Application of Collaborative Filtering and Explainable AI Methods in Recommendation System Modeling to Predict MOOC Course Preferences

Mugi Praseptiawan
Informatics Engineering
Institut Teknologi Sumatera (ITERA)
Bandar Lampung, Indonesia
mugi.praseptiawan@if.itera.ac.id

Nabila Muthia Putri
Informatics Engineering
Institut Teknologi Sumatera (ITERA)
Bandar Lampung, Indonesia
*Corresponding author:
nabila.120140023@student.itera.ac.id

M. Fikri Damar Muchtarom
Informatics Engineering
Institut Teknologi Sumatera (ITERA)
Bandar Lampung, Indonesia
mfikri.120140077@student.itera.ac.id

Mohd Hafiz Zakaria
Information and Communication Technology
Universiti Teknikal Malaysia Melaka
Melaka, Malaysia
hafiz@utem.edu.my

Ahmad Naim Che Pee
Information and Communication Technology
Universiti Teknikal Malaysia Melaka
Melaka, Malaysia
naim@utem.edu.my

2024 2nd International Symposium on Information Technology and Digital Innovation (ISITDI) | 979-8-3503-8924-1/24/$31.00 ©2024 IEEE | DOI: 10.1109/ISITDI62380.2024.10797073

***Abstract*—Unquestionably, the creation of MOOCs has transformed education by making learning accessible and inexpensive for people all over the world. However, consumers frequently find it difficult to select a course that fits their needs and interests due to the vast curriculum that MOOCs offer. This paper describes how a recommendation system that applies the Collaborative Filtering approach can be used to address these kinds of problems. This paper proposes a recommendation system modeling for MOOC platforms that allows users to be recommended courses based on their preferences. Using user interactions and choices, the Collaborative Filtering approach suggests courses that are specifically tailored to everyone. Utilizing a user-based method called collaborative filtering, users' preferences are predicted by comparing them to other users. The recommendation system's transparency and interpretability are improved by the incorporation of XAI using LIME. By gaining understanding of the reasoning behind course recommendations, users can build trust and make well-informed decisions. A quantitative assessment of the prediction accuracy is provided by the recommendation system's performance evaluation utilizing the RMSE measure. Over time, this statistic aids in system improvement and raises the caliber of course recommendations.**

***Keywords*—MOOC, recommender system, collaborative filtering, explainable AI**

I. INTRODUCTION

There is no doubt that MOOCs are a game changer in the field of education as they provide affordable and readily available learning options to the general public [1]. Suggestions for courses are difficult to enhance because of diverse needs and the rapid growth of MOOC pages. When confronted with this variety, it might be challenging for users to decide which course suits them best in terms of its curriculum due to the overabundance of resources on MOOC pages [2]. This problem often occurs since there is no user review on MOOC pages like APTIKOM or eLOK which can help algorithm recommend other courses.

One possible way of helping people pick courses that suit them well would be to model a recommendation system on MOOC pages [3]. The recommendation system within MOOC pages helps users make choices about similar courses from their top selections. Collaborative Filtering as a model for recommendation systems predicts user preferences based on their interactions with the course. According to this approach, users have the ability to express their short-term desires or requirements alongside long term preferences [4].

Collaborative Filtering applies its techniques to item-based, hybrid, and user-based recommendation system modeling. Combined, the three Collaborative Filtering techniques seek to increase suggestion accuracy by leveraging user interactions and preferences. Al-Ghuribi [5]pointed out that one effective tactic for recommendation systems is user-based Collaborative Filtering [6], [7]. This user-based strategy, which involves generating a similarity matrix for each course based on the user's rating history, centers on the notion that comparable users prefer to favor similar items in the context of MOOC pages. The system can produce recommendations for users based on their interactions with previously loved courses by identifying people who have a comparable sequence of ratings and reviews of a course. This method not only facilitates the discovery of relevant content, but also increases user engagement on the MOOC page [8].

