iJRASET
International Journal For Research in Applied Science and Engineering Technology

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

**Volume:** 7 **Issue:** VII **Month of publication:** July 2019

**DOI:** http://doi.org/10.22214/ijraset.2019.7095

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177*
*Volume 7 Issue VII, July 2019- Available at www.ijraset.com*

# Overall Web Rating into Single Standard Five-Scale Rating

P. Sharmila[1], J. Ilakkiya[2], S. Hemalatha[3], B. Ishwarya[4]
[1, 2, 3, 4]*Information Technology, Sri Sairam Institute of Technology*

***Abstract:*** ***Product recommendation is the key factors for convincing the customer in a right way, to buy the products in any e-commerce website. The online customer ratings are based on each customer actual experience which can incorporate meaningful and more user-oriented information. The recommendation might be based on cost, feature, previously purchased brand and many more. In this project we aim to enhance the recommendation service by evaluating standard rating out of all the ratings available in various e-commerce website for any particular product. The result is that the number of products will be reduced up to one third of total number of products available under the category the user is willing to purchase product. Finally very few products will be recommended to the user by applying user-provided filtering features and single five-scale rating.***
***Keywords: LDA, API, HFT***

## I. INTRODUCTION

The online consumer reviews are common and play a significant role in today's electronic commerce. For example, Amazon.com allows users to post their reviews to the web page of each product. The mobile application marketplaces like Google Play Store and Apple App Store offers user reviews on each update of every app. Travelers heavily rely on the online review information such as Trip Advisor to select their accommodations, and the room-sharing platform Airbnb also provides reviews written by the guests of each property. The online consumer reviews covey real experiences of using the product or staying at a hotel to help potential customers evaluate before placing the order. Unlike the product descriptions advertised by the vendor which aims to entice consumers to order and may not reflect all facts of products, the online consumer reviews are based on each consumer's actual experience which can incorporate meaningful and more user-oriented information.

## II. RELATED WORK

Collaborative filtering is widely implemented in recommender systems for e-commerce websites. Collaborative filtering utilizes the item (product) ratings received from users (consumers) to measure the similarity between users and recommends the items liked by similar users. Collaborative filtering suffers from the sparseness of the user-item rating matrix since most users only rate a small part of available items [10]. Content-based filtering represents user interests by the content of items they have shown interests. The items with similar content to the target user's interest are to be recommended [1].

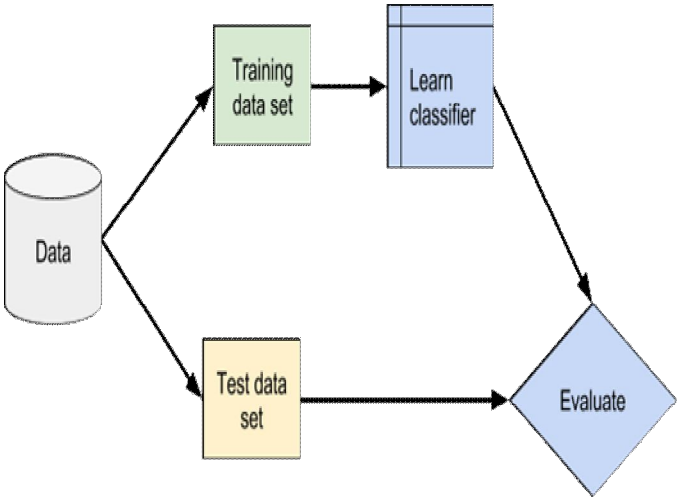



Fig1:Rating in existing systen

The context-aware recommender system [2] takes relevant contextual information such as time, location or click stream into account in the recommendation process to improve the performance. Consumer reviews are direct feedbacks of using the products, which not only provide opinions to the vendors for improving their products but also give suggestions to the potential customers. The work of Hu and Liu [6] proposed a consumer reviews summarization approach based on data mining and natural language processing methods to help the potential customers and vendors quickly read the huge amount of reviews. The user reviews have been utilized in recommender systems to help alleviate the sparsity and cold-start problem [4], [11], [13]. Among which the work of McCauley et al. [11] is closely related to our work. They applied LDA to learn latent topics from reviews, and combine the topics with hidden factors learned f r o m item ratings to develop a hidden factors as topics (HFT) model to predict the rating of an item for the target user.Existing works of review-based recommender system mostly focused on solving the sparsity problem of ratings in collaborative filtering. In the case of Airbnb where most users tend to give a high rating to the property, the ratings are less informative for collaborative filtering.

