

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 12, December 2015

A Survey on Product Aspect Ranking

Charushila Patil¹, Prof. P. M. Chawan², Priyamvada Chauhan³, Sonali Wankhede⁴

M. Tech Student, Department of Computer Engineering and IT, VJTI College, Mumbai, Maharashtra, India¹

Associate Professor, Department of Computer Engineering and IT, VJTI College, Mumbai, Maharashtra, India²

Asst. Professor, Department of Computer Engineering and IT, VJTI College, Mumbai, Maharashtra, India³

Asst. Professor, Department of Computer Engineering and IT, VJTI College, Mumbai, Maharashtra, India⁴

ABSTRACT: Consumers used to review regarding many products which are available on-line. These reviews are valuable and important for both firms and users. But many times it seems that these reviews are often disorganized making it difficult to do information navigation and knowledge extraction. So here product aspect ranking framework is proposed so as to identify the important aspects and improve the usability of reviews. There are 2 ways for important product aspect identification as: 1) they are commented on by a large number of consumers and 2) customer's opinions on the important aspects influence their overall opinions on that particular product. In particular, if reviews of the product are given by the customer, initially it identifies the product aspects by a standford parser then determines consumer opinions via sentiment classification. After that probabilistic aspect ranking algorithm will be used.

KEYWORDS: product aspect, sentiment classification, aspect ranking framework, probabilistic aspect ranking algorithm.

I. INTRODUCTION

In recent years there is rapidly expanding e-commerce. For example, any shopping site has indexed more than millions of products. Amazon.com archives a total of more than 36 million products. Shopper.com records more than five million products from over 3,000 merchants. Many of the websites encourages consumers to write reviews so that they can express their opinions on various aspects of the products such as usability, efficiency regarding many factors. Here, an aspect, also called feature which refers to an attribute of a certain product. Generally, a product may have many aspects. So, identifying important product aspects will improve the variability of reviews and is beneficial for consumers as well as firms. So consumers can make good purchasing decision by paying attentions to the valuable aspects, while firms can take care of quality improvement of aspects and thus product reputation will be enhance. However, it is bit difficult for people to manually go for identification of the important aspects of products from reviews. Therefore, a way to automatically identify the important aspects is highly demanded. Huge collections of consumer reviews are available on the Web expressing various opinions on multiple aspects of products. The important reviews are mostly disorganized hence creating problems in knowledge acquisition. To address this problem, product aspect ranking is explored to automatically identify important product aspects or features from on-line consumer reviews.

Paper is organized as follows. Section II describes literature review related to product aspect ranking. Section III presents system analysis with the help of Use case diagram and Activity diagram. After that system design consisting of Architecture and Component diagram is given in Section III. Finally, Section V presents conclusion.

II. RELATED WORK

Product aspect ranking framework is proposed to automatically identifying the important aspects of products from consumer reviews. Also develops an aspect ranking algorithm for the importance of various aspects by exploiting

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 12, December 2015

aspect frequency and the influence of consumers' opinions given to each aspect over their overall opinions on the product.

A. Product Aspect Ranking Framework

Here are the details of the proposed Product Aspect Ranking framework. We start with an overview of its pipeline (Fig. 1) consisting of three main components: (a) aspect identification; (b) sentiment classification on aspects; and (c) probabilistic aspect ranking. Given the consumer reviews of a product, we first identify the aspects in the reviews and then analyse consumer opinions on the aspects via a sentiment classifier then aspect ranking algorithm is applied so as to find out the ranked aspects.



Fig. 1: Product Aspect Ranking Framework

In overall terms, the ratings on some Websites might be a little higher or lower than those on others. Moreover, different Websites might offer different rating range. Hence, we here normalize the ratings from different Websites separately, instead of per-forming a uniform normalization on them.

B. Product Aspect Identification

Consumer reviews are composed in different formats on different forum Websites. CNet.com require consumers to give an overall rating on the product, describe positive and negative opinions on some product aspects, also write a paragraph of detailed review. Websites like Viewpoint, only ask for an overall rating and free-text review paragraph. The others such as Reevo.com require complete overall rating. So we can say a consumer review consists of positive and negative reviews also free text review, or both.

