International Journal of Research in Engineering, Science and Management
Volume-2, Issue-6, June-2019
www.ijresm.com | ISSN (Online): 2581-5792

E-Commerce Product Recommendations by Analyzing Customer Personality

Omkar Sinalkar[1], M. A. R.Shabad[2], Shubham Vaidya[3], Shubham Waghole[4], Prasad Gondhale[5]
[1,2,3,4,5]*Student, Department of Computer Engineering, Sinhgad Academy of Engineering, Pune, India*

***Abstract*: Recommender systems attempt to highlight items that a target user is likely to find interesting. Recommender systems apply knowledge discovery techniques to the problem of making personalized recommendations for information, products or services during a live interaction. Collaborative filtering, the most successful recommender system technology to date, helps people make choices based on the opinions of other people. Recommender system is a topic which falls under the domain of information retrieval, data mining and machine learning. Recommender systems are widely used by famous websites like Amazon, Flip-Kart and many others. There are various types of recommender systems like collaborative filtering, content based filtering and hybrid recommender system. Recommender systems can be used for recommending various products like books, movies, music and any products in general. Various researchers until now have developed various algorithms to improve the accuracy of recommender systems and provide good quality recommendations. Algorithm and approach used determines the quality of recommendations.**

***Keywords*: Collaborative filtering, content-based filtering, Vector Space Model (VSM), Knowledge Based Recommendation, Reciprocal Recommendation, Hybrid Recommendation**

1. Introduction

With the advent of emerging technologies and the rapid growth of Internet, the world is moving towards e-world where most of the things are digitized and available on a mouse click. Most of the commercial transactions are performed on Internet with the help of on-line shopping that makes e-commerce to become more popular. E-commerce is very much popular nowadays. Customers are buying more and more products on the Web and business organizations are selling more and more products on the Web. Whenever a user wants to buy a product on the Web, he visits an online store and looks for item of his interest. There are many popular e-commerce sites like ebay.com and amazon.com. Such online stores sell many items. For a single item, there are many brands and models available. The opportunity for the customer to select from a large number of products increases the burden of information processing before he decides which products meet his needs [1, 2]. If the customer is not sure about product of his choice, he may face the problem of information overload. He may come across a situation, where he may be unable to decide which product to buy. Whenever, a user visits a site and selects a product to buy, the sites recommend him some more products to buy. Product Recommender systems attempt to predict products in which a user might be interested, given some information about the product's and the user's profiles.

Ecommerce business is growing. There are a number of ecommerce websites; to name some of them we have flip kart, Amazon. If you see flip kart there is a menu which shows recommendations. If you see Amazon it shows “your browsing history”, Recommendations for you”, improve your recommendations”, “your recently viewed items and featured recommendations”.

These recommendations are similar to the item you have purchased before. So several ecommerce websites have recommendation engines. So recommendation systems are gaining importance these days. There is too much of information overload for the users and people spend too much time in selecting the product. Recommendation systems make their jobs easier. Most of the algorithms are based on users searching patterns, browsing history, previously purchased items. One researcher talked about how to reduce the traditional limitation of cold start and how it can be overcome using fuzzy collaborative Filtering. There are mainly 2 types of recommender systems-collaborative filtering and content based filtering and demographic Based. The main problem in collaborative filtering is, it is not able to provide recommendations if no ratings exist. This is known as cold start problem. In content based filtering there is no variety in recommendations. User will not get to see different items which he has not seen or purchased before. Another problem that is common in recommender systems is scalability.

2. Objectives of the work

- Increasing faith in E-commerce.
- Problem of Information Overload.
- Product selection problem.
- Personalized Recommendation System.

3. Problem statement

To build an accurate RC system based on user’s personal attribute. Providing user with the features like filtering the product based on various attributes. Improved search has great payoff.

