Journal of Engineering Sciences
Vol 12, Issue 7, July/2021
ISSN NO:0377-9254

HYBRID RECOMMENDATION SYSTEM

Rathna Jyothi Ch[1], V.S Pavan Mallik G[2], S.V Rajan D[3] and Nagul Baji Sk[4]
[1]*chrjyothi@aliet.ac.in Andhra Loyola Institute of Engineering and Technology, JNTUK, India.*
[2]*guttapavanmallik23@gmail.com Andhra Loyola Institute of Engineering and Technology, JNTUK, India.*
[3]*svaradarajan111@gmail.com Andhra Loyola Institute of Engineering and Technology, JNTUK, India.*
[4]*nagulnani786@gmail.com Andhra Loyola Institute of Engineering and Technology, JNTUK, India.*

***Abstract*— During this age of knowledge, it's extremely difficult to hunt out the right information from a large amount of data present on online platforms. Because the knowledge available electronically is drastically increasing every minute, the need for effective information retrieval and implementation of filtering tools became essential for easy access to relevant information. Recommendation systems have made finding the things easy that we'd like. Book recommendation systems aim at helping bibliophiles by recommending the books they might have an interest in. The proposed solution is introducing a model combining both content-based and collaborative-filtering approaches to realize more accurate recommendations.**

***Keywords*— Book Recommendation System, Hybrid Recommendation System, Machine Learning, Content-based Filtering, Collaborative-based Filtering, Cosine Similarity, K-Nearest Neighbors.**

1. INTRODUCTION

Recommender systems play a serious role in today's e-commerce industry. Recommender systems recommend items to users like books, movies, videos, electronic products, and lots of other products generally. Recommender systems help users to require correct decisions in their online transactions, increase sales and redefine the user's web browsing experience, retain the purchasers, enhance their shopping experience. The knowledge overload problem is solved by search engines, but they are not providing the personalization of knowledge. Recommendation engines provide personalization.

The user can gain the benefit because they interact with the Hybrid Recommendation System while checking out a specific book as the system also shows some recommendations which are quite almost like the one user is checking out.

The user just has got to access the interface and input a book name within the input field provided and click on the submit button. The Machine Learning model which we've developed retrieves the book details if the book is within the datasets provided. Alongside the small print, the model generates recommendations by applying two algorithms one among which is cosine similarity for generating content-based recommendations, and therefore the other is that the K-Nearest Neighbors algorithm for generating collaborative-based recommendations. Finally, the user can access the book details and a few recommendations which he could be curious about.

1.1 OVERVIEW OF RECOMMENDATION SYSTEM

Recommender systems aim to predict user's interests and recommend product items that quite likely are interesting for them. They are among the foremost powerful machine learning systems that online retailers implement to drive sales. Recommender systems function with two sorts of information:

- Characteristic information

  This is often information about items.

- User-item interactions

This is often information like ratings, number of purchases, likes, etc. Based on this, we will distinguish between three systems utilized in recommender systems:

1. Content-based systems

   These systems make recommendations employing a user's item and profile features. They hypothesize that if a user was curious about an item within the past, they're going to once more have an interest in it within the future. Similar items are usually grouped to support their features.

2. Collaborative filtering systems

Collaborative filtering is currently one of the foremost frequently used approaches and typically provides better results than content-based recommendations. These sorts of systems utilize user interactions to filter for items of interest. In short, collaborative filtering systems have supported the idea that if a user likes item A and another user likes an equivalent item A also as another item, item B, the primary user could even be curious about the second item.

3. Hybrid systems

They combine both sorts of information meaning to avoid problems generated when working with only one kind.

