ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 4, Issue 6 , June 2017

An Efficient Domain Specific Recommendation System for Subgroup Analysis

N. S. Barley, Dr. R. R. Keole

P.G. Student, Dept.of Comp Sci & Info Tech, H.V.P.M's C.O.E.T. Amravati, Maharashtra, INDIA
Associate Prof. P.G. Dept.of Comp Sci & Info Tech, H.V.P.M's C.O.E.T. Amravati, Maharashtra, INDIA

ABSTRACT: With the wide variety of products and services available on the web, it is difficult for users to choose the product or service that most meets their needs. In order to reduce or even eliminate this difficulty, recommender systems have emerged. A recommender system is used in various fields to recommend items of interest to users. One of the main areas where this concept is currently used is e-commerce that interacts directly with customers by suggesting products of interest with the aim of improving its sales. Motivated by the observation, a novel Domain-sensitive Recommendation (DsRec) algorithm is proposed, to make the rating prediction by exploring the user-item subgroup analysis simultaneously, in which a user-item subgroup is deemed as a domain consisting of a subset of items with similar attributes and a subset of users who have interests in these items. Collaborative Filtering (CF) is an effective and widely adopted recommendation approach. Different from content-based recommender systems which rely on the profiles of users and items for predictions, CF approaches make predictions by only utilizing the user-item interaction information such as transaction history or item satisfaction expressed in ratings, etc.

KEYWORDS*:* Recommender system, user-item subgroup, collaborative filtering.

I.INTRODUCTION

Collaborative Filtering (CF) is one of the most successful recommendation approaches to cope with information overload in the real world. However, typical CF methods equally treat every user and item, and cannot distinguish the variation of user's interests across different domains. This violates the reality that user's interests always center on some specific domains, and the users having similar tastes on one domain may have totally different tastes on another domain. Different CF techniques can be classified into two classes: memory-based methods and model based methods.

Memory-based methods focus on finding similar users or items [12] for recommendation. These algorithms generally comply with the following steps:
1) Calculate the similarity which reflects correlation between two users or two items.
2) Produce a prediction for the active user based on the ratings of similar users found, or based on the computed information of items similar to those chosen by the active user. Such recommender systems are easy to implement and highly effective. However, there are several limitations for the memory-based CF techniques, such as sparsity and scalability.

To overcome the shortcomings of memory-based methods, model-based approaches have been proposed and actively studied, which establish a model using the observed ratings that can interpret the given data and predict the unknown ratings. Sometimes the time required for finding the rating predictions is more but still don't create the quality predictions. The proposed approach is used to overcome such problem and utilize less time and give better predictions [11].

A. Limitations of Existing Methods

Some limitations of research existing techniques are as follows:

1. Memory based algorithm uses the entire database every time it makes a prediction, so it is very slow.

ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 4, Issue 6 , June 2017

2. The fact that Model-based system do not use all the information (the whole dataset) available in database, it is possible that with model-based systems don't generate accurate predictions.
3. Building a model is often a time- and resource-consuming process, so it is usually more difficult to add data to model-based systems, making them inflexible.

II. PROPOSED METHODOLOGY

Domain-sensitive Recommendation Algorithm is proposed in this paper. This method provides the efficient way to identify the domain and them grouping them into subgroups depending on area of interest. So the performance of the Recommendation Systems will improve and give better recommendations. This paper gives the efficient way to improve the performance of the existing methods, also gives efficient recommendation by combining the existing methods to improve the performance of the systems and better prediction can be made using proposed method.

