A Syntactic Approach for Aspect Based Opinion Mining

Chinsha T C
M Tech Student
Department of
Computer Science & Engineering
Govt. Engineering College, Thrissur
Thrissur, Kerala - 680009
Email: chinsha555@gmail.com

Shibily Joseph
Assistant Professor
Department of
Computer Science & Engineering
Govt. Engineering College, Thrissur
Thrissur, Kerala - 680009
Email: shibilyj@gmail.com

***Abstract*—Opinion mining or sentiment analysis is the process of analysing the text about a topic written in a natural language and classify them as positive negative or neutral based on the humans sentiments, emotions, opinions expressed in it. Nowadays, the opinions expressed through reviews are increasing day by day on the web. It is practically impossible to analyse and extract opinions from such huge number of reviews manually. To solve this problem an automated opinion mining approach is needed. This task of automatic opinion mining can be done mainly at three different levels, which are document level, sentence level and aspect level. Most of the previous work is in the field of document or sentence level opinion mining. This paper focus on aspect level opinion mining and propose a new syntactic based approach for it, which uses syntactic dependency, aggregate score of opinion words, SentiWordNet and aspect table together for opinion mining. The experimental work was done on restaurant reviews. The dataset of restaurant reviews was collected from web and tagged manually. The proposed method achieved total accuracy of 78.04% on the annotated test set. The method was also compared with the method, which uses Part-Of-Speech tagger for feature extraction; the obtained results show that the proposed method gives 6% more accuracy than previous one on the annotated test set.**

I. Introduction

Opinion mining or sentiment analysis is the field of study that analyses peoples opinions, sentiments, attitudes, and emotions towards entities such as products, services, organizations, individuals, issues, events and topics and which comes under the area natural language processing. We can say that opinion mining is a sub problem of traditional text classification problem because in traditional text classification the documents are clasified based on the topics, e.g., politics, sciences, and sports, but in opinion mining texts are classified based on the human sentiments, emotions or opinions.

Opinion mining represents a large problem space. There are different names with slightly different tasks, e.g., sentiment analysis, opinion mining, opinion extraction, sentiment mining, subjectivity analysis, affect analysis, emotion analysis, review mining, etc. However, they are now all under the area of sentiment analysis or opinion mining. The term opinion mining is commonly used in academia and sentiment analysis is in industry. The term sentiment analysis first appeared in [1], and the term opinion mining first appeared in [2].

Due to the explosive growth of social media (e.g., reviews, forum discussions, blogs, Twitter, comments, and postings in social network sites) on the Web, users now have many opportunities to express their opinions about a product or topic. Users express their opinion through the reviews. These reviews are used by the individuals and organizations for decision making. This trend makes attention of organizations businesses and researches around the world towards opinion mining area. This new trend provides a strong motivation to the researchers and also offers many challenging research problems.

Opinion mining is a hard problem to be solved due to the highly unstructured nature of natural language and the difficulty of a machine to interpret the meaning of a sentence. But the reviews and usefulness of the opinion from the reviews is increasing day by day. For solving this problem a system must be made to understand and interpret the human emotions and feelings. Opinion mining and sentiment analysis are approaches used for implementing the same.

Opinion mining can be done at three different levels, which are Document Level, Sentence Level and Aspect/Feature Level[3]. In document level, the overall opinion about the document is finding out and classifies them as positive or negative. In sentence level, each sentence in the document is analysed for finding the fine grained opinions about different topics in a document. Finally classify the opinion expressed in a sentence as positive, negative, or neutral. The product and restaurant reviews are a mixture of positive and negative opinion about different aspects. It needs more fine-grained analysis of reviews to mine these mixed opinions, aspect level perform this task. Hence aspect based opinion mining is preferred in this work.

This paper mainly focus on the aspect based opinion mining. The core tasks in aspect based opinion mining is aspect identification, aspect based opinion word identification and its orientation detection. For example, consider a review of a restaurant, "The environment is nice but food is bad". First step is to identify the aspects, which are environment and food, then find aspect related opinion word, which are nice and bad. Then detect its orientation, i.e. whether that opinion word expresses positive or negative opinion. By analysing the above

IEEE ICSC 2015, February 7-9, 2015, Anaheim, California, USA
978-1-4799-7935-6/15/$31.00 ©2015 IEEE

example we get that environment has positive opinion and the food has negative opinion. Current approaches for opinion mining, attempt to detect the overall polarity of a sentence, paragraph or text span regardless of the aspects mentioned in it. The work proposes a new syntactic based approach for aspect based opinion mining which uses syntactic dependency, aggregate score of opinion words, SentiWordNet and aspect table together for opinion mining process.