2. ABOUT THE PROPOSED WORK

2.1 LITERATURE SURVEY

This section overlooks similar existing solutions and examines the algorithms used and drawbacks. Approach Towards Hybrid Recommendation System using Content-Based and Collaborative Filtering Techniques [1] study used traditional machine learning concepts Pearson Correlation, K-Nearest Neighbors. A drawback of this study is that the Pearson Correlation used to generate recommendations isn't efficient with every type of operation. The reason Pearson Correlation Coefficient is invariant to adding any constant is that the means are subtracted out by construction. Review Paper on Collaborative Filtering [2] study incorporate the user's response models into the probabilistic matrix factorization (PMF), a popular matrix factorization CF model, to establish the response aware probabilistic matrix factorization (RAPMF) framework. Movie Recommendation System [3] study used Pearson Correlation and K-Means Algorithm to generate recommendations. This solution as using Pearson Correlation isn't effective at all times.

2.2 PROPOSED WORK

In this section, the proposed work is elaborated at a high-level scope.

A. DESIGN METHODOLOGY

The proposed work regarding Hybrid Recommendation System consists of two main modules namely

- Content-based Recommendation Module

The Content-based Recommendation module deals with code that generates similar books to a given book by considering the content of the books. This module uses the Cosine Similarity, which finds the similarity of things represented as vectors and pointed during a 2-d plane. Finally, the cosine angle distance between the vector points plotted on the plane decides the similarity between those points.

- Collaborative-based Recommendation Module

The Collaborative-based Recommendation module deals with code that generates recommendations for a given book based on the user-item collaboration. During this module, we use the K-Nearest Neighbor algorithm which finds k-nearest books. We consider the rating vector of a given book and find all the books whose ratings vectors are almost like the given book. Finally, these k-nearest books are recommended to the user because from the data we've, the model generates the foremost similar books whose rating vectors are clustered into one using the K-NN algorithm.

B. ALGORITHMS

We shall be using two algorithms to develop a Hybrid Recommendation model. They are

- Cosine Similarity

The cosine similarity of two objects is the similarity between those objects regarding the cosine angle between them when plotted on a 2-dimensional plane.

"If the angle Θ is a smaller amount i.e., 0 then the 2 objects are perfectly similar or same. If the angle Θ is 90o then the objects are least similar."

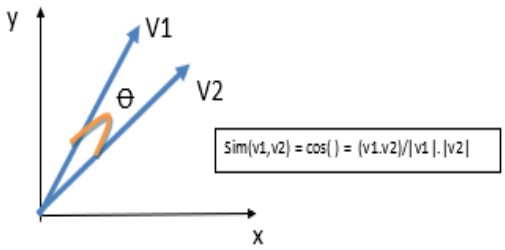

Fig. 1 Visualization of Cosine Similarity

Cosine similarity is computed using the following formula:

$$similarity(A,B) = \frac{A \cdot B}{\|A\| \times \|B\|} = \frac{\sum_{i=1}^{n} A_i \times B_i}{\sqrt{\sum_{i=1}^{n} A_i^2} \times \sqrt{\sum_{i=1}^{n} B_i^2}}$$

Fig.2 Cosine Similarity Formula

To find the similarity between the content (summary) of the books, we should convert the summary of books into vector notation. Then we can get n books whose contents(summary) are almost like the content of a given book using the cosine similarity.

- k-Nearest Neighbors

k-NN is a machine-learning algorithm to find clusters of similar items.

The k-NN working can be explained based on 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.

Here to implement collaborative filtering-based recommendations we use this algorithm to seek out books supported common book ratings, and make predictions using the typical rating of top-k nearest neighbors. For that, we'd like to preprocess the info we've from our datasets regarding books, users, ratings within the sort of a matrix during which each row represents a book and every column represents a user. For a given book, we then apply the k-NN model to seek out k nearest neighbors supported by the rating vector of books.

C. DESIGN AND WORKFLOW

The backbone of the proposed work is the data that we collected from web scraping, bookcrossing datasets.

Fig. 3 Architecture diagram of the proposed system.

The datasets are used to train the recommendation model which consists of a Content-based model and a Collaborative-based model. The content-based model requires the data obtained from web scraping. This data contains details of books such as Book name, author, ISBN, a summary of the book, image URL.

The collaborative-based model requires the book-crossing dataset which contains book name, ISBN, author, userID, ratings.

