ISSN 1870-4069

Opinion Summarization Using Product Feature Based Review Classification

K Vimal Kumar[1], Pawan Upadhyay[1], Abheet Kumar Gupta[2], Deepanshu Wadhwa[3]

[1]Jaypee Institute of Information Technology, India

[2]Valuefirst Digital Media Private Limited., India

[3]Morgan Stanley, India

vimal.kumar@jiit.ac.in, pawan.upadhyay@jiit.ac.in, {abheet5080, deepanshu1404}@gmail.com

Abstract. As the e-commerce websites are expanding on a large scale, the number of reviews of products are increasing at a very rapid pace along with it. The e-commerce websites as well as manufacturers ask their customers explicitly to review their product. It helps both the fellow customers and the manufacturers to get a better knowledge about the product along with the related services that they are going to offer/get. So, with the increasing count of reviews, it becomes difficult for a user searching for a particular product to analyze all the reviews of that product in a very short period of time and make an informed decision on whether to purchase the product or not. In this research paper, we aim to mine and provide a sentiment analysis of the reviews received for a product. The proposed work is different from others as the aim of the system is to provide a sentiment insight of the features of the product rather than the complete review/sentence and visualize it properly to provide an easy way for the customers to make a proper decision. Our approach is based on tools and techniques from the fields such as opinion mining, natural language processing, and sentiment analysis.

Keywords. Opinion mining, sentiment classification, part of speech tagging, opinion summarization, reviews

1 Introduction

With the rapid increase in the e-commerce and the number of products sold online, the number of reviews for these products keeps increasing at an exponential rate. These reviews act as a good source of information both for the manufacturers and the customers. However, the millions of data obtained in the form of reviews make it difficult for the customers to segregate the product reviews on the basis of their polarity (i.e. Positivity or Negativity) and to make a wise decision while buying the product. The overall rating of the product gives the customer an insight of the overall view of the product but not about its features. The customer still has to go through the large quantity of reviews to get feature-based polarity. The proposed method has been introduced to address these issues in a very efficient way. It deals with:

a) Extracting the polarity of the product reviews on the review level as well as feature level.
b) Summarization of the product features on the basis of their polarity.

Unlike the existing approaches that classify the reviews on the basis of polarity obtained on the product level, the proposed approach classifies the reviews on the feature level and not just on the sentence level. The current approaches extract the most common phrase/noun existing in the review to find the features of the product.

The proposed approach differs in the way that it extracts the features from the most popular e-commerce websites and use them as features. If any of the features or the most similar word to that feature appears in the review, they are used to find the opinion of users about that feature.

2 Related Work

The proposed work is closely related to the work done by Minqing Hu et al. [2], DuyKhang Ly et al. [3] on product review summarization from a deeper perspective. The existing method uses POS Tagger to identify the part-of-speech and further uses association rule mining to identify frequent explicit product features.

It then performs usefulness pruning on it to remove the useless words from the corpus. The method then identifies the sentences based on their positive and negative score. Based on the positive and negative score, the system summarizes the complete review. Proposed work is different from the existing in three ways:

i) The existing approach uses POS tagging, association rule mining, usefulness pruning to identify the product facets out of a given review related to the product. On the contrary, the proposed method identifies the frequent features of the product by scraping the features from various leading e-commerce websites. Furthermore, the features are added to the feature dictionary of the product based on the category in which the product lies. This saves a lot of processing time in the later processing phases.

ii) Apart from the technique that the existing method is using to score the opinion words, the system uses a well-defined corpus to identify feelings from the opinion word as happy, sad, arrogant, or excited. Thus, each of the opinion word will be scored not just on the basis of positivity or negativity that the word possesses, but according to various feelings that human beings possess.

iii) The opinion words are extracted from the reviews not just by identifying the nearest adjective word as proposed in the existing algorithm, but the proposed method introduces bigram and trigram approach to find the opinion word. This method first assumes the opinion word to exist as a bigram, then as a trigram