There is a rising need for recommendation systems that produce accurate results and offer clear rationale as artificial intelligence (AI) develops [9]. This is also true in the educational setting, where it is advantageous for students to comprehend the rationale behind the recommendations made for particular courses. Explainable AI (XAI) is included into recommendation systems to give users context for the suggestions they are given. In addition to increasing user trust, this transparency gives consumers the power to decide for themselves by providing information about their past behavior [10]. Users' worries regarding the lack of interpretation of findings in modeling MOOC course recommendation systems can be addressed by integrating XAI with Collaborative Filtering.

XAI seeks to give consumers the ability to take specific actions to impact the suggestions that will be made to them in the future, in addition to providing clarity on the outcomes produced by the system. The user will be able to make better and more appropriate judgments for themselves if they understand why specific recommendations were made for them. This might be the result of the user switching from law to computer science as their preferred course. Any modifications to the user's activity patterns should be communicated to the recommendation system [11].

Because they can concentrate on the numerical value of ratings rather than the quantity of ratings by users, Centered Cosine Similarity or Pearson Correlation calculations are frequently taken into consideration when analyzing the user activity patterns of MOOC recommendation system models with Collaborative Filtering methods. A metric that is simple to understand, Centered Cosine Similarity has values ranging from -1 to 1, with values near 1 denoting strong similarity, 0 denoting no similarity, and -1 denoting dissimilarity [12]. For computing similarity values in user preferences, the Centered Cosine Similarity calculation works well since it is resilient to the scatter that is typical of Collaborative Filtering matrices [13].

Evaluating recommender systems requires choosing the appropriate metric [14]. Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Mean Reciprocal Rank (MRR) are common metrics used. When utilizing Collaborative Filtering techniques to analyze user interactions and model recommender systems, including models using Centered Cosine Similarity, RMSE offers a comprehensive quantitative assessment of performance [15]. By examining the difference between predicted and actual user ratings, the RMSE metric systematically measures the accuracy of predictions.

II. Methodology

At the beginning of the research, it starts with the problem formulation process. Then proceed with the literature study process regarding related research that has been done before. Next, data collection is carried out through datasets from Kaggle, and proceeds to data preprocessing which includes deleting unused data fields or variables and dividing training and testing datasets. From the training dataset, Collaborative Filtering is implemented on the recommendation system built through the calculation of Centered Cosine Similarity to get the similarity value between each user. After that, XAI integration is carried out on the training data that has been processed previously. After passing through these stages, the model will provide output in the form of course recommendations that are predicted as user preferences in taking courses that they have never taken before. The last stage is the implementation of evaluation using the RMSE metric using the testing dataset.

In the process of problem formulation in modeling MOOC course recommendation systems using the Collaborative Filtering method, identification of research topics to be discussed is carried out. At this stage, the author narrows down the topic under study by using the user-based Collaborative Filtering method. The calculation of similarity between each user is done with Centered Cosine Similarity which does not ignore the value by the user. XAI integration is done with the LIME technique so that the system's reasoning in recommending courses can be understood by each user.

In the literature study process, a review of several studies that have been conducted on the same topic is conducted. In this case, the research review must also include objects, methods, and evaluations. Aspects of similar research also include several things used in modeling recommendation systems, such as the use of Collaborative Filtering, Centered Cosine Similarity, XAI, and RMSE methods. The topics reviewed mostly discuss the role of recommendation systems in education, such as learning style recommendations or personalized materials. Although there is already a lot of literature that examines recommendation systems in the field of education, there is a research gap where system modeling has not been equipped with an explainability aspect in each of its decision making. From this gap, XAI integration can be one of the solutions.

In the data collection stage, data collection is carried out through a dataset from Kaggle which is the result of web scraping of Coursera courses. The complete dataset used for this research is provided in the following link: https://www.kaggle.com/datasets/imuhammad/course-reviews-on-coursera?select=Coursera_reviews.csv. The Coursera dataset can also be used for system modeling as an example in other modeling on MOOCs that do not yet have a recommendation system but already have a place for users to rate the courses they have taken.

In the preprocessing stage, unnecessary data fields or labels are removed. For Collaborative Filtering, the required labels are user ID or name, course ID and name, and user ratings for a course. Other than these four variables, other variables in the dataset will be removed. In addition, because the dataset obtained initially consists of two different documents, they are made into one table to make it easier to divide the training dataset and the testing dataset.