A. Problem Statement

If R_1, R_2, R_3..... are the reviews about a product, then find its important aspect set A and for each aspects A1, A2, A3... in A, calculate its polarity scores and determine the final polarity (positive or negative) of them.

II. Related Work

The core tasks in aspect level opinion mining is aspect identification, aspect based opinion word identification and its orientation detection. Aspect identification is one of the most complex tasks in aspect based opinion mining (also known as topic, feature, or target extraction). It needs the use of Natural language processing techniques in order to automatically extract the aspects (features) in the opinionated documents. Aspect extraction is a special problem of general information extraction problem.

One of the earliest works to extract the aspects was done in[4], which was based on frequent nouns and noun phrases. They used association rule mining combined with pruning strategies to find the candidates of features. They assumed that the product features are nouns or noun phrases. They first performed Part-of-Speech (POS) parsing. A tag of POS was then given to each word. A transaction was built from the noun words of each sentence. All transactions were fed into the association rule mining algorithm to find the frequent item sets. The returned frequent item sets were used to identify product features.

Kessler and Nicolov[5] introduced a dataset of car and camera reviews in which opinion expressions and target aspects are annotated. They train a machine learning classifier (SVM) to finding related opinion expression and target aspect. The feature vectors were formed based on the syntactic and semantic relationship between the opinion expression and candidate aspect.

In[6] a double propagation method is used. It uses a set of seed opinion words to extract aspects and new opinion words and then uses them to extract more aspects and opinion words until no new words can be extracted.

In [7] aspect extraction aims to extract fine-grained opinion targets from opinion texts. They used Logic Programming, particularly Answer Set Programming (ASP), to implement the aspect extraction task. In [8] an unsupervised method was used for aspect detection using a set of heuristic rules.

Determining the opinion orientation expressed on each aspect in a sentence is the next task in aspect based sentiment analysis. It must determine whether the sentiment orientation on each aspect is positive, negative or neutral. Different approaches used for this purpose. In [9], AAC (adverb-adjective combinations) based sentiment analysis technique was used and propose a set of general axioms that all adverb scoring techniques must satisfy. Three specific AAC scoring methods that satisfy the axioms are presented. In [10], polarity of opinions expressed on product features in reviews are find out by using a set of opinion words and its semantic orientations are find out by using some linguistic rules.

Miniqing Hu et al.,[4] performed mining and summarization process to all the customer reviews of a product. In the proposed method he product features commented by the customer in the review are mined using natural language processing and Data mining. Then opinions in the review are identified and the opinions are classified as positive or negative and summarize the results. The objective of the study was to perform feature based summary of a large number of customer reviews of a product sold online.

In [11], authors proposed a system which provides a linguistic approach to aspect based opinion mining, which is a clause-level sentiment analysis of opinionated texts. For each sentence in a message post, it generates a dependency tree, and splits the sentence into clauses. Then it determines the contextual sentiment score for each clause utilizing grammatical dependencies of words and the prior sentiment scores of the words derived from SentiWordNet and domain specific lexicons. Automatic aspect extraction not included in his study. The proposed method included the same.

In[12] a combined approach was used, which combines rule-based classification, supervised learning and machine learning into a new combined method. In [13], they generate a reliable classification approach of hotel reviews based on an existing domain-specific corpus by applying a lexicon-based sentiment analysis. First, they build a domain specific lexicon then apply sentiment analysis based on the lexicon to generate a classification of hotel reviews.

In [14], authors proposed aspect based sentiment analysis using support vector machine classifier. They propose a different approach which combines the use of dependency parsing, co-reference resolution and SentiWordNet together for the sentiment analysis. The training of the system is done using the support vector machine. They did this work in a single domain and tests are done only for reviews about digital cameras. They consider only explicit aspects. Average accuracy of 77.98% is obtained. More training data is needed for this approach and it may fail when training data are insufficient. The proposed method does not use any training data. The use of emoticons in sentiment analysis is explained in[15], the results show that accuracy is improved. The emoticons are less in online reviews like restaurant reviews; hence it is not included in this study.

