

Personalized Book Recommendation System using Machine Learning Algorithm

Dhiman Sarma^{1*}, Tanni Mittra², Mohammad Shahadat Hossain³

Department of Computer Science and Engineering, Rangamati Science and Technology University, Rangamati, Bangladesh¹

Department of Computer Science and Engineering, East West University, Dhaka, Bangladesh²

Department of Electrical and Electronic Engineering, University of Science and Technology Chittagong, Chittagong, Bangladesh³

Abstract—As the amounts of online books are exponentially increasing due to COVID-19 pandemic, finding relevant books from a vast e-book space becomes a tremendous challenge for online users. Personal recommendation systems have been emerged to conduct effective search which mine related books based on user rating and interest. Most of these existing systems are user-based ratings where content-based and collaborative-based learning methods are used. These systems' irrationality is their rating technique, which counts the users who have already been unsubscribed from the services and no longer rate books. This paper proposed an effective system for recommending books for online users that rated a book using the clustering method and then found a similarity of that book to suggest a new book. The proposed system used the K-means Cosine Distance function to measure distance and Cosine Similarity function to find Similarity between the book clusters. Sensitivity, Specificity, and F Score were calculated for ten different datasets. The average Specificity was higher than sensitivity, which means that the classifier could re-move boring books from the reader's list. Besides, a receiver operating characteristic curve was plotted to find a graphical view of the classifiers' accuracy. Most of the datasets were close to the ideal diagonal classifier line and far from the worst classifier line. The result concludes that recommendations, based on a particular book, are more accurately effective than a user-based recommendation system.

Keywords—Personalize book recommendation; recommendation system; clustering; machine learning

I. INTRODUCTION

Most organizations have their recommendation system when they sell products online. But almost all the websites are not developed of the buyer interest; the organizations' force add-on sells to buyers by recommending unnecessary and irrelevant products. A personalized recommendation system (PRS) helps individual users find exciting and useful products from a massive collection of items. With the growth of the internet, consumers have lots of options on products from e-commerce sites. Finding the right products at the right time is a real challenge for consumers. A personalized recommendation system helps users find books, news, movies, music, online courses, and research articles.

The fourth industrial revolution emerges with a technological breakthrough in the fields like the internet of things (IoT), artificial intelligence (AI), quantum computing, etc. The economic boom improves the living standard of people and elevates the purchasing power of individuals. Nowadays, physical visits to shops and libraries have been

drastically reduced due to their busy schedules and COVID-19 pandemic. Instead, e-marketplaces and e-libraries became popular hotspots. E-book reading platforms and online purchasing tendencies made users discover their favorite books from many items. As a result, users tend to get swift and smart decisions from an unprecedented amount of choices using expert systems. Thus, recommendation systems came into the scene to customize users' searching and deliver the best-optimized results from a multiplicity of options. A personalized recommendation system was initially proposed by Amazon, which contributed to raising Amazon's sales from \$9.9 billion to \$12.83 billion in 2019 (second fiscal quarter) that was 29% more than the previous year [1].

The recommendation systems' algorithms were usually developed based on content-based filtering [2], associative rules, multi-model ensemble, and collaborative filtering. Multi-model ensemble algorithms can be used for personalized recommendation systems, but content-based filtering needs a massive amount of real-world data to train the predictive model. Apriori algorithm is used to find the association rules and degree of dependencies among rules. Multiple classifiers are typical for multi-model based RS. In that case, two different layers can be enforced. In the first layer, a few basic classifiers are trained, and in the second layer, the basic classifiers are combined by using ensemble methods like XGBoost or AdaBoost. A multi-model ensemble algorithm is also used in spatial pattern detection. It can calculate the spatial anomaly correlation with each other and can cluster the anomaly correlations. The clustering technique works as a filter to detect spatial noise patterns [3]. Collaborative filtering filters items based on the similar reactions. It searches a large group of people and can detect a smaller set of users who have a similar taste for collecting items. The similarity measure is a significant component of collaborative filtering. It can find the sets of users who show the behavior to select items [4].

Four main techniques are widely used to developed recommendation systems – collaborative, content-based, hybrid, cross-domain filtering algorithms. Firstly, collaborative filtering uses users' information and opinions to recommend products. It has narrow senses and general senses. It can make automatic predictions based on user preferences by collaborating information from many users in a narrow sense. For example, collaborative filtering could make predictions about a user that television shows a user like or dislike based on partial information of that user. In a general

*Corresponding Author

sense, collaborative filtering involves collaborating large volumes of multiple view-point, agents, and sources. It can be applied in mineral exploration, weather forecasting, e-commerce, and web applications where a massive volume of data needs to be processed to make the predictions. The drawback of collaborative filtering is that it needs a tremendous amount of user data, which is realistic for some applications where we do not use information.

