Journal of Theoretical and Applied Information Technology
28th February 2022. Vol.100. No 4
2022 Little Lion Scientific

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A HYBRID RECOMMENDER FRAMEWORK FOR SELECTING A COURSE REFERENCE BOOKS

SAMAR SALEH EBRAHEEM ALQALLAF[1], WALAA MOHAMED MEDHAT1[1,2],

Prof.Dr. TAREK AHMED EL-SHISHTAWY[1]

[1]Faculty of Computer and Artificial intelligence, Department of Information Systems, Benha University

13511, Egypt

[2]Information Technology and Computer Science, Nile University, Egypt

E-mail: [1]alqallafsamar@gmail.com, [2]walaa.medhat@fci.bu.edu.eg, [2]t.shishtawy@bu.edu.eg

ABSTRACT

Recommender systems are receiving great attention these days, as various researchers and major companies are conducting continuous research in this field. Companies like Google and Amazon have provided different effective models for video recommendation systems, but the educational field is poorly studied as other researchers explained. Different researchers proposed various approaches showing the challenges related to recommender systems and have proposed various effective recommender systems. This paper aims to propose a hybrid recommender framework that can recommend educational courses' books to study with high accuracy and efficiency. The proposed framework is a hybrid unified approach that helps those who desire to be taught to get suitable books related to a specific course description when a course description is used as an input. This work proposes three different recommendation algorithms for building a hybrid recommendation system. One of the algorithms uses an association rule algorithm to automatically and intelligently guide the end-user to find the most relevant materials.

Keywords: *E-Learning, Recommender System, Association Rule, Data Mining*

1. INTRODUCTION

Recommender Systems could be defined as software tools and techniques providing suggestions for items to be used by a user [1]. Recommender systems principally are directed towards individuals who lack sufficient competence or experience to evaluate the potentially overwhelming number of alternatives that a web site, for example, may offer [2]. This recommendation could be an online activity such as recommending favorite subjects to users, running an online simulation, reading posted messages on a conferencing system or could be simply a web resource [3]. Also, recommendation systems are widely used by big companies like Google; Google has introduced a YouTube recommendation system in 2010 to improve/increase users' search throughput by recommending/suggesting preferred videos based on past search patterns [4].

Recommender systems can be implemented using three different approaches: collaborative filtering approach which collects and analyze user's behavioral information in the shape of their activities, feedback, preferences, and ratings [5]; a content-based approach which grounded on a description of the item and a profile of the user's choices; a hybrid approach which combines and merges different approaches to enhance the overall recommendation results. In recommender systems, different types of data mining algorithms or techniques such as association rule, classification, and clustering algorithms are used to personalize the output data obtained [6].

Big data is a term applied to the data whose type or size is exceeding the ability of traditional relational databases to manage, process, and capture these huge volumes of data [7]. In the realm of Big Data, Machine Learning is used to keep up with the ever-growing and ever-changing stream of data and deliver continuously evolving and valuable insights [8]. An example of the big data is clearly shown in E-learning applications. E-learning is a multi-disciplinary field as the various fields of study interact to constitute it. It accumulates a vast amount of information which is very valuable for analyzing students' behavior and could create a gold mine of educational data [9]. But, one of the biggest challenges is the vast quantities of data these systems

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

can generate daily; as a result, it is very difficult to analyze this data manually [10].

Educational data mining can be explained as data mining application techniques that have a specific type of dataset that comes from educational environments to address important educational questions [11]. Educational data mining applications are applications dealing with the assessment of the student's learning performance, course adaptation, and learning recommendations. There are different algorithms that measure how strong a relationship is between two or more variables together such as Correlation [12].

The statistical measure Correlation is used to indicates how a relationship is a relative between two or more variables. There are several types of correlation coefficients such as the Pearson correlation coefficient. Different programming languages introduced different functions to calculate how strong the relationship is. Scatter plots are identical to Line graphs that show how much one variable is affected by the presence of another and this is considered a correlation. Python has a seaborn function that can achieve this task easily and also has Pandas Corrwith function that is used to calculate the pairwise correlation between columns or rows of two DataFrame objects.

Association rules mining is one of the most well-known data mining tasks based on machine learning algorithms to discover undiscovered relations or the co-occurrence between variables in large databases such as the Apriori algorithm and the FP-Growth algorithm. There are different researchers addressed association rule mining as shown in researches from [12] to [14]

Machine Learning technique involves the process of grouping objects into classes of similar objects. The unsupervised learning algorithm used for exploratory data analysis to find some unrevealed patterns which exist in data but cannot be categorized clearly as shown in researches [15] and [16].