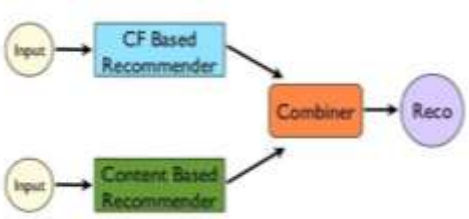

Fig. 4 Implementation of the proposed system

The proposed work consists of three parts as shown in Fig. 4

a. Generate content-based recommendations

From the data we've collected, the summary of books may be a vital feature to get content-based recommendations. Firstly, we'd like to convert the summary of books into vector notation which may be wont to find cosine similarity. To realize that, we are using the tensorflow_text module of Python which converts the multilingual sentences to a 512-byte vector. Then applying the pairwise cosine_similarity function, we will find n books whose contents are almost like the given book at a time. This is often actually done by the cosine_similarity function by applying the cosine similarity formula on all the vectors in a pairwise fashion.

b. generate collaborative-based recommendations

From the book-crossing dataset, we'd like to preprocess data to get a matrix notation during which each row contains a book name and every column represents a userID and therefore the values of the matrix are the ratings given by the user to the books. This usually would be a sparse matrix (Matrix whose values are mostly zeroes). To predict recommendations accurately, we need to predict the ratings which users haven't given to books. this is often achieved by using the Singular Value Decomposition (SVD) prediction algorithm by training with 80% of the data we have from each user and tested with the remaining 20% of data. And then, we use the k-NN model with the new matrix and now it generates K nearest neighbors whose rating vectors are nearer to the given book.

c. combine both recommendations

Finally, when a user interacts with the user interface and requests a book, the model processes the request and generates recommendations from both models, combines them, and displays them to the user.

D. RESULT ANALYSIS

The results obtained from the proposed system are pretty much accurate. As in Fig. 5, when a user accesses the interface, he will be having a choice to submit a book name as input.

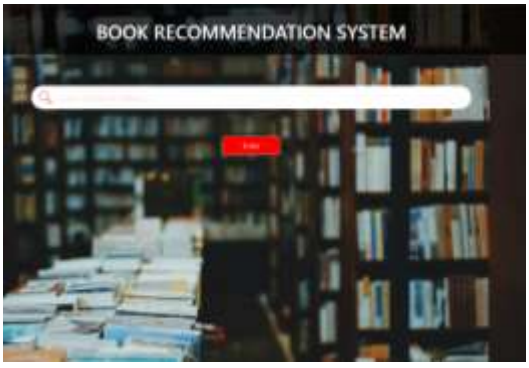

Fig. 5 User Interface

When a user clicks on submit button, the Hybrid Recommendation System processes the request and generates recommendations as shown in Fig. 6 which are a combination of content-based and collaborative-based models.

Fig. 6 Details of Book and Recommendations

These are the recommended books our proposed system is suggesting to the user.

Fig. 7 Recommendations and metrics

In Fig. 7 we can see the recommended books along with their respective cosine similarity value or distance metric of the k-NN algorithm.

The Mean Square Error and Root Mean Square Error of the ratings our model had predicted are pretty accurate as given in Table I.

Table I. Accuracy Metrics

MSE	RMSE
0.26654383619	0.51627883570

Thereby, as in Fig. 8, while generating recommendations for a book, we can see that the average of actual ratings that particular book has received from all users is almost close to the average of predicted ratings. That resembles our prediction of ratings using SVD and the recommendations generated are relevant to the requested book based on user-book ratings.

