IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-6, June 2018 E-ISSN: 2347-2693

Item Recommendation Using Hybrid Method

M. Munafur Hussaina[1*], R. Parimala[2],

[1] P.G and Research Dept. of CS, Periyar E.V.R.College, Bharathidasan University, Trichy-23, India.
[2] P.G and Research Dept. of CS, Periyar E.V.R.College, Bharathidasan University, Trichy-23, India.

[*]*Corresponding Author: munafurjalal@gmail.com, Tel.: +9942114566*

Available online at: www.ijcseonline.org

Accepted: 10/Jun/2018, Published: 30/Jun/2018

***Abstract*—** Recommender System provides various choices of the user preferences for suggesting the product/service to purchase. Collaborative filtering is one of the techniques in Recommender system used to find reviews and ratings of the users for similar products or users. To improve the performance of the recommendation, methods have been sometimes combined in hybrid recommenders. In this paper, the researcher have proposed an item based recommendation using Hybrid method called Item Recommendation Using Hybrid Method (IRHM), based on collaborative filtering approach that recommends the user for choosing the best item. The aim of the paper is to find the maximum value of precision and recall in hybrid method.

***Keywords*—** Item, Movie, Recommendation System, Hybrid, Collaborative Filtering, IRHM

I. INTRODUCTION

In today's world the challenge for online ecommerce lies on how to find a good product from variety of options. Users face lots of problems to look for a good opinion. Recommender systems analyze user preferences with the purpose of recommending items to users that meet their needs and satisfaction. Recommendation systems are used to help users to make decision about products, information or service among many for their preference to satisfy them. There are number of online recommendation systems available in the internet for the user to utilize in different ways. Examples of recommendation systems are recommendation on songs, movies, jokes, books, travel destination and e-learning etc. Recommender System's work is based on user's information from various sources and provides recommendation of good products. The ultimate goal is to bring out suggestions to users to make better decisions from many alternatives available over the web.

The decision and suggestion of the users can be stored in the recommender database and may be used for generating new recommender system in the next user-system interactions. Recommendation systems are very useful to both service providers and customers. They reduce online transaction cost when finding or selecting an item in a shopping environment. Recommendation system enhances revenues in ecommerce by effective way of selling more products. It also supports the users by allowing them to search beyond the catalogue. Recommender system can be classified as the collaborative based recommender, the content based recommender and the hybrid recommender.

The Collaborative recommendation is the most familiar and widely accepted technologies.The collaborative filtering is the process of filtering for information or patterns using collaborative techniques among viewpoints, data sources etc. Collaborative filtering methods have been applied to many different kinds of data including: sensing and monitoring data, such as in mineral exploration, environmental sensing over large areas or multiple sensors; financial data, such as financial service institutions that integrate many financial sources; or in electronic commerce and web applications where the focus is on user data, etc. The collaborative filtering based recommendation system generates ratings for active user based on rating given by recommender system users who are similar to active user [1]. The Collaborative filtering algorithms are based on information about similar users or similar items. They can be categorized as

Item-based collaborative filtering: This recommends to a user the items that are most similar to the user's purchases.

User-based collaborative filtering: This recommends to a user the items that are the most preferred by similar users.

Collaborative filtering system like user based collaborative filtering (UBCF) system recommends items based on similarity measures between users and/or items. The system recommends items that are prepared by similar kind of users. It provides unexpected recommendation based on user's similarity rather than items similarity. Since in Collaborative Filtering people make explicit ratings so that real quality

assessment of items are done. Content-based filtering defines a user profile and identifies the items that match it.

Content based filtering is based on the profile of user's preference and the item's description. In Content Based Filtering items are described by use of keywords apart from user's profile. CBF algorithm recommends best rated items that were liked in the past. Hybrid system is the combination of algorithms which provide more effective and accurate recommendations than a single algorithm done. Using of multiple recommendation techniques can suppress the weakness of an individual technique in a combined model. The combination of content based system and collaborative system is recommends products based on both the product attributes and user rating [2]. In proposed system, we combine the results of different models as hybrid to recommend the items/ products having incremental value of precision and recall for the user or customer to choose the right one. The structure of the paper is organized as follows: Section I contains the introduction of collaborative filtering, content based filtering and Hybrid method. Section II contains related work of collaborative filtering in Hybrid method. Section III discusses the Recommender approaches we have used in this work, Section IV describes the Proposed Model towards Item Recommendation using Hybrid System, Section V explains the Methodology with flow chart and algorithm, Section VI describes the Results and Findings of the proposed model and Section VII concludes the research work with Future directions.

