Intelligent Robot with Aspect based Sentiment Analysis

P.Sanju¹, T.T. Mirnalnee²

¹ Assistant Professor, Department of Computer Science and Engineering, University College of Engineering, Tindivanam, Tamilnadu, India. Email id: sanjuccev@gmail.com
² Professor, Department of Computer Science and Engineering, SSN College of Engineering, Chennai, Tamilnadu, India. Email id: mirna123@gmail.com

Abstract The robots are very useful in automotive industries to manufacture different types of products. However, the existing industrial robots are dumb robots which perform the works based on the instructions given by the users. It is necessary to provide more intelligence to the industrial robots by representing the knowledge using AI techniques. In order to make intelligent robots in automotive industry, the proposed approach constructs the aspects - multi polarity sentiment lexicon from the user reviews of the products. This aspects-multi polarity sentiment lexicon specifies the aspects of the products and its corresponding multi polarity sentiments such as strong positive, weak positive, strong negative and weak negative by analyzing the sentiment orientations of the reviews based on the linguistic hedges. Using this lexicon, intelligent robot acquires weakness of the products and takes decision itself to perform the changes in behavior of the system to satisfy the customer needs. Experiments on two different data sets show that the proposed approach not only identifies the aspects of the products and also determines the multi polarity of the aspects based on the linguistic hedges. In further quantitative evaluation, the proposed approach is proved to be more effective than that of manually constructed aspects- multi polarity sentiment lexicon.

1. INTRODUCTION

Now-a-days, robots play a vital role in automotive industry because they perform the specified tasks without getting tired and also work in risky places. As a result, more robots are implanted in manufacturing industries to increase the productivity. They are used in automotive industry for various applications such as welding, painting and assembling etc. Recently, they are designed with eyes which are very useful to assemble the parts of the car body. Although, the automotive industry uses a large number of industrial robots, they are not intelligent. To make the intelligent robots, the proposed aspect based sentiment analysis system constructs aspects – multi polarity sentiment lexicon.

In the present work, the aspect based sentiment analysis system is implanted that analyzes the online reviews of the products and spots out the strength and weakness for each aspect of the products. Using this system, the intelligent robot analyzes the reviews of the products daily or weekly basis and it takes decisions by itself and performs necessary changes in the manufacturing products based on the results of review analysis. For example, if people like gray color cars, the robot automatically paints more number of gray color cars instead of other colors. Generally, customers are writing the reviews with fuzziness while conveying the feedback about the products. For example, “The car shape is somewhat good” that diminishes the positive sentiment or opinion of the car. These types of reviews are available abundance in the web. Hence, it is necessary to extract the aspect of the products with its corresponding multi polarity sentiments such as strong positive (SP), weak positive (WP), strong negative (SN) and weak negative (WN). Sentiment analysis is used in various applications such as reviews summarization [Y.Lu et al., (2009) & Chien-Liang Liu et al., (2012)], aspect mining [T.K Fan et al., 2010] and contextual advertising [Jinbo Zhu et al., 2011].

The framework of aspect based sentiment analysis system involves various phases like (a) preprocessing (b) Feature extraction (c) computation of sentiment score (d) construction of aspects- multi polarity sentiment lexicon. In preprocessing phase, stop word removal and parts of speech (POS) tagging functions are applied to the reviews. In feature extraction phase, unigram, bigram and trigram features are extracted from the reviews and aspects are identified based on frequency occurrence of noun features. In computation of sentiment score phase, sentiment score is computed for all unigram and bigram features. Next, unigram and bigram features are clustered into strong positive (SP), weak positive (WP), weak negative (WN) and strong negative (SN) using K mean clustering based on the scores. The K mean clustering technique partitioning n features into k clusters by computing minimum distance between the sentiment score of the features and mean value of the clusters. Finally aspects –multi polarity sentiment lexicon is constructed using clustered results.

The rest of the paper is organized as follows, Section 2 describes related works, and Section 3 describes the methodology. The Experiments and Result Analysis are discussed in Section 4. Section 5 concludes with the direction of future work.