On the other hand, content-based filtering use objects information and recommendation are made based on object similarity. Generally, content-based filtering is useful when we do not have useful information. The Similarity among the products is considered while recommending. Both supervised and unsupervised machine learning algorithms are applied to measure the Similarity among products. The content can be structured, semi-structured, and unstructured, but it must be synchronized into a structured format to calculate the Similarity. A hybrid recommendation system combines two or more filtering techniques to produce the output. The performance of hybrid filtering is better comparing to collaborative and content-based filtering. Collaborative filtering does not consider domain dependencies, and content-based filtering does not consider people's preferences. A combined effort is required from both collaborative and content-based filtering techniques to make better predictions. The combined effort increases the common knowledge in collaborative filtering with content data and content-based filtering with user preferences. Cross-domain filtering algorithms can access information that belongs to different domains. Cross-domain filtering algorithms make predictions by exploring the source domain and increase the prediction in the target domain.

This paper proposed a clustering-based book recommendation system that uses different approaches, including collaborative, hybrid, content-based, knowledge-based, and utility-based filtering. Clustering allows regrouping all books based on the rating and user preference datasets. Such clustering shows remarkable prediction capability for a personalized book recommendation system. The core target of this research is to model an improved approach for customizing the recommendation system.

II. BACKGROUND AND RELATED WORK

Recommendation systems (RSs) or recommendation algorithms are immensely used by personal and corporate entities for searching news and information, pursuing online shopping, engaging in social dating, executing search optimization, etc. [5] [6]. Recommendation systems escalate user adhesion, elevate user experience, and accelerate the use of efficiency of the system. With the rising popularity of e-book reading tendency, and readers increasing demands for finding desired book, book recommendation system plays a significant role [7] while choosing books.

Table I shows a comparison of machine learning-based book recommendation systems with limitations, descriptions,

and used machine learning algorithms. Most of the researcher prefers collaborative filtering to the developed recommendation system. Collaborative filtering requires a vast amount of real-time user data that is not realistic for most recommendation systems. Besides, Table I shows that some researches have low accuracy, and some face overfitting due to small data size. In the paper, we proposed a cosine-distanced recommendation system that uses both user information and preferences.

Collaborative filtering is a very common technique for book recommendation [18] [19] [20]. But the accuracy of this technique was 88% [21] or 89% [22], which is comparatively low. However, a content-based recommendation system needs an enormous amount of training data set, which is not feasible for real-world scenarios [2]. When Jaccard similarity was added with collaborative filtering, it achieved the highest recall. The major drawbacks of a collaborative recommender system are sparsity and cold-start issues. These issues can be removed using a kernel-based fuzzy technique that scored a 95% accuracy rate [23].

The content-based filtering method [2] [24] was used to recommend items based on the Similarity among articles. The major drawback of this method is that it ignores current users' ratings when suggesting new items. But user rating is relevant for recommending new books or journals. As the user rating information is missing in the documents, the content-based filtering has low accuracy in the current book or journal recommendation.

Most of the systems are powered with Artificial Intelligence that search items on popularity, correlation, and content of books [25]. Other popular techniques for RSs are listed as influence discrimination model [26], linear mix model [27], transfer meeting hybrid for unstructured text [28], pseudo relevance feedback [29], fixed effect model [30], natural language processing with sentimental analysis [31], opinion leader mining [32], fuzzy c-mean clustering [33], knowledge graph convolution network, a personal rank algorithm using neural network [34], k-nearest neighbor, and frequent pattern tree [35]. Online search has an abnormal effect on the recommendation system. For example, clicking on high ranking books has no impact but clicking on low ranking books has a positive impact [30]. Data sparsity is another major problem for the traditional book recommendation system, which can be solved using a personal rank algorithm using a neural network [34]. Both k-nearest neighbor and frequent pattern tree are highly efficient for recommending scientific journals for academic journal readers [35]. Moreover, several context-aware rule-based techniques [36], and their recent pattern-based analysis [37] or classification-based techniques [38] [45] [46] or rule-based belief prediction [39] [40] [41] can be used to build the recommendation systems. In this paper, a clustering-based recommendation system was used to achieve the highest accuracy.

