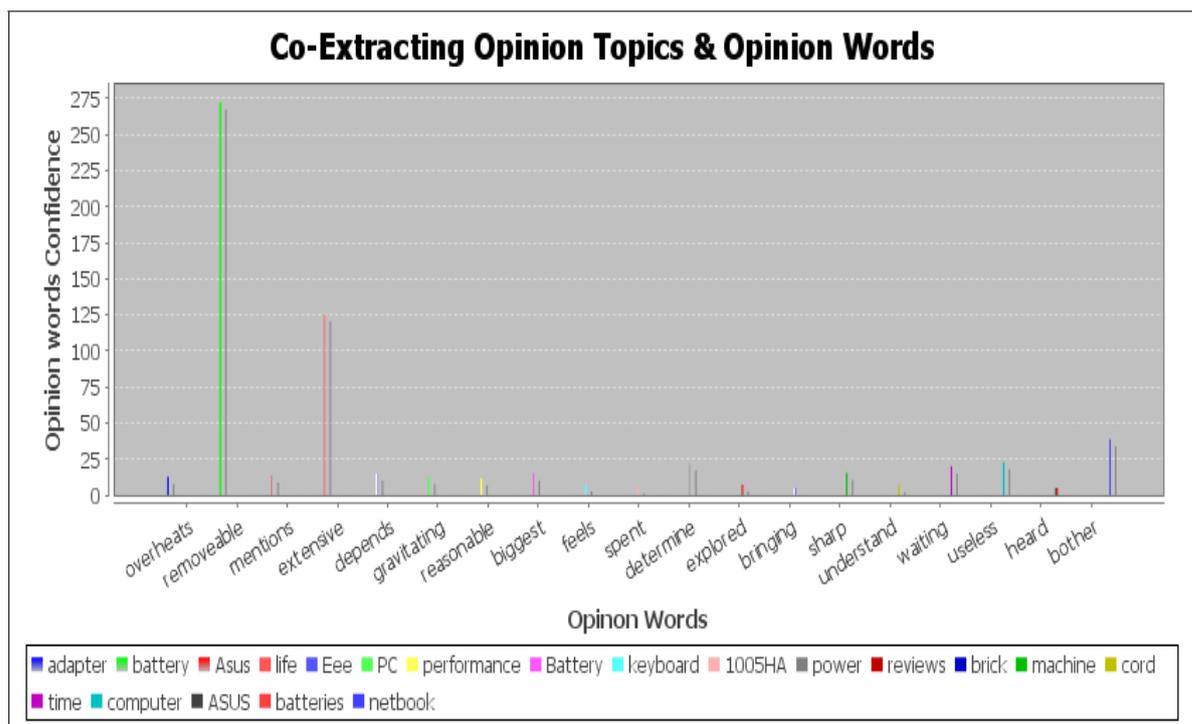


Figure 9: Co-extraction of opinion words and opinion targets

Graph 10 shows the comparison between nearest neighbor rule method and graph based method. In the graph, X-axis shows opinion targets of nearest neighbor rule method and graph based method and number of words extracted from the two methods is shown on Y-axis.

