ISSN: 2278-4632
Vol-13, Issue-05, No.04, May : 2023

ENHANCING ONLINE SHOPPING EXPERIENCE THROUGH WEB MINING: ANALYSIS AND RECOMMENDATIONS

Dr Abhijeet Kaiwade, HOD, Sadhu Vaswani Institute of Management Studies for Girls, Pune
kaiwade@gmail.com,
Mr Niraj D Bagwe, Research Scholar, Sadhu Vaswani Institute of Management Studies for Girls and Assistant Professor, Abeda Inamdar Senior College, Pune
bagweniraj@gmail.com,
Ms. Bhagyashree Shinde, Student, Sadhu Vaswani Institute of Management Studies for Girls, Pune.
bhagya.shinde012@gmail.com M 9834070088
Ms. Shreya Vibhute, Student, Sadhu Vaswani Institute of Management Studies for Girls, Pune.
shreyavibhute.sv@gmail.com

ABSTRACT

Today web is the best medium of communication in modern business. Many companies are redefining their business strategies to improve the business output. Business over the internet provides the opportunity to customers and partners where their products and specific business can be found. Nowadays online business breaks the barrier of time and space as compared to the physical office. The analysis of the user's current interest based on the navigational behavior may help the organizations to guide the users in their browsing activity and obtain relevant information in a shorter span of time. Since a user has a specific goal when searching for information, personalized search may provide the results that accurately satisfy user's specific goal and intent for the search. Personalization of web search is to retrieve information according to user's interests which may be inferred from user's action, browsed documents or past query history etc. The Paper focuses on recommender systems based on the user's navigational patterns and provides suitable recommendations to cater to the current needs of the user. This proposal contains the concept of Web Usage Mining which is the discovery and analysis of user access patterns through mining of log files and associated data from a particular website. The data most accessed by user will be stored in log files.

Keywords: Recommender Systems, Web Usage Mining, Collaborative Filtering, Content Based Filtering, Electronic commerce, data mining, web mining.

Introduction

The Web is taking an important role in human's life and day by day it increases the information based on the expectations of the customers using it. Updated information is necessary to fulfill the needs of the users. Web mining is the application of Data Mining to automatically fetch and evaluate information from the web services and documents. Automation is everywhere and, in every field, to avoid human work in creation of anything. Web mining utilizes the automatic way of information extraction from the World Wide Web according to the preferences. With the continuous development of electronic commerce, it is not easy for customers to select merchants and find the most suitable products when they are confronted with the massive product information on the Internet.

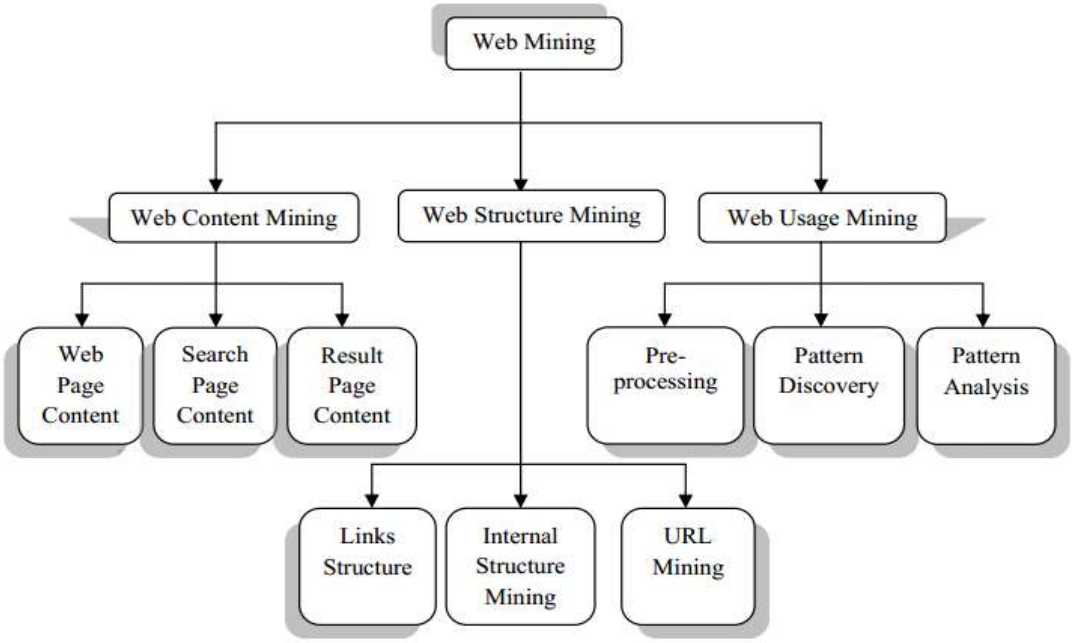
In the whole shopping process, customers still spend much time to visit a flooding of retail shops on Web sites and gather valuable information by themselves. This process is much time-consuming, even sometimes the contents of Web document that customers browse are nothing to do with those that they need indeed. So, this will inevitably influence customers' confidence and interests for shopping on the Internet.