A. Architecture of Proposed System

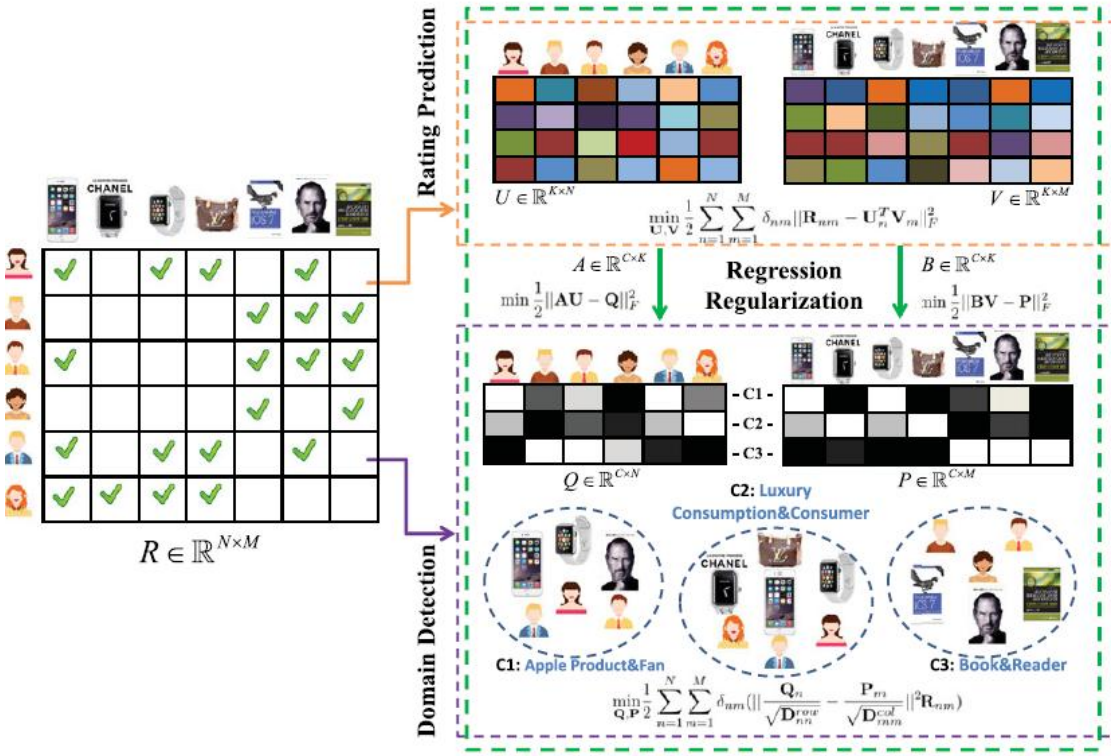

Fig 1: Architecture of the proposed DsRec

- **Rating Prediction Model**

As a typical solution, matrix factorization is adopted for rating prediction in our work. Suppose we have a user item rating matrix describing N users' numerical ratings on M items. Since in the real-world, each user usually rates a very small portion of items, the matrix R is extremely sparse. A matrix factorization approach seeks to approximate the rating matrix R by a multiplication of K-rank factors.

- **Domain Detection Model**

In this step, we will systematically interpret how to detect user-item subgroups (domains) with a bi-clustering model, which is also a two-sided clustering solution. It has been shown that the two-sided clustering often yields

ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 4, Issue 6 , June 2017

impressive performance over traditional one-sided clustering algorithms. More importantly, the resulting co-clustered subgroups may reveal valuable insights from the item attributes. For example, John likes both iPhone6Plus and a Louis Vuitton bag. If our bi-clustering model groups John, iPhone6Plus, and Louis Vuitton together, we can explain the domain is related to luxury consumption, since both iPhone6Plus and Louis Vuitton are luxury products.

- **Regression regularization**

In the regression regularization is that the latent factor representations of users (and items) are required to reflect the preferences of users (and the attributes of items) across different domains. In other words, the latent factor representations should be discriminative enough to find the subgroup distributions of users and items, and in turn some domain information should be embedded into the latent factor representations for rating prediction.

B. The Proposed Solution

The solution that proposed here is basically helps in reducing the extra efforts to recommend good products to users. The working of system in stepwise representation is as follows:

Step 1: (Initial step)
Creating new user account and searching product of interest
Step 2: (Storing Process)
Store user's history and user details in the database.
Step 3: (Evaluation Process)
Evaluate the recommendations using user history stored in database.
Step 4: (Clustering Process)
Forming High Rating product clusters
Step 5: (Analysis Process)
Creating the graph of count v/s no. of user reviews
Step 6: (Recommendation Process)
Recommend product to user according to user interest and the analysis.

