Journal of Theoretical and Applied Information Technology
15th February 2022. Vol.100. No 3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

SWITCHING HYBRID MODEL FOR PERSONALIZED RECOMMENDATIONS BY COMBINING USERS DEMOGRAPHIC INFORMATION

RAGHAVENDRA C K[1,2] , SRIKANTAIAH K.C[1]

[1]Computer Science & Engineering, SJB Institute of Technology,
Bangalore, Affiliated to Visvesvaraya Technological University,
Belagavi, Karnataka 590018, India
[2]Computer Science & Engineering, BNM Institute of Technology,
Bangalore, Affiliated to Visvesvaraya Technological University,
Belagavi, Karnataka 590018, India

[1]raghav.ck.clk@gmail.com, [2]srikantaiahkc@gmail.com

ABSTRACT

Recommendation systems are smart tools that are necessary to help users to find relevant items and for E-Commerce platforms to improve their revenues. List of items based on its rating and content are searched using collaborative filtering (CF) and content-based filtering (CBF). The recommendation system mainly depends on quality of item recommended and the rating of items given by the existing users. In this work a switching based hybrid model which is the combination of collaborative, content and demographic based model was built for recommending books. Cold start problem was solved using demographic data of users. Experimental analysis has been made on individual and hybrid models using RMSE and MAE metrics. Results show that hybrid model outperforms all the other traditional models in terms of predictions and quality of recommendations.
Keywords - *Book Recommendation System, Cold Start, Content Based Filtering, Demographic Based Recommender System, Collaborative Filtering Technique, Hybrid Filtering Technique*

1. INTRODUCTION

With the evolution of technology, the dependency on it has rapidly increased. Technology is seen in almost every sector. Technology usage in banking, finance, media, education, etc. has boomed up in the last two decades. Likewise, online shopping i.e. Electronic Commerce (E-Commerce) has made shopping more convenient. E-commerce is a platform that is primarily based on the internet for the purchase and sale of goods as well as the transmission of data across a network. The widespread usage of e-commerce platforms such as Amazon and eBay has contributed to significant growth in online retail during the previous decade. E-commerce amounted about 5.1 percent of overall retail sales in 2007, while it now accounts for 16.0 percent in 2019.

With the increase in online shopping, the variety of items sold online has increased exponentially. Also, the number people shopping online are rising day-by-day. Due to this, the users find it difficult to search the products of their interest. The tools such as Recommendation systems come into picture. Recommendation systems can be used to guide the users with items that they might prefer. They are basic tools that filter relevant and accurate information from a large pool of data to offer the most relevant and accurate items to the user. The usage of Recommendation systems make it easier for the users to spot relevant items and improve their experience in online shopping. The online shopping websites can improve their revenue with the usage of Recommendation systems.

1.1 Types of Recommendation Systems

Content-based Filtering: Most of the online retail sites are based on content-based filtering. Another name of Content-based filtering is “cognitive filtering”. It recommends products based on comparisons between the item and the user profile. Content-based systems are generally used with text documents. Content-based filtering algorithms usually fail if the user profile or the location descriptions are not well structured.
Collaborative Filtering: This method is also widely used by recommendation systems. It is

known as social filtering. This is one of the most evolved techniques available. The similarity between the pairs of items is generally computed using the cosine similarity metric. The drawback of using cosine similarity is that it does not consider the difference between the ratings of users. Hence adjusted cosine similarity is considered. This newly obtained cosine similarity subtracts the respective user's average rating from each pair. Other similarity metrics could be considered as well. Pearson's similarity metric could be used when the user rating scales are different, cosine metric when the data is sparse, Euclidean metric when the data is not sparse, and the magnitude of the attribute is high, adjusted cosine to adjust for user-bias for item-based approach.

Hybrid Recommender Systems: Hybrid RS combines two or more recommendation approaches in different ways to get the best from both approaches. Generally, in hybrid systems, collaborative filtering is combined with another technique in a weighted way.

Demographic Recommendation Systems: Recommendations based on the user's demographic attributes are generated in this approach. It categorizes the users based on their attributes and uses the demographic data for recommendations. This is being used by many websites because it is easy to implement and does not require user ratings.

Knowledge based Recommender System: This type of recommender system tries to make suggestions based on inferences about the requirements and preferences of the user.

1.2 Challenges of Recommendation System

Cold start: It is one of the most serious problems that Recommendation Systems encounter. When a person joins the system for the first time, he or she will have no history data. Thus, it is difficult to recommend books to him/her. This issue is called cold start problem [1,4]. The same problem arises when there is a new book in the system, where it is difficult to recommend that book to the existing users. Many researchers have tried to solve cold start problem in various approaches. Using the demographic data of the users, this issue can be overcome. Rather than using historical data, users' demographic data might be utilized to recommend books to them. This problem can also be solved with the help of a Knowledge Graph Convolutional Network. [11]. Many other solutions to this problem is necessary to provide significant results.

