"A Hybrid Recommendation System Based on Collaborative Filtering and Association Rule "

Khushboo Sahu, Prof. Mayank Bhatt

M-Tech Scholar, Assistant Professor, LNCT, Indore

Department of Computer Science & Engineering

sourabhsahu31@gmail.com*, mayankbhatt27@gmail.com**

Abstract

A recommendation engine is a feature (not a product) that filters items by predicting how a user might rate them. It solves the problem of connecting your existing users with the right items in your massive inventory (i.e. tens of thousands to millions) of products or content. Product recommendations are a must-have feature for all ecommerce websites, as they can drive sales, increase conversion rate and order value. Sending personalized product recommendations to your customers increases sales in just a few clicks. Add a Product Recommendations content block or merge tag to E-commerce Automation workflows to encourage subscribers to visit your store, re-engage inactive customers by promoting relevant items, suggest your best-selling products to new customers, and much more. Most of web portals have integrated this feature but web personalization is still a demanding issue. This project works target to integrate web personalization feature with product recommendation to enhance the sales and performance of E-commerce portal. A hybrid recommendation algorithm based on collaborative filtering and association rule will be used to recommend the products. Java technology will be used to develop these solutions.

Keywords: **Product Recommendation, Collaborative Filtering.**

1. INTRODUCTION

The growing internet world and hectic schedule of daily life create so much difficulty for internet Users to find desired information. This situation becomes worse when user try to search information and get irrelevant information. Inadequate knowledge of search tool and large amount of data gives poor performance to retrieve or extract desire information. Recommendation systems offer intellectual practice based on user preference. Recommendation systems offer separate and specialized set of information. In recent years, Web personalization has received much attention to help Internet users with the problem of information overload.

The complete study concludes that "An extensive application or tool that involves user preference or self collected knowledge for predicting user desire and explores the best possibility of relevancy among information is known Recommendation System." or it can be state that "Recommendation System is tool that provides pre specified knowledge based information". Recommendation System may useful in various fields such as news, marking, shopping, product search etc. News recommendation system offers collection of relevant news, articles, and suggestions based on user interest. They may offers news based on news popularity and visits. News ranking, priority, area, impact etc may be the core logic behind any news recommendation system according to

1. Offering news articles to on-line newspaper readers, based on a prediction of reader interests.
2. Offering customers of on-line retailer suggestions about what they might like to buy based on their past history of purchases and/or product searches.

A recommendation system can be classified according to their technique behind knowledge mapping and recommendation taught. They are explained as follows;

1. Knowledge based recommendation system
2. Content-based recommendation system
3. Collaborative-based recommendation system
4. Demographic recommender

1. Knowledge based recommendation system

Knowledge systems recommend suggestions or solution by generating manually or automatically a number of conclusions and decision rules. It emphasizes on explicit field knowledge about the requirements and user preference.

On the other hand, manually generated decision rules or drawn conclusions may be biased and not suitable for personalized systems. This system associated with different drawbacks such as bottleneck problem during knowledge processing and inherit problem during user profile creation and linking with existing information. A automatic knowledge based system is recommended where input of data may be subjective and can vary according to requirement.

2. Content based recommendation system

Traditional Content based recommendation system based on user preference and content exist at data source. It compares and extracts the information from web pages and data sources and match with user preference. It also uses popularity calculations and frequent uses to find most used and most demanding content. It uses this concept to evaluate and sort content according to demand and popularity. Generally, it observes the description associated with items or existing content and compare with user preference.

In many Web-based personalized applications such as e-commerce and e-learning sites, several techniques for document modeling, information filtering, and techniques for deriving information from the pages content are proposed. In such application, user profiles are generally described as vectors so that every entry of vectors represents a weight or an interest degree of each item in the Web pages.

3. Collaborative based recommendation system

Collaborative-based or so called social-based are an alternative approach to the previous approaches, aiming to improve the limitations of content-based approach. It exploits the other user's profiles in the same community and recommends new items not previously rated or seen by the user based on the assumption that similar users have similar interests in the same community. Therefore, recommendations take places based on the user similarity and recommend items from the interesting list of other people in the same community.

4. Demographic based recommendation system

A demographic recommender system provides recommendations based on a user's demographic profile which involves user's demographic data such as gender, age, date of birth, education, and other personal features.