This paper proposes a recommender system in the educational. The significant contributions of this paper are summarized as following: The first contribution is building an efficient recommender system in the educational field that will help those who desire to be taught to get suitable books related to a specific course description when a course description is used as an input. The second contribution is building a new hybrid recommender algorithm for improving the recommendation accuracy in the proposed framework. The third contribution is proposing three different techniques that make a recommendation with high accuracy.

The paper is organized as follows: section two presents other researchers' related works. Section three shows the methodology. Section four tackles the results and discussion while the conclusion is presented in section five.

2. RELATED WORKS

When Many approaches have been recently proposed to address the different models of recommender systems and many other proposed to address the challenges of recommender systems and how it's not easy to know the accurate relation between objects such as books, movies, products, and many different objects. These challenges need to be investigated in order to solve them. Other researchers proposed the use of machine learning techniques to group similar courses and how combining and making a hybrid resolve the challenge of grouping the objects. The work listed in this section help in constructing this paper and build a clear view of what is a recommender system.

In [17] authors explained how retrieving relationships or selecting appropriate objects from whole objects is the main idea behind the recommender system. The authors mentioned the different four types of recommender systems and these types are collaborative filtering, content-based filtering, demographic, and hybrid filtering. The authors have explained the content-based filtering model as the model that uses item specifications to make the recommendation with similar items. They have explained the collaborative filtering model as the model that uses users' past behavior such as purchased, viewed, or rated items to predict other interesting or related items. Also, the authors explained the demographic filtering model as the model where user profile data like gender, educational area, and age were used to find similarities with other profiles to make the recommendation that's besides explaining the hybrid model that combines the previous models to enhance the recommendation accuracy. The authors concluded the study shows how the recommender systems are important and how the data size is increasing rapidly and that raises the importance of big data analysis techniques. The authors recommended using big data algorithms to increase the recommender system performance.

In [18] authors show how it was challenging for the companies that worked online to develop an

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

effective recommender system. The authors proposed an integrated recommender system that identifies the significant nearest neighbors (SNN) and predicts an accurate recommendation at a cheaper computation cost. The authors also, show how the K-nearest neighbor algorithm worked with the similarity matrices and their drawbacks, and how the algorithms that considering the Wight metric are used to select only SNN. The authors concluded the study and explained how the use of unbiased similarity metrics improved aggregate diversity and accuracy. Also, the authors mentioned how the proposed recommender model contributes by meeting customers' interests.

In [19] authors showed how the ample and vast amount of web data leads to an increase in the importance of using recommendations and filtering techniques to find users' information interests. Authors show how the recommender systems are widely used in different fields such as e-learning, e-library, and e-business. The authors mentioned the most used recommendation models are collaborative filtering and content-based filtering and the authors explained each model's pros and cons. They have also shown the importance of using the new recommendation model (the hybrid model). The authors explained the different phases of the proposed recommendation engine including formatting and loading data, dataset formatting, and calculating similarity between the users. Also, the authors show how the Pearson Correlation Coefficient is used to calculate the similarity between the users.

In [20] authors showed how the use of recommendation systems increased after the spreading of online services. The authors mentioned the different challenges for recommender systems such as data sparsity, cold start problems, scalability, and similarity measures and also addressed how matrix factorization, deep learning, or both are used in solving data sparsity challenges. Also, the authors show the improvement to recommender systems because of the focus from big companies like YouTube and Netflix and show how the Netflix recommender system focuses on users' activities like liked videos, favorites, and watched that leads to increase company revenue. The authors showed how the similarity measures are used for determining the rating importance of the neighbor and selecting the nearest neighbor. That is beside the help of the Pearson Correlation Coefficient which enhances prediction performance and accuracy and also the importance of comments sentiment analysis that boosted the recommender system accuracy.

In [21] authors show how the recommender system helps mostly e-commerce users to retrieve related, similar, and relevant services, and goods and on the other hand there is poorly limited information in the education field. The authors made a systematic mapping study and select, analyze, and review 44 paper from 1181 papers. The results showed how the recommender systems are important and can support the educational field using various algorithms and techniques. The authors concluded the study with how their work helped in detecting some key areas and research gaps and recommended with more studies in introducing the artificial intelligence and data mining in educational recommender systems that improve the educational overall.

In [22] authors used a hybrid recommender model to resolve the recommender systems challenges using ontology and clustering. The authors showed how using the ontology is also challenging because replacing original terms with ontology terms sometimes caused information loss and can bring noise into the dataset. The authors solved these challenges by introducing a hybrid recommender model that combined ontology and frequent item clustering. Also, the authors made the document text clustering by using terms found in WordNet for adding or replacing terms by concepts. The authors proposed an enhanced k-means algorithm that used the WordNet hypernyms knowledge by enhancing the bag of terms used before the clustering process.

