International Journal of Computational and Experimental Science and ENgineering (IJCESEN)

Vol. 11-No.3 (2025) pp. 5142-5152
http://www.ijcesen.com

ISSN: 2149-9144

Research Article

Optimizing Online Educational Experiences through Semantic Ontology-Based Recommender Systems: A Case Study on Coursera

Diman Jalal Mustafa[1]*, Subhi R. M. Zeebaree[2]

[1*] Information Technology Management Department, Technical College of Administration, Duhok Polytechnic University, Kurdistan Region, Iraq
* **Corresponding Author Email:** diman.jalal@dpu.edu.krd- **ORCID:** 0009-0003-2161-8707

[2] Energy Engineering Department, Technical College of Engineering, Duhok Polytechnic University, Kurdistan Region, Iraq,
Email: subhi.rafeeq@dpu.edu.krd - **ORCID**: 0009-0003-2161-8706

Article Info:

DOI: 10.22399/ijcesen.3365
Received : 12 May 2025
Accepted : 07 July 2025

Keywords

Coursera, Web Scraping
Ontology Design
Content-Based Recommendation
Python.

Abstract:

With the rapid expansion of Massive Open Online Courses (MOOCs), learners face increasing challenges in identifying relevant educational content tailored to their interests and needs. This paper presents the design and implementation of a content-based recommendation system (CBRS) for Coursera courses, enhanced through the use of Semantic Web and ontology technologies. We constructed a domain-specific OWL ontology using real-world data extracted from over 1,000 Coursera courses, capturing key attributes. The system leverages semantic representations to improve the accuracy and relevance of recommendations by computing similarities between course contents and user preferences. The recommendation engine was evaluated using a test set of user queries and relevance judgments. Experimental results show strong performance, achieving Precision 0.98%, Recall 1.00, F1 Score ≈0.99% and Accuracy 0.98%. These findings demonstrate the system's effectiveness in delivering personalized, high-quality recommendations and underscore the value of integrating ontologies into educational recommender systems. This work contributes to the advancement of intelligent e-learning systems by enhancing resource discoverability, user engagement, and overall learning experience in MOOC platforms.

1. Introduction

The education sector is experiencing a profound digital transformation, with online learning becoming a central component in reshaping global education. Technological advancements such as big data, cloud computing, and artificial intelligence have empowered educational platforms with rich learning resources and detailed user behavior data, enabling a deeper understanding of learners' needs and preferences [1]. Among the most prominent innovations in this landscape are Massive Open Online Courses (MOOCs), which offer a vast array of online courses through platforms like Coursera. However, the abundance of options can overwhelm learners, making it difficult to identify the most suitable courses. Recommendation systems have emerged as a practical solution to guide learners toward content aligned with their interests and educational goals[2]. In this context, the Semantic Web plays a pivotal role in enhancing the organization and retrieval of educational content within MOOCs. As an extension of the conventional web, the Semantic Web aims to make data understandable by both humans and machines[3]. At its core, the Semantic Web relies on ontologies—formal knowledge models that semantically describe entities and their interrelations. Unlike traditional data structures that rely solely on raw text or numbers, ontologies capture contextual meaning, allowing for more accurate representation and interpretation of educational resources [4].

MOOCs face several persistent challenges, including content overload, difficulties in meeting learners' motivational and interest-based needs, and the lack of adaptive mechanisms for personalized learning. By developing educational ontologies, platforms can better classify course content, personalize recommendations, and enable cross-platform interoperability through standardized technologies such as OWL and RDF [5]. This paper proposes a content-based recommendation system that leverages a combination of web scraping, text preprocessing, ontology modeling, and similarity computation to deliver personalized course suggestions [6]. While online education provides convenient access to learning resources, it also introduces a new layer of complexity—users must navigate an overwhelming amount of information to find content that matches their specific needs. Without effective filtering, learners may become disoriented and struggle to choose appropriate starting points for their learning journey [7, 8].

