Iraqi Journal of Science, 2024, Vol. 65, No.1, pp: 487- 511
DOI: 10.24996/ijs.2024.65.1.39

ISSN: 0067-2904

# A Recent Trends in eBooks Recommender Systems: A Comparative Survey

**Abdullah Mohammed Saleh[1]*, Alaa Yaseen Taqa[2]**
[1]Computer Sciences Dept., College of Computer sciences and mathematics, University of Mosul, Nineveh, Iraq
[2]Computer Science Dept., College of Education for Pure Sciences, University of Mosul, Nineveh, Iraq

Received: 6/11/2022 Accepted: 22/2/2023 Published: 30/1/2024

**Abstract**
The great progress in information and communication technology has led to a huge increase in data available. Traditional systems can't keep up with this growth and can't handle this huge amount of data. Recommendation systems are one of the most important areas of research right now because they help people make decisions and find what they want among all this data. This study looked at the research trends published in Google Scholar within the period 2018-2022 related to recommending, reviewing, analysing, and comparing ebooks research papers. At first, the research papers were collected and classified based on the recommendation model used, the year of publication, and then they were compared in terms of techniques, datasets utilised, problems, contributions, and evaluation methods used. It was found that many in-depth studies of book recommendation systems directly affect how those systems grow. Many researchers interested in book recommendation systems can learn about the many parts of the field by looking at how the study was put together.

**Keywords:** Recommender System; Book Recommender system; Recommendation Models; Hybrid system; collaborative filtering; content-based filtering.

# البحوث الحديثة في انظمة توصية الكتب الالكترونية: دراسة مقارنة

**عبدالله محمد صالح[1]*, أ.م.د آلاء ياسين طاقة[2]**
[1]قسم علوم الحاسوب, كلية علوم الحاسوب والرياضيات, جامعة الموصل, نينوى, العراق
[2]قسم علم الحاسوب, كلية التربية للعلوم الصرفة, جامعة الموصل, نينوى, العراق

**الخلاصة**
أدى التقدم الكبير في تكنولوجيا المعلومات والاتصالات إلى زيادة هائلة في البيانات. لذلك لا تستطيع الأنظمة التقليدية مواكبة هذا النمو ولا يمكنها التعامل مع الكم الهائل من البيانات. تعد أنظمة التوصيات من أهم مجالات البحث في الوقت الحالي لأنها تساعد الأشخاص على اتخاذ القرارات والعثور على ما يريدون من بين كل هذه البيانات. نظرت هذه الدراسة في اتجاهات البحث المنشورة في Google Scholar خلال الفترة 2018-2022 المتعلقة بالتوصية بالكتب الإلكترونية ومراجعتها وتحليلها ومقارنتها. في البداية تم جمع الأوراق البحثية وتصنيفها بناءً على نموذج التوصية المستخدم ثم سنة النشر ، ثم تمت مقارنتها من حيث الأساليب ومجموعات البيانات المستخدمة والمشكلات والمساهمات وطرق التقييم المستخدمة. بعد ذلك، وجد أن العديد من الدراسات المتعمقة لأنظمة التوصية بالكتب لها تأثير مباشر على كيفية نمو هذه الأنظمة. يمكن للعديد من الباحثين

*Email: abdullah.20csp39@student.uomosul.edu.iq

المهتمين بأنظمة التوصية بالكتب التعرف على العديد من أجزاء المجال من خلال النظر في كيفية تجميع الدراسة.

## 1. Introduction

With increasing technological advancements, tons of data is available on the internet nowadays, making it completely tiresome for users to browse for the products they want. In addition, it has become difficult for digital service providers to engage multiple users for the maximum possible time on their applications. This is where the Recommender System (RMs) appears. RMs recommends content or various data types in accordance with the user's past actions and interactions with the system. Most internet users have seen RMs. Facebook recommends prospective friends, YouTube videos, Glassdoor jobs, TripAdvisor vacation spots, and Goodreads books.

RMs are popular in e-business[1]. E-commerce portals (e.g., eBay, Amazon) use RMs to entice customers with products they might like[2]. The system suggests to users the material of their choice and liking based on a vast set of goods and a description of their wants. Such systems help users interact better with the application and thus increase the amount of time spent on that application[3]. Bad decisions might lead to wasting of time and money. Traditionally, people have used various strategies to solve such problems, for example surfing the internet, taking suggestions from friends, or simply following others[4].

To put it another way, a RMs is a program that shows the most relevant items, products, or services to specific users by anticipating the items they are most interested in based on their past and current interactions with the system and other users[5]. Taking an appropriate decision within a constrained environment was more difficult because only a finite number of points of view could be considered. Thus, a systematic statistical approach is needed to analyse such a large data volume and extract only the most relevant information for the end user[6]. As a result, the internet and smart devices have evolved into environments where various types of user data can be collected.

RMs can use more than only user-provided data, such as their likes and ratings. They can also use information about the user's behaviour patterns, such as their visit logs, to make recommendations. Recently, researchers have employed implicit data in RMs to assess users' personalities or behaviour to predict their preferences [1] [7]. In this paper, we discuss popular ebook recommenders. These recommenders are helpful in libraries and schools. With the availability of ebooks on online learning platforms, readers can now access the resources they require at a lower cost and with less effort. Figure 1 describes the organisation of this paper.

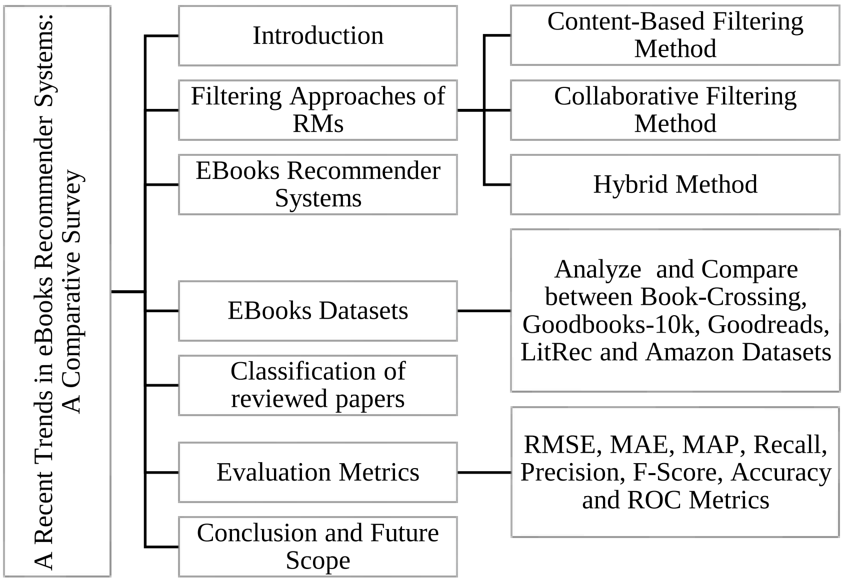



**Figure 1:** The Organization of This research paper

The rest of this review paper is structured as follows; in section 2, the main filtering approaches is explained, and the Content-Based Filtering (CBF), Collaborative Filtering (CF), and Hybrid methods are discussed respectively in 2.1, 2.2, and 2.3. The most popular fields of recommendation systems are discussed in section 3 to identify the eBooks recommendation systems for study in this research. Section 4 describes the essential datasets used in this field and compares them. In section 5, we do a literature survey on eBook RMs, and the evaluation metrics are described in section 6. Finally, section 7 concludes of this paper.

## 2. Filtering Approaches of RMs

RM models start with two types of data: user-item interactions, like ratings or buying habits, and attribute information about users and items, like textual profiles or descriptions full of keywords, are examples of user-item interactions. Also, most CBF use rating matrices[8]. The recommendations in knowledge based RMs are based on explicitly expressed user requirements. External RMs combine these many aspects to create hybrid systems rather than using rating or purchasing data from the past. Figure 2 shows how the different essential techniques for RMs are put into groups. These methods are often used to build recommender systems and have worked well in many situations. Innovative strategies can be developed by merging the advantages of several different RMs to produce hybrid systems [9][10].

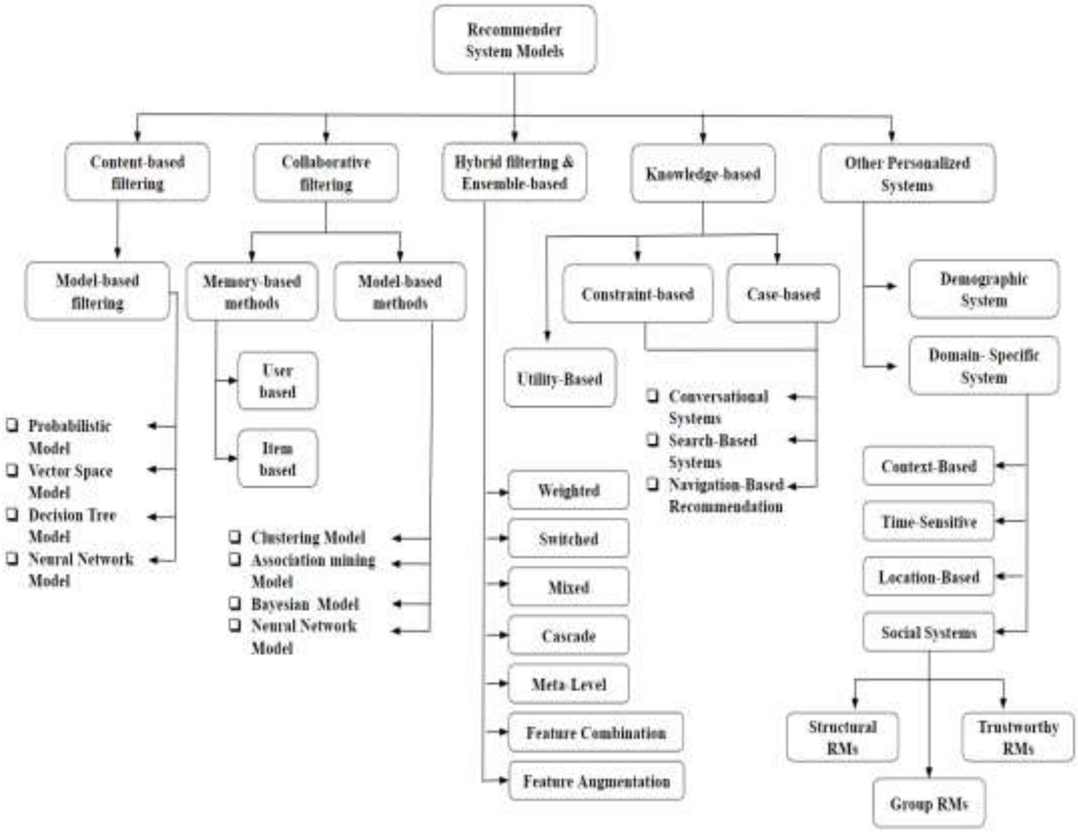



**Figure 2 :** Taxonomy of Recommendation Models [11][12]

### 2.1 Content-Based Filtering Method

Content-based methods illustrated in Figure 3 only analyzes the items and user profiles for a recommendation. It recommends items based on the user's browsing history, number of clicks, and viewed products. This approach can propose unrated items based on the user's rating; however, it does not function for new users who have not yet rated anything. In a CB approach, there is no recommendation of items that are unexpected to a user (serendipitous items), and it will not work if the system fails to distinguish the content that the user does not like[13] [9].

