

# A Survey on Various Techniques of Recommendation System in Web Mining

<sup>1</sup>Yagnesh G. patel, <sup>2</sup>Vishal P.Patel  
<sup>1</sup>Department of computer engineering  
<sup>1</sup>S.P.C.E, Visnagar, India

**Abstract** - Today internet has made the life of human dependent on it. Almost everything and anything can be used for discovering useful knowledge or information from the internet. Recommendation systems are defined as the techniques used for predict the rating of the social entity or items. These items can be music, videos, books, movies etc. To predict the user's interest from only past preference is not give particular result. The main objective of area is to solve challenges and issues regarding to finding proper recommend items for users. In this paper we describe the recommendation system related research and then introduces various techniques and approaches used by the recommender system User-based approach, Item-based approach, Hybrid recommendation approaches.

**Keywords** – Web mining; Recommendation system; Content based approach; Collaborative approach; Hybrid approach

## I. INTRODUCTION

Data mining is a process of extract useful information into a large dataset. The massive information available on the World Wide Web has search for data. To discover useful information from the web is called a web mining. E-commerce data are extract using a web mining. Web has deeply promoted the chance to create greater business opportunities and reach to customers easily. This 24\*7 online service has resulted in the large amount of choices, but customers are now faced to information over-load. To overcome this problem Recommendation system is used.

The First recommender system[1] was developed by Goldberg , Nichols , Oki &Terry in 1992. Recommendation system is a special type of Information Filtering (IF) technique that can be used for recommend items to the each specific users. Those items are like movies, books, music, news, images etc. These items are recommend to the users using recommendation techniques. Recommendation systems apply data mining techniques and prediction algorithms to predict users interest on products among the large amount of available items. Recommendation systems are used user's interest and preference for recommend products to particular users.

Most of these techniques are uses approach of Data mining. The process of data mining consist of 3 steps: Data preprocessing, Data analysis and Result interpretation. Example of recommendation systems are amazon.com, ebay[2].

Information are retrieve using 2 ways: Search case and recommendation case. In search case used user queries to find out product and in recommendation case used past preference and user purchase history are used to find product.

## II. BACKGROUND THEORY

Recommendation systems apply data mining techniques and prediction algorithms to predict users interest on information and products among the large amount of available items. Customers are faced the problem of information over-loaded. Recommendation system is a software of predict the useful information regarding product using user's past preference and interest. Using this improve the quality of search items. And benefit for both customers and E-commerce site.

Recommendation systems generally generate a list of recommendations in one of given ways(fig 1)

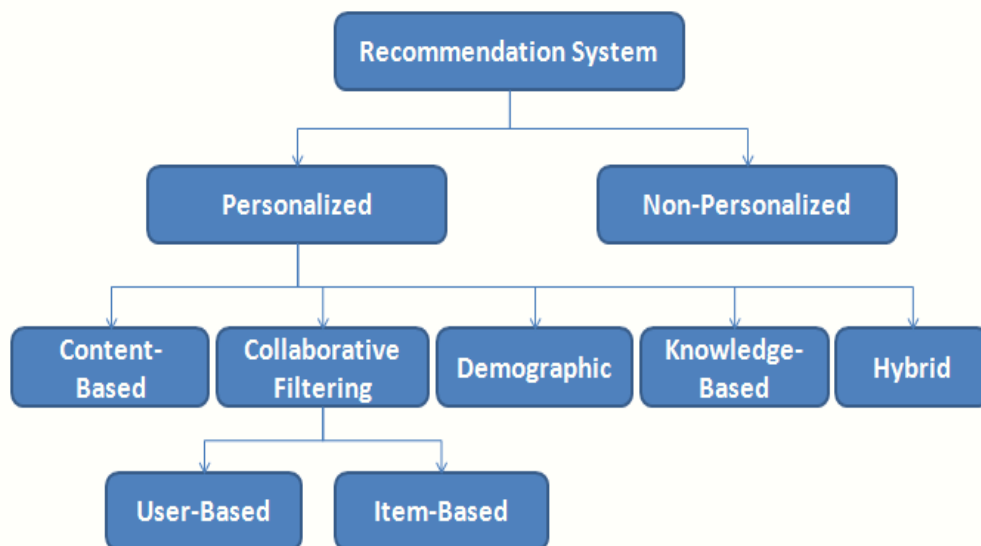


Figure 1 : types of recommendation system[3]

