Mining Explicit Features for Opinion Mining of Customer Reviews

Md. Safdar¹, Dr. Md. Jawed Ikbal Khan², Md. Daiyan³
¹,³Research Scholar, Inst. of Information Sciences & IT M. U. Bodh-Gaya, Gaya, Bihar, India
²Assistant Professor, Mirza Ghalib College, Gaya, Bihar, India

Abstract—Due to introduction of web 2.0 peoples are more familiar with internet. People buy the product and express their opinion in merchant website like amazon.com and opinon.com. There are various Web sites containing such opinions, e.g., consumer reviews of products, forums and blogs. Merchant website like amazon.com provides facility to give their opinion about the product. In this paper we purposed a new technique for explicit feature extraction. Datasets are crawled data from amazon.com of different modal of mobile. We use SVM and Naive Bayes classifier for our review documents. For SVM use SVMlight, Weka tools for Naive Bayes. Our purposed method result shows that it is very valuable over existing system.

Keywords—Opinion Mining, Feature Identification, Sentiment Classification, Feature Extraction, Supervised Techniques.

I. INTRODUCTION

Opinion mining can be described as a process in which customer reviews are mine for extraction of users opinions expressed on objects and their attributes. Opinion mining also called sentiment analysis is a process of finding users opinion from news, blogs, forum, online shopping etc. Opinion mining are Textual information can be divided into two main categories, facts and opinions. Facts are representing objective statements about entities and Opinions are representing subjective statements that reflect people’s sentiments [1]. Opinion mining express with identifying opinion words, e.g., excellent, amazing, bright, expensive, bad, and poor. Further opinion bearing words are classified into three orientations i.e., positive, negative or neutral. In [2], the authors identified several linguistic rules that can be exploited to identify opinion words and their orientations from a large corpus. In [6], most product features can be bringing into being by exploiting local information and their Parts-of-Speech (POS). As a result, the planned approach implements the information component extraction mechanism as a rule-based system, which applies both linguistic and semantic analysis on review documents. In opinion mining review is to be determined at three levels [4] document level, sentence level and feature level.

In a different way opinion mining at the word level focuses on sentiment polarities of opinion mining for rule based approach is used to categorize the documents based on text reviews as sentiment.

An information component is defined as a triplet format <f; m; o>, where f are represents a feature generally known as a noun phrase, o represents an opinion expressed over f generally recognized as adjective, and m is a modifier generally used to model the degree of expressiveness of o.

In [5] noted that produce feature can be explicit as well as implicit. Explicit features show as noun phrases in the review sentence whereas an implicit feature does not explicitly appear. For example, think the following review sentences wherein product features are represented through underlined with italic and the opinionated words are represented through bold with italic.

a) The camera quality is Excellent.
b) The battery life is amazing.
c) This camera is too small.

In above sentence a; product features camera quality and battery life is explicitly available along with opinionated words excellent and amazing. Extraction of explicit product feature and opinion word pairs from a review sentence is simple, i.e. if a sentence contains explicit feature, its adjacent opinion word can be associated using POS information and dependency relationship. In dissimilarity, many review sentences present opinionated words and the subsequent feature words are not present, i.e. explicit semantic or syntactic relationship between feature and opinion word at the sentence level is missing. For example, sentence b; contains only opinionated word small which is inferring to the implicit product feature size.

II. RELATED WORKS

Opinion mining or sentiment analysis Opinion mining is a type of natural language processing for tracking the frame of mind of the community about a particular product. There are lots of techniques for opinion mining.
Minqing Hu et al.[2,4] mine and to summarize all the customer reviews of a product based on data mining and natural language processing methods. Researcher performs this task in three steps (a.) Mining product features that have been commented on by customers, (b) identifying opinion sentences in each review and deciding whether each opinion sentence is positive or negative and finally (c) is summarizing the results. Experimental results show that author’s purposed techniques very effective.

In [4] author’s focused on mining opinion/product features that the reviewers have commented on. A number of techniques are presented to mine such features. For an experiment dataset are taken from merchant sites like cnet.com and amazon.com. Experiment results shows that authors purposed methods are very effective in performing their task. Bing Liu et al.[3] proposed two techniques i.e. a. novel framework for analysing and comparing consumer opinions of competing products and b. a new technique based on language pattern mining is proposed to extract product features from Pros and Cons in a particular type of reviews. Dataset are collected from opinions.com for an experiment. Author’s experiment results shows that their purposed method is very effective.

