

## BOOK RECOMMENDATION SYSTEM BASED ON COMBINE FEATURES OF CONTENT BASED FILTERING, COLLABORATIVE FILTERING AND ASSOCIATION RULE MINING

Ashish Fatarphekar<sup>1\*</sup>, Tejas Nashikkar<sup>2</sup>, Vivek Patil<sup>3</sup>, Gayatri Naik<sup>4</sup>

**\*Corresponding Author: -**

---

### **Abstract: -**

*Recommendation systems are widely used to recommend products to the end users that are most appropriate. Online book selling websites now-a-days are competing with each other by many means. Recommendation system is one of the stronger tools to increase profit and retaining buyer. The book recommendation system must recommend books that are of buyer's interest. This paper presents book recommendation system based on combined features of content filtering, collaborative filtering and association rule mining.*

**Keywords: -** Association rule, Collaborative filtering, Content based filtering, Recommendation system



Distributed under Creative Commons CC BY-NC 4.0 OPEN ACCESS

## I. INTRODUCTION

Now-a-days many book selling websites are available on the internet. Many of them are having their own recommendation system to recommend books to the buyers. Generally the book recommended by most of the websites is not of the buyer's interest. Generally lots of information and recommendations are pushed to buyers, but most of them are of not relevant. This paper presents a new approach for recommending books to the buyers. This system combines the features of content filtering, collaborative filtering and association rule mining to produce efficient and effective recommendations.

Our paper is further organized as follows. Section 2 describes content recommendation system; Section 3 presents collaborative filtering based recommendation system; Section 4 describes major point of association rule mining; Section 5 describes major tables used in book selling website. Section 6 discusses our novel approach for recommending books to the buyers.

### Content recommendation system

To provide relevant information to the buyers, Web Usage Mining (WUM) is used. Web Usage Mining stores the user's behavior on the internet and processes that data. WUM generate information that is generally most appropriate and relevant to the user. Recommendation system is one of the example Web Usage Mining. Content recommendation system recommends books to the buyers based on the content of the buyers past book buying history. Buying history gives the content of the *overview of the books*, in which buyer is generally interested from the large amount of books. Content recommendation system filter books based on the content of the book, in which buyer is interested. Content recommendation system uses Content based filtering for doing the separation.

Like other system Content based filtering is also having some limitations like finding the quality of the content. For example, Content based filtering cannot differentiate between good article and bad article if both of them are using same terminology.

## II. Collaborative filtering-based recommendation system

Content based filtering cannot find out the quality of the item. To overcome this problem collaborative filtering system are used because they are based on opinion of the other users. Collaborative filtering (CF) (sometimes called "social filtering" or recommender system) is one of the way to do recommendation on the web. One of the earliest used collaborative algorithms and results were GroupLens and Ringo. Item based collaborative recommendation algorithm looks in to the set of items the target user has rated and computes how much similar they are to the target item  $i$  and then selects  $k$ - most similar items  $\{i_1, i_2, \dots, i_k\}$  to the set of items the target user has rated, the recommendation is then computed by taking the weighted average of the target user's rating on these similar items.

To compute the similarity between the two items, each item is considered as vector in  $m$  dimensional user-space. The similarity between two item is measured by computing the cosine of the angle between the two vectors

. Suppose if there are  $m$  users ( $u_1, u_2, \dots, u_m$ ) and  $n$  items ( $i_1, i_2, \dots, i_n$ ). So in  $m \times n$  matrix in Fig.1, similarity between two item  $i$  and  $j$  is, denoted by

	$i_1$	$i_2$	$i_3$	.	$i_n$
$u_1$					
$u_2$					
$u_3$					
$u_4$					
.					
.					
.					
$u_m$					

Figure1: User-Item rating matrix

$sim(i,j)$  is given by

$$sim(i,j) = \frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\| \|\vec{b}\|}$$

where "." denotes the dot product of two vectors

## III. Association rule mining

Association rule mining finds interesting association and correlation relationship among large data set of items [9]. Market basket analysis is considered as a typical example of association rule mining. In market basket analysis customer buying habit is analyzed for finding association between different items customer put together in their shopping cart. Let  $I = \{i_1, i_2, \dots, i_m\}$  be a set of items. An association rule can be represented by this form  $A \rightarrow B$ , where  $A \subseteq I$ ,  $B \subseteq I$  and  $A \cap B = \emptyset$

. Association rule extract the pattern from the database based on the two measures minimum support and the minimum confidence. The support and confidence measures are described as stated in.