In order to overcome the limitations of collaboration filtering, the recommender system based on web mining is proposed in the synopsis. It utilized a variety of data mining techniques such as web usage mining, association rule mining etc. Based on these techniques, the system can trace the

ISSN: 2278-4632
Vol-13, Issue-05, No.04, May : 2023

customer's shopping behavior and learn his/her up-to-date preferences adaptively. All these pose a challenge to researchers to discover web management methods and effective extraction of information from the web.

To understand web mining, we should know all about the data mining techniques available. Figure 1 provides taxonomy of web mining:



Web Content Mining (WCM)

Web Content Mining is the process of extracting use full information from the contents of web documents. Web documents may consist of text, images, audio, video or structured records like tables and lists. Mining can be applied on web documents as well the results pages produced from a search engine.

Web Usage Mining (WUM)

Web Usage Mining is the process of extracting use full information from the secondary data derived from the interactions of the user while surfing on the Web. It extracts data stored in server access another page may be considered as a vote. However, not only the number of votes a page receives is considered important, but the "importance" or the "relevance" of the ones that cast these votes as well.

Web Structure Mining

The goal of Web Structure Mining is to generate structural summary about the Web site and Web page. It tries to discover the link structure of the hyperlinks at the inter-document level.

Related Work

This system involves integration of web mining techniques with an e-commerce application. This integration facilitates e-store owners to improve the features and services and also it will help to get information about the customer's or consumer's behavior of visiting web products and services.

There are many areas where data mining can be very helpful when integrating with e-commerce. Some of them are: Data mining in customer profiling (Customer profiling means searching for the data which is collected from existing customers of an business organization for patterns that will allow that business organization to predict about who are the potential customers are and how

ISSN: 2278-4632
Vol-13, Issue-05, No.04, May : 2023

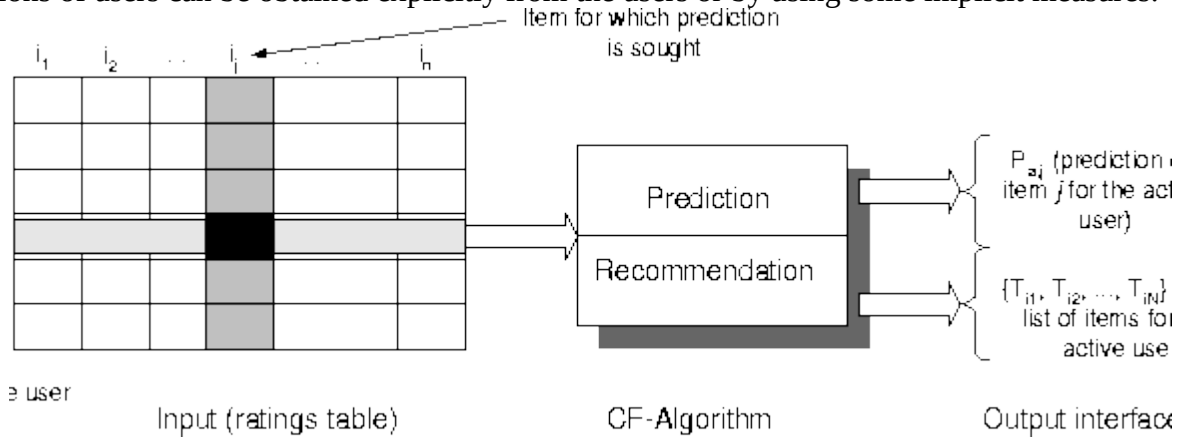
those customers are behaving), Data mining in recommendation systems, Data mining in web personalization, Data mining and multimedia e-commerce, Data mining and behavior of consumer in ecommerce. Online Recommendation exists and is being used by many shopping websites. Many algorithms, permutations and combinations, Associations will be used to create online recommendation. There are many different ways through which a system can be created.