For the Pros and Cons reviews, we need to identify the aspects by taking out the noun terms. It is shown that aspects are either nouns or noun phrases, and we can obtain accurate aspects by extracting frequent noun terms from the reviews. For aspects identification in the free text reviews, a solution is to employ an existing aspect identification approach. So first identify the nouns and noun phrases in the documents. Further the occurrence frequencies are counted and only the frequent occurring are kept as aspects.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 12, December 2015

Recently phrase dependency parser had been used to extract noun phrases, which form candidate aspects. To filter out the noises, a language model is used by an intuition that the more likely a candidate to be an aspect, the more closely it related to the reviews. The model was used on product reviews, and hence used for prediction of candidate reviews. The candidates with low scores are then filtered out. However, such language model might be biased to the frequent terms in the reviews and can't precisely sense the related scores of the aspect terms, as a result cannot filter out the noises effectively.

C. Sentiment Classification on Product Aspects

The task of analysing the sentiments on aspects is called aspect-level sentiment classification. Existing methods includes the supervised learning approaches and the lexicon-based approaches, which are typically unsupervised. The lexicon-based methods use a sentiment lexicon consisting of a list of sentiment words so as to determine the sentiment orientation on each aspect. The supervised learning methods used to train a sentiment classifier then used to predict the sentiments on each aspect. Many natural language learning-based classification models are applicable. Supervised learning is dependent on the training dataset and difficult to perform without training data set. Here labeling training data is time-consuming. In this work, the Positive and negative reviews have categorized positive and negative opinions on the aspects. These reviews are important training samples for learning a sentiment classifier. We thus exploit Pros and Cons reviews to train a sentiment classifier, which is in turning used to determine consumer opinions (positive or negative) on the aspects in free text reviews.

D. Probabilistic Aspect Ranking Algorithm

This is about a probabilistic aspect ranking algorithm to identify the important aspects of a product from consumer reviews. Generally, valuable aspects have the characteristics as they are mainly commented in consumer reviews and consumer's views on these aspects greatly influence their overall opinions on that particular product. The final opinion in a review is a generalization of the opinions given to specific aspects in the review, and various aspects have different contributions in the aggregation. That is, the opinions on important aspects have strong (weak) impacts on the generation of overall opinion.

E. Product Aspects Ranking

Once the set of product aspects is identified, we propose to order them according to their relevance. In this sense, we apply a methodology for modeling product aspects from a collection of free-text customer reviews. The proposal relies on a natural language modeling framework which is domain independent. Finally, we used this methodology for ranking our set of product aspects. Given a collection of customer reviews about a specific product and a free-text document d , which can be a sub collection of reviews or an individual review, our goal is to obtain a model for retrieving the product aspects from it. Specially, we consider modeling the set of aspects discussed in d as a statistical language model that assigns higher probability values to words defining aspects.

In the context of customer reviews, opinion words (e.g. "good", "bad", etc.) usually express sentiments about the valuable aspects of a product. Because of that the review texts to respect some entailment relationship from opinion words to aspect words. In this way, we consider the use of an (stochastic) entailment-based self-translation model between the words in the to reveal the probability distribution of words that approaches the language model of aspects expressed in d from a general probabilistic model of opinion words.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 12, December 2015