Sparsity: This issue arises when there is a lack of ratings and reviews by the users. With lesser ratings and reviews it is difficult to understand the user's taste and give recommendations. Thus, reducing the effectiveness of the recommendation system [1,4,7,10]. A possible solution is to use Knowledge Graphs to alleviate this problem [12]. Another solution would be to use matrix factorization. This issue opens up new opportunities for improvement.

Trust issues: This issue arises when certain users have lesser history and it is difficult to recommend books to them. Also, it occurs when it is difficult to decide the amount of weightage to be given to reviews and ratings. The users might have varying tastes, thus making it difficult to recommend books to them. That is, we do not know how much we can trust the existing ratings of users with respect to books. Social network data can be used to reduce the impact of this issue [9]. There is a scope of research in this issue.

Scalability: With the increase in the number of books and users in the system, the scalability issue arises. The system requires more and more resources for recommendation [8]. Also, the performance of the system might not be significant with the increase in number of books and users. Thus, it is necessary to develop recommendation models that can face the scaling up of data.

1.3 Need for Recommendation System

The use of an online recommendation system has grown commonplace. Nowadays, instead of going out and purchasing products for themselves, people prefer to use internet recommendations because it is a simpler and faster way to purchase items, and transactions are also faster when done online. Recommender systems are a strong new technology that assists consumers in locating products they want to purchase. A recommendation system is often used to recommend the best relevant products to end consumers. Nowadays, online book selling Web sites compete with one another based on a variety of factors. One of the most powerful strategies for increasing earnings and maintaining customers is a recommendation system. Existing systems result in the extraction of useless data as well as a lack of user satisfaction.

The demand for online shopping is increasing at a rapid rate. Hence, it is necessary to make it more relevant and convenient to the users. A book recommendation system is necessary in an online book selling website so that the readers'

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

experience with the website is appreciable. It is at the heart of the online book selling website, which makes the website efficient and more relevant to users. The book recommendation system provides personalized recommendations to the readers using a hybrid model so that readers don't find it cumbersome and difficult to search for books of their interest.
In this work a switching based hybrid model which is the combination of collaborative, content and demographic based model was built for recommending books. Cold start problem was solved using demographic data of users

2. LITERATURE REVIEW

Recommendation systems are trending applications nowadays and very important both from the user perspective and the company perspective. Users can conveniently shop with the recommendations being received from the recommendation systems. On the other hand, revenue of the company can be improved by using recommendation system. Many experts have tried to improve upon the current Book Recommendation Systems. They have proposed various new, innovative ideas for the creation of efficient recommendation systems.
Huayong Liu and Nianlai Jiao [1] have proposed a hybrid recommendation system with the usage of context awareness and social network. Various contextual factors that affect the user choice on books are obtained through the context aware layer such as gender of the reader, time of borrowing the book, etc. A user-book-context matrix is established to represent the contextual theme suitable for book recommendation. The contextual factors are associated with the book type in the matrix and then context aware computing is performed to obtain entropy and the weight of each contextual factor. On the other hand, user-to-user similarity is calculated based on Pearson similarity, based on which nearest k users are considered. The books to which the active user has not rated are scored using the corresponding scores of these k nearest users. The obtained scores are combined with context weights to obtain final scores, upon which recommendations are made. It is opined in the paper that with the usage multiple other context factors and multi-dimensional context factors, the system can be further improved.
JiabeiLi, TianweiXu, Juxiang Zhou [2] demonstrates how to use the hybridization method to effectively use content-based filtering and collaborative filtering techniques. The combination of popularity, inverse popularity with similarity and duration of borrowing of book is considered to measure the user's interest and likeability on those books. Inverse popularity highlights that the users who like unpopular books have similar interest. While applying for inverse popularity Borrowing Time has been calculated. Borrowing Time refers to the time between the borrowal and return of the book. It reflects the reader's interest on the book. If a person's borrowal duration on a book is short, he/she might be more interested on it. On the other hand, if it is longer, he/she might be less interested on it is the time. Including all the above mentioned parameters, the recommendation is done. The scalability issues are overcome in this system with the usage of cloud.