In [23] authors proposed a hybrid e-learning recommender system that combined association rule mining and clustering algorithm in order to recommend related university courses to students. The authors showed how the clustering was used in their work to group students based on the course grades and select the nearest neighbor group that related and most similar to the target student. Also, the authors showed how the association role mining was used to recommend the course to the target student.

3. METHODOLOGY

The proposed framework is an educational recommender system that takes a specific course description as an input and returns a set of suitable books related to the description of the course that was entered. The proposed recommender has three modules the first one is a Content-based recommender that takes the course description as an input and returns (**BooksList**) that appended to transaction tables (**TransactionsSetList**) that has all the items sets results. The second module is a collaborative Filtering recommender that takes each

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

item in BooksList (Content-Based result) as an input and returns another list of books that are also appended in (**TransactionsSetList**). The third module is an Association role recommender that uses all the transactions in (**TransactionsSetList**) as an input and returns top rules that previously appeared together and save them in **TopRules** class. Using the above modules we build the hybrid recommender and calculate the accuracy for the final result that returned to the user. The proposed framework follows the Algorithm (1) below that show how the system works and Figure (1) that shows the flowchart for the proposed recommender.

Algorithm 1: Framework Execution Flow
i. Course description as input ii. Preprocessing i. Clean text, remove stop words, tokenization and stemming ii. TF-IDF iii. ContentBase (course description) => return BooksList [book1, book2, book3, book4, …] i. Append results in TransactionsSetList iv. For book in BooksList i. CollaborativeFiltering ([book, book rate, user profile]) => return CF_Results[book, book, ….] ii. Append results in TransactionsSetlist v. AssociationRole (TransactionsSetList) => return TopRules vi. Merge the three results and eliminates the redundancy then build the hybrid recommender. vii. Calculate Accuracy viii. Recommend

3.1 Content Based Recommender

The content-based recommender uses the course description as input and follows these steps: loading data (course description and other needed data), creating TF-IDF Vectorizer, calculating the relevance or similarity, and in the end, the recommendation results will be ready to be appended in **TransactionsSetList** to be used with the other two modules (Collaborative Filtering and Association rule). The Content-Based follows steps in Algorithm (2) that shows how it's works and Figure (2) that shows the flowchart.

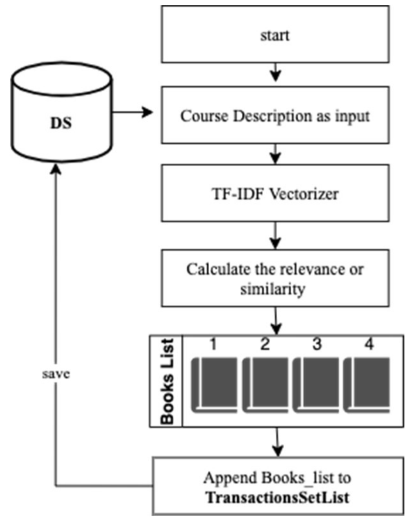

Figure 2: Content Based Recommender

Algorithm 2: Content Based Execution Flow
i. Course description as input ii. Create a TF-IDF Vectorizer(tfidf_rec_matrix). iii. Calculate the relevance or similarity i. Calculating Cosine Similarity iv. Recommendation and append results in class (**TransactionsSetList**)

As mentioned above in Algorithm (2) the Content Based recommender has different phases that built this recommender, sections from 3.1.1 to 3.1.3 below contains each phase and its description.

3.1.1 TF_IDF

The TF*IDF is used to weigh a keyword in any course description and assign the importance to that keyword based on the number of times it appears in the document and stores it in tfidf_rec_matrix. The TF (term frequency) of a word is the number of times it appears in a document. When it's calculated allows seeing if the term is used too often or too infrequently TF(x) = (Number of times term x appears in a document) / (Total number of terms in the document). The term IDF (inverse document frequency) the measure of how that term is significant in the whole corpus IDF(x) = log_e(Total documents number / Documents with term x in its number). The tfidf_rec_matrix is the matrix containing each word and its TF-IDF score concerning each book. Now, the representation of every book in terms of its description is ready. The relevance or similarity of one course to book is calculated using the cosine similarity.

3.1.2 Vector Space Model

The In the Vector space model, all items is stored as a vector creating an n-dimensional space angles that

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

determine the similarity between the vectors. **Cosine similarity:** The main purpose of using cosine is that the value of cosine will increase as the decrease of the value of the angle between vectors as shown in Figures (3), which means more similarity. The cosine similarity of each item with every other item in the dataset is calculated and arranged according to their similarity with the item i (the results are stored). The recommender system calculates cosine similarity based on the following Formula (1):

$$Similarity(Book_x, Book_y) = cos(\theta) = \frac{Book_x.Book_y}{||Boo\ \ _x||||B00k_y||} \quad (1)$$

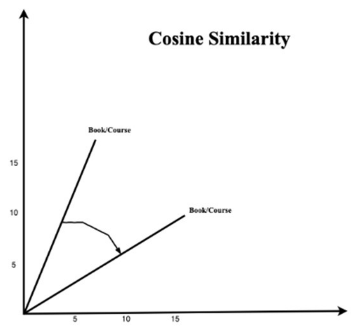

Figure 3: Cosine Similarity example

3.1.3 Recommendation

The recommendation is achieved using a function that takes the course description and other needed data as input to get the related books to the desired course. The results are appended in (TransactionsSetList) to be used with the Association rule module.