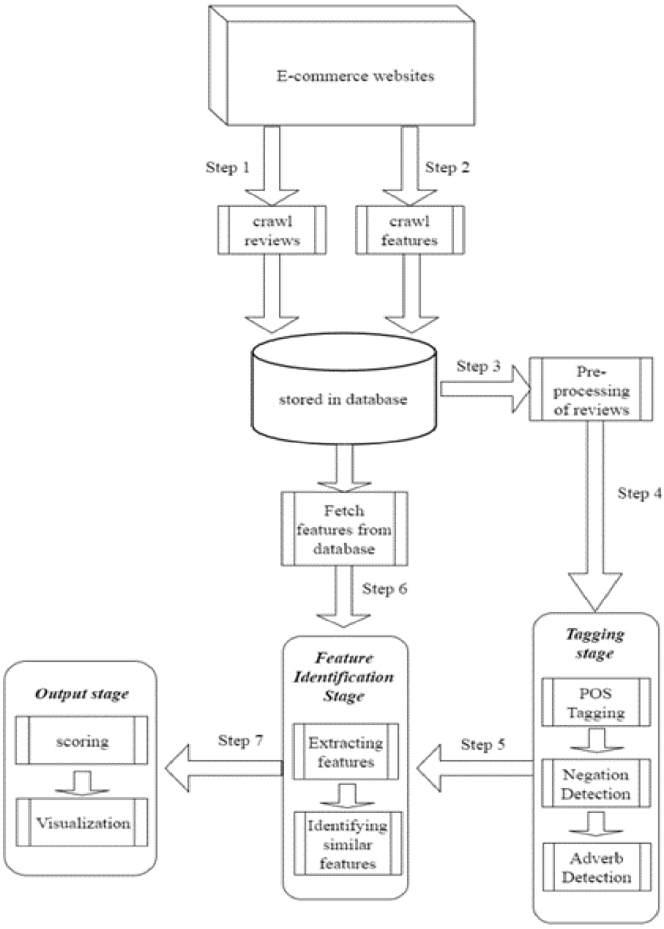

Fig. 1. The proposed architecture.

and so on up to n-gram. The opinion word is thus found and used to score the feature accordingly.

The existing work identifies features that appear explicitly as noun or noun phrase. For example, "The lens is awesome". Here lens is the direct feature obtained. In this proposed system, both the explicit feature and implicit features are identified and in the implicit feature identification, most semantically similar features are found from the predefined feature list using cosine similarity between the vectors.

3 Proposed Technique

Figure 1 shows the detailed architecture of the proposed system. The method considers product names as input and gives its detailed summary of each feature of the product

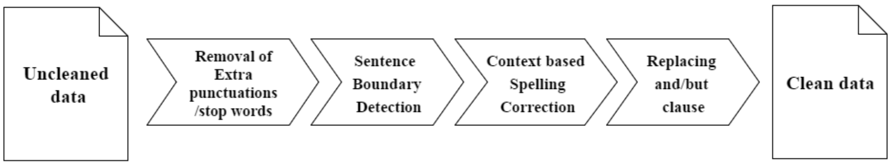

Fig. 1. Preprocessing steps.

Input Sentence:- the movie was awesome but the cast was not so good.

Output :- [('the', 'DT'),('movie', 'NN'), ('was', 'VBD'), ('awesome', 'JJ'), ('but', 'CC'), ('the','DT'), ('cast', 'NN'), ('was', 'VBD'), ('not', 'RB'), ('so', 'RB'), ('good','JJ')]

Fig. 3. Example of POS Tagging.

as output along with its polarity. The system performs the opinion generation of reviews as well as the features of the product in mainly five steps.

3.1 Review and Feature Extraction

The reviews of the product searched are extracted from different e-commerce websites along with their key features. This step gives us the advantage over the method proposed by Minqing Hu et al [2], D. Ly et al [3], in which the method does not identify the frequent feature dataset for the product and thus optimizing/reducing the steps involved in the process.

The step also involves general feature addition common to all products across the product category to the above feature dataset extracted from the websites. The extracted features as well as reviews are stored in the database which enables us to reduce the re-extraction of features every time a product is to be summarized.

3.2 Preprocessing Stage

Preprocessing stage is shown completely in figure 2. It consists of mainly four steps to clean the data. To describe about the preprocessing stage, consider the input sentence as:

"*The product has amazng features but its size is toooo big…!!!!!!*"

The reviews are taken and fed into first block which removes the extra punctuations and stop words out of it.

Output: *Product has amazng features but its size is toooo big.*

The second step involves sentence boundary detection in which different sentences in the same reviews are merged as a list.

