Ranking of products based on online customer reviews for a particular Aspect

Nitindeep Singh
nitindeep19069@iiitd.ac.in
Akshaj Anil Patil
akshaj19111@iiitd.ac.in
Ganesh Vilas Chaudhari
chaudhari19116@iiitd.ac.in

1 Problem Definition

With the hype of online shopping services from a quite long period, there has been seen tremendous increase in the online customer reviews for products. The data of reviews is increasing daily with exponential growth. These reviews are very important for users as well as corporates and if utilized properly and efficiently, then can help in better decision making as well as production decisions for the producers.After classification of the customer reviews, the products must be ranked but the ranking of the products has to be specific for a specific aspect.Taking an example that the user needs top k products which have best display, so here the aspects is display and the ranking should be done based on the reviews for that particular aspect. The problem is to design a IR system that takes aspect as an input parameter and the system outputs top k(choice of the developer) products which will have high rating for that aspect according to the customer reviews.

2 Introduction

Online reviews contain massive information that can be useful in many ways, one they can be used by other customers for decision purpose and also they can be used for understanding of business competition in the market. Ranking of products based on customer reviews is a task of information retrieval field and natural language processing. The reviews of the customers for a product are analyzed and classified into positive, negative and neutral classes using sentiment analysis and the polarity of the reviews are used for the classification. Then using various parsers mostly used being the Stanford parser we can perform the aspect extraction and relative sentiment analysis of the reviews. One of the important task is to identify important aspects. [6] Let's take an example of iPhone 8 mobile phone, it has many aspects like battery, processor, screen, speaker, camera etc. Only some of the aspects are very vital in effective decision making process and customers often feel overwhelmed by the complexity of the scattered reviews. These reviews should be computed in such a manner that producers as well as consumers can easily find out the important aspects and their reviews for the products because manual identification of aspects of a large number of items is not practically possible therefore we go for the automated approach by building a system that can identify important aspects efficiently and categorize the negative and positive feedback for a product based on a particular aspect.

3 Literature Review

There are various research articles on product ranking based on customer reviews on a specific aspect which have been considered before building the system. Various approaches that have been used in the existing systems are Hidden Markov Model, Conditional Random Fields approaches, Shallow dependency parser and much more. They collect user feedback for the products and the reviews are then displayed to other customers and producers. Some systems do provide feedback through scores (range of 1 to 5, 5 being excellent and 1 being very poor) which helps in showing the average score for a product but these score cannot be taken as exact measure to decide whether the product is good or bad for certain important aspects.This system has been build using hybrid approach that combines the efficient techniques from some of the existing systems.

Issues in Existing Systems:

- Manual identification of important aspects is almost practically impossible
- Accurate Classification of reviews as positive or negative is missing

- Aspect based polarity in the reviews is not available
- Users find it very difficult to take decision based on the scattered reviews.
- Ranking of products based on reviews for a particular aspect is missing.

4 Background

There is no doubt that today we have a large number of reviews present on a product on different platforms whereas this large set of data further increases the complexity in an effective decision making process. The users need a clear and classified review about a product on the basis of aspects whereas the current scenario about the systems present is somewhat very less, they only provide scattered reviews for the products and that too not related to any aspect. Using the reviews from the current systems, people often end up buying the wrong product because the ranking is overall and not aspect specified, therefore there is a strong need of a system that can classify the reviews with more accuracy and ranks the aspects of the product according to the opinion of the user for that aspect in the feedback, final rank of the product is based on the ranking score of the aspects of the product which gives a clear picture of the aspect performance for the product.

5 Dataset Used

The data-set that we have used is Multi-Domain Sentiment Dataset. The Multi-Domain Sentiment Dataset contains product reviews taken from Amazon.com for 4 domains: Electronics, Books, DVDs, and Kitchen. Each domain has several thousands of reviews. Reviews contain star ratings (1 to 5 stars).The data set is available in XML format, we have converted it to CSV and pre-processing has been done on the data-set.Several useless columns have been deleted and the columns that are finally kept are Unique product id, product name, product type, rating, title, review text. The link for the data set is given below:

https://http://www.cs.jhu.edu/~mdredze/datasets/sentiment/domain_sentiment_data.tar.gz

6 Proposed System

We propose a system that ranks the products on the basis of customer reviews for a particular aspect.

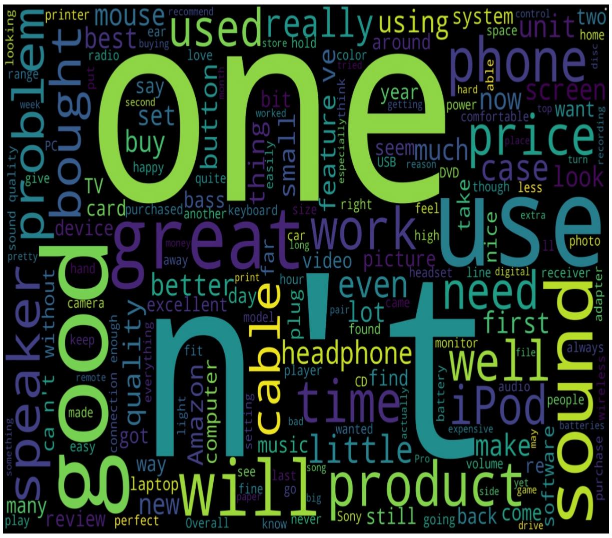

Figure 1: Word Cloud of the Reviews in Dataset

The system extracts the important aspects from the free text customer reviews using both the review title as well as the review text. The sentiment analysis methods and libraries are used which classify the reviews as negative, positive and neutral. After evaluating the polarity of the aspect for the product, static scoring is done which ranks the product based on that aspect.