TABLE I. COMPARISON OF MACHINE LEARNING-BASED BOOK RECOMMENDATION SYSTEMS

Description	Methods	Limitations	Ref.
* can rank the recommended book * opinion mining technique is used to improve the accuracy of the recommendation system	*opinion mining	* Limited to computer science-related books. * Only recommend 10 books for a particular query	[8]
* combined features from three widely used filtering techniques - content-based filter, association rule, and collaborative filtering	* content-based filtering * collaborative filtering * association rule mining	* These types of filtering techniques need a vast amount of real-time data	[9]
* combined features from two widely used filtering techniques - content-based filter and collaborative filtering	* Content-Based and Collaborative Filtering	* The authors use some unnecessary attributes like name of a registered user, password of the registered user, and email	[10]
* Scraping information is useful for making recommendations. * Consider the temporal aspects while recommending books * Overcome the problems of the content-based and collaborative filtering	* Item-Item Similarity Technique * Web Scraping Process	The authors failed to explain the impact of clustering in the recommendation system Web-based recommendation system needs to be secure	[11]
* Consider scholar reviews, which is helpful library user education * The authors explain how a recommendation system can be applied to grow the interest of the reader to a particular type of books	* problem-based learning (PBL) model intelligent mobile * location-aware book recommendation system	* Only limited to library and e-library book recommendation. * Not suitable for e-commerce-based book recommendation system	[12]
* Positional aggregation based scoring efficiently finds top-ranked books for a university student.	* aggregation based scoring; * fuzzy quantifiers, * Ordered Weighted Averaging	* Limited to university books recommendation system	[13]
* matrix sparsity problem in filtering is solved by the author * Can recommend books to newly admitted university's student with high accuracy	* collaborative filtering algorithm	* Cluster of books did not consider in recommendation; The authors did not consider borrowing the time and length of books	[14]
Use a user-based similarity matrix to increase the accuracy of the collaborative filtering algorithm	* User-Based collaborative filtering	* Cluster can improve the accuracy and performance of the recommendation system	[15]
support-vector machines are used to find the relationships between titles of the books or bibliographic of the authors	support-vector machines	* Dataset contains only 4612 books, which may lead to overfitting problems	[16]
users' behavior-based collaborative filtering recommends a series of books	users' behavior-based collaborative filtering	Low accuracy of the classifier, which is 59%	[17]

III. METHODOLOGY

The proposed system in Fig. 1 used a clustering technique to develop the recommender system. Fig. 1 shows three parts named data acquisition, preprocessing, and clustering techniques. The datasets were collected from the Goodreads-books repository of kaggle in this research. Though Goodreads-books repository of kaggle contains seven datasets, only four datasets (Books.csv, Book_tags.csv, Ratings.csv, and Max_Rating.csv) were considered for this experiment. The preprocessing technique was applied after merging all datasets where we removed the lower-rated books and developed a new dataset for analysis. Finally, the clustering technique was applied for recommending books to those users who stay in proximity to a specific cluster. Besides, a user can then search for a book through a query interface, and results in listing recommended books (Fig. 6).

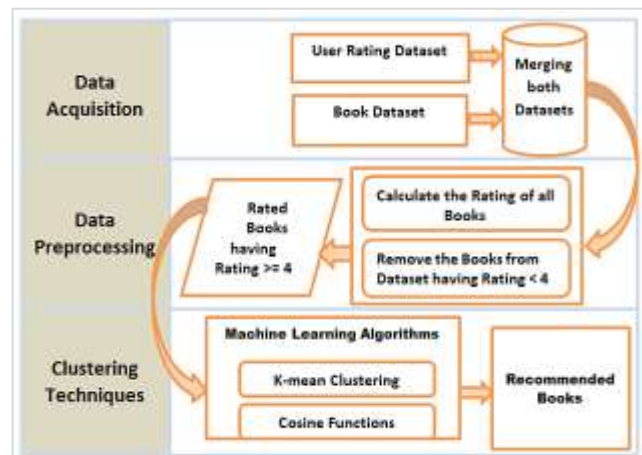


Fig. 1. Proposed Architecture for Book Recommendation.

A. Data Acquisition

The dataset was collected from the GoodReads book dataset repository. It has 10,000 rated data of popular books. This data set consists of 7 tables named Books.csv, Geners.csv, Book_tags.csv, Max_rating.csv, Ratings.csv, to_read.csv, and Tags.csv, where we used Books.csv and Book_tags.csv as book dataset and Ratings.csv and Max_Rating.csv as user rating dataset. The description of the datasets are as follows:

- Books.csv- it has attributes like an author, book_isbn number, rating and contains 10K books.
- Book_tags.csv- it has 596K rows and attributes are goodreader_book_id and tag_id.
- Ratings.csv- it has attributes like user_id,book_id, and rating and contains about 9,00,000 rows.
- Max_Rating.csv- it has similar attributes as Rating.csv. But the number of rows is about 500K.

B. Data Preprocessing

Unstructured noisy text in the data is needed to be preprocessed to make them analyzable. To do the analysis, the dataset needs to be cleaned, standardized, and noise-free. Fig. 2 shows that most of the books were rated 4 or above. We want to recommend only top-rated books. So we remove all the rows having a rating less than 4. It shows us that 68.89% of books were rated 4 and above. Thus our cleaned dataset becomes compact, standardized, and noise-free.

C. Clustering Techniques

K-mean algorithm is used as a cluster partition algorithm where each partition is considered as a k cluster. It is an agile algorithm applied in cluster assessment, feature discovery, and vector quantization. In this experiment, the k-mean algorithm begins with selecting the numbers of k cluster of books. Each book is assigned to the nearest cluster center and moved from the cluster center to cluster average and repeated until the algorithm reaches to convergence state.

Fig. 3 shows the cosine similarity function which calculates the cosine of the angle between two non-zero vectors (vectors A and B). When these vectors align in the same direction then they produce a similarity measurement of 1. If these vectors align perpendicularly then the similarity is 0, whereas two vectors align in the opposite direction will produce a similarity measurement of -1.