III. SYSTEM OVERVIEW WORKING

A. Flowchart

Following flowchart describes the proposed system step by step

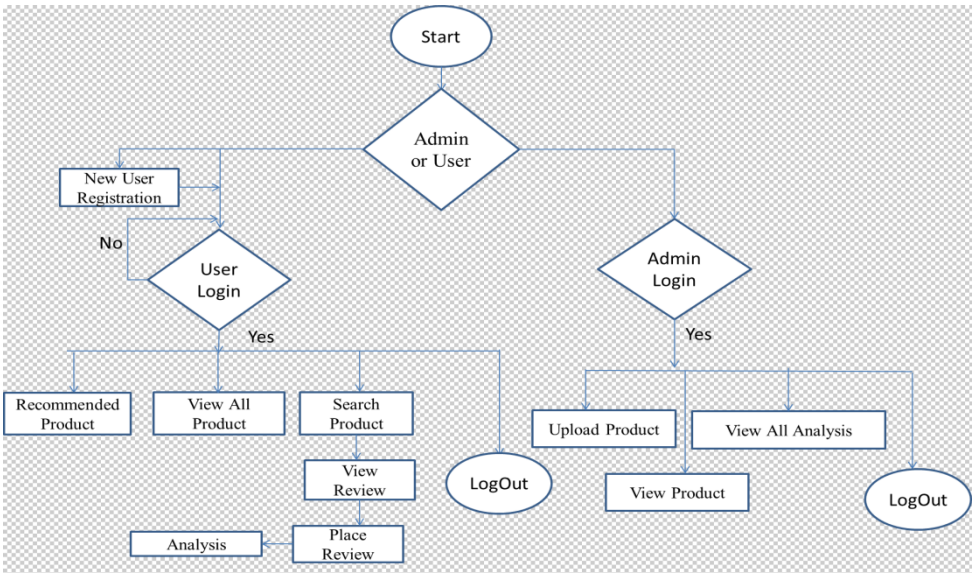

Fig 2: Flowchart of proposed system

ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 4, Issue 6 , June 2017

B. Modelling of Proposed Methodology

Modelling is the representation of a method which is used by simulation. Models may be mathematical, physical, or logical representations of a system, entity, phenomenon, or process. Models are, in turn, used by simulation to predict a future state. The Models of proposed method are as follows:

- **Domain-sensitive Recommendation Systems**

Input : Any real time database of e-commerce website.
Output:

a. Rating prediction model.
b. Domain detection model.
c. Regression regularization terms
d. Group of Item Analysis

Algorithm
Step 1: Create user account
Step 2: Login to user account
Step 3: Searching of items
Step 4: View product of interest
Step 5: Store user searching history in database
Step 7: If user=new then go to step 1 and repeat the procedure
Else
Step 8: Find the ratings and recommend top rating products

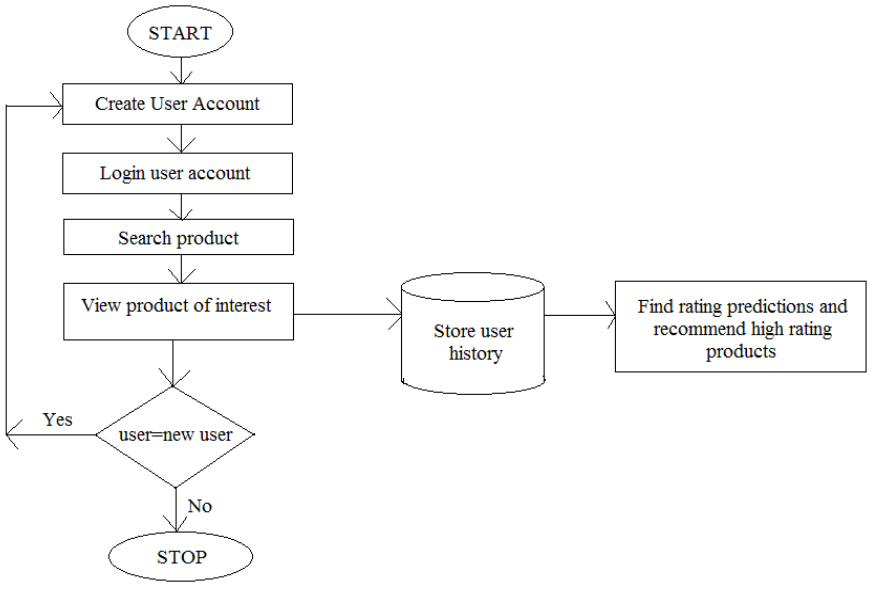

Fig 3: Flowchart of DsRec algorithm

C. Implementation of Proposed Work

Implementation encompasses all the processes involved in getting new software or hardware operating properly in its environment, including installation, configuration, and running, testing, and making necessary changes. As such, implementation is the action that must be followed by users as any preliminary things in order to see how it works.
The issues considered for implementing new approach are,

ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 4, Issue 6 , June 2017

1) Prediction Calculation Time

It considers the processing time of calculation using the contextual information to find the appropriate rating predictions. If the prediction calculation requires more time for ratings then performance of method will decrease.

2) Prediction Quality

It considers the quality of the predictions required in the recommendation systems. If the quality of the predictions is bad, then performance will decrease.

IV. RESULT ANALYSIS

This part focuses on result and its analysis based on the tools used for this method. The tools help in analyzing the user and the admin analysis graphs which help both user and the admin to know which products are the topic of discussion recently. To evaluate the behavior of the user the parameters like user interest, count, user reviews are used.
An example to show the analysis is given below.

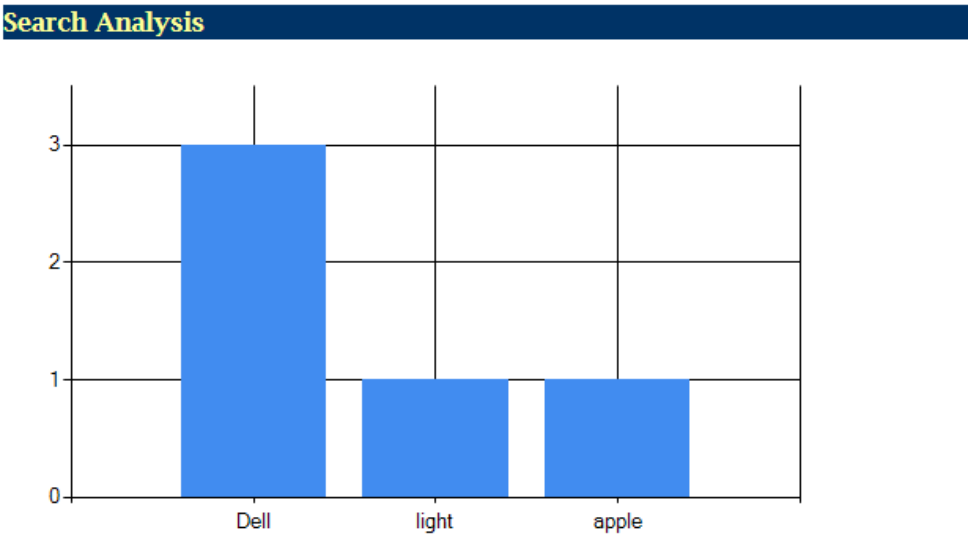

Fig4: Result of Analysis

This analysis is nothing but the graph between name of product and count. The count is no. of reviews on the product. These graphs shows the top products which are more discussed among users. This recommends users the top product and helps user to choose better product.

A. Advantages

1. The performance is better in terms of quality and time.
2. It provides the better use of the database which store user and product history.
3. Quality prediction, Scalability, Prediction, speed are the main advantages of the proposed scheme.

