IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 8, August 2021

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.542

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 8, August 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0908023 |

Hybrid Book Recommendation Engine: An Novel Approach for Better Efficiency of Suggestions

Shreya Bille, Sarvesh Kuthe

Software Developer-Accenture, Mumbai, India

Software Developer-Larsen & Toubro Infotech, Pune, India

ABSTRACT: Recommender Systems are nowadays utilized by the ecommerce websites and it is capable of understanding the user choices, preferences given by the users, also mutual likings of the users. Since the users are having the complex choices, the recommendation system provides suggestions of the products using the traditional algorithms for recommendation system such as Popularity based recommendation system Collaborative recommendation system, Content-based recommendation system, and Demographic-based recommendation system, Utility-based recommendation system, Knowledge-based recommendation system. These traditional algorithms have some flaws and doesn't give the perfect recommendations to the users. In this work we are proposing a hybrid book recommendation engine which will combine the Popularity Model, Collaborative Filtering and Content based algorithm which will be utilized for giving the recommendation of the books to the users also intends to solve the cold start problem. Since the pure collaborative and content based recommendations are not sufficient hence the hybrid approach is adopted for increasing the efficiency of the book suggestions to the user.

KEYWORDS: Recommender System, Hybrid Book Recommendation Engine, Collaborative filtering, content based recommendations, popularity model.

I. INTRODUCTION

Recommendation systems are the best for the users which spends time on the ecommerce websites, the recommendations system suggest the products to the users and it also keep tracks of their likings of the product, these can be utilized later for giving the similar kind of product recommendations. There are many types of the book readers on the internet, some likes the same genre books, some reads all types of books, hence there is need for the recommendation system which can be able to handle the several types of users, there product choice and relation between the product and the books. One of the approach is content based or cognitive filtering.Inthis technique it requires some detail about the items such as genre and it considers the item's content and the profile of the user and makes the comparison of user profile and content of item. Certain terms or descriptors represents the content of the item and profile of user can be represented with the similar terms and made by doing analysis of content of items seen by the user. This gives recommendation very much relevant to the user also it avoids the cold start problem for the user. There is no need of users to do interaction with products before the recommendation starts. Another approach is popularity based recommendation system. In this approach the user will get recommendation of the trending product, if there is one product which is purchased by every new user then it will be recommended to the user who is just new to the website. Collaborative is also previously used for recommendation systems which is also known as social filtering do the prediction of the products which are preferred by the users by analyzing the past relationship of the group of users who has the similar choices and preference of the products.

In Content based filtering the algorithm is not able to predict various products which user didn't liked ,content based filtering eliminates the issue of cold start, it is a problem which comes in personalized suggestions, also there is need of history give product suggestions , it is very complex if the history is very less for giving recommendation,

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 8, August 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0908023 |

also it leads to Data sparsity problems and it suggests items with ease for those users who has uncommon taste of products.

In Collaborative filtering algorithm new users faces cold start problem with sparse ratings issue. It becomes very complex to suggest products to users which uniquely selects items. This issue can be eliminated by decreasing the dimensions using singular value decomposition (SVD) by getting the similar users and products in item-user cluster.

1.2 Data Description

Dataset utilized for our work is goodbooks-10k. It consists 6 million ratings for 10,000popular books. Datasetdoesn't have tags of genre for every book. As content based filtering requires the genre of book, which is very important and hence I have done manual tagging ofthe first thousand books.

Some of the genres of the book dataset are:'Comedy', 'Young-Adult Fiction', 'Kids', 'Fantasy','History', 'Thriller','Poetry','Domestic Fiction', 'Biography', 'Fantasy'.

The Columns which are considered for the work are as shown in table 1.2

Column Name	Content
book_id	Unique Id for books.
title	Book Title
authors	Book author names
ratings	Book Ratings
genres	Genre of the book
tags	Book tags

Table: 1.2 Considered Attributes of dataset

II. LITERATURE SURVEY

The aim of most of the recommendation system is to predict user's likings and interest and suggest books accordingly. The below table 2.1 provides the researches about the book recommendation systems.

Reference	Technology	Analysis Of Research Gap
[1]	User based collaborative filtering, association rule mining.	This RecSys has Cold Start Problem as well as recommendation is more focused towards user budget and familiar publishers.
[2]	Combination of Content based filtering, collaborative filtering and association rule mining.	This RecSys considers more parameters but the recommendations are not accurate.
[3]	Recommendations using opinion mining	Obtained top 10 specific topic Books by sorting them according to maximum scores which are normalized scores in

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 8, August 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0908023 |

		descending order, thus the proposed model is time consuming with extra workload.
[4]	Content based and collaborative filtering	For Recommendation in this work combination used for improving predictions but the difference between performance of pure content-collaborative based algorithm and proposed system is 4% only.
[5]	Collaborative based filtering, content based filtering , key word filtering	Recommends the books based on the buyer'sinterest, hence the web based model's accuracy is less.