Jayanti Rathnavel and Kavita Kelkar [3] proposed a personalized recommender for recommending books to the users. In this experiment, they combined the two popularly, extensively used recommendation techniques i.e. collaborative and content based techniques to build a hybrid recommender. They personalized the system by trying to understand the interests of the users such as favorite author, favorite genre, etc. They have addressed the overspecialization problem. Overspecialization is a limitation in which the recommended books are similar to those that the current user has already read. Using lightfm model, overspecialization is overcome, due to which the recommended set of books also contains the type of books not explored by the active user. It gives the opportunity to the active user to explore new kinds of books. The recommeder can learn the new interests of the active user.
Madhuri Kommineni, P.Alekhya, T Mohana Vyshnavi, V.Aparna, K Swetha, V Mounika [4] discussed that User Based Collaborative Filtering technique along with the cosine rule is more effective to predict the desired books to the user. In User based filtering the system finds the similar preferences of several users and recommends the next book which like-minded user may like to read. This system is very helpful for the administration purpose as it collects the feedback from all the users, report them and analyze the items and recommends most desired output. User profile as well as item profile is maintained to find the "User Behavior" which is very effective in finding the desired output. In order to build collaborative filtering recommendation they have used Singular Value Decomposition (SVD) model

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

which helps to predict more efficiently and effectively. The quick sort algorithm is used to sort the dataset based on the keywords provided by the users after registering. Historical data should be maintained properly in this system.

Praveena Mathew, Bincy Kuriakose and Vinayak Hegde [5] According to the author, combining content-based filtering with collaborative filtering produces more effective and efficient results. Along with these two techniques, associative rule is used to predict the desired items from a large collection of items. This method aims to tackle the problem of sparsity by combining the techniques of Content Based Filtering, Collaborative Filtering, and Associative Rule Mining. The system also implements keyword based recommendation in which, the users enter keywords related to their interests and the system compares these words in the datasets to recommend the books. Equivalence class Clustering and bottom up Lattice Traversal are discussed in this paper (ECLAT), which aims to find frequently read sets of books in an efficient way. ECLAT performs using Depth First Search (DFS), thus scanning the dataset only once and consuming less time compared to other algorithms.

Anand Shanker Tewari and Kumari Priyanka [6] in their paper proposed a book recommendation system based on CF and Association Rule Mining (ARM) for College Students employs the user based CF technique to forecast the top n-rated books for students and academics. This system aims to help the students to find books based on the price ranges and publisher's name. The system employs categorization approaches, collaborative filtering based on user input, and association rule mining. Classification techniques are used to extract a set of rules and patterns in the data and classify the data to predefined classes, each class is processed independently while recommending. Similar people are detected using Pearson's similarity algorithm in user-based collaborative filtering. ARM determines the correlation of each users in the given dataset and associate the relation between users and finds the best suited items. ARM can also be used to discover interesting associations and relationships in the data, which can be used for user behaviour analysis. Based on these techniques, the system recommends books to the readers.

Kitti Puritat and Kannikar Intawong [7] have proposed a model for book recommendation system that uses Support Vector Machine (SVM). They took into account a variety of factors, including title similarity and book bibliographic information like author, year, category, number of books, etc. This model was specifically designed for usage in small libraries. SVM is a supervised machine learning model that can be used to solve classification and regression applications. The SVM is trained using three sources of data i.e. title similarity, Dewey Decimal Classification (DDC) for classification and bibliographic features. The model was found to perform considerably well.

Dharna Patel, Harish Patidar [8] have proposed a Recommendation Solution for Online Book Portal and have explained the need of cloud computing while recommending the books. The value-added feature in this paper is that the system gets the profession of the user while registering into the system. To recommend the book collaborative filtering technique and content-based filtering techniques are being utilized. Cloud computing has been used in this paper as dataset is very large and it is not always possible to store it in local disk, which is very difficult to recover when in case of any loss. In order to secure data, one can adopt cloud computing, which is also known as a storing centre because it maintains enormous datasets. This system of recommendation is more suitable to the readers who require the best book for general purpose rather than specific purpose. Raghavendra et al, [9, 10] provided an study of existing techniques, similarity metrics and research opportunities in this area.

3. METHODOLOGY AND SYSTEM ARCHITECTURE

3.1 Problem Definition

Given GoodReads dataset, the objective is to design and develop a Book recommendation system to suggest relevant books to the readers according to his/her interests using switching weighted hybrid technique.

3.2 Methodology.

The proposed system has the following five main features:

Hybrid model – The system combines both collaborative filtering and content based filtering in a weighted manner. The usage of hybrid model improves the performance and reduces the limitations of each approach when implemented individually.

Cold start for new readers – Both the approaches i.e. collaborative filtering and content based filtering need historical data. When a reader is new to the system, the recommender will not have any historical data of the reader and thus cannot recommend books. This problem is called cold start for readers. In the proposed system this problem is overcome using demographic recommender, which recommends books to readers based on his/her age, gender, location and profession.

Cold start for new books – When a new book is added to the system, the recommender will be unable to suggest that book to the readers. This problem is called cold start for books. In the proposed system, this limitation is tackled using knowledge based recommender, which associates the new books with existing ones using authors and publishers. In case the new book is authored by a new author that book will be recommended based on the matching of keywords.

