

Recommender Systems: Increasing Profits and Efficiency in Business

Charlotte Castelino¹, Reena Shaw Muthalaly², Aloma Lopes², Sweedal Lopes²

1. Undergraduate Student, Electronics And Telecommunication Engineering Department, Dwarkadas J. Sanghvi College of Engineering, Mumbai, India

2. Undergraduate Student, Computer Engineering Department, St. Francis Institute of Technology, Mumbai, India

Abstract - Today, in almost every field, when interacting with a computer or any other automated device, we are provided recommendations or suggestions of possible preferences related to a product or a service. These applications make use of recommender systems to entice users toward a particular product and increase user association with the application. In this paper, we present a detailed study on recommender system and their new features.

Keywords - Recommender systems, collaborative filtering, social media, content-based filtering, Apache Mahout

I. INTRODUCTION

A. What is a recommender system?

The software tools and techniques that generate effective and meaningful suggestions of products, that a user might be interested in, is called a Recommender System or a Recommendation System. These suggestions then improve a customer's browsing experience by offering them personalized suggestions and hence, influence their decision-making processes.

Websites that use recommendation systems include:

i) Amazon.com: The e-commerce giant makes note of the item that the consumer has just selected and suggests items frequently bought in conjunction with the selected item and the additional products which other consumers purchased together with the said item.

ii) LinkedIn: It is a professional social networking site. Based on the information that one enters in the 'Summary', 'Experience', 'Education' etc. fields, it offers recommendations, among several, regarding available job vacancies, groups that align with one's skills and interests, people that one may know.

iii) Netflix: An on-demand streaming service suggesting movies based on what one watches and the ratings that a user provides.

This wealth of information when mined, can lead to benefits such as cost cutting, increase in revenues, etc. On an average, sales increase at the rate of 8-12%. [1]

II. BASIC WORKING OF A RECOMMENDER SYSTEM

Companies store a vast repository of data for each user and their corresponding likes/dislikes. Ratings, reviews and feedback given by a customer concerning a particular product also help to garner data. Post data-collection, various machine-learning algorithms like collaborative filtering; clustering and classification are applied to group together the current customer with known customers having similar likes and affinities, or to group together users and products.

Now, a recommendation engine extracts recommendations catering specifically to the user. For example, based on previous browsing history and shopping logs, an e-commerce website may deduce that a customer is more aligned towards purchasing mobile phones, and hence, provide suggestions of the latest models of mobile phones the next time he visits.

To construct scalable recommendation models, a machine learning library like Apache Mahout or architecture such as Hadoop may be used [2], depending on the requirements of the system.

The diagram below explains the step-wise working of a recommender system from one stage to another along with the feedback which is generated in the process. It involves an active interaction between the various components that have been shown.