In [16] a new feature based heuristic was used for aspect level sentiment classification of movie reviews. The authors proposed AAAVC algorithm based on SentiWordNet to find the sentence level aspect score. This scheme calculates Adverb+Adjective and Adverb+Verb combined score using SentiWordNet. They used SentiWordNet scheme to compute

the document-level sentiment for each movie reviewed and compared the results with results obtained using Alchemy API. Accuracy of 78.7% obtained over movie review dataset. In this method, features are extracted by using POS Tagger. Firstly, aspect indicating term in a sentence is located, and then searches up to 5-gram forward or backward for the occurrence of features in it. Directly modified feature term of an aspect cannot be determined in it. For resolving it, the proposed method use dependency grammar for feature extraction.

In[17] authors extend the Bing Liu's aspect-based opinion mining technique to apply it to the tourism domain. The score of opinion on aspects is calculated based on some linguistic rules. But rule; too rule like different rules are used. They discover consumer preferences about tourism products using some statistics. In [18] a modular software is proposed named Opinion Zoom, that helps users to understand the vast amount of tourism opinions disposed all over the Web in an easy manner. Authors used the same method in [18]but it produces novel graphic summaries of opinions. They successfully implemented and tested Opinion Zoom, our proposed method use SetiWordNet and linguistic rules to find the aspect score. SentiWordNet based approach was used in [19] [20].

III. Design of Proposed Technique

The proposed design for opinion mining process is given in the Fig.1. The following section describes the proposed opinion mining model with each sequent steps in detail.

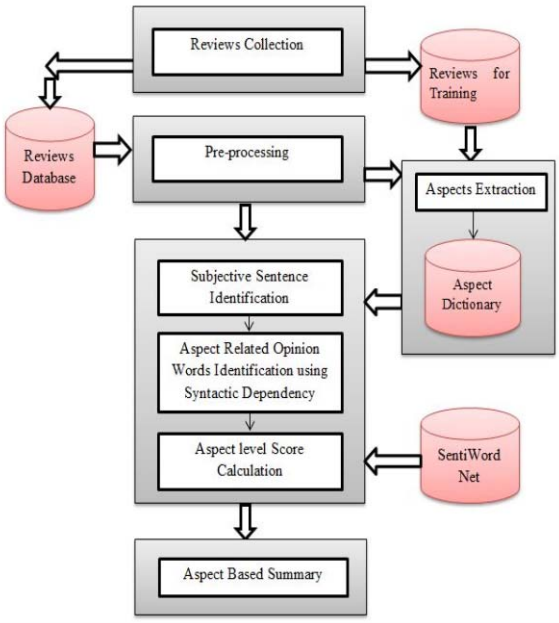

Fig. 1. Proposed Design for Opinion Mining

IV. Method

A. Reviews Collection

Input to the opinion mining process is the set of reviews. A review downloader should be developed for collecting reviews from the web instantly, it should parses the HTML webpage containing the reviews and should download the reviews.The collected reviews should store in a database which are used for Opinion mining process.

B. Aspects Extraction

Aspects are the important features rated by the reviewers. The aspects for a particular domain are identified through the training process. Aspect extraction task was implemeted based on the idea in [4]. An aspect may be a single word or a phrase. For example the service, atmosphere and quality of food are the different aspects of a restaurant. Authors in [4] says, in most cases nouns and noun phrases are the aspects. Hence in order to extract the aspect, searching of noun and noun phrases of the reviews are needed. The steps for aspect extraction task was given below. This task is performed in sentence wise for all reviews.

- Take each review and split it into the sentence and take each sentence for analysis.
- Find the Part-Of-Speech tag of the sentence.
- Extract NN, NNP, NNS etc. tagged words. This process is repeated for all sentences.
- Count Frequency of extracted words.
- Remove the unnecessary words like 'it', 'they' etc.
- Take most frequent one.
- Group synonyms and give a group label.
- Create aspect table.

An aspect table is a table of important aspects and their synonyms. Synonyms are the words having the same meaning to that of aspects. This aspect table is used for selecting the subjective sentences. Stanford POS Tagger[1] was used for tagging purpose.

C. Pre-Processing

Before starting the opinion mining process, pre-processing of reviews should be done, which improve the accuracy and also avoid the unnecessary processing overhead of opinion mining process, hence the collected reviews should be pre-processed. The pre-processing steps include stop words removal. The all unnecessary non alphabetic characters and smileys should be eliminated before doing opinion mining steps. In this way filtering or pre-processing of review sentences are to be done.

D. Subjective Sentence Identification

After pre-processing a review, the opinionated sentences in it should be identified because all the sentences in it does not contain an opinion. Such opinionated sentences are called subjective sentences, which should be analysed and other (objective) sentences should be removed. It helps for avoiding the further processing overhead. The proposed technique take a sentence as a subjective sentence only if it contains an aspect in aspect table(already created).