II. RELATED WORK

Prerna Agarwal et.al [3] the author recommends the Indian movie datasets for analysing the ratings in regional languages from various demographics of users. He analysed the data with supervised and unsupervised collaborative filtering techniques. This version of dataset with more number of ratings and users will help to improve the current state of recommender systems for the Indian audience.

Sarwar et.al [4] investigated several techniques for analysing large scale purchase and preference data for the purpose of producing useful recommendations to customers. He devises different techniques for different sub processes and applies their combinations on datasets to compare for recommendation quality and performance.

R. Suguna et.al [5] suggested that web usage mining is considered as the major source for web recommendation in association with collaborative filtering approach. Markov model and association rule mining are used to recommend the webpage to the user based on user's past history. The traditional apriori algorithm helps to improve the time duration spent on each web page.

Mohamed Koutheair et.al [6] developed a novel recommendation system for the students who are all themselves learning technologies through e-learning environment using web usage mining. Their recommendation system will automatically suggest the educational resources for the students based on browsing history.

Hamidreza Koohi et.al [7] proposed a fuzzy C-means approach for user-based Collaborative Filtering and its performance against different clustering approaches have been assessed. The Movie Lens dataset is used to compare different clustering algorithms to evaluate the recommendation accuracy, precision and recall using Pearson correlation coefficient can yield better recommendation results, compared to other techniques.

Ziming Zeng [8] conducted an experiment to evaluate its recommender quality and show that the system gives sensible recommendations and is able to help customers save enormous time for internet shopping. The system utilizes web mining techniques to trace the customer's shopping behavior and learn their up-to-date preference adaptively.

Shivani et.al [9] defined an algorithm called Behaviour based rational technique that uses brands and categories of each visited product for tracking user's dynamic behavior, uses popularity measures for accurate recommendation list of items. The proposed system yields good accuracy and diminishes the limitations of traditional system.

Amer Al-Badarenah et.al [10] discussed a new collaborative recommendation system that employed association rules algorithm to recommend university elective courses to students by exploiting courses that other similar students had taken.

Mohammad et.al [11] summarized a text mining approach to mine product features, opinions and their semantic similarity from web opinion sources. The consumer can clearly see the strength and weakness of each product in the minds of existing consumer. The system assists online shoppers by suggesting the most effective navigation products for their specified criteria and preferences.

Robin Burke [12] reviewed the landscape of actual and possible hybrid recommenders, and introduces a novel hybrid, EntreeC, a system that combines knowledge-based recommendation and collaborative filtering to recommend restaurants. The semantic ratings obtained from the knowledge-based part of the system enhance the effectiveness of collaborative filtering.

III. RECOMMENDER APPROACHES

Recommender systems have become omnipresent and important in various areas such as for recommending movies, music, news, books etc in recent years. The most popular algorithms used to produce a list of recommendations are collaborative filtering, which is used to calculate similarities among many items or users. Similarity measure is defined as the methods used to calculate the score that is expressed by similar users or items. These scores can be used as the foundation of user-or item-based recommendation generation. Similarity measures are measures between users and items with the help of filtering techniques. The models used in filtering techniques are random, popular, UBCF and

IBCF. The Random model is used to select the items randomly from the list and produce the recommendation. The popular model produces recommendation based on popularity of the items. The UBCF and IBCF display the recommendations based on user based collaborative filtering and IBCF is item-based collaborative filtering that helps to make a prediction [13]. These collaborative filtering methods are measured by similarity measures such as Cosine, Pearson correlation, Jaccard, Euclidean etc.

Cosine similarity is a measure of similarity between two non-zero vector of an inner product space that measure cosine of the angle between them. Two vectors with the same orientation have a cosine similarity of 1, two vectors at 90° have a similarity of 0, and two vectors diametrically opposed have a similarity of -1, independent of their magnitude. Cosine similarity gives a useful measure between similar items. It is very efficient to evaluate, especially sparse vectors, only for non-zero dimensions [14].

$$cosine\ similarity = \frac{crossprod(a,b)}{sqrt(crossprod(a) * crossprod(b))}$$

Pearson correlation is a measure used to check how correlated two variables (vectors) are. This measure is calculating the similarities between users in recommendation systems since all users can be represented as vectors of ratings. Euclidean distance is a metric used to measure distance between vectors. It is calculated by directly comparing how each pair of ratings differs [14]. The Jaccard distance measures the proximity of two vectors A and B, both are in binary features. It ranges from zero to one and viewed as a proportion of common features that the two vectors share in the total number of features present in both [15].