Suppose we put a type 'romantic' on the X-axis and 'adventure' on the Y-axis. Then, book B₁ (Sense and Sensibility) in the romantic type creates an angular difference of 90° to the book A₁ (Treasure Island) in the advancer type. Thus,

Cosine similarity between A₁ and B₁ is:
 $\cos 90^\circ = 0$ (1)

Cosine distance between A₁ and B₁ is:
 $1 - \cos 90^\circ = 1 - 0 = 1$ (2)

The angular difference between item A₁ (Treasure Island) and A₂ (Harry Potter) is 0°. Thus,

Cosine similarity between A₁ and A₂ is:
 $\cos 0^\circ = 1$ (3)
 Cosine distance between A₁ and A₂ is:
 $1 - \cos 0^\circ = 1 - 1 = 0$ (4)

where, cosine distance 0 represents that two objects are similar and adjacent, and cosine distance 1 suggests that the objects remain faraway.

The Cosine of two non-zero vectors can be derived by using the Euclidean dot product formula:

$$A \cdot B = \|A\| \|B\| \cos \theta$$
 (5)

Given two vectors of attributes, A and B, the cosine similarity, $\cos(\theta)$, is represented using a dot product and magnitude as

$$\text{Similarity} = \cos \theta = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}}$$
 (6)

where A_i and B_i are components of vectors A and B respectively [33] [47].

The resulting similarity ranges from -1 meaning exactly opposite, to 1 meaning the same, with 0 indicating orthogonality or decorrelation, while in-between values indicate intermediate Similarity or dissimilarity.

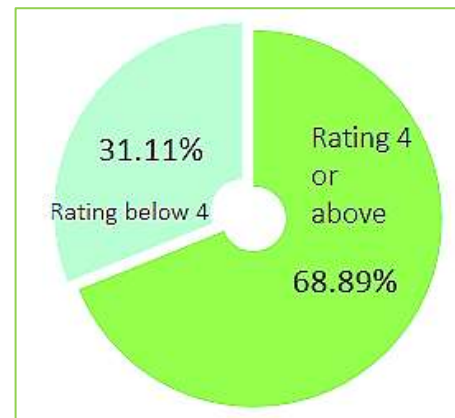


Fig. 2. The Pie Chart of Rating Dataset.

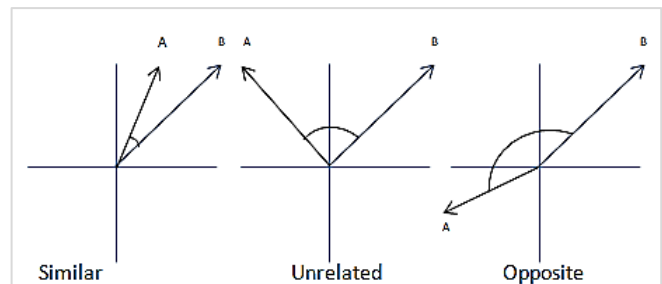


Fig. 3. Cosine Similarity and Cosine Distance Functions.

IV. RESULTS AND DISCUSSIONS

Assessment of predictive accuracy for the book recommendation system is a crucial aspect of evaluation. Receiver operation characteristic (ROC) is widely used for evaluating the accuracy of the classifiers [42][43]. Forecasting is an essential part of every financial department, atmospheric science, and machine learning algorithms. ROC curve gives a visual technique to summarize the accuracy of the classifiers. It is widely used in statistical education and training.

A. Binary Predictor

For the predictions, one of the standard techniques used is binary prediction. It contains beneficial building blocks of a ROC curve. Every classification problem has two classes. Each instance (I) belongs to two sets, (P) and (N), of positive and negative labels of class. A classifier instance has four possible types. If the positive instance is being classified correctly, it is considered as True Positive (TP).

On the other hand, it is regarded as a false negative (FN) if it is classified incorrectly. If the negative instance is classified correctly, it is regarded as true negative (TN). Otherwise, it is considered to be false positive (FP) if it is classified incorrectly. Table II shows performance evaluation results for our proposed system before splitting the training dataset. The test contains 1000 tuples where negative and positive tuples are 610 and 390, respectively. The proposed RS correctly identifies 760 tuples and wrongly classifies 240 tuples. The confusion matrix [44] is widely used to measure the performance of classifiers. Table I depicts the confusion matrix for this research.

We found an FR rate (FPR), FN rate (FNR), TN rate (TNR) or specificity, precision (P), recall (R), and F1 Score by using the following equations:

$$FPR = FP / (TN + FP) \quad (7)$$

$$FNR = FN / (TP + FN) \quad (8)$$

$$\text{Precision (P)} = TP / (TP + TN) \quad (9)$$

$$\text{Specificity or TNR} = TN / (FP + FTN) \quad (10)$$

$$\text{Sensitivity or Recall (R)} = TP / (TP + FN) \quad (11)$$

$$F1 \text{ Score} = 2 * (R * P) / (R + P) \quad (12)$$

We extend this definition to include sensitivity =1-FPR and specificity =1-FNR. Sensitivity is known as the true positive rate, and specificity is termed as the true negative rate.

