Literature Survey on Recommendation Systems

Dhvani Shah
CSE
PES University
Bangalore, India
dhvani.pranav.shah@gmail.com

Geethika Kommineni
CSE
PES University
Bangalore, India
kommineni.geethika@gmail.com

Manali Tanna
CSE
PES University
Bangalore, India
manalitanna29@gmail.com

Milinda KN
CSE
PES University
Bangalore, India
milimilindakn@gmail.com

***Abstract* - Recommender systems have grown to be quite popular in recent times. Providing useful suggestions of products to online users to increase their consumption on websites is the goal of many companies. People usually select or purchase a new product based on some friend's recommendations, comparison of similar products or feedbacks from other users. In order to do all these tasks automatically, a recommender system must be implemented. There is an extensive educational study on such systems. We have reviewed various research papers on recommender systems, and their evaluation methods, and how they perform in comparison with the others. They are categorized into different techniques of recommender systems and evaluated accordingly. We hope this paper gives researchers an insight into existing methods and what can be done to improve recommendation systems.**

Keywords - Recommendation systems; Literature review, Evaluation Methods

Problem Statement - The main purpose of the recommendation systems is to provide the users with the most accurate results of what the users may want to watch. Considering the current data security threats, our purpose is to provide the users with an efficient recommendation system without breaching the user's privacy and information.

Link to Google Sheet
https://docs.google.com/spreadsheets/d/1OVJuqUXmntfqBIqU2cTxWutCkqk_OXq3AACuFe5ZrAQ/edit?usp=sharing

Link to EDA
https://drive.google.com/file/d/1ApmE8CGnOtSeoRpCHeXikz-iRfTlPI6f/view?usp=sharing

I. Introduction

A recommender system is an algorithm that aims to suggest related items to users (movies to watch, texts to read, products to buy, etc., depending on the industry). They are very important in some industries because they generate huge revenues if they are efficient or clearly stand out from the competition. Recommender systems process large amounts of existing information by filtering the most important information based on the data provided by the user and other factors that take into account the user's preferences and interests. Examine the correspondence between users and articles, and infer the similarities between users and articles and recommend them.

This study is going to include extensive research into recommendation systems and their types. We will also be examining the performance of recommendation systems.

II. Type of Recommendation Systems

Over the years, recommender systems have been studied widely and are divided into different categories according to the approach being used. The categories are collaborative filtering (CF), content based and context-based.

1. **Collaborative Filtering**

   Collaborative filtering (CF) uses the numerical reviews given by the user to make recommendations. The historical data available helps to build the user profile and the data available about the item is used to make the item profile. It is considered the most basic and the easiest method to find recommendations.

2. **Content Based Recommender System**

   Recommendation systems help overcome sparsity problems faced in collaborative filtering based recommendation systems. Content-based systems focus on the features of the products and aim at creating a user profile. It is observed that reviews contain a product feature followed by his/her opinion about the product.

3. **Context Based Recommender System**

   Extending the user/item convention to the circumstances of the user to incorporate the contextual information is what is achieved in context-based recommender systems. This helps to abandon the cumbersome process of making the user fill a huge number of personal details.

III. Literature Survey

[1] This paper proposes user-based collaborative filtering based on kernel method and multi-objective optimization (MO-KUCF) which introduces kernel density estimation and multi-objective optimization. It can be increasing the diversity of the recommendation systems, improving concept drift in dynamic data, and the accuracy and diversity of the recommendation system.

The process used here is to find the optimal recommendation list set for users by finding the Pareto optimal solution set in the multi-objective function. The KUCF algorithm is used as an optimization algorithm of UCF, by describing the distribution of user interests in the item space, more accurately estimating the similarity of user interests, and improving the quality of recommendation results. the diversity of the recommendation system reduces the initial error caused by concept drift without reducing the system classification performance and enhances the system adaptability to concept drift.

[2] Analyze ratings as informative signals about the quality of movies. A structural Bayesian learning model links the revealed experience utilities of raters, who are prior consumers, to the product choice of the future consumers of the same good. Postulates that movies are chosen based on the consumer's prior beliefs and the precision of the signals provided by the ratings. Consumers use the rating signals more when more consumers have revealed their preferences in the ratings. Specifies and estimates a simulated maximum likelihood model using the Netflix data on rental choices and ratings. The very rich data set allows me to identify the effect of ratings of demand while controlling for the inherent popularity of each specific DVD using fixed effects. The results demonstrate that the ratings provide signals of quality to consumers.