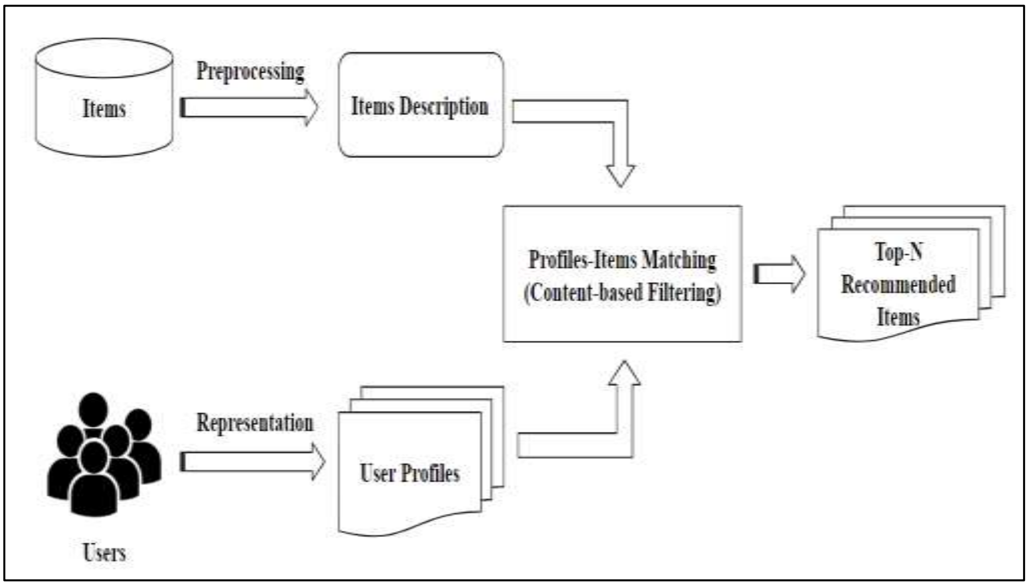



**Figure 3 :** Content-Based Filtering [14]

## 2.2 Collaborative Filtering Method

CF methods illustrated in Figure 4 (Inspired from a figure in [15]) work by finding similarities between different users and recommending their products. Methods of CF can be divided into two broad categories: Model Based and Memory Based. The Memory Based Approach works based on the user or item user-based identifying the nearest neighbors of the target user depending on the similarity of the training users and the target user. In other means, people who are a lot like the target user. While the item-based methods works by identifying the items similar to the active user's preferred item. Furthermore, generating a final list of recommendations. The model based approach considers the user rating behavior instead of directly using the data. The rating data is used to extract the model parameters, leading to better accuracy and performance[10][11].

The recommendation by the CF depends on the user's behavior and is content-independent. Because suggestions are based on user similarity rather than item similarity, it also gives unexpected recommendations. However, the problem with this approach is that it cannot recommend items to new users (Cold start Problem). This method also finds it challenging to recommend items to those users who have special interests and are different from most people. This is because they may not agree or disagree with the other users creating difficulty in producing relevant results. (Grey Sheep problem) [10].

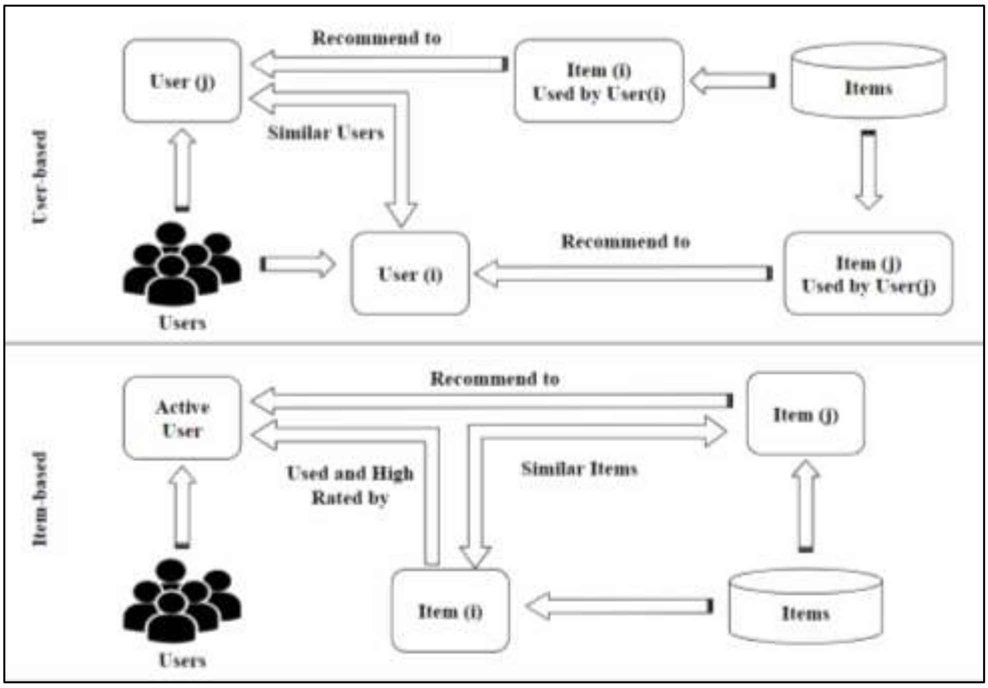



**Figure 4 :** Collaborative Filtering [15]

## 2.3 Hybrid Method

Because each approach mentioned earlier has its advantages and disadvantages, hybrid methods, as shown in Figure 5 (Inspired from a figure in [15]), combine the benefits of different approaches to create a system that performs well in a wide range of applications[16]. It is possible to combine the recommended methodologies (CF and CBF) in a hybrid strategy to receive the most advantage, generate better results, and decrease the risks and challenges connected with these applications[14][17].

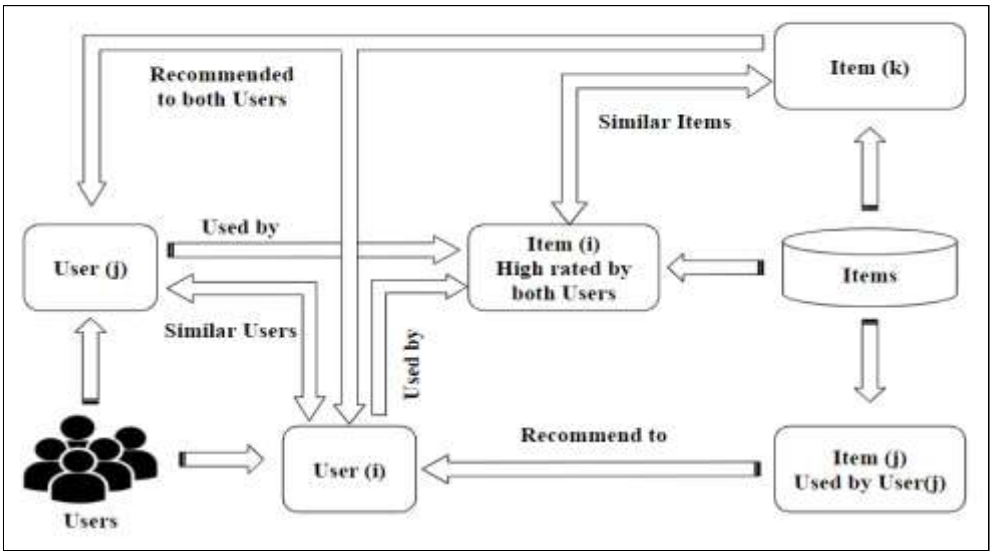



**Figure 5 :** Hybrid Approach [15]

*Hybrid Approaches Have Multi-Methods:*

- *Weighted***:** The system assigns a distinct score to each recommended component when combined numerically[18]. Several models may be developed in the framework of weighted recommendations, all of which can adequately interpret the dataset. The weighted RMs integrate the outputs of each model into the static weight that does not change throughout training and testing sets. For instance, if we combine a CBF model with an item-by-item CF model, each model contributes equally to the final forecast. The weighted hybrid allows us to use multiple linear models to support the recommendation process.
- *Switching***:** The system gives the user more than one option for each recommendation item and chooses the best one based on what the user wants. RMs are selected based on the context in which they are being used[19]. The recommender selector criteria for the item-level sensitive dataset should be determined by the user's profile or other characteristics, such as interests. In a hybrid switching approach, a new layer is added to the recommendation model, determining which model should be used.
- *Mixed:* The system simultaneously informs the user about several distinct topics. User profiles and features are used in the mixed hybrid strategy. To generate a diverse set of datasets, the RMs feed a different set of candidates into the recommendation model under the situation and then combine the prediction with the result to produce the suggestion. Mixture hybrid RMs can produce multiple predictions simultaneously while also fitting a partial dataset to the most suited model for improved performance [10].
- *Feature Combination:* To create RMs features, multiple knowledge sources are combined by incorporating a virtual recommendation model that can be used for feature engineering on the original user profile dataset. The feature combination hybrid improves upon the original system. A collaborative recommendation approach, for example, can be incorporated into a CB recommendation model. The hybrid model can consider collaborative data from the subsystem by relying simply on one model [9].
- *Feature Augmentation:* Feature augmentation is one of the important components required to construct a set of recommender system features. The primary RMs use the rating or classification generated by a contributing RM for the user/item profile to arrive at their predictions. The hybrid feature augmentation model can boost core system performance without making any changes to the primary recommendation model. Using the association rule, may enrich the user profile dataset, for instance. The increased dataset will enhance the CB recommendation model's performance [11].

