International Journal of Electrical and Computer Engineering (IJECE)
Vol. 15, No. 3, June 2025, pp. 3107~3117
ISSN: 2088-8708, DOI: 10.11591/ijece.v15i3.pp3107-3117

MOERID: enhancing open educational resource discoverability through an artificial intelligence-powered chatbot recommender system

Sandoussi Rima[1], Hnida Meriem[2,3], Daoudi Najima[2,1], Ajhoun Rachida[1]
[1]National Higher School for Computer Science and Systems Analysis (ENSIAS), Mohammed V University in Rabat, Rabat, Morocco
[2]ITQAN Team, School of Information Sciences (ESI), Rabat, Morocco
[3]Mohammadia School of Engineers (EMI), Mohammed V University in Rabat, Rabat, Morocco

Article Info

Article history:

Received Jun 4, 2024
Revised Jan 24, 2025
Accepted Jan 16, 2025

Keywords:

Chatbot
Open educational resources
Recommendation system
Recommender system
Teaching community

ABSTRACT

Open educational resources (OER) are valuable assets in learning and teaching. They ensure cost-effectiveness and customizability, and contribute to global collaboration in the education realm. Hence, education stakeholders face a critical challenge in locating suitable OERs that meet their needs and pedagogical objectives. To cater to this issue, researchers propose diverse digital solutions, one being recommender systems (RS). While a wide array of the suggested tools focused on learners only, this study introduces MOERID, an artificial intelligence (AI)-powered OER recommender Chatbot aimed at the teaching community. It aspires to facilitate resource discoverability, allowing instructors to save time and energy to concentrate on other pedagogical duties. MOERID engages natural language processing (NLP) and recommendation filtering techniques to locate and deliver relevant OERs to instructors. The study describes the implementation of MOERID, highlights its efficacy and provides actionable recommendations for future research by outlining the research gaps.

Corresponding Author:

Sandoussi Rima
National Higher School for Computer Science and Systems Analysis (ENSIAS), Mohammed V University in Rabat
Rabat, Morocco
Email: sandoussi.rima@gmail.com

1. INTRODUCTION

Conforming to UNESCO [1], open educational resources (OER) involves educational, instructional, and research assets available in different formats and mediums. These resources belong to the public domain, or are copyrighted but accessible under an open license. Such licenses make the access unexclusive, with the freedom to reuse, repurpose, modify, and distribute the materials. By providing a diverse range of educational settings and teaching methods, OERs play an inherent role in supporting inclusivity and accessibility in learning. In line with the global goal of improving educational resources, there is a concerted effort to develop, implement, and adapt high-quality OER. This movement encourages educators to collaborate and establish international standards for the quality and accessibility of OER. The focus on producing high-quality OER aims to establish a worldwide standard for educational resources in addition to encouraging collaboration within the teaching community [2].

Nevertheless, navigating through the massive volume of OER is challenging, and finding and discovering relevant ones is becoming complicated for the teaching community. For instance, many instructors spend a significant amount of time online, browsing through educational resource repositories to find materials that can help them construct their courses. To this end, recommendation systems (RS) have grown in popularity, in this field, and are attempting to address this issue by providing instructors with suitable learning materials matching their needs, contexts, and preferences. Recommender systems serve as an effective solution for navigating the search for learning materials within extensive repositories. They suggest items that might interest users by analyzing both their explicit and implicit preferences and those of users with similar profiles, and the characteristics of both users and items [3].

However, OER recommenders intended for the teaching community are way less common than OER RSs designed for the learning community [4], and to address this problem statement, this study presents and discusses an OER RSs chatbot-driven that recommends resources based on the instructors' input. Following a comprehensive comparative analysis of existing technologies, the system integrates cutting-edge Recommendation techniques and natural language processing (NLP) and deploys a sizable dataset of OERs for accurate recommendation. The remainder of the study is organized as follows: the second section provides a comprehensive review of existing literature, highlighting key advancements in the field of OER and RS. Section 3 outlines the methodology, design and implementation of the system. Results of the system's evaluation are discussed in section 4, and section 5 recapitulates the key findings and offers insights in the practical application of the system in teaching practices. Section 6, provides the research contributions, limitations, and future directions, bringing the study to a close.

2. RELATED WORKS

Recommender systems (RS) emerge as valuable information filtering technology. They provide tailored suggestions to users. These systems are based on various approaches, each offering strengths in uncovering items personalized to the user preferences. In the literature, many strategies have been proposed for the integration of RS into educational frameworks, emphasizing their potential to improve the development of courses. These methodologies can be classified into content-based (CBF), collaborative (CF), and hybrid models, which merge elements of the first two approaches. Content-based filtering (CBF) suggests related items based on attributes of previously liked items by a user, while CF recommends items by analyzing the past behavior of users with similar preferences. Hybrid systems generate recommendations by integrating both CF and CBF or more approaches.

Recently, a remarkable focus is paid on implementing large language model (LLM) and machine learning (ML) approaches to improve the efficacy and intelligence of these systems [5]. Table 1 compares (CBF), (CF), and hybrid recommendation systems by providing an overview of their advantages and limitations based on the [6] and [7] studies. CBF recommends similar items to those previously liked, handling new ones, yet, struggling with new users and emerging items. CF recommends items based on similar users' preferences, providing accurate predictions, But, struggling with "cold start," scalability, and data sparsity. Finally, hybrid recommenders combine diverse methods addressing different needs for better accuracy. However, this type of RS is computationally expensive and complex to manage.