IV. PROPOSED MODEL

Recommendation system is used to collect information of user preferences from various sources and provide better/best suggestion to make decisions of the products/items. The recommendation tasks are performed based on the outputs. They can be rating prediction or ranking prediction. Rating prediction fills the missing entries of the user item. Top-N recommendations produce the ranked list of items or users. In proposed system, we use the hybrid recommendation techniques by combing the values of recommendation model to evaluate the result. Our aim is to maximize the precision and recall value for the items.

V. METHODOLOGY

Different methodologies are used to find the best outcomes of Recommender System. All these methods are used to evaluate and perform the result for similarity measures to the user to make decision for the best suit / preference to select among the products. Figure 1 shows the flow of proposed system. The different models of recommender systems are calculated in the proposed system using IRHM algorithm. In proposed system we use the cosine similarity distance to measure the similarity between similar users or items.

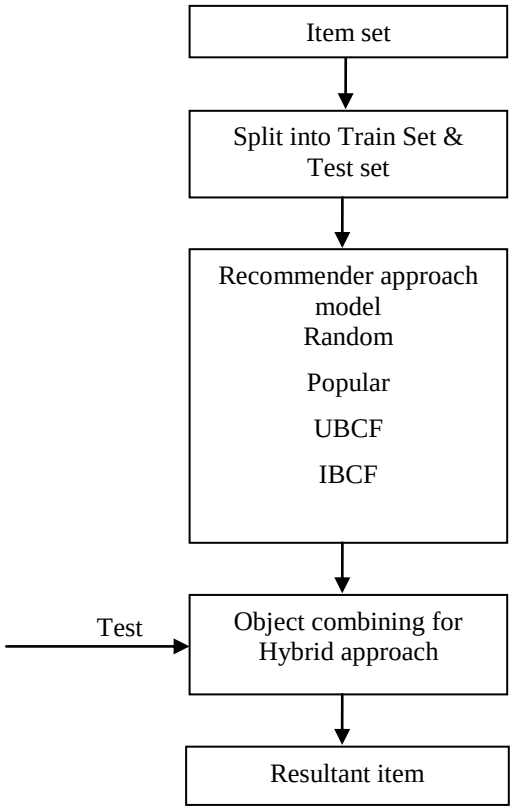

Figure1 Flow diagram of IRHM

A. Datasets:

Dataset like movie lens, jester 5K and MSWeb are used for recommendation system to recommend the output for Hybrid recommender systems. The movie lens dataset contains 943 users and 1664 movies with 100,000 ratings. The data was collected through the movie lens website. The jester5K dataset contains a sample of 5000 users and 100 jokes from the anonymous of 100 rating data from the jester online joke recommender system. The MSWeb data was created by sampling and processing the WWW.Microsoft.com logs [15].

B. Algorithm:

Algorithm of IRHM

Step1: Collect the item set and given as input
Step2: Split the items into train sets and test sets
Step3: Implement the recommender system approach
Step4: Output of step3 is combined for Hybrid technique
Step5: Display the output with Precision and Recall

VI. RESULTS AND FINDINGS

The datasets can be collected and split into training set and test set for the models of recommender system to implement. Filtering models like Random, popular, user based collaborative filtering and item based filtering are used to measure the similarity for split method. The splitting method randomly assigns a predefined proportion of the users to the training set and all others to the test set. The similarity measure is used to calculate the scores that express similar users or items. These scores can be used as the base for user or item based recommendation generation. The split method consists of 80% training data and 20% test data for predicting the results. The evaluation scheme which is used to perform the evaluation of predictive results for the given dataset of k fold values for given number of items/runs with good ratings for the users. The output of the evaluation result is calculated for finding average of precision, recall, true positive rate, false positive rate etc. Precision is a measure of exactness and recall is a measure of completeness.

$$\text{Precision} = \frac{|\text{test} \cap \text{top} - \text{N}|}{\text{N}}$$

$$\text{Recall} = \frac{|\text{test} \cap \text{top} - \text{N}|}{|\text{test}|}$$