Some algorithms used in general are:
1. Collaborative filtering:
- Item to Item.
- User to User.

2. Content Based Filtering:

A. Collaborative Filtering Based Recommender System:

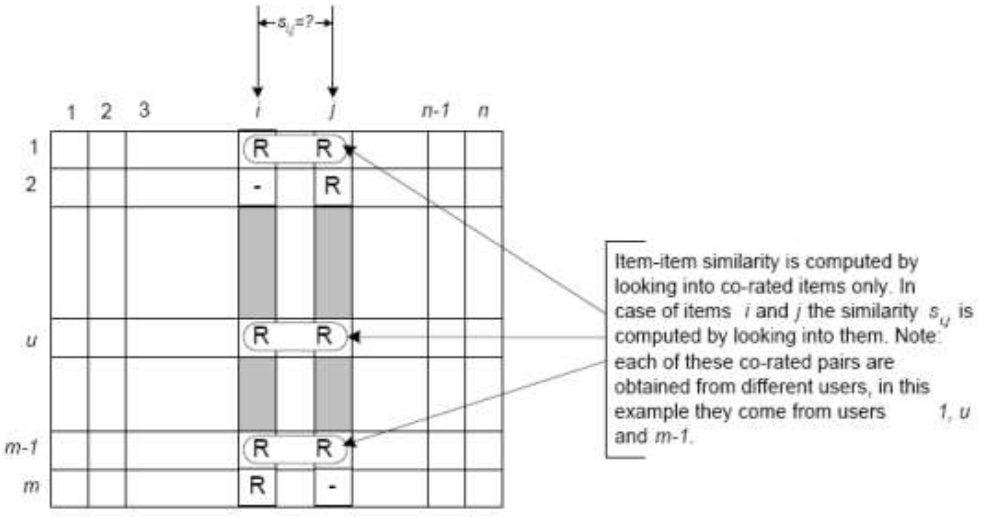
Recommender systems [1]-[2] apply data analysis techniques to the problem of helping users find the items they would like to purchase at E-Commerce sites by producing a predicted likeliness score or a list of top-N recommended items for a given user. Item recommendations can be made using different methods. Recommendations can be based on demographics of the users, overall top selling items, or past buying habit of users as a predictor of future items. Collaborative Filtering (CF) is the most successful recommendation technique to date. The basic idea of CF-based algorithms is to provide item recommendations or predictions based on the opinions of other like-minded users. The opinions of users can be obtained explicitly from the users or by using some implicit measures.



