International Journal of Digital Content Management (IJDCM)
Vol. 4, No. 7, Summer & Fall 2023; P 337-361
dcm.atu.ac.ir
DOI: 10.22054/dcm.2022.68818.1123

ATU PRESS

Original Research
Accept Date: 30 November 2022
Receive Date: 04 July 2022
eISSN: 2717-3313
ISSN: 2717-3305

Analyzing the Requirements of the Book Recommender System and Providing a Conceptual Model for Iranian Digital Libraries

Maryam Azimian kharazmi University
Nusrat Riahi Nia * Professor, kharazmi University
Ali Azimi Vaghar Assistant Professor, Kharazm University
Keyvan Borna Associate Professor, Kharazm University

Abstract

Purpose: The main purpose of this study is to design and evaluate a book recommender system in digital and public libraries. The solution has been provided by receiving and reviewing the preferences and experiences of users and profile information and studying the background of each user, as well as considering groups of features recorded in the recommendation process.

Method: This research is applied in terms of purpose and survey method. The statistical population studied in this research consists of 263 questionnaires of users and 30 questionnaires of librarian experts. In order to find similarity between users and books, clustering and grouping have been used.

Findings: There are two criteria for grouping: users grouping that can be used on the three indicators of age, gender, educational level, and thematic classification of books can be based on scope, branch, and sub-category. In analyzing the data in the descriptive statistics section,

* Corresponding Author: sara_purriahi@yahoo.com
How to Cite: Azimian, M., Riahi Nia, N., Azimi Vaghar, A., Borna, K. (2023). Analyzing the Requirements for Designing, Implementing, and Evaluating the Book Recommender System and Providing a Conceptual Model for Iranian Digital Libraries, *International Journal of Digital Content Management (IJDCM)*, 4(7), 337-361.

Excel software is used and in the analytical section, SPSS software. Findings indicate that the accuracy criterion has been improved by calculating MAE and RSME in the proposed method compared to the basic method in this field. The results also showed that classification can have a significant impact on the forecast and performance of book forecasting systems.

Conclusion: The evaluation of the conceptual design showed that by focusing on user characteristics and obtaining real feedback of Iranian libraries, the recommender can serve as a key and effective element in the service of the Iranian readership community and play a good role as a virtual reference librarian.

Keywords: book recommender system, recommender systems, item-based Collaborative Filtering, clustering.

Introduction

Today by increasing information, library users no longer face the problem of lack of information. Libraries, and especially digital libraries, are part of automated systems because they produce information much faster than information processed by users (Her, et al,2013).In this day and age, retrieving and finding the proper information for users among the huge quantities of information is the main challenge in libraries. For this reason, the focus of information retrieval systems has changed from inactive repositories to the proposed systems to be actively involved in retrieving useful information and have the ability to learn from user behavior (luj, et al,2015). These systems, introduced as recommender systems with the ability to discover users' interests and predict their priorities, filter books that are likely to be of interest to the user from a large volume of information resources and save users time by offering them what they seek. To put it plainly, in the book recommender system, an attempt is made to identify and suggest to the user the item that is most appropriate and closest to his taste by guessing the way of thinking. What is certain is that the proliferation of information resources in digital libraries and the lack of recommender systems can make members confused about meeting their information needs. This in turn can lead to a waste of time, more costs, and reluctance to re-enter the system (Hi, et al,2021). The occurrence and repetition of such events will have unfortunate consequences, the most important of which are the reduction of membership and departure from the main mission of the library, which is to attract more audiences. This research aims to pay more attention to digital libraries and create prototypes in Iran. Due to conditions such as the Coronavirus pandemic, the need to pay attention to digital libraries has increased. According to the researcher, addressing issues related to the design, implementation, and evaluation of digital libraries in a situation where the use of physical libraries is declining for various reasons can indicate the need for this research. Although a review of the literature shows that the book recommender system has not been implemented in Iranian libraries so far and the importance of identifying the needs of users has not been sufficiently considered, It seems that considering the necessity and need of Iranian libraries, identifying the levels of necessities and factors affecting customer satisfaction and how to communicate with them in a structural change, requires a prerequisite

for any change in Iranian public libraries. The structure of this article is as follows: First, the types of proposed systems are briefly described, then the evaluation criteria in these systems are examined. In the following, literature review will be briefly described. Then the background of the research, then the implementation environment, and the purpose of the study and the proposed method will be discussed, and finally, the result of the evaluation of the book recommender system and Discussion and Conclusion and Implications will be presented.

Types of recommender systems:

The classification of recommender systems is often based on different perspectives such as application, type of knowledge used to propose, and how to formulate and implement algorithms (Damianos,et al,2014). One of the most common classifications accepted by most researchers is the classification of recommender systems based on the type of information they use to make a recommendation: Collaborative Filtering-based systems are the most important and widely used in recommender systems (Ricci,et al,2011). In these systems, the participation and opinions of other users are utilized to predict the interests of the current user, and suggestions are made based on the ranking of other users in the system (Gupta , Gadge,2014) .For example, a user's favorite book can be figured out and predicted by other users' interests. In content-based recommender systems, the main focus is on the items and the similarities between them. For this reason, the design details in these systems depend on how the items are represented. Also, paying attention to user profiles (to the set of information that is provided to the system through the user's preferences, and interests) and learning the user model are two basic and important concepts that should be carefully considered in content-based recommender systems (Jannach,et al,2011) .Hybrid methods are used to cover the issues in the mentioned approaches. There are 7 different mechanisms in the field of hybrid methods called mixed algorithms, weighted switching, cascading, feature combining, feature completion, and superficial (Lu, et al,2015).