Output: *Product has amazng features but its size is toooo big.*

Next step involves spelling/slang correction to correct any incorrect spellings and slangs used in the reviews.

Output: *Product has amazing features but its size is too big.*

Final step involves but/and clause removal and breaks a complex sentence into two individual sentences.

Output: *Product has amazing features. Its size is too big.*

3.3 Review Tagging

After the preprocessing of reviews, the cleaned reviews are fed to the tagging stage. The tagging stage involves generation of a dictionary (*key-value pair*) which has the key as identified tokens of the reviews and values as a list of properties related to a token. The three properties are:

– Part of Speech Tagging,
– Negation Detection,
– Adverb Detection.

3.3.1 Part of Speech Tagging

Part of speech tagging refers to tagging each token of the sentence according to the part of speech represented by it (whether the token is a noun, a verb, a determiner, etc.). It is required so as to remove the insignificant part of speech as well as in the later feature extraction phase. It forms the list in the form of tuples containing individual tokens and their corresponding part of speech tag. Example is as shown in figure 3.

The extracted POS tagged data is further used in the next processing stage. The data is saved in the database in the form of dictionary containing each token of a sentence as a key and the value as a list containing POS tag along with other factors mentioned later.

3.3.2 Negation Detection

It is very important to detect the negation in the sentence because it can reverse the polarity of the adjective (used to identify the polarity of the sentence). Generally, the negating word is present before the opinion describing word (i.e. adjective), therefore the process involves the detection of negating word with a help of predefined list of negating words and adding a flag "neg y" to all the tokens value list till the end of the sentence. Else a "neg_n" flag is appended to the list to indicate that negating word is not present up to that token in the review. A list of negation words is maintained in advance and the tokens are compared with the list to find if the negation words are present or not.

3.3.3 Adverb Detection

The adverb's present before an adjective can have a very good impact on the sentiment expressed by a sentence. Adverbs can drastically increase the sentiment score as well as decrease it. The adverbs are extracted from the sentence through the part-of-speech tagged data and will be checked whether the adverb is an adverb of degree or not. If yes, that particular adverb is taken and checked further whether the adverb is a strong intensifying adverb or a weak intensifying adverb. For example, 'Very good' in a sentence will have a better score as compared to 'good'. So, adverb detection becomes a mandatory part of the process.

3.4 Feature Identification

This step involves identifying the key-category of features of the product on which the people have expressed their opinion. Before discussing the methods involved in identifying the features, let us first look at the kinds of examples that'll be handling. The aim of the proposed system is to find what feature of the product is liked or disliked by the user. It is a very important and crucial step to identify the feature about which the people are expressing their views. Let us discuss an example of a digital camera:

"*The lens of this digital camera is awesome.*"

In the above sentence, it is visible that a reviewer is satisfied with the lens of the digital camera; lens is the key-feature in this review that the reviewer has mentioned about. Whereas it can be observed from the sentence that the feature of the product is explicitly mentioned, there are some reviews where the features are hidden or implicit and hard to find. For example,

"The camera doesn't fit in the pocket."

The challenge in the above example is to identify the key-feature (i.e., size or dimension) of the camera about which the user has expressed his views. Thus, the proposed system deploys finding key-category of the feature mentioned in the review by the reviewer. Considering the above example, the key-category of the feature mentioned is dimension. This has an advantage that when it needs to summarize a key-category of feature that can include all those features which lie under a particular key-category of feature. For example, ['size', 'width', 'height'] all will lie under a key-category of dimension and thus saving us from re-computing the reviews which lie under a particular key-category of features in the later summarization stage.

The list of key-category of features is obtained from the same e-commerce websites from which the product reviews are obtained. This helps us in saving the process of defining the key-category of features. The basic idea is to extract all the nouns or noun phrases in the reviews and find the semantically similar key-category of feature from the obtained list, which is performed in the Word embedding stage that follows

3.4.1 Word Embedding Stage

The key-category of semantically similar features are identified in this stage. Word2vec creates vectors that are distributed numerical representations of word features, such as the context of individual words. The main advantage of Word2vec is that it performs the grouping of those vectors that are of similar words in vector space.

That is, it detects similarities mathematically. The method employed in the word2vec stage to obtain the key-category, involves finding the cosine similarity between the vector of keyword to be categorized and one by one vector of elements present in key-category list.