Personalized recommendation – Every reader will be recommended with a certain number of books after his/her login in the homepage based on their previous records of books bought by them.

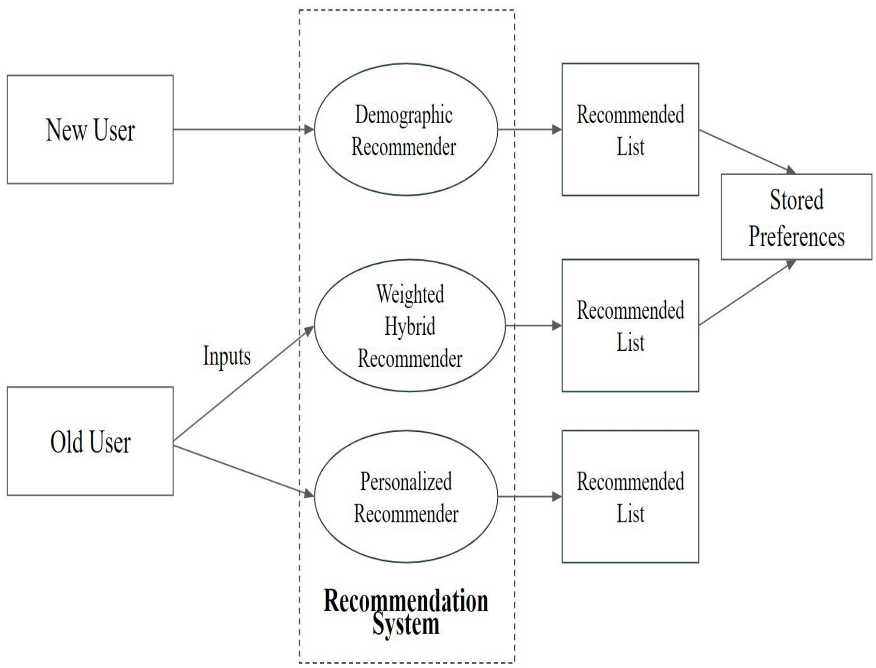

Figure 1. Block Diagram For The Proposed Recommender System.

The idea of the system as shown in figure 1 is to develop a recommendation engine that can more accurately propose books to users based on the user's interests and the features of the books. The data set in question is a massive collection of books, making it a big data set.. The proposed model tries to eliminate the problems like cold start problem by using demographic based recommendation, overspecialization problem by using hybrid model which tries to predict books in such a way that the recommendation list contains book which has not been explored by the user yet.

The proposed system discuss about the recommending list of books to the users by using some of the techniques of recommendation system. The system works for two users such as old users and new user. Old user is the one who has already used the system already and got the list of book recommendation to them and the system remembers the choice of its old user. New

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

user is the one who is using the system for the first time.

- New User: The system takes the demographic input from the user such as age and recommends the list of books which are relevant for that age. The system uses the demographic engine to recommend the books.
- Old User: The system uses the data which are already provided by the users when they provided it earlier and based upon the choices system recommends the list of books. Here, weighted hybrid recommender is used which is the combination of content based, collaborative based and popularity of the boos in the given datasets, The old user gets the personalized recommendation which can be based on age and different genres they have given earlier.

3.3 Content Based Filtering

The CBF system suggests items which are similar to the items bought by the user and based on their ratings in the past. These recommender systems relate various items based on their features. In this approach, the books are recommended based on the cosine similarity between the users and the books which users have preferred earlier. How Content Based System works:

3.3.1 Data preprocessing: The very first step is to prepare the data and clean it for next steps in content based filtering. The following data preprocessing steps are applied on the GoodReads dataset:

All the rows containing any null value in one or more columns are removed and all the unnecessary columns are deleted and the final data frame is created. Now, a new column is created which contains all relevant words from description of the book, genre, and genres. This string containing all words is used to compare different movies and to find similarity between them.

3.3.2. Cosine similarity matrix: It is a metrics used to find the similarity between the two items without depending upon the size of dataset. The two vectors have to be plotted on the multi-dimensional array, and measure the cosine angle between two vectors which can be helpful for the user to find the similarity between two items. Since, multi-dimensional space is used large-sized dataset can also support this technique. Sample code snippet is given below:

```
def content_based(self,md,ratings,user_id):
md['Genres']=md['Genres'].str.split(';')
print(md['Genres'])
md['soup'] = md['authors'] + md['Genres']
print(md['soup'])
md['soup'] = md['soup'].str.join(' ')
count=CountVectorizer(analyzer='word',ngram_ra
nge=(1,1),min_df=0, stop_words='english')
count_matrix = count.fit_transform(md['soup'])
print (count_matrix.shape)
cosine_sim = cosine_similarity(count_matrix,
count_matrix)
```