Table III shows Sensitivity, Specificity, F1 Score for the classifier. Sensitivity calculates the proportion of desired books for a user. Specificity calculates the proportion of boring books for an individual user. F1 Score calculates the harmonic mean of the desired and boring books that are correctly identified. The maximum values of the F1 Score can be 1. Table III shows that the highest sensitivity, Specificity,

and F1 Score are 73.14%, 74.28%, and 74.18%. The sensitivity in dataset-1 is higher than other datasets, which means that the prediction probability was high for an exciting book list. Specificity is 65% for dataset -6, which can detect boring books for a reader. F-score is more useful than accuracy. It finds harmonious relation between sensitivity and specificity.

B. ROC Curve

A receiver operating characteristic curve illustrates the trade-off between the five different datasets' sensitivity and specificity in Table III. It can be inferred from Fig. 4; all of our datasets have stayed close to the ideal diagonal line.

Table IV shows Sensitivity, Specificity, F1 Score for the classifier. The sensitivity in dataset-1 is higher than other datasets, which means that the prediction probability was high for an exciting book list. Specificity is 65% for dataset -6, which can detect boring books for a reader. F-score is more useful than accuracy. It finds harmonious relation between sensitivity and specificity.

C. ROC Curve

Fig. 5 presents a ROC curve that was plotted for sensitivity and specificity. Most of the datasets were closed to the diagonal ideal classifier line. None of the datasets crossed the worst classifier line.

D. User Interface

Fig. 6 shows the user interface for the proposed system. The input searching item was 'Sense and Sensibility,' a popular romantic and narrative book. As a result, the system showed all the similar books categorized into the romantic and narrative class.

TABLE II. CONFUSION MATRIX

Hypothesized Class	True Class		
	Positive(P)	Negative (N)	Total
True (T)	490	120	610
False (F)	120	270	390
Total	610	390	1000

TABLE III. SENSITIVITY, SPECIFICITY, AND F1 SCORE FOR 5 DIFFERENT DATASETS

Datasets	Sensitivity (%)	Specificity (%)	F1 Score (%)
1	68.25	67.15	70.57
2	70.25	71.15	73.55
3	48	58	52.52
4	73.14	74.28	74.18
5	55	60	57.39
Average	62.928	66.116	65.642

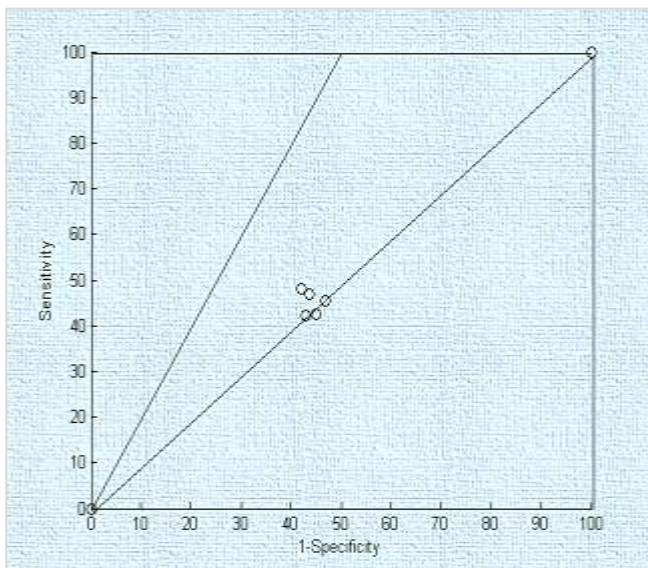


Fig. 4. ROC Curve for Five different Datasets.

TABLE IV. SENSITIVITY, SPECIFICITY, AND F1 SCORE FOR DIFFERENT DATASETS

Datasets	Sensitivity (%)	Specificity (%)	F1 Score (%)
Dataset 1	66	65.15	65.57225
Dataset 2	47	56.25	51.21065
Dataset 3	48	58	52.5283
Dataset 4	45.47	53	48.94709
Dataset 5	55	60	57.3913
Dataset 6	53.5	65	58.69198
Dataset 7	55.5	50	52.60664
Dataset 8	42.17	57	48.47615
Dataset 9	42.47	54.98	47.92202
Dataset 10	42.5	48	45.08287
Average	49.761	56.738	52.84293

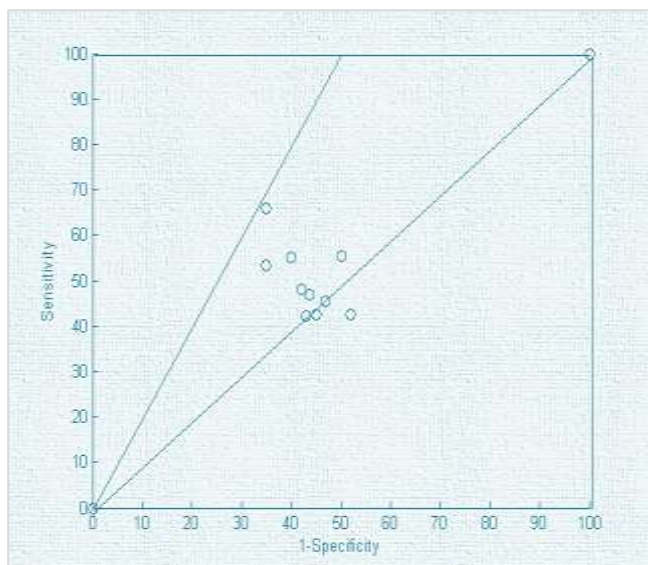


Fig. 5. ROC Curve for Sensitivity vs 1-Specificity (Different Datasets).