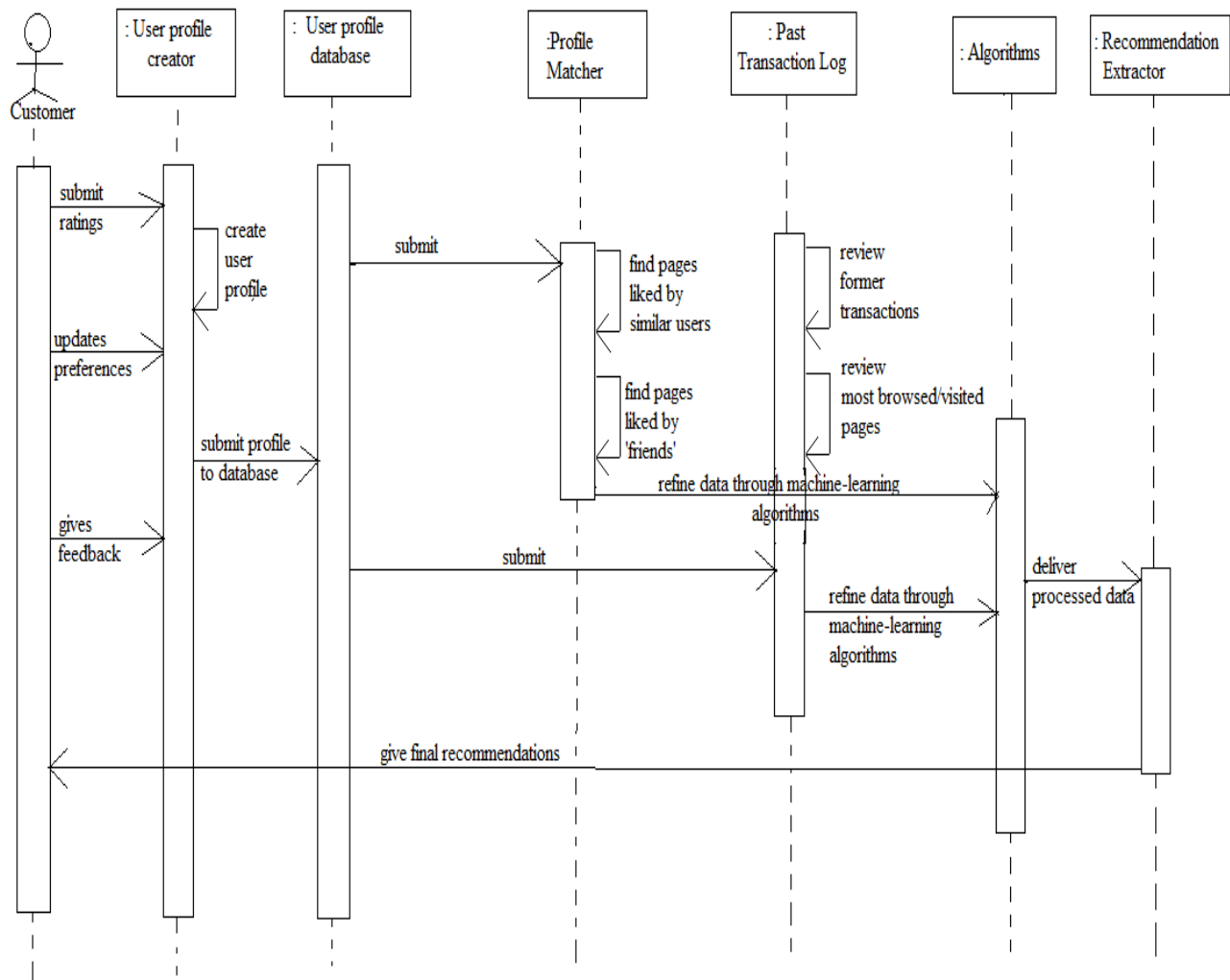


Fig. 1 Activities involved in the generation of recommendations

III. APPROACHES

A. Traditional approaches

i) *Collaborative filtering*: In Collaborative filtering, the past user behavior is collectively analyzed and on this basis, a model is created. It can be of the following types:

-User-based: Find similar users and recommend what they like

-Item-based: Find similar items to those that the current user has previously liked. [3]

than an item which is not. Popularity of the item is a major factor involved.

ii) *Content-based recommendations*: These approaches recommend items that are similar in features /content to items the user has liked in the past, or matched to the attributes of the user.[4]

B. Novel Approaches

There are several novel approaches. They include:

i) *Learning to Rank*: Assign a numerical or binary score to each item so that a relevant item receives a higher score

ii) *Context-aware recommendations*: Change in the context may produce a change in the user's preferences, and hence, in the suggestions generated. 'Context' includes

parameters like geographical location, season, time of day, variability in the type of options offered by a system.

iii) *Deep Learning*: Deep learning algorithms are based on distributed representations, a concept used in machine learning. The underlying assumption behind distributed representations is that observed data is generated by the interactions of many different factors on different levels. Deep learning adds the assumption that these factors are organized into multiple levels, corresponding to different levels of abstraction or composition. Varying numbers of layers and layer sizes can be used to provide different amounts of abstraction. [5]

iv) *Similarity*: Items may be classified as similar, if they possess similarity in their properties, similarity in the user-ratings of multiple items, and in the product features. In order to recommend reading material to readers, the former may be classified as similar on account of the articles that are read by people with similar reading habits, or the similarity of important words in the documents.

v) *Social Recommendations*: The concept of ‘trust’ is pivotal to this approach. Recommendations are generated based on the popularity of an item in the social proximity of the user. ‘Trust’ can be used to give more importance to the opinions of certain people.

C. Hybrid Approaches

Hybrid approaches are the combination of Content-based and Collaborative approaches.

IV. RECOMMENDER SYSTEMS USING APACHE MAHOUT

It is possible to develop customized recommender systems by making use of selective algorithms provided by Apache Mahout. Apache Mahout is a project which has been developed by the Apache Software Foundation. Since Mahout implements the three machine learning algorithms- collaborative filtering, clustering and classification, it is an excellent tool to be used for recommender systems. Mahout mainly focuses on collaborative filtering and most of its algorithms are implemented on the Apache Hadoop platform.

The collaborative filtering technique provides potentially interesting recommendations to users based on past preferences. The process of clustering is used to place similar items together into a related group. Various clustering techniques such as k-means algorithm, canopy, spectral clustering, etc. may be employed. Mahout also has several classification algorithms which can be used for developing an efficient recommender system. Classification learns from existing categorized documents, what documents of a specific

category look like and is able to assign unlabelled documents to the (hopefully) correct category. [8]

Several leading IT companies have made use of Apache Mahout to develop customized recommender systems. For example, in 2012, Cognizant developed an intelligent recommender system using Apache Mahout, which was designed to provide recommendations to scientists about relevant resources and research work. In this project, the demographic properties such as country, division, job title, department, etc. were used to establish similarity among users. [9]

V. RECOMMENDER SYSTEMS AND SOCIAL MEDIA

Today, social media plays a very important role in assisting recommender systems. The main reason for this is the fact that the social web transforms information consumers into active contributors, allowing them to share their status, comment or rate web content. [11]