III. PROPOSED SYSTEM

In this proposed work, hybrid book recommender engine is implemented on the goodbooks-10k dataset, we have considered the book-id, authors, titles, genres, user-id, rating attributes of the dataset. Since the data which is available cannot be used directly, hence performed some preprocessing operations on the dataset and filtering out only the useful information from the dataset attributes. The proposed model requires the average ratings of the books, so average ratings for the book-id's is calculated also the rating counts are obtained for the further implementation of the hybrid model. Further, we have decided to utilize 80% data for training purpose and 20% data for testing purpose.

The proposed model is visualized in the Figure 3.1, as the data is obtained and preprocessed then it is passed to the traditional algorithms of recommendation system.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 8, August 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0908023 |

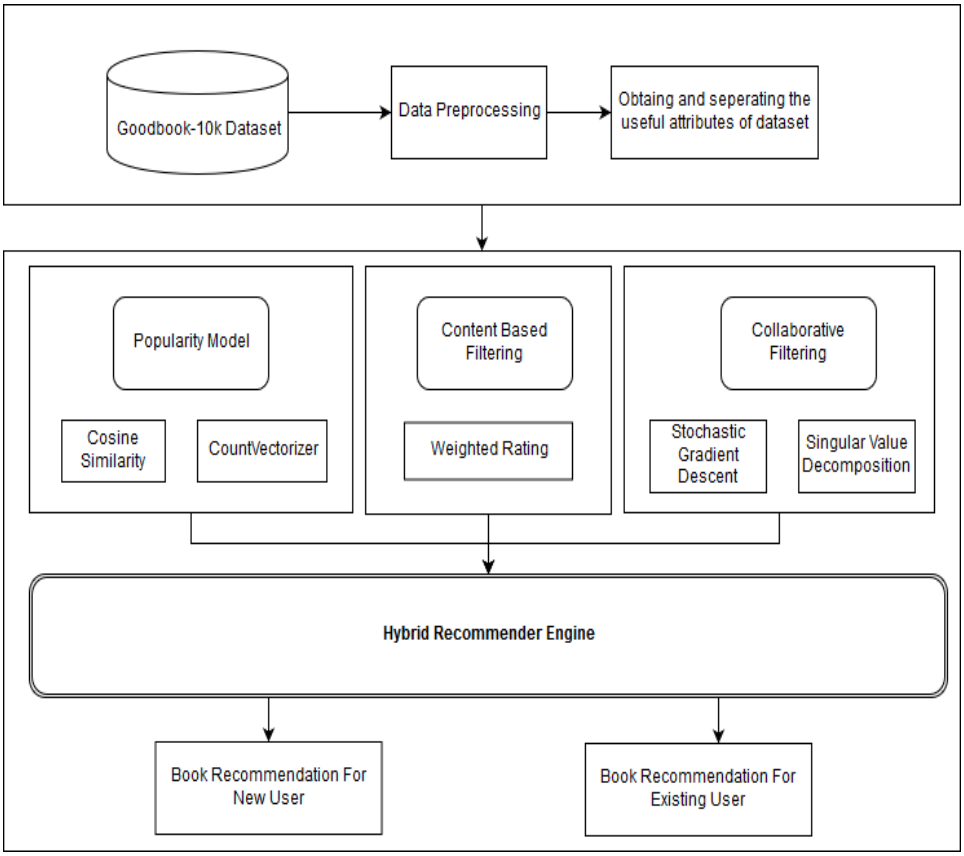

Figure 3.1: Hybrid Book Recommendation System.

Since recommendation is kept user dependent, Popular Model utilized to provide recommendation of those books which are popular among new users, hence we have done sorting as per the average of user ratings for any kind of particular book. This model suggest the top 25 books of multiple genres to new user and new user has to rate some of the books initially to get new recommendations hence the cold start problem is solved. Cold start problem occurs when recommendation engine gain a new user for the first time, due to no user history about user, the RecSys doesn't recognize the personal choices of the user. Also this helps the user to have choice from variety of genres in the beginning. Now for every book we have to calculate the weighted rating from rating matrix, rating count and ratings, hence to calculate the weighted rating I have considered the IMDB'S weighted rating formula mathematically represented as below:

$$[WR = (v/(v+m)*R)+(M/(v+m)*c)]$$

Where,

V is denoted as number of votes for the book.