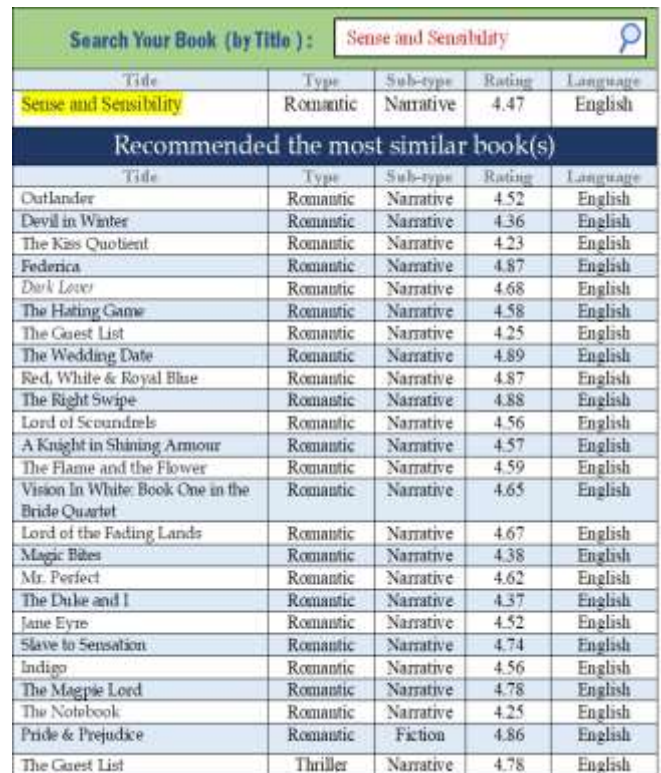


Fig. 6. User Interface for Book Recommendation System.

V. CONCLUSION

This research used clustering algorithms to increase the prediction capacity of the recommendation system. The datasets were collected from the Goodreads-books repository of Kaggle. About 900k ratings of 10k books were processed by using machine learning algorithms (k-means clustering and cosine function). Sensitivity, Specificity, and F1 Score were measured for the algorithms for the proposed model. The average sensitivity and average specificity were 49.76% and 56.74% respectively whereas the F1 Score was 52.84%. These results show that our proposed system can remove boring books from the recommendation list more efficiently. Finally, the ROC curve was plotted for sensitivity and specificity which shows that most of the datasets stay close to the diagonal ideal classifier line.

In our future work, we shall propose a suggestion system for recommending online courses using the convolutional neural network (CNN).

REFERENCES

- [1] N. Grover, "Enabling shift in retail using data: Case of Amazon," 2019.
- [2] Y. Lee, C. Wei, P. Hu, T. Cheng, and C. Lan, "Small Clues Tell: a Collaborative Expansion Approach for Effective Content-Based Recommendations," Journal of Organizational Computing and Electronic Commerce, pp. 1-18, 2020.
- [3] T. Zhao, W. Zhang, Y. Zhang, Z. Liu, and X. Chen, "Significant spatial patterns from the GCM seasonal forecasts of global precipitation," Hydrology and Earth System Sciences, vol. 24, no. 1, pp. 1-16, 2020.
- [4] A. Gazdar and L. Hidri, "A new similarity measure for collaborative filtering based recommender systems," Knowledge-Based Systems, vol. 188, p. 105058, 2020.
- [5] C. Zou, D. Zhang, J. Wan, M. M. Hassan and J. Lloret, "Using Concept Lattice for Personalized Recommendation System Design," in IEEE