Another type of Knowledge-Based Recommender System (KBRS) is based on a knowledge base and user profile. Computational intelligence-based systems also include methods such as fuzzy, genetics, neural networks of decision trees, and Bayesian techniques

(Lua, et al,2015). Also the basis of demographic information systems is based on the principle that users with similar demographic characteristics can have similar preferences. The structure of these algorithms, taking into account the characteristics of users that are usually asked at the time of registration, such as age, gender, education, and occupation, brings users who have closer characteristics together to form user groups with these workgroups then taking into account the preferences of the users, the group produces an offer for the chosen user.

Evaluation criteria of recommender systems:

Comparison between recommender systems requires the use of a standard and acceptable method for evaluating them. The two criteria of Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) are widely used in calculating forecast accuracy in recommender systems. The accuracy of the recommender systems is assessed by comparing the numerical deviation of the predicted scores from the actual corresponding scores of the users (WU, et al,2015). The method of calculating these two criteria is as follows:

MAE: The Mean Absolute Error is calculated from Equation 1.

The lower the value obtained from this relation is, the higher the accuracy of the prediction will be.

Based on this equation:

The rank that the system predicts the user to give to item i (Ru, i), minus the actual rank that user u gives to item i (Ru, i).

Equation 1:

$$MAE = \frac{\sum_{(u,i) \in R_{test}} \left| R_{u,i} - R'_{u,i} \right|}{\left| R_{test} \right|}$$

RMSE: The Root Mean Squared Error is used to calculate the prediction accuracy estimate and is derived from Equation 2. RMSE calculates a much stronger and more significant error than MAE and is, therefore, more suitable for situations where small errors are not very important (wang,et al,2014).

Equation 2:

$$RMSE = \sqrt{\frac{\sum_{(u,i)\in R_{test}} \left| R_{u,i} - R'_{u,i} \right|^2}{|R_{test}|}}$$

Proposed criteria of the collection:
The two accuracy classification criteria that are widely used in evaluating suggestions and user interests are Recall and Precision, the values of which are obtained from Equations 3 and 4.
Recall: Defines the degree of similarity between the number of dependent items and the list of suggestions in the test set.

Equation 3:

$$recall = \frac{\sum_u L(N,u)}{\sum_u L(u)}$$

L (N, u) and L (u) show the dependence of items on user u in the suggested list and test set, respectively. The upper limit of Recall is one, which means that all dependent items in the test set are found in the suggested list.

Precision: defines terms of dependent items in the proposed list, which is equal to the number of desirable (preferred) predictions divided by the total number of predictions (WU, et al,2015).

Equation 4:

$$precision = \frac{\sum_u L(\mathrm{N},\mathrm{u})}{UN}$$

Cosine similarity criterion:
The Cosine similarity criterion is obtained from Equation 5 as follows.

Equation 5:

$$sim(i,j)=\cos(i,j)=\frac{\vec{i}*\vec{j}}{\left\|\vec{i}\right\|^{2}*\left\|\vec{j}\right\|^{2}}=\frac{\sum_{k=1}^{n}R_{k,j}R_{k,j}}{\sqrt{\sum_{k=1}^{n}R_{k,j}{}^{2}\sum_{k=1}^{n}R_{k,j}{}^{2}}}$$

The Cosine similarity vector is calculated based on the angle between the two vectors. If the angle is close to zero, the cosine similarity value of sim (i, j) becomes one; That is, the angles between the two vectors are very similar. The angle between the alternating vectors cannot be more than 90 degrees, which is calculated using Equation 8, where i is the target item and j is another item. Rk,i is the target rate of item i by user k and Rk,j is the target rate of item j by user k and n is the sum of all users' points to items i and j.

Literature Review

Although the history of recommender systems in Iran dates back to 2010, Bakhshandeh Moghaddam in his article explicitly discussed the recommender systems based on trust and provided a method based on user expertise. In this article , It is believed that several methods have been introduced to use trust networks in referral systems, but in none of them has the concept of user expertise been used, because the opinion of experts on a subject is very valuable. Thus, a new method to determine the expertise of people about goods was introduced, and the amount of this expertise in the production of the proposed list by recommender systems was utilized (Bakhshandeh Moghaddam, 2013).

(zar,et al,2014) proposed a recommender system that improved their self-assessment criteria by using user textual comments reflecting the user's interests and feeling. Niknam(2015) dealt with recommender systems, their principles, and types, and examined the use of the recommender system in the database of Noor specialized journals. In the research of (Abbasi, 2018) a model for displaying and acquiring knowledge of processes using case-based inference was proposed. The validity of the proposed model was evaluated through interviews with experts and then, the proposed research model was

designed and implemented in the form of a software system. Kha,et al(2018) presented a web-based recommender system based on item-based Collaborative Filtering algorithms and neural networks. Their proposed method consists of 2 phases inline and online. This proposed method performs better in predicting and recommending items to the user than its basic methods and can provide better recommendations (accuracy and calling) according to the user's taste and needs. Nasouhi, Alizadeh(2021) in their research investigated the role and effect of ontologies in recommender systems. In their proposed method, they use the content information of the items, ratings, and tags between whose items the users want to get the most similarities . Their investigation of real databases shows that their proposed method performs better than previous works and reduces errors and increases accuracy in ranking predictions.