3.4 Collaborative Filtering

Collaborative filtering is the popularly used filtering technique which is used to find the like-minded users. In this filtering, the algorithm predicts the desired item for the user based on the past history. This filtering identifies similarities among users based on their ratings and suggests new items on the basis of inter-user comparisons. Collaborative filtering techniques are further categorize into two types, those are

- Memory based filtering techniques: This is a type of collaborative filtering which approaches to solve the problem by considering the entire dataset. This algorithm finds the users with the similar active users. Here, active users are the one who wants to make the predictions and add the preference to predict the desired output based the ratings for the active user.
- Model based techniques: In order to build the system much faster and scalable one can approach model based technique which is another type of collaborative filtering. In this technique “model” is the main keyword which refers to build the model based on the ratings present in the dataset.

```
def collaborative(self,ratings,user_id):
reader = Reader()
temp_ratings = ratings
data=Dataset.load_from_df(temp_ratings[['user_i
d', 'book_id', 'rating']], reader)
```

```
svd = SVD()
trainset = data.build_full_trainset()
algo = SVD()
cross_validate(algo, data, measures=['RMSE', 'MAE'], cv=5, verbose=True)
```

3.5 Hybrid Recommendation System

Hybrid approach to the recommender system involves combining two or more techniques Content based and Collaborative filtering is combined into one single model based on weighted hybridization technique. Here's how it works, when user enters the systems it verifies the type of user and based on that it switches the model for generating recommendations. The detailed algorithm is given below:

- Old user: if a user is old user then model switches to weighted hybrid model which is a combination of content and collaborative filtering model. Here ratings are predicted based on combined ratings of both the models by assigning optimal weights to each of them.
- New user: if a user is new user then model switches the demographic model where recommendations are generated based on demographic data of user. Here age attribute of user is considered for generating recommendations to tackle new user problem.

Algorithm1: Hybrid (u, b, P_{CBF}, P_{CF}, P_{clust})
Input: User *u*, Book *b*, P_{CBF} Combined Predictions of Content Based Filtering, P_{CF} Combined Predictions of Collaborative Filtering, P_{clust} Combined Predictions of clusters groped based on demographic data.
Output: $P_{u,b}$ – Predicted rating of user u for book b

1. If users u ∊ { U, set of existing users}
a. P(u,b) = w1 * P_{CBF} (u,b) + w2 * P_{CF} (u,b)
b. Return P(u,b)
End If
2. If user u in new user
a. Find the matching cluster based on demographic data
b. Fetch Predictions of users from the selected clusters
c. P(u,b) = 0
d. For each user *v* in cluster *C*
i. P(u,b) = p(u,b) + p(v,b)
End For
e. P(u,b) = P(u,b) / n // n is number of users in cluster
a. Return P(u,b)
End If

3.6 Cold Start Problem

The new user should register to the recommender system by providing the name, email id and password. If the new user enters to the system then based on the demographic data like gender, location, age and profession the recommender system process the data using demographic filtering and predicts the desired books. The data which is already provided by the news user gets stored in the metadata which can be helpful for the recommender system to recommend the books which the user has already read previously.

When the old user enters the system by providing the login credentials, the user can search the books by providing the keywords then the system matches those keywords in the dataset and recommends the books to the user. This type of recommender system is called as Personalized Recommender as it recommends based on the keywords provided by the user. If the user wants to give some prior information to the system then the system accepts that information and based on the previously stored preferences, the system recommends the books to the user. This type of recommendation is called Hybrid Filtering, when weights are used to filter the data then it is called Weighted Hybridization.

Cold start is the main problem while implementing this kind of techniques. This problem occurs when there is lack of information provided to the recommender system. Hence, the solution is provided below to overcome from this type of problem. To overcome from this, recommending books for new user and new book is needed. For new user, the system takes the input such as age, gender and profession. The recommender system analyses the given data, for example if the user is below 5 then the system recommends lower kg books for the users. And accordingly the systems process the data and outputs the recommended books to the user preferences.

Cold start problem occurs to the new book when the new book arrives to the system. When the new book arrives the system, the filtering technique identifies the respective authors which are already present in the database. If authors are not present in the database then system tries to find whether the publisher is already present in the dataset. Either author or publisher would be present in the dataset; hence by using knowledge based recommendation technique the system recommends the most preferred books. If author is new to the system, then based on the keywords provided by the user in the last login the system recommends the list of books as shown in figure 2.