2. RELATED WORKS

Aspect-based sentiment analysis is becoming more popular in recent years. A.M. Popesu et al., (2005) revealed the frequency based approach which extracted the high frequency noun phrases as aspects from the reviews. B. Liu et al., (2005) proposed relation based approach which identified the aspects based on the aspect-sentiment relation in the reviews. Victor C. Cheng et al., (2014) proposed probabilistic aspect mining model for drug reviews. They created Probabilistic Aspect Mining Model (PAMM) which identified the aspects.
based on the class label. Jo Y et al., (2011) proposed aspect and sentiment unification model which generated sentiment-topic pair for a single sentence rather than for a word as Joint Sentiment/Topic model. Sanj et al., (2014) proposed construction of enhanced sentiment sensitive thesaurus which aligns sentimentally similar features of various domains as well as semantic features from wiktorary.

Kennedy et al., (2006) proposed contextual valence shifter which predicted the sentiment polarity of the sentences by increasing or decreasing the sentiment score based on the presence of intensifier or diminishing features used in the reviews. The work on the papers (Cacilia Zirn et al., 2011) and (Axel Schulz et al., 2013) discussed the fine grained sentiment analysis of online reviews of the products. They analyzed the sentiment of the reviews for each aspect or topic. Chenghua Lin et al., (2012) proposed a novel probabilistic modeling framework called Joint Sentiment/Topic model which detected sentiments and topics simultaneously from the text. Xueke et al., (2013) proposed aspect based opinion mining from customer reviews. They proposed a multi aspect bootstrapping method which learnt aspect-related terms of each aspect. Aurangzeb kh et al., (2011) and Chihli Hung et al., (2013) discussed sentiment classification using SentiWordNet (A. Esuli et al., (2006)).

Yue lu et al.,(2011) proposed automatic construction of context-aware sentiment lexicon by optimization approach. This sentiment lexicon was created based on the aspects of the contexts. For example, in electronics domain, the word “large” is positive if the sentence as “laptop screen is large”. But the same word “large” is negative if the sentence as “battery of the laptop is large”. To solve this problem, they constructed context-aware sentiment lexicon using online reviews based on the contexts of the aspects. Mita K. Dalal et al., (2014) proposed opinion mining technique which classified the sentiment of online reviews into strong positive, weak positive, strong negative, weak negative and mixed based on the linguistic hedges present in the reviews. Moreover, the work in the paper (Mita K. Dalal et al., 2014) maintains the lookup table in which features and linguistic hedges are stored manually and they are not constructed any sentiment lexicon to predict the sentiment of the reviews. They only classify the reviews as SP, WP, WN and SN. The proposed aspect based sentiment analysis system constructs aspects – multi polarity sentiment lexicon automatically using the reviews. Moreover, this aspects – multi polarity sentiment lexicon spots out multi polarity sentiment for each aspect of the products.

3. METHODOLOGY

This section discusses the framework of aspect based sentiment analysis system which automatically constructs aspects – multi polarity sentiment lexicon using the reviews of the products. The intelligent robot collects online reviews of the products which were manufactured by itself in timely manner and construct aspects- multi polarity sentiment lexicon using the proposed aspect based sentiment analysis system. For example, if robot is used in Honda car manufacturing industry, it collects only the reviews of Honda car. Using this aspect based sentiment analysis system, the intelligent robot identifies the aspect of the product and its multi polarity sentiment such as SP, WP, WN and SN of car products. Using this aspects- multi polarity sentiment lexicon, strong and weak negative opinion features of the products are identified by the robot. In order to satisfy the customer needs, the intelligent robot takes decision itself and performs some changes in the products while manufacturing. The Fig.1 shows the framework of aspect based sentiment analysis system which constructs aspects- multi polarity sentiment lexicon using the reviews of the products.

![Fig.1. Framework of Aspect based Sentiment Analysis](image)

The various stages of aspect based sentiment analysis system are (a) Preprocessing (b) Feature Extraction (c) Computation of sentiment score (d) Construction of aspects - multi polarity sentiment lexicon.