The training dataset is the result of dividing the initial dataset with a ratio of 70:30 or 80:20 for the training dataset to the testing dataset. The training dataset will usually be larger than the testing dataset, because the model needs more data to train the system itself, especially to read patterns and user interactions. The larger the training dataset, the more accurate the system training will be.

After the dataset is split into a training dataset and a testing dataset, a similarity calculation is performed using the Centered Cosine Similarity algorithm. The Cosine Similarity algorithm is an algorithm that is quite often used in modeling recommendation systems, but often ignores the numerical value component of the user's assessment of an item. So Centered Cosine Similarity is used which not only calculates the similarity of users through the fact that they rate the same item, but also pays attention to the numerical value of the rating itself.

After the similarity calculation using Centered Cosine Similarity is complete, XAI integration is carried out to complete the recommendation system with a transparent explanation to the user. There are several choices of techniques that can be used for XAI integration in the recommendation system model. Among them are LIME and SHAP. LIME focuses on explaining predictions to each individual, providing a simple and interpretable model. Whereas SHAP offers an understanding of the impact of the system as a whole on the model's predictions. The scope of this MOOC recommendation system modeling is to provide personalized recommendations to each user, so the suitable technique for this modeling is LIME.

Furthermore, the model will provide output in the form of Centered Cosine Similarity calculation results and predictions, as well as an explanation of the prediction results using XAI LIME. The predictions issued by the model are some courses that are deemed suitable as user preferences to be taken in the future. From the results of this prediction, it will be evaluated with a testing dataset that is 30% of the initial dataset.

In the last stage, an evaluation of the recommendation system is conducted based on the performance of the model in providing predictions of MOOC course preferences to users. Some evaluation metrics that are often used in testing the performance of recommendation systems are RMSE and MAE. However, RMSE is closely related to the concept of accuracy and precision. In Collaborative Filtering where the goal is to provide accurate recommendations, RMSE is preferred because it encourages the model to reduce errors and make predictions that are close to the actual user preferences.

III. RESULTS

Recommender systems are very important in today's world with various machine learning models to improve user engagement on websites. Among the methods for modeling recommender systems, the Collaborative Filtering approach is particularly important. We will discuss Collaborative Filtering in depth in this chapter, including its nuances, approaches, and assessment strategies. By leveraging the interaction between users and items, a technique called Collaborative Filtering can provide customized suggestions. This method works on the assumption that people tend to exhibit the same preferences or behaviors in the future if they have done so in the past. Furthermore, we use LIME to investigate the interpretability of the Collaborative Filtering model by examining XAI subjects.

The modeling stage begins with the collection of data with parameters required in the application of the Collaborative Filtering method. We collected relevant data from the selected MOOC page–Coursera, including username data, course attributes, and feedback by users. The dataset obtained from Kaggle is a document with ZIP extension, and includes two datasets with CSV extension. The first dataset contains 622 courses labeled with course title, teaching institution, course URL, and course ID. While the second dataset contains 1.45 million reviews and ratings given by users with labels of review, user name, review date, rating, and course ID. Due to the size of the second dataset that is too large for computation so that it takes a long time, random sampling is carried out to 10,000 reviews.

To guarantee accuracy and consistency, the gathered data was cleaned. Parameters like institution and course_url that are not utilized in the Collaborative Filtering approach are present in the course dataset that was retrieved from Coursera web scraping results. As a result, the table was cleared of both parameters.

Following the removal of the two parameters, the following new table with the course_id and name parameters is obtained (Table II).

In the second dataset containing reviews from users, user IDs are allocated with a unique value for each user. This is done to make it easier to create a pivot table between users and items. The output is 5 initial user IDs, and 5 final user IDs. The total number of user IDs is 264.000 out of a total of 1.45 million reviews which states that many users review at least two courses.