In [5] Qi Su et al. Proposed a novel mutual reinforcement approach for feature-level opinion mining problem. This approach clusters product features and opinion words simultaneously and iteratively by fusing both their content information and sentiment link information. Under the same framework, based on the product feature categories and opinion word groups, they construct the sentiment association set between the two groups of data objects by identifying their strongest n sentiment links. Author’s purposed modal provides a more accurate opinion evaluation. Experimental results make obvious researchers method outperforms the state-of-art algorithms.

Peter D. Turney [9] purposed a unsupervised learning algorithm for recognizing synonyms, based on statistical data acquired by querying a Web search engine. They evaluated using 80 synonym test questions from the Test of English as a Foreign Language (TOEFL) and 50 synonym test questions from a collection of tests for students of English as a Second Language (ESL) used Pointwise Mutual Information (PMI) and Information Retrieval (IR) to calculate the resemblance of pairs of words. For both tests, author’s algorithm obtains a score of 74%.

Usually, document-level opinion mining systems not succeed to expose the product features liked or disliked by the customer, slightly they classify the reviews as positive or negative. In general a positive review does not mean that the opinion possessor has positive opinion on all aspects or features of the product.

Correspondingly, a negative review does not mean that the opinion proprietor dislikes everything about the product.

In [8] author purposed a OPINE system which is based on an unsupervised pattern mining approach, extracts explicit product features using feature assessor and web PMI statistics. Author’s results shows that purposed method is very effective over existing system.

Nozomi Kobayashi et al. [7] aspect-evaluation and aspect-of relation extraction. Author used Japanese weblog posts for an experiment. Experimental results shows that author purposed model using contextual clues improved the performance for both tasks.

In [10] A. Shoukry et all show an application on Arabic sentiment analysis for Arabic tweets at the Sentence level in which the aim is to classify a sentence whether a blog, review, tweet , etc. They purposed an approach that differs and improves those existing works. In this approach the pre-processing of the tweets is different from the pre-processing done in Arabic sentiment analysis as different stop words list will be used, particularly built for the Egyptian dialect. This approach uses different machine learning classifier Naïve Bayes and Support vector machine. The feature used is unigram and bigrams. The process starts by getting the tweets from twitter, then passes by each tweet and labels it as positive, or negative. After that the features in each tweet will be extracted and represented in a feature vector. Then, these feature vectors will be used in the training phase of the classifier. For each tweet the following feature vector was constructed using term frequency.

\[ (\text{word}:\text{frequency1}, \text{rd2}:\text{frequency2} \ldots, \text{"polarity"}) \]

However authors find results of SVM and NB in both cases (before removing stop word and after removing stop word) SVM has better results. The improvement between the best accuracy results of both models is almost 4-6% for SVM.

III. DATASET

For experiment we use product reviews which are obtained from merchant sites www.amazon.com. It is very large and covers all range of products. In amazon site, there is a facility for users can evaluate the posted review after the review is posted. Dataset is crawled using crawler4j [11] and pre-processed by some filtering to smoothen the noise and chunking to decompose the text into individual meaning full chunks. Stanford Parser API [12] we used for text chunks are late broken down to crumble the text into individual chunks. . The user can provide achieve to symbolize if this review is helpful, or write comments for the reviews. To review spam, we manually build a review spam corpus. In this work we used different products of mobile and camera.
The Dataset consist of 600 product reviews which are crawled from amazon.com. Each amazon.com’s reviews consist of 8 fields:

- **Product ID**
- **Reviewer ID**
- **Rating**
- **Date**
- **Review Title**
- **Review Body**
- **Number of helpful Feedbacks**
- **Number of feedbacks**

We have used 4 product of digital camera i.e., Canon, Kodak, Nikon, Panasonic, iPhone.

### IV. PROPOSED METHOD

![Proposed architecture for feature extraction](image)

Our proposed method consists of following steps:-

- **A. Pre-process and Stemming data**
- **B. Pos Tagging**
- **C. Feature Extraction**
- **D. Classifier**

#### A. Pre-process and Stemming data;

![Table 1: A partial list of extracted features, modifiers, and opinions](table)

<table>
<thead>
<tr>
<th>Feature</th>
<th>Modifier</th>
<th>Opinion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Camera</td>
<td>compliment, very, right, surprisingly, extremely</td>
<td>perfect, adaptable, easy, light, simple, sharp, heavy, incredible, easy</td>
</tr>
<tr>
<td>Lens</td>
<td>too, quite, very</td>
<td>good, great, fine, heavy, amazing, clear</td>
</tr>
<tr>
<td>Battery</td>
<td>life, too, very</td>
<td>great, long, extra</td>
</tr>
<tr>
<td>Pictures</td>
<td>Very</td>
<td>good, excellent, awesome, so-so, outstanding, poor, glorious, great, excellent, fantastic</td>
</tr>
</tbody>
</table>

#### D. Classifier;

Classifies the message or review documents are feature, modifier and opinion using supervised classification techniques.