M is denoted as minimum voted required to be the book in top charts.

R is denoted as average rating of book.

C is denoted as mean vote of book.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 8, August 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0908023 |

Since, in thisproposed work new user has to rate some books initially, Content based filtering works with the data which is given by the user and by considering that data it do processing and give recommendations also along with the Content based filtering we are utilizingCosine similarity to calculate the quantity (numeric) which shows the similarity between the two books. Countvectorizer is used in this algorithm for converting textual data into vectors and these vectors can be used for calculations of matrix. The predictions can be mathematically represented as follows:

$$P_{u,i} = \frac{\sum_{\text{all similar items, N}}(s_{i,N} * R_{u,N})}{\sum_{\text{all similar items, N}}(|s_{i,N}|)}$$

Since the content based filtering algorithm is able to giving suggestions of books which are certainly close to particular book.so content based filtering cannot considers the choice of user which are of different genres. So collaborative filtering uses the similar users like the user, according to the choices and likings of them we can predict how much I like a certain book and implementing this for calculating the user similarity according to the books and ratings. Singular value decomposition is utilized for making calculation of subsequent matrix in a simpler way and reduces the Root mean square error. Also we have used the stochastic Gradient descent instead of Gradient descent only because it is comparatively slow for identifying the values of co-efficient of a particular function and reduce the RMSE.Since some flaws are there in these traditional algorithms a hybrid recommendation engine is proposed which gives the better recommendation than the pure content, collaborative and popular based filtering. The Hybrid model can be mathematically represented as:

$$S_{hybrid} = (1\text{-}2\alpha) * S_{popularity} + \alpha * S_{collaborative} + \alpha * S_{content}$$

IV. RECOMMENDATIONS AND RESULTS

The proposed model gives the recommendation to the new user as well as existing user, the user has given option to choose whether he is existing user or new user as shown in figure 4.1.

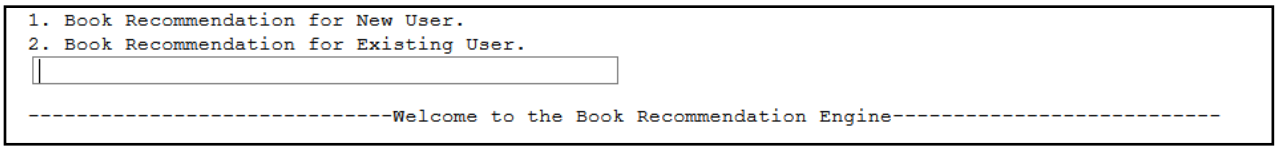

Figure 4.1: Type of Recommendation

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 8, August 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0908023 |

The snapshots of the recommendation for new user is as follows:

```
1. Book Recommendation for New User.
2. Book Recommendation for Existing User.
1

 Rate from books

ID   Author                          Title
Genre

2.   J.K. Rowling, Mary              Harry Potter and the Sorcerer's Stone (Harry Potter, #1)
Fantasy;Young-Age
127. Malcolm Gladwell                The Tipping Point: How Little Things Can Make a Big Differenc
e Self-Help
239. Max Brooks                      World War Z: An Oral History of the Zombie War
Horror;Fiction
```

Figure 4.2: Book Ratings given by user

In Figure 4.2, More 10 books are shown to user and instructed user to rate 5 books from above listed books. Once user rated 5 books then recommendation is given to user as shown in figure 4.3.

```
Book ID                                          Title  `
    627                The True Story of the 3 Little Pigs
    296             The Very Hungry Caterpillar Board Book
     85                                  The Giving Tree
    339                                   Goodnight Moon
     93                               The Secret Garden
    601                         The Tale of Peter Rabbit
    391                                       The Lorax
    245                            Bridge to Terabithia
    321                         Where the Red Fern Grows
     50                          Where the Sidewalk Ends
    133   Anne of Green Gables (Anne of Green Gables, #1)
    549                                        Corduroy
    454                      If You Give a Mouse a Cookie
    386  Island of the Blue Dolphins (Island of the Blu...
    358                      Oh, The Places You'll Go!
    157                            Green Eggs and Ham
    251                            The Cat in the Hat
    540                             A Little Princess
    497                                   Black Beauty
    413                          The Three Musketeers
    158  Charlie and the Chocolate Factory (Charlie Buc...
    545                          The Velveteen Rabbit
    544     Little House on the Prairie (Little House, #2)
```