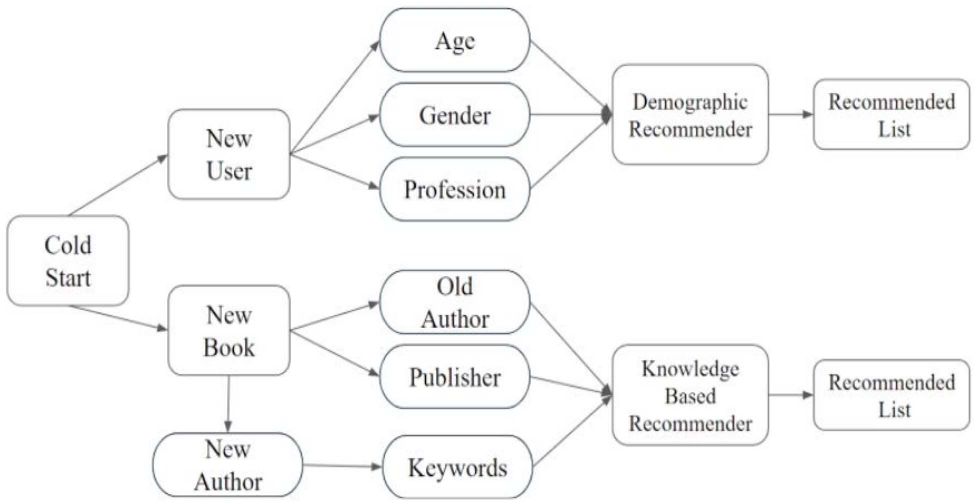

Figure 2. Tackling of Cold Start Problem

4. EXPERIMENTAL RESULTS

This section of the experimental results contains results of the recommender system that gives an idea about project to a layman who does not have good technical knowledge, followed by the recommendations that are given as the output by the selected model, in this case 3 models: Collaborative Filtering, Content-Based Filtering and The Hybrid model.

4.1 Experimental setup

The Experiments were carried out on minimum hardware requirements that include Processor of Intel core i3, Disk space should be at least 1GB, and Operating System must be windows 7 or later. The machine learning algorithms were implemented using Python 3.8 and Jupiter Notebook. Experiments are carried out with GoodReads dataset which contains 10k books, over 53k users and over 9 lakh ratings. The dataset can be downloaded from https://www.kaggle.com/zygmunt/goodbooks-10k which contains two csv files books.csv and ratings.csv. The attributes of GoodReads dataset is given in table 1.

Table 1: Attributes Of Books from GoodReads Dataset

Column	Description
bookId	Unique Id of each book
Title	Titile of the book
authors	Authors details
average_rating	Average rating given for the book
Isbn and isbn13	Unique number that identifies book
Language_code	Language used in book
#num_Pages	Number of pages
Ratings_count	Total ratings received
Text_reviews_count	Total reviews given

4.2 Performance Evaluation

RMSE**:** It is a standard performance measurer for errors in any model. It is used to calculate the errors while predicting quantitative data.

It is defined as:

$$RMSE = \sqrt{\frac{\sum i, j \in X \,|\, r_{ij} - r_{ij}^{'} \,|^2}{|\, X \,|}}$$

(1)

r_{ij} is the actual rating of user, r'_{ij} is predicted rating $|X|$ is test data.

MAE: It is used to measure absolute error, or the difference between two continuous variables.

$$MAE = \frac{1}{n}\sum_{j=1}^{n} |p_j - a_j| \quad (2)$$

Where, n is number of samples, p_j is predicted value and a_j is actual value.

Cross Validation: Cross-validation is a resampling technique for evaluating machine learning models on a small sample of data. Steps include.
1. Randomly shuffle the dataset.
2. Divide the data into a total of k groups.
3. for each one-of-a-kind group:
- Create a holdout or test data set for the group
- Create a training data set for the remaining groups
- Fit a model to the training set and evaluate it on the test set
- Keep the evaluation score but toss out the model.

4. using the sample of model evaluation scores, summarize the model's skill.

4.3 Results and Screenshots

The GUI of developed web based book recommendation application contains the options for viewing books recommended, list of books rated by the user, list of books which are not read by the user and list of books recommended based on popularity. In order to view the details of the system, user has to login with username and password. If user is new to the system he has to sign up using the sign up page by filling the fields given in the form. Some of the key snapshots of the systems are also shown. Table 2 shows the lists of recommendations based on popularity. Table 3 shows the lists of recommendations for new user based on demographic model where top 5 books with genres thriller and young-age are recommender for new user with age 23. Table 4 shows the list of top 5 books recommended for existing user based on weighted hybrid model.