It is very important steps in text mining, takes more than 60 percent of the process. In our dataset which is crawled from amazon.com mostly data are incomplete, noisy and inconsistent, and is probable to contain many errors. Data pre-processing is a demonstrated method of resolving such issues. So that Data must be pre-processed in order to perform any data mining functionality. Data pre-processing includes Data cleaning, data integration, Data transformation, Data reduction and data discrimination. Kotsiantis et al. (2006) present a well-known algorithm for each step of data pre-processing [13].

#### B. POS (Parts-of-Speech) Tagging:

In this paper main aims to find explicit feature that appear explicitly as nouns phrases from the product review. So we use the parts of speech tagging. In this phase takes record size chunks generated by pre-processor as input to assign parts-of-speech tags to each word. For pos tagging we used Stanford parser [11].

#### C. Feature Extraction;

In a review feature and opinion words that is used by user to express positive and negative opinions. In a text categorization for a classification we have to care about steps. It is a very tedious task because if features extractor has been made in spite of the context, whatever astounding classification algorithm the accuracy will be always not good enough. Noun phrases are generally product feature adjective and adverbs are generally refers to opinion and modifiers respectively [7]. Stemming and fuzzy matching we use to take care of word variants and misspellings. In this way we build up feature, opinion and modifier list.
In a review data classification is a two-step process. In the first step, a classification algorithm builds the classifier by “learning from” a training set made up of our corpus and their associated class labels. In a second step, the model is used for classification. A different set called test set is used to evaluate the correctness of the built model.

We used 10-fold cross validation calculate the recital of our system. Our train dataset is divided into 10 sub samples with the same number of instances.

V. EXPERIMENT

For an experiment we used customer reviews of mobile products which are crawled from merchant site amazon.com. Each of the product reviews includes text review and title. We used 300 review documents on different model of mobile in our experiment.

The features are generally explicit in customer review opinion sentences. For example “The picture quality is very excellent”. We conduct the experiments using and SVM light tools to train Naïve Bayes and Support Vector Machine classifiers respectively. The data has been described in Section 3. We divided the data set into training set and test set and conducts 10-fold cross-validation: the data set is randomly split into ten folds, where nine folds are selected for training and the tenth fold is selected for test.

We determine experimental results using standard Information Retrieval (IR) metrics Precision, Recall and F-score that are defined in equations 1, 2, and 3 respectively. Where TP indicates true positive FP indicates false positive and FN indicates false negatives.

\[
\text{precision} = \frac{TP}{TP + FP} \quad (1)
\]

\[
\text{recall} = \frac{TP}{TP + FN} \quad (2)
\]

\[
F - score = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (3)
\]

In Table 2 we present the result summary found for our dataset of different products of mobile. Our results show that our purposed method is very effective over existing methods.

<table>
<thead>
<tr>
<th>Product</th>
<th>TP</th>
<th>TN</th>
<th>FN</th>
<th>Precision</th>
<th>Recall</th>
<th>F-Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nokia</td>
<td>73</td>
<td>35</td>
<td>28</td>
<td>65.66</td>
<td>69.74</td>
<td>69.73</td>
</tr>
<tr>
<td>Sony</td>
<td>71</td>
<td>29</td>
<td>23</td>
<td>71.67</td>
<td>71.23</td>
<td>72.21</td>
</tr>
<tr>
<td>LG</td>
<td>61</td>
<td>28</td>
<td>19</td>
<td>61.44</td>
<td>55.74</td>
<td>58.67</td>
</tr>
<tr>
<td>Samsung</td>
<td>51</td>
<td>23</td>
<td>29</td>
<td>66.75</td>
<td>68.66</td>
<td>68.15</td>
</tr>
</tbody>
</table>

Fig 2: A comparison of Precision, Recall and F-score of different mobile models

VI. CONCLUSION AND FUTURE WORK

In this paper we purposed a technique for explicit feature of review documents. Data are crawled from amazon.com and Support Vector Machine and Naïve Bayes classifier. Our experiments results show that our method is very effective over existing method. In future work, we will improve our results and we will work on implicit features.

REFERENCES


A. Shoukry, A. Rafea, “Sentence- Level Arabic Sentiment Analysis”, 978-1-4673-1382, 2012 IEEE.

http://code.google.com/p/crawler4j/.

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