Researchers of paper [8] focus on developing a CBF recommender for suggesting elective courses aligned with students' abilities. It deploys preprocessed word items from users' academic history, calculating weighted cosine similarity between elective course syllabi and user profiles. The system is evaluated using questionnaires and validation methods, with results indicating a high level of recommendation diversity and an average accuracy. Julianti *et al.* [9] built a RS to make the process of navigating the right skills for a career in a specific field or domain less complex. This system makes recommendations based on user data and preferences. It automates filtering, allowing users to receive an online personalized experience. this system utilizes CBF and ML.

The potential for streamlining the course production process made the integration of RS within learning management systems (LMS) a subject of interest in the rapidly evolving realm of digital education. The creation of MoodleRec [10], a plugin for the LMS Moodle, is a noteworthy development in this field. It assists teachers in the process of courses' creation by using keyword-based searches to find learning objects (LOs) across repositories. The experiment showed a 61.4% satisfaction rate, with challenges like the cold start problem and limited social feature usage. Future suggestions include non-experimental use, a central server setup, and improvements for more meaningful queries, addressing LO specification variations. The study [11] presents a CBF model for a personalized learning RS. It calculates similarities between learning materials and course descriptions to recommend learning content. The system's goal is to meet learning objectives and prevent dropouts by responding to individual learning needs in online learning environments.

In recent studies, RS are becoming increasingly valuable in education. They have the ability to help instructors choose resources and adjust their teaching practices. Dhahri and Khribi [12] pointed to the significance of these systems in providing personalized recommendations for teaching enhancements, often through hybrid approaches. The systematic review analyzed many scientific papers and found that educational portals and repositories were commonly used environments for gathering educational resources and making recommendations. The study also proposed a teacher recommender system that aims to provide teachers with the most relevant OER retrieved from collections of resources aligned with the UNESCO ICT competency framework for teachers (CFT) [13].

Table 1. Overview of the advantages and limitations of recommendation techniques

	Approach definition Advantages Limitations	Approach definition Advantages Limitations	Approach definition Advantages Limitations
CBF RECOMMENDERS	CBF recommenders use item characteristics, user preferences, and machine learning algorithms to analyze user behavior patterns, recommending products with similar features to those previously positively reviewed.	CBF recommendation systems deploys NLP to analyze text reviews, addressing issues with new products through assigned keywords.	These algorithms exhibit resilience to the cold start problem but manifest a bias toward established categories, often overlooking newly emerging, potentially relevant ones due to their recent introduction.
CF RECOMMENDERS	CF engines use ML algorithms to analyze users' preferences, suggest items based on similar tastes, and gather user data through explicit or implicit data collection. Used massively in E-commerce and movies Recommendations	The ability to accurately predict user interest in unfamiliar products without requiring detailed item descriptions and subsequently improving recommendation accuracy.	CF faces challenges like cold start problem, scalability, rich-get-richer effect, data sparsity, and shilling attacks, affecting recommendation accuracy and enhancing popular items at the expense of new ones.
HYBRID RECOMMENDERS	A combination of the two previous techniques or more	Hybrid recommendation systems enhance user experience by merging various techniques, offering tailored suggestions and better handling new users or items with limited data (cold start issue). They leverage the strengths of multiple methods for more accurate recommendations.	It can be a disadvantage of hybrid systems to need more processing power and resources by requiring high computational complexity. Also, handling an extensive database of ratings from various sources is part of managing a hybrid system.

Researchers in paper [14] presented a quality-driven, ontology-based framework to provide educators with the access to high-quality OERs and assist them make more informed decisions in educational contexts. Furthermore, research [4] demonstrated a growing interest in OER recommenders by proposing a chatbot-based system designed for the teaching community. Herrera-Cubides *et al.* [15] have examined the influence of pedagogical quality scores in OER recommenders, pointing that these scores could improve the selection of high-quality resources without diminishing relevance. These studies jointly underline the remarkable role of advanced RS in enriching educational practices and resource utilization.

In summary, the integration of advanced recommendation systems in learning management systems represents a significant advancement in optimizing digital education. These systems use different methods, like content-based filtering and machine learning, to improve how courses are created and to give learners a more tailored experience. However, there is a gap in literature about the current trends on OER recommenders including limitations and opportunities in teaching practices. In this regard, the following sections report the implementation of an educator-centered Chatbot-driven OER recommender based on CBF techniques.

3. METHODOLOGY

In this paper, we introduce "MOERID" (My Open Educational Resource for Instructional Delivery- the acronym is derived from the Arabic word "resource"), a chatbot-driven RS for OER. Unlike other models that rely on static recommendations, this system dynamically exchanges with educators through a conversational interface, for a better understanding of their preferences leveraging natural language processing (NLP) and machine learning (ML) techniques. In this section we outline the systematic process of MOERID's implementation. The research outlines the system architecture by presenting their main components: Authentication system to secures access using user credentials, chatbot interface as a GUI-based interactive system built using the Tkinter library for user-friendly operation and the recommender engine that Combines NLP and CBF techniques to generate tailored recommendations. The research also highlights the data collection process by evaluating multiple datasets for a comparative analysis to select a dataset with rich metadata that supports semantic similarity computations necessary for CBF. The research also emphasizes data preprocessing by using Metadata Standardization to fill missing values, text cleaning through NLP preprocessing (tokenization, lemmatization) using `SpaCy` to ensure data consistency and feature extraction and vectorization for CBF.

