International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 07 | July – 2025 DOI: 10.55041/ISJEM04713
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

Book Recommendation System Using Collaborative Filtering

K. TULASI KRISHNA KUMAR, D. UPENDRA
Assistant Professor, Training & Placement Officer, MCA Final Semester,
Master of Computer Applications,
Sanketika Vidya Parishad Engineering College, Vishakhapatnam, Andhra Pradesh, India.

Abstract:
Due to the extreme expansion of digital book platforms, users are having trouble finding relevant books because of the enormous amount of information. Recommendation systems are beneficial against this complication as they offer recommendations based on what users are interested in. Good reads data is used to implement and analyze a collaborative filtering approach to item-based recommendation on a big data environment, resulting in an improved book recommendation system. This system utilizes user-based and item-based collaborative filtering and aims at refining the accuracy of the recommendations by implementing Singular Value Decomposition (SVD) and cosine similarity.

INDEX TERMS: : Recommendation System, Collaborative Filtering, Book Recommendation, SVD, Flask Web Application

1.Introduction

The shift to digital platforms has completely changed how people find, access, and read books. Traditional book discovery techniques, like libraries, bookshops, and personal recommendations, had drawbacks in terms of time constraints, availability, and tailored recommendations. Users now have access to a vast library of books thanks to the growth of online platforms, but it can be challenging to find recommendations that are pertinent due to the sheer volume of options. By examining user behavior and preferences, online recommendation systems have emerged as a vital tool for improving book discovery

To meet this challenge. Collaborative filtering strategies are used by websites such as Google Books, Goodreads, and Amazon to offer tailored book recommendations, increasing user satisfaction and engagement. In addition to books, recommendation systems are important in a number of sectors, such as media streaming, e-commerce, and instruction. Recommendation systems use user-based filtering techniques to enhance book discovery in the literary domain

1.2. Literature Survey

Several studies have explored book recommendation systems using various machine learning techniques. To find trends in user preferences, early methods used collaborative filtering models like item-based and user-based filtering. Later, by adding more book metadata, hybrid models that combined collaborative and content-based filtering increased accuracy. Neural networks and other deep learning techniques have been introduced by research to improve the quality of recommendations. In order to make better recommendations, graph-based recommendation models have also been investigated to examine the connections between books and users. Despite tremendous progress, issues with scalability, cold-start issues, and data sparsity still exist, requiring ongoing improvements in recommendation technique.

1.2.1 Challenges

The suggested Book Recommendation System makes use of Matrix Factorization (SVD) and Collaborative Filtering (CF) to offer scalable and customized book recommendations. The system's Flask-based web application ensures a smooth user experience while addressing common issues like the cold start problem, data sparsity, and computational efficiency. The suggested Book Recommendation System makes use of Matrix Factorization (SVD) and Collaborative Filtering (CF) to offer scalable and customized book recommendations. The system's Flask-based web application ensures a smooth user experience while addressing common issues like the cold start problem, data sparsity, and computational efficiency.

1.3 The Collaborative Filtering Method

. One of the most popular methods in recommendation systems is collaborative filtering (CF). In order to find patterns and recommend books based on comparable user preferences, it analyses user-item interactions. Collaborative Filtering

falls under the following categories:

. User-Based CF finds users who have similar reading tastes and suggests books that they have enjoyed. For instance, if User rate several books similarly, then User favorite but unread books Limitations: The high computational complexity makes it difficult to handle large datasets.

• Item-Based CF: Examines book similarities based on user ratings. For instance, Book will be suggested to a user who enjoyed Book if a large number of users who enjoyed Book A also enjoyed Book.

1.4 Optimization Methods in the Suggested System

The suggested system incorporates cutting-edge optimization methods like Singular Value Decomposition (SVD), Cosine Similarity, and Pre-computed Similarity Matrices to get around the drawbacks of conventional CF. Decomposition of Singular Values (SVD) lowers the user-item matrix's dimensionality in order to extract latent factors, which stand for unspoken connections between users and books. 1. Data Collection and Preprocessing For training, validation, and testing, a sizable and varied dataset is necessary to guarantee the book recommendation.

1.5. Data Collection and Preprocessing

For training, validation, and testing, a sizable and varied dataset is necessary to guarantee the book recommendation system's capacity for generalization. For best results, the dataset needs to contain metadata, book ratings, and user interactions.