- *Cascade:* The recommender's list includes a weighted priority item with the highest rating appearing first, followed by items with lower ratings descending. Cascade hybrid establishes a rigorous hierarchical framework for RMs, in which the primary RMs create the immediate output. The second model is used to correct minor flaws in the preliminary outcome, such as breaking a tie in the scoring. Since most datasets are sparse, the secondary recommendation model can successfully address the difficulties, including equal scoring or missing data [10].
- *Meta-level:* This is one of the methods used to make a model for the next step of the recommender system's algorithm. The feature augmentation hybrid is like the meta-level hybrid in that the contributing model gives the primary recommendation model an expanded dataset. In contrast to feature augmentation, the meta-level utilizes a learned model from the contributing model as the input to the primary recommendation model. When these methods are used together, problems and challenges are solved, and performance is improved. When only content-based or CF is used, performance is not good [9].

## 3. Recommendation System Fields

More and more service sectors are implementing the RMs. These models and technology are examined to see how they may be applied to the specifics of the service industry in this research. The service fields where RMs were utilized are categorized as follows: Streaming Service (video and audio), E-Commerce Service (eBook, Advertisement, and video game), Healthcare Service (Food), Education Service (library), News (News articles), and Trust-based (trust-based, reputation-based, and intrusion detection). [9]. In this study, the focus will be on research trends related to recommending eBooks.

### 3.1 EBooks Recommender Systems

As the process of informatization continues to grow, people from all walks of life are making changes to better use of this technology. In this situation, the process of digitizing information about library management is getting better over time [20]. Digital libraries are very popular with readers because they make it easy and quick to find documents, make personalized suggestions, and offer other unique services. With so many bibliographies, it's hard for readers to find interesting books in a short amount of time. So, the traditional way of borrowing from a library is bad for the people who try to use it [21].

The RMs forecasts whether a potential customer will be interested in a product they are not yet informed about. In general, for an RMs to make a recommendation, it needs user data, items, and user ratings on those items. After making a suggestion, either explicitly or implicitly, user feedback on the item is obtained and used to make future recommendations [9].

## 4. EBooks Datasets

Many academics are unaware of the numerous accessible datasets and APIs available for a book recommendation. Several datasets can be used for evaluation due to the availability of attributes. This can be seen in Table 1, which is inspired by the "features of BRMs datasets" from a table in [2], the datasets that include user ratings, making them appropriate for CF. The demographic properties of context aware RMs can be used to evaluate them. However, a book's metadata, summary, and complete text are helpful for CB assessment [2]. You can learn more about your readers' tastes by performing syntactic and semantic analysis on the full text of a book, such as topic modelling or genre identification. Tag-based can benefit from user-generated tags. A recommendation system's (RS) capacity to meet users' demands can be evaluated by the data they provide through its requests for recommendations. The following is a list of datasets with brief descriptions [22].

1- *Book-Crossing dataset:* The CF literature frequently makes use of this dataset. It was retrieved in 2004 in under four weeks via the Book-Crossing website. A total of 278,858 users have contributed 1,149,780 ratings for 271,379 books. Titles, authors, publishers, and covers is the metadata that was added for specific publications. The scale goes from 1-10. (10 is the highest). Researchers have the option of expanding the dataset by including book details (e.g., book summaries and reviews from other Websites) [2][23].

2- *Good books-10k:* It was originally scraped from the Goodreads API in September 2017 by Zygmunt Zając. Additional fields are included in the books_enriched.csv file. The biggest advantage of this new version is that it adds a text description field for the 9943 books. The dataset contains six million numerical ratings of the platform's ten thousand most popular books, with data collected from 53,424 users [2][23].

3- *Goodreads:* According to the website, there are 876,145 subscribers, 2,360,655 books, and 112,131,203 reviews. It offers demographics, tags, reviews, friend lists, reading groups, and user quotes. This approach can also obtain all the book's metadata, such as ratings and reviews. Author information can be obtained via the API user [2][24].

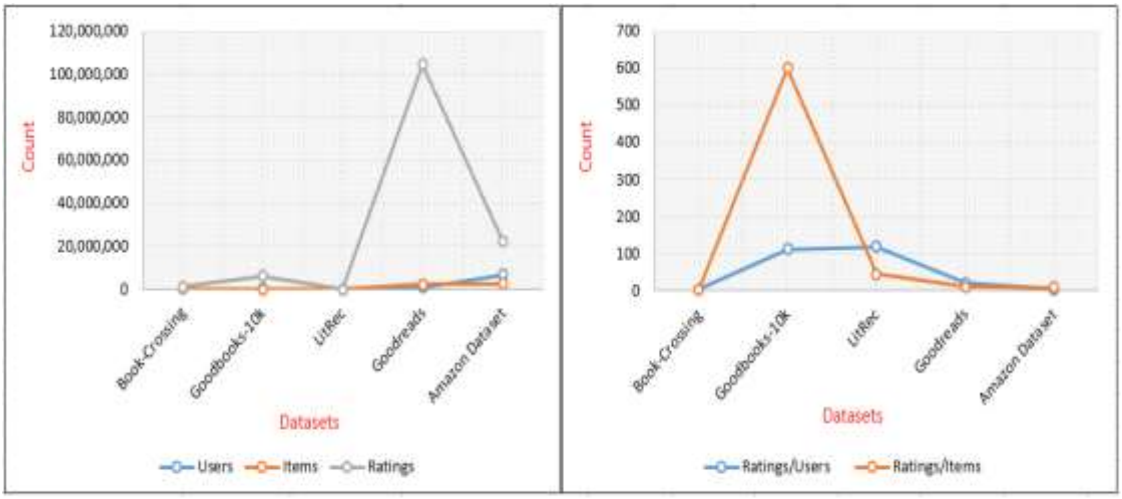
4- *LitRec Goodreads* and Project Gutenberg data are also included in the dataset. Book reviews, author biographies, and the dates on which a book was added rated and read can all be found on Goodreads. Here you will find a book's star rating, synopsis, and full text, all marked up with part-of-speech tags [2][22].

5- *Amazon Dataset* spanning 18 years is publicly available. The dataset includes book reviews, book information, users, and ratings. Each review is timed and graded for utility. If you want to incorporate additional information about books but not users, this set, like Book-Crossing, may be easily expanded [2][25].

**Table 1:** Description of eBook-Recommendation Datasets

| *Features/ Datasets* | *Book-Crossing* | *Goodbooks-10k* | *Goodreads* | *LitRec* | *Amazon Dataset* |
|---|---|---|---|---|---|
| ***Rating range*** | 1-10 | 1-5 | 1-5 | 1-5 | 1-5 |
| ***Demographics*** | Locations and ages | N/A | N/A | Locations | Amazon user-id |
| ***Metadata*** | Title, authors, year, publisher, and image of the cover | bestbookid , goodreadsbookid, books_count | number of ratings, reviews received, and average rating | title, authors | Amazon book id, title, price |
| ***Description*** | N/A | Involved in 9943 book | N/A | N/A | Involved |
| ***Users*** | 278,858 | 53,424 | 876,145 | 1,927 | 6,643,669 |
| ***Items*** | 271,379 | 10 000 | 2,360,655 | 3,710 | 2,441,053 |
| ***Ratings*** | 1,149, 780 | 6 million | 112,131,203 reads and 104,551,549 ratings | 38,591 | (22,507,155 ratings) (8,898,041 reviews) |
| ***Ratings/Users*** | 4.123 | 112.309 | 119.33 | 20.026 | 3.387 |
| ***Ratings/Items*** | 4.236 | 600 | 44.28 | 10.401 | 9.220 |

In this study, the most well-known data sets accepted by most researchers in the field of recommending ebooks were considered, as they were analyzed and compared in the previous Table. The primary dataset features are depicted in Figure 6-(a), along with their relative proximity. For example, the users and item features in Goodreads are close to those in Amazon, whereas, in bookcrossing, the average data is suitable for most recommender systems. LitRec is used in models based on simple Artificial Intelligence (AI) techniques such as K-Nearest Neighbor(KNN). Large ratings values are used to determine user preferences, especially in collaboration-based recommendation systems. Figure 6-(b) shows the ratio of rating values to users/items. Figure 7 shows how the reviewed publications are categorized using the dataset.




**(a)** **(b)**

**Figure 6:** This figure describes datasets (a) Statistical number of features (b) Percentage of Ratings to Users/Items for Given Datasets.

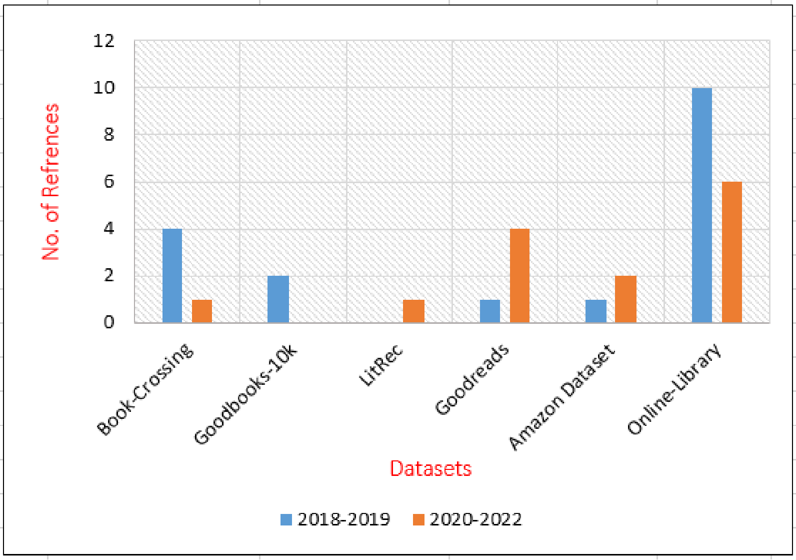



**Figure 7:** Number of papers using datasets during the period (2018-2022).

## 5. Literature Survey

A considerable amount of work has been done in the area of BRMs using various publishing forums. This section provides an extensive survey of the various approaches (traditional and advanced) to support future research in this area. The research papers adopted in this study are classified in Figure 8 based on the recommendation model used and then on the year of publication from 2018 to 2022.

