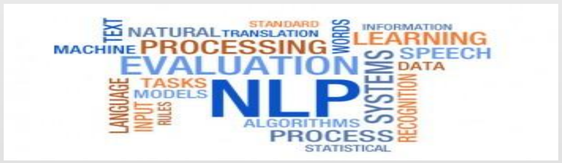
Ranking of Products Based on Customer Reviews on Particular Aspect.

INDRAPRASTHA INSTITUTE of INFORMATION TECHNOLOGY DELHI

NitinDeep Singh (MT19069)
Akshaj Patil (MT19111)
Ganesh Chaudhari (MT19116)

Abstract

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 corporate and if utilized properly and efficiently, then can help in better decision making as well as production decisions for the producers. This paper explains various approaches for ranking of products based on customer reviews and that too for some specified aspects. Various Information Retrieval and Natural Language Processing techniques for the process. The paper aims to provide best knowledge for building a system that can use the customer reviews to rank products according to a specific aspect of it.



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.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.

Motivation

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.

Proposed System

By the means of this review article, we propose a system that ranks the products on the basis of customer reviews for a particular aspect. The system extracts the important aspects from the free text customer reviews using the aspect frequency and the influence of aspect opinion on the overall opinion of the product. The sentiment analysis methods and libraries are used which classify the reviews as negative, positive and neutral.

POS Tagging is used to identify important aspects and their corresponding adjectives, and w.r.t corresponding adjective polarity is assigned to that aspect.

Methodology

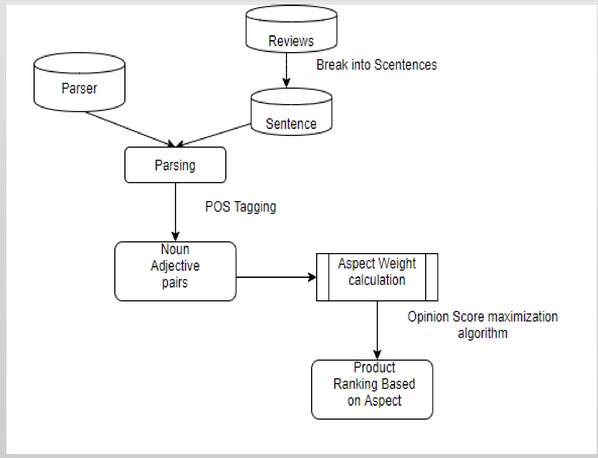
. Using the Stanford parser and parts of speech tagging, for each product the adjective noun pairs are extracted from the reviews.

.Then the aspect count/frequency and the opinion score of the aspect is calculated for both review title as well as review text.

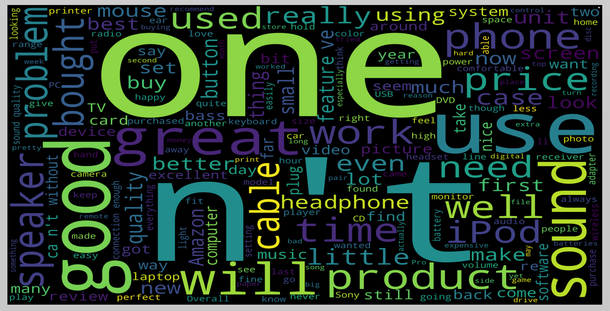
.Finally, static scoring of the aspects is done using the score of aspects from review title text, score from review body text and the average rating of the product.

.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.

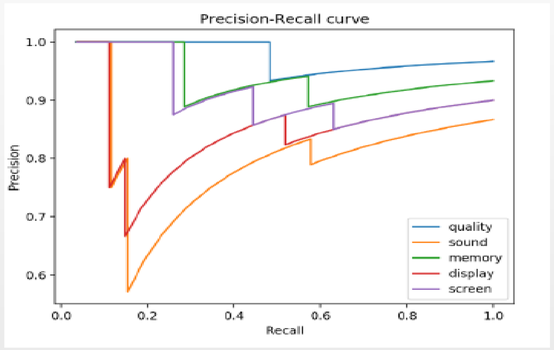
.The final list of the top 30 products for entered product in the output is shown.



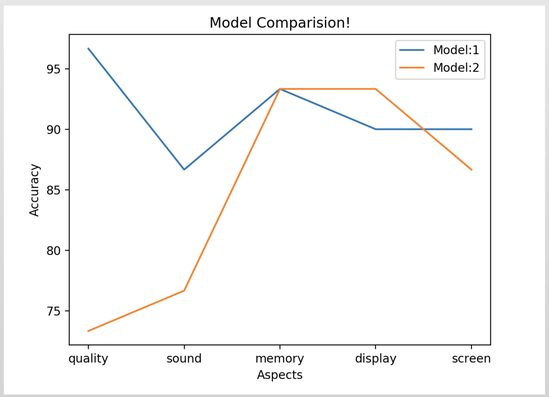
Word Cloud (Review Dataset)



Results



PR Curve is plot for mentioned 5 aspects. We can see that value of PR curve is increasing at the end.



Comparison of accuracy score between two models is plotted. Model 1: Considering Review Title. Model 2: Not Considering Review Title.

Conclusion

Model 1 considers Title and gives 60% weightage to title and 40% weightage to body, and Model 2 does not consider title. From above graph we can see that Model 1 performs better. By Model 1 efficient ranking of products is done according to aspect based in product reviews.