The research depicts the recommendation engine algorithms that are: Word2Vec for embedding textual features and cosine similarity for similarity scoring. For the design and implementation, the system leverages intent and entity recognition. Its workflow includes parsing user input to extract key information, matching user preferences with OER metadata via the recommendation engine and finally presenting a ranked list of OER links to the user and capturing feedback for iterative improvements. To evaluate the system's efficacy, we generated synthetic user preferences to simulate real-world interactions to compute precision, recall, and F1 score for offline evaluation. Thus, to mitigate Cold Start issues, we deployed warm-start techniques to recommend items with limited user history. Figure 1 depicts the detailed methodology:

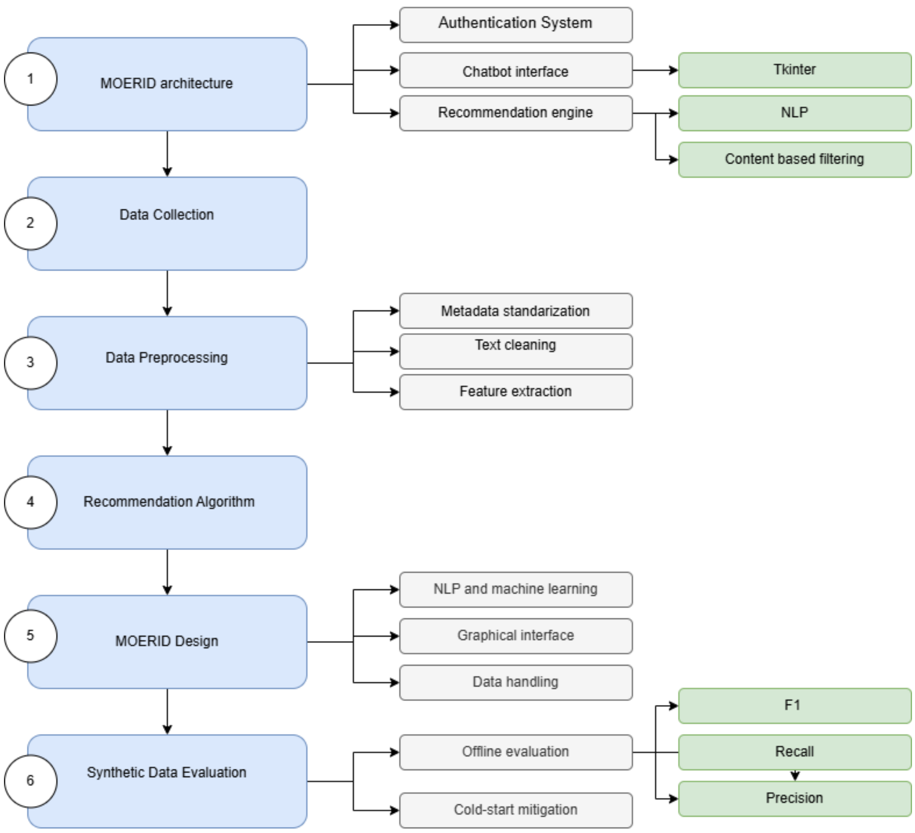


Figure 1. Research methodology

3.1. MOERID architecture

MOERID is an artificial intelligence (AI)-driven chatbot that uses state-of-the-art recommendation techniques and NLP. Access to the chatbot is secured through an authentication process, which requires users to authenticate themselves using login credentials. For instance, Figure 2 presents an updated overview of the system architecture previously elaborated in [16]. Authors in this preceding version rely on the OpenAI API, instead of an OER Dataset to get the recommended OER URLs. It also uses a static conversational mood that is based on a set of changeless questions to extract the intents and entities based on GPT3. Sandoussi *et al.* [16] utilized GPT3 since its GPT algorithms are the only compatible ones with the API method to extract recommendation.

MOERID relies on NLP techniques and is an AI based recommender that is designed for the teaching community. Unlike other systems that focus on the learners only and adopt static search methods, this system is conversational and deploys an OER dataset.

a. User logs to the system using personal credentials.
b. User gets a greeting message once logged.
c. User types and enters a paragraph to express the specificities of the needed OERs.
d. System uses NLP to assimilate the semantic meaning of the user's query.

e. System deploys specifically the NLU to extract Intents and Entities to understand the semantic meaning of the user's paragraph.
f. System sends a list of OERs links from the available OER in the dataset, by utilizing CBF filtering based on the Intents and Entities previously extracted, this process is ensured in the Core recommender engine of the chatbot.
g. System presents the list of links to the chatbot interface for the users via NLP and asks for the user's feedback about the recommended list of OERs links.
h. User gives its feedback by rating the recommendation.