[1]http://nlp.stanford.edu/software/tagger.shtml

E. Opinion Words Identification

Opinion words are the words which express opinion towards aspects. In the aspect based opinion mining, the aspect related opinion words should be identified, i.e. the words, which modify the meaning of aspects. The opinion words used in this work was adjectives, verbs, adverb adjective and adverb verb combinations. The proposed technique uses a syntactic based approach with the help of dependency parser for opinion words extraction. Stanford dependency parser was used for this purpose[2].

The adjectives are the main opinion words in a sentence. Proposed technique uses adjectives, verbs, adverb adjective combinations and adverb verb combinations as opinion words. In [16] Opinion words are extracted by searching 5-gram forwards and backwards from the aspect position in a sentence. This was done based on POS tag information; it is not a suitable method if a sentence contains multiple aspects and also will not get contextual polarity scores accurately. Hence the proposed technique uses syntactic dependency relations between words. Some algorithms are written for the extraction of the opinion words and for the score calculation of opinion words. Based on this score, the final polarity of the aspects are found out.

1) Extraction of Adjectives: The extraction process is given in Algorithm 1. The function GETADJECTIVE(A,P_s) extracts the aspect related adjectives of a subjective sentence using the proposed conditional rules. The input A is the aspect of the subjective sentence and P_s is the parsed form of that subjective sentence (i.e. output of a Stanford dependency parser).

2) Extraction of Verbs: The aspect related verb extraction is given in Algorithm 2. Function GETVERB(A, P_s) extracts the aspect related verbs using dependency relations.

3) Extraction of Adverbs: Adverb extraction is given in Algorithm 3. Function GETADVERB(*op*, P_s) extracts the adverb corresponding to the opinion word *op*. If the adjective is given as *op*, the above function return the adverb of that adjective, such a way adverb adjective combinations in a sentence are obtained. In the same way, adverb verb combinations can be extracted.

F. Aspect Level Score Calculation

Upto this step, aspect related opinion words are extracted from the subjective sentence of a review. In aspect level score calculation step, the polaity score of an aspect in a sentence is calculated by aggregating opinion words scores in that sentence.

The proposed technique assigns the pripority scores to the opinion words using SentiWordNet. SentiWordNet is a dictionary of sentiment words. In it each synset of terms in Wordnet has given positive and negative scores. Positive score shows the positive polarity of the corresponding word and negative score shows the negative polarity of the corresponding word. Word sense disambiguation is not considered here. The categories a, r, n, v in the SentiWordNet stands for adjective,

Algorithm 1 Extraction of Adjectives

function GETADJECTIVE(Matchedaspect A, ParsedSentence P_s))
  $Adjectives$: Array of Adjectives
  X, Y, Y_1: String
  if (cc(Y, "but")) $\in P_s$ **then**
    if nsubj(Y, X)& (X = "it" or $X = A$) **then**
      Add Y to $Adjective$
    end if
  end if
  if (amod(A, Y) or rcmod(A, Y) or advmod(A, Y)) $\in P_s$ **then**
    Add Y to $Adjective$
    if (conj_and(Y, Y_1)) $\in P_s$ **then**
      Add Y_1 to $Adjective$
    end if
  end if
  if (prep_like(Y, A)) $\in P_s$ **then**
    Add "Like" to $Adjective$
  end if
  if (prep_of(X, A)) $\in P_s$ **then**
    $Adjective$ = getAdjective(X, P_s)
  end if
  if (root(Root, Y) & nsubj(Y, A) & cop(Y, X)) $\in P_s$ **then**
    Add Y to $Adjective$
  end if
  if (nsubj(X, A) & acomp(X, Y)) $\in P_s$ **then**
    Add Y to $Adjective$
  end if
  Return $Adjective$
end function

Algorithm 2 Extraction of Verbs

function GETVERB(MatchedAspect A, ParsedSentence P_s))
  $Verb$: Array of verbs
  Y : String
  if (root(Root, Y) & nsubj(Y, A)) $\in P_s$ **then**
    Add Y to $Verb$
  end if
  if (root(Root, Y) & (dobj(Y, A) or pobj(Y, A))) $\in P_s$ **then**
    Add Y to $Verb$
  end if
  if (root(Root, Y) & (nsubj_pass(Y, A) or xsubj(Y, A))) $\in P_s$ **then**
    Add Y to $Verb$
  end if
  Return $Verb$
end function