- Systems Journal, vol. 11, no. 1, pp. 305-314, March 2017, doi: 10.1109/JSYST.2015.2457244.
- [6] K. Lu, "Research on Intelligent Detection System for Information Abnormal Defect Based on Personalized Recommendation of E-Book," in The International Conference on Cyber Security Intelligence and Analytics, 2019: Springer, pp. 1110-1117.
- [7] T. Y. Tang, P. Winoto and R. Z. Ye, "Analysis of a multi-domain recommender system," The 3rd International Conference on Data Mining and Intelligent Information Technology Applications, Macao, 2011, pp. 280-285.
- [8] S. S. Sohail, J. Siddiqui and R. Ali, "Book recommendation system using opinion mining technique," 2013 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Mysore, 2013, pp. 1609-1614, doi: 10.1109/ICACCI.2013.6637421.
- [9] A. S. Tewari, A. Kumar and A. G. Barman, "Book recommendation system based on combine features of content based filtering, collaborative filtering and association rule mining," 2014 IEEE International Advance Computing Conference (IACC), Gurgaon, 2014, pp. 500-503, doi: 10.1109/IADCC.2014.6779375.
- [10] P. Mathew, B. Kuriakose and V. Hegde, "Book Recommendation System through content based and collaborative filtering method," 2016 International Conference on Data Mining and Advanced Computing (SAPIENCE), Ernakulam, 2016, pp. 47-52, doi: 10.1109/SAPIENCE.2016.7684166.
- [11] S. Kanetkar, A. Nayak, S. Swamy and G. Bhatia, "Web-based personalized hybrid book recommendation system," 2014 International Conference on Advances in Engineering & Technology Research (ICAETR - 2014), Unnao, 2014, pp. 1-5, doi: 10.1109/ICAETR.2014.7012952.
- [12] Chih-Ming Chen "An intelligent mobile location-aware book recommendation system that enhances problem-based learning in libraries, Interactive Learning Environments", Journal Interactive Learning Environments, 2013, pp 469-495, DOI: 10.1080/10494820.2011.593525.
- [13] S. S. Sohail, J. Siddiqui, R. Ali, "OWA based Book Recommendation Technique", Procedia Computer Science, 2015, pp. 126-133, doi: 10.1016/j.procs.2015.08.425.
- [14] J. Qi, S. Liu, Y. Song and X. Liu, "Research on Personalized Book Recommendation Model for New Readers," 2018 3rd International Conference on Information Systems Engineering (ICISE), Shanghai, China, 2018, pp. 78-81, doi: 10.1109/ICISE.2018.00022.
- [15] M. Kommineni, P. Alekhya, T. M. Vyshnavi, V. Aparna, K. Swetha and V. Mounika, "Machine Learning based Efficient Recommendation System for Book Selection using User based Collaborative Filtering Algorithm," 2020 Fourth International Conference on Inventive Systems and Control (ICISC), Coimbatore, India, 2020, pp. 66-71, doi: 10.1109/ICISC47916.2020.9171222.
- [16] K. Puritat and K. Intawong, "Development of an Open Source Automated Library System with Book Recommendation System for Small Libraries," 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), Pattaya, Thailand, 2020, pp. 128-132, doi: 10.1109/ECTIDAMTCON48261.2020.9090753.
- [17] R. Rahutomo, A. S. Perbangsa, H. Soeparno and B. Pardamean, "Embedding Model Design for Producing Book Recommendation," 2019 International Conference on Information Management and Technology (ICIMTech), Jakarta/Bali, Indonesia, 2019, pp. 537-541, doi: 10.1109/ICIMTech.2019.8843769.
- [18] C. Sipio, D. Ruscio, and P. Nguyen, "Democratizing the Development of Recommender Systems by means of Low-code Platforms."
- [19] Y. Pan, D. Wu, C. Luo, and A. Dolgui, "User activity measurement in rating-based online-to-offline (O2O) service recommendation," Information Sciences, vol. 479, pp. 180-196, 2019.
- [20] S. B. Shirude and S. R. Kolhe, "Improved Hybrid Approach of Filtering Using Classified Library Resources in Recommender System," in Intelligent Computing Paradigm: Recent Trends: Springer, 2020, pp. 1-10.
- [21] J. Das, S. Majumder, P. Gupta, and K. Mali, "Collaborative Recommendations using Hierarchical Clustering based on Kd Trees and Quadrees," International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, vol. 27, no. 04, pp. 637-668, 2019.
- [22] S. Parvatikar and B. Joshi, "Online book recommendation system by using collaborative filtering and association mining," 2015 IEEE International Conference on Computational Intelligence and Computing Research (ICCCIC), Madurai, 2015, pp. 1-4, doi: 10.1109/ICCCIC.2015.7435717.
- [23] S. Kanetkar, A. Nayak, S. Swamy, and G. Bhatia, "Web-based personalized hybrid book recommendation system," in 2014 International Conference on Advances in Engineering & Technology Research (ICAETR-2014), 2014: IEEE, pp. 1-5.
- [24] A. Gholami, Y. Forghani, and M. Branch, "Improving Multi-class Co-Clustering-Based Collaborative Recommendation Using Item Tags Improving Multi-class Co-Clustering-Based Collaborative Recommendation Using Item Tags."
- [25] E. Cho and M. Han, "AI Powered Book Recommendation System," in Proceedings of the 2019 ACM Southeast Conference, 2019, pp. 230-232.
- [26] F. Eskandarian, N. Sonboli, and B. Mobasher, "Power of the Few: Analyzing the Impact of Influential Users in Collaborative Recommender Systems," in Proceedings of the 27th ACM Conference on User Modeling, Adaptation and Personalization, 2019, pp. 225-233.
- [27] Y. Gao, C. Huang, M. Hu, J. Feng, and X. Yang, "Research on Book Personalized Recommendation Method Based on Collaborative Filtering Algorithm," in IOP Conference Series: Earth and Environmental Science, 2019, vol. 252, no. 5: IOP Publishing, p. 052099.
- [28] G. Hu, Y. Zhang, and Q. Yang, "Transfer meets hybrid: a synthetic approach for cross-domain collaborative filtering with text," in The World Wide Web Conference, 2019, pp. 2822-2829.
- [29] R. Kumar, G. Bhanodai, and R. Pamula, "Book search using social information, user profiles and query expansion with Pseudo Relevance Feedback," Applied Intelligence, vol. 49, no. 6, pp. 2178-2200, 2019.
- [30] Q. Liu, X. Zhang, L. Zhang, and Y. Zhao, "The interaction effects of information cascades, word of mouth and recommendation systems on online reading behavior: An empirical investigation," Electronic Commerce Research, vol. 19, no. 3, pp. 521-547, 2019.
- [31] C. Musto, P. Lops, M. de Gemmis, and G. Semeraro, "Justifying recommendations through aspect-based sentiment analysis of users reviews," in Proceedings of the 27th ACM Conference on User Modeling, Adaptation and Personalization, 2019, pp. 4-12.
- [32] H. Pasricha and S. Solanki, "A New Approach for Book Recommendation Using Opinion Leader Mining," in Emerging Research in Electronics, Computer Science and Technology: Springer, 2019, pp. 501-515.
- [33] S. Vimala and K. Vivekanandan, "A Kullback-Leibler divergence-based fuzzy C-means clustering for enhancing the potential of an movie recommendation system," SN Applied Sciences, vol. 1, no. 7, p. 698, 2019.
- [34] G. Wen and C. Li, "Research on Hybrid Recommendation Model Based on PersonRank Algorithm and TensorFlow Platform," in Journal of Physics: Conference Series, 2019, vol. 1187, no. 4: IOP Publishing, p. 042086.
- [35] B. Ye, Y. Tu, and T. Liang, "A hybrid system for personalized content recommendation," Journal of Electronic Commerce Research, vol. 20, no. 2, pp. 91-104, 2019.
- [36] Iqbal H. Sarker, "Context-aware rule learning from smartphone data: survey, challenges and future directions," vol. 6, no. 1, p. 95, 2019.
- [37] I. H. Sarker, A. Colman, and J. J. J. o. B. D. Han, "Recencyminer: mining recency-based personalized behavior from contextual smartphone data," vol. 6, no. 1, p. 49, 2019.
- [38] S. Hossain, D. Sarma, and R. J. J. M. L. Chakma, "Machine Learning-Based Phishing Attack Detection," International Journal of Advanced Computer Science and Applications(IJACSA), vol. 11, no. 9, 2020.
- [39] S. Hossain, D. Sarma, F. Tuj-Johora, J. Bushra, S. Sen and M. Taher, "A Belief Rule Based Expert System to Predict Student Performance under Uncertainty," 2019 22nd International Conference on Computer and

