An Unsupervised Approach for Feature Based Sentiment Analysis of Product Reviews

Sherin Molly Babu1, Shine N Das2
1MTech Scholar, Department of CIS, College of Engineering Munnar, Kerala
2Associate Professor, Department of CSE, College of Engineering Munnar, Kerala

Abstract
Opinion mining or sentiment analysis is the task of mining polarity of opinions which comprises of the area such as natural language processing, data mining. As E-shops grow at rapidly increasing rates, customers on web write product reviews. These product reviews are useful for a customer in his decision making process on whether to purchase the product. Thousands or Hundreds of reviews may appear for a popular product. It is difficult for a customer to read and analyze large number of reviews and form an opinion on product reviews. In order to solve this problem we use an automated approach to mine polarity of reviews. Feature based sentiment analysis focus on different aspects or features of a product. Features of camera are battery life, zoom etc. Earlier works are based on document level and sentence level opinion mining which give a polarity of the review of the product as a whole rather than considering its features. This paper focus on an unsupervised model for analysis of product reviews at feature level and combined use of co-reference resolution, dependency parsing, linguistic rules, adverb adjective combinations, adverb verb combinations and SentiWordNet together for sentiment analysis.

Keywords - Feature Based Sentiment Analysis, Opinion Mining, Sentiment Analysis, SentiWordNet

1. Introduction

Opinion mining or sentiment analysis consists of natural language processing, computational linguistics and data mining. It is actually a modified version of traditional text classification. Due to the large development in web applications, a large number of user reviews or suggestions on everything are available on the web. Web may contain the reviews of products, services or critic reviews on movies etc. which help other users in their decision making. Now E-shops play a significant role in product marketing. Reviews are increasing in a faster rate because every person likes to give their opinion on the Web and improve the performance of every product on web. Large numbers of reviews are available for a single product which makes difficult for a customer to read all the reviews and make a decision. Thus, mining this data, identifying the user opinions and classifying them are important tasks. These opinionated data are of large amount and hence reading of each opinion and finding the essence of it is a tedious job for a reader. Therefore an automated approach is essential. Using opinion mining technique, we can classify an opinion in natural language as positive, negative or neutral. This is very useful because a lot of user opinions can be considered and classified within a very short amount of time which is impossible for a human to do. So customers and organizations now rely on various automated opinion mining applications. Due to these reasons now opinion mining has become an area of challenging research problems.

Opinion mining can be done in various levels: document level, sentence level, phrase level, feature or aspect level. Feature based sentiment analysis incorporate more fine-grained analysis of various features contained in online reviews. Polarity of various features can be mined using this fine-grained approach. In this paper, we use the importance of adjective adverb combinations and adverb verb combinations in determining the polarity of a sentence.

Moreover we use linguistic rules and co-reference resolution to improve performance of the system.

2. Related Works

In opinion mining, the emotional polarity of a given text is analyzed and classified as positive, negative or neutral and its summary is presented to the user. In this section, we present a review of the existing and related works on Opinion Mining (OM) and Sentiment Classification (SC).

Dave et al. [1] first introduced the term opinion mining (OM) in 2003. In this document level opinion mining is done by processing a set of search results for a particular
entity, generating a list of product attributes or features and summarizing opinions about each of them (poor, mixed, good).

Tetsuya Nasukawa and Jeonghee Yi [2] first introduced the term sentiment analysis (SA) in 2003. This paper illustrates a document level sentiment analysis approach and extracts polarity of specific subjects in documents, instead of classifying the whole document into positive or negative.

Pang, Lee and Vaithyanathan [3] discussed on document level opinion mining with supervised learning. They classified movie reviews into two classes, positive and negative based on the opinion it contains.

Turney [4] proposed a simple unsupervised learning algorithm for classifying a review as recommended (thumbs up) or not recommended (thumbs down). Classifications are done using some fixed syntactic patterns, which are composed on part-of-speech (POS) tags. He computed words point wise mutual information (PMI) and determined whether words are positive or negative. The algorithm was evaluated on 410 reviews from Epinions, sampled from four different domains and achieved the average accuracy of 74%.