## III. PROPOSED WORK

This method uses a kind of content-based filtering approach which relies only on the rating content. The topic modeling is performed on consumer ratings to derive the features of products, and the preferences of consumers are inferred from topics generated from the past rating history of the consumer for making recommendations.

### *A. Rating Algorithm*

Consider a scenario where only 2 people bought a product and has given 3 out of 5 stars as rating. The overall rating for that product will now be 3 which is calculated as follows:

[(Total number of 5 star *5)+ (Total number of 4 star *4)+ (Total number of 3 star*3)+ (Total number of 2 star *2)+ (Total number of 1 star *1)] /(Total number of ratings)]

The above formula is used to calculate overall rating of any product out of 5 scale. This rating is displayed along with the product to provide a quick judgment over the quality of the product.

### *B. WEB Scraping*

Data extraction is data scraping used for extracting data from websites. Web scraping software may access the World Wide Web directly using the Hypertext Transfer Protocol, or through a web browser.

This scraped data will provide the overall rating information. This data is combined into a single standard rating. Web pages are built using text-based mark-up languages (HTML and XHTML), and frequently contain a wealth of useful data in text form. However, most web pages are designed for human end-users and not for ease of automated use. A web scraper is an Application Programming Interface (API) to extract data from a web site. The scraped data will be loaded to the database. There will be separate columns to store count of five stars, four star, three star, two star and one star rating. Separate column will specify the website from which the count is scraped. Query can be used to calculate the overall rating by applying the traditional formula on the collaborated count. Finally the overall single standard rating will be calculated for the product and filtering or other sorting of

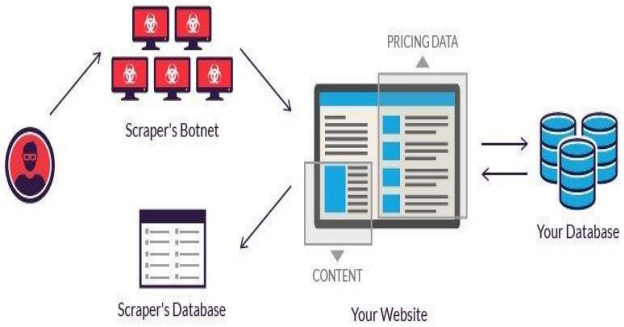



Fig1: Web Scraping

products are done based on this standard rating. In this paper, we propose a recommender system for effectively recommendation of consumers' preferences based on the ratings in e-commerce websites. We semantically model the topics in the consumer ratings to derive the product features and consumer preferences.