Where top-N denotes the recommendation set, test denotes the test set and N denotes the number of recommendation. Since increasing the size of recommendation set leads to an increase in recall but at the same time a decrease in precision [16]. The algorithm of IRHM (Item Recommendation using Hybrid Method) to be implemented for predicting the item recommends to user. The results of evaluation scheme for different datasets have been recorded in the Table 1. The dataset Jester5k and MSWeb have realRatingMatrix of 362106 ratings with good ratings>=4 for training set of 80% and test set of 20%. The Movie Lens dataset having realRatingMatrix of 99392 ratings with good ratings>=4 for same training and test set with one k fold value. List of recommender techniques are tested and stored in the algorithm variable to evaluate. Cosine similarity is used with z-score normalize form for normalizing the datasets. Evaluate the prediction result for random of n items (like 1, 3, 5, 10, 15, 20) are taken. Finally the average results of precision and recall values are combined for hybrid recommender system. The predictive result of hybrid recommender systems of precision and recall curve are displayed in the figure 2. The table1 shows the model time and prediction time for different datasets with one number fold /run value. The prediction time for jester5k dataset records 0.5sec, movie lens records 1.78sec and MSWeb records 0.65sec of time for random model. Like the same other models also records different prediction time for different data sets. Table 2 shows the precision and recall value for proposed system. The precision value gradually starts from 0.16 at random model, having a difference of three for popular model as 0.19, UBCF having 0.21 and IBCF ends with 0.24. Like precision, the recall also records the value of 0.030 for random model, 0.034 for popular model, 0.036 for UBCF and 0.038 for IBCF models.

Table 1 Display of Model time and Prediction time

Models	Datasets		
	Jester5k	***Movie Lens***	***MSWeb***
Random	0sec/0.5sec	0.01sec/1.78sec	0.02sec/0.65sec
Popular	0.4sec/4.58sec	0.11sec/1.71sec	0.36sec/4.02sec
UBCF	0.28sec/10.55sec	0.11sec/5.06sec	0.28sec/10.68sec
IBCF	0.92sec/0.54sec	80.12sec/0.19sec	1.3sec/0.74sec

Table 2 Display of Precision and Recall value for proposed model

Models	Precision	Recall
Random	0.167964142	0.030426814
Popular	0.190060998	0.034568065
UBCF	0.212298946	0.036843861
IBCF	0.240488798	0.038705652

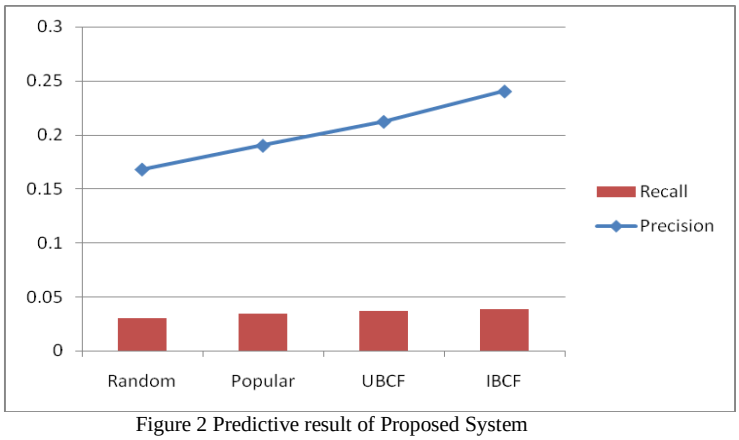

Figure 2 Predictive result of Proposed System

VII. CONCLUSION

Recommender System recommends the products or service to the user for avoiding confusion when taking decision among many. Using of ratings gives clarity to the user for the products. The system focuses on recommendation of items /products to the users. Proposed systems for item recommendation using Hybrid method have competence to attract new customers and retain existing ones. Through experiments we find the maximum value of precision and recall performance for the Hybrid items recommendation system. Future recommendations for other datasets are taken for further research consideration.

ACKNOWLEDGMENT

We first thank the Almighty and thank all those who have helped us to make this paper a successful one and giving

several openings in research. I also thank the R Community for providing this open source software towards successful implementation of this research work.

REFERENCES

[1] Saurabh Kumar Tiwari, Shailendra Kumar Shrivastava, "*An Approach for Recommender System by Combining Collaborative Filtering with User Demographics and Items Genre*s", International Journal of Computer Applications, Vol.128, No.13, pp. 16-24, 2015.