CBF is the simplest model of recommendation. Due to the disadvantage of suggesting only biased items (the problem of over-specialization), the number of studies employing this approach has dropped progressively in recent years. However, it continues to be explored and applied in the disciplines of books and news, which are text-centered application areas. CF is the most popular and extensively researched filter in recommendation models, accounting for the biggest proportion of published works. Although CF has limits, researchers have continued to examine them. However, recently, the use of hybrid models has been made in the RMs that combine the advantages of each method to solve the problems of each other, taking into account the complexity of the resulting model.

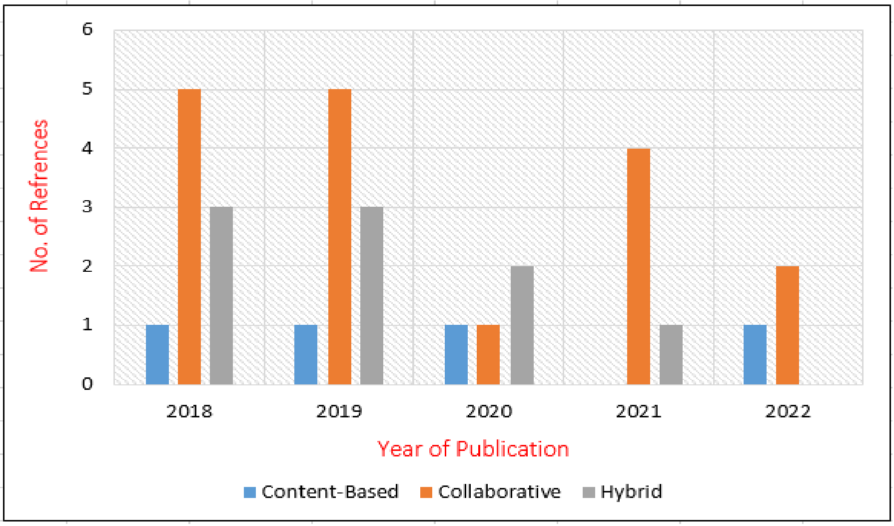



**Figure 8:** Papers based on year of publication (2018-2022) and recommendation models.

### 5.1 Research Trends in Ebooks Content-Based Filtering

Edward Rolando et al. [26] described RMs for eBooks that are based on the implicit feedback that is derived from the user's engagement with electronic content, and evaluated the quality of the output supplied by the machine learning algorithms (IBk, K*, and random forests) that are used to create these suggestions.

Reza Rahutomo et al.[27], generated a recommendation-based embedding model using explicit feedback to increase the accuracy of the recommendation. The proposed model was trained to comprehend each user's pattern of highly rated books to compute the desired books as precisely as possible. Corporate learning and development at Bonus University keeps track of this research to utilize it in Beelajar, an internal online course platform.

Yiu-Kai Ng [28] created a web application that suggests reading books for kids. They combined the matrix factorization method with the content-based method to address the cold-start issue. This model made some grade-level predictions on the books too. In addition, Tulasi Prasad Sariki and G Bharadwaja Kumar [29] developed a new model to improve the recommendations generated in the book domain by using a judicious combination of the Natural Language Processing and Deep Learning techniques. The current study offered a three-module system to improve the suggestion process. The Named Entities-Based Recommender (NER) module retrieves the critical semantic units from the book material. The visual feature extraction module examines the front cover of the book to identify items and text thereon, as well as the cover's description. The Stylometry module provides an additional feature set, strengthening features used in the literature to identify comparable writers. Over the baseline CBF approach, the proposed three modules enhanced the total suggestion accuracy by 18 percent. Table 2 compares recently published research that relies on content-based filtering models.

**Table 2 :** Research Trends in Content-based filtering for eBooks RMs

| Ref. | Year | Techniques / Similarity Methods | Dataset | Findings / Contribution | Problems | Evaluation Metrics |
|---|---|---|---|---|---|---|
| [26] | 2018 | CART; multilayer perceptrons; SVR..etc. And NLP | Sample of Real data | Analysis of the state-of-the-art in relation to widely-used algorithm families (tree-based, function-based, rule-based) | Using NLP to classify the comments and other actions as positive, negative, and neutral. | Mean Absolute Error (MAE) |
| [27] | 2019 | Word embedding model, PCA, and ANN | Goodbooks-10k | Using an improved method called Collaborative Knowledge Base Embedding, the performance of embedding models is drastically different. | Cold start problem, no serendipitous recom-mendations | Adam and Mean Squared Error (MSE) |
| [28] | 2020 | MF, content-based approach, ReLAT Tool | book-crossing dataset, 30% of the children's books | Combined the CBF method with matrix factorization to address the issue of information overload | serendipitous, scalability and complexity | Root Mean Squared Error (RMSE), MAE |
| [29] | 2022 | SVD, XGB, K-NN, Decision Tree, Linear Discriminant Analysis, Gaussian NB and Random-Forest | LitRec dataset and goodreads.com | The aggrandized framework can improve the overall recommendation accuracy by 18% than the baseline models | New Item Cold start problem, no serendipity, and increased complexity | MAE, RMSE, Mean Absolute Precision (MAP), and Mean F1-Score |

**5.2 Research Trends in eBooks Collaborative Filtering**

Chaloemphon Sirikayon et al. [30] conducted a recommendation experiment was using CF for university students. Each student's book recommendation was generated using borrowing records with a time stamp. Data sparsity and high dimensionality were solved using matrix factorization. According to the results, these book recommendations were accurate enough to help the library increase book use. Nevertheless, it suffers from new user cold-start that can be addressed using CF and CBF techniques. To increase user satisfaction, Chaloemphon Sirikayon [31] calculated student ratings from the records of students who borrow and return books with time stamps. It appears that the system's recommendation of a book was accurate and that students were satisfied with it.

To improve prediction accuracy in future studies, researchers used both book bibliographies and student profiles like faculty members, academic year, and significance. In both Jiayun Wang et al. [32] and Yongen Liang et al.[33] Records of books borrowed and returned with time stamps were used to calculate student ratings. According to the system's recommendations, students were satisfied with the book. Prediction accuracy will be improved in future studies by using both book bibliography and student profile information such as professors, academic year, and majors. Rohit et al.[34] provide a solution to new user problems based on three separated user parameters (age, location, and interest); the anticipated user ratings are utilized to propose three different models.

In addition, to help students find online courses that will meet their needs, Raghad Obeidat et al. [35] implemented a system that analyses their prior academic performance and makes recommendations. Clustering datasets has a profound effect. Improvements in the RMs are achieved through improved performance and the selection of a high level of coverage. Kaivan Shah [36] implements an item-based CF approach on the "goodbooks10k" dataset found on Kaggle. The paper discusses various methods for building a recommender system.

Noor Ifada et al. [37] and Zhi Hui Wang and De Zhi Hou [21] used cosine similarity to implement CF model. However, the first solves the sparsity issue in a library book RMs by developing a probabilistic-keyword CF method. Both book circulation records and keyword data are considered. Based on the user's keyword model, it uses a probabilistic method to predict the list of books recommended to the user.

In Thi Thanh Sang Nguyen [38], Naive Bayes for book recommendation was implemented with acceptable runtime and accuracy. For classifier models, numeric and string types are inefficient. The word embedding method can be used to represent book titles better. Search engines, digital libraries, and e-commerce sites that sell books all need book RMs. Avi Rana and K. Deeba [39] proposed a recommendation that utilizes Jaccard similarity to give more accurate recommendations by utilizing CF. Compact datasets proved to be more accurate than complete datasets in the proposed algorithm.

Missi Hikmatyar and Ruuhwan [40] RMs are built using CF that uses centered cosine similarity and the number of KNN to create a data matrix that will be used to calculate an algorithm. While M. Fatih Adak and Metehan Uçar [41] designed an example of RMs tailored to customers of the Amazon online bookstore. Due to the enormous number of characteristics involved with the data acquired from the data sets, an application based on fuzzy models could be developed. It was realized that the parameter "number of pages" played a pivotal role in deciding which rules to extract from the decision tree model.

Addanki Mounika and S. Saraswathi [42] utilized sentiment analysis for recommending books to the target user based on the reviewers' clustered data to show the finding of books in the RMs. It is all about providing the most excellent books possible to users and proposes an approach that will enhance accuracy to address the shortcomings of current RMs. In addition, Dhiman Sarma and Tanni Mittra et al. [43] clustering algorithms were used to improve the RMs prediction capacity. The datasets were obtained from Kaggle's Goodreads-books repository and processed by machine learning algorithms, including approximately 900,000 ratings of 10,000 books. Sensitivity, Specificity, and F1-Score were calculated for the proposed model's algorithms. The average sensitivity and specificity were 49.76% and 56.74%, respectively.

Rui Sun, Chuyang Wei, et al. [44] built an expert system-based dataset to train a good book classifier for elementary school students. The study provides a more detailed grade division. Elementary schools can be divided into lower grades (1st, 2nd, and 3rd) and upper grades (4th, 5th, and 6th). Four algorithms of machine learning and deep learning were used to classify the text which were Logistic Regression (LR), Support Vector Machine(SVM), AdaBoost (ADB), Naive Bayes (NB), and Text Convolutional Neural Network (CNN). It was found that the linear regression method had an accuracy rate of 98.5 percent.

Furthermore, Taushif Anwar and V. Uma [45] presented a new approach CD-SPM that combines Wpath, CF, and Sequential Pattern Mining (SPM) to recommend the most popular items from different domains with better recommendation accuracy. Wpath aids in the discovery of semantic similarity among items from various domains in this study. To find the most common sequences, the PrefixSpan algorithm and Topseq rules are used. Cross-Domain

Sequential Pattern Mining (CD-SPM) outperforms the CF-KNN approach in terms of performance and alleviates the new user and sparsity problems to some extent because one domain's knowledge (rating) is applied to another. Table 3 compares recently published research that relies on CF models.

