**Library Book Recommendation System**

**Abstract: -**

Due to wide application of management system, information data grows rapidly. On one hand, student have a large number of information resources. On the other hand, the time cost and difficulty of student to finding the proper information increases. To tackle the problems, book recommendation is one of the solutions for college libraries which possess huge volumes of books and reading-intensive users.

Recommender systems usually provide the student with a list of recommendations that they might prefer, or supply predictions on how much the student might prefer each item. Choosing what book to read next has always been a question for many. Even for students, deciding which textbook or reference book to read on a topic unknown to them is a big question.

In this report we try to present a model for a personalized recommendation system for books that uses hybrid recommendation approach which is combination of user-based collaborative filtering and KNN algorithm. The proposed recommendation system tries to learn the student’s preferences, the rating of books provided by students and recommends the books to the student based on their preferences.

**Introduction: -**

Recommendation systems are software programs that help a student to find products according to their needs and interests by using the student’s rating of each item and the student’s preferences. A recommender system works as a helper in finding relevant and related items based either on their explicitly mentioned preferences or objective behaviours. This way, it is a big source of reducing information overload in finding relevant items in several domains including books.

In order to recommend items including books to the student. The book recommendation system would help the student to borrow a book, by recommending books based on collaborative filtering and k-nearest neighbour classification algorithm.

**Literature Survey**

Recommender systems have become a vital research field since the emergence of the first paper on collaborative filtering in the mid- 1990s. In general, these systems are stated as the support systems which help users to find content, products, or services (such as books, movies, music, TV programs, and websites) by gathering and examining suggestions from other users, which means reviews from various establishments, and users. These systems are broadly classified into collaborative filtering (CF) and content-based filtering (CB). CF is an information filtering practice that is based on the user’s evaluation of items or previous purchases records. However, this method has been known to expose two major issues that are sparsity problem and scalability problem. CB examines a set of items rated by an individual user and then uses the content of these items, as well as the provided ratings, to deduce a profile that can be used to recommend additional items of interest. However, the syntactic nature of Content based filtering to detect the similarity between items that share the same attributes or features causes overspecialized recommendations that only comprise very similar items to those the user already knows.

**Tools used**

* Anaconda
* PyCharm
* VSCode
* MS-Excel
* Excel to CSV convertor

**Technology used**

* Python
* Scikit-learn

Scikit-learn provides a range of supervised and unsupervised learning algorithms via a consistent interface in Python. It is licensed under a permissive simplified BSD license and is distributed under many Linux distributions, encouraging academic and commercial use. Scikit-learn was initially developed by David Cournapeau as a Google summer of code project in 2007.

* KNN
* Cosine similarity
* Train-test-split
* SciPy

SciPy is an Open Source Python-based library, which is used in mathematics, scientific computing, Engineering, and technical computing. SciPy is the most used Scientific library only second to GNU Scientific Library for C/C++ or MATLAB’s.

* SVD
* Matplotlib
* Histogram
* Pandas
* Dataframes
* Read\_csv

**Issues or Challenges faced by Recommendation System:**

1.Cold-start problem: It's difficult to give recommendations to new users as his profile is almost empty and he hasn't rated any items yet so his taste is unknown to the system. This is called the cold start problem. Items can also have a cold start when they are new in the system and haven’t been rated before.

2. Scalability: With the growth of numbers of users and items, the system needs more resources for processing information and forming recommendations. Majority of resources is consumed with the purpose of determining users with similar tastes, and goods with similar descriptions.

3. Sparsity: In online shops that have a huge number of users and items there are almost always users that have rated just a few items. i.e. Students in general rate only a limited or no number of books. Using collaborative and other approaches recommender systems generally create neighborhoods of users using their profiles. If a user has evaluated just few items then it’s pretty difficult to determine his taste and he/she could be related to the wrong neighborhood. Sparsity is the problem of lack of information.

**Dataset Collection**

The dataset of books is collected from IIF (Institute for Informatik Freiburg). The dataset Contains 278,858 users (anonymized but with demographic information) providing 1,149,780 ratings (explicit / implicit) about 271,379 books.( <http://www2.informatik.uni-freiburg.de/~cziegler/BX/> )

The dataset comprises 3 tables. They are as follows:

BX-Users: Contains the users. Note that user IDs (`User-ID`) have been anonymized and map to integers. Demographic data is provided (`Location`, `Age`) if available.

BX-Books: Books are identified by their respective ISBN. Invalid ISBNs have already been removed from the dataset. Moreover, some content-based information is given (`Book-Title`, `Book-Author`, `Year-Of-Publication`, `Publisher`).

BX-Book-Ratings: Contains the book rating information. Ratings (`Book-Rating`) are either explicit, expressed on a scale from 1-10 (higher values denoting higher appreciation), or implicit, expressed by 0.

**Preliminaries:**

This section summarizes related work and briefly defines the fundamental concept needed to facilitate the presentation of the proposed algorithm.

A. Related work:

Different models have been developed in order to generate book recommendation. Many approaches rely on collaborative filtering (CF) methods based on the main idea that people have similar preferences and interests, so similarities of users or books are calculated. Basically, each user gives rating scores for a list of items and these scores are used to predict the rating active user.

B. User-based collaborative filtering:

Collaborative Filtering algorithm is based on the main idea that people have similar preferences and interests. One user’s behavior is compared with other user’s behavior to and his/her nearest neighbors, and according to his/her neighbor’s preferences or interest to predict his/her preferences or interest.

Suppose that U = {u1, u2, ..., um} is a list of m users and I = {i1, i2, ..., in} is a list of n items. Each user Ui gives rating scores for a list of items Iui. The prediction problem is to predict the rating active user Ua will give to an item Iua from the set of all items that Ua has not yet rated. The CF technique composes of 3 steps as follows: 1) users similarity calculation 2) top N nearest neighbors selection(knn) and 3) prediction.

1. Similarity
2. k-nearest neighbour classification
3. Recommendation

**Conclusion**

This system aims to provide personalized recommendation of books to the students. This system considers big data of books. The system makes use of collaborative filtering algorithm and knn classification algorithm to provides the user with recommendation list. The system tries to predict the ranking by considering the item's similarity as well as user's similarity so that a user can get recommendations of new books.

**REFERENCES**

[1] X. Su and T. M. Khoshgoftaar, “A survey of collaborative filtering techniques,” Advances in Artificial Intelligence archive, 2009

[2] L. Xin, E. Haihong, S. Junde, S. Meina, and T. Junjie, “Collaborative book recommendation based on readers’ borrowing records,” In Proceeding of the 2013 International Conference on Advanced Cloud ang Big Data, Dec. 2013, pp. 159-163.

[3] Y. Koren, R. Bell, and C. Volinsky, “Matrix Factoirzation techniques for recommendaer systems,” Computer, vol. 42, Aug. 2009, pp. 42-49

[4] H. Polat and W. Du, “SVD-based collaborative filtering with privacy,” ACM Symposium on Allied Computing, Mar. 2005, pp. 142-146

[5] 2018 5th International Conference on Business and Industrial Research (ICBIR), Bangkok, Thailand, A Collaborative Filtering Based Library Book Recommendation System.