**Support:** The rule (A B) holds in transaction set D with *support*  $s$ , where  $s$  is the percentage of the transactions in D containing  $A \cup B$ .

$$\text{Support}(A \cup B) = P(A \cup B) \quad (1)$$

**Confidence:** The rule (A B) has *confidence*  $c$  in the transaction set D, where  $c$  is the percentage of transactions in D containing A that also contain B. than or equal to the user defined minimum confidence.

#### IV. Major tables used in book selling website

While buying books from a website customer is using the following tables to finish the buying action. The website will store the data of each customer order in its database. It is the data from which information is extracted. Major tables used in the book selling website are as follows.

**Customers' Information Table:** Table of customer's information stores information of all the customers like login Id, password etc. It is shown in table 1.

**Customer's Information**  
**TABLE I. TABLE**

Field Name	Description
Login Id	Customer login Id
Customer Name	Customer Name
Password	Customer Password
email	Customer email used for sending push information
Mobile number	Customer mobile number

**Book Information Table:** Book information table stores the information about all the books sold by the website. It is shown in the table 2.

**TABLE II. BOOK INFORMATION TABLE**

Field Name	Description
BookId	Book ID
BookName	Book's Name
CategoryId	Book Category Id
SubCatId	Book Sub Category Id
Author	Book's Author Name
Publisher	Book's Publisher Name
BookPrice	Book's Price
BookOverview	Book's Overview/ Table of content
Edition	Book's Edition
BookRating	Book's average rating given by the registered customers who have bought the book.

**Category Table:** The category table stores all the categoryId and the corresponding book's category name. It is shown in table 3.

Confidence (A B) = P (B|A) (2)

In general association rule mining can be thought of as two step process.

(i) Generating all item sets having support greater than or equal to the user defined minimum support. (i) Generate all the rules having the confidence factor greater

**TABLE III. CATEGORY TABLE**

Field Name	Description
CategoryId	Book Category Id
Category Name	Book'sCategory Name

placed. It stores the ordered, BookId, quantity etc. It is shown in table 4 and 5.

**ORDER INFORMATION****TABLE IV. TABLE**

Field Name	Description
Login Id	Customer login Id
OrderId	Customer's order Id
OrderDate	Date of Purchase
TotalAmount	Total amount
Email	Email of the Customer.

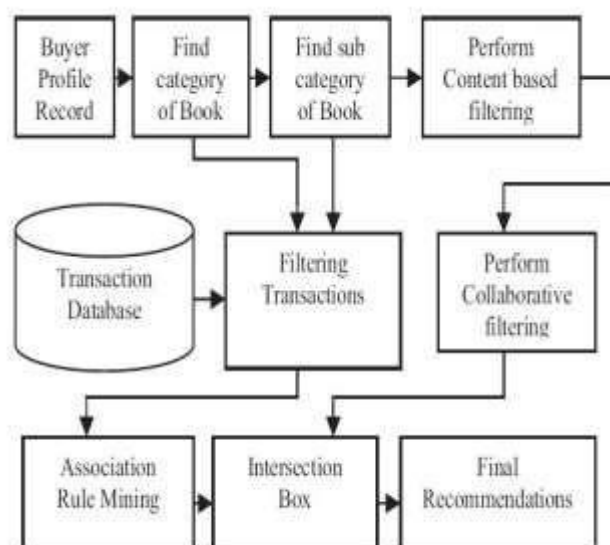
**TABLE V. ORDER DETAILS TABLE**

Field Name	Description
OrderId	Customer order Id
BookId	Book's Id
Quantity	Total Number of books bought by the customer

**Vi. Book recommendation system**

Purpose of this book recommendation system is to recommend books to the buyer that suits their interest. This recommendation system works offline and stores recommendations in the buyer's web profile. This system has following seven steps:

1. Find out the category of the book that the buyer has bought earlier like novel, science, engineering etc from the buyer's web profile.
2. Find out the subcategory of the book if there is any in the step1 found category.

**Figure2. Block diagram of Book recommendation system**

2. Perform content based filtering in category / subcategory found in step1 and 2, to find out the books that are much similar to the books that the buyer has bought earlier based on the books overview content from the buyers past history record.