III. SYSTEM ANALYSIS

A. Use Case Diagram

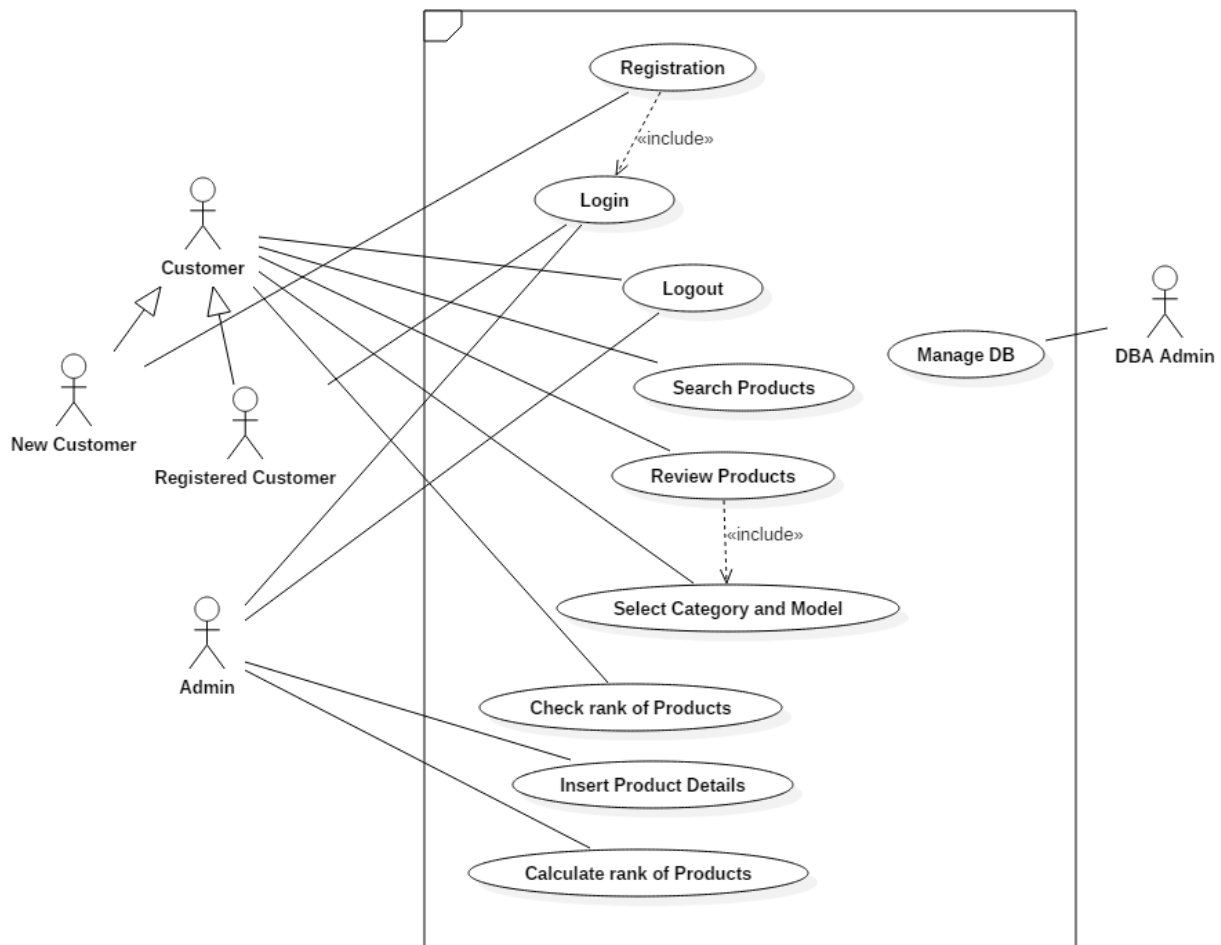


Fig. 2: Use Case Diagram

Use Cases description:

Use case ID: UC 1.0

Use case Name: Registration

Actor: New Customer

Description: Customer needs to register first if he wants to do login.

Precondition: None

Post Condition: Customer is able to login.

Use case ID: UC 1.1

Use case Name: Login

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 12, December 2015

Actor: Registered Customer
Description: Customer is able to login.
Precondition: Customer must register first.
Post Condition: Customer is able to access home page.

Use case ID: UC 1.2

Use case Name: Logout
Actor: Registered Customer, Admin
Description: Customer is able to logout.
Precondition: Customer or Admin must have logged in.
Post Condition: Customer or Admin is able to access index page.

Use case ID: UC 1.3

Use case Name: Search Product
Actor: Customer
Description: Customer can search the products.
Precondition: Customer must have logged in.
Post Condition: Customer will get information of products.

Use case ID: UC 1.4

Use case Name: Review Product
Actor: Customer
Description: Customer is able to review product.
Precondition: Customer must have logged in and category and model must be selected.
Post Condition: Customer is able to fill different aspects for review.

Use case ID: UC 1.5

Use case Name: Check rank of product
Actor: Customer
Description: Customer can check rank of products.
Precondition: Customer must have logged in.
Post Condition: Customer is able to see rank based on respective aspects.

Use case ID: UC 1.6

Use case Name: Insert Product details
Actor: Admin
Description: Admin is able to insert respective product details.
Precondition: Admin must have logged in.
Post Condition: Admin is able to check different aspects for detailing information.

