Book Recommendation System Using Matrix Factorization

Mercy Milcah Y[1], Moorthi K[2]
[1] Dept of Computer Science and Engineering
[2] Assistant Professor, Dept of Computer Science and Engineering
[1, 2] Jansons Institute of Technology, Tamil Nadu, India

Abstract- *Recommender systems have become evidently common in recent years, and are applied to a diverse range of applications. Recommendations typically speed up searches and make it easier for users to access content that they are interested in, and surprise them with offers they would have never searched for. The purpose of this project is to demonstrate the development of book recommendation systems through the usage of Matrix Factorization. In recent years, matrix factorization has become increasingly popular due to its accuracy and scalability. Recommender systems typically produce a list of recommendations either through collaborative filtering or through content-based filtering.Matrix Factorization is a collaborative filtering solution for recommendations. Using common training methods, the matrices and vectors can be trained to yield a product that is very similar to the original user-item matrix, thus creating an accurate book recommendation model.*

Keywords- Recommender systems (RS),collaborative filtering, matrix factorization, training methods.

I. INTRODUCTION

Recommender Systems are software tools and techniques of machine learning that provides suggestions for items to an individual user. Recommender systems enable an advanced access to the specific products and their information by providing suggestions that are personalized based on the examples of a similar user's likes and dislikes. Recommendation systems emerge into intelligent algorithms, which can generate results in the form of recommendations to the users. The popular suggestions are related to various decision-making processes, such as what items to buy, what music to listen to, or what online news to read. Whatever the system suggests the user is termed as, the "item". A system normally focuses on a specific type of item, say product or utility range and accordingly its design, its graphical user interface, and the core recommendation technique used to provide the recommendations are all trained to generate useful and effective suggestions for that unique type of item. Recommendation systems are initially designed for the utilization of individuals who lack sufficient personal experience to evaluate the overwhelming number of alternative items that a resource might offer.

We consider a case, specifically a book recommendation system (RS) that assists users to select a suitable book to read. A popular Web site, Amazon.com, employs a RS to personalize the online store for each customer [11]. Recommendations are usually personalized where, varying individual users and user groups receive diverse suggestions.

Additionally non-personalized recommendations are also provided which are much simpler to generate with typical examples that include the top ten selections of books, CDs etc, and are normally featured in magazines or newspapers. While being useful and effective in certain situations, these non-personalized recommendations are not usually addressed by RS research, whereas personalized recommendations are offered as ranked lists of items. In performing this ranking, the system tries to predict the most suitable products or services for the user based on each person's preferences and constraints. In order to complete such a computational task, the system collects preferences from the users, which can be explicitly expressed, such as ratings for products in our case, books or inferred by interpreting user actions. For instance, according to the user preference for items shown on the main page, a recommender system may consider the navigating to a particular product page consequently.

A. Recommendation Systems:

Recommendations typically speed up searches thereby making it easier for users to access content that they are interested in, and surprise them with offers they would have never searched for. Recommender systems in recent years have become extremely common and are applied in a variety of popular applications. The most famous ones are probably movies, music, news, books and products in general. Over the years, collaborative filtering had emerged as the most prominent approach for recommendations. There has been an explosion of methods that are introduced in the area of recommendations, in the recent years. Moreover, the

development of recommender systems has also increased the complexity of the modern systems when compared to the traditional or basic systems that utilize methods such as collaboration and content based filtering.
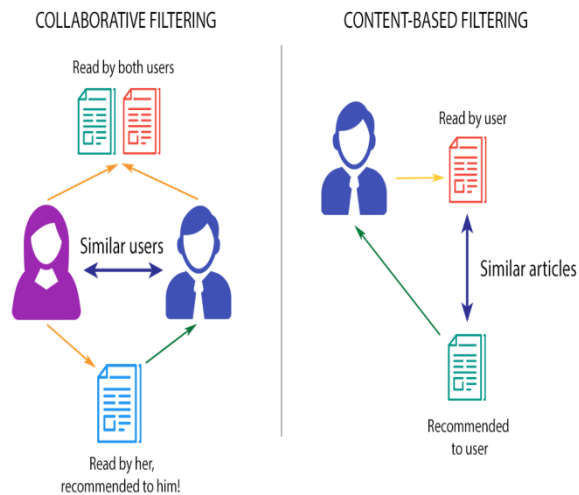

Fig.1. Basic methods used in RS

B. Types of RS:

There are mainly six types of recommender systems that are used by the user friendly resources or websites. They are Collaborative Recommender system, Content-based recommender system, Demographic based recommender system, Utility based recommender system, Knowledge based recommender system and Hybrid recommender systems.