In 2017, Hristakiva and his colleagues designed and implemented a recommender system in the Mendeli system that was able to predict users 'information need based on users' profiles, activity levels, and academic rank in the system. The use of four methods of recommender systems (collaboration filtering, content-based filtering, order based on user choice and user interest) was the chosen method of the Mendeli system that was able to help improve the performance of this system by using a combination of the above methods and applying implicit feedback on users' choices (Hristakeva, et al,2017).

In 2018, Parekh and his colleague developed a book-based hybrid recommender system resting on genetic algorithms. In this article, several different recommenders (with random parameters) are used to predict users' opinions about items. At the end of each stage, the results are evaluated and the necessary strategy to improve the results is considered according to the genetic algorithm, and the predictions are followed again with a new strategy to obtain satisfactory prediction results. Puritat et.al(2021) in their research developed an open-source library system using a recommender system. Therein, Puritat and Intawong used the machine learning method and used the similarity of titles and bibliographic information of books such as author, year of publication, number of visits, etc. They also used the school library, which includes two collections of 2,555 and 4,612 books separately, to measure and evaluate their proposed method. The results of this study show that using a combination of two methods of similarity of titles and bibliographic

information in book recommender systems can be an effective step in small libraries.

Studying the previous literature on recommender systems in Iran, it can be said that in most articles, for designing recommender systems, a model has been described in a different way, each of which has its advantages and disadvantages. It is important to note that the applications of recommender systems in various fields such as education, e-commerce, medicine, and online stores have reached the implementation phase in Iran. However, due to the urgent need for this system in library information retrieval systems, none of the digital and public libraries have used these systems. While outside Iran, many researchers have implemented these systems in digital and electronic libraries and have helped their users to find the resources they need with the help of various methods of data retrieval. Therefore, considering the numerous benefits that the development and growth of these systems will have in improving the situation of public and digital libraries in the country, the presence of such systems is considered vital and practical and will certainly be welcomed by users.

Implementation environment:

To implement, test, and evaluate the proposed algorithm, Java programming language was used with the help of the CF4J open-source library, which is a part of open-source software. The CF4J open-source library is used by several reputable and up-to-date articles as a suitable tool for implementing and evaluating recommender systems.

Purpose of the Study

This research aims to create a book recommendation system to solve all the problems and issues existing in guiding users to obtain their information needs in digital libraries. These problems can be the non-attendance of the reference librarian on time, their lack of information in guiding users, fatigue, and even forgetfulness due to the large volume of their daily activities. For this reason, the use of recommendation systems equipped with artificial intelligence will be able to prevent human mistakes and errors and perform better than humans in conducting reference interviews and guiding users. Needless to say, due to the wide range of information resources in

digital libraries and their increasing growth, the lack of recommendation systems can cause confusion for members in meeting their information needs. This leads to time wastage, imposing more costs on them, and unwillingness to visit libraries again.

Suggested method:

The following steps were taken to implement a book recommender system Building user data sets. Due to the lack of a book recommender system in public libraries in Iran so far, 3 datasets are required to be built. An important issue for the first dataset is understanding the needs of a user and identifying the user. When the user's basic information is requested at the time of registration, the accuracy of entering this information can be considered a crucial step in recognizing a user's needs.

For this reason, the first data set, which contains the profiles, was collected by online questionnaire from the public library members. 263 questionnaires were collected among 3 age groups young, adult, and middle-aged as acceptable questionnaires. Efforts were made to collect the dataset as standard. Also, our target population had to be presented following the research idea. Geographic and demographic information such as national code, age, gender, and marital status are the explicit data asked from users. As the education level, the field of study, and the occupation can affect users' choices, this issue has been addressed in part of the design of the questionnaire. The results of the user questionnaires indicate that the users in this research include 64 men (24%) and 199 women (76%), of whom 110 (42%) have a bachelor's degree, 85 (32%) have master's degree, 27 (10%) high school diploma, 19 (7%) doctorate, 14 (5%) undergraduate and 11 (4%) postgraduate. Also, using the demographic information of users, it was determined that the highest age group of users is less than 37 years old with 130 people as total number.

Questions were also asked about the user's interest in various topics, and the book selection criteria. Based on the user's interest in various topics, a classification of 9 main categories (generalities, philosophy, psychology, religion, social sciences, language, science, technology, art, literature, history, and geography) was provided to the users to categorize different topics according to their priority. Statistical data from the user dataset based on "Measurement of users' favorite categories" show that the categories of psychology with an

average rank of 3.53, literature with an average rank of 5.1, and generalities with 5.11 are considered the most popular categories of users. Accordingly, it can be said that technology, science, history, and geography with the highest average rank are at the lowest level of selection of their target audience.

In this questionnaire, users' criteria for selecting books were asked. Based on this question, 8 main criteria were provided to users to organize them according to their priority. These 8 criteria are book author, book title, translator, subject, publisher, friends and acquaintances' suggestions, list of popular books, and a list of newly published books, respectively. Statistical information from the user dataset based on "measuring the priority of the book selected by users from the mentioned methods" based on the output of the questionnaire indicates that the subject of the book is the priority of the audience with an average rank of 2.47. The book title is the second priority of users with an average rank of 2.82. The author of the book is the third criterion for selecting a book by users with an average rank of 3.19. Users have also prioritized the use of the suggested options of friends and acquaintances, book publishers, and a popular books list with averages of 3.4, 4.82, and 5.41, respectively. Publication year with an average of 57.8% is the last selected rank of users.