## IV. PERFORMANCE EVALUATION

The performance evaluation of proposed system with existing is listed in Table:1

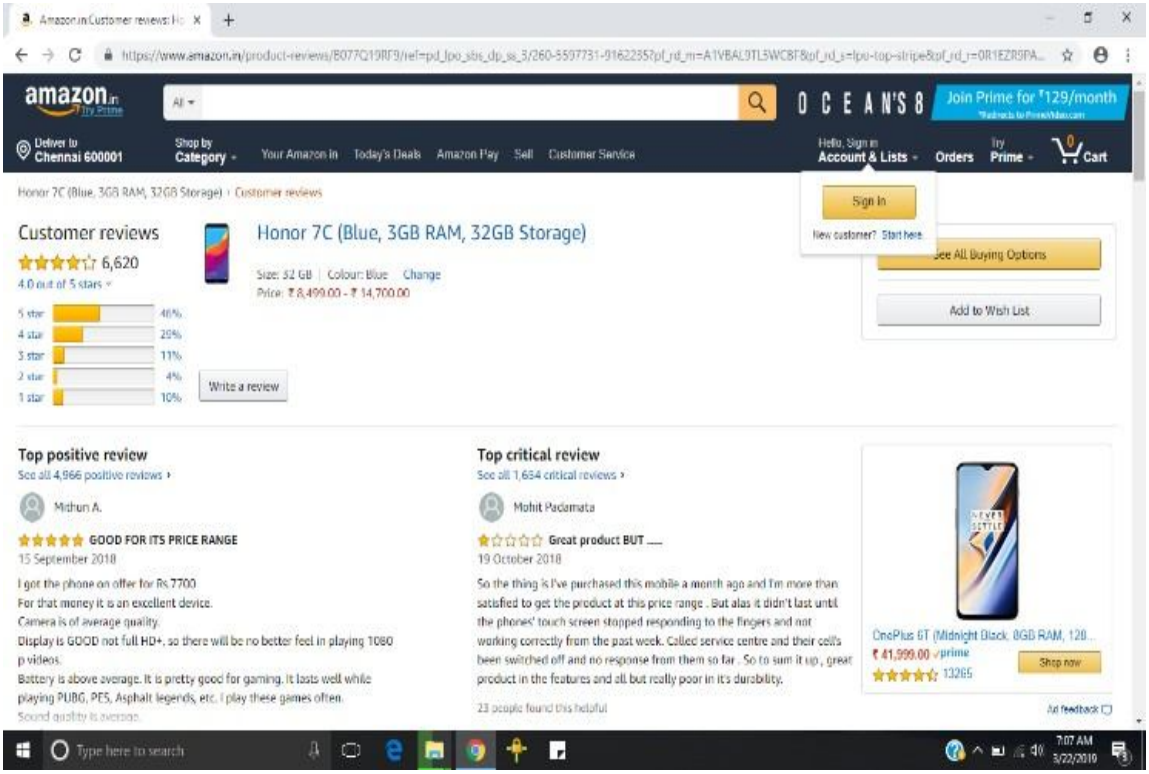



Fig3:Result of Existing system

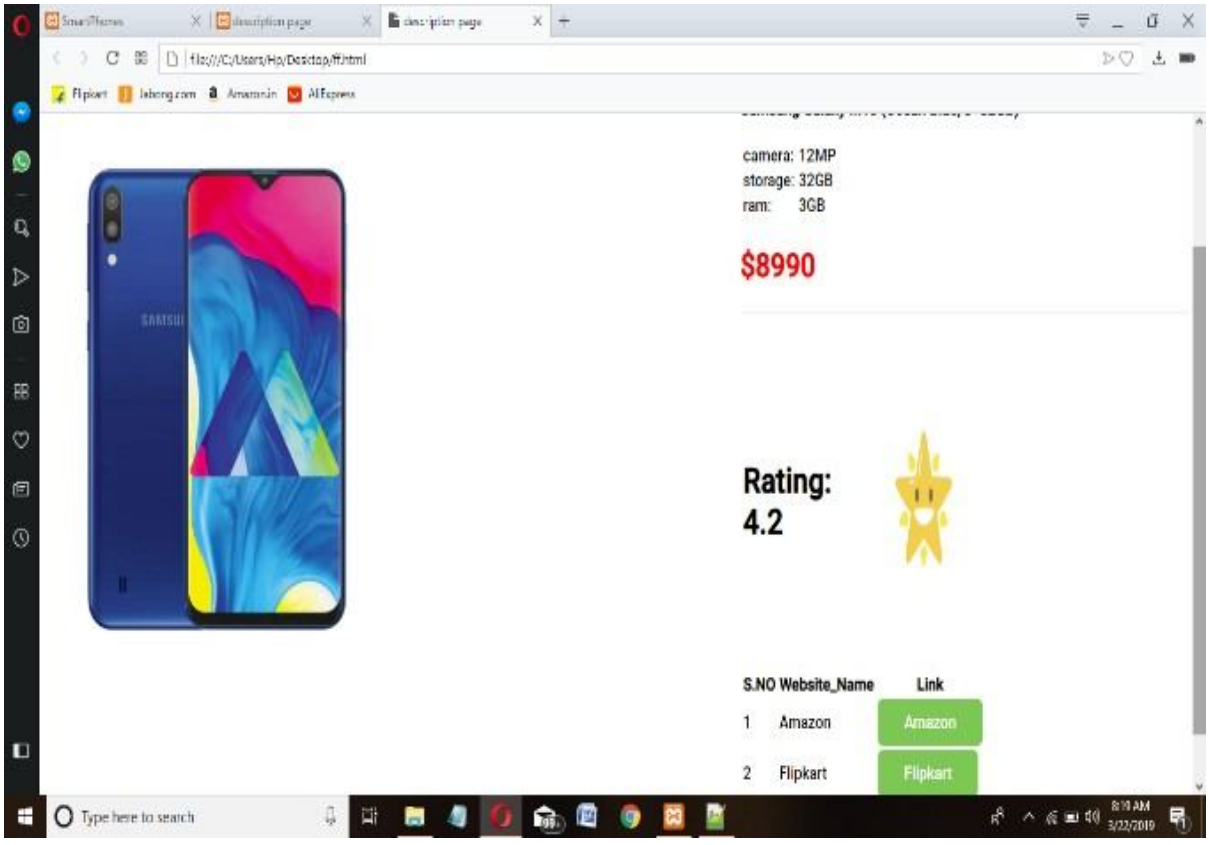



Fig4:Result of Proposed Rating

Table 1: Comparison between existing and proposed system

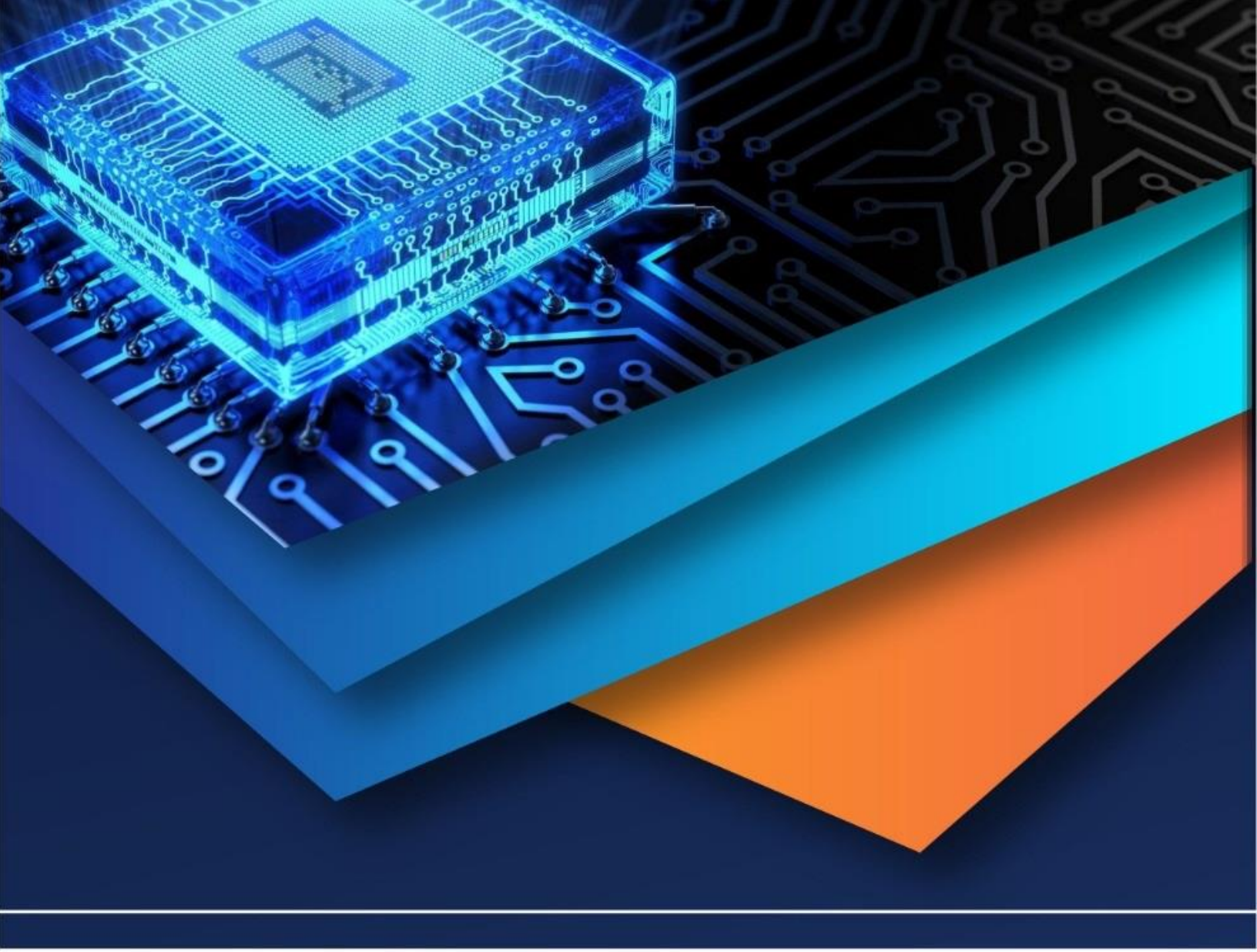
| Rating in Existing System | Rating in Proposed System |
|---|---|
| The rating is specifically provided omer who bought the product via this site. | The rating is collaboration of all the ratings available on the web for this particular product. |
| It is website specific. | It is product specific. |
| It has low accuracy as count of rating is less when compared to the proposed system. | It has high accuracy as all the rating from all the site that sells the product is collaborated. |
| Sorting of product on the page is done using less accurate ratings of the product. | Sorting is done using more accurate rating by combining all the available ratings using web scraping technology |