### PERSONALIZED RECOMMENDATION

Personalized recommendation enables the online introduction insertion, suggestion of data in any format that is relevant to each and every user. [3] Personalized recommendation systems are classified into five types depends on their approach to recommendation [3]:

#### 1. Content-Based Filtering

Content based recommendation systems recommend an item to users based on a description of the item and a profile of the user's interests. The recommendation of a content based system are based on personal information and ignore the contributions of other users. The recommendation engine then can find items with the preferred in the past as illustrated in Fig 2.[4]

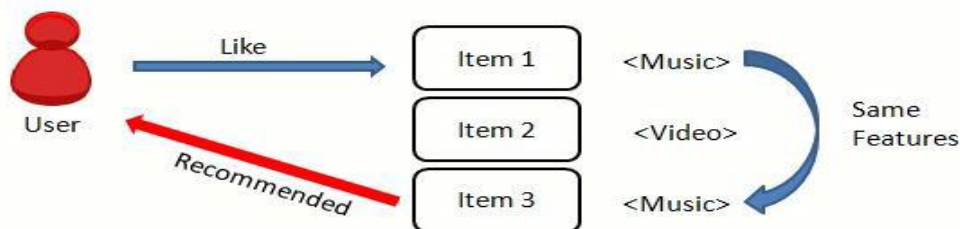


Figure 2: content based recommendation system

There are two types of users implicit users and explicit users. Implicit users means automatically updates as the user interacts with the system. Explicit users means a common feedback technique is the one that allows users to express their opinions by selecting a value of range.

#### 2. Collaborative Filtering:

Collaborative filtering technique based on users history in the form of rating given by the user to an item as their information source [4].

Collaborative filtering approaches often suffer from three problems: cold start, scalability, and sparsity

Collaborative filtering is categorized into three types: user-based, item-based, model-based.

#### User-based approach:

User-based Approach makes recommendation based on interest of the users having the similar taste. It correlates user as per the rating given to the items.

From the Fig. 3 [4], first user related to third user because the rating given by third user is someone similar to the first one. So item 3 also recommend to user 1.

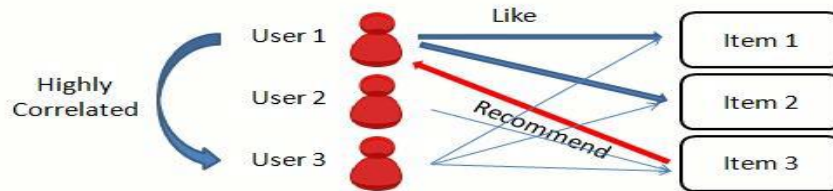


Figure 3: user-based recommendation system

Item-based approach:

Item-based Approach is based on the items as the user rated items similarly. From Fig. 4[4], 2nd and 3rd user rated item 1 and 3 so it assumes that item 1 and 3 are become similar. if first user like item 1 then item 3 is recommended

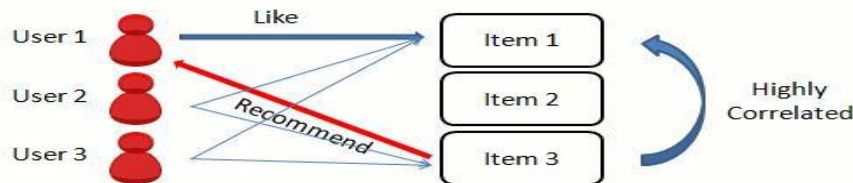


Figure 4: item-based recommendation system

### 3. Demographic approach:

Demographic recommendation technique uses knowledge about user only. The demographic types of users include gender, age, knowledge of languages, disabilities, ethnicity, mobility, employment status, home ownership and location. The system recommends items according to the demographic similarities of the users.[4]

### 4. Knowledge-Based Filtering approach:

Knowledge based recommendation system is based on the explicit knowledge about item classification, user interest. It is an alternative approach to the collaborative filtering and content-based filtering.[4]

### 5. Hybrid approach:

Hybrid approach is a combination of all above types. Combining collaborative filtering and content-based filtering could be more effective in some cases.