[2]http://nlp.stanford.edu/software/lex-parser.shtml

Algorithm 3 Extraction of Adverbs

function GETADVERB(OpinionWord *op*, ParsedSentence P_s))
  Adverb : Array of adverbs
  Y : String
  if (advmod(*op*, Y) or advcl(*op*, Y) or amod(*op*, Y)) $\in P_s$ **then**
    Add Y to $Adverb$
  end if
  Return $Adverb$
end function

adverb, noun and verb respectively. Priority scores are assigned by searching the opinion words in the SentiWordNet. If it is present then retrieve the corresponding scores, positive or negative. If the matched opinion word is more than one time in the SentiWordNet, then take the average positive or negative score, which one is higher.

1) Negation Handling: Negations should be handled appropriately to get the contextual information of a sentence, which is given in Algorithm 4. The proposed technique use dependency relations for it. The function CHECKNEGATION(Op, P_s) check that the input opinion word Op is in some negative relations. If the opinion word is in negative relation then its priority score is reversed for negtion handling purpose.

Algorithm 4 Negation Handling

function CHECKNEGATION(OpinionWord Op, ParsedSentence P_s)
  Y : String Variable
  $flag$: Boolean Variable
  $flag = FALSE$
  if (neg(Op, Y)) $\in P_s$ **then**
    $flag = TRUE$
  end if
  if (conj_negcc(Y, Op)) $\in P_s$ **then**
    $flag = TRUE$
  end if
  if (pobj("not", Op) $\in P_s$ **then**
    $flag = TRUE$
  end if
  Return $flag$
end function

2) Aggregation scheme of Adverb and Opinion words: The polarity score of adverb and Opininon word (adjective or verb) is aggregated by adding or subtracting some weightage of adverb score to the Opinion word score as in [16], which is given in Algorithm 5. sf is the weightage or scale factor, which was fixed by assigning different values to sf. The taken value was 0.3.

3) Aggregation Scheme of Adjectives: This aggregation scheme is used in two cases. In the first case, the extracted opinion words of a sentence are adjectives. In such cases, the scores of all adjective are to be aggregated. In the second case,

Algorithm 5 Aggregation scheme of Adverb and Opinion words

function AGGREGATESCORE(AdverbScore $AdverbScore$, OpinionWordScore $OpScore$))
  $AggregateScore$: Variable which store the aggregate score of adverb and opinion word
  if $OpScore = 0$ **then**
    $AggregateScore = 0$
  else
    if $AdverbScore > 0$ **then**
      if $OpScore > 0$ **then**
        $AggregateScore$ = min(1, $OpScore + sf * AdverbScore$)
      end if
      if $OpScore < 0$ **then**
        $AggregateScore$= min(1,$OpScore - sf * AdverbScore$)
      end if
    end if
    if $AdverbScore < 0$ **then**
      if $OpScore > 0$ **then**
        $AggregateScore$= max(-1,$OpScore + sf * AdverbScore$)
      end if
      if $OpScore < 0$ **then**
        $AggregateScore$= max(-1,$OpScore - sf * AdverbScore$)
      end if
    end if
  end if
  Return $AggregateScore$
end function

opinion words are adjectives and adverb adjective combinations. In such cases, the scores of all adjectives and the scores of adverb adjective combinations are to be aggregated. The Average($AdjectiveScore$) in the algorithm find the average adjectives score. The algorithm for verb score calculation is same as adjective score aggregation (not included).

4) Sentence wise Score of an Aspect: The polarity score of the aspect A in the sentence S is find out by aggregating the verb scores and adjective scores, which is given in Algorithm 7. The scale factor sf is used for giving the score weightage, here some weightage of verb score is added to the adjective score. The value of sf was fixed by trial and error method. The value is 0.3.

G. Calculation of Aspects Score in All Reviews

The total score of an aspect from all reviews is find out by aggregating the sentence wise score of that aspect. Positive and negative scores are seperatly aggregated. For example, after analysing the 5 sentences, if we get three positive and two negative scores for the aspect food. Final polarity is find out by aggregating this three positive and two negative scores separately. If the positive score is high, the opinion about the