3. On the result of step 3 perform item based collaborative filtering and find out the list of books in the descending order of recommendations. In this step system actually evaluate the quality of the recommending books based on the rating given to those books by the other buyers.
4. From the book transaction database find all those transactions whose category and sub category (if there is any) is same as found in step1 and step2. Apply association rule on those transactions and find out the books that the buyer can buy afterward. Adjust the support and confidence parameters to get stronger rules.
5. Find out the intersection of the result of step 4 and 5. Arrange the intersection result in the descending order of recommendations as given by the step 4. This step is actually more refining the recommendations generated by the step 4
6. Outcome of the step 6 is the final recommendations for the buyer. All these steps are performed when the buyer is offline and the results are stored in the buyers web profile. When the buyer comes online next time the recommendations will be generated automatically

This book recommendation system is represented by block diagram in the Fig 2.

## Vii. Future Work

### 1) Movie recommendation system:

The proposed system uses suggested approach in movies or songs recommendation system as it enlists the user's genres, stores it in the database and then suggests users with similar genres and types.

### 2) Product recommendation system:

This is the most popular and widely used application. In this type of application a user or a visitor searches through various products available online and buys if anything interests the user. Popular examples of these are flipkart, amazon etc.

### 3) Online library:

This is similar application, when a particular user requests a particular book of a specific genre then types are enlisted and next time when user is browsing through list then at that time different books are suggested according to the stored result.

## Viii. Conclusion

The goal of the most recommendation system is to predict the buyer's interest and recommends the books accordingly. This book recommendation has considered many parameters like content of the book and quality of the book by doing collaborative filtering of ratings by the other buyers. This recommender system also uses associative model to give stronger recommendations. This system does not have performance problem since it built the recommendations offline.

## References

- [1] FOLTZ, P. W. AND DUMAIS, S. T. 1992. Personalized information delivery: an analysis of information filtering methods. *Comm. ACM* 35(12), pp. 51–60.
- [2] SHARDANAND, U. AND MAES, P. 1995. Social information filtering: algorithms for automating “word of mouth”. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*
- [3] Resnick, P., and Hal, R. V., 1997. Recommender Systems, *Communications of the ACM*, 40, 3, pp. 56-58.
- [4] Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., and Riedl, J., 1994. GroupLens: An Open Architecture for Collaborative Filtering of Netnews, *Proceedings of ACM 1994 Conference on Computer Supported Cooperative Work*, Chapel Hill, NC, pp.175-186.
- [5] Baraglia R., Silvestri F.: Dynamic Personalization of Web Sites Without User Intervention. *Comm. ACM*, Vol. 50, 2 (Feb., 2007): pp. 63-67.
- [6] Eirinaki, M. and Vazirgiannis, M.: Web Mining for Web Personalization. *ACM Trans. On Internet Technology*, Vol. 3, 1 (Feb., 2003): pp. 1-27.
- [7] Thede, L., Marshall, V.A., Rick W.: An Economic Answer to Unsolicited Communication. *EC'04*. (2004).
- [8] SARWAR, B., KARYPIS, G., KONSTAN, J., AND REIDL, J. 2001. Item-based collaborative filtering recommendation algorithms. In *Proceedings of the 10th International Conference on World Wide Web (WWW'01)*. ACM, New York, NY, pp.285– 295.
- [9] J. Han, M. Kamber, *Data Mining: Concepts and Techniques*, The Morgan Kaufmann Series, 2001.
- [10] Agrawal, R., Imielinski, T., Swami, A. N. "Mining association rules between sets of items in large databases". In *Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data*, 207-216, 1993.
- [11] Luo Zhenghua, "Realization Of Individualized Recommendation System On Books Sale" *IEEE 2012 International Conference on Management of e-Commerce and e-Government*. pp.10-13.

## Author's Profile

Ashish Fatarphekar, Computer engineering, Mumbai University/ yadhavrao Tasgaonkar Institute of engineering & Technology/ Saraswati Education Society, bhivpuri karjat, India, Mobile No.7276461458, (e-ashishf6@gmail.com).  
Tejas Nashikkar, Computer engineering, Mumbai University/ yadhavrao Tasgaonkar Institute of engineering & Technology/ Saraswati Education Society, bhivpuri karjat, India, Mobile No.9987769669, (e-mail:tj.tejas007@gmail.com).

Vivek Patil, Computer engineering, Mumbai University/ yadhavrao Tasgaonkar Institute of engineering & Technology/  
Saraswati Education Society, bhivpuri karjat, India, Mobile No.7208203610, ( e-mail: vivekvp.patil8 @gmail.com).  
Gayatri Naik, Computer engineering, Mumbai University/ yadhavrao Tasgaonkar Institute of engineering & Technology/  
Saraswati Education Society, bhivpuri karjat, India, Mobile No.7208203610, ( e-mail: gayatri.naik@gmail.com).