Harb et al. [5] presented blog classification by starting with the 2 sets of seed words with positive and negative semantic orientations respectively and used Google search engine to create association rules that find more. Then they counted the number of positive versus negative adjectives in a document to classify the documents.

Wiebe et al. [6] proposed an unsupervised method for subjectivity classification, which simply used the presence of subjective expressions in a sentence to determine the subjectivity of a sentence.

Benamara et al. [7] discussed on the use of adverb-adjective combinations (AACs) in sentiment analysis. They proposed a set of general axioms that all adverb scoring techniques must satisfy. They proposed an axiomatic treatment of AACs based on the linguistic classification of adverbs, instead of aggregating scores of both adverbs and adjectives using simple scoring functions. Three specific AAC scoring methods that satisfy the axioms are presented.

Xiaowen Ding and Bing Liu [8] discussed on polarity of opinions expressed on product features in reviews and used a set opinion words for this analysis. The semantic orientations of many words are context dependent. So they proposed some linguistic rules with a new opinion aggregation function for finding semantic orientation of reviews.

Dietmar Grabner et al. [9] discussed on how to generate a reliable classification approach of customer reviews based on an existing domain-specific corpus by applying a lexicon based sentiment analysis.

Raisa Varghese and M Jayasree [10] proposed aspect based sentiment analysis using support vector machine classifier and SentiWordNet. For the co-reference resolution Stanford Deterministic Co-reference Resolution System was used.

VK Singh et al. [11] proposed a new feature based heuristic for aspect level sentiment classification of movie reviews. They proposed AA AVC algorithm using SentiWordNet to find the aspect score of sentences.

Edison Marrese-Taylor et al. [12] extended the Bing Lius 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.


3. Methodology

3.1 Problem Definition

This work focused on feature based sentiment analysis. Current approaches for sentiment analysis, attempt to detect the overall polarity of a sentence, paragraph or text span regardless of the aspects mentioned in it. The proposed method is focused on the combined use of co-reference resolution, domain specific aspect dictionary, SentiWordNet, linguistic rules, adjectives, verbs, adverb adjective combinations and adverb verb combinations together for sentiment analysis of features. The camera reviews are taken as domain for feature dictionary which contain camera features. The score of each feature is calculated by the proposed model. Finally produce a sentiment profile of features of a product. This sentiment profile can be used by a customer in his decision making purpose.

3.2 System Design

The block diagram of proposed feature based sentiment analysis model is given in the Fig.1.
3.2.1 Reviews Collection:

Reviews about camera from online website amazon.com are collected as the data-set for sentiment analysis. Role of reviews collection module is to download the opinions and reviews about camera from the specified URL. Then these collected reviews are stored in the database.

3.2.2 Pre-Processing:

The second step of the technique involves pre-processing or filtering of reviews, which improve the accuracy and also avoid the unnecessary processing overhead of opinion mining process. The pre-processing steps include stop words removal. Non alphabetic characters like numbers and symbols and smileys are removed before sentiment analysis. This can increase the speed of the opinion mining process.

3.2.3 Feature Extraction:

The aspects or features of a product can appear as a single word or a phrase. For example, picture quality of a camera is one among its features while size is another feature. In this, we have a feature dictionary which is domain specific. We manually add the known features of a camera that is to be needed for the creation of sentiment profile. These features are also stored in the database for further use.

3.2.4 Co-reference Resolution:

Co-reference Resolution is the task of identifying the mentions to entities that they refer to. The Stanford Deterministic Co-reference Resolution System is used for resolving all noun phrases that refer to the same entities. For instance, consider the two sentences given below.

Picture quality of the camera is very good. It is amazing. We could not relate the aspect used in second sentence with the aspect in first sentence without using co-reference resolution. The co-reference resolver produces an output that “Picture quality” in first sentence and “It” in second sentence as it is co-referred. So we replace the pronouns that got resolved, with the corresponding nouns. This replacement is limited to the pronouns that got resolved to aspect names of the product.