3.2 Collaborative Filtering Recommender

The collaborative filtering recommender is used when two users' profiles are similar to each other. Collaborative Filtering recommender must follow these steps: loading data (user ratings and other needed data), creating a data frame, applying Pearson correlation coefficient, getting the number of ratings for each book, relationship checker, and recommending and appending the results in (TransactionsSetList) that will be used with the next module (Association rule). All details are shown in the below Algorithm (3) and Figure 4.

Algorithm 3: Collaborative Filtering Execution Flow
i. **Collaborative Filtering** ([book, Book ratings, user profile]) => i. Create a data frame ii. Pearson correlation coefficient measure iii. Get the number of ratings i. create number of ratings column ii. set threshold for the minimum number of ratings iv. Relationship checker v. Return list of books and append it in (**TransactionsSetList**)

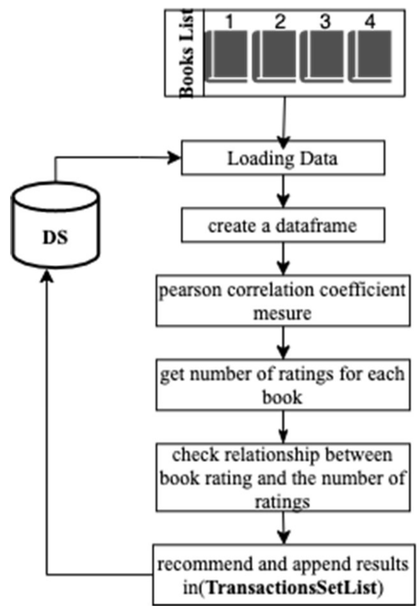

Figure 4: Collaborative Filtering Recommender

As mentioned above in Algorithm (3) the collaborative filtering recommender has different phases that built this recommender, sections from 3.2.1 to 3.2.5 below contains each phase and its description.

3.2.1 Data frame

We create a data frame with the average rating for each book and the number of ratings. These ratings are used to calculate the correlation between the books later

3.2.2 Correlation

Pearson correlation coefficient is used and as is mentioned in the introduction section; Correlation is a statistical measure that indicates how strong a relationship is between two or more variables together. Books that have a high correlation coefficient are the books that are most similar to each other. The Pearson correlation coefficient number will lie between 1 and -1. Minus one correlation coefficient means a negative correlation, for example, the power in someone decreases in

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

(almost) perfect correlation with running. While 1 indicates a positive linear correlation. For example, shirt sizes go up in (almost) perfect correlation with body size. Zero means that: books with a zero correlation are not related at all.

3.2.3 Book Rating

Getting the number for rating for each book by creating a number_of_ratings column that helps on making the relationship obvious between the average rating of a book and the number of ratings the book got. This number is critical because some times a 10-star book was rated only once so it's statistically wrong to classify that book, therefore, a need to set a threshold for the minimum number of ratings is also critical.

3.2.4 Relationship Checker

Relationship checker usage is to check the relationship between the rating of a book and the number of ratings. This process was coded with the help of using plotting a scatter plot using the jointplot() seaborn function mentioned in the introduction section.

3.2.5 Recommend

Recommend is the process where the recommendation happens and to make the recommendation the dataset needs to be converted into a matrix. The matrix is used to measure/compute the correlation between the ratings of a single book and the rest of the books in the matrix. This process was coded with the help of using the pandas pivot_table utility mentioned in the introduction to create the book matrix. The recommendation finishes through computing the correlation between two dataframes and this process was programmed with the help of pandas corwith functionality that computes the pairwise correlation of rows or columns of two dataframe objects.

3.3 Association Role Recommender

Association Role Recommender calculates the similarity between the objects to enhance the poorly understood rules by integrating rule grouping into association mining to enhance the management, distribution and retrieval of the relevant books. Association Role Recommender starts after completing both content-based and collaborative filtering recommenders because it uses all the transitions whose relation matches the books that were founded in these approaches. This recommender stores its results in (TopRules) and it follows the details that shown in the Algorithm (4) below.