A. Preprocessing and Feature Extraction

The online review documents of the products are collected from e-commerce websites. First, review documents are split into sentences and preprocessing techniques such as stop words removal and Parts of speech (POS) tagging are applied on the reviews using Natural Language Tool Kit (NLTK). Using stop word removal algorithm unwanted words like ‘is’, ‘was’, ‘when’, ‘where’, ‘that’ etc., are removed from each review. Second, POS tagging is applied to the reviews. This POS tagger attaches POS tag to each word of the sentences in the given review document. The features nouns, verbs and descriptors such as adverbs and adjectives are extracted
from each sentence of the reviews. Next, unigram, bigrams and trigram features are formed for each sentence of the reviews. Table 1 shows the extraction of unigram, bigram and trigram features for the sample review by applying preprocessing techniques such as stop word removal and POS tagging using NLTK. In Table 1, the first row shows an example of review sentence for Honda city car model. Second row shows the stop words removed sentences. Third row shows the POS tagged sentences in which [n] indicates noun and [a] indicates adjective. Fourth, fifth and sixth row show the extraction of unigram, bigram and trigram features from each sentence of the review. Bigram and trigram features are very useful to determine the semantic meaning of the sentences. Moreover, bigram and trigram features are used to determine the multi polarity sentiment for each aspect of the products. The aspects of the products are determined using trigram or bigram features. Generally noun or noun phrases are taken as aspect of the products. In Table 1, ‘Exterior’, ‘toxic smell’, ‘height’ are taken as aspect and its multi polarity sentiments are determined based on the presence of linguistic hedges used in the reviews.

Table 1. Example for Preprocessing and Feature Extraction

<table>
<thead>
<tr>
<th>Review</th>
<th>Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>The exterior is extremely beautiful. But toxic smell is not good. Lower height is less attractive.</td>
<td>Exterior extremely beautiful / toxic smell not good / lower height less attractive</td>
</tr>
<tr>
<td>Stop word removal</td>
<td>Exterior extremely beautiful / toxic smell not good / lower height less attractive</td>
</tr>
<tr>
<td>Unigram</td>
<td>Exterior, extremely, beautiful / toxic smell, not, good / lower height, less, attractive.</td>
</tr>
<tr>
<td>Bigram</td>
<td>Exterior+extremely, extremely+beautiful/ Toxic smell+not,not+good/lower height+less, less+attractive</td>
</tr>
<tr>
<td>Trigram</td>
<td>Exterior+extremely+beautiful / toxic smell+not, not+good, lower height+less+attractive</td>
</tr>
</tbody>
</table>

More significantly, the bigrams and linguistic hedges features play a vital role to identify the multi polarity sentiment for each aspect of the product. Hence, four different types of bigram features such as ‘Adjective with Noun’, ‘Noun with Adjective’, ‘Adjective with Adjective’ and ‘Linguistic hedge with descriptors’ (adjective or adverb) are extracted from each sentence of the given reviews. Next, sentiment score is computed for all bigram features of each sentence. The sentiment score of the adjective feature is directly assigned when it is preceded /followed with noun features. The average score is assigned when the bigram features are adjectives. The sentiment score is computed when the descriptor feature is preceded by Linguistic hedges. More specifically, the presence of concentrating /dilating/ modify linguistic hedges are useful to identify the multi polarity sentiment of the reviews such as strong positive (SP), weak positive (WP), Weak Negative (WN) and Strong Negative (SN). Moreover, these hedges are used to spot out the strength and weakness of the features of the products. Table 2 shows some examples of concentrator, dilator and modifier linguistic hedges.