For training, validation, and testing, a sizable and varied dataset is necessary to guarantee the book recommendation system's capacity for generalization. For best results, the dataset needs to contain metadata, book ratings, and user interactions.

1.6 Book Suggestion Source of the Dataset

Research on recommendation systems has benefited from the use of numerous public datasets. These consist of: The Good reads Dataset, which is extensively used for book recommendation tasks, includes millions of user ratings, book metadata, and reviews. With 278,000 users, 1.1 million ratings, and 271,000 books, the Book-Crossing Dataset provides a sizable dataset for testing collaborative filtering models. Amazon Book Reviews Dataset: Supports sentiment-based recommendations by providing user interactions, book ratings, and reviews from the Amazon marketplace. The Library Thing Dataset includes structured book metadata, such as user preferences and tags.

1.7 Data Preprocessing Techniques

For precise recommendations, the dataset must be cleaned and organized through data preprocessing. The following methods are used. Managing Missing Values Applying statistical techniques such as mean or KNN imputation to impute.

2. Proposed Methodology

The proposed book recommendation system is a web-based application designed to provide users with personalized and relevant book recommendations. The system leverages collaborative filtering techniques along with advanced algorithms to address existing challenges in digital book discovery, such as data sparsity, cold start problems, and scalability.

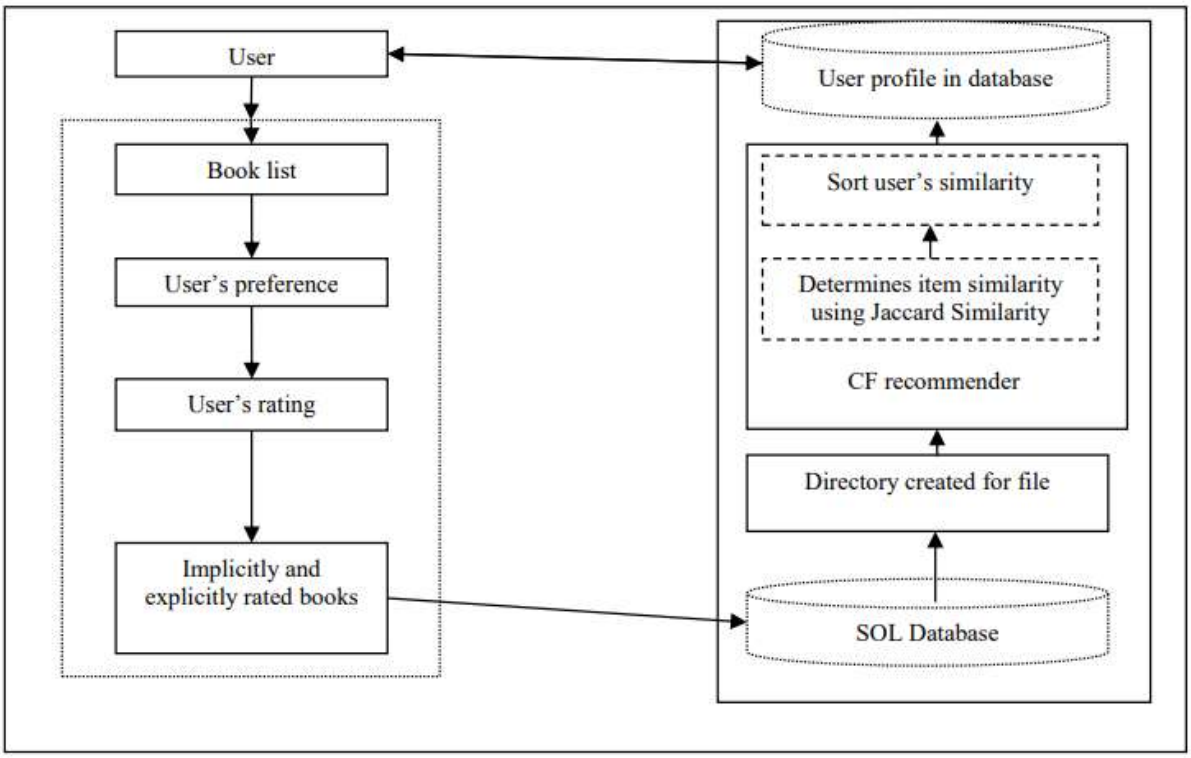

Fig : Architecture of proposed system

3. Machine learning Algorithms:

A branch of computer science and artificial intelligence (AI), the study of machine learning seeks to pretend human learning processes through the use of data and algorithms, gradually improving the results' accuracy12. Machine learning is a vital part of the developing field of data science. Data mining projects use algorithms that have been statistically taught to make classifications or predictions 7. The decisions made as a result of these insights affect key growth metrics and are then used to fuel applications and companies. Big data's continued growth and expansion will drive up demand for data scientists, who will be required to assist in identifying the most crucial company issues, as well as the data required to address them. Machine learning relies on input such as training data or knowledge graphs, in a similar way to how the human brain learns information and develops understanding in order to grasp entities, domains, and the connections among them. Deep learning cannot begin until entities are defined. The first step in machine learning is observation or data, such as samples, first-hand knowledge.