LITERATURE REVIEW:
Collaborative filtering (CF):

Collaborative filtering (CF) is a technique used by recommender systems.[1] Collaborative filtering has two senses, a narrow one and a more general one.[2]

In the newer, narrower sense, collaborative filtering is a method of making automatic predictions (filtering) about the interests of a user by collecting preferences or taste information from many users (collaborating). The underlying assumption of the collaborative filtering approach is that if a person A has the same opinion as a person B on an issue, A is more likely to have B's opinion on a different issue x than to have the opinion on x of a person chosen randomly. For example, a collaborative filtering recommendation system for television tastes could make predictions about which television show a user should like given a partial list of that user's tastes (likes or dislikes).[3] Note that these predictions are specific to the user, but use information gleaned from many users. This differs from the simpler approach of giving an average (non-specific) score for each item of interest, for example based on its number of votes.

In the more general sense, collaborative filtering is the process of filtering for information or patterns using techniques involving collaboration among multiple agents, viewpoints, data sources, etc.[2] Applications of collaborative filtering typically involve very large data sets. 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 remainder of this discussion focuses on collaborative filtering for user data, although some of the methods and approaches may apply to the other major applications as well.

Association Rule Mining:

Association rule mining is known as finding frequent patterns, associations, correlations, or causal structures among sets of items in transaction databases. Consider an example of customer buying

habits by finding associations and correlations between the different items that customers place in their "shopping basket"

Here, Table 1 gives a set of transactions, find rules that will predict the occurrence of an item based on the occurrences of other items in the transaction.

TID	Items
1	Bread, Milk
2	Bread, Diaper, Beer, Eggs
3	Milk, Diaper, Beer, Coke
4	Bread, Milk, Diaper, Beer
5	Bread, Milk, Diaper, Coke

Example of Association Rules

$\{Diaper\} \rightarrow \{Beer\}$,
$\{Milk, Bread\} \rightarrow \{Eggs, Coke\}$,
$\{Beer, Bread\} \rightarrow \{Milk\}$,

Definition: Frequent Itemset

Itemset
A collection of one or more items
Example: {Milk, Bread, Diaper}
k-itemset
An itemset that contains k items
Support count (σ)
Frequency of occurrence of an itemset
E.g. $\sigma(\{Milk, Bread, Diaper\}) = 2$
Support
Fraction of transactions that contain an itemset
E.g. $s(\{Milk, Bread, Diaper\}) = 2/5$
Frequent Itemset
An itemset whose support is greater than or equal to a *minsup* threshold
Association Rule
An implication expression of the form $X \rightarrow Y$, where X and Y are itemsets
Example:
$\{Milk, Diaper\} \rightarrow \{Beer\}$
Rule Evaluation Metrics
Support (s) : Fraction of transactions that contain both X and Y

Confidence (c): Measures how often items in Y appear in transactions that contain X

POBLEM STATEMENT

The act of purchasing products or services over the Internet is becoming more common today. Online shopping has grown in popularity over the years, mainly because people find it convenient and easy to bargain shop from the comfort of their home or office. One of the most enticing factor about online shopping, particularly during a holiday season, is it alleviates the need to wait in long lines or search from store to store for a particular item.

Online shopping has revolutionized the business world by making everything anyone could want available by the simple click of a mouse button. Product Recommendation is like recommending similar product based on customer transactions or current selection nature. This approach uses complex algorithms to analyze large volumes of data and determine what products that potential customers might want to buy based on their stated preferences, online shopping choices, and the purchases of people with similar tastes or demographics. However, the technology has become considerably more sophisticated and is now an essential part of many online retailers' economic models.

E-Commerce based online shopping portals concentrate to allow huge range of product variety and simplify the way of purchasing. Consumers do not need to worry to visit various shops, compare and buy a product. Although, this sites provide wide range of shopping feature but also increase the consumer expectations. Now a day's consumer wants to purchase lots of product in minimum shopping time. This generates the demand of personalization of shopping portals and recommends products or customized product view based of consumer interest. Customer behavior can play very crucial role in this dimension. Subsequently, previous shopping transactions can also help to observe the shopping demand and consumer interest.

The complete work raises the demand to integrate collaborative filtering algorithm with association rule to create a web personalization environment.

PROPOSED SOLUTION

Web mining is the application of data mining techniques to extract knowledge from Web data, i.e. Web Content, Web Structure and Web Usage data. Collaborative Filtering is a way to extract

knowledge according to users' content preference and content usage. Association rules examine properties of the substance recommended. In a recommendation-system application there are two classes of entities, which we shall refer to as users and items. Users have preferences for certain items, and these preferences must be teased out of the data. Here, User preference will be section selection or keyword to be searched into news content and item will be news and articles. The proposed solution will integrate association rule and collaborative filtering to expand the perfection of product suggestions.