Table 2. Examples of modifier, concentrator and dilator linguistic hedges

<table>
<thead>
<tr>
<th>Modifier Hedges</th>
<th>Concentrator Hedges</th>
<th>Dilator Hedges</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not, never, not so</td>
<td>Very, Extremely, Absolutely, Highly, incredibly, Positively, significantly</td>
<td></td>
</tr>
<tr>
<td>Quite, hardly, somewhat, almost, less</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

When the descriptor has a preceding hedge, its sentiment score is computed using the equation (1) as in [Mita K. Dalal et al., 2014].

\[
F(\mu(s))=1-(1-\mu(s))^\delta 
\]  

B. Computation of Sentiment Score

To detect multi polarity sentiment for each aspect of the products in the given reviews, it is necessary to compute sentiment score for all the features of the reviews using SentiWordNet3.0. The sentiWordNet3.0 is an enhanced lexical resource used for sentiment analysis and opinion mining. It has triplet score which indicates positive, negative and objective sentiment scores for each synset of WordNet. For example, the sentiment score of the feature “good” is taken from SentiWordNet3.0 as [0.75, 0.0, 0.25] which indicates its positive, negative and objective score. Generally, adjectives are very useful to identify the sentiment polarity of the products. Hence, the ‘adjective’ unigram features are taken from each sentence of the reviews and its sentiment score is directly assigned from SentiWordNet3.0. But in practical, sentiment score is not available for some of the features. For those features, the corresponding definitions are taken from WordNet and sentiment scores of unigram features of the definition are averaged.

In order to identify the multi polarity sentiment for each aspect of the products, sentiment score is computed when the descriptor is preceded by fuzzy linguistic hedges. Fuzzy concentrator/dilator operator is used to compute the sentiment score. When the hedge is concentrator, \(\delta\) value is taken as 2 in equation (1) which gives fuzzy concentrator equation (2) as in [Mita K. Dalal et al., 2014] and when the hedge is dilator, \(\delta\) value taken as \(\frac{1}{2}\) which gives fuzzy dilator equation (3) as in [Mita K. Dalal et al., 2014].
\[ F_4(\mu(s)) = 1 - (1 - \mu(s))^2 \quad (2) \]

\[ F_d(\mu(s)) = 1 - (1 - \mu(s))^{1/2} \quad (3) \]

For example, consider the review sentence “The interior of the Honda city car is extremely good”. The sentiment score of the descriptor “good” is \( \mu(s) = [0.75, 0.0, 0.25] \) which is taken from SentiWordNet 3.0. The maximum score in the triplet determine the sentiment polarity of the features. The sentiment score for “good” feature has high positive score, hence it is considered as positive sentiment. The proposed work aims to determine the multi polarity sentiment for each aspect of the product based on the presence of the linguistic hedges. In the example, the descriptor “good” is preceded by concentrator linguistic hedge “extremely”. The maximum value in triplet score is \( \mu(s) = 0.75 \) which is modified as 0.93 by applying fuzzy concentrator equation (2) and remaining scores in the triplet scores also modified to get summation of triplet score equal to 1. Hence, the score of the bigram feature “extremely good” is computed as \( \mu[s] = [0.93, 0.0, 0.07] \) using fuzzy concentrator and important property which is held by the fuzzy function is summation of these three score should not exceed 1. The positive sentiment score is increased due to the occurrence of concentrating hedge “extremely”. When the descriptor is preceded by modifier hedge “not good”, sentiment score is modified by swapping both positive and negative score of the feature. The sentiment score of the feature “not good” is modified as \( \mu(s) = [0.0, 0.75, 0.25] \). Similarly, when the descriptor is preceded by dilator linguistic hedge “somewhat good”, sentiment score is computed by fuzzy dilator operator using equation (3) and the remaining scores are also modified to get summation value 1. The sentiment score for the feature “somewhat good” is modified as \( \mu(s) = [0.50, 0.0, 0.50] \). Here, positive value of the good is diminished due to the occurrence of “somewhat dilating hedge. Similarly, the sentiment score is calculated for all bigram of the features. Generally, concentrating hedges are used to determine strong positive or negative features. Dilating hedges determine weak positive or negative features. The modifier hedges are used to reverse the sentiment of the products.