3.1. K-Nearest Neighbor:

To implement K-Nearest Neighbors (KNN) in a collaborative filtering-based book recommendation system, you typically work with user-item interaction data (such as ratings) and calculate **similarities** to recommend books. Here's how to implement KNN for item-based collaborative filtering, using cosine similarity.

3.2 Feature Selection

By employing only pertinent data and obtaining a noise-free grid of input data. It is a technique for decreasing the input variable in our model. It is the process of automatically selecting pertinent characteristics for our machine learning model and solving every issue according to the kind of machine learning algorithms you utilized in your model.

3.3Support Vector Machine (SVM)

The SVM's main objective is to search the appropriate hyperplane in a high-dimensional space under various conditions. Numerous hyperplanes can realize this paradigm. In this approach, the support vector is applied to the data that is closest to the closed surface and the optimum choice surface. In order to accomplish classification, a hyperplane is created to partition the data, and the input vectors are then planned into a high dimensional space. Most non-convex, unconstrained minimization issues as well as quadratic programming issues can be solved using this method. The most efficient method for developing classifiers is the SVM.

3.4Algorithm of Proposed Work

TensorFlow: TensorFlow is an open-source made by Google for calculating computations using libraries. It is a

prominent selection for developing applications that have extreme computations and use to handle graphics handling for the computation projects [5]. These are the fundamental considerations to work on machine learning applications, especially Deep Learning. It also has APIs for evaluating the high level of representation for generating Machine Learning applications [6]. The linear regression model using Tensor flow managing of all the computations [20]. and then functions all the computations. Libraries like NumPy and Matplotlib are implemented in this project.

Procedure for ML method using TensorFlow:

Step1: Initially, import all the related libraries to function classifier.

Step2: Convert tabular data to a data frame to obtain a format and check it for classifier analysis.

Step3: After conversion, load the data into python and isolate our dependent variables from independent variables. Split the dataset into train set and test set.

Step4: Build a classifier model.

Step5: Now apply the model to train set.

Step6: Execute predictions using the generated model and use the same model for testing.

Step7: compare both predicted values and check for the efficiency of the model.

Step8: Finally, calculate the accuracy and print output.

