

Recommendation System Using Machine Learning

Suhasini Parvatikar
Computer Engineering, Assistant Professor
SCOE, Kharghar, India
er.suha@gmail.com

Dr. Deepa Parasar
CSE, ASET, Amity University Mumbai
Panvel, India
dparasar@mum.amity.edu

Abstract— Due to increase of data on internet, there is an increased dependency on internet by people. Thus, recommendation systems help people by suggesting products where there is overload of information on ecommerce websites. There are various methods for recommendation. This paper studies about various techniques used in designing of recommendation system with machine learning algorithm.

Keywords— Content filtering, Collaborative Filtering, Naïve Bayes, KNN clustering.

I. INTRODUCTION

The use of internet and availability of data has increased enormously due to which people are relying more on internet nowadays. Most used service of the internet [10] is online shopping, where many companies sell their products over the internet by using ecommerce applications like Amazon, MyNtra. An ecommerce website provides large database of specified items to a user to choose from, which leads to overload problem of information.

In this scene, the users find difficulty to make decision about which product to choose with vast amount of information available to them. Recommender Systems [10] have come up in response to this problem by learning from users' behavior and recommending items that are similar to their interest. Thus, the Recommender system helps user to find perfect choice of item required on ecommerce sites.

There are various reasons to use recommendation system, but the reason behind why companies tend to use recommender systems are:

1. To increase sell of their items/products
2. To understand users behaviour for product purchase.

Recommendation systems provide filtering of information by cutting down large amount of database and recommend items according to their interest. Recommendation systems generally use one of the methods as shown in Fig 1[1] for recommendation.

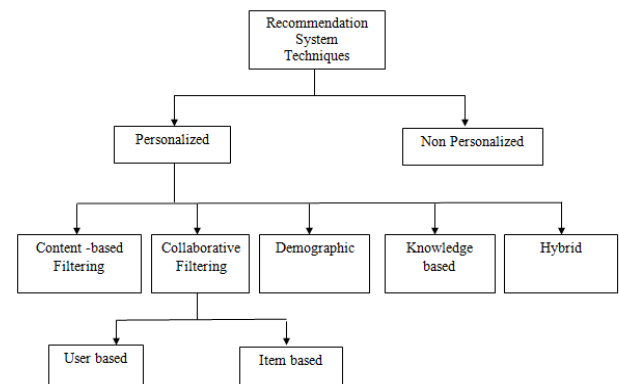


Figure 1: Recommendation System Techniques

A. Personalized Recommendation: Personalized recommendation [1] depends on users' behavior for selecting item/product on ecommerce website. Generally, personalized RS are of 5 types:

a) Collaborative Based Filtering:

Collaborative filtering systems generally work by collecting users' ratings for items in a specified field and then find likeness in rating among many users to recommend an item. There are two ways of Collaborative filtering:

1) **User based CBF:** User-based CBF gives suggestion based on behaviour of users who are having similar kind of taste in options. It correlates users' ratings given to the items as shown in Fig.2 [1]

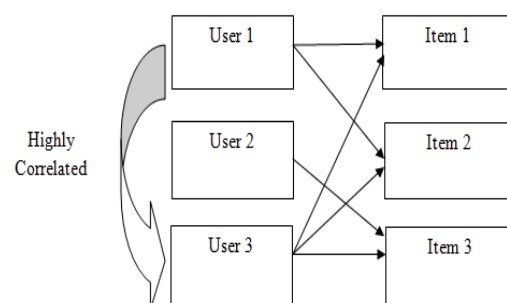


Figure 2: User based CBF

2) **Item based CBF:** Item-based CBF [1] gives suggestion based on the items rated similarly by user as shown in Figure 3[1].

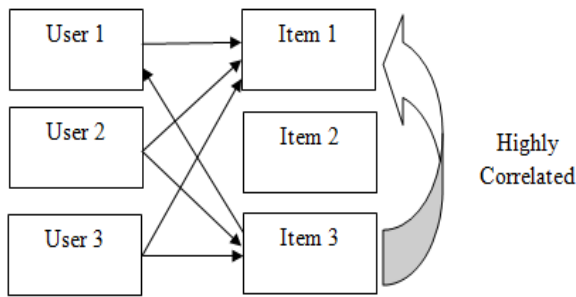


Figure 3: Item based CBF

b) Content Based Filtering

This type of Recommender System as shown in Fig 4 is dependent on inputs provided by user in past. Examples of content based recommendation are Google, Wikipedia etc.