[3] The research develops an integrated recommender system model with the ability to provide personalized recommendations. The K-nearest neighbors (KNN) algorithm uses similarity matrices for performing the recommendation system; however, multiple drawbacks associated with the conventional KNN algorithm have been identified. Thus, an algorithm considering weight metric is used to select only significant nearest neighbors (SNN). Using a secondary dataset on MovieLens and combining four types of prediction models, the study develops an integrated recommender system model to identify SNN and predict accurate personalized recommendations at lower computation costs. A timestamp used in the integrated model improves the performance of the personalized recommender system. The research contributes to behavioral analytics and recommender system literature by providing an integrated decision-making model for improved accuracy and aggregate diversity.

[4] The recommendation system is broken down into the Content Filtering Approach, Neighbourhood Method, and Latent Factor Model. The tools and technique used to decipher this is SVD currently. This paper also talked about running the dataset in a small environment and then with nearly 1M entries in the dataset. This paper focused on comparing four different collaborative filtering algorithms, in which the aim was to find out which one produced the best prediction rate. The four algorithms were KNN, SVD, ALS, and Slope One. The mathematical models used were weighted mean and arithmetic mean. Out of the 4, best was SVD and worst was ALS, however, by increasing the dataset WAM produced the worst results.

[5] CRS allows users and systems to communicate dynamically through natural language interactions. This provides an unprecedented opportunity to explicitly determine the exact preference of the user. Considerable effort has been put into developing CRS distributed across different environments and applications. CRS's existing models, technologies, and evaluation methods are not yet mature. This white paper provides a systematic overview of the techniques currently used in CRS. CRS development is mentioned in these directions. (1) Evaluation of Question Prediction (2) Evaluation of Search/Rec Performance. (3) Effect of Conversation Length. (4) Effect of Embedding Size.

The approach taken here is that the system can actively ask relevant questions. Understanding User Needs-Basic Search Goals And recommender system. Initially, the system proposed was a system ask–user response (SAUR) paradigm. Also, a multi-store With a network architecture and a personalized version for dialogue-oriented search and recommendations based on this paradigm was proposed. Both sequential modeling and attention mechanisms have been experimented with. It was assumed the user-system conversation to be about aspect-value pairs, while in the future, it is necessary for the system to perform more flexible exchanges and to handle unexpected user responses appropriately.

[6] Users can utilize a recommender system to deal with information overload and identify objects that are relevant to them. Existing CBF methods suffer from over-specialization as a result of the lack of appropriate data for determining item similarity. In terms of accuracy and resilience, the system outperformed conventional methods.

In contrast, CBF–MN allows various items to be recommended to a user based on network analysis, and this improves the system's performance. CBF–MN solves the issue of over-specialization because a number of attributes are used as criteria for characterizing items. Ultimately, it is highly desirable that a recommender system should not recommend excessively similar items, but rather diverse items by considering diverse criteria. Furthermore, CBF–MN adopts a network analysis that considers the relationships between all items and examines the structural and indirect relationships among them. CBF–MN increases the system's performance by allowing different products to be recommended to a user based on a network analysis here movieLens data is used.

More text features in future research to solve the problem of over-specialization in recommender systems.

[7] The paper considers that the user's rating has subjectivity and can be easily affected by the surroundings, while the attributes of the project are much more stable. The user-project rating matrix can be substituted with project attributes for calculating similarity and project attributes can reflect the interests of users in many ways and increase the recommended accuracy in a way. Therefore, this paper uses the time decay function to portray the users' interests and its changes in multidimensionality at the same time as introducing project attributes to represent the project and proposes an improved clustering-based user collaborative filtering algorithm.