Eventually, users were asked to write a list of books they had recently read and to rate them on a scale of 1 to 5 based on their level of interest in their content. The scale was designed by Likert in 1932 and is now used as the most common scale in research. The reason for this question is to implicitly get familiar with the interests of users and the content of the books they have read. The total number of books submitted by users was 1920, which reached 1863 based on the elimination of items for which the selection was not made correctly or the book in question was not part of the book collection. Figure (1) shows an example of the final output of the Excel file. Each line contains a user's input information, which includes users' national code, field of study, educational group, occupation, age, gender, and final grouping.

B	I	J	K	L	M	N	O	P	Q
user_id	major	degree_group	job	age	age_year	age_group	gender	degree_code	roup_nam
3871431060	زبان	LAN	کارمند	46	1354	MID	F	H	group7
3860498487	تربیت بدنی	EXE	کارمند	28	1372	YNG	F	H	group3
3874655970	خیاطی	ART	خانه دار	45	1355	MID	F	L	group8
3873734362	خیاطی	ART	خانه دار	70	1330	AD	F	L	group12
3873734362	ریاضی	MATH	آزاد	74	1326	AD	M	L	group10
4051430790	مدیریت	BUS	کارمند	22	1378	YNG	F	H	group3
0059935529	مهندسی الکترونیک	ELE	آزاد	42	1358	MID	F	H	group7
3873536161	علم اطلاعات	LIB	دانشجو	38	1362	MID	F	H	group7
0947178120	امار	MATH	خانه دار	36	1364	YNG	F	H	group3
0068193963	مدیریت مالي بانکي	BUS	کارمند	39	1361	MID	F	H	group7
5110245371	علوم سیاسی	HU	سایر	39	1361	MID	F	H	group7
3873466481	کامپیوتر	INFO	سایر	40	1360	MID	F	H	group7
0056496206	ادبیات	LIT	خانه دار	49	1351	MID	F	H	group7
3860976702	تکنولوژی آموزشی	ES	دانشجو	24	1376	YNG	F	H	group3
3874053032	تربیت بدنی وعلوم ورزشب	EXE	کارمند	38	1362	MID	F	H	group7
3873443236	مهندسی کامپیوتر	INFO	آزاد	40	1360	MID	F	H	group7
1210133865	پزشکی	MED	دانشجو	24	1376	YNG	F	H	group3
3873881047	اقتصاد	SCI	خانه دار	56	1344	AD	F	L	group12
3861015315	روانشناسی	CON	دانشجو	23	1377	YNG	F	H	group3
1828329445	سینما	ART	سایر	23	1377	YNG	M	L	group2
3860662589	مهندسی معماری	ARC_BU	آزاد	27	1373	YNG	F	H	group3
3874519367	روانشناسی	SCI	خانه دار	42	1358	MID	F	H	group7

Figure (1) An example of the final output of users' Excel file

Creating a database of topics and books:

In the second and third steps of the data collection stage, there is a need to separate and divide the books based on the main topics and sub-topics. Due to the lack of a complete reference for categorizing the thematic and content list of books in the Iranian library system and the use of Dewey classification, the need to create a clear and orderly classification as a necessary step in data collection must be met. For this reason, 5 Iranian sites (Fidibo, Taghcheh, Gisom, Ketabrah, and Iran Kitab that are pioneers in thematic cataloging were selected as reference sites for creating thematic book datasets. Then, by reviewing each book within all five sites and obtaining the necessary subscriptions, the thematic database of the book was created. Accordingly, for each book, one topic was collected as the main topic and three topics as sub-topics with a priority of 1 to 3. 2379 main topics and sub-topics were presented as an Excel file. An example of an Excel file created is shown in Figure 2.

زیر شاخه		شاخه		حیطه	رده اصلی	نام کتاب
شهیدان مسلمان - سوریه	08420922	مدافع حرم	955	تاریخ	900	درعا
جنگ ایران و عراق	842092	داستان و رمان انقلابی	955	تاریخ	900	سگ سالی
عرفان	83	عرفان و تصوف	297	دین و آیین	200	سماع ققنوس
تفسیر قرآن	158	علوم قرآنی	297	دین و آیین	200	روان‌شناسی در قرآن
عاشورا	9534	اسلام	297	دین و آیین	200	در کربلا چه گذشت؟
اهل البیت	974	اسلام	297	دین و آیین	200	زیباترین شکیب
فرهنگ اسلامی	4833	حقوق بشر	297	دین و آیین	200	حقوق بین‌الملل اسلام
علمای شیعه	996	زندگی نامه و خاطرات	297	دین و آیین	200	شیخ آقا بزرگ تهرانی
انسان شناسی	48	عرفان و تصوف	297	دین و آیین	200	کوبریات
جهان بینی اسلامی	285	کلام و عقاید	297	دین و آیین	200	لاصه تاریخ انبیاء: از آدم تا خاتم انبیاء (ص
ازدواج	64	خانواده و روابط	155	روانشناسی	150	ازدواج بدون شکست
روان درمانی	63	خواب دیدن	154	روانشناسی	150	تعبیر خواب
مدیریت ذهن	1	خودسازی	158	روانشناسی	150	درباره‌ی خوب بودن
مثبت اندیشی	2	موفقیت	158	روانشناسی	150	هنر رهایی از چالش‌ها
اختلالات شخصیتی	2	خودسازی	155	روانشناسی	150	بیشعوری
موفقیت شغلی	20	موفقیت	158	روانشناسی	150	ثروتمندترین مرد بابل
شخصیت سازی	3	خودسازی	158	روانشناسی	150	جوجه اردک زشت درون

Figure (2) is an example of the created book thematic dataset

Implementation method:

The implementation phase of this research includes a 6-step method that includes problem analysis and statistical set selection, data collection, test set construction, data analysis and processing, algorithm selection, and implementation. The most important part of this method is the implementation phase of the proposing system algorithm, the core of which includes a combination of methods and techniques to create a suitable user profile.