4. Literature survey

Collaborative filtering is considered to be one of the most successful product recommendation methods. Collaborative filtering identifies previous customers whose interests were similar to those of a given customer and recommends products to the given customer that was liked by previous customers. But, application of collaborative filtering to ecommerce has exposed some well-known limitations such as sparseness and scalability [3, 4]. As collaborative filtering requires explicit non-binary user ratings for similar products, the number of ratings already obtained is very small compared to the number of ratings that need to be predicted. Therefore, collaborative filtering based recommendations cannot accurately identify the products to recommend. Moreover, Algorithms to find the similar customers/products usually require very long computation time that grows linearly with both the number of customers and the number of products. With a large number of customers and products in a real world situation, collaborative filtering based product recommendations suffer serious scalability problems. These problems lead to poor recommendations. The quality of the recommendations has an important effect on the customer's future shopping behaviour. If an online store recommends products that are not to be liked by the customer, customer may become angry and it is unlikely that he will visit the online store again [4]. If the online store target customers who are likely to buy recommended products and recommend products to only them, then this situation may be avoided. Web mining is the application of data mining techniques to extract knowledge from the Web data [5]. Data mining refers to extracting unseen, hidden, novel and useful informative knowledge from a large amount of data. Web mining can be broadly divided into three distinct categories according to the kinds of data to be mined namely Web content mining, Web structure mining and Web usage mining. Web content mining is the process of extracting useful information from the contents of Web documents. Content data corresponds to the collection of information on a Web page, which is conveyed to users. It may consist of text, images, audio, video, or structured records such as lists and tables. Web structure mining is the process of discovering structure information from the Web. The structure of a typical Web graph consists of Web pages as nodes and hyperlinks as edges connecting related pages. Web usage mining is the application of data mining techniques to discover interesting usage patterns from Web data, in order to understand and better serve the needs of Web-based applications. Usage data captures the identity or origin of Web users along with their browsing behaviour at a Web site. Some of the studies have suggested web usage mining as an alternative to collaborative filtering since it will reduce the need for obtaining subjective user ratings or registration based personal preferences [6, 7]. One of the e-commerce data is click stream that means visitor's path through a web site. Click stream in an online store provides information essential to understand shopping patterns or purchasing behaviours of customers such as what products they see, what products they add to the shopping cart, and what products they buy. Through analyzing such information (i.e., web usage mining), it is possible to make a more accurate analysis of customer's interest or preference across all products than analyzing the purchase records only. Another approach of product recommendation is to benefit from the experience of the others. It is natural that whenever a person intends to buy an item, it takes views of his friends or relatives to select the brand and the model. The business companies also advertise their products highlighting their features. But, a normal person never blindly trusts these advertisements. In this era of e-commerce, customers are turning towards online opinions for the purpose. There are many online opinion resources such as online news, forums, blogs and reviews. Opinions are subjective statements that reflect user's sentiments or perceptions towards an item. It is possible that by reading other's posted opinions, a customer can take decision on buying a product. On the web, there are hundreds of opinion sources available. A user may like to search for opinions on a particular item by utilizing a search engine such as Google. But, the search engine may provide the user not only the desired information, but also a large amount of irrelevant information. Hence, the user again has to face the problem of information overload. Opinion mining is a subclass of Web content mining, where reviews of various products are mined to extract people's opinion. Opinion extraction allows Web users to retrieve and summarize people's opinions scattered over the Internet. Automated opinion mining can provide quick search [8] and analysis [9] results to both consumers and manufacturers [10] area of product recommendation. In earlier works, collaborative filtering has been used successfully in a number of different applications such as recommending web pages, movies, articles and products [11]-[13]. Since, collaborative filtering has some major limitations, researchers investigated to use Web mining techniques for product recommendation.