C. Construction of Aspects- Multi polarity sentiment Lexicon

To construct aspects- multi polarity sentiment lexicon, all unigram and bigram features are clustered into four clusters such as strong negative, weak negative, strong positive and weak positive based on its sentiment scores of features using K Mean Clustering algorithm. The K mean clustering algorithm partitions the given features into K clusters by finding minimum distance between the sentiment score of the features and cluster mean values. All unigram and bigram features sentiment scores are given to K Mean Clustering algorithm which partitions the features into K clusters using equation 4 as in [Tapas Kanungo et al., (2002)].

\[
\arg\min_{\mu_i} \sum_{k=1}^{K} \sum_{x_j \in \mu_i} ||x_j - \mu_i||^2
\quad (4)
\]

where \( x_j \) indicates triplet score of features and \( \mu_i \) indicates mean value of the clusters. Here initial clusters mean values are given to K Mean algorithm to cluster the features into four categories such as SP, SN, WN and WP. The initial mean values are taken as \([1.0, 0.0, 0.0, 0.0], [0.5, 0.0, 0.5, 0.5], [0.0, 1.0, 0.0], [0.0, 0.5, 0.5] \) for SP, WP, SN and WN clusters respectively. Here, SP cluster mean value is taken as \([1.0, 0.0, 0.0, 0.0] \) because strong positive features require highest positive value in the triplet score. Similarly SN cluster mean value is taken as \([0.0, 1.0, 0.0, 0.0] \) because strong negative features require highest negative value in the triplet score. The WP cluster mean values are set as \([0.5, 0.0, 0.0, 0.0] \) which indicate positive and objective scores are set to 0.5. Similarly WN cluster mean values are set as \([0.0, 0.5, 0.0, 0.5] \) which indicate negative and objective scores are set to 0.5. Here k value is taken as 4 to make four clusters. The K Mean clustering algorithm performs partitioning the given n features into four clusters by finding the minimum distance between the sentiment score of the feature and mean value of the clusters.

Aspects of the products are identified based on the term frequency of the noun features appeared in the reviews. The noun features which occur more than threshold values are considered as aspects. Here, the proposed approach set the threshold value as 5. The multi polarity sentiment of each aspect of the product is identified based on the presence of the descriptors with linguistic hedges. Here, aspects (nouns) are extracted from bigram and trigram features whenever the aspects are followed by descriptor or linguistic hedges with descriptor and its corresponding multi polarity sentiment is determined using clustered features. For example, the trigram features, ‘noun+hedges+descriptor’, multi polarity sentiment of the noun or aspect is determined by the presence of concentrating/dilating hedges before the descriptor. Similarly, ‘noun+adjective’ bigram features, multi polarity sentiment of the noun is determined based on the sentiment score of the adjective features. Finally aspects-Multi polarity sentiment lexicon is constructed by specifying aspect with its multi polarity sentiment such as SP, WP, WN and SN using clustered features. Using this lexicon, the intelligent robot
identifies the weakness of the products and it automatically makes decision and performs some changes in the manufacturing products.

**Algorithm 1**

**Input:** Reviews of the products  
**Output:** Construction of Aspect – Multi Polarity Sentiment Lexicon

1. Unigram, bigram and trigram features are extracted from each sentence of online reviews after preprocessing such as stop words removal and POS tagging.
2. Sentiment scores are assigned for all unigram features directly from SentiWordNet.
3. Sentiment scores of bigram features are calculated by applying the following rules
   - When the descriptor is preceded by concentrating hedges, the sentiment score is modified using equation (2).
   - When it is preceded by dilating hedges, the sentiment score is modified using equation (3).
   - When it is preceded by modifier hedges, the sentiment score is modified by swapping both positive and negative sentiment score.
   - When it is preceded or followed by noun feature, the sentiment score of the adjective feature is assigned.
   - When the bigram features are adjectives, the sentiment score is determined by averaging sentiment scores of these two adjective features.
4. Features are clustered into SP, WP, WN and SN using equation (4).
5. Aspect- multi polarity sentiment lexicon is constructed by extracting the aspects from both bigram and trigram features and its corresponding multi polarity sentiment are identified using clustered results.