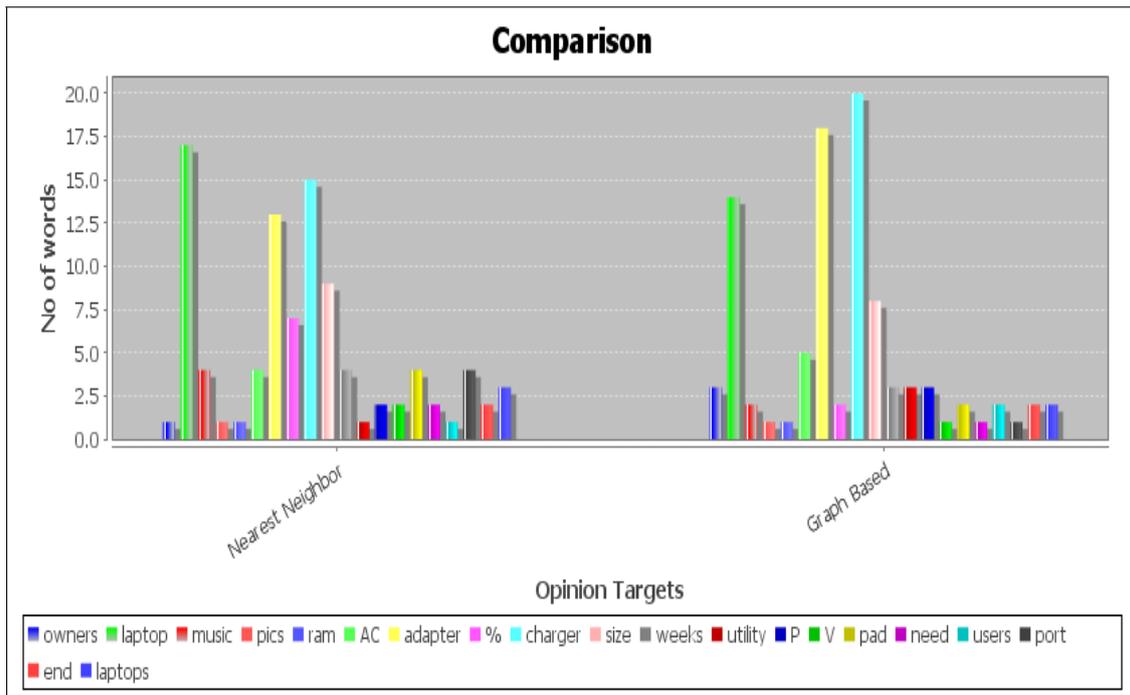


Figure 10: Comparison between nearest neighbor rule method and graph based method

## VI. CONCLUSION

A novel method word alignment model is used for co- extracting opinion targets and opinion words. To mine opinion relations among the words, monolingual word alignment model(WAM) is used. Easily obtained a portions of the links of full alignment in the sentences by using PSWAM. To estimate the confidence of each candidate a graph co-ranking algorithm is used. The items with higher ranks are extracted out. For co- extraction opinion targets and opinion words, topical relations are used. Algorithm is used for capturing topical relations. Also sentiment analysis is used. Because of this, user can get positive, negative and neutral opinions. In the future we can use the more advance techniques for extraction of opinion words and opinion targets.

## ACKNOWLEDGNT

We would like to thank the publishers, researchers for making their resources available and teachers for their guidance. We are also thankful to reviewer for their valuable suggestions. Lastly, we would also like to extend a heartfelt gratitude to friends and family members who have inspired, guided and assisted us in performing this work.

## REFERENCES

- [1] Liu, Kang, Liheng Xu, and Jun Zhao. "Co-extracting opinion targets and opinion words from online reviews based on the word alignment model." *IEEE Transactions on knowledge and data engineering* 27.3 (2015): 636-650.
- [2] Liu, Bing, and Lei Zhang. "A survey of opinion mining and sentiment analysis." *Mining text data*. Springer US, 2012. 415-463.
- [3] Hu, Mingqing, and Bing Liu. "Mining opinion features in customer reviews." *AAAI*. Vol. 4. No. 4. 2004.
- [4] Popescu, Ana-Maria, and Orena Etzioni. "Extracting product features and opinions from reviews." *Natural language processing and text mining*. Springer London, 2007. 9-28.
- [5] M. Hu and B. Liu, Mining opinion features in customer reviews, in Proc. 19th Nat. Conf. Artif. Intell., San Jose, CA, USA, 2004, pp. 755760.
- [6] B. Wang and H. Wang, Bootstrapping both product features and opinion words from chinese customer reviews with crossinducing, in Proc. 3rd Int. Joint Conf. Natural Lang. Process., Hyderabad, India, 2008, pp. 289295.

- [7] B. Liu, *Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data*, series Data-Centric Systems and Applications. New York, NY, USA: Springer, 2007.
- [8] K. Liu, L. Xu, and J. Zhao, Opinion target extraction using wordbased translation model, in Proc. Joint Conf. Empirical Methods Natural Lang. Process. Comput. Natural Lang. Learn., Jeju, Korea, Jul. 2012, pp. 13461356.
- [9] R. C. Moore, A discriminative framework for bilingual word alignment, in Proc. Conf. Human Lang. Technol. Empirical Methods Natural Lang. Process., Vancouver, BC, Canada, 2005, pp. 8188.
- [10] 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, 2010, pp. 653661.
- [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, 2009, pp. 465472.
- [12] T. Ma and X. Wan, Opinion target extraction in chinese news comments. in Proc. 23th Int. Conf. Comput. Linguistics, Beijing, China, 2010, pp. 782790.
- [13] 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, 2012, pp. 410419.
- [14] 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, 2005, pp. 339346.
- [15] 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. 11991204.
- [16] 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. 927, 2011. [17] L. Zhang, B. Liu, S. H. Lim, and E. OBrien-Strain, Extracting and ranking product features in opinion documents, in Proc. 23th Int. Conf. Comput. Linguistics, Beijing, China, 2010, pp. 14621470. [18] J. M. Kleinberg, Authoritative sources in a hyperlinked environment, *J. ACM*, vol. 46, no. 5, pp. 604632, Sep. 1999. [19] Brown, Peter F., et al. "The mathematics of statistical machine translation: Parameter estimation." *Computational linguistics* 19.2 (1993): 263311.