Figure (3) shows the steps of implementing a book recommender system.

Performance details of the proposed CF algorithm:

In the proposed method, the UserKnn participatory algorithm with cosine similarity matrix with group-book similarity is used. This method consists of two parts: group similarity and book similarity. In group similarity, users are divided into 12 groups. This division is based on 3 main criteria, namely age category (young, middle-aged, and adult), gender (male and female), and education (undergraduate and postgraduate). It should be noted that the selection of these three criteria is based on determining the priority of the above categories through a questionnaire obtained by 30 professional librarians. Accordingly, user groups are

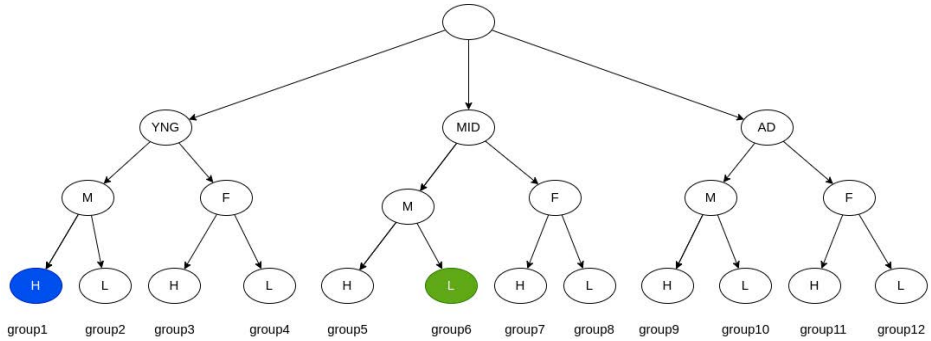

Figure (4) User groups created by the proposed algorithm

In the proposed method, the formula of relation one is used to calculate the distance between two groups (target user and neighboring user).

Equation 5:

$$, u_k) = 1 - \frac{1 - D(u_i , u_k)}{\Sigma(G)} W(u_i$$

In this relation:

$W(u_i, u_k)$ is equal to the group distance score between the target user and the neighboring user k

$\Sigma(G)$ is equal to the number of groups

$D(u_i, u_k) = |G_i - G_k|$ is equal to the distance between group i and group k

In book similarity, the average score is calculated between all the books of the neighbor user and the target book that the neighbor user did not rate. To calculate this similarity, the category and subcategory of the neighboring user books and the target book are used, so that if the neighboring user has a book with both the subcategory and the target book, the system calculates the similarity. To calculate this similarity, Equation 6 is used:

Equation 6:

$$,\beta) = \frac{\frac{1}{2} * \Sigma_{i=1}^{m} (\Sigma_1^l b_{i,l} \cap \beta_L) + (\Sigma_1^s b_{i,S} \cap \beta_S)}{m} \quad W(u_k$$

In this equation:

$W(u_k,\beta)$ is equal to the average of the target book in the neighboring book list
M: is equal to the number of neighboring user books
Bi,l: is equal to the fruit of l's book i
Bi,s: is equal to the s subdivision of the i book
β: equal to the target book
Uk: is equal to the neighboring user
β_S : is equal to the subset of the target book
β_L: is equal to the fruit of the target book
$\left(\sum_1^l b_{i,l} \cap \beta_L\right)$ is equal to one when the neighboring user has a book aligned with the target book, otherwise, it is zero.
$\left(\sum_1^s b_{i,S} \cap \beta_S\right)$ is equal to one when the neighboring user of a book has a subcategory with the target book, otherwise, it is equal to zero.
$\left(\sum_1^s b_{i,S} \cap \beta_S\right)$ is checked if $\left(\sum_1^l b_{i,l} \cap \beta_L\right)$ is equal to one.

Finally, the maximum between-group similarity and book similarity is established as group-book similarity.

Up to this point, the group-book similarity has been achieved. The working method of the proposed method is similar to the original algorithm. In the Knn algorithm, the product of the cosine similarity matrix, multiplied by the user score, is calculated when the neighboring user has rated that book. Otherwise, nothing is done but based on the proposed method, the similarity between the target user and the neighbors' books can be obtained, which is the group-book similarity. Therefore, in the proposed method when the neighboring user has not scored, the group-book similarity is calculated and multiplied by the maximum average score of the target user with the neighboring user. According to the Knn algorithm, where the product of the similarity matrix, multiplied by the user score, is divided by the sum of the value of the similarity matrix, the resulting value is divided by the sum of the group-book similarity. Eventually, the final formula concerning the similarity matrix and the group-book similarity matrix for predicting book B for the target user u_i is obtained from Equation 7 as follows:

$$P(u_i,B) = \frac{\sum_1^k S(u_i,u_k) * R(u_{k,B}) + \left(max(W(u_i,u_k) * W(u_k,B)) * max\left(\mu(R_i),\mu(R_k)\right)\right)}{\sum_1^k S(u_i,u_k) + max\left(W(u_i,u_k) * W(u_j,B)\right)}$$

$P(u_i, B)$ Predicting book B for the target user u_i

$S(u_i, u_k)$ The similarity value of the target user with the neighboring user k

$R(u_{k,B})$ The value of the neighboring user's score to book B.