**Table 3 :** Research Trends in Collaborative Filtering for eBooks RMs

| Ref. | Year | Techniques / Similarity Methods | Dataset | Findings / Contribution | Problems | Evaluation Metrics |
|---|---|---|---|---|---|---|
| [30] | 2018 | Matrix factorization, Pearson correlation and cosine similarity | 124,406 library records of Dhurakij Pundit University obtained from 2014 to 2017 | The matrix rating is constructed using the time-stamped book borrowing records, and The sparseness problem solved using SVD. | Low users' satisfaction and accuracy can be enhanced by utilizing book bibliography such as category, publisher, author | Accuracy |
| [31] | 2018 | Matrix factorization technique, Pearson correlation, Cosine similarity and | Contains borrowing records from 2014 to 2017 of the Dhurakij Pundit University Library | The students' rating matrix is constructed based on the borrowing and corresponding returning records with time stamps | Low accuracy, Grey sheep, cold-start, and data sparsity problems | Accuracy Measure and Student Satisfaction |
| [32] | 2018 | Restricted Boltzmann machine RBM | Ritsumeikan Art Research Center (ARC) database | Using the RBM of some of the digital archive datasets offered by ARC to provide better cultural treasures. | Scalability and complexity | MAE and RMSE |
| [33] | 2018 | Cosine similarity , expert recommendati on function | BookCrossing dataset | Used CF algorithm to propose books for new readers and new novels. | Low accuracy, Grey Sheep Problem, Data Sparsity, and big data set required | N/A |
| [34] | 2018 | k-NN, Pearson Similarity, and Cosine Similarity | Book-Crossing community | Proposing three different models using personal attributes, age, location, and interest for new users | Categorizing datasets based on different genres and recommending books based on an area of interest. | RMSE, MAE |
| [35] | 2019 | Association Rule Mining, SPADE, Euclidean distance, and *k*-means algorithms | Dataset from Open Online Courses | Selecting a dataset with high coverage improves performance, and that clustering dataset has a major impact. | Grey sheep problem, cold start problem, data sparsity problem | Coverage Measure |
| [36] | 2019 | Cosine similarity, weighted sum method and Correlation Matrix | goodbooks10k | It does not require a prior user profile, personalized recommendations, expands the user interest area | Cold-start, data sparsity, filling the unknown cells in the matrix with zeros reduces accuracy and increases the bias problem | MAE and Statistical Accuracy Metric |

| | | | | | | |
|---|---|---|---|---|---|---|
| [37] | 2019 | Cosine similarity | UTM library dataset | probabilistic-keyword CF method is developed to solve the sparsity problem of a library BRMs | Grey sheep problem, cold start problem, data sparsity problem | Average Precision (AP) and F1-Score |
| [38] | 2019 | Naïve Bayes, decision tree, Word2Vec model | Book-Crossing dataset | Applying the word embedding method to represent book titles can improve title representation and make better predictions | Low accuracy, Grey Sheep Problem, Data Sparsity | Accuracy, Precision, Recall, and RMSE |
| [39] | 2019 | Jaccard Similarity | BookCrossing dataset | Using CF with JS with a compact dataset was more accurate than existing algorithms with full datasets | Scalability, Grey Sheep Problem, and Data Sparsity | RMSE |
| [40] | 2020 | Cosine similarity, and KNN | Local Datab ase in TasikmalayaPe juang University Library | This system provides book title solutions to users according to their profile | Low accuracy and new user cold-start problem | Blackbox testing method by Trial and Error. |
| [41] | 2021 | Decision Tree, Fuzzy Model | Amazon e-commerce site and GoodReads | Proposed a fuzzy-logic RMs for online book shoppers | Grey Sheep Problem, Data Sparsity and vulnerability to attacks | MAE, Recall, Precision and F-Measure |
| [42] | 2021 | POS tagger, word embedding, CNN and KNN | Kaggle website and Amazon website | Using Sentiment Analysis and deep learning increases the accuracy | Cold-start problem and vulnerable to attacks | Accuracy |
| [43] | 2021 | K-means, Cosine Distance function | GoodReads book dataset repository | Recommendations based on a particular book are more accurately effective than a user-based recommendation system | Grey Sheep Problem, Data Sparsity and vulnerability to attacks | Recall, Precision and F-Measure |
| [21] | 2021 | cosine similarity | data of the library of Wuxi Vocational College of Science and Technology | the proposed method converges faster than the traditional method. | The cold-start problem, low accuracy and vulnerability to attacks | MAE and RMSE |
| [44] | 2022 | Logistic Regression (LR), SVM, AdaBoost (ADB), Naive Bayes (NB), and TextCNN | Constructed dataset (Chinese Books) | Due to the low discrimination between datasets, it is difficult for deep learning to automatically extract feature information, which | Specific model, low accuracy and vulnerable to attacks | Accuracy, F1-score and Precision |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | in turn leads to low accuracy. | | |
| [45] | 2022 | WpathSimilarity, SVD, Cosine similarity PrefixSpan algorithm and Topseq Rules | Using the Book domain dataset from github.com | PrefixSpan algorithm and Topseq rules are applied to improve the accuracy and sparsity problem can be effectively addressed | Cold-start, Grey Sheep problem and vulnerable to attacks and take more time | Precision, Recall and F1Score |

### 5.3 Research trends in eBooks hybrid filtering

Aleksandar Simović [46] presented an approach to managing large amounts of differential data from many sources using Hadoop-based intelligent libraries. Library customers' satisfaction and some distinctive aspects of library administration are created by integrating smart RMs into a big data environment. Similarly, Rohit Darekar [47] proposed a system that can take advantage of both content-based and collaborative algorithms. The method used to reduce the user history while generating a recommendation for the users by Neglecting the data of books that users used in recent times helps the system to generate efficient results faster.

In Sivaramakrishnan N et al. [48] a neighbour-based approach's correlation coefficients were compared (Pearson Correlation (PC) coefficient, Constrained Pearson correlation (CPC) coefficient, Spearman Rank Correlation (SRC) coefficient). All three relationships for neighbours between 2 and 9 were tested. The system PC coefficient and CPC were identical, with minor mean differences.

Abhay E. Patil et al. [49] proposed a system that recommends various books to consumers using CF and association rule mining. Hybrid RMs were built using these strategies, which address the issue of data sparsity and the issue of a cold start. Both algorithms produce correct results. Meanwhile, Erin Cho and Meng Han [50] presented a study that aimed to illustrate the use of AI in the development of book RMs. Personalised reading lists could be created for each user based on the books they want. The researchers drew on a Goodreads dataset and user data to create these recommender structures.

In Yonghong Tian et al. [51] CF and CBF algorithms were examined for use in university textbooks. In order to create a personalised book RMs. The Spark big data platform and the hybrid algorithm were utilised to generate the results. There is evidence to suggest that hybrid techniques can deliver more precise suggestions.

Nida Khairunnisa et al. [52] proposed a web-based RM system for the Open Library at Telkom University. The user must borrow the book from the library before receiving a recommendation. After returning the book, the patron should rate it on their account library's web page. Madhuri Kommineni et al. [53] The study provided user-based CF approaches for a book recommendation and assessed the effectiveness of several similarity measures. The overall architecture of the proposed system is modelled, and its implementation is illustrated via model design. Sunny Sharma et al.[14] a hybrid approach was proposed for predicting book suggestions. The suggested system combines CF and CBF techniques to identify users who are similar to the active user by matching their profiles. In the final step, the intended user offers items based on the prediction value calculated for each item using the Resnick prediction equation. Table 4 compares recently published research that relies on hybrid recommendation models.

**Table 4 :** Research Trends for Hybrid eBooks RMs

| Ref. | Year | Techniques / Similarity eq. | Dataset | Findings / Contribution | Problems | Evaluation Metrics |
|---|---|---|---|---|---|---|
| [46] | 2018 | Big Data technology | Collected from multiple sources, including LMS moodle[54], educational institution IS, social networks and online bookstore server logs | Integrating recommender systems to the smart library in the Big Data environment | Low user satisfaction and trustless problem | Big Data Analysis (Hadoop Ecosystem result with library recommended books), An online questionnaire |
| [47] | 2018 | Cosine similarity formula | *http://snap.stanford.edu/data/web-Amazon-links.html .* | Combines CF, CBF with the demographic filtering approach | Data sparsity, scalability and Reliable Integration | N/A |
| [48] | 2018 | KNN, Pearson, cosine similarity, Kendall" s Tau correlation, Jaccard similarity, Spearman Rank Correlation, Mean-squared distance | Local Book Dataset | In this book recommender system spearman correlation coefficient works best and having mean absolute error less than 1. | Data Sparsity Problem, Cold Start Problem and increased complexity | Mean Absolute Error (MAE) |
| [49] | 2019 | Pearson correlation | Local Dataset | The usage of both CF and association rule mining can help to control the data sparsity and cold start problem in recommendation systems | Scalability, Protect the system data against attacks and Using NLP Techniques | N/A |
| [50] | 2019 | TF-IDF, Pearson Similarity, Cosine similarity | Goodreads | Users could compile personalized reading lists with books recommended to them. | comparing the whole texts of the books themselves, not just the short description | Pearson correlation coefficient |
| [51] | 2019 | Cosine similarity, KNN, and K-means | Datasets of Library(Mongolia University) | Library website improvements more accurate recommendations possible than with pure approaches | Data Sparsity Problem, Cold Start Problem | Precision |
| [52] | 2020 | cosine similarity, weighted sum method | Telkom University Open Library | Item-based matching employs book attributes to determine | Reliable Integration, Efficient Calculation | Beta Testing, Accuracy and MAE |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | similarities between titles | | |
| [53] | 2020 | Similarity techniques (e.g., Cosine, Jaccard) and Machine Learning (SVD) | Kaggle Goodreads books data | Decision-making can benefit from the model's instruction, feedback, administration, reporting, and setup. | Data sparsity, scalability and take more time | Mean Absolute Precion (MAP) , Recall and Precision |
| [14] | 2021 | Nearest Neighbor Approach, TF-IDF, Cosine similarity | IIF's repository | propose a hybrid system that solves Cold start and sparsity problems | Data sparsity, scalability | MAE, Recall, Precision, and F-Measure |

Figures 9 and 10 are visualizations of Table 4, used to assess the research trends of approaches employed in recommendation systems. Figure 9 represents the number of recommended techniques utilized in the publications reviewed based on the survey gathering criteria established in this research (On the three recommendation models). In addition, Figure 10 depicts the usage of each recommended technique in a certain recommendation model based on the flow throughout the year.