In literature, we find that majority of works on product recommendation using Web mining techniques are based on Web usage mining. Web usage mining is the process of applying data mining techniques to the discovery of behaviour or patterns from web data. The pattern discovery tasks include the discovery of association rules, sequential patterns, usage clusters, page clusters, user classifications or any other pattern discovery method [6], [7]. In [14], Cho et al. proposed a personalized recommendation system based on Web usage mining. They recommended products based on web usage data as well as product purchase data and customer related data. In [15], Kim et al. discussed personalized recommendation based on Web usage mining. Their method focused on the problem of helping customers to get recommendation only about the products they would like to buy. For this, they suggested a list of top-N recommended products for a customer at a particular time. They performed experiments with the Web usage data of a leading Internet shopping mall in Korea for the evaluation of

their methodology. Experimentally, they deduced that choosing the right level of product taxonomy and the right customers increases the quality of recommendations. In [16], Liu and Shih developed a product recommendation methodology that combined group decision-making and data mining techniques. They applied the analytic hierarchy process (AHP) to determine the relative weights of frequency, monetary (RFM) variables in evaluating customer lifetime value or loyalty. They then applied clustering techniques to group customers on the basis of the weighted RFM value. Finally, product recommendations to each customer group were provided using association rule mining. They concluded that recommending more number of items helps to improve the quality of recommendation for more loyal customers, but not do so for less loyal customers. Zeng also discussed the development of a personalized product recommendation system in [17]. The recommender system utilized the web mining techniques to trace the customer's shopping behaviour from his/her click streams and learned his/her up-to-date preferences adaptively. Here, the customer preference and product association were automatically mined from click streams of customers. Then, the matching algorithm which combined the customer preference and product association was used to score each product. The system then produced the recommended product lists for a specific customer. Experimentally, they showed that their system provides sensible recommendations, and enabled customers to save enormous time for Internet shopping.

One of the earlier works on opinion extraction was reported by Hu and Liu in [18]. They considered three things in opinion extraction namely (i) extraction of Subject (the product), (ii) aspect (the attributes or features), and (iii) semantic-orientation. Semantic-orientation was binary valued either positive or negative. Popescu and Etzioni in [19] additionally annotated Hu and Liu's corpus with tags. In [20], Kobayashi et al. discussed how customer reviews in web documents can be best structured. They proposed to structure opinion unit as a quadruple, that is, the opinion holder, the subject being evaluated, the part or the attribute in which it is evaluated, and the evaluation that expresses positive or negative assessment. They used a machine learning-based method for opinion extraction. Aciar et al. in [21] used prioritized consumer product reviews to make product recommendations. Using Web content mining (also, called sometimes text mining) techniques, they mapped each piece of each review comment automatically into an ontology. Scaffidi et al. implemented a prototype system called Red Opal to score each product on each feature for the users to locate products rapidly based on features. Sun et al. in [23] proposed an automated system to compare and recommend products for customers from both subjective and objective perspectives. For subjective comparison of products, they used results of opinion mining. They also included product technical details to improve the comparison results from the objective perspective.

5. Requirements

System: Intel Core i3 2.4 GHz
Hard Disk: 40 GB
RAM: 512 Mb
Operating system: Windows 7
Coding Language: JAVA, JSP and Servlets IDE: JAVA Eclipse
Database: MySQL