$\mu(R_i)$ Average user weight i

$\mu(R_k)$ The average weight of the neighboring user k

For example, suppose user A has 4 neighbors N1, N2, N3, and N4. The algorithm wants to get the amount of prediction for user A compared to book B. Of user A's neighbors, only N1 and N2 rated Book B. Therefore, for neighbors N3 and N4, group-book similarity should be calculated. For this purpose, the prediction value for user A and book B is calculated from the following formula.

$$P(A,B) = \frac{\begin{array}{c} S(A,N1) * R(N1,B) + S(A,N2) * R(N2,B) + \\ GS(A,N3) * \max(avg\_rate(A), avgrate(N3)) + \\ GS(A,N4) * \max(avg\_rate(A), avg\_rate(N4)) \end{array}}{S(A,N1) + S(A,N2) + GS(A,N3) + GS(A,N4)}$$

This suggestion method will have a higher success rate than the original method of CF (UserKnn) because the neighboring user may not have rated the target book but the neighboring user may have rated the books with the same target book. For example, if a user read and rated foreign literature but Book B, which also deals with foreign literature but has not rated, can be a good suggestion for the target user. That is, based on the tastes of their neighbors and books that the neighbors did not read, the book can be recommended to the target user. This similarity between the user and the item is done based on the categories of the book that exist in the book category of the main category, the subcategory. On the other hand, due to the grouping of users, group similarity will have a great impact on the success of this method.

Algorithm evaluation results:

To calculate the evaluation results of the proposed algorithm along with the basic algorithm with cosine similarity matrix, the number of executions and the number of neighbors were performed as described in Table (1).

Table (1) shows the number of neighbors and implementations of both the proposed and basic algorithms

	اجرای ۱	اجرای ۲	اجرای ۳	اجرای ۴	اجرای ۵	اجرای ۶	اجرای ۷	اجرای ۸	اجرای ۹	اجرای ۱۰
تعداد همسایه	۲۰	۴۰	۶۰	۸۰	۱۰۰	۱۲۰	۱۴۰	۱۶۰	۱۸۰	۲۰۰

Figure (5) shows the performance of the book recommender system model in retrieving recommendations. According to this figure, more than 86.89% of the proposals in the implemented system use group similarity, and 12.78% use book similarity.

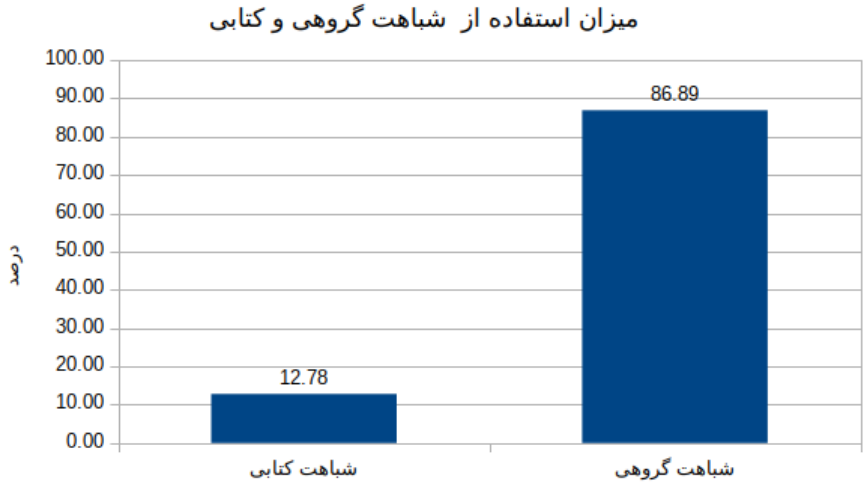

Figure (5) Percentage of group and book similarity use

Examining the MAE diagram between the proposed method and the base method, it is concluded that the Mean Absolute Error of the proposed method is on average 15.89% more improved than the base method.

Figure (6) shows this difference.

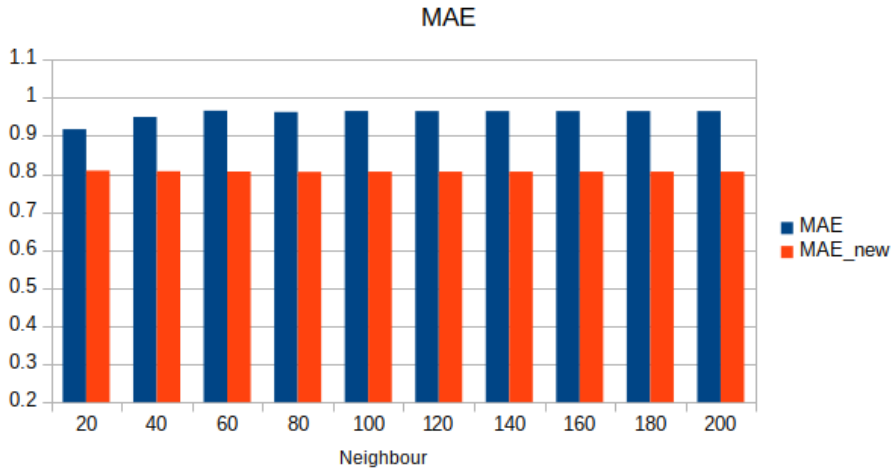

Figure (6) MAE review of the proposed method with the basic method

Also, the studies in RMSE in the proposed method and the basic method show a decrease of 16.28% on average of this evaluation criterion.