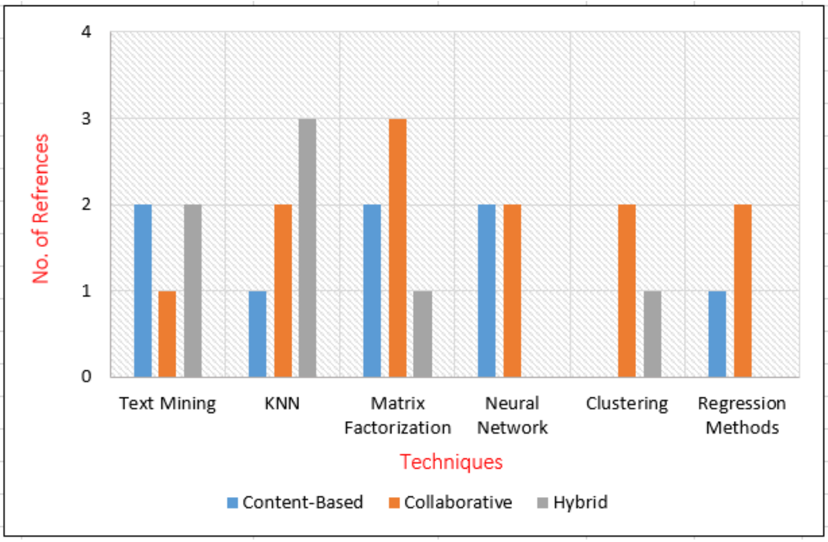



**Figure 9:** Trend of research papers by recommendation model

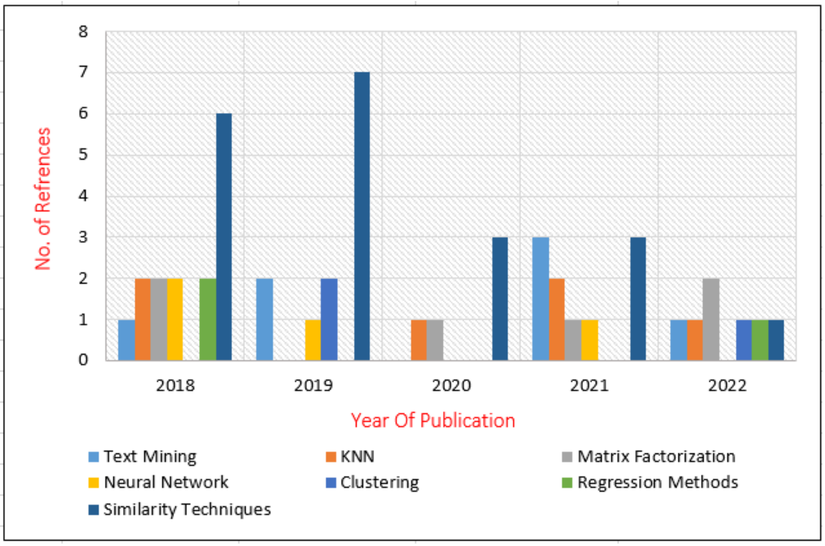



**Figure 10:** Trend in recommendation technique papers by year during the period (2018–2022).

Text mining is an essential approach for RMs, particularly those that analyse the attributes of user-specified items using a CBF model (that uses natural language processing techniques). Additionally, it is utilised in the item-based CF model and the hybrid recommendation model. This technique may be employed in various recommendation system models and is in demand in fields containing vast amounts of textual data, such as medical data (the healthcare industry), academia, and tourism. Figure 9 demonstrates that text mining is a technology actively employed in the RMs study. In addition, Figure 10 demonstrates that text mining is consistently utilised in RMs study. However, because of the ineffective search process for K values, the bias issue for K, and the issue that it cannot be utilised when the data size is large, the explicit use of KNN was restricted in the RMs search gathered in this study. However, the majority of the research that has been discussed heavily relies on similarity approaches.

In the travel industry, CBF and the hybrid recommendation model commonly utilise clustering to analyse the similarity of location-based data. In modern apps and web services, the usage of clustering as a suggestion strategy is declining because of the popularity of likes, star ratings, and quantitative data for user evaluation. However, Figure 9 shows that clustering has been employed consistently. MF technology has seen widespread application, particularly in the collaborative filtering model, as it seeks potential factors that express user preference for the items provided by the service. It is possible to analyse the numerical data of specific items and the collected case data. Since it is computed by decoding a matrix consisting of user and item ratings, the amount of time spent processing external mouse traffic data is also minimised. Since the method ultimately solves KNN difficulties, notably the Sparsity problem, its popularity has grown.

Studies of the RMS started using neural network technology to examine various data sets. For instance, a neural network technique often employed for evaluating photos and making predictions about photographs submitted by a user or an item purchased by a user was not particularly relevant to our investigation. While Recurrent Neural Network (RNN) is commonly employed in text-based systems like news and ebooks, CNN is more used in image-based ones. Numerous technical issues with MF are addressed, including the cold start problem, over-specialisation, and difficulty employing side features. The use of neural networks over the research period is shown in Figures 9 and 10.

## 6. Evaluation Metrics for eBook RMs

It is important to quantify what makes a good RMs and evaluate the RMs model. Evaluating models in RMs research is an important component of the field. Thus, this study looked at how those models are traditionally measured. The RMSE is the simplest indicator of an RMs's performance. Evaluation of prediction accuracy is done using the Square Root of the Mean Squared Error (RMSE). It is important to quantify what makes a good RMs to evaluate the RMs model [55][9]. These Metrics are divided by the total number of predicted grades. Precision, Recall, Accuracy, F-Measure, Receiver Operating Characteristic Curve (ROC Curve), and Area Under the ROC Curve (AUC) are qualitative indicators of RMs. This model's qualitative evaluation index was calculated using the confusion matrix.

Table 5 is a confusion matrix used to measure RMs performance. This matrix shows if the system recommends the user's preferred item. The rows display the user's preferences, and the columns indicate whether or not the recommendation model suggested that option [56].

**Table 5 :** Confusion Matrix of RMs [9]

| Preference | Recommended | Not Recommended |
|---|---|---|
| User-preferred item | True Positives (TP) | True Negatives (TN) |
| User-non-preferred item | False Positives (FP) | False Negatives (FN) |

The TP column in the confusion matrix represents the number of items that meet the user's desire when the RMs proposes an item. TN is the total number of user-preferred items for which the recommender system made no recommendations. How often a system suggests an action the user does not wish to take is a measure of its FP. When the algorithm fails to suggest products consumers have said they do not like, this is called an FN. Rates of actual positivity can be measured using terms like True Positive Rate (TPR) and False Positive Rate (FPR). The TPR is equivalent to the Recall [9]. The ROC curve shows how the FPR relates to the TPR graphically. This approach provides a visual representation of the Precision and Recall metrics. The ROC curve is a graph, so deriving a numerical value from it is quite challenging. To solve this, the AUC index is typically applied. A recommendation model's efficacy can be evaluated by calculating its AUC curve. The AUC value is close to 1, indicating that the model's performance is very good. AUC values of 0.8 or higher indicate a very accurate model [56].

Diverse and unique recommendations are evaluated by their diversity score. A recommender with limited diversity would only suggest items from the same provider. To determine serendipity, compare the probability that item *i* will be suggested to a specific user with the probability that it will be suggested to any user. The likelihood of a recommendation is simply proportional to its rank among n items. The hardest part of solving this problem is verifying its applicability. Fewer items will be used in the test set if a user does not rate many goods, and serendipity will be low. Novel suggestions are ones that the user has never seen before. Because it indicates how well recommendations meet users' desires for known and new information, measuring novelty is challenging. Improved accuracy and increased efficiency benefit from incorporating new information into a recommendation system [57]. Table (6) includes the equations and descriptions of the general qualitative measuring metrics used to evaluate the performance of RMs.

**Table 6 :** Description of Evaluation Metrics

| Metric Equation | Description | Used by |
|---|---|---|
| $\text{RMSE} = \sqrt{\frac{1}{n}\sum_{u,i}(p_{u,i} - r_{u,i})^2}$ | Is the standard deviation of the forecast errors | [28] [29] [32] [34] [38] [39] [21] |
| $\text{MAE} = \frac{\sum_{i=1}^{N}\lvert \text{Pi} - \text{Qi}\rvert}{\text{N}}$ | Relates to the average size of errors in a series of projections. It is the average absolute difference between predicted and actual observation over the test sample. | [26] [28] [29] [32] [34] [36] [41] [21] [41] [48] [52] [14] |
| $\text{MAP} = \frac{1}{n}\sum_{i=1}^{n} P(u_i)$ | Is the average of Average Precision of each class | [29] [53] |
| $\text{Recall} = \frac{\text{TP}}{\text{TP+FN}}$ | It is the proportion of the current user's favourite recommended items to his or her overall favourite items. | [14] [38] [41] [43] [45] [53] |
| $\text{Precision} = \frac{\text{TP}}{\text{TP+FP}}$ | The percentage of goods, out of all those that are recommended to the user, which are compatible with the user's preferences. | [14] [29] [37] [38] [41] [43] [44] [45] [51] [53] [58] |
| $\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+ FN}}$ | Asymmetrical mean value can be seen in both precision and recall measurements . | [30] [31] [36] [38] [42] [44] [59] |
| $\text{F-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$ | Precision and recall's harmonic mean values. | [14] [29] [37] [41] [43] [44] |
| ROC curve metric<br>$\text{Sensitivity} = \text{TP Rate} = \frac{\text{TP}}{\text{TP+FN}}$ and<br>$\text{Specificity} = \text{FP Rate} = \frac{\text{FP}}{\text{FP+ TN}}$ | A graph illustrating the connection between the FPR and TPR. a representation in visual form of the performance outcomes for Precision and Recall as a ratio [60]. | [61] |
| The area under the ROC curve, or AUC | AUC estimates the likelihood that a random relevant item is ranked above a random irrelevant item. | [62] |
| Diversity = 1 – Similarity | This means that the user has approved of a wide range of options that differ from their usual preferences. | [63] |
| $\text{Pi} = \frac{\text{n} - \text{ranki}}{\text{n} - 1}$<br>$\text{Serendipity}_u = \frac{1}{n}\sum_{i\in n}^{n} \max(\text{Pi(user)} - \text{Pi(allUser)}, 0)\ \_\ \text{reli(user)}$ | A serendipity metric gauges how relevant or surprising recommendations are made to the user. | [57] |
| $\text{Novelty} = \frac{\text{Number of recommended items unknown}}{\text{Total number of provided recommendations}}$ | Influences how unknown things recommended by other users are recommended to a user. | [57] |

Some of the evaluation metrics listed in the preceding table were not included in most book recommendation systems; nonetheless, common criteria like RMSE, MAE accuracy and recall were utilised. The assessment metrics in RMs may change based on the data supplied since conventional procedures are employed when working with common datasets. However, when the work is posted on the internet without any prior information, the questionnaire might be utilised in the rating process. Figure 11 depicts the usage of these metrics in the research under consideration.