Algorithm 6 Aggregation Scheme of Adjectives

function AGGREGATEADJECTIVESCORE($Adjectives$, P_s)
  $Adjectives$: Array of adjectives
  $AdjectiveScore$: Priority Score of adjective retrieved from SentiWorNet.
  $Adverbs$: Array of Adverbs
  Agg_AdvAdj_Score : Aggregate score of adverb and adjective(Ex: 'very good').
  $Agg\_Adjective\_Score$: Aggregated scores of adjectives and Adverb Adjective combination.
  $Agg\_Adjective\_Score = 0$
  $AdjectiveScore = 0$
  if $Adjectives \neq null$ **then**
    for all Adjective $adj \in Adjectives$ **do**
      $AdjectiveScore$ = Score(adj)
      if CheckNegation(adj, P_s)$\neq False$ **then**
        Reverse the $AdjectiveScore$
      end if
      $Adverb$ = getAdverbs(adj, P_s)
      if $Adverb \neq null$ **then**
        $AdverbScore$ = Score($Adverb$)
        if CheckNegation($Adverb$, P_s)$\neq False$ **then**
          Reverse the $AdverbScore$
        end if
        Agg_AdvAdj_Score+= AggregateScore($AdverbScore$, $AdjectiveScore$)
      else
        $AdjectiveScore+ = AdjectiveScore$
      end if
    end for
    if $Adverb \neq null$ **then**
      Agg_AdvAdj_Score = Average(Agg_AdvAdj_Score)
      $AdjectiveScore$ = Average($AdjectiveScore$)
      $Agg\_Adjective\_Score = Agg\_AdvAdj\_Score + sf * AdjectiveScore$
    else
      $Agg\_Adjective\_Score$ = Average($AdjectiveScore$)
    end if
  else
    $Agg\_Adjective\_Score = 0$
  end if
  Return $Agg\_Adjective\_Score$
end function

Algorithm 7 Sentence wise Score of an Aspect

function FINALSENTENCESCORE(AggregateAdjectiveScore Tas, AggregateVerbScore Tvs)
  $SentencewiseAspectScore = Tas + sf * Tvs$
  Return SentencewiseAspectScore
end function

food from the set of reviews is classified as positive otherwise negative. Final aspect score from all reviews can be found out by using the formula given below.

For each aspect j of the Product,

$$Aggregate\_Positive\_Polarity[j] = \sum_i Positivepol_{i,j} \quad (1)$$

$$Aggregate\_Negative\_Polarity[j] = \sum_i Negativepol_{i,j} \quad (2)$$

In the above formula the $Positivepol_{i,j}$ is the positive score of the j^{th} aspect in i^{th} sentence(j^{th} aspect matched sentence). $Negativepol_{i,j}$ is the negative score of the j^{th} aspect in i^{th} sentence(j^{th} aspect matched sentence). For normalization, the final scores are averaged. Positive and negative scores of aspects are separately aggregated; hence we get an aggregate positive score and negative score of each aspect. Formula for the normalization is given below.

$$Normalized\_Positive\_Polarity[j] = \frac{\sum_i Positivepol_{i,j}}{\sum_i} \quad (3)$$

$$Normalized\_Negative\_Polarity[j] = \frac{\sum_i Negativepol_{i,j}}{\sum_i} \quad (4)$$

Using these values we can summarize the results by graphical representations. Visualization tools can be used for this purpose, which will be very useful for customers for knowing the aspect level opinions of others about a product. The users also get an idea that, opinion of each aspect is how much positive and how much negative.

Algorithm 8 PolarityDetection

for all Aspect $A_d \in AspectTable$ **do**
  if $AggPosScore.A_d > AggNegScore.A_d$ **then**
    $FinalAspectScore.A_d = AggPosScore.A_d$
    $A_d = POSITIVE$
  else
    $FinalAspectScore.A_d = AggNegScore.A_d$
    $A_d = NEGATIVE$
  end if
end for

V. EXPERIMENTAL RESULT

The proposed opinion mining model is implemented using java with Eclipse IDE and Apache Tomcat server.

A. Dataset

Restaurant reviews collected as test dataset, which are available in TripAdvisor. The reviews were collected using review crawler, which was implemented using Jsoup Parser. The collected reviews are manually annotated with positive

and negative polarities in sentence wise. The dataset contains 170 reviews was tagged manually. For example if a sentence has positive polarity of food, then it is tagged as FP., if it is negative then tagged as FN.

B. Performance Evaluation

The performance measures used for the evaluation purpose was Accuracy, Precision and Recall. Formulas[20] are given below.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

$$Precision = \frac{TP}{TP + FP} \quad (6)$$

$$Recall = \frac{TP}{TP + FN} \quad (7)$$

In the above formula, TP is the true positive instances. TN is the true negative instances. FP is the false positive instances and FN is the false negative instances.