B.1. Collaborative Recommender System:

This is the most desired type of RS, being widely implemented and one of the most mature technologies that is available in the market. Collaborative recommender systems generate new recommendations based on inter-user comparisons considering the ratings and recognize commonalities between various users, on the basis of their ratings and precisions. The greatest ability of collaborative techniques is that they are completely independent of any machine-readable representation of the objects that are being recommended and they also work acceptably well for complex objects in which variations in taste are responsible for the variation in preferences. Collaborative filtering performs based on the assumption that people who liked a product in the past will like in the future and that they will agree similar kind of objects as they liked in the past.

B.2. Content based Recommender System:

Content based RS mainly is classified as an outgrowth and continuation of information filtering research. In this system, the objects are mainly defined by the products' associated features. A content-based recommender system collects and learns the profile of a new user's interests based on the features present in objects the user has rated. This is basically a keyword specific recommender system where keywords are used in describing the items. Thus, in a content-based recommender system the algorithms used are such that it recommends users with items similar to the ones liked by the user in the past or is examining currently.

B.3. Demographic based Recommender System:

This system aims in categorizing the users based on attributes and provide recommendations based on demographic classes. Many industries have taken this kind of approach as it is not very complex and is easy to implement. In Demographic-based recommender system the algorithms first need a proper market research in the specified region that is accompanied with a short survey to gather data for categorization. Demographic techniques form "people-to-people" correlations like collaborative ones, but here, they use different data. The major benefit of a demographic approach is that it does not require any user ratings like that in collaborative and content based recommender systems.

B.4. Utility based Recommender System:

Utility based RS will make suggestions based on computation of the utility of each object for the user. Evidently, the central problem for this type of system is that how can a utility be created for individual users. In a utility based system, every industry will have a unique technology to create a user specific utility function and apply it to the objects under consideration. The main advantage of using a utility based recommender system is that it can work with non-product attributes, such as vendor reliability and product availability, while computing utility. By this, it possible to check for real time inventory of the object and then display it to the user.

B.5. Knowledge based Recommender System:

Knowledge based RS attempts to suggest objects based on inferences about a user's needs and preferences. Knowledge based recommendation works on functional knowledge, where they have knowledge about how a particular item meets a particular user need, and can therefore explain the relationship between a need and a possible recommendation.

B.6. Hybrid Recommender System:

Combining any of the two types of recommender systems, in a manner that suits a particular industry is known as Hybrid Recommender system. This is the most demanded Recommender system that many companies and resources look after, as it combines the strengths of more than two Recommender systems and also eliminates the weaknesses which exist when only one recommender system is used.

II. LITERATURE SURVEY

The paper[1] by Dheeraj Bokde, et al, discuss various factorization models such as Singular Value Decomposition(SVD),Principal Component Analysis(PCA) and Probabilistic Matrix Factorization (PMF). They attempt to present a comprehensive survey constituting Matrix Factorization model like SVD so as to address the challenges of Collaborative Filtering algorithms, which can be served as a road map for research and to practice in that specific area.

Aditya Krishna Menon and Charles Elkan in their research[2], propose to solve the link prediction problem in graphs using a supervised matrix factorization approach. The model learns about the latent features from the topological structure of a (possibly directed) graph, and is proved to make better predictions than popular unsupervised scores. They show how these latent features may be combined with optional explicit features for nodes or edges that yield better performance than using either type of feature exclusively. Finally, a novel approach is put forward to address the class imbalance problem which is common in link prediction by directly optimizing for a ranking loss. The model has been optimized with stochastic gradient descent and scales to large graphs.

Yu Wang et al, in their work [3], they propose a two-way multinomial logistic model for recommender systems for categorical ratings. Specifically, as we are treating the possible ratings as mutually exclusive events, whose probability is determined just by the latent factor of the users and the items through a two-way multinomial logistic function. This proposed method hasa compatibility with categorical ratings and the advantage of incorporating both the covariate information and the latent factors of the users and items uniformly.

Ruiqin Wang et al,[4] discuss about Collaborative Filtering (CF) algorithms that have been widely used to provide personalized recommendations in e-commerce websites and social network applications and Among which, Matrix Factorization (MF) is one of the most popular and efficient techniques. However, most MF-based recommender models relyonly on the past transaction information of users, hence there is inevitably a data sparsity problem. So they propose a novel recommender model based on matrix factorization and semantic similarity measure. Initially, a new semantic similarity measure is created based on the semantic information in the Linked Open Data (LOD) knowledge base, which is a hybrid measure that is based on feature and distance metrics.

Ritu Rani et al, in their paper [5], variety of algorithms such as k-mean clustering, collaborative filtering are used for the information suggestion. With the increase in demand of items amongst customer the growth is enhanced in information technology and ecommerce websites. The proposed algorithm mainly involves K-mean clustering and CF, as they first explain about this involved algorithm and then describe the specific purpose of the used algorithm in their study.