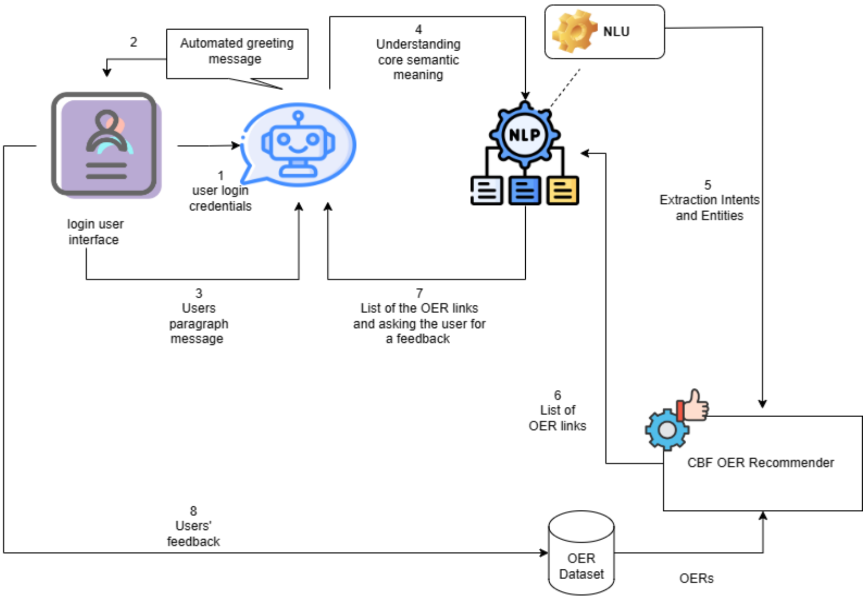


Figure 2. MOERID architecture

3.2. Data collection

Obtaining free datasets from OER repositories like Merlot and Common OER to test educational RS is not a commonly used practice. Due to legal and ethical concerns, we opted for real-world datasets freely available on online community platforms such as ResearchGate, Kaggle, and GitHub. A first example is the "Travel well dataset" [17], sourced from the learning resource exchange portal, comprising data from the EC-funded MELT project's pilot phase (August 2008 to February 2009) involving 98 users. The dataset contains explicit interest indicators, allowing inference of resource relevance based on user ratings and tags. The second findings are the CiteUlike-a [18] and CiteUlike-t [19] datasets, containing articles information from CiteULike and Google Scholar. Though, these two datasets are not labeled as OER. The third example is the ITM-Rec dataset, gathered from student questionnaires at the ITM department, Illinois Institute of Technology, USA, offers preferences on final project topics in database and data science classes [20].

To select the most suitable dataset and make an informed decision about which one aligns best with our objectives, we have established these criteria:

a. Overview and User Engagement: This criterion presents the dataset's characteristics, structure, and content. The goal is to gain insights into various aspects of the dataset, considering the interaction, interest, and involvement that users exhibit with the data.
b. Personalization: This criterion focuses on understanding how each OER dataset supports a personalized learning environment. We look for information that reflects how the dataset caters to individualized educational experiences.
c. Metadata and context: Here, we examine information related to metadata and contextual details associated with a set of features in the dataset. It helps us understand the additional information provided alongside the core data elements.
d. Data quality: This criterion assesses the quality and suitability of the three datasets and how each dataset aligns with the needs of a different type of recommenders.

Table 2 evaluates the datasets (TEL, CiteULike-a, and ITM) based on their characteristics. TEL Record user activities for European courses with rich metadata (19 features) and high data quality, suitable for CBF and CF recommenders. CiteULike-a emphasizes Google Scholar articles with minimal to poor metadata. Valid for CF recommenders but unsuitable for CBF testing. ITM focuses on students' preferences for project topics with limited metadata. It fits context-aware and CF recommenders but not CBF.

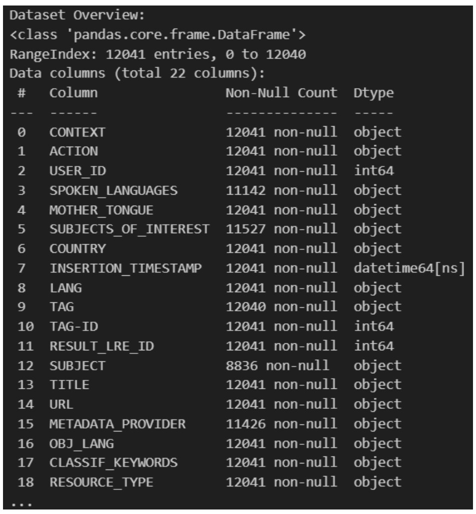
Table 2. Data comparison

Dataset/characteristic	Overview and user engagement	Personalization	Metadata and context	Data quality
TEL	98 users 13.922 user activities were recorded on 1.923 resources - xls format	Diverse primary and secondary courses from all over Europe in diverse subjects	A range of 19 feature of metadata	Valid and accurate with little to no missing data - fits CBF and CF recommenders
CiteUlike-a	5551 users and 16980 items with 204987 user-items pairs - .dat format	Resources articles from google scholar	No to low metadata	Accurate for social and CF recommenders with regression problems Invalid for CBF recommenders testing
ITM	476 users and 70 items with 5230 rating - csv format	student and group preferences on the topics of final projects in database and data science classes, collected from student questionnaires at the ITM department, Illinois Institute of Technology, USA	No to low metadata	Valid for Context-aware, group and CF recommenders inaccurate for CBF recommenders testing