## V. CONCLUSION

The result is achieving accuracy in the rating by combining all the rating available online. The process is done by static scraping of rating data from various e-commerce websites. Future work may be a try to acquire this data dynamically whenever the web page is loaded. This will help update the rating table dynamically without the administrators intervention as done in the current project.

## REFERENCES

[1] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions," IEEE Transactions on Knowledge and Data Engineering, 2005.
[2] G. Adomavicius and A. Tuzhilin, "Context-aware recommender systems," Recommender Systems Handbook, 2011.
[3] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," Journal of Machine Learning Research, 2003.
[4] S. Garcia Esparza, M. P. O'Mahony, and B. Smyth, "A multi-criteria evaluation of a user-generated content based recommender system," In Proceeding of the 3rd Workshop on Recommender Systems and the Social Web in RecSys'11, 2011.
[5] T. L. Griffiths and M. Steyvers, "Finding scientific topics," In Proceedings of the National Academy of Sciences of the United States of America (PNAS), 2004.
[6] M. Hu and B. Liu, "Mining and summarizing customer reviews," In Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), 2004.
[7] D. Jannach, M. Zanker, A. Felfernig, and G. Friedrich, Recommender Systems: An Introduction, 2010.
[8] L. Le Cam and G. L. Yang, Asymptotics in Statistics: Some Basic Concepts, 2000.
[9] E. Loper and S. Bird, "NLTK: The Natural Language Toolkit," In Proceedings of the ACL-02 Workshop on Effective Tools and Methodologies for Teaching Natural Language Processing and Computational Linguistics, 2002.
[10] X. Luo, M. Zhou, Y. Xia, and Q. Zhu, "An efficient nonnegative matrix factorization-based approach to collaborative filtering for recommender systems," IEEE Transactions on Industrial Informatics 2014.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY