

E-Commerce Recommender System using Data mining Algorithms

Amogh D¹, Mohammed Fuwad Anjum², Navya N³, B N Kiran⁴

^{1,2,3,4}Eighth Semester, Dept. of ISE, The National Institute of Engineering, Mysuru

Abstract - E-Commerce businesses are constantly working on finding better ways to run the business more efficiently and effectively by improvising their businesses in terms of attracting new customers and also by retaining old customers. Several Recommendation systems have been developed and implemented to improve customer satisfaction. This paper sheds light on some of the problematic situations related to customer satisfaction such as Limited resource, cold start situations and also on data valid time constraint, which have been less considered as an important issue to resolve. This paper also proposes the improved version of Hybrid and Random algorithms along with collaborative filtering method using data mining technique for encountering the above stated problems.

The Hybrid algorithm is used for recommending the products which are in limited quantity only to a particular customer who is keen on buying it. This is achieved by analyzing the transaction history of the user. An add up to the standard Hybrid algorithm is that we have also used a collaborative filtering method which also takes into consideration the transaction history of other customers with similar area of interest with that of the customer's. The Random algorithm is used to recommend products in random to a new customer by analyzing the transaction history of other customers. The proposed approach contributes to the betterment of E-Commerce businesses.

Key Words: Recommender System, E-Commerce, Data mining, Hybrid Algorithm, Random Algorithm.

1. INTRODUCTION

The Internet, which was a vague idea 50 years ago is now a reality and it is getting more popularity amongst people around the world. Most people often tend to know things which are often spoken about in terms of its effectiveness and greater usability. Internet stands out to serve both the purposes well enough to attract a major group of population towards it. Because of this emerging trend, more businesses are going online and this intern has lead to the creation of a lucrative E-Commerce Industry. Now, because of more businesses going online, there is a huge chunk of data being generated every second. Even recent businesses which are online are moving towards making data driven decisions. Hence the use of Data mining techniques makes a greater deal in analyzing data and giving out the better business decisions and recommendations.

A successful E-Commerce business is the one which is more customer friendly. Many recommendations have been developed and proposed in order to ease the transaction process of the customers. But still many problems related to customer satisfaction have not been resolved yet because of the complexity of the E-Commerce businesses. The main issues are Limited resource and cold start situations along with the data valid time constraint.

2. RELATED WORK

Sanjeevan Sivapalan, Alireza Sadeghian, Hossein Rahnama [1] 2014 proposed a filtering method for e-commerce recommendation system which carries out analysis on the relevant email documents of the customers. This filtering method carries out analysis of properties which are common to two or more documents. Some of the properties which were proposed to be analyzed included the messages, replies and its annotations. This analysis led to the discovery of the individuals interest and recommend products according to their interest.

Ms. Shakila Shaikh, Dr. Sheetal Rathi [2] 2017 carried out a survey on a variety of e-commerce sites according to some factors such as precision of recommendation products, recommended products, semantic recommendation, speed and numerous sort of products recommended.

Shaffy Goyal and Namisha Modi [3] proposed data minig techniques customer relationship management business module which is utilized to evaluate internal processes and functionalities of an organization.

Zhipeng Gao, Zhixing Li and Kun Niu [4] proposed a recommender system based on content based filltreing approach.

$$sim(T_i, T_j) = \frac{|S_i \cup S_j|}{|S_i \cap S_j|} \tag{2}$$

Based on (2) and Table 1, we can calculate the similarity between T_1 and T_2 , T_1 and T_3 , T_1 and T_4 , T_1 and T_5 as follows

$$sim(T_1, T_2) = \frac{|S_1 \cup S_2|}{|S_1 \cap S_2|} = \frac{3}{5} = 0.6 \tag{3}$$

$$sim(T_1, T_3) = \frac{|S_1 \cup S_3|}{|S_1 \cap S_3|} = \frac{3}{5} = 0.6 \tag{4}$$

$$sim(T_1, T_4) = \frac{|S_1 \cup S_4|}{|S_1 \cap S_4|} = \frac{2}{5} = 0.4 \tag{5}$$

$$sim(T_1, T_5) = \frac{|S_1 \cup S_5|}{|S_1 \cap S_5|} = \frac{1}{5} = 0.2 \tag{6}$$

If the value of threshold θ is set to be 0.5, then the neighbors of T_1 are T_2 and T_3 .

Ru Jia, Ru Li, Meiju Yu, Shanshan Wang [5] proposed methods to predict item buying patterns based on user behavior.

3. SYSTEM DESIGN

1. Hybrid Algorithm for Limited Resource Situation

This algorithm is used to handle limited resource situation with transaction history. Basically its a mixture of content and collaborative filtering algorithms. It uses user previous orders and rating to handle the limited resource situation. Top 'n' products based on the business decisions will be recommended for the users.

1.1. Finding out most bought out products by a user

Firstly we will find out the products which the customer is interested on buying the most. This is done based on analyzing the customer's transaction history by calculating the support function value ($S_1, S_2, S_3, S_4, \dots$) based on customer transaction history.

That is support function(S_1) for customer T_1 :

$$S_1 = \frac{\text{Number of transactions of a particular product}}{\text{Total number of transactions}}$$

Total number of transactions

The same calculation method is used for all the customer transactions listed in table 1.

A threshold value is set and whichever support function value crosses this threshold value, will be shortlisted.

Table 1: Purchasing History of the Customer

TABLE I. Purchasing History of the customer

Customers	Attractions				
	A_1	A_2	A_3	A_4	A_5
T_1	1	0	1	1	1
T_2	0	1	1	1	1
T_3	1	0	1	1	0
T_4	0	1	1	0	1
T_5	0	1	0	1	0

1.2. Generation of Neighbors

The next step is to identify nearest neighbors with similar area of interest. This is carried out by a correlation similarity method called cosine method.

Top 'n' products based on the business decisions will be recommended for the users.

1.3. Generation of Recommendations

Recommendations of attractions are computed by the purchasing times of neighbors. According to the calculation above, we know that the neighbors of customer T1 are T2 and T3, so we can list all the purchasing history of all the attractions so as to summarize the most popular ones. As listed in Table 2, we can find that the maximal purchasing times of neighbors are attraction A3 and attraction A4.

2. Random algorithm for Cold start situation

When new customers enter the system, there is usually insufficient information to produce recommendation for them, because there is no purchasing history of the new consumers. The usual solution of the cold start problem is similarity calculation between each user by profile information, such as user area of interest, gender, Age, etc.

The Random algorithm uses user profile and matches with the existing users to identify the neighbors and based on their transaction history, the cold start problem will be encountered. The listed out customer's transaction history is considered and the trading volume is calculated for short listing the top 'n' products for recommendation to the new user.

Volume = transaction containing that product/total number of transactions.

Administrator

Administrator is a business entity who maintains the entire application. An application can contain more than one admin. The admin has the following functionalities:-

1) Login Module

In this module, the admin/business entity gets login to the application by providing the credentials such as login id and password. The login id and password are set in the database manually.

2) Member Profile Module

Here the admin can view the registered users/members profiles. Member profile contains the information such as member id, first name, last name, address, age, gender, occupation, education, income level, marital status etc.

3) Product Module

In this module, admin of the application manages the product categories, subcategories and products. Management means adding the new categories/subcategories/products, edit, update and delete operations. Admin also can view the existing categories/subcategories/products.

4) Transaction Module

In this module, admin can view the customer/member transactions and can dispatch the transactions. Admin also can filter the member transactions based on age, gender, marital status, occupation, education and income level of the member/customer.

5) Rating Module

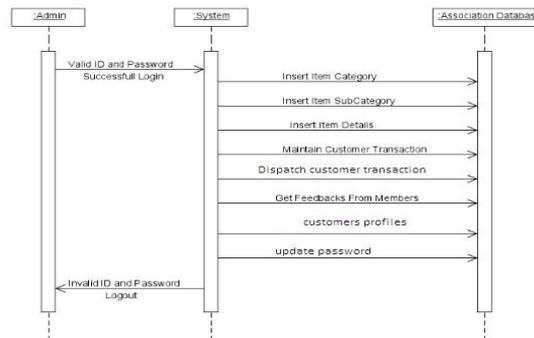
In this module, admin can view the ratings for the products given by the members. The ratings shows the demand for the products which helps other members for purchasing.

6) Feedback Module

In this module, admin can view the feedbacks posted by the customers.

7) Change Password Module

In this module, admin manages the requests of application members to change the password.



Sequence diagram of Administrator functionalities

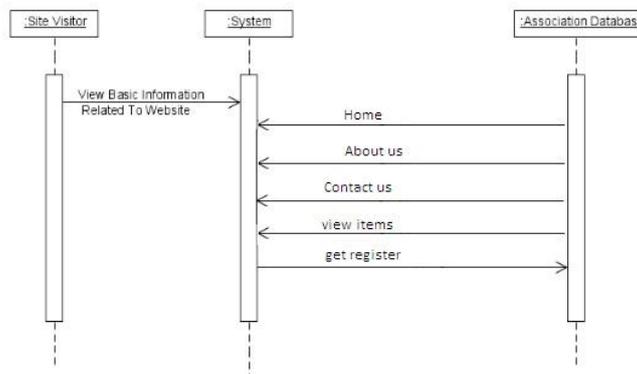
Visitor

Visitor is a customer who visits the application. Guest is the one who can access the basic information related to the application.

Visitor has the following functionalities:

- 1) View basic information related to application [home page, about us page, contact us page, services page, login page, registration page]
- 2) Browse Products

Here guest can browse the products based on category, subcategory.



Sequence diagram of Visitor functionalities

Member

Member is a registered user who receives the major services from the application. The major service given by the system limited resource situation, data valid time and clod start problem.

Member has the following functionalities:

- 1) Registration Module

In this module, customer gets register to the application by providing the profile details such as member id, first name, last name, address, age, gender, occupation, education, income level, martial status etc. During registration process customer gets an unique id and password for future user.

2) Login Module

In this module, customer gets login to the application by providing the credentials such as login id and password. The login id and password were generated during registration time.

3) Browse Products Module

Here member can browse the products based on category, subcategory.

4) Recommendation Module

This is the core module that we have in the application; here system recommends the products based on customer's interest. This module provides gives customers what they want.

This module handles 3 major issues of the existing system:

- Limited resource situation (association rules mining - hybrid collaborative Filtering Algorithm)
- Data valid time
- Cold start problems (Random algorithm/AOI)

5) Cart Module

In this module, member adds the selected products into the cart for order placement.

6) Review Order Module

Here member can review the order of items/products, can finalize before placing the order.

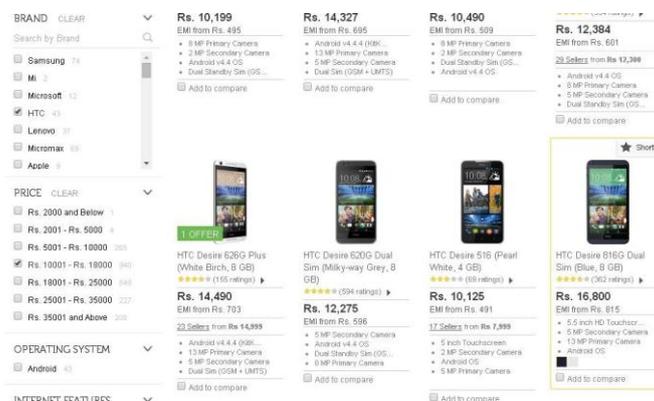
Place Order Module

In this module, member after reviewing the products in the cart, clicks on the place order button in order to place the order.