Figure 4.3: Book Recommendation to new users

For existing user recommendation the user has to give his user id to the model then the model will evaluate ratings of books from dataset that user has rated as shown in figure 4.4 and based on those ratings and book id the recommendation is given as shown in figure 4.5. Snapshots of recommendation for existing user is as follows:

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 8, August 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0908023 |

```
1. Book Recommendation for New User.
2. Book Recommendation for Existing User.
2

Please Enter User Id: 55

----------------Welcome User55------------
        user_id  book_id  rating
401          55      212       5
418          55      264       5
419          55      397       2
420          55      653       5
421          55      520       4
424          55       43       5
436          55      217       3
7817         55      895       4
12969        55       81       4
14319        55      179       5
14320        55      100       5
14321        55      516       4
16510        55      483       4
16511        55      709       4
16512        55      301       4
17830        55       60       1
17831        55       11       4
```

Figure 4.4: Evaluation of user ratings of books

```
Book ID                                                Title  \
    212                                            Atonement
    179                  Angela's Ashes (Frank McCourt, #1)
     43                                           Jane Eyre
     81                                  The Glass Castle
    100                               The Poisonwood Bible
     87                      Night (The Night Trilogy #1)
    199  Marley and Me: Life and Love With the World's ...
     15                            The Diary of a Young Girl
    101                               Me Talk Pretty One Day
    121                                              Lolita
    653                                          Blindness
    264                                The Sun Also Rises
     33                               Memoirs of a Geisha
    102                            Where the Wild Things Are
```

Figure 4.5: Book Recommendation for existing user

V. CONCLUSION

In this work, we have implemented Hybrid recommender system by combining the traditional popular model, content based and collaborative filtering algorithm. These algorithms gives better recommendations if we combine them, the hybrid book recommendation system can be utilized in various ecommerce websites or for more accurate results the parameters such as language short descriptions can be considered and this model can also be combined with deep learning models.We have obtained RMSE of hybrid 0.6960 which is better than the pure content or

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 8, August 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0908023 |

collaborative filtering.Also Stochastic Gradient Descent can be used for learning weights of content based, collaborative and popular ratings.

REFERNCES

[1] Anand Shanker Tewari **,** Kumari Priyanka **"***Book recommendation system based on collaborative filtering and association rulemining for college students"in*2014 International Conference on Contemporary Computing and Informatics (IC3I) ,Electronic ISBN:978-1-4799-6629-5.

[2] Anand Shanker Tewari , Abhay Kumar,Asim Gopal Barman "*Book recommendation system based on combine features of content based filtering, collaborative filtering and association rule mining*" in 2014 IEEE International Advance Computing Conference (IACC),Electronic ISBN**:** 978-1-4799-2572-8

[3] Shahab Saquib Sohail ,Jamshed Siddiqui , Rashid Ali "Book recommendation system using opinion mining technique" in 2013 International Conference on Advances in Computing, Communications and Informatics (ICACCI) , Electronic ISBN**:** 978-1-4673-6217-7.

[4] Prem Melville, Raymod J. Mooney ,Ramadass Nagarajan "Content-boosted collaborative filtering for improved recommendations" in Eighteenth national conference on Artificial intelligence Pages 187-192, ISBN:0-262-51129-0.

[5] Praveena Mathew ,Bincy Kuriakose ,Vinayak Hegde, "Book Recommendation System through content based and collaborative filtering method" in 2016 International Conference on Data Mining and Advanced Computing (SAPIENCE) ,Electronic ISBN**:** 978-1-4673-8594-7.

[6] Priyank Thakkar ,Krunal Varma ,Vijay Ukani ,Sapan Mankad ,Sudeep Tanwar, "*Combining User-Based and Item-Based Collaborative Filtering Using Machine Learning"*ISBN978-981-13-1746-0.

[7] Rohan Passi,Surbhi Jain,Pramod Kumar Singh, "*Hybrid Approach for Recommendation System*" ISBN978-981-13-1609-8.

[8] Guy Shani,Asela Gunawardana, "*Evaluating Recommendation Systems*" ,ISBN978-0-387-85819-7.

[9] Zhen Zhu , Jing-Yan Wang, "*Book Recommendation Service by Improved Association Rule Mining Algorithm*",ISBN**:** 978-1-4244-0972-3.

[10] Dharmendra Pathak, Sandeep Matharia ,C. N. S. Murthy, "*NOVA: Hybrid book recommendation engine",*ISBN**:** 978-1-4673-4527-9.

INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 7.542

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details