3.3. Data preprocessing:

Selecting a suitable dataset is a crucial phase for implementing a RS. The dataset for the purpose of this research ought to contain a wide range of resources, covering diverse subjects and levels of complexity, to ensure that the system can cater to a broad expanse of user preferences. To efficiently compute similarity between resources and generate relevant recommendations, the dataset should include meticulous and diverse metadata for each resource and should be sufficient in size to allow the learning of significant patterns. The most accurate dataset for a CBF task that relies on different textual features and metadata from Table 1 is TEL. It includes a wide range of metadata, which allow users to obtain optimal responses to their queries. Figure 3 illustrates an overview of TEL metadata. This dataset containing 22 columns and 12,041 entries, with most columns are fully populated. Its key features are presented as follows:

a. General columns: Includes USER_ID, ACTION, CONTEXT, and INSERTION_TIMESTAMP to track user interactions.
b. User attributes: Columns like SPOKEN_LANGUAGES, MOTHER_TONGUE, SUBJECTS_OF_INTEREST, and COUNTRY provide user-specific details.
c. Content metadata: Includes TAG, TITLE, SUBJECT, CLASSIF_KEYWORDS, RESOURCE_TYPE, and METADATA_PROVIDER for describing and categorizing resources.



```
Dataset Overview:
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 12041 entries, 0 to 12040
Data columns (total 22 columns):
 #   Column               Non-Null Count  Dtype
---  ------               --------------  -----
 0   CONTEXT              12041 non-null  object
 1   ACTION               12041 non-null  object
 2   USER_ID              12041 non-null  int64
 3   SPOKEN_LANGUAGES     11142 non-null  object
 4   MOTHER_TONGUE        12041 non-null  object
 5   SUBJECTS_OF_INTEREST 11527 non-null  object
 6   COUNTRY              12041 non-null  object
 7   INSERTION_TIMESTAMP  12041 non-null  datetime64[ns]
 8   LANG                 12041 non-null  object
 9   TAG                  12040 non-null  object
 10  TAG-ID               12041 non-null  int64
 11  RESULT_LRE_ID        12041 non-null  int64
 12  SUBJECT              8836 non-null   object
 13  TITLE                12041 non-null  object
 14  URL                  12041 non-null  object
 15  METADATA_PROVIDER    11426 non-null  object
 16  OBJ_LANG             12041 non-null  object
 17  CLASSIF_KEYWORDS     12041 non-null  object
 18  RESOURCE_TYPE        12041 non-null  object
...
```

Figure 3. Dataset overview

3.4. Recommendation algorithm

To optimize the choice of the CBF algorithm, a comprehensive comparison of the CBF algorithms and techniques, mentioned in the technical framework paragraph is conducted. Measuring the utility of CBF is commonly calculated by using heuristic functions, and one of the most used is the cosine similarity metric [21]. Hence, the method deploys a score similarity comparison. For this purpose, ground truth and concatenate text columns for a cosine comparison analysis are established. Mean cosine similarities for the previously cited techniques, namely TF-IDF, N-gram, Word2Vec, LSA, LDA, FastText, GloVe, and BERT are implemented and computed. Each of the techniques used in the script, generates numerical vectors that present documents.

In other words, the comparison uses mathematical approaches to estimate the similarity between items. starting by converting textual data to numerical models. Each document d is represented as a vector vd in which every component represents the TF-IDF score of a term in the document. The cosine similarity between document di and document dj is represented as follows:

$$cosine\_similarity(di, dj) = \| vdi \| \times \| vdj \| \; vdi \cdot vdj$$

As depicted in Figure 4, word2Vec reveals superior performance in terms of mean cosine similarity. This finding emphasizes the effectiveness of Word2Vec embeddings in grasping nuanced semantic relationships within the text corpus. The observed distinction of Word2Vec suggests its potential as a robust and reliable choice for the task of the study.

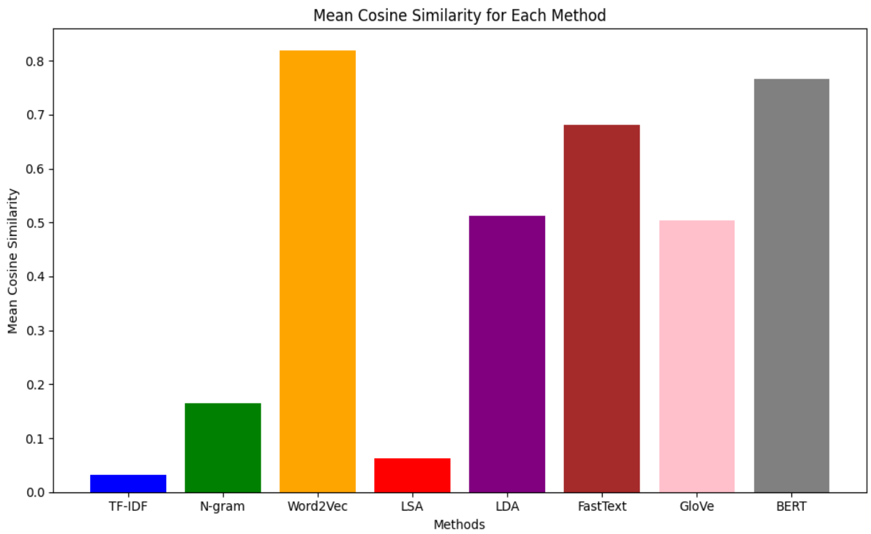


Figure 4. Algorithms comparison

3.5. System implementation

MOERID is built as follows:

a. User authentication: To secure access to the system and to assure a personalized user-based experience, MOERID is endowed with an authentication system requiring identification credentials. It is also a security protocol regulating access to the MOERID's interface.
b. Entity extraction using \`SpaCy\`: To categorize the Intents and Entities within users' input, we utilize natural language processing (NLP), deploying the SpaCy library. It ensures the chatbot's semantic comprehension by comprehending and identifying entities.
c. Language detector and handler: A part of the NLP that handles different languages since the OERs are originally coming from different European countries adopting different languages.
d. Content-based filtering (CBF) recommendation using \`Word2Vec\`: To tailor OER recommendations based on semantic content similarity. The RS deploys the Word2Vec model from the Gensim library. It creates vector representations for user queries and resource texts. The cosine similarity metric is calculated then to quantify the parity between preferences and recommendations. The reason to pair CBF algorithms and cosine similarity is that this approach has demonstrated promising results in maximizing recommendation accuracy in several cases, such as in [22].

e. User feedback mechanism: To enable user input for continuous progress and improvement, we opt for a systematic user feedback process. This mechanism is based on prompting the user for input that is collected and stored as rating.

The system opted for a modular design based on several classes. It is an architectural approach to enhance the code organization, comprehension, and maintainability. This modular aspect is respectively responsible for the previous actions: authentication class, input processing class, recommendation generation class, user feedback class and main chatbot class. To ensure user interaction, the chatbot deploys a graphical user interface (GUI) based on the Tkinter library [23].

Figure 5 showcases MOERID's graphical user interface (GUI). The login screen demonstrates that users must enter their username and password to access the chatbot. The chat Interface proves that users can interact with MOERID by entering text queries or paragraphs in the input field. It also features options to rate recommendations and submit inputs for further assistance.

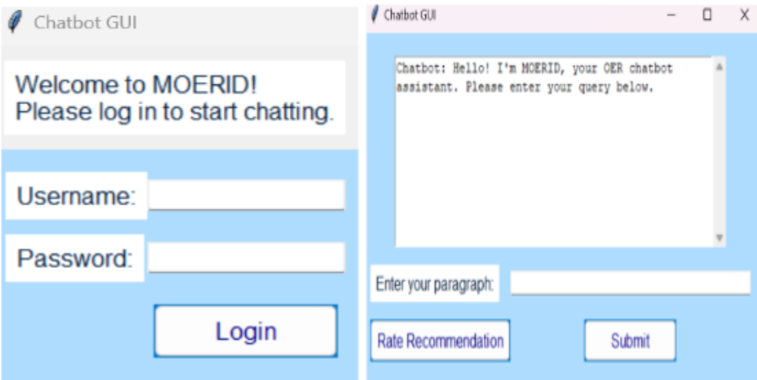


Figure 5. MOERID's interface

3.5. Synthetic data evaluation

The usefulness of synthetic data extends from the efficiency of synthetic reviews in evaluating RS [24] to examining synthetic datasets for conversational RS [25]. The generation of this type of data, is defined as the approach to create data with statistically equivalent properties to real-world data. These techniques deem considerably efficient in evaluating and testing several systems [26], [27]. Inspired by the research [28], synthetic logic is used to generate simulated user preferences combining URLs and interest scores. Calculating precision, recall and F1 score as they are the most used evaluation metrics [29]. Technically, the used method utilizes synthetic user preferences stored in the synthetic_user_preferences dictionary to simulate user interest in a set of URLs. Figure 6 depicts the performance metrics of the system. Precision is defined by the accuracy of positive predictions, and reported at 0.93. Recall, represented by the system's relevance to capture pertinent items and observed at 0.80. And the F1 score is interpreted as a ballast of precision and recall capturing the overall performance of the system and is reported as 0.86. These values suggest an accurate and reliable model with balanced precision and recall.

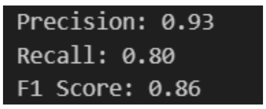


Figure 6. Evaluation metrics

4. DISCUSSION

The MOERID chatbot is deemed to be efficient as demonstrated by the evaluation metrics. The features used for this model are 'SUBJECT', 'TITLE', 'TAG', 'SUBJECTS_OF_INTEREST', 'SPOKEN_LANGUAGES', and 'MOTHER_TONGUE', as they are the most relevant metadata. MOERID is user-friendly and simple to manipulate. The obtained values indicate that the recommendations are relevant and accurate. We applied a warm start method inspired by Sandoussi *et al.* [16], to mitigate the cold start problem, the CBF model aims to recommend items similar to those previously uploaded by the user to the dataset, ensuring relevance even with limited user history [30]. The system addresses literature gaps highlighted in [4]. It implements the envisioned architecture in [16]. To make modification and upgrading unproblematic, the system is modulable through a class-based structure, endowed with a chatbot to support an interactive and engaging user experience, motivating educators to utilize the system for their instructional needs.

Despite its promising performance, MOERID requires further validation under real-world conditions. For more efficient testing, we expect to experiment with a real-world review evaluation to better understand the system's practical implications. We also aim to add a downloading feature to empower educators to contribute new OER, enriching the dataset and expanding the diversity of recommendations.

Future work will focus on integrating CF as a valuable technique to fine-tune and enhance recommendations by incorporating user-item interactions and group preferences. By adding CF, a widely used recommendation technique, we aim to improve the system's accuracy and transition it into a hybrid recommender. Additionally, as a future task, we plan to optimize MOERID's response time for better performance.