- **Item-Based Collaborative Filtering Algorithm:**

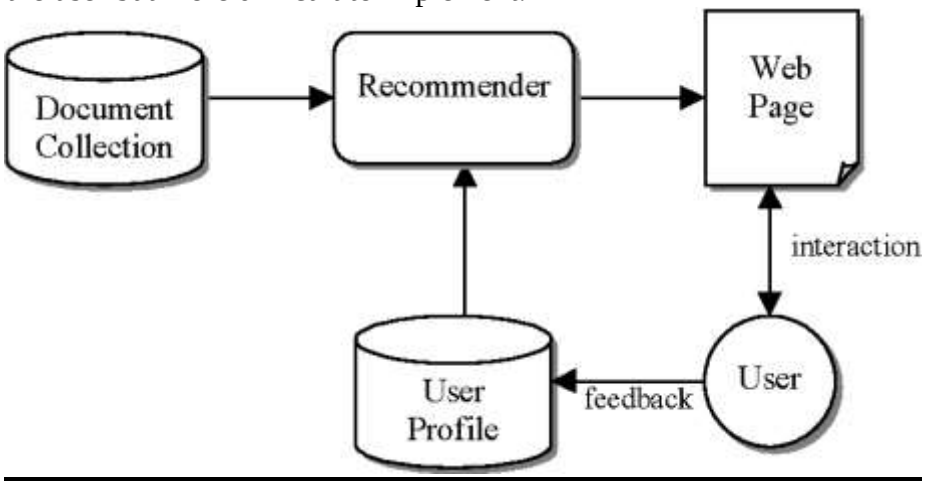
In this section we study a class of item-based recommendation algorithms for producing predictions to users. Unlike the user-based collaborative filtering algorithm, the item-based approach looks into the set of items the target user has rated and computes how similar they are to the target item i and then selects k most similar items {i1, i2 ,….ik}. At the same time their corresponding similarities {S1, S2….Si k} are also computed. Once the most similar items are found, the prediction is then computed by taking a weighted average of the target user's ratings on these similar items. We describe these two aspects, namely, the similarity computation and the prediction generation in details here

ISSN: 2278-4632
Vol-13, Issue-05, No.04, May : 2023



B. Content Based Filtering:

A content-based filtering system selects items based on the correlation between the content of the items and the user's preferences as opposed to a collaborative filtering system that chooses items based on the correlation between people with similar preferences. PRES is a content-based filtering system. It makes recommendations by comparing a user profile with the content of each document in the collection. The content of a document can be represented with a set of terms. Terms are extracted from documents by running through a number of parsing steps. First all HTML tags and stop words (words that occur very often and cannot be used as discriminators) are removed. The remaining words are reduced to their stem by removing prefixes and suffixes [Porter 1980]. For instance, the words "computer", "computers" and "computing" could all be reduced to "compute". The user profile is represented with the same terms and built up by analyzing the content of documents that the user found interesting. Which documents the user found interesting can be determined by using either explicit or implicit feedback. Explicit feedback requires the user to evaluate examined documents on a scale. In implicit feedback the user's interests are inferred by observing the user's actions, which is more convenient for the user but more difficult to implement.



Proposed System:

Online recommendations are already in use. But Different ways and algorithms are used for building it. We have tried to make a comparative study of the techniques which were used in the previous systems. We have combined algorithms and made it efficient and other measures of efficiency such as high speed, minimum memory usage.

ISSN: 2278-4632
Vol-13, Issue-05, No.04, May : 2023

Our Proposed System works in 5 phases:

Phase 1: Sort Phase

The database (D) is sorted, with customer-id as the major key and transaction-time as the minor key. This step implicitly converts the original transaction database into a database of customer sequences.

Phase 2: Large Item-set Phase

In this phase we find the set of all L-item sets. We are also simultaneously finding the set of all large 1-sequences, since this set is just {(l)| l L}.

Phase 3: Transformation Phase

In a transformed customer sequence, each transaction is replaced by the set of all L-item sets contained in that transaction. If a transaction does not contain any l-item set, it is not retained in the transformed sequence. If a customer sequence does not contain any l-item set, this sequence is dropped from the transformed database. However, it still contributes to the count of total number of customers.

Phase 4: Sequence Phase

Use the set of l-item sets to find the desired sequences. Algorithms for this phase are:

1. Apriori All

The Apriori All algorithm uses each pass to find large sequences from the previous pass to generate the candidate sequences and then measure their support by making a pass over the database. At the end of the pass, the support of the candidates is used to determine the large sequences.