To address this, recommendation systems are increasingly seen as essential tools to combat information overload. By analyzing learners' profiles, interests, and feedback, these systems can dynamically adapt and provide high-relevance course recommendations. Such intelligent recommendation frameworks not only enhance user satisfaction but also contribute to improved learning outcomes through more targeted and engaging educational pathways.

The remainder of this paper is presented as follows: Section 2 the related work. In section 3 show the proposed model methodology with experimental results. Section 4 Comparison and Discussion. Finally in section 5,6 present conclusion and future work.

2. Related Work

This section begins by examining recent research on the use of ontology and recommendation systems in the education field, highlighting their role in improving personalization and supporting educational decision-making. It then reviews the most prominent studies that focused on predicting personality traits using recommendation system explaining the relationship between personality traits and learning preferences. Finally, the section discusses the latest developments in personalized learning systems, including recent based recommendations**.**

[9]By integrating educational psychology and artificial intelligence theories, the researchers developed a successful strategy for providing students in a distance learning environment with appropriate learning materials. First, students' learning behaviors are examined to assess their skills and classify them into three main categories. A recommendation algorithm based on the LinUCB model is then created using features extracted from the learning resources, such as difficulty level. Experimental results of the basic functions demonstrate that their recommendation system is able to accurately find appropriate learning resources, with the highest accuracy reaching nearly 70%. This algorithm also features a personalized exploration parameter that is dynamically adjusted based on the student's ability and attention span, balancing exploration and exploitation during the recommendation process. Thus, their system successfully provides learning resources tailored to students' individual abilities and needs, while encouraging them to develop their capabilities through appropriate challenges. In addition, the researchers[10] in this paper explore the development of a personalized course recommendation system to improve e-learning experiences by employing machine learning (ML) and collaborative filtering (CF) algorithms. Data was collected from coursera and Udemy platforms focusing on students' academic performance, interests, and learning preferences to recommend the most useful courses, they found that the system analyzed students preferences and academic performance to generate personalized course recommendations from a selected set of 20 courses, achieving a student engagement rate exceeding 86% Overall, the study highlights the potential of personalized recommendations to improve learning outcomes and student satisfaction in smart education environments. On other hand [11] focused on enhancing online education by integrating semantic analysis with deep machine learning to create an Intelligent Content-Based Recommendation System (ICRS) helping learners select suitable materials based on their existing knowledge and experience the proposed system also introduces a novel method that combines semantic graphs and context to generate a semantic matrix for better resource classification. Four machine learning (Logistic regression, Random Forest, SVM, and MLP) models were evaluated, along with an augmented deep learning model (LSTMM), which outperformed others with an Accuracy of 0.8453 and F1 Score of 0.7731. The results demonstrate that LSTMM is highly effective in improving e-learning recommendations.

[12]presented an intelligent web platform the system employed algorithms such as PageRank for ranking articles, neural networks for document classification, and the ROCK algorithm for friend recommendations based on article similarity, the neural network achieved the highest accuracy (95.2%) in classifying articles, and the system effectively recommends articles and friends based on user interests the platform facilitates knowledge sharing and social interaction among users, improving access to information and learning opportunities in rural areas

[13] In this study, the authors present a personalized learning system that uses the MBTI framework to classify individuals into 16 personality types, enabling a deeper understanding of information processing and decision-making. They also present a system based on ontology-based frameworks to provide personalized recommendations for educational resources. Results from the (Mbti_1 and mbti_full_pull) datasets, using the BERT methodology and deep learning models, showed an accuracy of 97%. however, [14]the authors of this research paper explore machine learning applications in online learning systems to understand learner characteristics, a model examines learning behaviors (length, frequency, and interactions),results indicate that course preferences can be predicted with approximately 70% accuracy, personalized recommendations extend learning time and increase course completion rates by 30%,For behavior analysis, the system uses methods such as random forest and logistic regression model interpretability and data completeness pose challenges future developments for greater flexibility and visualization are recommended. However, researchers[15] have used the E-SWT framework and Semantic Web technologies to improve educational experiences by facilitating data-driven decision-making, encouraging collaboration, and simplifying administrative tasks using intelligent agents and knowledge graphs, the framework addresses traditional educational problems, including data fragmentation and rigid learning plans this methodology has demonstrated significant advantages in real-world classroom situations by encouraging individualized learning, efficient resource allocation, and enhanced student engagement.