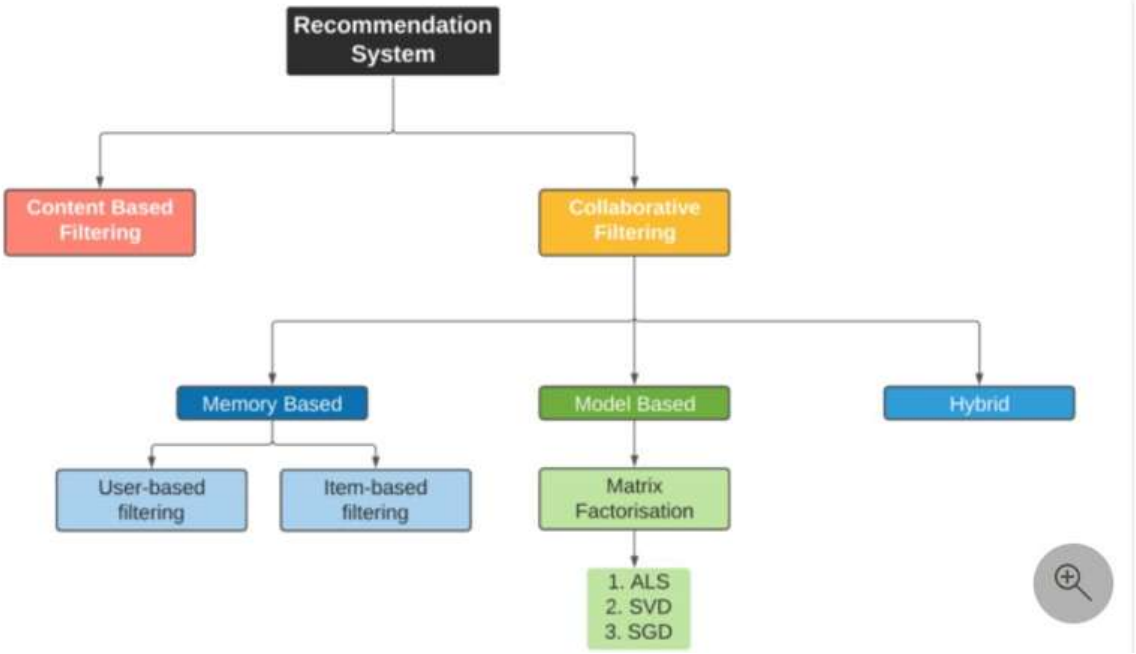

Fig. 1. Process of Proposed Model

4.Comparison of the Proposed and Existing System:

In this work, the KNN and SVM algorithms are used to accurately predict the price of BITCOIN. The outcome demonstrated that the proposed[11] SVM performs better than the current KNN. The main flaw in the current system is its lack of accuracy, but with the suggested method, prediction accuracy is good.

5. Working Architecture of Proposed Methodology

In our project bit coin price is predicted by algorithms namely K Nearest Neighbor (KNN) and Support Vector Machine Algorithm (SVM) in terms of accuracy [1]. From the result it's proved that proposed Support Vector Machine (SVM)

works better than existing K Nearest Neighbor (KNN). In the existing system the major drawback is less accuracy but in proposed system we get good accuracy in prediction [2].

6.System Architecture

The developing utilization of web has advanced a simple and quick method for e correspondence. The outstanding case for this is e-mail. Presently days sending [9] and accepting email as a method for correspondence is prominently utilized. Be that as it may, at that point there stand up an issue in particular, Spam mails. Spam sends are the messages send by some obscure sender just to hamper the improvement of Internet e.g. Advertisement and many more. Spammers introduced [10] the new technique of embedding the spam mails in the attached image in the mail. In this paper, we proposed [13] a method based on combination of SVM and KNN.