7 Proposed Algorithm

The system flow of the IR system being built is shown using the block diagram below.
The Steps of the algorithm are:

- The procedure starts from collecting consumer reviews for each product.
- After that manual annotation of the reviews is done aspect wise for every product. This forms the ground truth of the system.
- After fetching the reviews, average ranking of the products is calculated using the ratings given by the user.
- Using the Stanford parser and parts of speech tagging, for each product the adjective noun pairs are extracted from the reviews.
- Then the aspect count/frequency and the opinion score of the aspect is calculated for both review title as well as review text.
- Finally, static scoring of the aspects is done using the score of aspects from review title text,

score from review body text and the average rating of the product.

- Using these final scores, aspects are ranked based on aspect frequency and the influence of aspect opinion using the opinion score on the overall opinion of the product.

- The final list of the top 30 products for entered product in the output is shown.

8 Methodology of the baseline created

There are various ways and techniques by which the aspect based ranking of the products can be done using the user feedbacks. Some of the effective and powerful techniques in sequence of development of this efficient ranking system following the algorithm discussed above are discussed below:

8.1 Average Product Ranking

The products are assigned an average product ranking based on the ranking given by the users. For e.g. there are 5 reviews about iPhone X and the ratings given are 4,5,3,4,5, then the average product rating given will be 21/5 = 4.2. This average product rating is used in the calculation of the final ranking score of the product.

8.2 Identification of Aspect

The user reviews are in the form of free text and they need to be parsed in order to get the aspects of the products. The popular POS (parts of speech) technique is used to get noun and noun phrases from the reviews and Stanford parser is also used to fetch aspects from the free text reviews. Then these extracted aspects contain a lot of noise in them hence synonym clustering can also be done in the process so that unique aspects are fetched from the reviews [11]. For e.g. the review being "good speaker", here the noun is speaker which is also an aspect that needs to be fetched. The aspects with high RF (Review Frequency) (count of in how many reviews this aspect has been mentioned and reviewed by the user) are kept using a threshold of the RF of the aspect and the rest of them are ignored.

8.3 Sentiment Analysis of the reviews and polarity with respect to aspect in the review.

While fetching the aspects, using the POS tagging the noun adjective pairs are fetched. The polarity of the aspects is computed using the adjectives used to describe the aspect in the review. The Text Blob library of python is used to classify the reviews as positive or negative using its sentiment function that returns two values polarity and subjectivity. Polarity value ranges from [0,1] in which 1 means positive and 0 means negative. Subjectivity value also ranges in [0,1] and refers to whether the opinion is public or it is factual information. Additionally, another parallel process is being done which is checking whether there is a negation to the adjective used for the aspect in the review, if found then the polarity of the aspect is multiplied by -1. For e.g. review text being "not good speaker" hence the polarity of good speaker is multiplied by -1 because not is being used before good. Any negative word will be taken in consideration.

8.4 Opinion Score Maximization algorithm

Till the precious step we have calculated the opinion score of the aspects where we first referred the opinion words that describe the aspects and then we also considered the negation or inversion words to the opinion words that will negate the opinion score of the aspects of a product. This process is done separately for review title as well as the review body and finally we have two scores for an aspect of a product i.e. body score and title score. Now for the final computation of the scores we will introduce the average product rating that was computed in the beginning. We have developed two models, one that takes into account title score and the other without accounting title score in the final computation. **The expression for the final computation of scores is:**
Model 1:

$$\alpha.TitleScore + \beta.BodyScore + Avg.rating$$

,where α and β are weightage coefficients that provide importance / weightage to the title score and body score respectively of the review for a product. The values of alpha and beta are 0.6 and 0.4 respectively

Model 2:

$$FinalScore = BodyScore + Avg.rating$$

9 Results and Evaluation Matrix

The following are the evaluation measure graphs of the system for both the model using aspects(quality, sound, memory, display, screen):