The authors[16] also appreciated that by effectively integrating user ratings and sentiment analysis from reviews, the hybrid recommendation system demonstrated improved performance by addressing issues such as cold starts and data scarcity, the system produced more accurate and personalized course recommendations, experiments confirmed the system's effectiveness, showing a significant increase in user satisfaction and interest in the suggested courses ,the overall similarity rate of the dataset, 98%, indicated a high degree of agreement between user ratings and the preferences expressed in text reviews. But The authors[17] used AI models, specifically Mistral-7B-Instruct-v0.2 and GPT-4-based ChatGPT, to develop a personalized, multilingual course recommendation system that optimizes course recommendations based on users' interests and competencies to create personalized learning paths, the system used a hybrid strategy combining collaborative filtering, content-based filtering, and AI-based matching. By delivering personalized courses.

Table 1. *Comparative Analysis of Prediction Recommendations System Using Different Techniques*

Reference	Dataset	Ontology	Recommendation Type	Model	Results
[9]	Coursera platform	Not applied	User-based recommendation	LinUCB-based recommendation model	precision has arrived at nearly 70%
[10]	Coursera and udemy platform	Not applied	Collaborative Filtering	(Random Forest, Decision Tree, K-Nearest Neighbors, Singular Value) ML	precision has 0.8621%
[11]	NPTEL and Coursera	Not applied	content-based recommender system	(LSTMM)ML and DL	Accuracy= 0.8453 and F1 Score= 0.7731 respectively.
[12]	social e-learning platform	Not applied	similarity-based recommendation	(ROCK , Naïve Bayes, Decision Tree, Neural Network.)ML	Accuracy= 95.2%
[13]	Mbti_1 and mbti_full_pull	Applied	recommended resources (Hybrid recommend)	DL models, BERT	Accuracy= 97%
[14]	online education platform	Not applied	personalized recommendation system	(CFSFDP)ML MODELS	Accuracy= 75%
[15]	open data platform	Applied	Not applied	E-SWT (Educational Semantic Web Technology)	Accuracy= 97.68%
[16]	622 Course online	Not applied	content-based recommender system and Collaborative Filtering	(CNN)DL	similarity score of 98%
[17]	800 online courses	Not applied	content-based filtering, collaborative filtering	mistralai/Mistral-7B-Instruct-v0.2	Accuracy more than 78%
Prosed model	Coursera platform	Applied	content-based recommendation (50 courses)	Ontology Course Recommender	98.99%

3. Methodology

This section covers the steps we took to build and implement a recommendation system on Ontology. Figures (1, 2 and 3) represent the main steps of the: System architecture, Ontology Class and Property Structure, and System Architecture of the Semantic Recommender System).