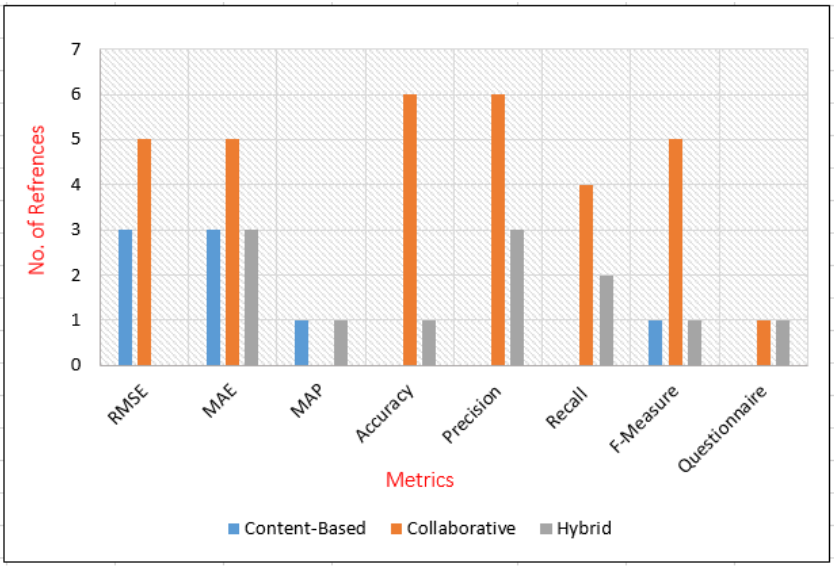



**Figure 11:** Statistical numbers of reviewed papers based on evaluation metric

**7. Conclusions and Future scope**

The internet is home to a plethora of unstructured data, making RMs a fertile topic for study. Users can read the books online or purchase them. Nevertheless, on any digital ebooks platform, a huge number of books are available, creating a problem for the user in finding books matching their interest. To resolve this problem, book RMs is introduced.

Several approaches are used to design the book RMs. Recommendation systems in application domains incorporating mouse clicks, browsing or viewing period, commenting, and rating values through user interactions with the system (feedback) have been found to provide superior outcomes. User behaviour affects how much the data will change.

This paper reviewed the recent five years of research on book RMs for the period from 2018 to 2022. The study was based on the basic recommendation models CBF, CF, and hybrid filtering model. The techniques used in these models were classified into seven categories: text mining, KNN, matrix factorisation, neural network, clustering, regression methods, and similarity metrics. Each model has advantages and limitations and may be used in conjunction with the others. For example, when no description or keywords are provided in the CBF model, the CF model can rely on the ratings matrix to discover comparable users with the same taste. When there are no ratings for new products or users, CF models suffer from scalability, sparsity, and cold start problems, which can be rectified using CBF filtering. As a result, the hybrid model was employed to overcome these challenges and constraints, albeit at times at the price of the model's complexity.

In addition, the available datasets, the recommendation model used, and the contributions and shortcomings of each work were analysed and compared to help the curious researcher. the paper also discussed the evaluation metrics used to evaluate the performance of book RMs.

In the future, many new features and technologies may be created and tested for the effective implementation of RMs. Also, by integrating RMs with ML and NLP, we can create effective and powerful RMs considering many factors. Through the study, it was also determined that the majority of research in this field did not account for the reliability of the items and users. That made these systems vulnerable to attack by a variety of malicious attacks; for instance, there are individuals who provide large numbers of elements that are of no value to confuse the system and provide false assessments to reduce the trust in efficient providers. To address these issues, a hybrid model that combines classic recommendation models with trust-based models may be employed.

**References**


**[1]** S. A. Qader and A. R. Abbas, “Survey of User to User Recommendation System in Online Social Networks,” *Eng. Technol. Journal, Univ. Technol. Baghdad, Iraq*, vol. 37, no. 10, pp. 7, 2019.

**[2]** H. Alharthi, D. Inkpen, and S. Szpakowicz, “A survey of book recommender systems,” *Journal of Intelligent Information Systems*, vol. 51, no. 1. pp. 139–160, 2018. doi: 10.1007/s10844-017-0489-9.

**[3]** P. Nitu, J. Coelho, and P. Madiraju, “Improvising Personalized Travel Recommendation System with Recency Effects Paromita,” *Big Data Min. Anal.*, vol. 4, no. 3, pp. 139–154, 2021, doi: 10.26599/BDMA.2020.9020026 Improvising.

**[4]** S. Verma and K. Patel, “Weighted product Taxonomy for Mobile-Commerce site in Recommendation of Product based on Heuristic Approach,” *Proc. Int. Conf. Intell. Comput. Control Syst. , IEEE Xplore*, no. CFP19K34-ART; ISBN: 978-1-5386-8113-8, pp. 6, 2019.

**[5]** N. Hariri, Carlos Castro-Herrera, J. Cleland-Huang, and B. Mobasher, “Recommendation systems in software engineering,” in *Recommendation Systems in Requirements Discovery*, Chicago, IL, USA: *Springer-Verlag Berlin Heidelberg*, pp. 455–476, 2014. doi: 10.1007/978-3-642-45135-5.

**[6]** Z. Shahbazi and Y.-C. Byun, “Agent-Based Recommendation in E-Learning Environment Using Knowledge Discovery and Machine Learning Approaches,” *MDPI, Math.*, vol. 10, pp. 19, 2022, doi: 10.3390/math10071192.

**[7]** P. Pavan Kumar, S. Vairachilai, S. Potluri, and S. N. Mohanty, "Recommender Systems Algorithms and Applications," *First. Caslon: CRC Press is an imprint of Taylor & Francis Group*, LLC, 2021.

**[8]** S. Chinchanachokchai, P. Thontirawong, and P. Chinchanachokchai, “A tale of two recommender systems: The moderating role of consumer expertise on artificial intelligence based product recommendations,” *J. Retail. Consum. Serv. , Elsevier Ltd*, pp. 12, 2021.

**[9]** S. Malik, A. Rana, and M. Bansal, “A Survey of Recommendation Systems: Recommendation Models, Techniques, and Application Fields,” *Information Resources Management Journal*, vol. 33, no. 4. pp. 53–73, 2020. doi: 10.4018/IRMJ.2020100104.

**[10]** Z. Fayyaz, M. Ebrahimian, D. Nawara, A. Ibrahim, and R. Kashef, “Recommendation systems: Algorithms, challenges, metrics, and business opportunities,” *Applied Sciences (Switzerland)*, vol. 10, no. 21. pp. 1–20, 2020. doi: 10.3390/app10217748.

**[11]** B. B. Sinha and R. Dhanalakshmi, “Evolution of recommender paradigm optimization over time,” *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 4. pp. 1047–1059, 2022. doi: 10.1016/j.jksuci.2019.06.008.

**[12]** C. C. Aggarwal, *Recommender Systems*. NY, USA: Springer International Publishing Switzerland, 2016. doi: DOI 10.1007/978-3-319-29659-3.

**[13]** G. Geetha, M. Safa, C. Fancy, and D. Saranya, “A Hybrid Approach using Collaborative filtering and Content based Filtering for Recommender System,” *Natl. Conf. Math. Tech. its Appl. (NCMTA 18)*, pp. 7, 2018, doi: 10.1088/1742-6596/1000/1/012101.

**[14]** S. Sharma, V. Rana, and M. Malhotra, “Automatic recommendation system based on hybrid filtering algorithm,” *Education and Information Technologies*, vol. 27, no. 2. pp. 1523–1538, 2022. doi: 10.1007/s10639-021-10643-8.

**[15]** M. Sharma, R. Mittal, A. Bharati, D. Saxena, and A. Kumar, “A Survey and Classification on Recommendation Systems,” *Int. Conf. Big Data Mach. Learn. Appl.*, pp. 1–17, 2021.

**[16]** A. K. Sahoo, C. Pradhan, and B. S. P. Mishra, “SVD based Privacy Preserving Recommendation Model using Optimized Hybrid Item-based Collaborative Filtering,” *Int. Conf. Commun. Signal Process. IEEE Adv. Technol. Humanit.*, pp. 5, 2019.