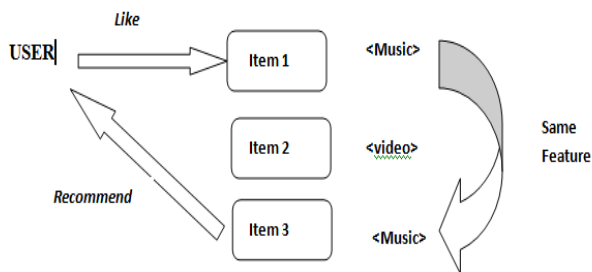


Figure 4: Content based filtering

c) Demographic Based Filtering

Demographic Based filtering uses user profile information like for eg: gender, marital status, age, postal-code [9]. But, this method is time consuming process as it requires users information.

d) Knowledge Based Filtering

These types of recommendation system ask user for preferences. No interaction of data is needed.

e) Hybrid Filtering

Hybrid Approach includes the combination of all above methods.

B. Non Personalized Recommendation: These RS systems don't rely on user's behaviour; they just suggest items randomly to customers based on average feedback about product. Therefore, all customers get the same suggestions on product.

The most common recommender systems applications [2] are:

- Recommendation system in Music domain
- Recommendation system in E-Commerce
- Recommendation system In Social Network

II. LITERATURE SURVEY

M.Sunitha Reddy, T.Adilakshmi, V.Swathi [5] proposed hybrid recommendation system which combines clustering and association mining. The main focus of authors was to find solution on CBF challenge i.e. cold start problem. The proposed algorithm works as follows.: In first step, user clusters are formed. These clusters are formed on the basis of similarity calculated by cosine similarity method, thus using collaborative filtering. In second step, each cluster is converted to transactional database. This transactional database is used to find frequent item set by using extended FP tree as shown in Figure 6[5]. From the frequent item sets strong association rules are formed by traversing the tree when it is required. This method improves the quality of recommendation.

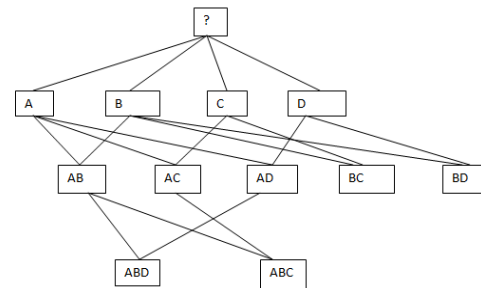


Figure 6: Extended FP-tree

Pijitra Jomsri [8], designed a system which uses association rule technique. This method identifies relationship between books items. The motive of this project is to help user for searching books. This project proposed various steps for book recommendation. In first step, to collect dataset, apply association rule to recommend books. The results are evaluated by using precision and recall.

Monali Gandhi, Khusali Mistry, Mukesh Patel [9], discusses about the reason behind using hybridization of algorithms. The authors described the different methods of

recommendation . Association mining alone cannot be effective when the data set is not sufficiently available. The proposed system divides its modules of work into 4 phases .In first phase, database preselection is done then, association rule mining is used and at last collaborative filtering is done. Finally top N items are recommended to users.

Hongwu Ye [15] proposed RS self-organizing maps using combination of collaborative filtering, association mining and to remove data sparsity problem in recommender system.

III. PROPOSED METHODOLOGY

The proposed framework consists of following modules: 1) Database 2) ML Algorithm 3) Item based CBF as shown in Fig 3. The database maintains content-based information i.e. It uses the log details of books which are rated by user positively in the past. Our system computes clustering/classification using ML algorithm and then finds similarity between items and gives recommendation. When the user searches for any item from database, the database runs these algorithms and recommends item with user’s interest.

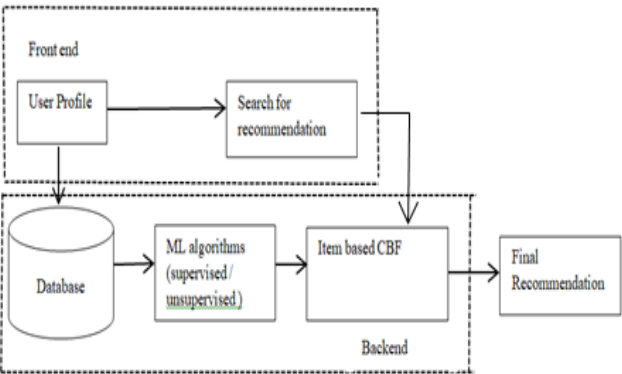


Figure 7: Recommendation architecture

A)ML Algorithm :

We can use unsupervised method ie Clustering or supervised method ie Naïve bayes algorithm for classification .

Clustering is one of machine learning technique to find hidden patterns in a given dataset. Its an unsupervised learning.

In this step the proposed system forms the user clusters based on accessing the frequent book itemset.