Use case ID: UC 1.7

Use case Name: Calculate rank of product
Actor: Admin
Description: Admin is able to calculate rank of product.
Precondition: Admin must have logged in.
Post Condition: Admin is able to find rank and do other activities.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 12, December 2015

B. Activity Diagram

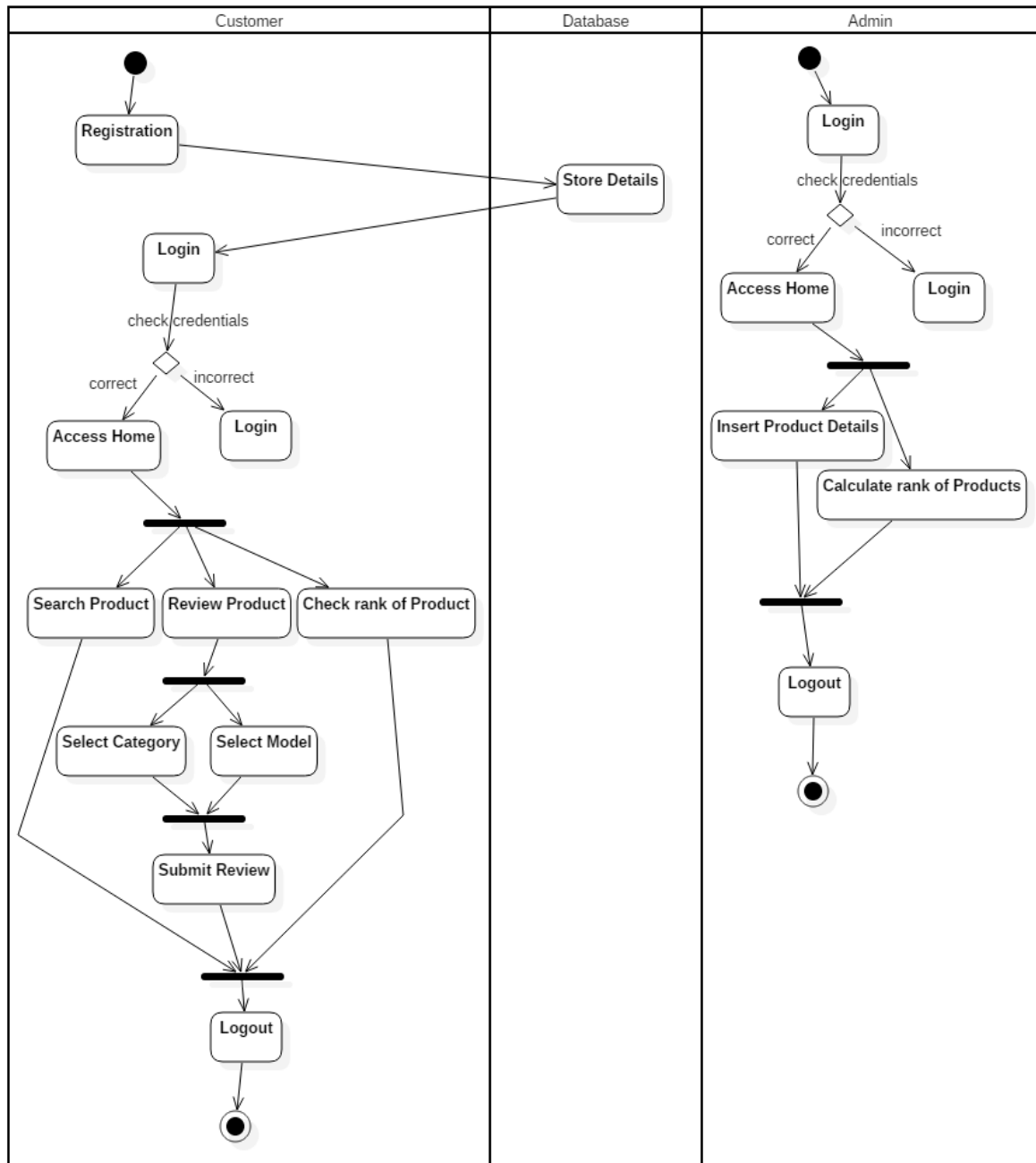


Fig. 3: Activity Diagram

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 12, December 2015

Activity diagram description:

- Initially, customer will do registration and details will be stored in database.
- As soon as customer will do login credentials will be checked, and if credentials are correct then customer will be able to access home page and if not then it will be redirected to the login page itself.
- After accessing home page, customer will be able to search products, review products and check rank.
- For reviewing products customer need to select category and model and then all the reviews are submitted.
- Now customer can logout or continue.
- As soon as admin will do login credentials will be checked, and if credentials are correct then admin will be able to access home page and if not then it will be redirected to the login page itself.
- After accessing home page, admin will be able to insert product details, calculate rank.