In their paper, Rupali Hande et al, [6] a movie recommender system is proposed called, Movie Mender. The objective of Movie Mender is to provide accurate movie recommendations to users. Usually the basic recommender systems consider any one of the following factors for generating recommendations; the preference or rating of a user (i.e. content based filtering) or the preference of similar users (i.e. collaborative filtering). To build a stable and accurate recommender system a hybrid system of content based filtering as well as collaborative filtering was being used.

Aditya G. Tate et al, in their paper [7] present a book recommender system that mines frequently hidden and useful patterns from the data in book library records and make recommendations based on the pattern that is generated using associated rule mining technique.

The paper [8] by Ms. Sushama Rajpurkar et al put forward a book recommendation system based on the combined features of content filtering, collaborative filtering and association rule mining respectively. This book recommendation has considered many parameters such as content of the book and quality of the book, by applying collaborative filtering. This recommender system also uses associative model to give stronger recommendations. This system does not have performance problem since it built the recommendations offline.

III. METHODOLOGY

The book recommender system uses collaborative filtering to achieve effective efficient level of recommendation.

A. Collaborative filtering:

The usual method analyzes the nature of each item. In our case, recommending books to a user by performing Natural Language Processing on the content of each book. Collaborative Filtering, on the other hand, does not require any information about the items or the users themselves. It recommends books based on users' past behavior. Among the various types of collaborative filtering techniques, this system uses model-based method. An efficient model-based CF is, matrix factorization.

A.1.Matrix Factorization:

Book recommendation system has been developed at a great rate due to the emerging web technology and library modernization. The purpose of this project is to demonstrate the development of book recommendation systems through the usage of Matrix Factorization. Recommender systems typically produce a list of recommendations either through collaborative filtering or through content-based filtering. Matrix Factorization is a collaborative filtering solution for recommendations. Using the data set of user-item pairings, one can create a matrix such that all rows represent different users, all columns represent different items, and each entry at location (i, j) in the matrix represents user i's rating for item j.
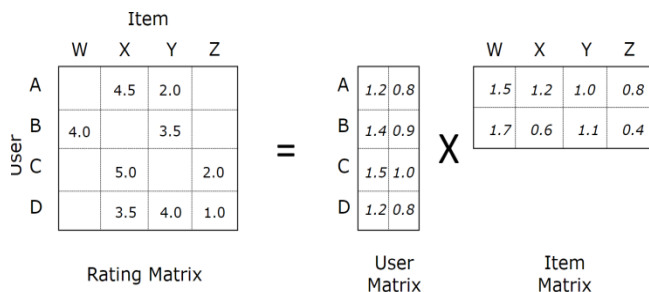

Fig 2.User-Item factorization matrices

When these two separate matrices are created, they each carry separate information about users, items, and the relationships between them. Precisely, one of the matrices will store information that characterizes users, while the other will store information about the items. Using common training methods, these two matrices, and the embedding vectors within them, can be trained to yield a product that is very similar to the original user-item matrix, thus creating an accurate recommendation model. In recent years, matrix factorization has become increasingly popular due to its accuracy and scalability.

IV. PROCESS MODEL

The book recommendation process consists of three major phases or steps. They are namely; Information collection, learning and prediction or final recommendation phase.
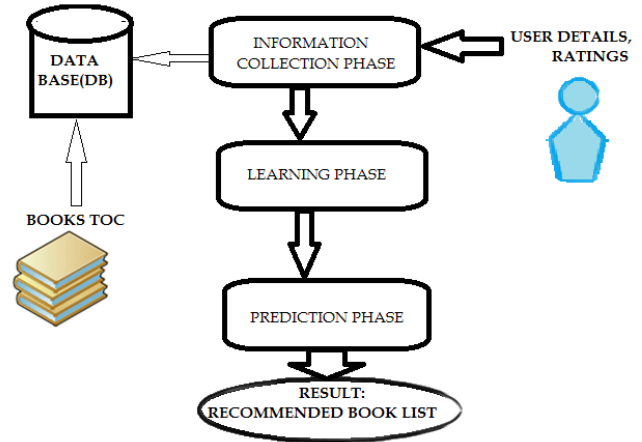

Fig 3. Process Model

A. Information collection phase:

The initial process of data collection is performed where, the user details are collected. Details include, the user profile data of each individual in order to provide personal recommendations. Every user has to create a simple profile in order to login to the system. The user is asked for the profile name, email id, age, and country in order to be used for providing a more personal set of recommendations with best matching books. The other important information to be collected from the user should include the reviews marked by the user for the viewed and read books. As it is really essential in knowing the reading taste of every individual user.

The collected review will be used further in listing a set of similar and favorable books for that reader.