This similarity would result in a similarity score and if the similarity score is greater than a threshold value (in our case it was 0.30) and maximum score among all key-categories, then the review will be categorized to a particular key-category of features.

The nouns or noun phrases which are used to categorize the review to a particular key-category of features, are saved for identification of opinion words in the next step (III.E).

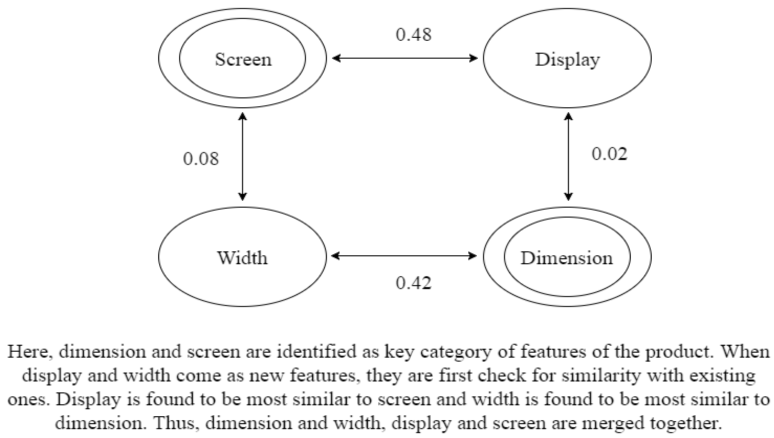

Fig. 4. Similar feature identification.

3.5 Opinion word extraction and summarization

The next task is to identify the opinion words. These opinion words give a clear view of what a particular person thinks about the product as well as about the particular feature of the product.

The method involves the identification of adjectives which will be used as the opinion word in this proposed method. For each key-category of feature, its opinion word is extracted and saved for computation purpose in further stages. Let's first see how the opinion words are extracted (as shown in Figure 5).

Let's understand the above algorithm with the help of an example:

The image quality of this camera is poor.

```
for each review in the review database:
  if (it contains a  particular feature keyword recorded in the dictionary
    in the previous stage)
    for each feature determining keyword in the dictionary:
      extract the nearby adjective to be used as an opinion word for
      that particular feature
```

Fig. 5. Steps for opinion mining.

Here if feature extraction is applied on the above example, a dictionary will be created which will contain the specific feature-word present in the review (e.g. {review1: [quality]}). This feature-keyword from the dictionary will be used in this stage to extract the nearby opinion word (adjective) which in above example is poor.

3.5.1 Score Calculation and Visualization

The opinion word obtained in the previous stage are saved for each review and for each feature, which helps in giving the opinion summary at sentence level as well as at product feature level.

A dictionary of all the words tagged as positive and negative is taken and for each opinion word in the review, the opinion word is looked up in the dictionary for classifying the word as positive or negative, negation detection is done to identify the exact polarity of the opinion word (negation can revert the polarity of the opinion word) and finally adverb detection is applied which can increase the score of a particular opinion word, if present. The pseudo code mentioned below shows how the scoring is done:

```
Procedure ReviewScoring():
    positive = 0; negative = 0; neutral = 0;
    for each review ri
      score_review = 0;
      for each opinion word op in ri
          score_review += OpinionWordScore(op)
          if (score_review> 0) positive += 1;
          else if (score< 0) negative += 1;
          else neutral += 1;
      end..for;
    end..for;
end
```

For feature level Opinion identification, the opinion word obtained for word is passed into OpinionWordScore (opinion_word) function and the obtained score is used to identify the polarity of the review.

```
Procedure OpinionWordScore(word):
  Score = 0;
  if word in dictionary:
     if word in dictionary is tagged as positive:
       score = 1;
     else
       score = -1;
     end..if;
     if adverb_word is an adverb_of_degree:
       if adverb_word is strong intensifying adverbs
          Score = Increase (Score)
       else
          Score = Decrease (Score)
       end..if;
     end..if;
     if (there is NEGATION_WORD that appears closely
          around word in sentence)
          orientation = Opposite_magnitude(Score);
     end..if;
 end..if;
end
```

The next step is opinion summarization which involves the summarization of overall user review's opinion as well as feature level product opinion. The score from opinion word about a particular feature present in the review is calculated as either positive, negative or neutral. Percentage of each is calculated and is used to visualize the data in form of pie chart or gauge chart. The frequency (occurrence) of all the features in reviews is calculated and is used to rank the features. The feature with a greater number of reviews is ranked higher and is given more preference as compared to others. This is done to throw light on that feature which is highly talked about by the customers.