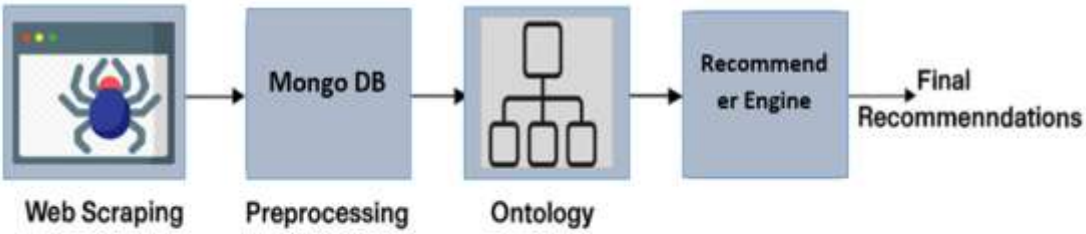

Figure 1. *System Architecture.*

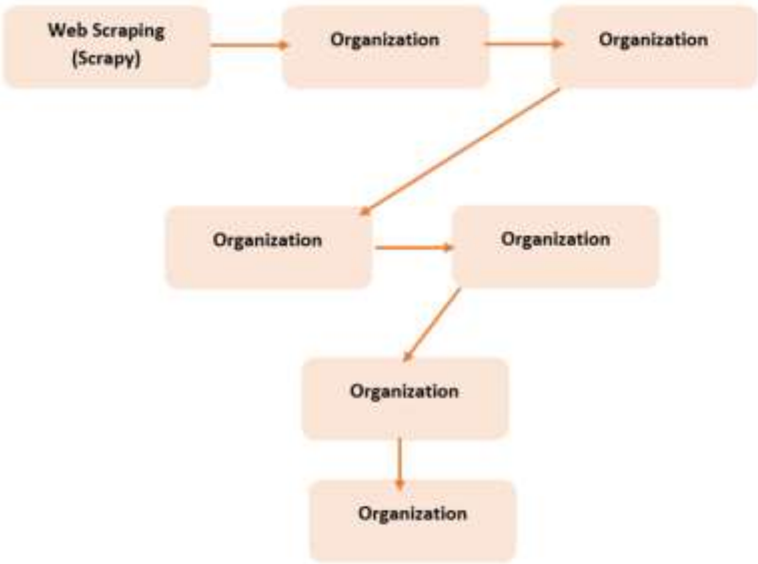

Figure 2. *Ontology Class and Property Structure.*

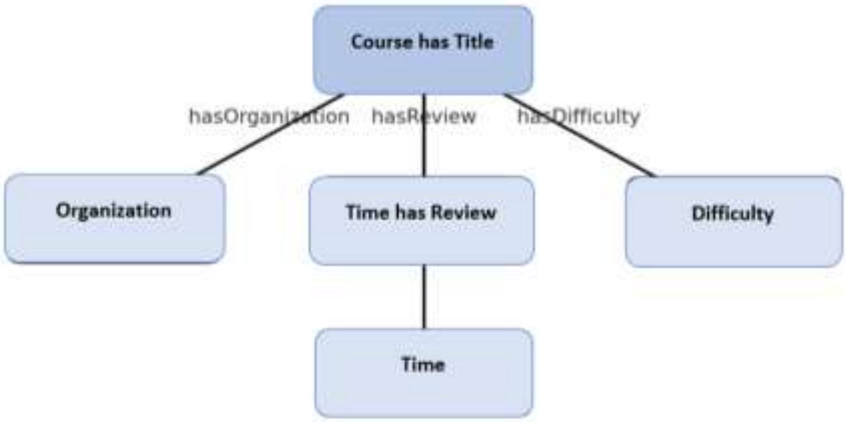

Figure 3. *System Architecture of the Semantic Recommender System.*

3.1 Data Collection

To build a personalized recommendation system, relevant course data was collected from the Coursera platform using an automated web data extraction method, as described previously[18]. The Scrapy framework, a powerful Python web data extraction library, was used within the PyCharm development environment to extract structured information from Coursera course listing pages. Key attributes of each course were retrieved, such as the course title, URL, instructor organization, course description, difficulty level, user ratings, number of enrolled students, and more. The data extraction process was designed to track pagination links to ensure comprehensive coverage of available courses. To efficiently store the extracted data, a MongoDB database was integrated via a custom Scrapy pipeline. MongoDB was chosen for its flexibility in handling semi-structured data, making it ideal for the evolving and diverse nature of educational content. This integrated dataset served as the basis for further semantic modeling and development of the recommendation system. I then stored this data from MongoDB on my computer, where JSON records are stored. Locally extracted for further processing.