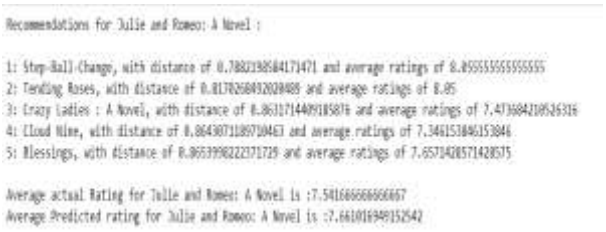

Fig. 8 Actual Ratings and Predicted Ratings

3. CONCLUSION AND FUTURE WORK

In this paper, the proposed and implemented Hybrid Recommendation System is a Machine Learning Application and may be used at any Book Store, E-Commerce application, Libraries. This application recommends
books to users supported available past data of item description and user-item interaction. It takes the advantage of overcoming the sparsity problem. Thus, this application makes the task of a user easy by recommending books he could also be curious about supported the user's past data. The proposed system has been developed in such a fashion, which helps for future development. The Recommendation System which we developed recommends books to users supported static data available to the system, which is additionally an obstacle to be utilized in real-time. The longer-term enhancements to the present system could be by including real-time user interaction like user registration, user login, and integrating this data with the book's details, which can generate more accurate user-specific recommendations within the desired application domain.

ACKNOWLEDGMENTS

This research was partially supported by The Dept. of Computer Science and Engineering, Andhra Loyola Institute of Engineering and Technology, JNTU Kakinada. We are grateful to our Professor Dr. Ch. Rathna Jyothi for leading us to develop and contribute a paper to the journal.

REFERENCES

[1] Debadrita Roy and Arnab Kundu, "Design of Movie Recommendation System by Means of Collaborative Filtering", International Journal of Emerging Technology and Advanced Engineering, vol. 3, no. 4, 2013.

[2] Ananya Agarwal, S. Srinivasan, "Movie Recommendation System", International Research Journal of Engineering and Technology (IRJET), Volume: 07 Issue: 07

[3] Mayuri Dalvi 1, Prof. S.V. Gumaste, "REVIEW PAPER ON COLLABORATIVE FILTERING", International Research Journal of Engineering and Technology (IRJET), Volume: 02 Issue: 09.

[4] Siddhesh Jagtap, Yash Mane, Tushar Kadam, Trupti Dange, "Approach Towards Hybrid Recommendation System using Content-Based and Collaborative Filtering Techniques", International Research Journal of Engineering and Technology (IRJET), Volume: 07 Issue: 08.

[5] Y. Shi, M. Larson, and A. Hanjalic, "Collaborative Filtering Beyond the User-Item Matrix: A Survey of the State of the Art and Future Challenges," ACM Comput.Surv., vol. 47, no. 1, pp. 3:1–3:45,2014.

[6] Kumar, Manoj & Yadav, D.K. & Singh, Ankur & Kr, Vijay, "A Movie Recommender System: MOVREC", 2015 International Journal of Computer Applications. 124. 7- 11. 10.5120/ijca2015904111.

[7] S. G Walunj, K Sadafale, "An online recommendation system for e-commerce based on Apache Mahout framework", 2017 ACM SIGMIS International Conference on Computers and People Research, pp. 153-158,2013.

[8] H. W. Chen, Y. L. Wu, M. K. Hor and C. Y. Tang, "Fully content-based movie recommender system with feature extraction using neural network," 2017 International Conference on Machine Learning and Cybernetics (ICMLC), Ningbo, China, 2017, pp. 504-509. doi: 10.1109/ ICMLC.2017.8108968.

[9] RSML Patibandla, AP Gopi, BT Rao (2021), Comparative Study on Analysis of Medical Images Using Deep Learning Techniques, Deep Learning for Biomedical Applications, 2021.

[10] BT Rao, RSML Patibandla, (2021), Comparative Study on Security and Privacy Issues in VANETs, Cloud and IoT-Based Vehicular Ad Hoc Networks, 2021

[11] A Peda Gopi, (2020),” Design and Analysis of CMOS LNA with Extended Bandwidth For RF Applications”, Journal of Xi'an University of Architecture & Technology, Vol. 12, Issue. 3,pp.3759-3765.

[12] Z. D Zhao, M. S Shang, "User-Based collaborative filtering recommendation algorithms on Hadoop", Proc. of Third International Workshop on Knowledge Discovery and Data Mining, pp. 478-481, 2016.