Algorithm 4: Association Role Execution Flow

i. **TransactionsSetList as input**
ii. **TransactionsSetList encode**
iii. **AprioriAlgorithm()**
iv. $Support(book_x, book_y) = \frac{N(book_x \cdot book_y)}{N}$
v. $Lift = \frac{N(book_x, book_y)}{N(book_x) * N(book_y)}$
vi. $Confidence(book_x -> book_y) = \frac{N(book_x, book_y)}{N(book_x)}$
vii. Return **TopRules**

4. Dataset

The proposed framework used three datasets, the first dataset (BFCI) was created specifically for my university recommendation system (Faculty of Computer and Artificial Intelligence, Benha university), and the other two datasets (Goodreads, Book-croosing) were used only to test the accuracy of the proposed framework. The books' description was required to test the proposed work with the other two datasets, so the books' description was extracted from Goodreads' website and what we got has been saved into both datasets (Goodreads, Book-Crossing) in books table.

4.1 BFCI Dataset

The BFCI dataset was created specifically for the hybrid proposed framework, and it consists of five tables: courses, books, ratings, students, and transactions. Table (1) below show how the BFCI dataset tables looks like.

Table 1: BFCI dataset tables description

Table	Description
courses	This table has basic information about the faculty courses and its consists of two columns (course_name, course description). The dataset contains 14 course in total and all courses are in the field of computer science.
books	This table contains basic information about the books the faculty studies, and consists of 5 columns (book_id, book_name, book_description, total_ratings, grade). The dataset contains 100 books in total and all books are in the field of computer science.
ratings	This table has three attributes (book_id, student_id, rating) the rating attribute is from 1 as a minimum rating to 10 as a maximum rating. Both user IDs and book IDs are contiguous and all users have rated at least twice.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table	Description
students	This table contains basic information about the faculty students, and consists of 13 columns (student_id, student_name, student_address, student_email, etc.). The dataset contains 500 student in total
transactions	This table contains all the previous results that came from running the content-Base and collaborative filtering modules, it stores all the books that returned to the end user as a recommendation in order to build the third module (AssociationRule)

4.2 Goodreads Dataset

The Goodreads dataset consists of five tables: books, ratings, book_tags, and to_read. The dataset contains six million ratings for ten thousand books. Table (2) below show how the Goodreads dataset tables looks like.

Table 2: Goodreads dataset tables description

Table	Description
books	This table has twenty-three attributes (id, book_id, best_book_id, work_id, books_count, isbn, authors, etc.), and has each book metadata that extracted from Goodreads XML files. Each book may have many editions. bestbookid and goodreadsbookid generally goodreads work_id refers to the book in the abstract sense while bestbookid point to the most popular edition of a given book. As mentioned in the above paragraph this table has been modified and the unuseful attributes were eliminated and each book' description has been added using Amazon API under a new attribute (book_description).
ratings	This table has three attributes (book_id, user_id, rating) the rating attribute is from 1 as a minimum rating to 5 as a maximum rating. Both user IDs and book IDs are contiguous and all users have rated at least twice.
book_tags	This table has two attributes (tag_id, tag_name) and the table contains tags/genres/shelves specified by users to books. Tags in this table are represented by their IDs.
to_read	This table translates tag IDs to names.

4.3 Book-Crossing Dataset

The dataset was collected by Cai-Nicolas Ziegler in 2004, he extracted it in 4 weeks starting August to September from the Book-Crossing [ref] community after getting permission from the community. It holds anonymized 278,858 users that rated about 271,379 books with a total of 1,149,780 ratings. The Book Crossing dataset contains three tables and Table (3) below show how the Book Crossing dataset looks like.

Table 3: Book-Crossing dataset tables description

Table	Description
BX-Users	Contains the users. Demographic data is provided (`Location`, `Age`) if available. Otherwise, these fields contain NULL values.
BX-Books	Contains books that are identified by their respective book_ISBN and these ISBNs have been check and the invalid ones were removed from the dataset. Besides a few content-based information was given such as `Year-Of-Publication`, `Book-Title`, `Publisher`, and `Book-Author`. As mentioned above this table has been modified and the unuseful attributes were eliminated and each book's description has been added using Amazon API under a new attribute (book_description). Also, the book_id attribute was added and this was done by copying the book_id from the Goodreads dataset using the corresponding book_ISBN.
BX-Book-Ratings	contains all information about book rating. Book-Rating (ratings) is a scale from 1-10 higher values that refer to a higher appreciation.

5. RESULTS AND DISCUSSIONS

In the proposed framework each module was tested separately then the three algorithms were tested as one framework at the end to make the idea clear and show different framework parts clearly. Firstly, a sample from the BFCI dataset will be printed to show how it looks like. Table (4) below shows a computer science related courses sample from courses table.