3.2 Data Preprocessing

To ensure the reliability and quality of the collected dataset, a structured data cleaning process was applied using Python and the Pandas library. The raw dataset, exported in CSV format from the Coursera scraping phase, included potentially noisy and inconsistent entries across multiple fields. A batch-wise processing strategy was employed to handle large volumes of data efficiently, avoiding memory overflow. Key text-based fields such as course titles, descriptions, organizations, and difficulty levels were normalized by converting them to lowercase, stripping whitespace, and removing non-alphanumeric characters using regular expressions. Numeric fields such as course reviews, and enrollment counts, were also cleaned by removing non-numeric symbols and converting valid entries into proper numerical formats. Duplicate entries, based on course titles, were filtered out to preserve unique course records. The cleaned and structured dataset was then saved in a new CSV file, ensuring it was ready for semantic transformation and recommendation system training.

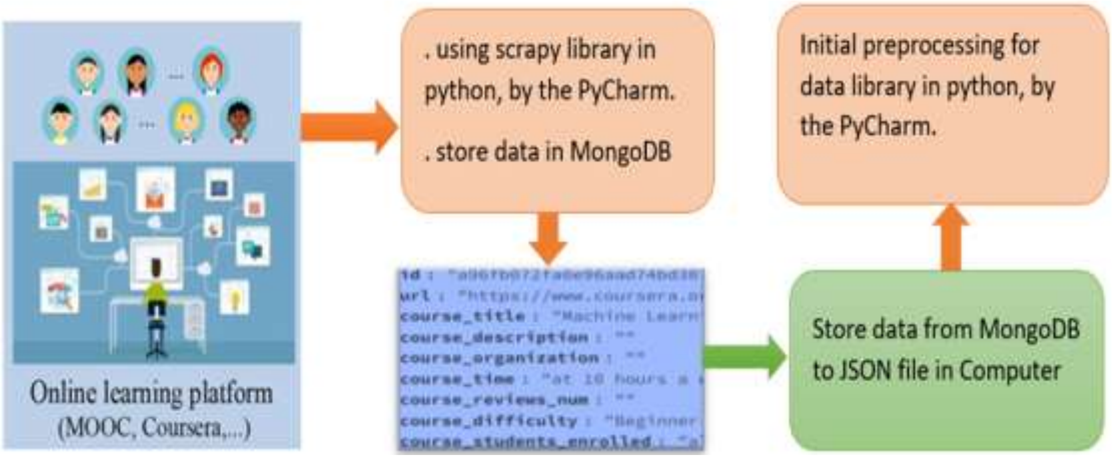

Figure 4. *Data collection process.*

3.3 Ontology Design

To semantically represent online course data and facilitate advanced reasoning and recommendations, an ontology was created using the Owlready2 Python library. The cleaned dataset, previously saved in CSV format, was transformed into OWL classes and properties. Key entities, such as Course, Organization, Review, Time, and Difficulty, were defined as OWL classes. Related data properties (e.g., has Title, has Description, has Students Enrolled, and has Price) and object properties (e.g., has Organization, has Time, has Review, and has Difficulty) were also defined to reflect course characteristics and their real-world relationships. Each course record was transformed from the CSV file into an individual OWL file with semantic links to its associated properties. Review scores were converted from percentages to numerical ratings, while free courses were coded at zero. The final ontology was serialized and saved in RDF/XML format for integration into semantic reasoning and content-based recommendation tasks, which I also presented in PyCharm as shown in figure 5.

3.4 Recommendation System Implementation

Recommendation systems are integral to personalized learning environments, helping tailor educational content based on user preferences, behaviors, and learning goals the system can provide options for students or users to express

```xml
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
         xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
         xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
         xmlns:owl="http://www.w3.org/2002/07/owl#"
         xml:base="D:/Coursera_Courses.owl"
         xmlns="D:/Coursera_Courses.owl#">

<owl:Ontology rdf:about="D:/Coursera_Courses.owl"/>

<owl:DatatypeProperty rdf:about="#hasCourseID">
  <rdfs:domain rdf:resource="#Course"/>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
</owl:DatatypeProperty>
```

Figure 5. *Part of the Ontology.*

their opinions about course recommendations based on their needs, taking into account time and quality, and they can find a list of massive open online courses [19]. This section explores the types of recommendation systems and focuses on the implementation of a content-based approach.