Aspect level accuracy, precision and recall of the proposed method on the annotated dataset are calculated given in the Table 6.1 and in Fig. 2.

TABLE I
ACCURACY, PRECISION AND RECALL OF PROPOSED TECHNIQUE

Aspects	Accuracy	Precision	Recall
Food	80.48	86.11	91.17
Service	92.85	100	91.66
Atmosphere	86.95	100	84.21
Price	58.33	50	100
Menu	66.66	66.66	100
Staff	81.81	88.88	88.88
Restaurant	61.53	66.66	75

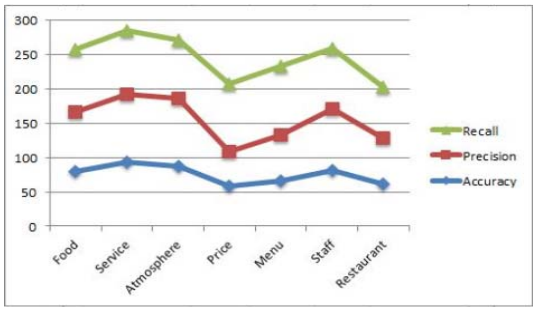

Fig. 2. Accuracy, Precision and Recall of Proposed Technique

For comparison, first implement a system using the existing method in[16], which use adjective+adverb adverb+verb as features and they are extracted using POS Tagger; the dependency relations are not used. For the identification, this method is named as POS. Then the evaluation was done on the annotated dataset. The total accuracy of 72.56% was obtained. It was modified by using syntactic dependency with same features, this new modified method is named as AOMR1. AOMR1 is also evaluated on the annotated dataset. The total accuracy of 71.79% was obtained. AOMR1 is again modified by adding negation handling and some more relations rules are added(i.e. some positive words in restaurant domain has negative score in SentiwordNet, which is also corrected) and evaluation was done on the same annotated dataset, then got the total accuracy of 78.04% on annotated dataset. Final version is called AOMR2.

The comparison result of the above three method on total accuracy, precision and recall is given in Table.6.2.

TABLE II
COMPARISON RESULT OF PROPOSED METHOD

Method	Accuracy	Precision	Recall
POS	72.56	79.56	83.05
AOMR1	71.79	77.78	87.5
AOMR2	78.04	83	89.25

The errors are happened in the results. The errors are occurred due to

1) The user written reviews are highly unstructured. Spelling mistakes are also the problem. Hence we will not get correct syntactic dependency.
2) This proposed method takes only the explicit aspects matched sentences for processing. Implicit aspects are not identified.
   Example: "The restaurant was expensive". Here the implicit aspect 'price' cannot be determined using proposed method.
3) The correct opinion can not be obtained from complex sarcasm sentences.
   Example : "The food was fantastic; I was hospitalized for two days".

VI. CONCLUSION

Opinion mining or sentiment analysis is a comprised area of natural language processing, computational linguistics and data mining, in it reviews about a topic is analysed and expressed opinions are extracted. Most of the previous work is in the field of document or sentence level analysis. This paper proposes a new different syntactic approach to aspect level opinion mining, which use aspect dictionary, SentiWordNet, Dependency parsing, adverb adjective, adverb verb combinations, adjectives and adverbs together for opinion mining process with automatic acquisition of aspects. It is a syntactic based approach hence there is no need any training data.

In the proposed method aspect matched opinion Words are extracted using dependency parsing. Polarity of opinions i.e. positivity or negativity of an aspect is find out using SentiWordNet, adjective, adverb adjective and adverb verb combinations using this aspect based visual summary can be produced, which shows positivness and negativeness of each aspect from total reviews.

The performance of the proposed method is evaluated by building a annotated test set of restaurant reviews. Evaluation is done by measuring accuracy, precision and recall. The total accuracy of 78.04% was obtained on manually annotated test dataset.

In this work only explicit aspects are considered and aspects are extracted using training. If modified the aspect extraction task without training will improve the accuracy. Word sense disambiguation is ignored in it. Next work aims to include the same and also the analysis of different type of sentences like conditional, comparative sentences in order to improve the accuracy of opinion mining and also include automatic grouping of aspect synonyms. In this work the priority scores of opinion words are assigned using SentiWordNet, future work aims to avoid the SentiWordNet score, and find an unsupervised approach.

REFERENCES

[1] T. Nasukawa and J. Yi, "Sentiment analysis: Capturing favorability using natural language processing," in *Proceedings of the 2nd international conference on Knowledge capture*. ACM, 2003, pp. 70–77.