Table 4: BFCI Courses Table Sample

Course	Course Description
Operations Research for Engineers	Operations research helps in solving problems in various environments that needs decisions. etc
Machine Learning	This course provides a broad introduction to statistical pattern recognition and machine learning. Including: supervised learning etc

5.1 Content-Based Recommender

The content-based recommender follows the flow mentioned in the framework architecture section; This recommender steps were getting the data, Create a TF-IDF Vectorizer, calculate the relevance or similarity and finally recommend. Table (5) shows the results when we chose the “Machine Learning” course as an input.

Table 5: Content_Based Recommender input output sample

Input	Machine Learning course description
Output	Books Related to Machine Learning Course are : ['Data Mining: Practical Machine Learning Tools and Techniques', 'Python Machine Learning: Machine Learning and Deep Learning with Python scikit-learn and TensorFlow 2', 'Understanding Machine Learning', 'Machine Learning in Action', …]

5.2 Collaborative Filtering Results

Collaborative Filtering Recommender follows the flow mentioned in the framework architecture section; Collaborative Filtering recommender steps were getting the data, create a dataframe, calculate the correlation, similarity, and finally recommend. Table (6) shows the results when we chose the "Machine Learning" course as an input.

Table 6: Collaborative Filtering Recommender input output sample

Input	BooksList (Content_Based Recommender output)
Output	Books Related to BooksList are : [['Machine Learning for Dummies', 'Machine Learning in Action', 'Python Machine Learning: Machine Learning and Deep Learning with Python scikit-learn and TensorFlow 2', 'Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms Worked Examples and Case Studies' …], ['Python Machine Learning: A Technical Approach to Machine Learning for Beginners', 'Machine Learning for Dummies', 'Bayesian Reasoning and Machine Learning', 'Python Machine Learning: Machine Learning and Deep Learning with Python scikit-learn and TensorFlow 2',…], ['Python Machine Learning: A Technical Approach to Machine Learning for Beginners', 'Machine Learning for Dummies', 'Python Machine Learning: Machine Learning and Deep Learning with Python scikit-learn and TensorFlow 2', 'Understanding Machine Learning', 'Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms Worked Examples and Case Studies', …], […], […], …]

5.3 Evaluation Metrics

In the proposed framework each module was tested separately. The regularly used evaluation metrics are recall, f-measure, and precision which are used for testing the efficiency of the system. Recall, Precision, balanced F-score measures are used for measuring the effectiveness of the information retrieval system. The recall is the ratio of the number of the relevant retrieved documents to the total number of relevant documents in the dataset. While precision is the ratio of the number of relevant retrieved documents to the total number of documents. The Recall and Precision of retrieved items are shown in Table (7) below. Precision, recall, F-Measure uses the Equations (2), (3), (4) as shown below. These accuracy results were calculated Manually after running the three recommenders (100 times in total and taking the average) and getting the relevant and unrelated books for only 13 courses and 100 books from the BFCI dataset.

$$\boldsymbol{Precision} = \frac{|\boldsymbol{relevantDocuments} \cap \boldsymbol{retrievedDocuments}|}{\boldsymbol{retrievedDocuments}} \quad (2)$$

$$\boldsymbol{Recall} = \frac{|\boldsymbol{relevantDocuments} \cap \boldsymbol{retrievedDocuments}|}{\boldsymbol{relevantDocuments}} \quad (3)$$

$$F - Measure = 2.\frac{precision.recall}{precision+rec} \quad (4)$$

Table 7: Accuracy Results

Recommender	Precision	Recall	F-Measure
Content-based	.75	.68	.71
Collaborative	.78	.70	.74
Hybrid	.86	.80	.83

The hybrid recommender helps those who desire to be taught to get suitable books related to a specific course description when a course description is used as an input. The hybrid recommender achieves the best F-Measure results with 83% as shown in Figure (5) below.

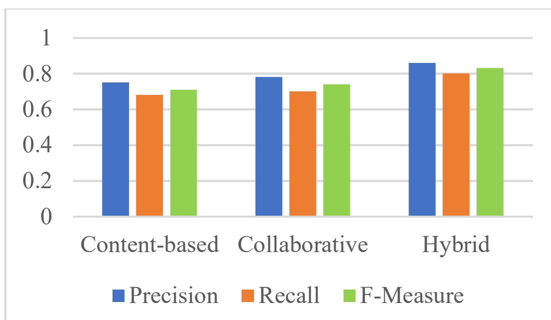

Figure 5: Proposed System Accuracy Results

6. CONCLUSION

Recommender systems are receiving great attention these days, as various researchers and major companies are conducting continuous research in this field. The proposed framework is a hybrid unified approach that helps those who desire to be taught to get suitable books related to a specific course description when a course description is used as an input. This work proposes three different recommendation algorithms for building a hybrid recommendation system. One of the algorithms uses an association rule algorithm to automatically and intelligently guide the end-user to find the most relevant materials. The hybrid recommender achieves the best F-Measure results with 83%.