Table 2: The List of Top 10 Recommended Books Based On Popularity

Id	Author	Title	Genre
2	J K Rowling	Harry Potter and the Sorcerer's Store	Fantasy; Young-Age
24	Den brown	The Da Vinci Code	Thriller; Drama
84	Michael Crichton	Jurassic Park	SciFi; Thriller; Fantasy
86	John Grishm	A Time to Kill	Thriller
42	Louisa May Aloott	Little Women	Young-Age; Romance; Drama
44	Nicholas Sparks	The Notebook	Romance; Drama
94	Douglas Adams	The Hitchhiker's Guide to the Galaxy	Fantasy; Fiction
127	Malcolm Gladwell	The Tipping Point	Self-Help
239	Max Brooks	World War Z	Horror; Fiction
966	Scott Turow	Presumed Innocent	Thriller; Crime

Table 3: The List of Top 5 Recommended Books for New User With Age=23

Id	Title	Genre
12	1984	Thriller
29	Gone Girl	Thriller
85	A Time to Kill	Thriller
93	The Husband's Secret	Thriller
24	The Da Vinci Code	Thriller; Drama

Table 4: Hybrid Recommendations For Existing User

Id	Collaborative Rating	Popularity Rating	Content Based Rating	Hybrid Rating
182	5.0	3.91	4.2	4.49
731	3.91	3.92	5.0	4.35
637	4.27	3.92	2.6	3.54
367	4.40	3.94	1.4	3.14
608	4.86	3.95	1.6	3.06

All the three models are evaluated based on prediction evaluation metrics i.e., RMSE and MAE. Table 5 shows the values of RMSE and MAE of all three models where cross validation is applied to evaluate the models with number of folds=5.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 5: RMSE And MAE Values From Cross Validation

CF Model	Fold 1	Fold 2	Fold 3	Fold 4	Mean	Std
RMSE (test set)	0.9306	0.9290	0.9328	0.9280	0.9295	0.0020
MAE (test set)	0.7352	0.7346	0.7382	0.7322	0.7346	0.0021
Fit time	13.70	13.68	13.71	13.82	13.71	0.06
Test time	0.44	0.45	0.37	0.38	0.42	0.03
CBF Model	**Fold 1**	**Fold 2**	**Fold 3**	**Fold 4**	**Mean**	**Std**
RMSE (test set)	0.9306	0.9290	0.9328	0.9380	0.9394	0.0028
MAE (test set)	0.7552	0.7446	0.7582	0.7522	0.7597	0.0022
Fit time	13.76	13.78	13.74	13.77	13.76	0.15
Test time	0.44	0.45	0.47	0.44	0.45	0.03
Hybrid Model	**Fold 1**	**Fold 2**	**Fold 3**	**Fold 4**	**Mean**	**Std**
RMSE (test set)	0.9027	0.8941	0.8961	0.9028	0.8984	0.0036
MAE (test set)	0.7114	0.7075	0.7092	0.7135	0.7100	0.0022
Fit time	14.36	13.67	13.58	14.05	13.85	0.30
Test time	0.50	0.50	0.45	0.49	0.48	0.02

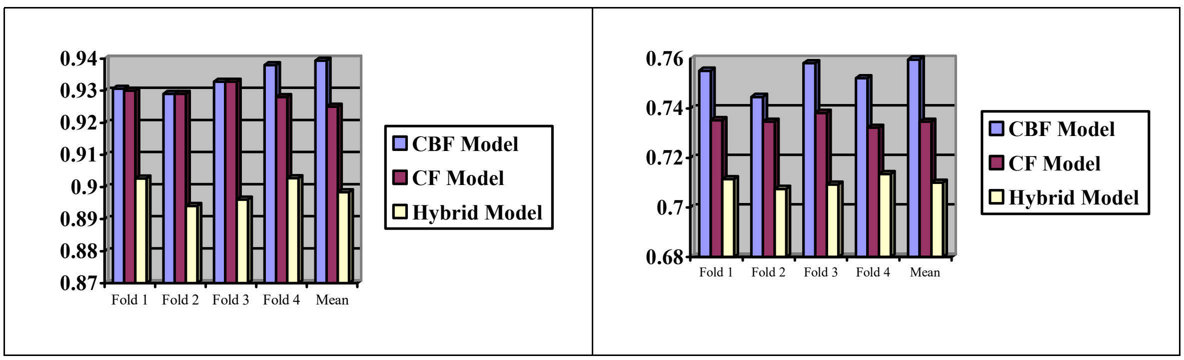

Figure 3. RMSE and MAE Values

Figure 3 shows RMSE and MAE values of all three models and their variations in each fold of cross validation.

Table 4: Comparison of Hybrid Models

Hybrid Model	MAE	RMSE
CF- 0.4, CBF-0.6	0.728	0.9012
CF- 0.6, CBF-0.4	0.710	0.8984
CF- 0.5, CBF-0.5	0.721	0.9078

Table 5: Comparison of All Three Models

Methods	MAE	RMSE
Content Based	0.7597	0.9394
Collaborative	0.7346	0.9295
Hybrid	0.710	0.8984

Hybrid model is experimented with 3 combinations of weights as shown in table 6. Hybrid model with 60 % weight for CF and 40% weight for CBF will give better results compared to other models. Experimental results show that Hybrid model outperforms the content and collaborative models as shown in table 7.