- Information Technology (ICCIT), Dhaka, Bangladesh, 2019, pp. 1-6, doi: 10.1109/ICCIT48885.2019.9038564.
- [40] F. Ahmed, et al., "A Combined Belief Rule based Expert System to Predict Coronary Artery Disease," in 2020 International Conference on Inventive Computation Technologies (ICICT), 2020, pp. 252-257.
- [41] S. Hossain, D. Sarma, R. J. Chakma, W. Alam, M. M. Hoque, and I. H. Sarker, "A Rule-Based Expert System to Assess Coronary Artery Disease Under Uncertainty," in Computing Science, Communication and Security, Singapore, 2020, pp. 143-159: Springer Singapore.
- [42] Y. Nieto, V. Gacía-Díaz, C. Montenegro, C. C. González, and R. G. Crespo, "Usage of machine learning for strategic decision making at higher educational institutions," IEEE Access, vol. 7, pp. 75007-75017, 2019.
- [43] J. Chen, X. Luo, Y. Liu, J. Wang and Y. Ma, "Selective Learning Confusion Class for Text-Based CAPTCHA Recognition," in IEEE Access, vol. 7, pp. 22246-22259, 2019, doi: 10.1109/ACCESS.2019.2899044.
- [44] C. Chang, "Statistical Detection Theory Approach to Hyperspectral Image Classification," in IEEE Transactions on Geoscience and Remote Sensing, vol. 57, no. 4, pp. 2057-2074, April 2019, doi: 10.1109/TGRS.2018.2870980.
- [45] M. N. Alam, D. Sarma, F. F. Lima, I. Saha, R. -E. -. Ulfath and S. Hossain, "Phishing Attacks Detection using Machine Learning Approach," 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 2020, pp. 1173-1179, doi: 10.1109/ICSSIT48917.2020.9214225.
- [46] I. Saha, D. Sarma, R. J. Chakma, M. N. Alam, A. Sultana and S. Hossain, "Phishing Attacks Detection using Deep Learning Approach," 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 2020, pp. 1180-1185, doi: 10.1109/ICSSIT48917.2020.9214132.
- [47] W. Usino et al., "Document similarity detection using k-means and cosine distance," International Journal of Advanced Computer Science and Applications(IJACSA), vol. 10, no. 2, pp. 165-170, 2019.