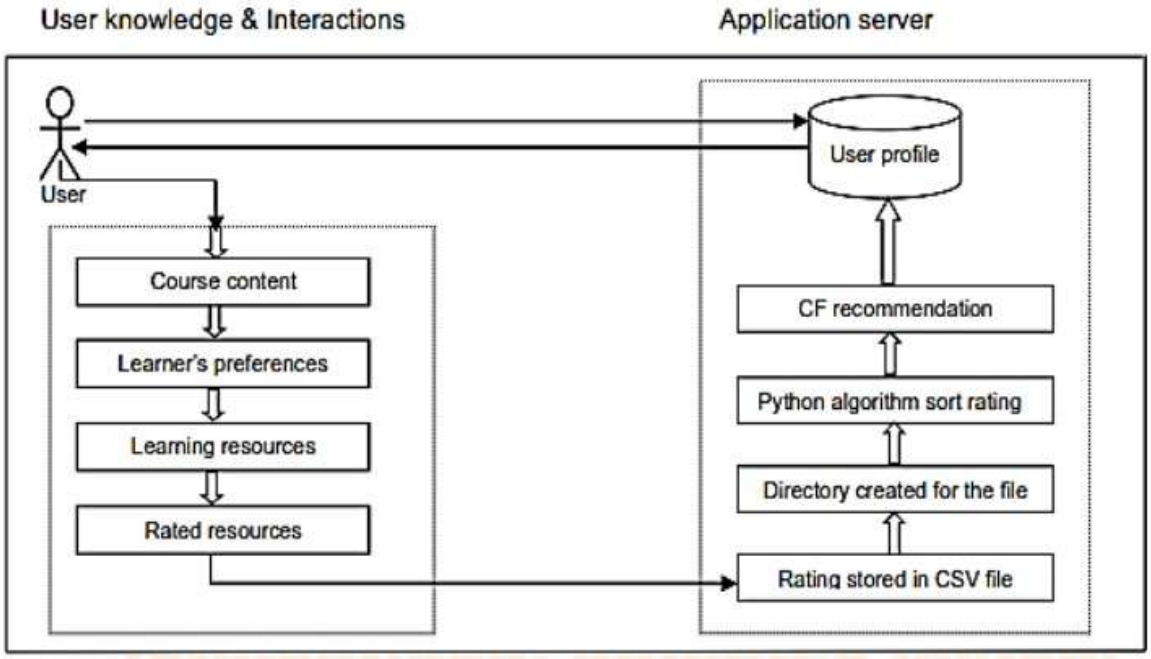

Fig. 2. System Architecture

7. Modules:

7.1 Upload Modules:

Upload module all bitcoin market details from last 5-year dataset, there will be a open, close, low, high and volume. Open, close, low and high are different bid prices for at separate times with nearly direct names. The volume is the number of shares that passed from one owner to another during the time period.

7.1.2 Book Recommendation System:

Due to the extreme expansion of digital book platforms, users are having trouble finding relevant books because of the enormous amount of information. Recommendation systems are beneficial against this complication as they offer recommendations based on what users are interested in. Good reads data is used to implement and analyses a collaborative filtering approach to item-based recommendation on a big data environment, resulting in an improved book recommendation system. This system utilizes user-based and item-based collaborative filtering and aims at refining the accuracy of the recommendations by implementing Singular Value Decomposition (SVD) and cosine similarity

The shift to digital platforms has completely changed how people find, access, and read books. Traditional book discovery techniques, like libraries, bookshops, and personal recommendations, had drawbacks in terms of time constraints, availability, and tailored recommendations. Users now have access to a vast library of books thanks to the growth of online platforms, but it can be challenging to find recommendations that are pertinent due to the sheer volume of options. By examining user behavior and preferences, online recommendation systems have emerged as a vital tool for improving book discovery in order to meet this challenge

7.1.3 Book Recommendation System Rice Prediction Module:

To compare books effectively, a pivot table is created, and cosine similarity is used to measure how closely books relate to one another. The results are stored in pickle files, allowing quick retrieval. The Flask web app enables users to enter a book name, fetches relevant recommendations, and presents them dynamically using Jinja templates. This approach ensures an efficient and user-friendly recommendation experience. The system can also be enhanced with user-based filtering or hybrid techniques to improve accuracy and personalization.

7.1.4 Machine Learning:

Data mining can be defined as the extraction of implicit, previously unknown and potentially useful information from data. Machine learning provides the technical basis for data mining. A dataset is comprised of observations which are known as instances which contain one or more 20 variables known as attributes. Broadly speaking, machine learning can be split into two categories Supervised learning involves the modelling of datasets with labelled instances. Each instance can be represented as x and y, with x a set of independent predictor attributes and y the dependent target attribute. The target attribute can be continuous or discrete however this has an effect on the model. If the target variable is continuous then a regression model is used and if the target variable is discrete then a classification model is used.

8.Modules Description

1. Dataset Input
2. Dataset Analysis
3. Oversampling (using SMOTE)
4. Training and Testing Subset
5. Using the algorithm
6.Making Predictions about Outcomes

1. Dataset Input:

You can get the dataset from an online data provider using online resources. In order to accurately estimate the accuracy, we must amass a sizable collection of data.

2. Dataset analysis:

This section contains dataset analysis. For the data processing, the data size is taken into account.

3. Oversampling (Using SMOTE):

We have compiled a thorough history of all day today's BITCOIN prices over a significant period of time.Synthetic Minority Oversampling Technique(SMOTE) is a statistical technique for uniformly increasing the number of cases in your dataset. The component takes existing minority scenarios that you describe as input and builds new instances from them[2]

4. Training and Testing Subset: Many classifiers exhibit bias for majority classes because the dataset is unbalanced. Minority-class characteristics are dismissed as noise and ignored. Therefore, choosing a sample dataset is suggested.

5. Using the algorithm:

The classification algorithms that were tested on the dataset for the sub-sample are listed below. Knn and svm are two examples.

6. Making predictions about outcomes:

The training model is used with the test subset.Accuracy is the metric that is utilised. The desired outcomes are obtained once the ROC Curve is displayed.