5. CONCLUSION

In this work, a switching hybrid model is proposed which is the weighted combination of collaborative-based and content-based models was built. Switching of RS depends on the type of user. If user is old predictions are made using weighted hybrid model and if user is new predictions are made using demographic based recommender to tackle cold start problem. It was verified that weighted hybrid model performed better than the traditional models with CF weight 0.6 and CBF weight 0.4. Further as an enhancement, we have used GoodReads 10K dataset it can be replaced with 100K dataset for better results. Demographic recommendation model is based on age attribute of the user which can be replaced with multiple attributes like gender, location etc., and Using the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

sentiment analysis of reviews, the performance of the Hybrid model can be even more personalized yielding better predictions.

REFERENCES

[1] Huayong Liu, Nianlai Jiao. "A Hybrid Book Recommendation Algorithm Based on Context awareness and Social Network". In 2020 3rd International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE)

[2] JiabeiLi, TianweiXu, Juxiang Zhou. "A personalized Recommendation Algorithm for College Books based on User Interest". In 2020 IEEE 2nd International Conference on Computer Science and Educational Informatization (CSEI).

[3] Jayanti Rathnavel, Kavita Kelkar. "Personalized Book Recommendation System". In International Journal Of Engineering And Computer Science ISSN: 2319-7242, Volume 6 Issue 4 April 2017, Page No. 21149-21153.

[4] MadhuriKomminen,P.Alekhya,T.MohanaVyshnavi,V.Aparna,KSwetha,VMounika. "Machine Learning based Efficient Recommendation System for Book Selection using User based Collaborative Filtering Algorithm". International Journal of Advanced Research in Computer and Communication Engineering, 2018.

[5] Ms.PraveenaMathew, Ms. BincyKuriakose, Mr.VinayakHegde."Book Recommendation System through Content Based and Collaborative Filtering Method". In Department of Computer Science And Engineering, international journal of innovations in engineering and technology (IJIET), ISSN:2319-1058,vol.2, issue 2, pp1-7, April 2017.

[6] AnandShanker Tewari, Kumari Priyanka. "Book Recommendation System Based on Collaborative Filtering and Association Rule Mining for College Students". In 2014 nternational Conference on Contemporary Computing and Informatics (IC3I).

[7] Kitti Puritat, Kannikar Intawong. "Development of an Open Source Automated Library System with Book Recommendation System for Small Libraries". In 2020 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT & NCON).

[8] Dharna Patel, Dr. Harish Patidar. "Hybrid Recommendation Solution for Online Book Portal". In International Journal for Research in Applied Science & Engineering Technology (IJRASET) ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887 Volume 6 Issue V, May 2018.

[9] Raghavendra C K, Srikantaiah K.C, Venugopal K. R, " Personalized Recommendation Systems (PRES): A Comprehensive Study and Research Issues", International Journal of Modern Education and Computer Science(IJMECS), Vol.10, No.10, pp. 11-21, 2018. DOI: 10.5815/ijmecs.2018.10.02

[10] R. C K and K. C. Srikantaiah, "Similarity Based Collaborative Filtering Model for Movie Recommendation Systems," *2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS)*, 2021, pp. 1143-1147, doi: 10.1109/ICICCS51141.2021.9432354.

[11] Hongwei Wang, Miao Zhao, Xing Xie, Wenjie "Knowledge Graph Convolutional Networks for Recommender Systems" In arXiv:1904.12575v1 [cs.IR], 18 Mar 2019.

[12] Chang-You Tai, Meng-Ru Wu, Yun-Wei Chu, Shao-Yu Chu, Lun-Wei Ku "MVIN: Learning Multiview Items for Recommendation" In arXiv:2005.12516v1 [cs.IR], 26 May 2020.

[13] HUAXUAN ZHAO, HONGCHEN WU, JINGZHI LI, HUAXIANG ZHANG AND XINJUN WANG. "Dtree2vec: A High-Accuracy and Dynamic Scheme for Real-Time Book Recommendation by Serialized Chapters and Local Fine-Grained Partitioning". SPECIAL SECTION ON ADVANCED DATA MINING METHODS FOR SOCIAL COMPUTING, January 2020.

[14] Cheng Zhao, Chenliang Li, Rong Xiao, Hongbo Deng, Aixin Sun. "CATN: Cross-Domain Recommendation for Cold-Start Users via Aspect Transfer Network". In arXiv: 2005.10549v2 [cs.IR] 23 May 2020.

[15] Singh, P.S., Srikantaiah, K.C. "ATSR: distance and time based advanced tour sequence recommendation system", Int.j. inf. Technology, 13,pp.1531-1538,2021.