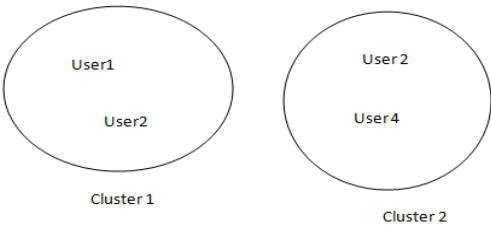


Figure 8: Example of User based Clustering

This clustering technique represents each user with partial participation in several clusters. If any new user, access any of these items we will be added to corresponding cluster. The use of this type of clustering is just to predict the user’s interest in future.

B) Collaborative Filtering

Collaborative Filtering (CF) [19] is most common technique used for recommendation. The main motive of CBF is to find similar users. The users having similar taste i.e. almost similarly rated items, form a group. For example, Table 1 shows that 3 users rated the books positively with similar rating. I.e. they have common taste and thus they form group. The user A hasn’t rated the book ASP.Net, might be he hasn’t watched it yet. As other users have positively rated that , user A will get this item recommended.

Books User →	Java	C++	C#	ASP.Net
A	8	7	9	-
B	9	7	9	9
C	9	-	9	8

Table 1: Example of Collaborative filtering

The main functions [10] of Collaborative filtering recommendations system are:

1. Recommendation: It displays items to a user. The items are listed in the order of usefulness to the user.
2. Prediction: In prediction a calculation of predicted rating is made for a particular item.

User based CBF systems perform poorly when they had many items and comparatively few ratings are present. So item based algorithm is usually recommended. The Item based CBF algorithm can be summarized in the following steps:

Step 1: Compute Similarity between items from above cluster by using equation 1 as given below.

Step 2: Select n active items that have the highest similarity.

Step 3: Compute a prediction $P_{u,t}$ from a weighted combination.

a) Item Similarity Computation

This is most important step in Item based CBF. It finds similarity between items and selects the most similar item. There are different kind of ways to find the similarity between items using formula:

1. Pearson Correlation: It finds the linear correlation between two vectors of ratings the target item t and the remaining item r.

$$sim(t, r) = \frac{\sum_{i=1}^m (R_{it} - A_t)(R_{ir} - A_r)}{\sqrt{\sum_{i=1}^m (R_{it} - A_t)^2 \sum_{i=1}^m (R_{ir} - A_r)^2}} \dots \dots \dots (1)$$

Where R_{it} is the rating of the target item t by user i, R_{ir} is the rating of the remaining item r by user i, A_t is the average rating of the target item t for all the co-rated users, A_r is the average rating of the remaining item r for all the co-rated users, and m is the number of all rating users to the item t and item r.

2. Cosine Vector Similarity: It is the angle between two vectors of ratings as target item t and the remaining item r.

$$sim(t, r) = \frac{\sum_{i=1}^m R_{it} R_{ir}}{\sqrt{\sum_{i=1}^m (R_{it})^2 \sum_{i=1}^m (R_{ir})^2}} \dots \dots \dots (2)$$

Where R_{it} is the rating of the target item t by user i, R_{ir} is the rating of the remaining item r by user i, and m is the number of all rating users to the item t and item r.

b) Prediction using item based Collaborative filtering

The most important step [18] during a collaborative filtering system is to get prediction. Since we have similar items supported calculation of similarity measures, subsequent step to urge prediction of rating of target user. The rating [12] of the target user u to the target item t is as following:

$$P_{u,t} = \frac{\sum_{i=1}^c R_{ui} \times sim(t, i)}{\sum_{i=1}^c sim(t, i)} \dots \dots \dots (3)$$

Where R_{ui} is the rating of the target user u to the neighbour item i, $sim(t, i)$ is the similarity of the target item t and the neighbour item i for all the co-rated items, and m is the number of all rating users to the item t and item r.

IV. CONCLUSION

Recommendation System is most important part of life. Due to recommendation system, online shopping has gained much popularity. Adequate services are being provided by recommender system. Thus, recommender system has proved a gift for the society which helps customer to decide among their choices. The recommender system uses several approaches for recommendation.

REFERENCES

- [1] Yagnesh Patel, Vishal Patel, "A Survey on Various Techniques of Recommendation System in Web Mining", International Journal of Engineering Development and Research Volume 3, Issue 4, 2015, pp 696-700.
- [2] Jaimeel Shah, Lokesh Sahu, "Recommendation based on Clustering and Association Rules", International Journal of Advance Research and Innovative Ideas in Education, Vol-1, Issue-2, 2015, pp 49-52.
- [3] Anand Shanker Tewari and Kumari Priyanka, "Book Recommendation System Based on Collaborative Filtering and Association Rule Mining for College Students", IEEE International Conference on Contemporary Computing and Informatics, 2014, pp 135-138.
- [4] Anand Shanker Tewari, Abhay Kumar and Asim Gopal Barman, "Book Recommendation System based on Combine Features of Content Based Filtering, Collaborative Filtering and Association Rule Mining", IEEE International Advance Computing Conference, 2014, pp 500-503.