5. CONCLUSION

The recommendation system has revolutionized educational practices. However, most of the current OER RSs have focused on developing learner-centered OER recommenders but have not yet focused much on instructors-centered OER recommenders. In this paper, we built a recommendation system for the teaching community by combining CBF and NLP techniques to endow it with a chatbot. First, we used cosine similarity to identify the best algorithm that suites our case. Next, we selected word2vec as a recommendation algorithm and we added an NLP component to build our chatbot. The used data are from a real-life dataset for accurate results and optimal recommendations. We also deployed synthetic data to evaluate the system and we obtained outstanding results. However, we still encounter limitations, such as the optimization of data processing speed due to the machine resource limits and the volume of the dataset.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Sandoussi Rima	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓			
Hnida Meriem		✓		✓	✓	✓		✓		✓		✓	✓	
Daoudi Najima		✓		✓		✓		✓		✓		✓	✓	
Ajhoun Rachida				✓		✓				✓		✓	✓	

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The authors confirm that the data supporting the findings of this study are available within the article and references.

REFERENCES

[1] *Open educational resources*. GEM Report UNESCO, 2023. doi: 10.54676/lbiv3337.
[2] "Recommendation on open educational resources (OER) - Legal affairs," *Underground Pipeline Corrosion: Detection, Analysis and Prevention*. https://www.unesco.org/en/legal-affairs/recommendation-open-educational-resources-oer (accessed Nov. 15, 2024).
[3] X. N. Lam, T. Vu, T. D. Le, and A. D. Duong, "Addressing cold-start problem in recommendation systems," in *Proceedings of the 2nd International Conference on Ubiquitous Information Management and Communication, ICUIMC-2008*, Jan. 2008, pp. 208–211. doi: 10.1145/1352793.1352837.
[4] S. Rima, H. Meriem, D. Najima, and A. Rachida, "Systematic literature review on open educational resources recommender systems," *International Journal of Interactive Mobile Technologies*, vol. 16, no. 18, pp. 44–77, Oct. 2022, doi: 10.3991/ijim.v16i18.32197.
[5] E. Çano and M. Morisio, "Hybrid recommender systems: A systematic literature review," *Intelligent Data Analysis*, vol. 21, no. 6, pp. 1487–1524, Nov. 2017, doi: 10.3233/IDA-163209.
[6] M. T. Sharad Phalle and P. S. Bhushan, "Content-based filtering and collaborative filtering: A comparative study," *Journal of Advanced Zoology*, pp. 96–100, Mar. 2024, doi: 10.53555/jaz.v45is4.4158.
[7] S. S. Lakshmi and T. A. Lakshmi, 'Recommendation Systems:Issues and challenges.' [Online]. Available: https://www.ijcsit.com/docs/Volume%205/vol5issue04/ijcsit20140504207.pdf
[8] Y. Adilaksa and A. Musdholifah, "Recommendation system for elective courses using content-based filtering and weighted cosine similarity," in *2021 4th International Seminar on Research of Information Technology and Intelligent Systems, ISRITI 2021*, Dec. 2021, pp. 51–55. doi: 10.1109/ISRITI54043.2021.9702788.
[9] M. R. Julianti, Y. Heryadi, B. Yulianto, and W. Budiharto, "Recommendation system model for personalized learning in higher education using content-based filtering method," in *Proceedings of 2022 International Conference on Information Management and Technology, ICIMTech 2022*, Aug. 2022, pp. 1–6. doi: 10.1109/ICIMTech55957.2022.9915109.
[10] C. De Medio, C. Limongelli, F. Sciarrone, and M. Temperini, "MoodleREC: A recommendation system for creating courses using the moodle e-learning platform," *Computers in Human Behavior*, vol. 104, p. 106168, Mar. 2020, doi: 10.1016/j.chb.2019.106168.
[11] S. R. Goudar and G. Shidaganti, "Multi-parameter based educational recommendation system for online courses: A content-based filtering approach," in *2023 International Conference on Network, Multimedia and Information Technology, NMITCON 2023*, Sep. 2023, pp. 1–7. doi: 10.1109/NMITCON58196.2023.10276153.
[12] M. Dhahri and M. K. Khribi, "A review of educational recommender systems for teachers," *18th International Conference on Cognition and Exploratory Learning in Digital Age, CELDA 2021*, pp. 131–138, 2021, doi: 10.33965/celda2021_202108l016.
[13] M. Bellal, A. Mejri, and L. Cheniti-Belcadhi, "Towards a quality-driven and ontology-based recommender framework for open educational resources," in *17th International Conference on INnovations in Intelligent SysTems and Applications, INISTA 2023 - Proceedings*, Sep. 2023, pp. 1–6. doi: 10.1109/INISTA59065.2023.10310521.
[14] A. Gordillo, D. López-Fernández, and K. Verbert, "Examining the usefulness of quality scores for generating learning object recommendations in repositories of open educational resources," *Applied Sciences (Switzerland)*, vol. 10, no. 13, p. 4638, Jul. 2020, doi: 10.3390/app10134638.
[15] J. F. Herrera-Cubides, P. A. Gaona-García, C. E. Montenegro-Marín, and S. Sánchez-Alonso, "Improving OER descriptions to enhance their availability, reuse, and enrichment," *Education and Information Technologies*, vol. 27, no. 2, pp. 1811–1839, Aug. 2022, doi: 10.1007/s10639-021-10641-w.
[16] R. Sandoussi, M. Hnida, N. Daoudi, and R. Ajhoun, "MOERID: Design and implementation of a GPT-3-based generative chatbot for personalized open educational resource recommendations in teaching," in *Digital Technologies and Applications*, Springer Nature Switzerland, 2024, pp. 64–73. doi: 10.1007/978-3-031-68653-5_7.
[17] "R. Vuorikari, 'Travel well dataset,' Nov. 2017. doi: 10.13140/RG.2.2.27742.74562".
[18] C. Wang and D. M. Blei, "Collaborative topic modeling for recommending scientific articles," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Aug. 2011, pp. 448–456. doi: 10.1145/2020408.2020480.
[19] H. Wang, B. Chen, and W. J. Li, "Collaborative topic regression with social regularization for tag recommendation," in *IJCAI International Joint Conference on Artificial Intelligence*, 2013, pp. 2719–2725.
[20] Y. Zheng, "ITM-Rec: An open data set for educational recommender systems," *arXiv:2303.10230*, 2023.
[21] Z. Fayyaz, M. Ebrahimian, D. Nawara, A. Ibrahim, and R. Kashef, "Recommendation systems: Algorithms, challenges, metrics, and business opportunities," *Applied Sciences (Switzerland)*, vol. 10, no. 21, pp. 1–20, Nov. 2020, doi: 10.3390/app10217748.
[22] V. Thada and V. Jaglan, "Comparison of Jaccard, Dice, Cosine Similarity," *International Journal of Innovations in Engineering and Technology (IJIET)*, vol. 2, no. 4, 2013.
[23] A. D. Moore, *Python GUI Programming with Tkinter: Design and build functional and user-friendly GUI applications*. 2021.
[24] M. F. Abdurrafi and D. H. U. Ningsih, "Content-based filtering using cosine similarity algorithm for alternative selection on training programs," *Journal of Soft Computing Exploration*, vol. 4, no. 4, pp. 204–212, Dec. 2023, doi: 10.52465/joscex.v4i4.232.
[25] S. Ouyang, A. Lawlor, F. Costa, and P. Dolog, "Improving explainable recommendations with synthetic reviews," *arXiv:1807.06978*, 2018.
[26] H. Lara and M. Tiwari, "Evaluation of synthetic datasets for conversational recommender systems," *arXiv:2212.08167*, Dec. 2022.
[27] M. D. Ekstrand, A. Chaney, P. Castells, R. Burke, D. Rohde, and M. Slokom, "SimuRec: Workshop on synthetic data and simulation methods for recommender systems research," in *RecSys 2021 - 15th ACM Conference on Recommender Systems*, Sep. 2021, pp. 803–805. doi: 10.1145/3460231.3470938.
[28] F. K. Dankar and M. Ibrahim, "Fake it till you make it: Guidelines for effective synthetic data generation," *Applied Sciences (Switzerland)*, vol. 11, no. 5, pp. 1–18, Feb. 2021, doi: 10.3390/app11052158.
[29] E. Stavinova, A. Grigorievskiy, A. Volodkevich, P. Chunaev, K. Bochenina, and D. Bugaychenko, "Synthetic data-based simulators for recommender systems: A survey," *arXiv:2206.11338v1*, Jun. 2022.
[30] I. Barjasteh, R. Forsati, D. Ross, A. H. Esfahanian, and H. Radha, "Cold-start recommendation with provable guarantees: A decoupled approach," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 6, pp. 1462–1474, Jun. 2016, doi: 10.1109/TKDE.2016.2522422.