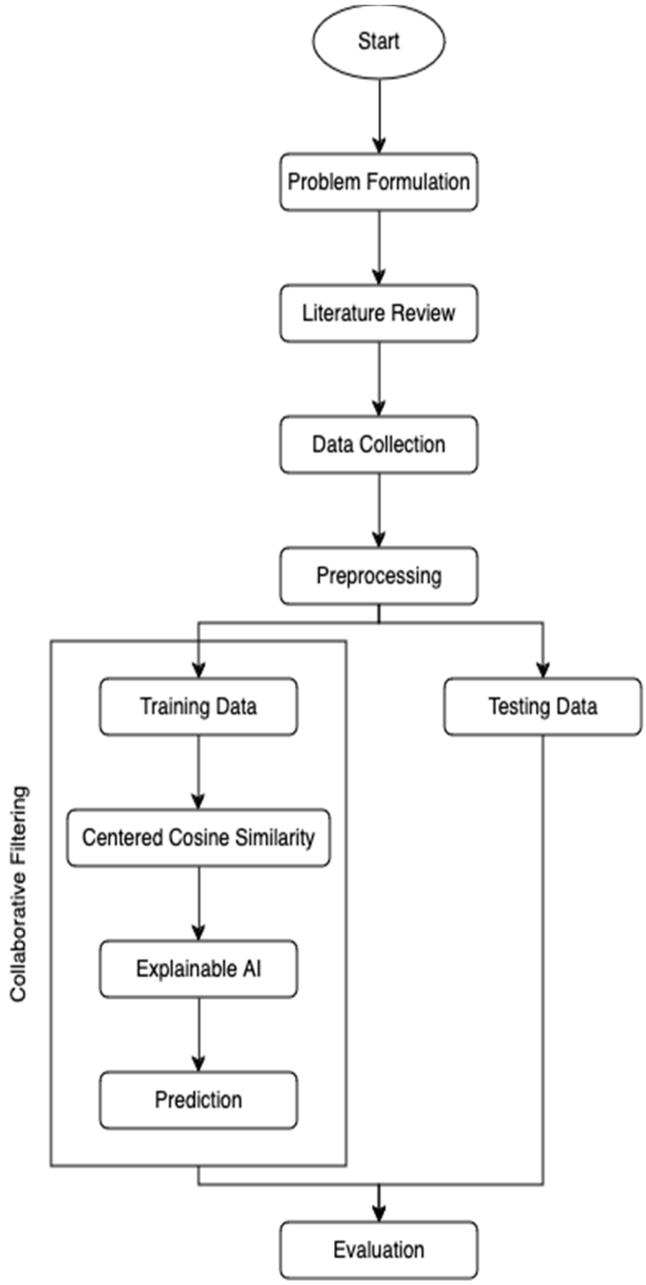


Fig. 1. Research methodology flowchart

TABLE I. COURSERA_COURSES.CSV INITIAL TABLE

	name	institution	course_url	course_id
0	Machine Learning	Stanford University	https://www.coursera.org/learn/machine-learning	machine-learning
1	Indigenous Canada	University of Alberta	https://www.coursera.org/learn/indigenous-canada	indigenous-canada
2	The Science of Well-Being	Yale University	https://www.coursera.org/learn/the-science-of-well-being	the-science-of-well-being
3	Technical Support Fundamentals	Google	https://www.coursera.org/learn/technical-support-fundamentals	technical-support-fundamentals
4	Become a CBRS Certified Professional Installer by Google	Google – Spectrum Sharing	https://www.coursera.org/learn/google-cbrs-cpi-training	google-cbrs-cpi-training

TABLE II. COURSERA_COURSES.CSV TABLE AFTER REMOVING *INSTITUTION* AND *COURSE_URL* PARAMETERS

	name	course_id
0	Machine Learning	machine-learning
1	Indigenous Canada	indigenous-canada
2	The Science of Well-Being	the-science-of-well-being
3	Technical Support Fundamentals	technical-support-fundamentals
4	Become a CBRS Certified Professional Installer by Google	google-cbrs-cpi-training

Next, mapping the reviewers variable which is the user name into a unique user ID is done. After that, unnecessary fields can be deleted, including reviewers, reviews, and date_reviews so that it becomes the following new table.