- [5] M.Sunitha Reddy, T.Adilakshmi, V.Swathi, "A Novel Association Rule Mining and Clustering based hybrid method for Music Recommendation System", International Journal of Research in Engineering and Technology ,Vol 3,Issue 5 ,May 2014, pp 55-59.
- [6] V.Manithva, M.Sunitha Reddy, "Music Recommendation System using Association Rule Mining and Clustering Technique to address Cold Start problem ", International Journal of Engineering and Computer Science ,Vol 3 ,Issue 7, July 2014 , pp 6855-6858.
- [7] ZHAO Kai , LU Peng-yu, " Improved Collaborative Filtering Approach Based on User Similarity Combination", IEEE International Conference on Management Science & Engineering (21th), Helsinki ,Finland ,August 2014 ,pp 238-243.
- [8] Pijitra Jomsri , " Book Recommendation System for Digital Library Based on User Profiles by using Association Rule ", IEEE 4th International Conference on Innovative Computing Technology , London , UK , August 2014 , pp 130-134.
- [9] Monali Gandhi , Khushali Mistry, Mukesh Patel , " A Modified Approach towards Tourism Recommendation System with Collaborative Filtering and Association Rule Mining", International Journal of Computer Application, Vol 91, No 6, April 2014, pp 17-21.
- [10] Noraswaliza Abdullah, Yue Xu, Shlomo Geva, " Integrating collaborative filtering and matching-based search for product recommendations", Journal of Theoretical and Applied Electronic Commerce Research, Vol. 8 No. 2 ,2013,pp. 34-48
- [11] Rohini Nair and Kavita Kelkar , " Implementation of Item and Content based Collaborative Filtering Techniques based on Ratings Average for Recommender Systems" ,International Journal of Computer Applications , Vol 65 , No 24 ,March 2013 ,pp 1-4.
- [12] Ujwala Wanaskar, Sheetal R. Vij, Debajyoti Mukhopadhyay, "A Hybrid Web Recommendation System Based on the Improved Association Rule Mining Algorithm", Journal of Software Engineering and Applications, Vol.6 No.8, 2013, pp 396-404.
- [13] Mohd Abdul Hameed, Omar Al Jadaan, S. Ramachandram, " Collaborative Filtering Based Recommendation System : A survey " , International Journal on Computer Science and Engineering , Vol 5, No 5 , May 2012 ,pp 859-876.
- [14] Nana Yaw Asabere, "Towards a Perspective of Hybrid Approaches and Methodologies in Recommender System", International Journal of Computer Science and Telecommunication, Vol 3, Issue 11, Nov 2012, pp 23-32.
- [15] Hongwu Ye, "A Personalized Collaborative Filtering Recommendation using Association rule mining and Self Organizing Map", Journal of Software , Vol 6 , No 4 ,April 2011,pp 732-73
- [16] Francesco Ricci, Lior Rokach and Bracha Shapira, "Introduction to Recommender Systems Handbook", Springer, 2011.
- [17] SongJie Gong, "Employing User Attribute and Item Attribute to enhance the Collaborative Filtering Recommendation", Journal of Software, Vol 4, No 8, 2009, pp 883-890.
- [18] Binge Cui, Xin Chen, "An Online Book Recommendation System based on Web Service", IEEE 6th International Conference on Fuzzy Systems and Knowledge Discovery, 2009, pp 520-524.
- [19] HengSong Tan ,HongWu Ye , "A Collaborative filtering Recommendation Algorithm based on Item Classification " , IEEE Pacific Asia Conference on Circuits , Communication and System , 2009 , pp 694-697.
- [20] Greg Linden , Brent Smith, Jeremy York , "Amazon.com Recommendations Item to Item Collaborative Filtering " , IEEE Internet Computing ,Feb 2003,pp 76-80.
- [21] Sarwar, B., Karypis, G., Konstan, J.A., Reidl, J., "Item-Based Collaborative Filtering Recommendation Algorithms", Proceedings of the 10th international conference on World Wide Web ,Hong Kong. ACM Press, 2011, pp. 285-295.
- [22] <http://www.ibm.com/developerworks/library/os-recommender1/>
- [23] www.google.com