3.2.5 Subjectivity/Objectivity Classification:

All the sentences in the reviews do not contain an opinion. A sentence of the review is analyzed only when it contains an opinion. Such sentences are called subjective sentences and non-opinionated sentences are called objective sentences. The subjective sentences should be identified and other (objective) sentences should be removed before the analysis. It helps for avoiding the further processing overhead. This is done using feature dictionary containing feature words. If the sentences taken for this purpose contain the feature words in the feature dictionary, then these sentences are taken as subjective. Then these subjective sentences are taken for further analysis.

3.2.6 Opinion Words Identification:

Opinion words are usually adjectives, adverbs, and verbs which expresses the polarity of a feature of a product in online reviews. These opinion words can express positive or negative polarity of opinions. In-order to identify the opinion words by using syntactic dependency between the aspect word and the opinion word, we use Stanford dependency parser. By the incorporation of Stanford dependency parser, we could get dependency relations of opinion words related to a particular feature in feature dictionary. This parser outputs the adverb, verb and adjectives having dependency relations with the aspect in a sentence.
These opinion words are further used for calculating the polarity of different features of the product in reviews.

3.2.7 Orientation Detection:

In Orientation Detection, positive or negative score of every feature in the reviews are determined. Reviews contain one or more sentences. First we calculate the sentence level score of the opinion by analyzing each sentence in the reviews. By combining all the sentences in a review and calculate the overall score of features in reviews. Then calculate the overall score of each feature by combining all reviews. For calculating the opinion score of each feature, we have used the algorithm in [11] for determining the score of adverb adjective and adverb verb combinations in sentences.

Algorithm: For each aspect in a sentence, extract adverb+adjective and adverb+verb combinations related to that aspect if present.

1) For each extracted adv+adj combine do:
   - If adj score=0, ignore it.
   - If adv is positive, then
     - If score (adj)>0
       \[ f_{adv,adj}=score(adv)+sf*score(adj) \]
     - If score (adj)<0
       \[ f_{adv,adj}=score(adj)-sf*score(adv) \]
   - If adv is negative, then
     - If score (adj)>0
       \[ f_{adv,adj}=score(adj)+sf*score(adv) \]
     - If score (adj)<0
       \[ f_{adv,adj}=score(adj)-sf*score(adv) \]

2) For each extracted adv+verb combine do:
   - If verb score=0, ignore it.
   - If adv is positive, then
     - If score (verb)>0
       \[ f_{adv,verb}=score(verb)+sf*score(adv) \]
     - If score (verb)<0
       \[ f_{adv,verb}=score(verb)-sf*score(adv) \]
   - If adv is negative, then
     - If score (verb)>0
       \[ f_{adv,verb}=score(verb)+sf*score(adv) \]
     - If score (verb)<0
       \[ f_{adv,verb}=score(verb)-sf*score(adv) \]

3) \[ f_{AAAAC}(aspect)=f_{adv,adj}+0.3*f_{adv,verb} \]

Score of adverb, verb and adjective are calculated by using SentiWordNet containing sentiment words. Each word in SentiWordNet is extracted and stored in the database. When an opinion word is encountered, score of that opinion word is taken from the SentiWordNet and used in algorithm to calculate the scores of the combination. If there is no adverb+adjective combination or adverb+verb combination we calculate adverb, adjective, verb score alone for calculating the score of the feature in the sentence. “sf” used in the algorithm is scaling factor and is taken as 0.3 from [11]. Negation rules and but clauses rule from [12] are also used in orientation detection.

3.2.8 Feature Based Summary:

Feature based summary calculates the overall score of each feature in all reviews and produce a Feature based summary. Positive and negative scores of aspects are separately calculated. Hence we get a total positive score and negative score of aspect. By using this result a sentiment profile of camera can be created. The scores of opinions on each feature in all reviews can be calculated using the formula given below.