IV. SYSTEM DESIGN

A. System Architecture

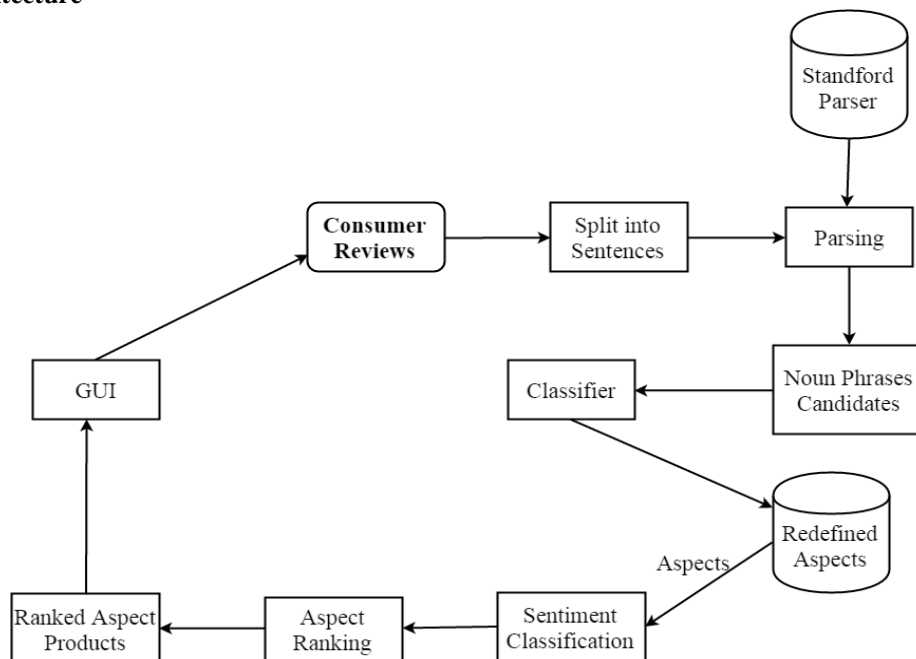


Fig. 4: System Architecture

Architecture description:

- GUI:** Graphical User Interface consists of consumer reviews which will be further split into sentences and parsing is done on that.
- Consumer Reviews:** Contains valuable information which will provide appropriate knowledge to the parser so that it will help for the product's aspect ranking.
- Parsing:** Parsing is done with the help of parser and noun phrase candidates are separated which will be further passed to the classifier.
- Classifier:** Classifier takes noun phrase candidates as input and further it stores redefined aspects into Database.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 12, December 2015

5. **Redefined Aspects stored in Database:** Redefined aspects are stored in Database and further passed for the sentiment classification and from that aspect ranking is done.
6. **Sentiment Classification:** Sentiment classification is a special task of text classification whose objective is to classify a text according to the sentimental polarities of opinions it contains. e.g., favourable or unfavourable, positive or negative.
7. **Aspect Ranking:** Based on the output of the sentiment classifier aspect ranking is done.
8. **Ranked Aspect of Products:** Now ranked aspects of products are displayed on GUI.