5. APPLICATION DOMAIN
1. E-commerce Portals
2. Shopping Malls
3. Sales Prediction
4. Prediction of stock ordering
5. Forecasting of Breakdown situation

6. CONCLUSION
The complete work expects an integrated solution based on product characteristics and customer behavior for product recommendation. Few points to describe the expectations are illustrated below;
1. To develop customer classification algorithm for customer behavior analysis.
2. To calculate product similarity and popularity index to estimate sales trend and popular product list.
3. Product recommendation based on product and customer nature.

REFERENCES

[1] Schafer, J. Ben, Joseph Konstan, and John Riedl. 1999. "Recommender Systems In E-Commerce." In 1St ACM Conference On Electronic Commerce, 158-166.

[2] Su, Xiaoyuan, and Taghi M. Khoshgoftaar. 2009. "A Survey Of Collaborative Filtering Techniques." Advances In Artificial Intelligence 2009: 1-19. doi:10.1155/2009/421425.

[3] Melville, Prem, and Vikas Sindhwani. 2010. "Recommender Systems." Encyclopedia Of Machine Learning.

[4] Xingyuan Li.2011 "Collaborative Filtering Recommendation Algorithm Based on Cluster", International Conference on Computer Science and network Technology(ICCSNT), IEEE, 4: 2682-2685.

[5] Francesco Ricc, Lior Rokach, Bracha Shapira.2011. Recommender Systems Handbook. NY:Springer, 1-35

[6] Sarwar, George Kaypi, Joseph Konstan and John Riedl.2000. "Application of Dimensionality Reduction in Recommender Systems -- A 6 Study." In ACM WebKDD Workshop. [7] Sarwar, George Kaypi, Joseph Konstan, John Riedl.2000. "Analysis of recommendation algorithms for E-commerce." In the Second ACMConference on Electronic Commerce, 158–167 42

[8] Sarwar, George Kaypi, Joseph Konstan, John Riedl.2001. "Item-based Collaborative Filtering Recommendation Algorithms." In the 10th International World Wide Web Conference, 285-295 [9] Gong, Songjie. 2010. "A Collaborative Filtering Recommendation Algorithm Based On User Clustering And Item Clustering." JSW 5 (7). doi:10.4304/jsw.5.7.745-752.

[10] Yan Shi, Xiao, HongWu Ye, and SongJie Gong. 2008. "A Personalized Recommender Integrating Item-Based And User-Based Collaborative Filtering." ISBIM '08 International Seminar On Business And Information Management 1 (2008): 264-267.

[11] Mangalindan, JP. 2012. "Amazon'S Recommendation Secret." Fortune. http://fortune.com/2012/07/30/amazons-recommendation-secret/.

[12] Koren, Yehuda.2009. "Collaborative Filtering with Temporal Dynamics." 15th ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining (KDD 09), ACM (2009): 447-455.

[13] Liu, Jiahui, Dolan, Peter, Pedersen, Elin Rønby.2010. " Personalized news recommendation based on click behavior", In: Rich, et al. (eds.) In the 14th Int. Conf. on Intelligent User Interfaces (IUI), ACM, (2010): 31–40. [14] Owen, Sean, Anil, Robin, Dunning, Ted, Friedman, Ellen. 2011 . Mahout in action. Shelter Island NY: Manning. () 43 [15] Walunj, Sachin, Sadafale, Kishor. 2013. "An online recommendation system for e-commerce based on apache mahout framework." Proceedings of the 2013 annual conference on Computers and people research, ACM (2013): 153–158.

[16] Verbert, Katrien, Drachsler, Hendrik,

Manouselis, Nikos, Wolpers, Martin, Vuorikari, Vuorikari, Riina, Duval, Erik.2011. "Dataset-driven Research for Improving Recommender Systems for Learning." 1st International Conference Learning Analytics & Knowledge, ACM (2011): 44-53.

[17] Wikipedia,. 2015. "Cosine Similarity." https://en.wikipedia.org/wiki/Cosine_similarity.

[18] Duryee, Tricia. 2014. "Amazon Adds 30 Million Customers In The Past Year - Geekwire." Geekwire. http://www.geekwire.com/2014/amazon-adds-30-millioncustomers-past-year/. [19] Grey, Paul. 2013. "How Many Products Does Amazon Sell? | Exportx." Exportx. https://export-x.com/2013/12/15/many-products-amazon-sell/.

[20] Resnick, Paul, Iacovou, Neophytos, Suchak, Mitesh, Bergstrom, Peter, Riedl, John.1994. " GroupLens: an open architecture for collaborative filtering of netnews." CSCW conference, ACM (1994).

[21] GroupLens,. 2013. "Movielens'. http://grouplens.org/datasets/movielens/.