TABLE III. COURSERA_REVIEWS.CSV TABLE AFTER REMOVING *REVIEWS* AND *DATE_REVIEWS* PARAMETERS

	rating	course_id	user_id
0	4	google-cbrs-cpi-training	1
1	4	google-cbrs-cpi-training	2
2	4	google-cbrs-cpi-training	3
3	4	google-cbrs-cpi-training	4
4	4	google-cbrs-cpi-training	5

The same thing is also done to the course_id variable which originally has a string data type to be allocated to an integer with a unique value. From the result of the allocation, a new variable named course_id_encoded is created, so that the initial course_id variable can be deleted into a new table as follows.

TABLE IV. MAPPING COURSE NAME TO COURSE ID

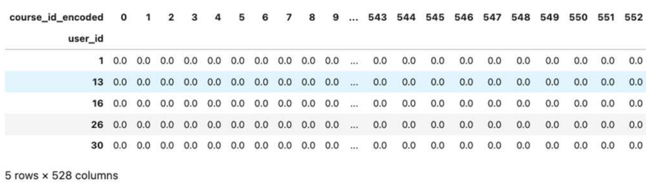
	name	course_id
0	Machine Learning	356
1	Indigenous Canada	283
2	The Science of Well-Being	556
3	Technical Support Fundamentals	550
4	Become a CBRS Certified Professional Installer by Google	259

After allocating the variable, random sampling is carried out for the second CSV into 10,000 tuples. The sampled review dataset was combined with the new course table so that it becomes the following tail table.

TABLE V. PREPROCESSED DATASET

	rating	user_id	course_id_encoded	name
9995	5	170705	443	Privacy Law and Data Protection
9996	2	235682	180	Epigenetic Control of Gene Expression
9997	5	123416	20	Antibiotic Stewardship
9998	3	235402	84	Competencias digitales. Herramientas de ofimtica (Microsoft Word, Excel, Power Point)
9999	5	257335	161	e-Learning Ecologies: Innovative Approaches to Teaching and Learning for the Digital Age

The merged dataset is divided into 7000 training data, and 3000 testing data. Next, the core of the Collaborative Filtering method is the creation of a user-item matrix which is the foundation of the recommendation system. This matrix summarizes the user's interaction with the course represented by ratings. Null values will be replaced with value 0.



course_id_encoded	0	1	2	3	4	5	6	7	8	9	...	543	544	545	546	547	548	549	550	551	552
user_id																					
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
13	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
16	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
26	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
30	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

5 rows × 528 columns

Fig. 2. Pivoted table of user ID and encoded course ID

Then, normalization is performed for each table by reducing each value with the average value of the rating by each user. After that, an ordinary cosine similarity calculation is performed with the following result shape.

```
[[ 1.          0.43961645 -0.00268607 ...  0.70643558  0.70643558
  -0.00268607]
 [ 0.43961645  1.         -0.00266958 ...  0.62379099  0.62379099
  -0.00266958]
 [-0.00268607 -0.00266958  1.         ... -0.00189753 -0.00189753
  -0.00189753]
 ...
 [ 0.70643558  0.62379099 -0.00189753 ...  1.          1.
  -0.00189753]
 [ 0.70643558  0.62379099 -0.00189753 ...  1.          1.
  -0.00189753]
 [-0.00268607 -0.00266958 -0.00189753 ... -0.00189753 -0.00189753
   1.        ]]
Shape of Centered Cosine Similarity Matrix: (6710, 6710)
```

Fig. 3. Matrix of Centered Cosine Similarity calculation result

To get the predicted rating value from the user, dot product is done between the prediction result array and the pivot table so that it becomes the following array.

```
array([[-0.01343033, -0.08595414, -0.05372134, ..., -0.00537213,
        -0.0644656 , -0.08595414],
       [-0.01334788, -0.08542644, -0.05339153, ..., -0.00533915,
        -0.06406983, -0.08542644],
       [-0.00948767, -0.06072106, -0.03795066, ..., -0.00379507,
        -0.0455408 , -0.06072106],
       ...,
       [-0.00948767, -0.06072106, -0.03795066, ..., -0.00379507,
        -0.0455408 , -0.06072106],
       [-0.00948767, -0.06072106, -0.03795066, ..., -0.00379507,
        -0.0455408 , -0.06072106],
       [-0.00948767, -0.06072106, -0.03795066, ..., -0.00379507,
        -0.0455408 , -0.06072106]])
```