Netflix is a good example of the use of hybrid recommender systems. They make recommendations by comparing the watching and searching habits of similar users (i.e. collaborative filtering) as well as by offering movies that share characteristics with films that a user has rated highly (content-based filtering).

Seven hybridization techniques:

- **Weighted:** The score of different recommendation components are combined numerically.
- **Switching:** The system chooses among recommendation components and applies the selected one.
- **Mixed:** Recommendations from different recommenders are presented together.
- **Feature Combination:** Features derived from different knowledge sources are combined together and given to a single recommendation algorithm.
- **Feature Augmentation:** One recommendation technique is used to compute a feature or set of features, which is then part of the input to the next technique.
- **Cascade:** Recommenders are given strict priority, with the lower priority ones breaking ties in the scoring of the higher ones.
- **Meta-level:** One recommendation technique is applied and produces some sort of model, which is then the input used by the next technique.[9]

## NON-PERSONALIZED RECOMMENDATION

Non-personalized recommendation system recommend items to customers based on what other customers have said about the product in an average. The recommendations are independent of the customer, so all customers gets the same recommendation.

## 1. ISSUES IN RECOMMENDATION SYSTEM

### 1. Data Collection[5]

The data used by recommendation systems can be categorized into explicit and implicit data. Explicit is all data that user themselves fill into the system. Implicit data source in e-commerce is the transaction data including the purchase information. Implicit data needs to be analyzed first before it can be used to describe user-item ratings.

### 2. Cold Start[5]

The cold start problem occurs when too little rating data is available in the initial state. The recommendation system then lacks data to produce appropriate recommendations. Two cold start problems are new user problem and new item problem.

### 3. Stability vs. Plasticity[5]

The converse of the cold start problem is the stability vs. plasticity problem. When consumers have rated so many items, their preferences in the established user profiles are difficult to change.

#### 4. Performance & Scalability[5]

Performance and scalability are important issues for recommendation systems as e-commerce websites must be able to determine recommendations in real-time and often deal with huge data sets of millions of users and items. The big growth rates of e-businesses are making the sets even larger in the user dimension.

#### 5.Sparsity: [5]

The number of items sold on major e-commerce sites is extremely large. The most active users will only have rated a small subset of the overall database. Thus, even the most popular items have very few ratings.

#### 6.Scalability:[5]

In many of the environments in which these systems make recommendations, there are millions of users and products. Thus, a large amount of computation power is often necessary to calculate recommendations.

#### 7. Privacy[5]

Privacy is an important issue in recommendation systems. To provide personalized recommendations, recommendation systems must know something regarding to users. In fact, the more the systems know, the more precise the recommendations can get. Users are concerned about what information is gathered, how it is used.

### III. MERITS/DEMERITS

TABLE I MERITS AND DEMERITS OF RECOMMENDATION TECHNIQUES

Sr. No.	Techniques	Merits	Demerits
1	Content-Based Filtering	1.No need for data of other users. 2.No cold start problem. 3.Able to recommend new and unpopular items.	1. Content analysis is necessary to determine the item features. 2. The quality of items cannot be evaluated.
2	Collaborative Filtering	1. For demographic recommendations no knowledge about the item features is needed. 2. It able to make recommendations outside the preferences of the individual user.	1. The quality of the recommendations depends on the size of the historical rating data set. 2. The technique suffers from the cold start problem
3	Demographic	1. It is not based on user-item ratings, it gives recommendation before user rated any item.	1. Gathering of demographic data leads to privacy issues. 2. Stability vs. plasticity problem.