BIOGRAPHIES OF AUTHORS

Sandoussi Rima received an engineering degree in data sciences from the School of Information Science (ESI) in Rabat, Morocco, she is now a Ph.D. student at the National Superior School of Computer Science and Systems Analysis Ensias, Mohammed V University in Rabat, Morocco. Her research areas are artificial intelligence, recommender systems and chatbots, deep learning, and natural language processing and their application to real-world problems. She is currently working on the recommendation algorithms and chatbots applied to the teaching field. She can be contacted at email: sandoussi.rima@gmail.com.

Hnida Meriem is a professor at the School of Information Sciences (ESI). She holds a Ph.D. in computer science from Mohammadia School of Engineers (EMI), Mohammed V University in Rabat, Morocco. Her research focuses on MLOps, TinyML blockchain technology, software engineering, and educational technologies. She can be contacted at email: mhnida@esi.ac.ma.

Daoudi Najima is a full professor at the School of Information Sciences, Rabat,Morocco. She has an engineering degree from the National Institute of Statistics and Applied Economics (INSEA), Rabat, Morocco, and a Ph.D. in Computer Science from Ensias, Rabat, Morocco. Her research interests include artificial intelligence, e-learning, recommendation systems, and natural language processing. She can be contacted at email: ndaoudi@esi.ac.ma.

Ajhoun Rachida is a professor of higher education at the National Superior School of Computer Science and Systems Analysis (ENSIAS), Mohammed V University in Rabat, Morocco. She holds a Ph.D. in Computer Science and E-learning from the Mohammadia School of Engineers (EMI), Mohammed V University in Rabat, Morocco. Her research interests include e-learning, recommendation systems, and educational technology. She can be contacted at email: ajhoun@gmail.com.