With this set of information, a matrix made of user-item pairings is created, one can create a matrix such that all rows represent different users, all columns represent different items, and each entry at location (i, j) in the matrix represents user i' s rating for item j.

$$d\left(x,y\right)=\sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$$

Fig 4 .Matrix formula

B. Learning Phase:

After the required details are collected from the user as input to the system, the learning phase begins by using several efficient methods. The user's interests are known to the system and similar interests are matched to improve the efficiency. A matrix is formed where, for every user the specific rating for each viewed book is arranged in that matrix respectively.

Fig 5. User rating matrix for books

C. Prediction or Recommending Phase:

For further results, we use an effective prediction formula with which books that are not rated and reviewed by a user will be predicted ratings accordingly. By this prediction of rating for user's unrated books, the list of top most rated books are recommended to that user as a result.

We compare two books i and j , in order to measure the similarity between them. The book that has been rated already by a user 'x' is considered to be book 'j'. The books that are not yet been rated by user is considered to be 'i'. With 'N', taken as the set of similar books to the specific book 'i'.

$$r_{xi} = \frac{\sum_{j\in N(i;x)} s_{ij} \cdot r_{xj}}{\sum_{j\in N(i;x)} s_{ij}}$$

Where, S_{ij} – similarity of books i and j

r_{xj} – rating of user x on book j

$N(I;x)$ – set of books rated by x similar to i

Fig 6.Formula for prediction

With the two matrices s_{ij} and r_{xj}, the prediction formula is applied after which ratings for other similar books are predicted and there which top rated book list is recommended.

V. RESULT AND CONCLUSION:

This book recommendation uses one of the filtering techniques known as collaborative filtering(CF) and under which is a model base CF method called matrix factorization that is mainly used in this system to provide an efficient recommendation, as a result. Thus as a result, a list of books that are not yet visited or viewed by a specific user is predicted the respective rating values and are presented to the reader as final recommended set of books.

S.NO	UNRATED BOOK TITLE	PREDICTED BOOK RATING
1	Harry Potter	4.5
2	Goosebumps	4
3	The Time Keeper	3.73
4	Prison To Praise	3.7
5	To Kill a Mocking Bird	3

Fig 7. Recommended resultant book list

This technology used has thus proved to be effective in suggestions, and useful in providing personalized recommendations. This system can be used in websites and stores to be applied for user interaction and service providence

REFERENCES

[1] DheerajBokde, SheetalGirase, DebajyotiMukhopadhyay, "Matrix Factorization Model in Collaborative Filtering Algorithms: A Survey", Procedia Computer Science 49 (2015) 136 – 146.

[2] Aditya Krishna Menon and Charles Elkan "Link Prediction via Matrix Factorization", c Springer-Verlag Berlin Heidelberg 2011.

[3] Yu Wang, Xuan Bi and Annie Qu, "A Logistic Factorization Model for Recommender Systems with Multinomial Responses" , Accepted author version posted online: 11 Sep 2019, Published online: 25 Oct 2019.

[4] Ruiqin Wang, Yunliang Jiang, Hsing Kenneth Cheng and Jungang Lou, "A Novel Matrix Factorization Model for Recommendation with LOD-based Semantic Similarity Measure" , in Expert Systems with Applications 123 • June 2019

[5] International journal of engineering sciences & research technology, "book recommendation using k-mean

clustering and collaborative filtering" Ritu Rani*, Rahul Sahu.

[6] International journal of engineering sciences & research technology, "Moviemender- A Movie recommender system" RupaliHande, AjinkyaGutti, Kevin Shah, Jeet Gandhi, VrushalKamtikar.

[7] Aditya G. Tate, YogeshR.Guptaand Sharedchandra P. Kawane, "Intelligent Book Recommendation System Based on Collaborative Filtering and Association Rule Mining", International Journal of Emerging Technologies in Engineering Research (IJETER) Volume 5, Issue 4, April (2017).

[8] Jinny Cho, Ryan Gorey, Sofia Serrano , Shatian Wang and JordiKai Watanabe-Inouye , "Book Recommendation System", winter 2016 .

[9] ShahabSaquibSohail, JamshedSiddiqui, Rashid Ali, "Book recommendation system using opinion mining technique", International Conference on Advances in Computing,Communications and Informatics (ICACCI),2013

[10] Kesavaraj, G. ;Sukumaran, S. , " A study on classification techniques in data mining", International Conference inn Computing, Communications and Networking Technologies(ICCCNT),2013.

[11] PriyankThakkar, KrunalVarma, Vijay Ukani, SapanMankad and SudeepTanwar, "Combining User-Based and Item-Based Collaborative Filtering Using Machine Learning" Conference paper, First Online: 15 December 2018.