**Book Recommendation System**

**IRSHAD**

**Data science trainees,**

**Alma Better, Bangalore**

**Abstract:**

Today the amount of information on the internet grows very rapidly and people need some instruments to find and access appropriate information. One of such tools is called a recommendation system. Recommendation systems help to navigate quickly and receive necessary information. Generally they are used in Internet shops to increase the profit. This paper proposes a quick and intuitive book recommendation system that helps readers to find appropriate books to read next. The overall architecture is presented with its detailed description. We used a collaborative filtering method based on Pearson correlation coefficient. Finally the experimental results based on the online survey are provided with some discussions.

**1.Problem Statement**

Online data is the collection of structured and unstructured data organized into relational or non-relational ways. It comes with a very large volume and velocity with a variety of types. Massive nature makes it complex to analyze and derive relevant conclusions. . Data is becoming one of the most important technology trends that have the potential for dramatically changing the way organizations. Subsequently, the growth of e-commerce portals creates a huge attention of users on online shopping and it changes the trend of business. Big investors are trying to get into this market and launching their online shopping apps and web applications.

A recommender system, or a recommendation system, is a subclass of information filtering system that seeks to predict the "rating" or "preference" a user would give to an item. They are primarily used in commercial applications.

**2. Introduction**

A recommendation engine is a class of machine learning which offers relevant suggestions to the customer.  Before the recommendation system, the major tendency to buy was to take a suggestion from friends. But Now Google knows what news you will read, Youtube knows what type of videos you will watch based on your search history, watch history, or purchase history

A recommendation system helps an organization to create loyal customers and build trust by their desired products and services for which they came on your site. The recommendation system today is so powerful that they can handle the new customer too who has visited the site for the first time. They recommend the products which are currently trending or highly rated and they can also recommend the products which bring maximum profit to the company.

## **3.Data Description**

We have 3 files in our dataset which are extracted from some books selling websites.

* Books – first are about books which contain all the information related to books like an author, title, publication year, etc.
* Users – The second file contains registered user’s information like user id, location.
* ratings –  Ratings contain information like which user has given how much rating to which book.

So based on all these three files we can build a powerful collaborative filtering model. let’s get started

### .

**4. DATA MINING**

Data mining is sorting through data to identify patterns and establish relationships. Data mining parameters include: ∙ Association - looking for patterns where one event is connected to another event ∙ Sequence or path analysis - looking for patterns where one event leads to another later event ∙ Classification - looking for new patterns (May result in a change in the way the data is organized but that's ok) ∙ Clustering - finding and visually documenting groups of facts not previously known ∙ Forecasting - discovering patterns in data that can lead to reasonable predictions about the future (This area of data m

**5. RELATED WORK**

Book Recommendation System (BRS) based on combined features of content based filtering (CBF), collaborative filtering (CF) and association rule mining to produce efficient and effective recommendation. For this we are proposing a hybrid algorithm in which we combine two or more algorithms, so it helps the recommendation system to recommend the book based on the buyer's interest.

Steps involved in the algorithm,

**Step 1:**

Scan the Books Dataset In this step scan the entire storage server and simultaneously perform the data cleaning, which include removing the irrelevant data and keeping the relevant data for mining.

**Step 2**:

Data Preprocessing According to our application, it includes the extraction of data that are needed for mining, which means extraction of only book categories and subcategories.

**Step 3:**

Filtering Transactions For filtering the transactions categorize the book based on category and subcategories.

**Step 4:**

Perform Content based Filtering

**Step 5:**

Perform Collaborative Filtering Here we consider the quality of the book content. In our example, recommending the books B3 and B4. This will perform based on the registered user’s opinions and rating.

**Step 6:** Final Recommendations

In the final recommendation, the order of the book is considered based on the book which holds the highest rating compared to remaining books and arranged in descending order. This book recommendation system is recommending books to the buyers that suits according to their interest and stores recommendations in the buyer’s web profile.

This system will store the details of the books which users have bought earlier and find the category of book from users buying history. It uses content based filtering and collaborative filtering and finds out the list of books based on content and ratings.

The system actually evaluates the quality of the recommended books dependent on the rating given by the existing users. Also, they use association rule mining algorithms to find interesting associations and relationships among a large data set of books and provide an efficient recommendation for the book.