For each aspect \( j \) of the product,

\[
\text{Total}_\text{Pos}_\text{Score}[j] = \sum \text{PositiveScore}_{i,j} \tag{3.1}
\]

\[
\text{Total}_\text{Neg}_\text{Score}[j] = \sum \text{NegativeScore}_{i,j} \tag{3.2}
\]

In the above formula, the \( \text{PositiveScore}_{i,j} \) is the positive score of the \( j \)th feature in \( i \)th sentence. The \( \text{NegativeScore}_{i,j} \) is the negative score of the \( j \)th feature in \( i \)th sentence. In order to normalize the overall score of the features, we use the following equations.

\[
\text{Norm}_\text{Pos}_\text{Score}[j] = \frac{\text{Total}_\text{Pos}_\text{Score}[j]}{\sum_i} \tag{3.3}
\]

\[
\text{Norm}_\text{Neg}_\text{Score}[j] = \frac{\text{Total}_\text{Neg}_\text{Score}[j]}{\sum_i} \tag{3.4}
\]

4. Results

Result visualization tools like tables and bar charts are employed. Customers can easily identify the overall feature score of a product. Result consists of the representation of scores of each feature of the product. Each feature of the product has a positive and negative score. Both the normalized negative and positive scores of features are shown in table as output as given in the fig. 2. The highest of these scores is taken to visualize result using bar chart. Y-axis of the bar chart is plotted with normalized scores. Bar chart is shown in fig. 3.
5. Evaluation

Dataset used for evaluation of the work is the reviews about cameras from amazon.com. Dataset consists of 300 reviews of three different digital cameras, 100 reviews of each camera. Evaluation measures used in this paper are accuracy, precision, recall. The accuracy, precision, recall varies with number of reviews and dataset used. The method used in [11] is named as AAAVC method. The method used in our system is named as AAAVCCR as it has used AAAVC method and co-reference resolution, SentiWordNet, adverb+adjective and adverb+verb combinations, linguistic rules. Combination of these improves the performance of the feature based sentiment analysis. Automatic aspect extraction is not considered and the sentiment analysis is domain specific.

Table 1: Evaluation of AAAVC method

<table>
<thead>
<tr>
<th>Product</th>
<th>Accuracy</th>
<th>Precision</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Camera1</td>
<td>65%</td>
<td>65.54%</td>
<td>93.97%</td>
</tr>
<tr>
<td>Camera2</td>
<td>71.1%</td>
<td>71.1%</td>
<td>97.95%</td>
</tr>
<tr>
<td>Camera3</td>
<td>70.34%</td>
<td>70.31%</td>
<td>94.73%</td>
</tr>
</tbody>
</table>

Table 2: Evaluation of AAAVCCR method

<table>
<thead>
<tr>
<th>Product</th>
<th>Accuracy</th>
<th>Precision</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Camera1</td>
<td>68%</td>
<td>69.49%</td>
<td>94.25%</td>
</tr>
<tr>
<td>Camera2</td>
<td>74.3%</td>
<td>73.28%</td>
<td>98.2%</td>
</tr>
<tr>
<td>Camera3</td>
<td>73.10%</td>
<td>71.87%</td>
<td>96.84%</td>
</tr>
</tbody>
</table>

From these tables we infer that an average accuracy of 68.81% and 71.8% is obtained from AAAVC method and proposed AAAVCCR method respectively. Approximately, an increase of 3% in accuracy is obtained by AAAVCCR method when compared with AAAVC method.

6. Conclusion

In this paper, we use an unsupervised model for feature based sentiment analysis. So training of data set is not needed in this approach. Previous works using parsing and sentiwordnet are done in movie and tourism domains. This work is done in product domain. This approach make use of the syntactic dependencies of words for orientation detection with the help of co-reference resolution, SentiWordNet, adverb+adjective and adverb+verb combinations, linguistic rules. Combination of these improves the performance of the feature based sentiment analysis. Automatic aspect extraction is not considered and the sentiment analysis is domain specific.

6.1 Future Work

Score of explicit features can be determined but score of implicit features cannot be obtained by this model. Future work includes mining of polarity of implicit features. Comparative sentences have major significance in reviews of product domain. Analysis of comparative sentences can also be included in future work.

References