Fig. 4. New array after dot products

Users will be given recommendations with the following output, in this case for user ID 12.

```
course_id_encoded
374    558.916495
0       -0.000000
364     -0.000000
378     -0.000000
377     -0.000000
Name: 102, dtype: float64
```

Fig. 5. Course recommendation for user ID 12

To get the clarity aspect about the choice of the system, XAI will be integrated using LIME for tabular data. This process requires clustering using K-Means with 3 clusters for variable X in the training data. The explanation will use 3 features namely user ID, course ID, and rating with the following output results for user ID 12.

```
Intercept 0.7336514155342808
Prediction_local [1.00687087]
Right: 1
```

Fig. 6. Intercept, Prediction_local, and Right value for user ID 12

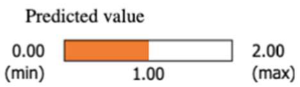


Fig. 7. Visualization of Predicted value

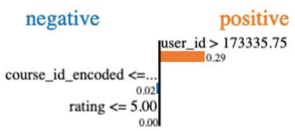


Fig. 8. Visualization of probability for each feature

Feature	Value
user_id	198266.00
course_id_encoded	42.00
rating	4.00

Fig. 9. Values for each feature

To scale the prediction results to ratings from 1 to 5, scaling is done using Reader from the surprise library in Python. This scale is needed for RMSE calculation, with the following results.

```
RMSE: 0.6655612975285078
```

Fig. 10. RMSE evaluation result of the recommender system

IV. Conclusion

This research highlights the importance of AI models that provide transparency so that they can be understood by users. This plays an important role in increasing user trust and engagement with a choice made by the system. Future research directions could explore advanced recommendation techniques, and incorporate additional features for personalization.

In developing a recommendation system model to predict MOOC course preferences for each user using the user-based Collaborative Filtering method, there are methods to ensure that modeling is carried out systematically and continuously.

From the existing dataset, preprocessing has been done to divide the data into training data and testing data. The training data is used to calculate the similarity between one user and another using Centered Cosine Similarity. After performing the calculation, the model can predict users' preferences for MOOC courses that they have never taken before. The testing data is used in the evaluation using RMSE calculation which calculates the prediction result with the original state. The predicted result of the model is considered as the course that will be preferred by the users, while the original state is obtained from the testing dataset that shows the assessment of the courses that have not been taken based on the training data.

When discussing a recommendation system using user-based Collaborative Filtering, it is important to consider the user's rating of the course. Therefore, Centered Cosine Similarity is a suitable choice because it does not treat courses that have not been rated by users as negative. The Centered Cosine Similarity algorithm will calculate the average of the ratings given by each user, then subtract a user's rating from the average. From the normalized value, it can be calculated using Cosine Similarity as usual.

After the model provides a similarity value between each user, it will predict the MOOC course preferences for each user. From this prediction result, XAI integration can be done using the LIME technique with the built-in library from Python. LIME can explain the prediction results from tabular data, image data, and written data. In this modeling, the data used for prediction output is tabular data. The parameters used in creating a sample LIME explainer for tabular data are the training dataset, the prediction mode which can be either classification or regression, a list of column names in the training dataset, and a verbose value. After that, the explainer object that has been previously exemplified should explain the test using samples using the parameters of the testing data row, prediction function, and the maximum number of features used for decision making.

The calculation of the RMSE value in this modeling starts after the output of the MOOC course preference prediction to the user. The prediction results must be used together with the actual assessment of the testing dataset to calculate the squared error value. From the results of the similarity calculation with Centered Cosine Similarity, it is necessary to normalize to get the numerical value of the user assessment between 0.5 to 5. Then, calculate the RMSE by taking the square root of the average squared difference between the predicted and actual assessments.