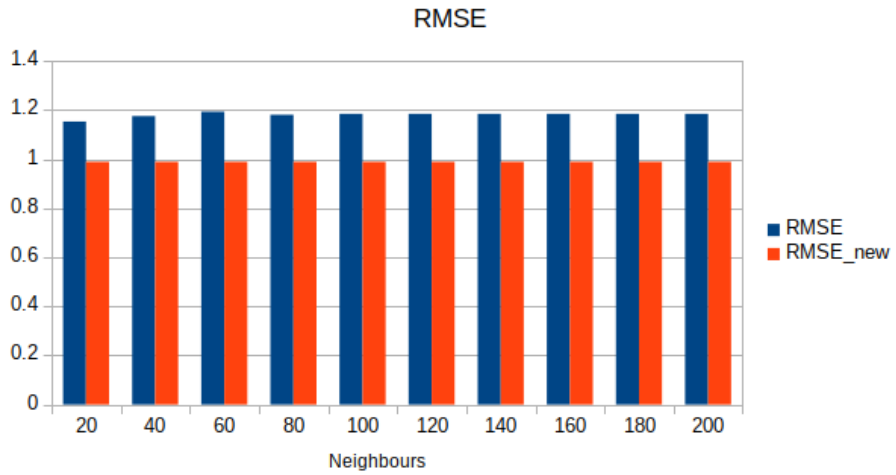

Figure (7) RMSE review of the proposed method with the basic method

Examination of the results from Figure (8) which are related to Precision performance shows that this algorithm has improved 0.1415 more than the basic algorithm.

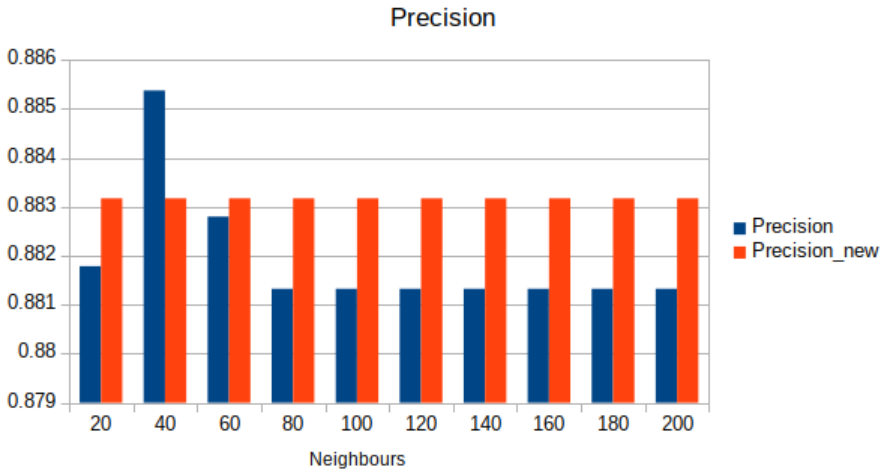

Figure (8): Checking the Precision Value of the Proposed Method with the Basic Method

Discussion:

The large volume and the increasing growth of information and the expansion of scientific productions in addition to the web platform in libraries have also made the process of decision-making and information selection difficult for users and it has caused a problem called data overflow. Recommendation systems, which are referred to as one of the most important information retrieval systems in the current era to eliminate the above challenge, have also entered the field of libraries and are providing recommendations to users. It is expected that the use of book recommendation systems can help to retrieve users' information sources based on their tastes and interests. Using the collaborative filtering approach is one of the most effective approaches to create a book recommendation system. But the use of the CFKnn method is the distinguishing point of this research from other studies proposed in this field. The use of this approach in the current research is significant in two aspects: Firstly, the use of special methods of data analysis in this approach has led to the creation of a book recommendation system with real data sets of books and real users. Secondly, it is possible to look at the data obtained in this research from a different perspective and reuse them. Based on the findings of this research, the characteristics of a user such as the titles of the books read by the user, the rating given to the read books, the

user's interests in different subjects of books, the preference of users in choosing the main criteria of a book, as well as individual characteristics such as age, level of education and gender can be considered as the most important influencing factors in the recommendation. The findings show that the proposed method has better suggestions than the basic method (CFKnn) due to the use of group and book similarity. Also, as stated in the results of this evaluation, the proposed method has been able to improve the result of an important criterion such as RMS compared to the basic method, and also has increased the accuracy of MAE. Although the main goal of this research was to identify the influencing factors on book recommendations and also to discover the existing relationships between the characteristics of the target user and his group users, the collected data shows the fact that there is a need for more specialized investigations and sometimes even with a psychological approach in this field.