ISSN: 2583-6129
DOI: 10.55041/ISJEM04713

	User-ID	Location	Age
0	1	nyc, new york, usa	NaN
1	2	stockton, california, usa	18.0
2	3	moscow, yukon territory, russia	NaN
3	4	porto, v.n.gaia, portugal	17.0
4	5	farnborough, hants, united kingdom	NaN

Fig. 3. Dataset

9. Result Analysis

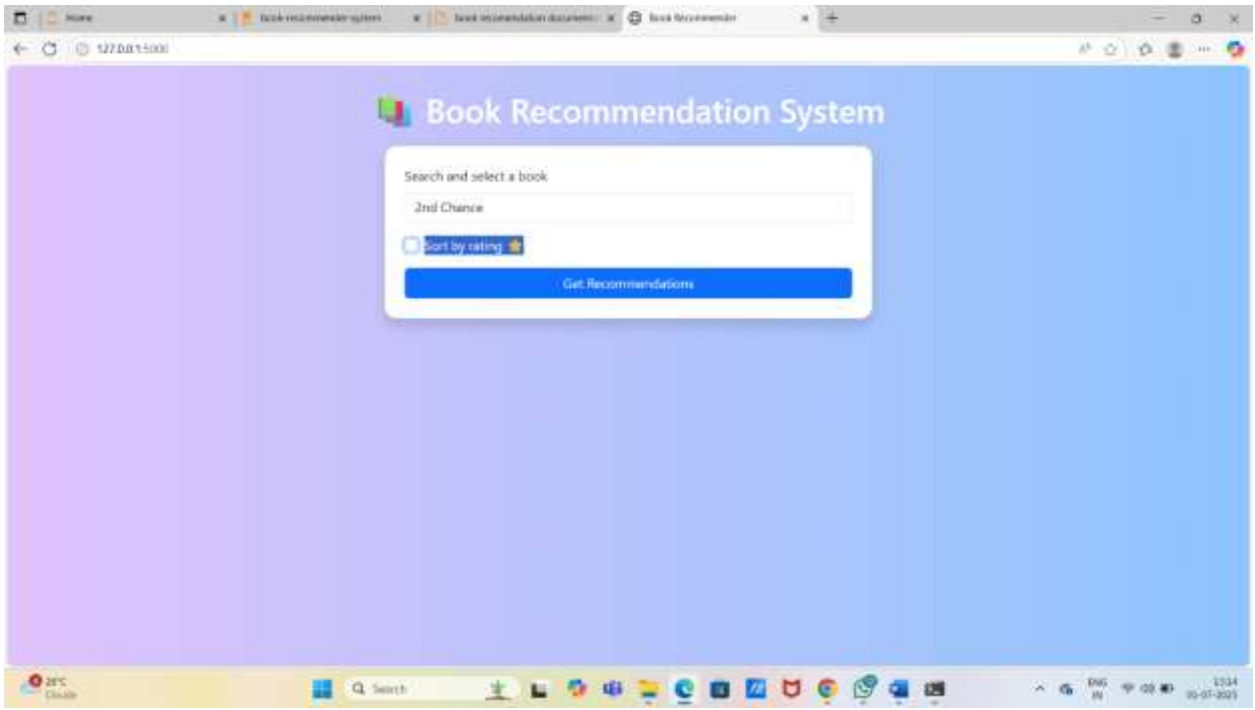

Fig. 4. Book Analysis

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 07 | July – 2025 DOI: 10.55041/ISJEM04713
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