Table 1. Categories used for testing.

	Tagged as positive	Tagged as negative
Predicted as positive	TP	FP
Predicted as negative	FN	TN

4 Result Evaluation

The proposed system has been implemented in python language through the use of various Natural Language Processing techniques. This system is evaluated by crawling reviews from different leading e-commerce websites. Next, these reviews are crowd sourced to find the opinion of people towards various features present in the review. Then, the results were compared to find the accuracy of the proposed method. The main difficulty in the review opinion generation is the presence of sarcasm in the reviews. The reviews can be subjective and the satire in the reviews make them difficult to judge whether a review is positive or negative. The experimental results are shown in the Table 1 & 2 in the form of accuracy, precision, and recall.

Table 2. Result of the proposed system.

Accuracy	81.5%
Precision	0.8558
Recall	0.8079

5 Conclusion and Future Work

In this research paper, the proposed method tends to obtain the features in a review and the opinion of people towards those features. The main objective is to get the viewpoints of people for the products sold online and generate a report of the features along with opinions of people towards them. This is beneficial both for the customers and the manufacturers selling their products online. The results of the experiments suggest that the algorithm for the proposed method works quite efficiently.

Our system finds the features of the products stored as nouns and merges similar features to make them single. The future work plans to identify the solutions to the problems yet not covered in the proposed work. Sentences with sarcasm will be checked for the identification of judgment of people towards the product. Furthermore, strength of adverbs is not identified efficiently in the proposed method. We also plan to improve the efficiency for finding the key-category of features from the reviews containing the implicit features.

References

1. Mishra, N., Jha, C.K.: Classification of Opinion Mining Techniques, International Journal of Computer Applications, vol. 56, no. 13 (2012), pp 1–6. DOI:10.5120/8948-3122
2. Hu, M., Liu, B.: Mining and Summarizing Customer Reviews. In: Proceedings of the ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining (KDD04), pp. 168–177 (2004)
3. Ly, D., Sugiyama, K., Lin, Z., Kan, M.: Product review summarization from a deeper perspective. In: Proceedings of the Joint Conference on Digital Libraries, pp. 311–314 (2011)
4. Kokkoras, F., Lampridou, E., Ntonas, K., Vlahavas, I.: Summarization of multiple, metadata rich, product reviews. In: Workshop on Mining Social Data (MSoDa), 18th European Conference on Artificial Intelligence (ECAI '08), Patras, Greece (2008)
5. Mugdha, M., Bharat T.: A framework for summarization of online opinion using weighting scheme. In: Proceedings of the Pervasive Computing (ICPC). International Conference, pp 1–6 (2015)
6. Krishna, G.R., Jothi S., Minojini, N., Sowmiyaa, P.: Survey of Various Opinion Mining Approaches, in Proceedings of the International Journal of Innovative Research in Computerand Communication Engineering. vol. 3, no. 3 (2015)
7. Kumar, P., Mohd, H.: Analytical study of Feature Extraction Techniques in Opinion Mining. In: International Conference on Advance in Computing and Information Technology, page 85–94 (2013)
8. K. Ly, K. Sugiyama, Z. Lin, M. Kan.: Product review summarization based on facet identification and sentence clustering. CoRR'11

9. Vinodhini, G., Chandrasekaran, RM.: Sentiment Analysis and Opinion Mining: A Survey. International Journal of Advanced Research in Computer Science and Software Engineering, vol. 2, no. 6 (2012)
10. Brisson, L., Jean-Claude, T.: Opinion mining on experience feedback: A case study on smarphones reviews, Conference: RCIS 2015. In: IEEE 9th International Conference on Research Challenges in Information Science (2015). doi: 10.1109/RCIS.2015.7128879
11. Dalal, M.K., Zaveri, M.A.: Semi supervised learning bases opinion summarization and classification for online product reviews. Applies Computational Intelligence and Soft Computing, vol. 2013 (2013)