7) Transaction Module (previous transactions)

Here member can view the previous transaction details and can filter the transactions based on the status [pending, dispatched]

8) Rating Module



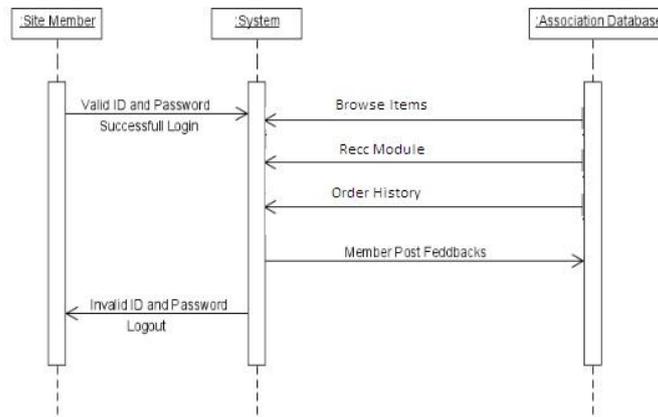
The member is given permission to rate the product.

9) Feedback Module

In this module, member can post the feedback for the services given by the website to the administrator. This will help the administrator to maintain the website in a better way.

10) Change Password Module

The member is allowed to change his/her password.



Sequence diagram of Member functionalities

Experimental results and analysis

The proposed system has been developed using the using technologies such as ASP.NET 4.0 for front end development, C#.NET used for coding algorithms, MS SQL server for back end development, Visual Studio 2010 an Integrated Development Environment. The main objective of the system is to give personalized recommendations to the customer. The input given to the system is recent transactions of the customer, User details and information related to Price and offers of products.



Personalized recommendations during cold start situations

CONCLUSION

This paper has shed light on some of the undermining problems that the E-commerce businesses are facing related to customer satisfaction which intern leads to an inefficient business output. Considering the above stated problems and the references used from various sources as a base for our study, we have developed a recommender system and other methodologies using data mining which is used for encountering the problems stated, in an efficient manner. The Hybrid algorithm is used for encountering limited resource situation where a product of limited quantity is only recommended to a person who is keen on buying it and not to any other. This is done by analyzing the information about that particular person’s previous transaction history and by implementing the collaborative filtering approach which even accounts the transaction’s history of other users based on similar interest matching. The Random algorithm is used in the cases of recommendation for new customers where the products are randomly recommended to the new user by considering the transaction history of other people whose area of interest is matching with that of the new user. A correct implementation of the above described techniques and methodologies will lead to improved efficiency and effectiveness in the performance of E-commerce businesses.

REFERENCES

- [1] Sanjeevan Sivapalan, Alireza Sadeghian, Hossein Rahnama, "Recommender Systems in E-Commerce", WAC 2014 1569901611.
- [2] Ms. Shakila Shaikh, Dr. Sheetal Rathi, "Recommendation system in E-commerce websites: A Graph Based Approach", 2017 IEEE 7th International Advance Computing Conference.
- [3] Shifty Goyal and Namisha Modi, "Data Mining Using Enhanced Decision Table Classifier for Online Shopping", 978-1-5090-3519-1/17/\$31.00, 2017 IEEE.
- [4] Zhipeng Gao, Zhixing Li and Kun Niu, "Solutions for Problems of Existing E-Commerce Recommendation System", 2016 International Conference on Consumer Electronics-Taiwan.
- [5] Ru Jia, Ru Li, Meiju Yu, Shanshan, "E-commerce Purchase Prediction Approach By User Behavior Data", 978-1-5090-5957-7/17/\$31.00 ©2017 IEEE.
- [6] Greg Linden, Brent Smith, and Jeremy York, "Amazon.com Recommendations Item-to-Item Collaborative Filtering", Published by the IEEE Computer Society, 1089-7801/03/\$17.00©2003 IEEE.
- [7] Shikhar Kesarwani, Astha Goel, Dr. Neetu Sardana, "MSD-Apriori: Discovering Borderline-rare items using Association Mining", Proceedings of 2017 Tenth International Conference on Contemporary Computing (IC3), 10-12 August 2017, Noida, India.