**[17]** F. Trabelsi, A. Khtira, and B. El Asri, “Hybrid Recommendation Systems: A State of Art,” *Int. Conf. Eval. ofNovel Approaches to Softw. Eng.*, pp. 281–288, 2021, doi: 10.5220/0010452202810288.
**[18]** M. H. Mohamed, M. H. Khafagy, and M. H. Ibrahim, “Recommender Systems Challenges and Solutions Survey,” *Proc. 2019 Int. Conf. Innov. Trends Comput. Eng. ITCE 2019*, pp. 149–155, 2019. doi: 10.1109/ITCE.2019.8646645.
**[19]** X. Li, J. Xing, H. Wang, L. Zheng, S. Jia, and Q. Wang, “A hybrid recommendation method based on feature for offline book personalization,” *arXiv*, no. Xx, pp. 1–15, 2018.
**[20]** R. K. Behera, A. Gunasekaran, S. Gupta, S. Kamboj, and radip K. Bala, “Personalized digital marketing recommender engine,” *Elsevier Ltd*, pp. 24, 2019. doi: 10.1016/j.jretconser.2019.03.026.
**[21]** Z. H. Wang and D. Z. Hou, “Research on Book Recommendation Algorithm Based on Collaborative Filtering and Interest Degree,” *Wirel. Commun. Mob. Comput.*, pp. 7, 2021.
**[22]** H. Alharthi, D. Inkpen, and S. Szpakowicz, “Authorship identification for literary book recommendations,” *School of Electrical Engineering and Computer Science*, 2019.
**[23]** C.-N. Ziegler, S. M. McNee, J. A. Konstan, and G. Lausen, “Improving Recommendation Lists Through Topic Diversification.” *International World Wide Web Conference Com- mittee (IW3C2)*, Chiba, Japan, pp. 11, 2005.
**[24]** P. G. Sudasinghe, “Enhancing Book Recommendation with the use of Reviews,” *University of Colombo School of Computing*, 2019.
**[25]** K. S. Srujan, S. S. Nikhil, H. Raghav Rao, K. Karthik, B. S. Harish, and H. M. Keerthi Kumar, “Classification of amazon book reviews based on sentiment analysis,” *Advances in Intelligent Systems and Computing*, vol. 672. pp. 401–411, 2018. doi: 10.1007/978-981-10-7512-4_40.
**[26]** E. R. Núñez-Valdez, D. Quintana, R. González Crespo, P. Isasi, and E. Herrera-Viedma, “A recommender system based on implicit feedback for selective dissemination of ebooks,” *Information Sciences*, vol. 467. pp. 87–98, 2018. doi: 10.1016/j.ins.2018.07.068.
**[27]** R. Rahutomo, A. S. Perbangsa, H. Soeparno, and B. Pardamean, “Embedding Model Design for Producing Book Recommendation,” *Int. Conf. Inf. Manag. Technol. 537, IEEE,* 2019.
**[28]** Y. K. Ng, “CBREC: A book recommendation system for children using the matrix factorisation and content-based filtering approaches,” *International Journal of Business Intelligence and Data Mining*, vol. 16, no. 2. pp. 129–149, 2020. doi: 10.1504/IJBIDM.2020.104738.
**[29]** T. P. Sariki and G. B. Kumar, “An aggrandized framework for enriching book recommendation system,” *Malaysian J. Comput. Sci.*, vol. 35, pp. 17, 2022.
**[30]** C. Sirikayon, P. Thusaranon, and P. Pongtawevirat, “A collaborative filtering based library book recommendation system,” *Proc. 2018 5th Int. Conf. Bus. Ind. Res. Smart Technol. Next Gener. Information, Eng. Bus. Soc. Sci. ICBIR*, pp. 106–109, 2018.
**[31]** C. Sirikayon, P. Thusaranon, and W. Sukpongthai, “A preprocessing matrix factorization on collaborative filtering based library book recommendation system,” *ACM International Conference Proceeding Series*. pp. 33–37, 2018. doi: 10.1145/3239283.3239302.
**[32]** J. Wang and K. Kawagoe, “A recommender system for ancient books, pamphlets and paintings in ritsumeikan art research center database,” *ACM International Conference Proceeding Series*. pp. 53–57, 2018. doi: 10.1145/3192975.3193018.
**[33]** Y. Liang and S. Wan, “The Design and Implementation of Books Recommendation System,” *Proceedings of the IEEE International Conference on Software Engineering and Service Sciences, ICSESS*. pp. 305–308, 2018. doi: 10.1109/ICSESS.2018.8663914.
**[34]** Rohit, S. Sabitha, and T. Choudhury, “Proposed approach for book recommendation based on user k-NN,” *Springer Nature Singapore Pte Ltd*, vol. 554. pp. 543–558, 2018. doi: 10.1007/978-981-10-3773-3_53.
**[35]** R. Obeidat, R. Duwairi, and A. Al-Aiad, “A Collaborative Recommendation System for Online Courses Recommendations,” *Int. Conf. Deep Learn. Mach. Learn. Emerg. Appl. IEEE*, 2019.
**[36]** K. Shah, “Book Recommendation System using Item based Collaborative Filtering,” *Int. Res. J. Eng. Technol.*, 2019, [Online]. Available: www.irjet.net
**[37]** N. Ifada, I. Syachrudin, M. K. Sophan, and S. Wahyuni, “Enhancing the performance of library book recommendation system by employing the probabilistic-keyword model on a collaborative

filtering approach," *Procedia Computer Science*, vol. 157. pp. 345–352, 2019.
**[38]** T. T. S. Nguyen, "Model-Based Book Recommender Systems using Naïve Bayes enhanced with Optimal Feature Selection." *Association for Computing Machinery*. ACM ISBN 978-1-4503-6573-4/19/02, Penang, Malaysia, 2019.
**[39]** A. Rana and K. Deeba, "Online book recommendation system using collaborative filtering (with jaccard similarity)," *Journal of Physics: Conference Series*, vol. 1362, no. 1. 2019. doi: 10.1088/1742-6596/1362/1/012130.
**[40]** M. Hikmatyar and Ruuhwan, "Book Recommendation System Development Using User-Based Collaborative Filtering," *Journal of Physics: Conference Series*, vol. 1477, no. 3. 2020. doi: 10.1088/1742-6596/1477/3/032024.
**[41]** M. F. Adak and M. Uçar, "A Book Recommendation System Using Decision Tree-based Fuzzy Logic for E-Commerce Sites," *3rd Int. Congr. Human-Computer Interact. Optim. Robot. Appl. (HORA), IEEE Xplore*, pp. 5, 2021, doi: 10.1109/HORA52670.2021.9461319.
**[42]** S. Saraswathi and A. Mounika, "Design of Book Recommendation System Using Sentiment Analysis," *Springer Nat. Singapore*, pp. 7, 2021, doi: 10.1007/978-981-15-5258-8_11.
**[43]** D. Sarma, T. Mittra, and S. Hossain, "Personalized Book Recommendation System using Machine Learning Algorithm," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 1. pp. 212–219, 2021. doi: 10.14569/IJACSA.2021.0120126.
**[44]** R. Sun, C. Wei, Z. Li, J. Zhao, S. Li, and J. Ma, "Chinese Books Recommendation for Elementary Students: A Comparison between ML and DL Methods," *Int. Conf. Neural Networks, Information, Commun. Eng.*, pp. 7, 2022, doi: 10.1117/12.2639345.
**[45]** T. Anwar and V. Uma, "CD-SPM: Cross-domain book recommendation using sequential pattern mining and rule mining," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 3. pp. 793–800, 2022. doi: 10.1016/j.jksuci.2019.01.012.
**[46]** A. Simović, "A Big Data smart library recommender system for an educational institution," *Libr. Hi Tech*, p. 26, 2018, doi: 10.1108/LHT-06-2017-0131.
**[47]** R. Darekar, K. Dayma, R. Parabh, and Swapnali, "A Hybrid Model for Book Recommendation," *Int. Conf. Inven. Commun. Comput. Technol. IEEE*, pp. 5, 2018.
**[48]** N. Sivaramakrishnan, V. Subramaniyaswamy, S. Arunkumar, A. Renugadevi, and K. Ashikamai, "Neighborhood-based approach of collaborative filtering techniques for book recommendation system," *Int. J. Pure Appl. Math.*, vol. 119, Special Issue, 2018.
**[49]** A. E. Patil, S. Patil, K. Singh, P. Saraiya, and A. Sheregar, "Online book recommendation system using association rule mining and collaborative filtering," *Int. J. Comput. Sci. Mob. Comput.*, vol. 8, no. 4, pp. 83 – 87, 2019.
**[50]** E. Cho and M. Han, "AI powered book recommendation system," *ACMSE 2019 - Proceedings of the 2019 ACM Southeast Conference*. pp. 230–232, 2019. doi: 10.1145/3299815.3314465.
**[51]** Y. Tian, B. Zheng, Y. Wang, Y. Zhang, and Q. Wu, "College Library Personalizsed Recommendation System Based on Hybrid Recommendation Algorithm," *Procedia CIRP, ScinceDirect , Elsevier*, vol. 65, pp. 32–37, 2019, doi: 10.1016/j.procir.2017.04.009.
**[52]** N. K. Kusumawardhani, M. Nasrun, and C. Setianingsih, "Web Recommended System Library Book Selection Using Item Based Collaborative Filtering Method," *IEEE Xplore*, pp. 8, 2020, [Online]. Available: Auckland University of Technology
**[53]** M. Kommineni, P.Alekhya, T. M. Vyshnavi, V.Aparna, K. Swetha, and V. Mounika, "Machine Learning based Efficient Recommendation Collaborative Filtering Algorithm," *IEEE Xplore*, pp. 6, 2020, [Online]. Available: Cornell University Library
**[54]** F. G. K. Yilmaz and R. Yilmaz, "Student Opinions About Personalized Recommendation and Feedback Based on Learning Analytics," *Technol. Knowl. Learn. Springer Nat. B.V.*, pp. 16, 2020, doi: 10.1007/s10758-020-09460-8.
**[55]** Z. Duan, W. Xu, Y. Chen, and L. Ding, "ETBRec: a novel recommendation algorithm combining the double influence of trust relationship and expert users," *Springer Nat. Singapore*, 2021, doi: 10.1007/s10489-021-02419-x.
**[56]** R. Soleymania, E. Grangera, and G. Fumerab, "F-Measure Curves: A Tool to Visualize Classifier Performance Under Imbalance," *Elsevier Ltd*, 2019, doi: 10.1016/j.patcog.2019.107146.

**[57]** O. Stitini, S. Kaloun, and O. Bencharef, “An Improved Recommender System Solution to Mitigate the Over-Specialization Problem Using Genetic Algorithms,” *Electronics (Switzerland)*, vol. 11, no. 2. 2022. doi: 10.3390/electronics11020242.
**[58]** A. K. A. Hassan and A. B. A. Abdulwahhab, “Proposed collaborative filtering recommender system based on implicit and explicit user’s preferences,” *Iraqi J. Sci.*, vol. 59, no. 2, pp. 771–785, 2018, doi: 10.24996/IJS.2018.59.2A.15.
**[59]** A. R. Abbas and S. Ashor, “Design recommendation system in e-commerce site,” *Iraqi J. Sci.*, vol. 57, no. 4, pp. 2549–2556, 2016.
**[60]** E. Eziama, L. M. . Jaimes, A. James, K. S. Nwizege, A. Balador, and K. Tepe, “Machine Learning-Based Recommendation Trust Model for Machine-to-Machine Communication,” *IEEE Int. Symp. Signal Process. Inf. Technol. Mach.*, p. 6, 2018.
**[61]** L. Yan and Y. Liu, “An ensemble prediction model for potential student recommendation using machine learning,” *Symmetry (Basel).*, vol. 12, no. 5, p. 17, 2020, doi: 10.3390/SYM12050728.
**[62]** P. Han, Shuo Shang, A. Sun, P. Zhao, K. Zheng, and X. Zhang, “Point-of-Interest Recommendation with Global and Local Context,” *IEEE Trans. Knowl. Data Eng.*, p. 12, 2020, doi: 10.1109/TKDE.2021.3059744.
**[63]** J. Kim, I. Y. Choi, and Q. Li, “Customer satisfaction of recommender system: Examining accuracy and diversity in several types of recommendation approaches,” *Sustainability (Switzerland)*, vol. 13, no. 11. p. 20, 2021. doi: 10.3390/su13116165.