Conclusion and Implications :

This research seeks to create a conceptual model of the book recommender system for users in which the real dataset is obtained using online questionnaires. recommender system based on participatory filtering algorithms uses similarity recognition algorithms to form user groups. Using these algorithms, the degree of similarity between two users or the degree of similarity between two items is calculated. The larger the system, the greater the cost of running the algorithm. In this paper, the clustering algorithm is used to classify users by adding a new criterion of group similarity and similarity between the book and the neighboring. This new criterion was calculated ten times with a different number of neighbors and a cosine similarity matrix. The results of this evaluation show that the proposed method has been able to provide better results and has improved the accuracy of criteria such as MAE, and RMSE in the basic method, given that the main purpose of this study was to identify the factors affecting the suggestions of the book recommender system with a focus on user information. Based on the findings of this study, a user's characteristics include the titles of the user's read books, the score he gives to his read books, the user's area of interest in different topics of the books, users' preferences in choosing the main criteria of a book as well as personal characteristics such as age, education, and

gender, and they are assumed as the most influential factors in recommendations. In addition, the collected data indicate the need for more specialized studies, even with a psychological approach in this area.

Research Limitations:

One of the challenges in creating a book recommender system was in the field of user datasets.

Despite the fact that the country's public library system has a relatively good foresight towards the placement of items needed to create a bidding system, the lack of justification of respected librarians for the sensitivity and accuracy of the data has led to the neglect of filling these items in the forms and this dataset cannot be used as a valid reference. Another important challenge in this research is the lack of appropriate datasets in the subject area of books. Because the performance of the recommender systems depends on the subject set of the books, this data set does not exist as a thematic classification of the books and a complete reference. Thematic division of books into 3 levels of category, classification, and scope should be done so that the thematic dataset can be used to conduct this research. For this reason, choosing the topics of the books, considering their level of content, was a very time-consuming and sensitive task. The third major constraint on the research process was the existence of the coronavirus pandemic and the closure of libraries, resulting in the unavailability of users. For this reason, designing an online questionnaire with the minimum items not being boring and being justifiable for the user is a very important task.

Future suggestions and work:

The results of this study showed that the quality of the proposed level of recommender systems has important factors such as the choice of algorithm and dataset.

- Since the volume of data in the dataset has a very decisive role in the selection of the algorithm, in this regard, the suggestions are given as follows:
- Experimental results showed that information retrieval performance in recommender systems is improved by adding effective elements. This finding suggests that further work should be done with more reflection to extract the user's features.

- The test set of the present research includes a questionnaire that was filled in randomly and in a limited way by users; Future research can be developed by increasing the actual dataset of users to perform multiple tests of this algorithm or other algorithms that appear more powerful with the above dataset.

Reference

Abbasi, K. (2018). Comparison of the impact of sentiment analysis and user ratings on the performance of suggestion systems. *The 16th International Conference on Management*.

Bakhshandeh Moghaddam, F. (2013). Recommender systems based on trust:Considering the concept of user expertise. *Graduate University of Basic Sciences-Zanjan.*

Gupta, J. & Gadge, J. (2014). A framework for a recommendation system based on. *Communication and Information Technology Applications (CSCITA),*, 300-304.

Herrera-Viedma, E., & López-Gijón, J. (2013). Libraries' social role in the information age. *Science*, 339(6126), 1382-1382.

Himeur, Y. A.-K. (2021). A survey of recommender systems for energy efficiency in buildings: Principles, challenges and prospects. *Information Fusion,*, 72, 1-21.

Hristakeva, M., Kershaw, D., Rossetti, M., Knoth, P., Pettit, B., Vargas, S., & Jack, K. (2017). Building recommender systems for scholarly information. *In Proceedings of the 1st workshop on scholarly web mining . ACM.*, 25-32.

Jannach, D., Zanker, M., Felfernig, A., & Friedrich, G. . (2011). An introduction to recommender systems. *New York: Cambridge.*

Khamisi, S. a. (2018). presenting a recommender system on the web, based on collaborative filtering algorithm based on item and neural network. *International Conference on Science, Engineering, Technology and Technological Businesses*.

Lu, J., Wu, D., Mao, M., Wang, W., & Zhang, G. (2015). Recommender system application developments: a survey. *Decision Support Systems, 74, 12-32*.

Nasouhi, E. H. (2021). Improving tag-based recommender systems using ontology. *7th International Web Research Conference.*

Niknam, Q. (2015). recommender systems; User assistant in the digital age. *Islamic databases and internet studies*.

Parekh, P., Mishra, I., Alva, A., & Singh, V. (2018). Web Based Hybrid Book Recommender System Using Genetic. *International Research Journal of Engineering and Technology (IRJET)*.

Puritat, K. &. (2021). Development of an open source automated library system with book recommedation system for small libraries. *Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications*.

Ricci, F., Rokach, L., & Shapira, B. (2011). Introduction to recommender systems handbook. *In Recommender systems handbook (pp. 1-35). Springer, Boston, MA.*

Wang, L. C., Zeng, X. Y., Koehl, L., & Chen. (2014). Intelligent fashion recommender system: Fuzzy logic in personalized garment design. *IEEE Transactions on Human-Machine Systems*, 45(1), 95-109.

Wu, D., Zhang, G., & Lu, J. (2015). A fuzzy preference tree-based recommender system for personalized business-to-business e-services. *IEEE Transactions on Fuzzy Systems*, 23(1), 29-43.

Zare, M. S. (2014). providing a solution for improving the recommender system by creating user profiles and optimizing evaluation criteria. *the first national conference of computer engineering research.*.

How to Cite: Azimian, M., Riahi Nia, N., Azimi Vaghar, A., Borna, K. (2023). Analyzing the Requirements for Designing, Implementing, and Evaluating the Book Recommender System and Providing a Conceptual Model for Iranian Digital Libraries, *International Journal of Digital Content Management (IJDCM)*, 4(7), 337-361.
DOI: 10.22054/dcm.2022.68818.1123