The combination of Collaborative Filtering and XAI with LIME techniques can create course recommendations according to user preferences with clear information about the decision selection process, which promotes a broader user learning experience on MOOC platforms. Finally, the performance of the system through evaluation between the predicted value and the original value shows a low margin of error. This indicates that the prediction results of the system have high accuracy.

REFERENCES

[1] S. Al Shaqsi and R. T. Syed, "Massive open online courses and entrepreneurship education in higher education institutions," in *Technology and Entrepreneurship Education: Adopting Creative Digital Approaches to Learning and Teaching*, 2022. doi: 10.1007/978-3-030-84292-5_8.

[2] S. Bansal and N. Baliyan, "A study of recent recommender system techniques," *International Journal of Knowledge and Systems Science*, vol. 10, no. 2, 2019, doi: 10.4018/IJKSS.2019040102.

[3] R. Duan, C. Jiang, and H. K. Jain, "Combining review-based collaborative filtering and matrix factorization: A solution to rating's sparsity problem," *Decis Support Syst*, vol. 156, 2022, doi: 10.1016/j.dss.2022.113748.

[4] Z. Li, H. Zhao, Q. Liu, Z. Huang, T. Mei, and E. Chen, "Learning from History and Present: Next-item Recommendation via Discriminatively Exploiting User Behaviors," 2018, doi: 10.1145/3219819.3220014.

[5] S. M. Al-Ghuribi and S. A. Mohd Noah, "Multi-Criteria Review-Based Recommender System-The State of the Art," *IEEE Access*, vol. 7. 2019. doi: 10.1109/ACCESS.2019.2954861.

[6] M. Á. García-Cumbreras, A. Montejo-Ráez, and M. C. Díaz-Galiano, "Pessimists and optimists: Improving collaborative filtering through sentiment analysis," *Expert Syst Appl*, vol. 40, no. 17, 2013, doi: 10.1016/j.eswa.2013.06.049.

[7] W. Zhang, G. Ding, L. Chen, C. Li, and C. Zhang, "Generating virtual ratings from chinese reviews to augment online recommendations," *ACM Trans Intell Syst Technol*, vol. 4, no. 1, 2013, doi: 10.1145/2414425.2414434.

[8] S. Li *et al.*, "Quantification and prediction of engagement: Applied to personalized course recommendation to reduce dropout in MOOCs," *Inf Process Manag*, vol. 61, no. 1, Jan. 2024, doi: 10.1016/j.ipm.2023.103536.

[9] Z. Cui *et al.*, "Personalized Recommendation System Based on Collaborative Filtering for IoT Scenarios," *IEEE Trans Serv Comput*, vol. 13, no. 4, 2020, doi: 10.1109/TSC.2020.2964552.

[10] J. Manyika, J. Silberg, and B. Presten, "What do we do about the biases in AI?," *Harvard Business Review*. 2019.

[11] A. Adadi and M. Berrada, "Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)," *IEEE Access*, vol. 6, 2018, doi: 10.1109/ACCESS.2018.2870052.

[12] A. M. Ridwan, D. R. Ramdania, D. S. A. Maylawati, A. Wahana, E. Mulyana, and M. I. Al-Amin, "Leather Product Recommendation System using Collaborative Filtering Algorithm," in *Proceeding of 2022 8th International Conference on Wireless and Telematics, ICWT 2022*, 2022. doi: 10.1109/ICWT55831.2022.9935395.

[13] R. Bharti and D. Gupta, "Recommending Top N Movies Using Content-Based Filtering and Collaborative Filtering with Hadoop and Hive Framework," in *Advances in Intelligent Systems and Computing*, vol. 740, 2019. doi: 10.1007/978-981-13-1280-9_10.

[14] D. Shin, "How do users interact with algorithm recommender systems? The interaction of users, algorithms, and performance," *Comput Human Behav*, vol. 109, 2020, doi: 10.1016/j.chb.2020.106344.

[15] I. Karabila, N. Darraz, A. El-Ansari, N. Alami, and M. El Mallahi, "Enhancing Collaborative Filtering-Based Recommender System Using Sentiment Analysis," *Future Internet*, vol. 15, no. 7, 2023, doi: 10.3390/fi15070235.