4. **EXPERIMENTS AND RESULT ANALYSIS**

The proposed work employs two different data sets reviews such as car reviews and Amazon data set reviews for evaluation of aspects – multi polarity sentiment lexicon. To construct aspects- multi polarity sentiment lexicon for car domain, 5,000 reviews are collected from the various car products web sites by web crawling. Next, pre processing techniques such as POS tagging and feature extraction are applied to the online reviews. Finally, aspect- multi polarity sentiment lexicon is constructed by computing sentiment score for all the features based on linguistic hedges present in the reviews. Table 3 shows the sample results in aspects- multi polarity sentiment lexicon constructed by proposed framework using online reviews of the cars. This aspect – multi polarity sentiment lexicon is automatically created whenever giving the input reviews. Using this lexicon, the intelligent robot identifies the strong negative and weak negative features.

<table>
<thead>
<tr>
<th>Table 3. Sample of Aspect-Sentiment Lexicon</th>
</tr>
</thead>
<tbody>
<tr>
<td>Riding quality</td>
</tr>
<tr>
<td>comfort</td>
</tr>
<tr>
<td>Mileage</td>
</tr>
<tr>
<td>Spare parts</td>
</tr>
<tr>
<td>price</td>
</tr>
<tr>
<td>Interior</td>
</tr>
</tbody>
</table>

A. **Evaluation of Aspects- multi polarity Sentiment Lexicon Quality**

There is no existing data set available to validate the quality of aspects- multi polarity sentiment lexicon and hence the proposed work evaluates the effectiveness of the aspect- multi polarity sentiment lexicon as compared to manually constructed aspects- multi polarity sentiment lexicon.

1) **Car Reviews Description**

To evaluate the proposed work, 5000 online reviews about a well-known car models are collected from various web sites such as carwale, cartrade, gaddi, and autoportal. Each review has an overall rating between 1 and 5 stars in addition to the review texts. They also specified rating ranges from 1 to 5 stars for 10 aspects of the car model such as interior, comfort, exterior, mileage, spacing, riding quality, gear shifting, spare parts, breaking and maintenance.

2) **Amazon product Data set reviews**

The aspect based sentiment analysis is applicable to all domain reviews. The proposed work also uses Amazon data set Reviews to construct aspects- multi polarity sentiment lexicon for each domain. Amazon product reviews are bench mark data set consisting of four different product types such as Books, DVDs, Electronics
and Kitchen appliances. Each review is assigned with rates ranging from 1 to 5 stars. This benchmark data set has been used in previous work (Sanju et al., 2013). The data sets structure is shown in Table 4. For each domain, there are 1000 positive reviews, 1000 negative reviews and also some unlabeled reviews.

**Table 4: Amazon product data sets**

<table>
<thead>
<tr>
<th>Domain</th>
<th>Positive</th>
<th>Negative</th>
<th>Unlabeled</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kitchen</td>
<td>1000</td>
<td>1000</td>
<td>16746</td>
</tr>
<tr>
<td>DVDs</td>
<td>1000</td>
<td>1000</td>
<td>34377</td>
</tr>
<tr>
<td>Electronics</td>
<td>1000</td>
<td>1000</td>
<td>13116</td>
</tr>
<tr>
<td>Books</td>
<td>1000</td>
<td>1000</td>
<td>5947</td>
</tr>
</tbody>
</table>

The present work has chosen 1000 reviews from each domain and constructed aspect–multi polarity sentiment lexicon.

**B. Manual construction of Aspects- Multi polarity sentiment Lexicon**

To evaluate the aspects–multi polarity sentiment lexicon, 500 reviews sample were randomly selected from each domain and lexicon was constructed by manual. Similarly, the same reviews sample were used by the proposed to construct aspects–multi polarity sentiment lexicon. Finally, the lexicon constructed by the proposed framework was validated with manually constructed lexicon.