V. CONCLUSION AND FUTURE WORK

The proposed method provides the efficient approach to find the appropriate rating predictions. In other words provides the better recommendations to the users of same interest. This approach calculates the ratings of the items searched by the users. The reviews on the product help in the analysis phase. The use of few recommendations optimize the less

ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 4, Issue 6 , June 2017

time required for calculation of ratings using all recommendations. It makes the analysis phase easy. It also takes less time to find quality recommendations. In future the investigation of how to use the changes in customer behaviour based on the long and short term preferences in order to improve recommendations can also help to achieve the goal.

REFERENCES

[1] Y. Zhang, B. Cao, and D.-Y.Yeung, "Multi-domain collaborative filtering," in Proc. 26th Conf. Annu.Conf. Uncertainty Artif.Intell., 2010, pp. 725–732.
[2] X. Zhang, J. Cheng, T. Yuan, B. Niu, and H. Lu, "TopRec: Domainspecific recommendation through community topic mining in social network," in Proc. 22nd Int. Conf. World Wide Web, 2013, pp. 1501–1510.
[3] Y. Jiang, J. Liu, X. Zhang, Z. Li, and H. Lu, "TCRec: Product recommendation via exploiting social-trust network and product category information," in Proc. 22nd Int. Conf. World Wide Web, 2013, pp. 233–234.
[4] S. Han, S. Chee, J. Han, and K. Wang, "RecTree: An efficient collaborative filtering method," in Proc. 3rd Int. Conf. Data Warehousing Knowl. Discovery, 2001, pp. 141 151.
[5] B. M. Sarwar, J. Konstan, and J. Riedl, "Recommender systems for large-scale e-commerce: scalable neighborhood formation using clustering," in Proc. 5th Int. Conf. Comput. Inf. Technol., 2002, pp. 1–6.
[6] G.-R. Xue, C. Lin, Q. Yang, W. Xi, H.-J.Zeng, Y. Yu, and Z. Chen, "Scalable collaborative filtering using cluster-based smoothing," in Proc. Res. Develop. Inf. Retrieval, 2005, pp. 114–121.
[7] A. Bellogin and J. Parapar, "Using graph partitioning techniques for neighbour selection in user-based collaborative filtering," in Proc. 6th ACM Conf. Recommender Syst., 2012, pp. 213–216.
[8] M. O'Connor and J. Herlocker, "Clustering items for collaborative filtering," in Proc. ACM SIGIR Workshop Recommender Syst., 1999, pp. 1–4.
[9] L. Ungar, D. Foster, E. Andre, S. Wars, F. S. Wars, D. S. Wars, and J. H. Whispers, "Clustering methods for collaborative filtering," in Proc. Workshop Recommendation Syst., 1998, pp. 114–129.
[10] L. Si and R. Jin, "Flexible mixture model for collaborative filtering," in Proc. 20th Int. Conf. Mach. Learn., 2003, pp. 704–711
[11] T. George and S. Merugu, "A scalable collaborative filtering framework based on co-clustering," in Proc. 5th IEEE Int. Conf. Data Mining, 2005, pp. 625–628.

[12] B. M. Sarwar, G. Karypis, J. A. Konstan, and J. Reidl, "Item-based collaborative filtering recommendation algorithms," in Proc. 10th Int. World Wide Web Conf., 2001, pp. 285–295.
[13] T. Hofmann and J. Puzicha, "Latent class models for collaborative filtering," in Proc. 6th Int. Joint Conf. Artif.Intell., 1999, pp. 688–693.
[14] S. Zhang, W. Wang, J. Ford, and F. Makedon, "Learning from incomplete ratings using non-negative matrix factorization," in Proc. 6th SIAM Int. Conf. Data Mining, 2006, pp. 549–553.
[15] I. Esslimani, A. Brun, and A. Boyer, "A collaborative filtering approach combining clustering and navigational based correlations," in Proc. 5th Int. Conf. Web Inf. Syst. Technol., 2009, pp. 364–369.
[16] Jing Liu, Yu Jiang, Zechao Li, Xi Zhang. "Domain-Sensitive Recommendation with User-Item Subgroup Analysis" In: *IEEE Transaction on Knowledge and Data Engineering*, Vol. 28, No. 4, April 2016.