REFERENCES:

[1] Bansal, Saumya & Baliyan, Niyati. (2019). A Study of Recent Recommender System Techniques. International Journal of Knowledge and Systems Science. 10. 13-41. 10.4018/IJKSS.2019040102.

[2] Ricci F., Rokach L., Shapira B. (2011) Introduction to Recommender Systems Handbook. In: Ricci F., Rokach L., Shapira B., Kantor P. (eds) Recommender Systems Handbook. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-85820-3_1

[3] Prasad, RVVSV. (2012). A Categorical Review of Recommender Systems. International Journal of Distributed and Parallel systems. 3. 73-83. 10.5121/ijdps.2012.3507.

[4] Kumar, Vikas, et al. "Collaborative filtering using multiple binary maximum margin matrix factorizations." Information Sciences 380 (2017): 1-11.

[5] Debashis Das, Laxman Sahoo,Sujoy Datta." A Survey on Recommendation System". International Journal of Computer Applications, vol 160(7), February 2017, pp.6-10

[6] Ahmed Memon, Mashooque & Soomro, Safeeullah & Jumani, Awais & Kartio, Muneer. (2017). Big Data Analytics and Its Applications. Annals of Emerging Technologies in Computing. 1. 10.33166/AETiC.2017.01.006.

[7] Hadioui, Abdelladim & El Faddouli, Nour-eddine & Yassine, Benjelloun & Bennani, Samir. (2017). Machine Learning Based On Big Data Extraction of Massive Educational Knowledge. International Journal of Emerging Technologies in Learning (iJET). 12. 151. 10.3991/ijet. v12i11.7460.

[8] R. Baker, K. Yacef, The State of Educational Data Mining in 2009: A Review and Future Visions, Journal of Educational Data Mining, Article 1, Vol 1, No 1, Fall 2009.

[9] El-Seoud, Samir & Taj-Eddin, Islam & Seddiek, Naglaa & Mohamed, Mahmoud & Nosseir, Ann. (2014). E-Learning and Students' Motivation: A Research Study on the Effect of E-Learning on Higher Education. International Journal of Emerging Technologies in Learning (iJET). Vol. 9. 20-26. 10.3991/ijet.v9i4.3465.

[10] Adam, Khalid. (2018). Big Data and Learning Analytics: A Big Potential to Improve e-Learning. Advanced Science Letters. 24. pp. 7838-7843(6). 10.1166/asl.2018.13028.

[11] Aldowah, Hanan & Al-Samarraie, Hosam & Fauzy, Wan. (2019). Educational Data Mining and Learning Analytics for 21st century higher education: A Review and Synthesis. Telematics and Informatics. 10.1016/j.tele.2019.01.007.

[12] Zhang, Huaifeng & Zhao, Yanchang & Cao, Longbing & Zhang, Chengqi. (2008). Combined Association Rule Mining. PAKDD08. 1069-1074. 10.1007/978-3-540-68125-0_115.

[13] Zhan, Foxiao & Zhu, Xiaolan & Zhang, Lei & Wang, Xuexi & Wang, Lu & Liu, Chaoyi. (2019). Summary of Association Rules. IOP Conference Series: Earth and Environmental Science. 252. 032219. 10.1088/1755-1315/252/3/032219.

[14] Jayshree Jha & Leena Ragha2. (2013) Educational Data Mining using Improved Apriori Algorithm, International Journal of Information and Computation Technology. ISSN 0974-2239 Volume 3, Number 5 (2013), pp. 411-418.

[15] K. Bindra and A. Mishra, "A detailed study of clustering algorithms," 2017 6th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, 2017, pp. 371-376.

[16] Gupta, Tanvi & Panda, Supriya. (2019). A Comparison of K-Means Clustering Algorithm and CLARA Clustering Algorithm on Iris Dataset. 7. 4766-4768. 10.14419/ijet. v7i4.21472.

[17] Mohamed, Marwa & Khafagy, Mohamed & Ibrahim, Mohamed. (2019). Recommender Systems Challenges and Solutions Survey 10.1109/ITCE.2019.8646645, International Conference on Innovative Trends in Computer Engineering (ITCE'2019), Aswan, Egypt, 2-4 February 2019.

[18] Bag, Sujoy & Ghadge, Dr Abhijeet & Tiwari, Manoj. (2019). An integrated recommender system for improved accuracy and aggregate diversity. Computers & Industrial Engineering. 130. 187-197. 10.1016/j.cie.2019.02.028.