The nearest neighbor selection accuracy to a certain extent determines the quality of the recommendation algorithm, namely the similarity measurement method for collaborative filtering algorithm. The rating of the current users on un-rated projects can be forecasted according to the rating information of the nearest neighbor of the current user and the recommendation Top-N is produced. The clustering model is introduced into collaborative filtering algorithm, the clustering of users or items, the similar users or items clustered into the same cluster, looking for the nearest neighbor

[8] A recommender system is a specific type of intelligent system, which exploits historical user ratings on items and/or auxiliary information to make recommendations on items to the users. It plays a critical role in a wide range of online shopping, e-commercial services, and social networking applications. Collaborative filtering (CF) is the most popular approach used for recommender systems, but it suffers from complete cold start (CCS) problem where no rating records is available and incomplete cold start(ICS) problem where only a small number of rating records are available for some new items or users in the system. Two recommendation models, which are based on a framework of tightly coupled CF approach and deep learning neural network. A specific deep neural network SADE is used to extract the content features of the items. The State of the art CF model, timeSVD++, which models and utilizes temporal dynamics of user preferences and item features, is modified to take the content features into a prediction of ratings for cold start items.

[9] Describes the various web recommender systems in use by some popular websites on the internet like Amazon.com, LinkedIn.com, and YouTube.com, etc. Further, it describes the various approaches used in the various recommender systems such as Content-based, Collaborative, and Hybrid recommender systems. At the end of this paper, we focus on some of the main challenges faced by the web recommender systems and analyze some techniques to overcome them.

[10] Proposes a novel collaborative-based recommender system that provides a user with the ability to control the process of

constructing a list of suggested items. This control is accomplished via explicit requirements regarding the rigorousness of identifying users who become a reference base for generating suggestions. Uses a new way of ranking items rated by multiple users. The approach is based on Pythagorean fuzzy sets and takes into account not only assigned rates but also their number. The proposed approach is used to generate lists of recommended movies from the Netflix competition database.

[11] The research has proposed a hybrid framework recommendation system to be applied on two-dimensional spaces (User × Item) with a large number of Users and a small number of Items. It uses the favorite and unfavorite items of a user. The framework is built on the integration of association rules mining and a content-based approach.

First, they evaluate favorite Items that a user has seen in the past, and based on those items, the system uses the association rules mining technique to recommend new items to a user. Then secondly they consider Non-Favorite Items that a user has seen before and apply an item-based approach to find similar items to those on the Non-Favorite Items category.

The results of the experiment show that the proposed framework can provide accurate recommendations to users.

[12] Discusses the different approaches that follow to deal with large streams of data in order to extract information for personalizing their service. Also describe some of the machine learning models used, as well as the architectures that allowed them to combine complex offline batch processes with real-time data streams

IV. Future Work

Using Content-based algorithms to ensure user privacy.

V. References

[1] Tie-min Ma, Xue Wang, Fu-cai Zhou & Shuang Wang: Research on diversity and accuracy of the recommendation system based on multi-objective optimization - 2020

[2] Ivan Maryanchyk: Value of Information from Ratings: Evidence from the

Netflix DVD Rental System - 2020

[3] Sujoy Bag, Dr. Abhijeet Ghadge, Manoj Kumar Tiwari: An integrated recommender system for improved accuracy and aggregate diversity - 2019"

[4] Ankur A. Ranjan, Amod Rai, Saiful Haque, Bhanu P. Lohani and Pradeep K. Kushwaha: An Approach for Netflix Recommendation System using Singular Value Decomposition - 2019"

[5] Yongfeng Zhang Xu Chen, Qingyao Ai, Liu Yang3, W. Bruce Croft: Towards Conversational Search and Recommendation: System Ask, User, Respond - 2018

"[6] Jieun Son, Seoung Bum Kim: Content-Based Filtering for Recommendation Systems Using Multiattribute Networks - 2017

[7] Liu Xiaojun: An improved clustering-based collaborative filtering recommendation algorithm - 2017

[8] Jian Wei, Jianhua He, Kai Chen, Yi Zhou, Zouyin Tang: Collaborative filtering and deep learning-based recommendation system for cold start items - 2016

[9] Dr. Sarika Jain, Anjali Grover, Praveen Singh Thakur, Sourabh Kumar Choudhary: Trends, Problems, And Solutions of Recommender System - 2015

[10] Marek Z. Reformat, Ronald R. Yager: Suggesting Recommendations Using Pythagorean Fuzzy Sets illustrated Using Netflix Movie Data - 2014

[11] Ahmed Al Salama: A Hybrid Recommendation System Based on Association Rules - 2013

[12] Xavier Amatriain: Big & Personal: data and models behind Netflixrecommendations - 2013