[2] K. Dave, S. Lawrence, and D. M. Pennock, "Mining the peanut gallery: Opinion extraction and semantic classification of product reviews," in *Proceedings of the 12th international conference on World Wide Web*. ACM, 2003, pp. 519–528.

[3] B. Liu, "Sentiment analysis and opinion mining," *Synthesis Lectures on Human Language Technologies*, vol. 5, no. 1, pp. 1–167, 2012.

[4] M. Hu and B. Liu, "Mining and summarizing customer reviews," in *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2004, pp. 168–177.

[5] J. S. Kessler and N. Nicolov, "Targeting sentiment expressions through supervised ranking of linguistic configurations." in *ICWSM*, 2009.

[6] G. Qiu, B. Liu, J. Bu, and C. Chen, "Opinion word expansion and target extraction through double propagation," *Computational linguistics*, vol. 37, no. 1, pp. 9–27, 2011.

[7] Q. Liu, Z. Gao, B. Liu, and Y. Zhang, "A logic programming approach to aspect extraction in opinion mining," in *Web Intelligence (WI) and Intelligent Agent Technologies (IAT), 2013 IEEE/WIC/ACM International Joint Conferences on*, vol. 1. IEEE, 2013, pp. 276–283.

[8] A. Bagheri, M. Saraee, and F. de Jong, "An unsupervised aspect detection model for sentiment analysis of reviews," in *Natural Language Processing and Information Systems*. Springer, 2013, pp. 140–151.

[9] F. Benamara, C. Cesarano, A. Picariello, D. R. Recupero, and V. S. Subrahmanian, "Sentiment analysis: Adjectives and adverbs are better than adjectives alone." in *ICWSM*, 2007.

[10] X. Ding and B. Liu, "The utility of linguistic rules in opinion mining," in *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*. ACM, 2007, pp. 811–812.

[11] T. T. Thet, J.-C. Na, C. S. Khoo, and S. Shakthikumar, "Sentiment analysis of movie reviews on discussion boards using a linguistic approach," in *Proceedings of the 1st international CIKM workshop on Topic-sentiment analysis for mass opinion*. ACM, 2009, pp. 81–84.

[12] R. Prabowo and M. Thelwall, "Sentiment analysis: A combined approach," *Journal of Informetrics*, vol. 3, no. 2, pp. 143–157, 2009.

[13] D. Gräbner, M. Zanker, G. Fliedl, and M. Fuchs, *Classification of customer reviews based on sentiment analysis*. na, 2012.

[14] R. Varghese and M. Jayasree, "Aspect based sentiment analysis using support vector machine classifier," in *Advances in Computing, Communications and Informatics (ICACCI), 2013 International Conference on*. IEEE, 2013, pp. 1581–1586.

[15] A. Hogenboom, D. Bal, F. Frasincar, M. Bal, F. de Jong, and U. Kaymak, "Exploiting emoticons in sentiment analysis," in *Proceedings of the 28th Annual ACM Symposium on Applied Computing*. ACM, 2013, pp. 703–710.

[16] V. Singh, R. Piryani, A. Uddin, and P. Waila, "Sentiment analysis of movie reviews: A new feature-based heuristic for aspect-level sentiment classification," in *Automation, Computing, Communication, Control and Compressed Sensing (iMac4s), 2013 International Multi-Conference on*. IEEE, 2013, pp. 712–717.

[17] E. Marrese-Taylor, J. D. Velásquez, F. Bravo-Marquez, and Y. Matsuo, "Identifying customer preferences about tourism products using an aspect-based opinion mining approach," *Procedia Computer Science*, vol. 22, pp. 182–191, 2013.

[18] E. Marrese-Taylor, J. D. Velásquez, and F. Bravo-Marquez, "Opinion zoom: A modular tool to explore tourism opinions on the web," in *Web Intelligence (WI) and Intelligent Agent Technologies (IAT), 2013 IEEE/WIC/ACM International Joint Conferences on*, vol. 3. IEEE, 2013, pp. 261–264.

[19] B. Ohana and B. Tierney, "Sentiment classification of reviews using sentiwordnet," in *9th. IT & T Conference*, 2009, p. 13.

[20] B. Monalisa Ghosh, Animesh Kar and Bu, "Unsupervised linguistic approach for sentiment classification from online reviews using sentiwordnet 3.0," *International Journal of Engineering Research and Technology (IJERT)*, vol. 2, no. 9, pp. 2278–0181, 2013.