In the first pass, the output of the L item set phase is used to initialize the set of large 1-sequences. The candidates are stored in hash-tree to quickly find all candidates contained in a customer sequence.

2. Apriori Some

1. Two of the proposed algorithms, Apriori some and Apriori-All, have comparable performance, albeit Apriori Some performs a little better when the minimum number of customers that must support a sequential pattern is low. 2. The major advantage of Apriori Some over Apriori-All is that it avoids counting many non-maximal sequences.

Phase 5: Maximal Phase (Output Phase)

Find the maximal sequences among the set of large sequences. In some algorithms, this phase is combined with the sequence phase to reduce the time wasted in counting non-maximal sequences.

Conclusion

Online Recommendation System for a shopping cart, an application that enables the user to search for and receive recommendations to find the product based on the previously searched products. The search result comprises a large number of products. Hence the recommendation system cuts down the burden on the user by recommending the products that best fits his preferences by placing them at the top. This recommendation is retrieved from the log files. Explicit feedback leads to an increase in the user's frustration level. The recommendation process uses implicit feedback. The feedback is obtained by making some observations on the user's purchase history. The online recommendation system for shopping cart makes successful recommendations by using the technique of assigning feature weights and user specific preference-based recommendations.

References

- Rakesh Agrawal , Ramakrishnan Srikant , "Mining Sequential Pattern" IBM Almaden Research Center,650 Harry Road, San Jose, CA 95120
- C.P. Sumathi, R.Padmaja Valli, T. Santhanam ,"Automatic Recommendation of web pages in Web Usage Mining", (IJCSE)International Journal of Computer Science and Engineering, Vol 02,No 09,2010,3046-3052

ISSN: 2278-4632
Vol-13, Issue-05, No.04, May : 2023

- Zhu Xinjuan, Huang junfang, Shi Meihang," An Intelligent on-line recommendation system in B2c apparel e-commerce ", Department of Computer Science, Xi'an Polytechnic, University, International Conference on E-Business and E-government,2010
- Guandong Xu, " Web Mining Techniques for Recommendation and Personalization " , Dissertation submitted to the School of Computer Science and Mathematics, Victoria University, Australia, March 2008.
- http://scholarworks.sjsu.edu/cgi/viewcontent.cgi?article=1111&context =etd_projects
- Dipa Dixit, Jayant Gadge," Automatic Recommendation for online users using Web Usage Mining", International Journal of Management Information Technology(IJMIT),Vol 02,No 03, August 2010 [9] http://en.wikipedia.org/wiki/Web_mining
- Jaideep Srivastava, Robert Cooley, Mukund Deshpande, "Web Usage Mining- Discovery and Application of Usage Pattern", Department of Computer Science and Engineering, University of Minnesota, MN55455
- DB Desai, Abhijeet Kaiwade, "The effects of mobile games on male adolescents using data mining techniques-A review" International Journal of Computer Applications Technology and Research, Vol 09, Issue 01, 2020, https://www.researchgate.net/profile/Deelip-Desai/publication/338221840_The_Effects_of_Mobile_Games_on_Male_Adolescents_using_Data_mining_techniques-_A_Review/links/5e089143299bf10bc381226a/The-Effects-of-Mobile-Games-on-Male-Adolescents-using-Data-mining-techniques-A-Review.pdf
- B. Bhamangol, A. Kaiwade, B. Pant, A. Rana, A. Kaiwade and A. Shaikh, "An Artificial Intelligence based Design and Implementation for classifying the Missing Data in IoT Applications," 2022 5th International Conference on Contemporary Computing and Informatics (IC3I), Uttar Pradesh, India, 2022, pp. 1376-1382, doi: 10.1109/IC3I56241.2022.10072634.
- Kadam, S. C., & Kaiwade, A. (2019). Data Mining: Process of Knowledge Invention. INTERNATIONAL JOURNAL OF ADVANCES IN COMPUTING AND MANAGEMENT, 38, 25.