[2] Manoj Kumar, D.K. Yadav, Ankur Singh, Vijay Kr. Gupta, "*A Movie Recommender System: MOVREC*", International Journal of Computer Applications, Vol.124, No.3, pp.7-11, 2015.

[3] Prerna Agarwal, Richa Verma, Angshul Majumdar , "*Indian Regional Movie Dataset for Recommender Systems*", https://arxiv.org/pdf/1801.02203.pdf, arXiv:1801.02203v1 [cs.IR], pp.1-7, 2018.

[4] Badrul Sarwar, George Karypis, Joseph Konstan, John Riedl, "*Analysis of recommendation algorithms for e-commerce*", In the Proceedings of the 2nd ACM Conference on Electronic Commerce, Minnesota, USA, pp.158-167, 2000.

[5] R. Suguna, D. Sharmila, "*An Efficient Web Recommendation System using Collaborative Filtering and Pattern Discovery Algorithms*", International Journal of Computer Applications, Vol. 70, No.3, pp.37-44, 2013.

[6] Mohamed Koutheair Khribi, Mohamed Jemnil, Olfa Nasraoui, "*Automatic Recommendations for E-Learning Personalization based on Web Usage Mining Techniques and Information Retrieval*", Educational Technology and Society, 12(4), pp.30-42, 2009.

[7] Hamidreza Koohi, Kourosh Kiani, "*User based Collaborative Filtering using fuzzy C-means*", Science Direct, Measurement 91, pp.134–139, 2016.

[8] Ziming Zeng, "*An Intelligent E-Commerce Recommender System Based on Web Mining*", International Journal Business and Management, Vol.4, No.7, pp.10-14, 2009.

[9] Shivani Diwan, Komal Dani, Sahil Desai, "*Dynamic Recommendation System for E-Commerce Users*", International Research Journal of Engineering and Technology, Vol.03, Issue. 05, pp.141-144, 2016.

[10] Amer Al-Badarenah, Jamal Alsakran, "*An Automated Recommender System for Course Selection*", International Journal of Advanced Computer Science and Applications, Vol.7, No.3, pp. 166- 173, 2016.

[11] Mohammad Daoud, S.K. Naqvi, Asad Ahmad, "*Opinion Observer:Recommendation System on E-Commerce Website*", International Journal of Computer Applications, Vol.105, No.14, pp.37-42, 2014.

[12] Robin Burke, "*Hybrid Recommender Systems: Survey and Experiments*", User Modeling and User - Adapted Interaction", Vol.12, Issue.4, pp. 331-370, 2002.

[13] Senthil Kumar Thangavel, Neetha Susan Thampi, Johnpaul C I, "*Performance Analysis of Various Recommendation Algorithms Using Apache Hadoop and Mahout*", International Journal of Scientific & Engineering Research, Vol.4, Issue 12, pp.279-287, 2013.

[14] Suresh K. Gorakala, Michele Usuelli, "*Building a Recommendation System with R*", Packt Publishing Ltd, UK, pp. 9-10, 2015.

[15] Michael Hahsler, "*Recommenderlab: A Framework for Developing and Testing Recommendation Algorithms*", Southern Methodist University, pp.1-40, 2011.

[16] Mojtaba Salehi, "*An effective recommendation based on user behaviour: a hybrid of sequential pattern of user and attributes of product*", International Journal of Business Information Systems, Vol.14, No.4, pp.480-496, 2013.

Authors Profile

M.Munafur Hussaina is a Research Scholar and currently working as Assistant Professor in Aiman College of Arts and Science for Women, Tiruchirapalli. She receieved her Master degree of Computer Science in 2002 from Madurai Kamaraj University, Madurai and M.Phil Computer Science in 2011 from Bharathidasan University, Tiruchirapalli. Her area of research work focuses on Data Mining, Web Usage Mining and Web Analytics.

R.Parimala graduated with M.Sc. Applied Science at the National Institute of Technology (formerly Regional Engineering College) Tiruchirappalli in 1990. She received her M.Phil Computer Science at Mother Teresa University, Kodaikanal in 1999. She started teaching in 1999 at National Institute of Technology and is currently working as Assistant Professor in Department of Computer Science, Periyar E.V.R. College (Autonomous), Tiruchirappalli. She completed her Ph.D. at National Institute of Technology, Tiruchirappalli. Her area of research interests includes Neural Networks, Data Mining and Optimization Techniques.