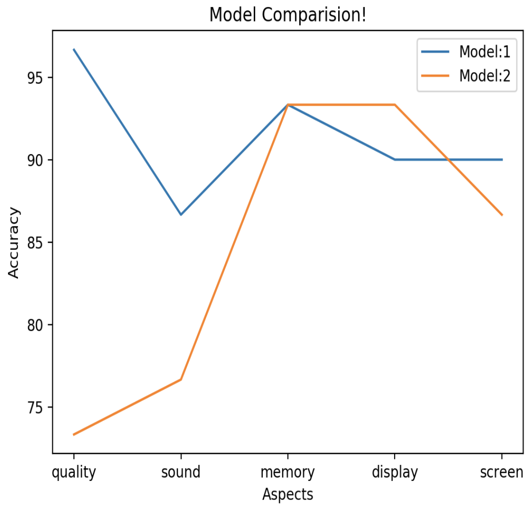

Figure 2: Accuracy of both models

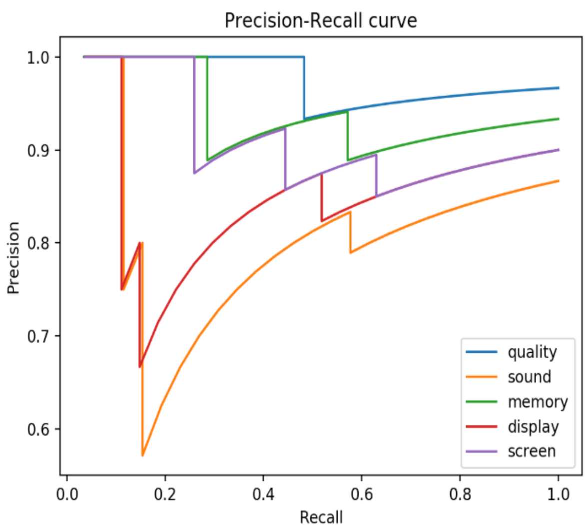

Figure 3: Precision Recall Curves for Model 1

10 Conclusion and Further Scope

As we can clearly see in the accuracy graph for both the models, the model 1 that takes into account title text opinion score performs better than the model 2. The accuracy values either remain same or mostly increase in the model 2. Giving

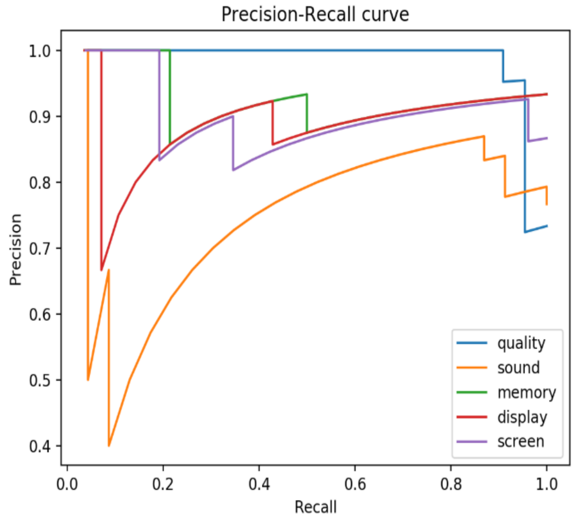

Figure 4: Precision Recall Curves for Model 2

```
***** Aspect: quality *****
tn= 0   tp= 22
fp= 0   fn= 8
Accuracy= 73.33333333333333

***** Aspect: sound *****
tn= 0   tp= 23
fp= 0   fn= 7
Accuracy= 76.66666666666667

***** Aspect: memory *****
tn= 0   tp= 28
fp= 0   fn= 2
Accuracy= 93.33333333333333

***** Aspect: display *****
tn= 0   tp= 28
fp= 0   fn= 2
Accuracy= 93.33333333333333

***** Aspect: screen *****
tn= 0   tp= 26
fp= 0   fn= 4
Accuracy= 86.66666666666667
```

Figure 5: Confusion Matrix and Accuracy Score for Model 1

the title text a weightage of 60 percent is therefore a better decision as the title of the review has literally more influence on the product rating. The system build is very efficient when working with large datasets whereas some parts of the implementation will differ when the structure of the datasets is different. Different aspect extraction techniques that are not described in this developed system can be used according to the data format of the reviews

```
***** Aspect: quality *****
tn= 0   tp= 29
fp= 0   fn= 1
Accuracy= 96.66666666666667

***** Aspect: sound *****
tn= 0   tp= 26
fp= 0   fn= 4
Accuracy= 86.66666666666667

***** Aspect: memory *****
tn= 0   tp= 28
fp= 0   fn= 2
Accuracy= 93.33333333333333

***** Aspect: display *****
tn= 0   tp= 27
fp= 0   fn= 3
Accuracy= 90.0

***** Aspect: screen *****
tn= 0   tp= 27
fp= 0   fn= 3
Accuracy= 90.0
```

Figure 6: Confusion Matrix and Accuracy Score for Model 2

to be computed. Machine learning algorithms and techniques can be integrated for classifying the reviews into different classes and the model can be trained using existing old datasets to increase the efficiency of the model when tested for current online review data available. The k fold technique can be used in the classification task by splitting the train data into test and train portions. The system build using the above-mentioned methodology is capable of classifying the user reviews into three classes, positive negative and neutral and the system extracts important aspects from the reviews and their relative opinion from the user feedback. The aspects are then ranked for the product and the overall rating is also computed using the aspect score and the review polarity average for the product.