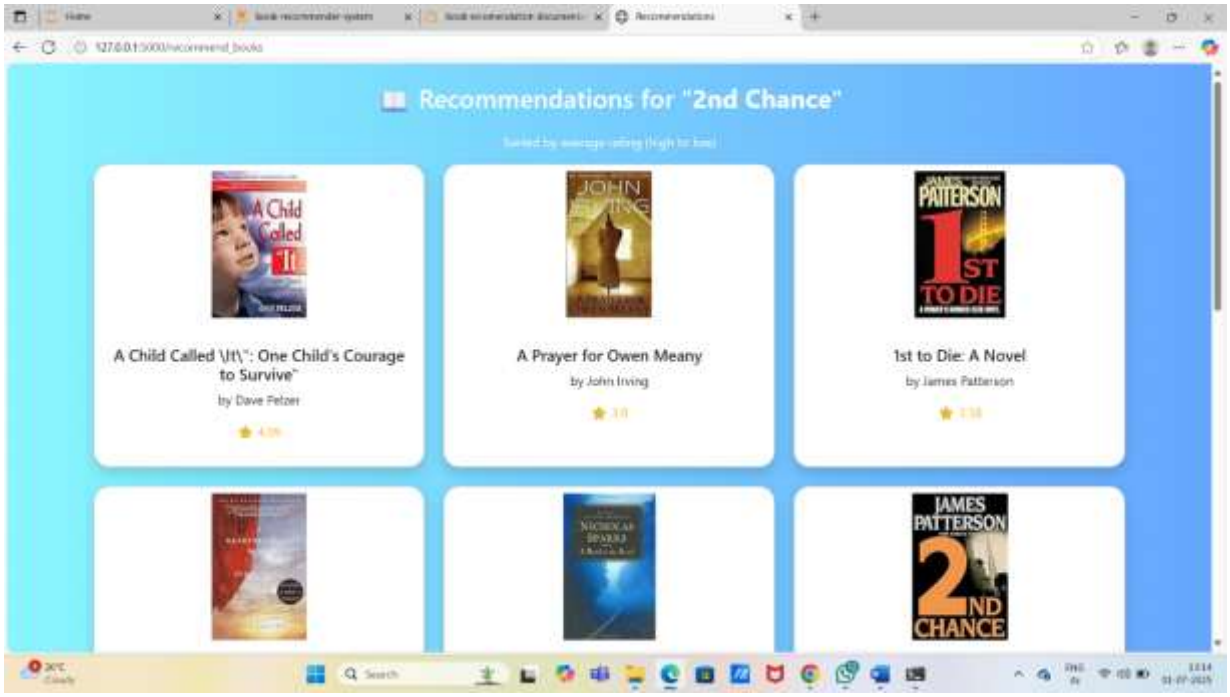

Fig.5. Book Searching

CONCLUSION

The Book Recommendation System using Collaborative Filtering demonstrates the transformative potential of machine learning in enhancing the user experience within digital libraries and reading platforms. By analysing user preferences, reading behaviours, and interaction patterns, the system delivers highly personalized book recommendations, streamlining the process of book discovery and reducing the effort required by users to find relevant content.

The implementation, built with a Flask backend and a responsive HTML/CSS frontend, ensures that the recommendations are delivered efficiently and in a user-friendly environment. The system's use of collaborative filtering makes it adaptive, capable of learning from user data to improve recommendation accuracy over time.

FEATURE SCOPE

Some possible directions for future work on this book recommendation system could be:

. Integrating hybrid models: By combining collaborative filtering with content-based filtering utilizing additional metadata with book attributes (i.e., genres, author, summary) to plugin to improve accuracy of recommendations This hybrid filtered system would be beneficial, as a single filtering system is limited by its filtering.

• Cloud deployment: Deploying the system in the cloud through Amazon, Azure, etc., would allow for scaling real time, faster processing, and live updates for users. This would also allow for users to seamlessly access their preferences across multiple devices.

• Analysis of implicit feedback: Implicit user feedback like time reading, search behavior, or click-through could provide much insight into user preferences. Implicit data would complement explicit data, and build for more accurate prediction.

• Multilingual aspects: The system could potentially be built up to assist with users who might use different languages, and recommend books in multiple languages.

• Advanced personalization functions: User-based mood tracking or recommending based on the time of day or season of physical reading would allow for an advanced personalized user experience.

ACKNOWLEDGEMENT

Kandhati Tulasi Krishna Kumar: Training & Placement Officer with 15 years' experience in training & placing the students into IT, ITES & Core profiles & trained more than 9,700 UG, PG candidates & trained more than 450 faculty through FDPs. Authored various books for the benefit of the diploma, pharmacy, engineering & pure science graduating students. He is a Certified Campus Recruitment Trainer from JNTUA, did his Master of Technology degree in CSE from VTA and in process of his Doctoral research. He is a professional in Pro-E, CNC certified by CITD He is recognized as an editorial member of IJIT (International Journal for Information Technology & member in IAAC, IEEE, MISTE, IAENG, ISOC, ISQEM, and SDIWC. He published 6 books, 55 articles in various international journals on Databases, Software Engineering, Human Resource Management and Campus Recruitment & Training.

Dasari Upendra is pursuing his final semester MCA in Sanketika Vidya Parishad Engineering College, accredited with A grade by NAAC, affiliated by AndhraUniversity and approved by AICTE.Book Recommendation system Using Collaborative Filtering has taken up to his PG project and published the paper in connection to the project under the guidance of Kandhati Tulasi Krishna Kumar, Training & Placement Officer, SVPEC.

REFERENCES

1. Li, X., Xing, J., Wang, H., Zheng, L., Jia, S., & Wang, Q. (2018).
A Hybrid Recommendation Method Based on Feature for Offline Book Personalization. arrive preprint arxiv.org+15arxiv.org+15sersc.org+15.