**C. Evaluation Measures**

The proposed aspects–multi polarity sentiment lexicon quality is evaluated by various measures such as precision, Recall and F-measure.

\[
\text{Precision} = \frac{N_{\text{correct}}}{N_{\text{lexicon}}} \tag{5}
\]

\[
\text{Recall} = \frac{N_{\text{correct}}}{N_{\text{manual}}} \tag{6}
\]

\[
\text{F measure} = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \tag{7}
\]

Where \(N_{\text{correct}}\) represents the total number of aspect-multi polarity sentiment pairs are correctly labeled in both proposed sentiment lexicon and manual lexicon. \(N_{\text{lexicon}}\) represents total the number of aspect-multi polarity sentiment pairs in proposed lexicon. \(N_{\text{manual}}\) is the total number of aspect-multi polarity sentiment pairs in manual lexicon. Table 5 shows the lexicon quality evaluation for various domain Reviews such as Car, DVD, Books, electronics and Kitchen appliances reviews.

**Table 5. Lexicon quality evaluation of various domain reviews**

<table>
<thead>
<tr>
<th>Domain</th>
<th>Precision</th>
<th>Recall</th>
<th>F-Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Car</td>
<td>0.7834</td>
<td>0.7532</td>
<td>0.7680</td>
</tr>
<tr>
<td>DVD</td>
<td>0.8318</td>
<td>0.7877</td>
<td>0.8091</td>
</tr>
<tr>
<td>Books</td>
<td>0.7665</td>
<td>0.7345</td>
<td>0.7501</td>
</tr>
<tr>
<td>Electronics</td>
<td><strong>0.8767</strong></td>
<td><strong>0.8277</strong></td>
<td><strong>0.8514</strong></td>
</tr>
<tr>
<td>Kitchen</td>
<td>0.8517</td>
<td>0.8227</td>
<td>0.8369</td>
</tr>
</tbody>
</table>

The aspects–multi polarity sentiment lexicon is constructed for each domain by the proposed method as well as manual. Next, proposed lexicon is validated by various measures such as precision, recall and F-measure. In Table 5, the proposed work obtains high accuracy in electronics domain than other domains. Using this aspect based sentiment analysis system, two different car products reviews such as Honda amaze and Chevrolet sail are analyzed based on its computed sentiment scores of all aspects of the car models. The Fig.2 shows the comparison of aspects of two cars such as Honda and Chevrolet based on its sentiment score. From the analysis, Honda car has best features than Chevrolet car. Similarly, industry people can analyze their different products by using the proposed aspect based sentiment analysis system. This aspect–multi polarity sentiment lexicon has 1257 features for 500 sample reviews. The size of the lexicon vary depends on the number of reviews used to construct lexicon.
CONCLUSION AND FUTURE WORKS

The proposed aspect based sentiment analysis system automatically constructs the aspects- multi polarity sentiment lexicon using the reviews of the products. To construct this lexicon, the unigram, bigram and trigram features are extracted from the reviews after performing the preprocessing functions such as stop word removal and POS Tagging. The sentiment scores are assigned for features from SentiWordNet and scores are modified for bigram and trigram features based on the presence of linguistic hedges. These features are clustered into four categories such as SP, WP, WN and SN by K mean clustering algorithm. Aspects are considered as frequently occurred nouns and they are extracted from bigram and trigram features. Finally, aspects- multi polarity sentiment lexicon is constructed using these aspects and their multi polarity sentiment is determined by using clustered features. Experiments were carried out on online reviews of cars and Amazon data sets to evaluate the performance of the proposed work and it was validated by the manually constructed lexicon. In Future, the proposed work will be extended to summarize the reviews.

REFERENCES


Chihli Hung and Hao-Kai Lin,“Using Objective Words in Sentiwordnet to improve the Sentiment Classification Word of Mouth,” IEEE Intelligent System,vol.28, no. 2, pp. 47-54, 2013, doi: 10.1109/MIS.2013.1


Jo Y, Oh A H. Aspect and Sentiment Unification Model for OnlineReview Analysis Proceedings of the 4th
ACM International Conference on Web Search and Data Mining: February 9-12, 2011.