This system may be helpful for lots of people as well as students who need the best books available from the database for both general and academic purposes. observe that due to wide application of management systems, information data grows rapidly. On one hand, people have a large number of information resources. On the other hand, the time cost and difficulty of people finding the proper information increases. To tackle the problems, book recommendation is one of the solutions for university libraries which possess huge volumes of books and reading-intensive users. This paper proposes a library book recommendation system based on user profile loaning and applying association rules to create models. The result shows a new association rule algorithm suitable to apply for recommended books in libraries. Recommender Systems have been around for more than a decade now. 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 paper, we try to present a model for a web-based personalized hybrid book recommender system which exploits varied aspects of giving recommendations apart from the regular collaborative and content-based filtering approaches. Temporal aspects for the recommendations are incorporated. Also for users of different age, gender and country, personalized recommendations can be made on these demographic parameters. Scraping information from the web and using the information obtained from this process

**6.1. Algorithms:**

.

**K nearest Neighbor**

The K-Nearest Neighbors (KNN) algorithm is a simple, easy-to-implement supervised machine learning algorithm that can be used to solve both classification and regression problems. The KNN algorithm assumes that similar things exist in close proximity. In other words, similar things are near to each other.

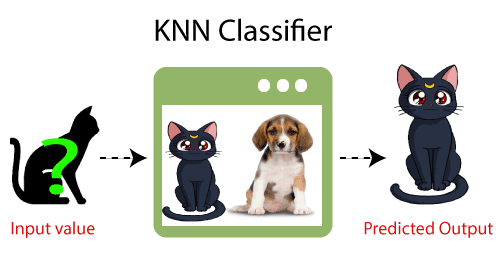
#### **Create Pivot Table**

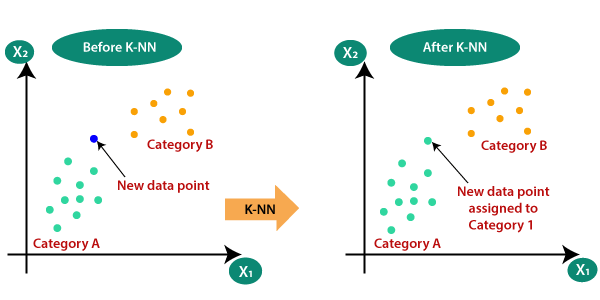
As we discussed above we will create a pivot table where columns will be user ids, the index will be book titles and the value is ratings. And the user id who has not rated any book will have value as NAN so impute it with zero.

## **How does K-NN work?**

The K-NN working can be explained on the basis of the below algorithm:

* **Step-1:** Select the number K of the neighbors
* **Step-2:** Calculate the Euclidean distance of **K number of neighbors**
* **Step-3:** Take the K nearest neighbors as per the calculated Euclidean distance.
* **Step-4:** Among these k neighbors, count the number of the data points in each category.
* **Step-5:** Assign the new data points to that category for which the number of the neighbor is maximum.
* **Step-6:** Our model is ready.

****

****

**SOLUTION STATEMENT**

Recommendation system is a method of filtering that uses rating, similarity score or preference score to predict the frequency between item and elements. Recommendation systems have become increasingly popular in recent years due to wide area of applications and use into movies, books, articles etc. prediction and suggestions. Here, a customized recommendation algorithm has been proposed using the DBSCAN algorithm to recommend a Book based on user nature. Following Points are listed to describe the approach of the proposed solution.

Initially, Dataset will be used as the input source for user and book information.

2. A customized classification method will be used to extract the nature and important credentials from the dataset.

3. A parsing module has been proposed to clean the data and remove unnecessary information and attributes.

4. Afterwards, a Stop word removal and lemmatization approach is proposed to extract the nature of the user and book from user review and other information. This activity is required to convert review text into meaningful word format.

5. A DBSCAN clustering algorithm is proposed to create a set of relevant users and optimize the recommendation process. Before doing this clustering, a binary matrix generation is expected to map matching factors between book and users.

**Conclusion**

This book provides an introduction to recommender systems. In the context of ever-increasing amounts of available information and data, it is difficult to know what information to look for and where to look for it. Computer-based techniques have been developed to facilitate the search and retrieval process; one of these techniques is recommendation, which guides users in their exploration of available information by seeking and highlighting the most relevant information.

Recommender systems have their origins in a variety of areas of research, including information retrieval, information filtering, text classification, etc. They use techniques such as machine learning and data mining, alongside a range of concepts including algorithms, collaborative and hybrid approaches, and evaluation methods.