6. Proposed architecture

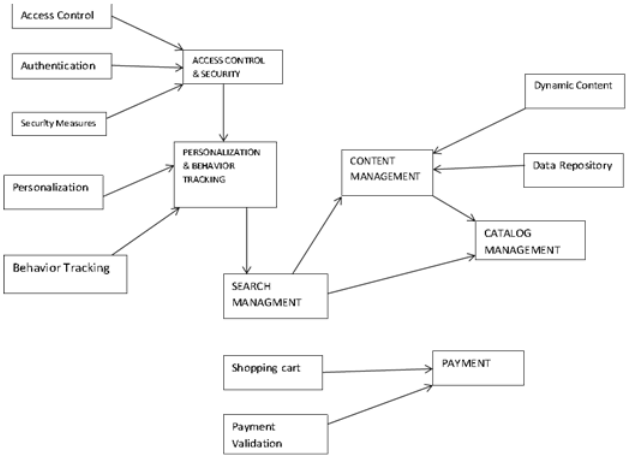

Fig. 1. System Architecture

7. Algorithm analysis

Collaborative Filtering systems can produce personal recommendations by computing the similarity between your preference and the one of other people one of the most promising such technologies is collaborative filtering [shardanand et al., 1995]. Collaborative Filtering works by building a database of preferences for items by users. A new user, Neo, is matched against the database to discover neighbours, which are other users who have historically had similar taste to Neo. Items that the neighbors like are then recommended to Neo, as he will probably also like them. Collaborative Filtering has been very successful in both research and practice, and in both information filtering applications and E-commerce applications. However, there remain. Important research questions in overcoming two fundamental challenges for Collaborative Filtering recommender systems.

Content Based Recommendation: These algorithms try to recommend items that are similar to those that a user liked in the past (or is examining in the present). Various items are compared with items previously rated by the user and the best-matching items are recommended. In content based filtering say for ex a person likes gadgets we would recommend him some other gadgets in future. A widely used algorithm is the tf–idf representation in content based recommendation engine. Items are recommended based on items the user previously bought, based on users browsing history or clicks and then similar items are recommended to the user.

Let us see what is TF-IDF Fij=frequency of the term(i) in a document (j) TFij=Fij/Max number of times the word occurs in diff docs Ex: a word pear appeared 5 times in one document and the total number of times it occurred in several other documents is say 25. TFij=5/25=0.2. The more the TF score of a word the more important is the word. e.g. A word apple appears 5 times in a document and in other document it appears only once.So the word apple is more important in one document as compared to other documents. But how do you calculate the weight of different words. Ni=the number of docs that mention the term i N=total number of documents IDFi=logN/Ni TF-IDF score=TFij x IDFi.

IDF score gives lower weight to less common words and higher weight to rarer words.

Hybrid approach combines collaborative and content based approach.

Demographic based Recommender Systems Recommender systems that are based on demographic characteristics of consumers and recommend a list of items that have good feedback from the consumers that are demographically similar to the target consumer.

8. Results

Products are recommended by three ways viz., Search based, review based and slope one algorithm.