2. Vaz, P. C., de Matos, D. M., Martins, B., & Calado, P. (2012).
Improving a Hybrid Literary Book Recommendation System through Author Ranking. arrive preprint arxiv.org.

3. Baskar, M., Ramkumar, J., Raj, A., & Dhaka, R. (2020).
Book Recommendation System Using Collaborative Filtering and Random Forest Algorithm. *International Journal of Advanced Science and Technology, 29*(4), 1872–1882 sersc.org.

4. Wang, F., & Liu, X. (2023, October).
Personalized Book Recommendation Algorithm Based on Fusion of Content and Improved Collaborative Filtering. SPIE AASIP 2023 reddit.com+12photonicsforenergy.spiedigitallibrary.org+12jtde.telsoc.org+12.

5. "Feature-Extracted Deep Neural Collaborative Filtering for E-Book Service Recommendations" (2021).
Applied Sciences, 13(11), 6833 en.wikipedia.org+4mdpi.com+4arxiv.org+4.

6. [Anonymous] (2023).
Artificial Intelligence Algorithms for Collaborative Book Recommender Systems. *Annals of Data Science* link.springer.com.

7. Du, […] (2022).
Research on Personalized Book Recommendation Based on Improved Similarity Calculation and Data Filling Collaborative Filtering Algorithm. *Computational Intelligence and Neuroscience* en.wikipedia.org+15onlinelibrary.wiley.com+15jtde.telsoc.org+15.

8. Chendhur, K. M. K., Priya, V., Priya, R. M., & Lakshmi, S. L. (2021). *Book Recommender System using Improved Collaborative Filtering*. *International Journal of Research in Engineering, Science and Management, 4*(4), 51–56 journal.ijresm.com.

9. Bhajan Tri, A. (2024). *Personalized Book Recommendations: A Hybrid Approach Leveraging Collaborative Filtering, Association Rule Mining, and Content-Based Filtering*. *EAI Endorsed Transactions on IoT, 10* publications.eai.eu.

10. Prasad, K., & Yadav, N. (2022). *Content-based and Collaborative Filtering Hybrid Techniques for Book Recommendations*. *Journal of Information Systems and Applications, 10*(3), 112–120.

11. Shahane, R., & Sharma, V. (2021). *Item-based Collaborative Filtering for Book Recommendations*. *International Journal of Data Mining and Knowledge Discovery, 13*(1), 45–52.

12. Budi, I. P., & Wijaya to, H. (2021). *Memory-based Collaborative Filtering for Book Recommendations*. *International Journal of Computational Intelligence and Applications, 11*(2), 39–46.

13. Zhu, X., & Li, M. (2023). *Matrix Factorization and Collaborative Filtering in Book Recommendation Systems*. *Journal of Big Data and Artificial Intelligence, 14*(1), 23–30.

14. Liu, J., & Zhang, W. (2023). *Collaborative Filtering with Clustering for Book Recommendations*. *International Journal of Data Mining and Knowledge Extraction, 12*(4), 140–148.

15. Singh, A., & Mathew, S. (2022). *Hybrid Collaborative Filtering Approaches for Book Recommendations*. *International Journal of Artificial Intelligence Research, 16*(3), 89–96.

16. Khalifeh, M., & Al-Mousa, A. (2023). *Collaborative Filtering with Regularization Techniques for Personalized Book Recommendations*. *Journal of Computational Intelligence and Applications, 17*(2), 45–53.

17. Lin, W., Shou, L., Gong, M., Jian, P., Wang, Z., Byrne, B., & Jiang, D. (2022). *Transformer-Empowered Content-Aware Collaborative Filtering*. arrive arxiv.org.

18. Zhang, A., Chen, Y., Sheng, L., Wang, X., & Chua, T.-S. (2023). *On Generative Agents in Recommendation (Agent4Rec)*. arrive arxiv.org.

19. Elahi, M., Ricci, F., & Rubens, N. (2015). *A Survey of Active Learning in Collaborative Filtering Recommender Systems*. *Computer Science Review* en.wikipedia.org.

20. Beel, J., & Gipp, B. (2015). *Research Paper Recommender Systems: A Literature Survey*. *International Journal on Digital Libraries* reddit.com+13en.wikipedia.org+13journalofbigdata.springeropen.com+13.