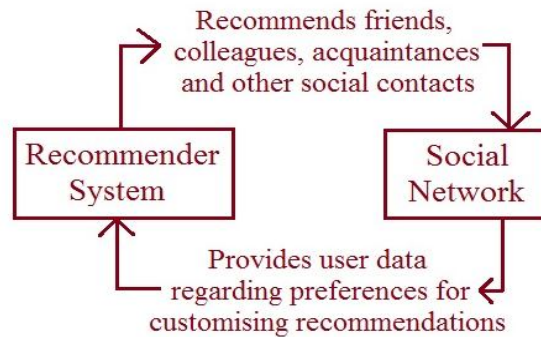


Fig. 2 Two-way interaction with social networks

A. Obtaining inputs from social media

The interaction between users regarding the features of a product or a service provides a direct input to the recommender system about the preferences of users. This preference may then be manipulated by the system to suggest other products or services from a given website which may meet the user’s requirements. Thus, even without a direct feedback from the user, preferences may be known for providing favorable suggestions.

B. Recommender systems in social networking

Recommender systems are widely used by social networking sites such as Facebook to suggest possible acquaintances or web pages which may be of interest to the user. This leads to an increased network activity, which in turn, leads to more refined suggestions for the user.

VI. FACTORS INFLUENCING RECOMMENDER SYSTEMS

A. Demographics

Many a times, the demographics of a region are important while deciding preferences. A region may be dominated by a particular community or culture and thus users may have a proclivity towards particular styles or products. The recommender system must be able to recognize and modify these.

B. Fidelity

Any recommender system must be built to improve user fidelity towards a product or website. As the association of the user with the product or site matures over time, the preference model of the user becomes more distinguished and recommendations may be more refined.

C. Persistence

It is suggested that recommendations must be persistent. It is possible that when a suggestion was first provided, the user might not have paid attention or may not have required the product or even did not have the required funds.

D. Serendipity

Serendipity in a recommender can be seen as the experience of receiving an unexpected and fortuitous item recommendation, therefore it is a way to diversify recommendations. [10] An ideal recommender system must blend accuracy, novelty, serendipity, and diversity.

E. Privacy

While providing suggestions to a user, the recommender system manipulates user data and tries to gather as much data as possible to improve suggestions. However, while doing so, user privacy must not be invaded extensively.

F. User satisfaction

The system must also be able to learn from previous experiences. If a user continuously overlooks certain preferences, then the system must modify its suggestions till user satisfaction is ensured.

VII. LIMITATIONS

A. Sparsity

Most users do not rate most items and hence, the user ratings matrix is typically very sparse. This is a problem for

collaborative filtering systems, since it decreases the probability of finding users with similar ratings. [4]

B. Changing data

A style and fashion website, leveraging a user's past data may not be a good idea as fashion trends are always changing. [6]

C. New beginners

The system may not give accurate or enough recommendations if the user is new or the product is new.

D. Fraud

Fraudulent vendors may indulge in 'profile-injection attacks', wherein they create dummy profiles and inflate the perceived desirability of their own products and lower the ratings of their rivals. The average attack assumes knowledge of the average rating for each item; and the attacker assigns values randomly distributed around this average, along with a high rating for the item being pushed. [4]

VIII. FUTURE SCOPE

With the advent of Big Data, recommender systems are looking at possible real time updation of user profiles. Synchronization of user profiles requires high network bandwidth as well as an increased processing capability. Hence, there is a need for efficient algorithms which would reduce the complexity involved and thus reduce the memory, bandwidth and processor requirements.

IX. CONCLUSION

A good recommender system must be able to provide suitable suggestions as per the user profile. These suggestions must not be too monotonous, but at the same time relevant to the user profile and aligned with one's tastes. Also, the system must store only as much data as is required to prevent information overload.

REFERENCES

- [1]<http://www.practicalecommerce.com/articles/1942-10-Questions-on-Product-Recommendations>
- [2]<http://inside-bigdata.com/2014/10/01/ask-data-scientist-recommender-systems/>

[3]<http://technocalifornia.blogspot.in/2014/08/introduction-to-recommender-systems-4.html>

[4] <http://vikas.sindhvani.org/recommender.pdf>

[5] Y. Bengio, A. Courville, and P. Vincent., "Representation Learning: A Review and New Perspectives," IEEE Trans. PAMI, special issue Learning Deep Architectures, 2013

[6]readwrite.com/2009/1/28/5_problems_of_recommender_systems

[7] Shyong K. Lam and John Riedl. Shilling recommender systems for fun and profit. In WWW '04: Proceedings of the 13th international conference on World Wide Web, pages 393–402, New York, NY, USA, 2004. ACM.

[8] <http://mahout.apache.org/>

[9]<http://www.cognizant.com/InsightsWhitepapers/How-to-Develop-Online-Recommendation-Systems-that-Deliver-Superior-Business-Performance>

[10] {R}icci, {F}., {R}okach, {L}., {S}hapira, {B}., {K}antor {B}. {P}. (2011). "Recommender systems handbook". *Recommender Systems Handbook* (Springer): 1–35.

[11] Recommender Systems using Social Network Analysis: Challenges and Future Trends Johann Stan¹, Fabrice Muhlenbach², Christine Largeron²