B. Component Diagram

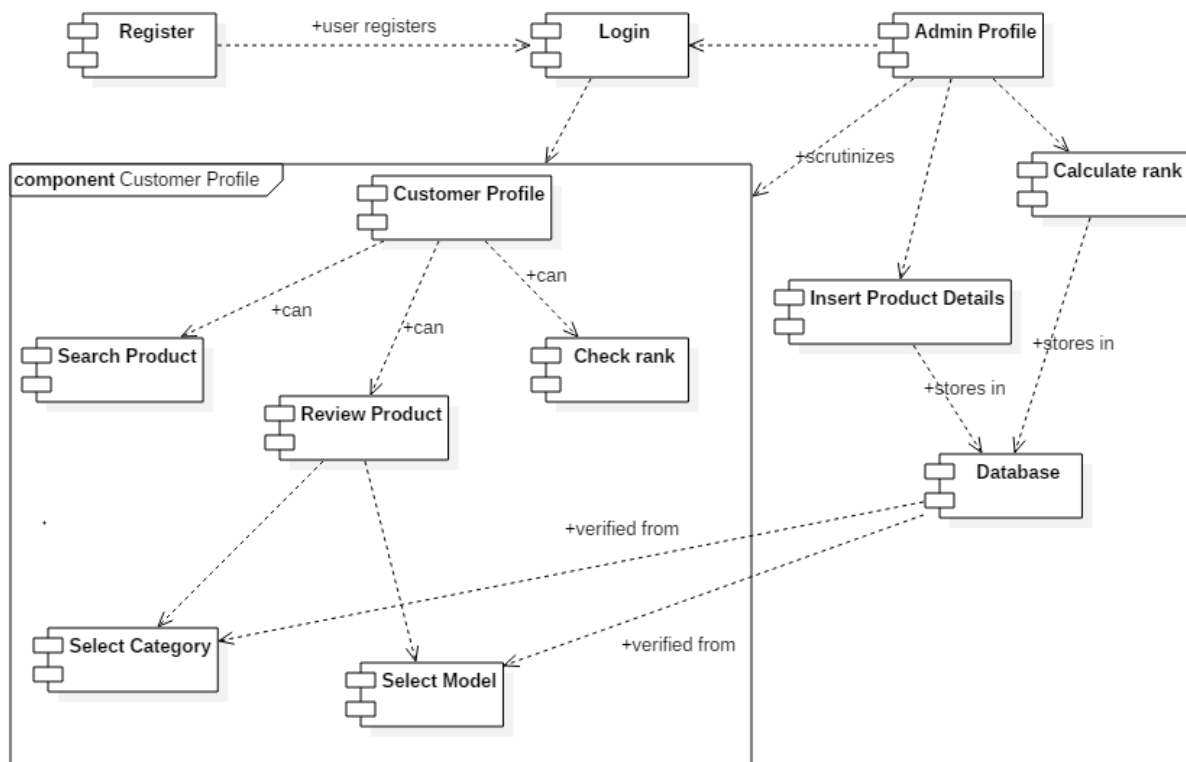


Fig. 5: Component Diagram

Component diagram description:

- Components of the system like register, login, admin profile, customer profile.
- Further, customer profile can search products, review products, check rank.
- Admin profile scrutinizes customer profile and also can insert product details and calculate rank.
- Product and rank details are stored in database.
- Category and model details are verified from database required at the time of submitting reviews.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 12, December 2015

V. CONCLUSION

In this project, It's been proposed a framework to identify the important aspects of a product from on-line consumer reviews. And develop an aspect ranking algorithm to identify the important aspects by simultaneously considering the aspect frequency and the influence of consumer's opinions given to each aspect on their overall opinions. Further, apply the ranking results to the application of sentiment classification, and will try to improve the performance.

So, it's been planned to identify the important aspects of a product from online consumer reviews. Our supposition is that the important aspects of a product should be the aspects that are frequently commented by consumers and consumers' opinions on the important aspects greatly pressure their overall opinions on the product. Based on this assumption, will try to develop an aspect ranking algorithm which will identify the important aspects by concurrently considering the aspect frequency and the pressure of consumers' opinions given to each aspect on their overall opinions.

REFERENCES

- [1] http://www.academia.edu/9236895/Product_Aspect_Ranking_and_Its_Applications
- [2] <https://prezi.com/mhcrxfw7rq-h/product-aspect-ranking/>
- [3] <http://en.wikipedia.org>
- [4] <http://www.slideshare.net/papithavelumani/product-aspect-ranking-and-its-applications>
- [5] Zheng-Jun Zha, Member, IEEE, Jianxing Yu, Jinhui Tang, Member, IEEE, Meng Wang, Member, IEEE, and Tat-Seng Chua, "Product Aspect Ranking and Its Applications", IEEE Transactions on Knowledge and Data Engineering, Vol. 26, pp. 1211-1224, may 2014.
- [6] Jianxing Yu, Zheng-Jun Zha, Meng Wang, Tat-Seng Chua, "Aspect Ranking: Identifying Important Product Aspects from Online Consumer Reviews" Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics, pp. 1496-1505, June 19-24, 2011.
- [7] S. Revathi Manju, E. V. R. M. Kalaimani, R. Bhavani, "Product Aspect Ranking Using Semantic Oriented Sentiment Classifier", IJSER, Volume 2, pp. 25-28, October 2014.
- [8] Shahuraj Patil, Jyoti Raghatwan, "An Efficient Product Aspect Ranking and its Application: A Review", IJSR, Volume 3, pp. 2048-2051, December 2014.