### IV. RELATED WORK

There are many Researchers have worked in this area for finding problem and solutions in recommendation system. Aprojyoti Lopes, Bididha Roy[6] considered the research challenges in collaborative filtering is used user's past data for finding user's behavior. Authors suggest that used action based rational technique that provide recommendation as per changing user's behavior. XI-Ze Heng zhang[7] considered the problem of user's past preference is only used for recommendation. So authors proposed a personalized recommendation system using association rule mining and classification(CBA-CB). CBA-CB classifier is used to predict the item labels for new customers requirement and then assigns the class labels to the new customers.

Ya-min,Xue-ling,Xiao-wei LIU[8] considered the problem that case base reasoning is not giving accurate result for matching past and present cases. So authors suggest that used case based reasoning with web log mining for improving performance.

ZHAO Kai, LU Peng-yu[10] find out the problem that in traditional collaborative filtering methods are inefficient especially when the user-rating data is extremely sparse. So authors suggests ab approach to compuye the user similarity based on user-rating items and then use improved collaborative filtering algorithm.

Xia Min-jie, Zhang Jin-ge[11] considered the problem of information overload in the websites. So customers are face the problem that which product will I select or not. So authors suggest the new methodology is used userID-URL associated matrix and distance matrix to cluster users into groups. Based on visitor's browsing patterns, the site can efficiently and automatically dynamic adjust web page's content and recommend right items.

### V. CONCLUSION

WWW is a source of information where large amount of data is stored. Recommendation systems are helps to user to select any products from the web. Recommendation systems are use to overcome information overload problem as e-commerce continuously growing. Recommendation systems can benefits for both customers and provider. Customers use it by finding new interesting products and provider can enhance their sales. There are various techniques used for recommendations and their

related issues are discussed. Two recommendation systems are widely used that are YouTube.com and Amazon.com for videos and product respectively.

## VI. REFERENCES

- [1] Witten I. H. and Frank I. Data Mining, Morgan Kaufman Publishers, San Francisco, 2000.
- [2] Paul Resnick, Neophytos Iacovou, Mitesh Suchak, Peter Bergstrom, John Riedl, GroupLens: an open architecture for collaborative filtering of netnews, Proceedings of the 1994 ACM conference on Computer supported cooperative work, p.175-186, October 22-26, 1994, Chapel Hill, North Carolina, United States
- [3] Lalita Sharma, Anju Gera, A Survey of Recommendation System: Research Challenges, International Journal of Engineering Trends and Technology (IJETT) - Volume4Issue5- May 2013.
- [4] Tejal Arekar, Mrs. R.S. Sonar, Dr. N. J. Uke, A Survey on Recommendation System, International Journal of Innovative Research in Advanced Engineering (IJIRAE) ISSN: 2349-2163 Volume 2 Issue 1 (January 2015)
- [5] Mukta kohar , Chhavi Rana, Survey Paper on Recommendation System, (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 3 (2) , 2012,3460-3462
- [6] Aprojyoti Lopes,Bididha Roy, "Dynamic recommendation system using web usage mining for e-commerce users", 2015 SCIENCE DIRECT
- [7] XI-Ze Heng zhang," Building personalize recommendation system in e-commerce using association rule-based mining and classification" 2007 IEEE
- [8] Ya-min,Xue-ling,Xiao-wei LIU, "E-commerce recommendation system based on CBR and web log mining"2011 IEEE
- [9] Robin Burke , Hybrid Web Recommender Systems, pp. 377-408, The Adaptive Web, Peter Brusilovsky, Alfred Kobsa, Wolfgang Nejdl (Ed.), Lecture Notes in Computer Science, Springer-Verlag, Berlin, Germany, Lecture Notes in Computer Science, Vol. 4321, May 2007, 978-3-540-72078-2.
- [10] ZHAO Kai, LU Peng-yu, Improved Collaborative Filtering Approach Based on User Similarity Combination, 2014 International Conference on Management Science & Engineering (21th), 2014 IEEE
- [11] Xia Min-jie, Zhang Jin-ge, Research on Personalized Recommendation System for e-Commerce based on Web Log Mining and User Browsing Behaviors, 2010 International Conference on Computer Application and System Modeling (ICCASM 2010), 2010 IEEE