[19] Tanya Maan, Shikha Gupta, Dr. Atul Mishra." A Survey On Recommendation System",international conference on recent innovations in management,engineering,science and technology (RIMEST 2018), (2018).pp:543-549.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[20] Low, Jia-Ming & Tan, Ian & Ting, Choo-Yee. (2019). Recent Developments in Recommender Systems. 10.1007/978-3-030-33709-4_4.

[21] Rivera, Abdon & Tapia-Leon, Mariela & Luján-Mora, Sergio. (2018). Recommendation Systems in Education: A Systematic Mapping Study. 10.1007/978-3-319-73450-7_89.

[22] Hotho, A., S. Staab and G. Stumme, 2003. Wordnet improves text document clustering. Proceedings of the SIGIR 2003 Semantic Web Workshop, July 28-August 1, 2003, Toronto, Canada, pp: 541-544.

[23] Amer Al-Badarenah, Jamal Alsakran an Automated Recommender System for Course Selection, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 7, No. 3, 2016.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

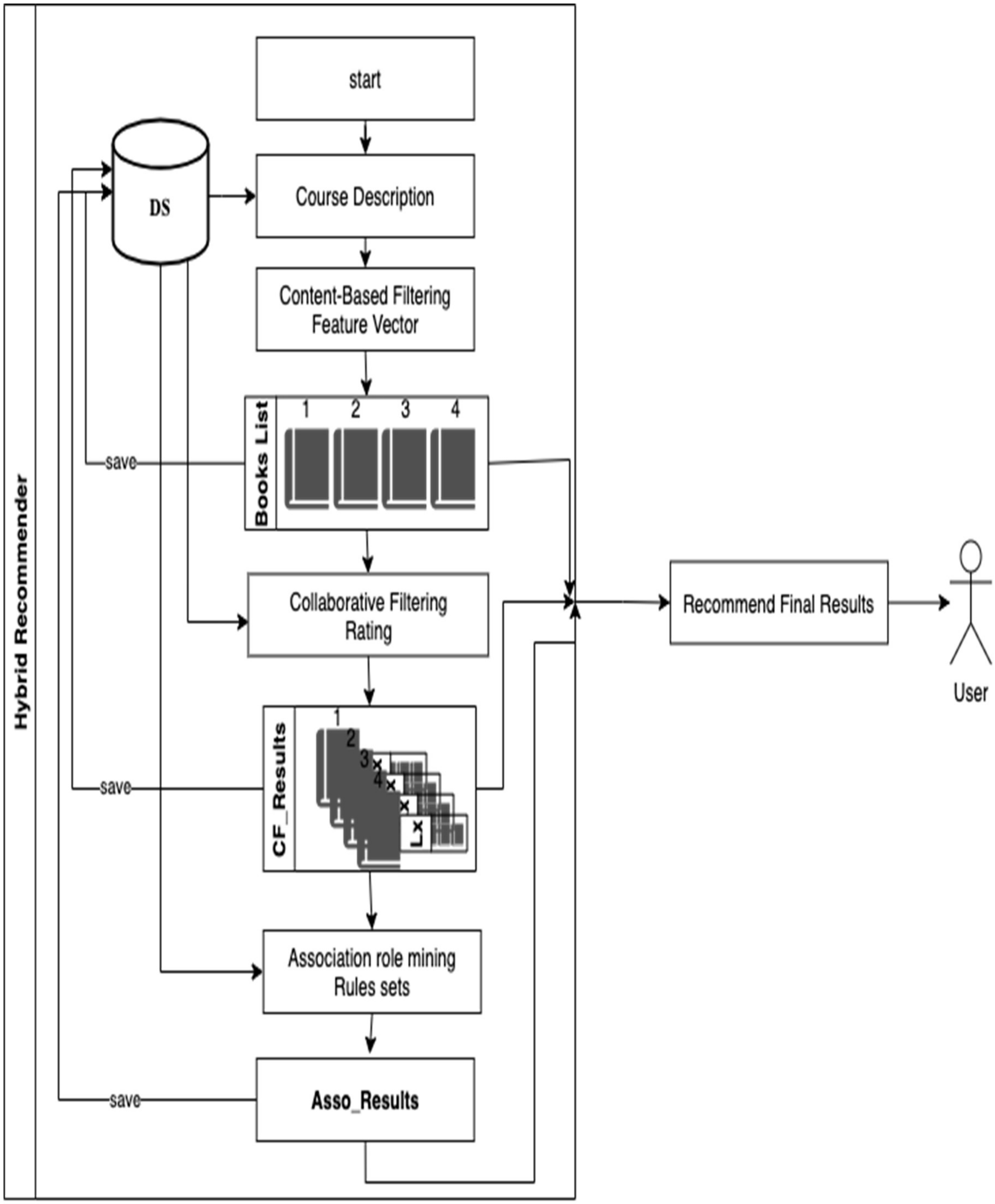

Figure 1: Hybrid Recommender Flowchart