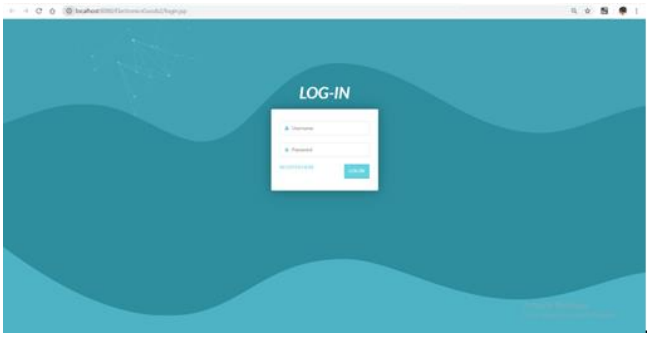

Fig. 2. Login page

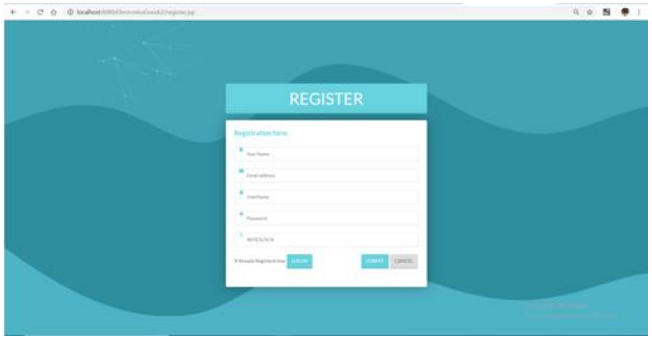

Fig. 3. Register page

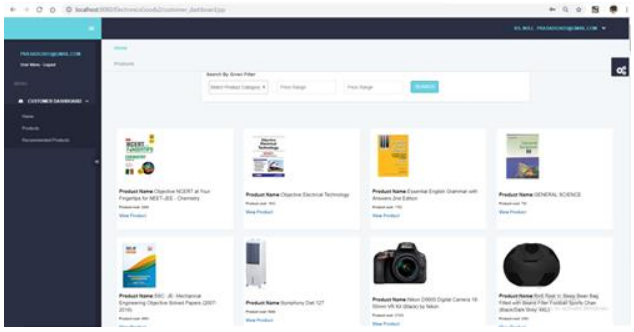

Fig. 4. Dashboard 1

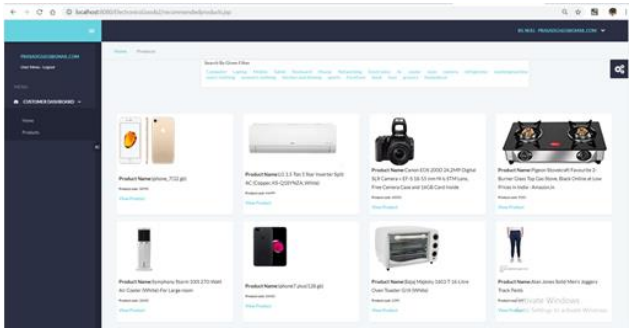

Fig. 5. Dashboard 2

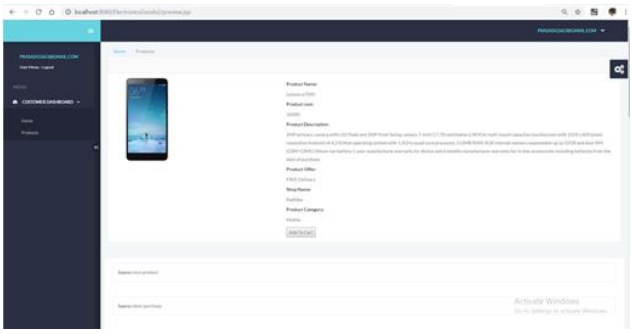

Fig. 6. Product review

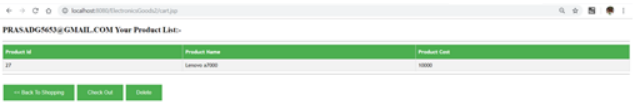

Fig. 7. Cart

Fig. 8. Payment

9. Conclusion

With an extra information overload over internet, users need good and sound recommendation techniques. In this paper, we describe various recommendation techniques and briefly their advantages and limitations are elaborated. This gives a clear idea about the recommendation approaches and easy to understand the phenomena of recommendation, even for a native user. Finally, we conclude that there is a need to make a lot of efforts to overcome the limitations of the existing techniques. Thus we aim to decrease the limitations and increase the enhancement by collaborating the above mentioned algorithms.

References

[1] E. Kim, W. Kim, and Y. Lee (2000). "Purchase propensity prediction of EC customer by combining multiple classifier based on GA", In Proceedings of International Conference on Electronic Commerce 2000, pages 274–280.

[2] J. B. Schafer, J. A. Konstan and J, Riedl (2001). "E-commerce recommendation applications", Data Mining and Knowledge Discovery, volume 5, issue 1–2, pages 115–153.

[3] M. Claypool, A. Gokhale, T. Miranda, P. Murnikov, D. Netes and M. Sartin (1999). "Combining content-based and collaborative filters in an online newspaper". In Proceedings of ACM SIGIR ‟99 Workshop on Recommender Systems, Berkeley, CA.

[4] B. Sarwar, G. Karypis, J. Konstan and J. Riedl (2000). "Analysis of recommendation algorithms for e-commerce", In Proceedings of ACM ECommerce 2000 Conference, pages 158–167.

[5] G. Xu (2008). "Web Mining Techniques for Recommendation and Personalization", Ph.D. Thesis submitted to The School of Computer Science & Mathematics, Faculty of Health, Engineering & Science, Victoria University, Australia.

[6] B. Mobasher, R. Cooley, and J. Srivastava (2000). "Automatic personalization based on web usage mining", Communications of the ACM, volume 43, issue 8, pages 142–151.