3.4.1 Recommendation System Categories

There are several main types of recommender systems, each with distinct methodologies and applications, as shown in Table 2:

Table 2. *Recommendation Systems Types and Description.*

Recommendation system type	Description
Content-Based Filtering	Recommends items similar to those a user has previously engaged with, based on item features.
Collaborative Filtering	Relies on user behavior patterns and preferences, identifying similarities between users or items
Hybrid Systems	Combine content-based and collaborative approaches to improve accuracy and overcome individual limitations.
Knowledge-Based Systems	Use domain knowledge and user-specific constraints or goals to make recommendations.
Context-Aware Systems	Factor in contextual information such as time, location, or device to personalize recommendations

Each type can be adapted to educational settings depending on the objectives of the system, the available data, and the level of customization required. In this paper, we focused on Content-Based Filtering

3.4.2 Content-Based Recommendation System

With the significant growth in the number of online courses available, it has become imperative to develop intelligent recommendation systems that help users choose the content most appropriate to their interests and backgrounds. In this project, a content-based recommendation system (CBRS) was designed using an OWL ontology created from real data extracted from Coursera, with the goal of enhancing the e-learning experience and improving the quality of recommendations provided to users. Figure 6 illustrating this process.

3.5 Experimental Results

Therefore, there is a specific system for determining the relevance of recommendations to a user type. In this case, the validity of the content-based system's performance evaluation is based on data extracted from the Coursera database, totaling 1,136 courses. We tested the system in multiple experiments to ensure its accuracy, which we will explain in this section.

3.5.1 System Testing and Enhanced Evaluation Approach.

The evaluation focused on four key performance metrics: Precision, Recall, F1-Score, and Accuracy.
Precision. In relation to the total number of resources suggested by the model, it is the percentage of accurate classifications in the recommendation where the learner is genuinely interested, calculated by:

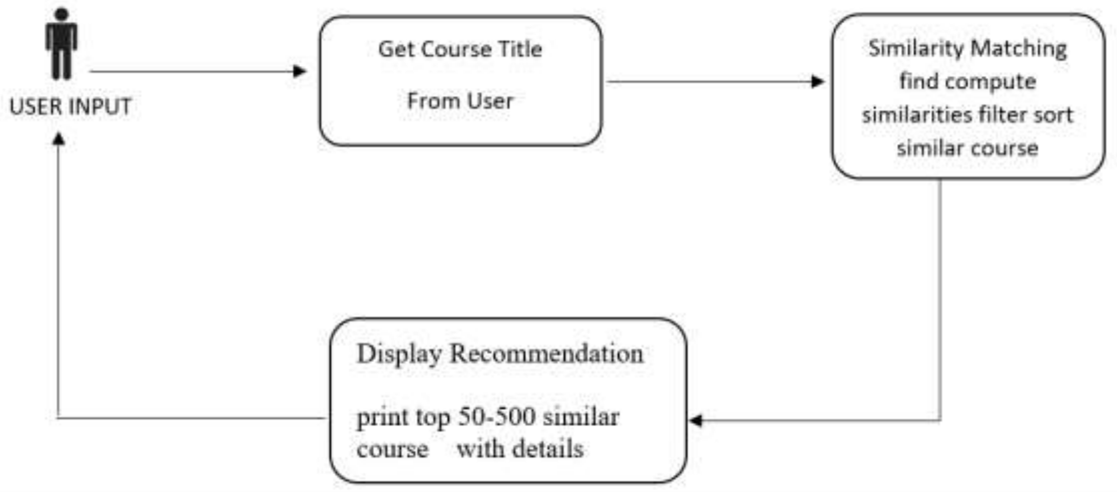

Figure 6. *Recommender System Structure.*

Precision= $\frac{\text{TP}}{(\text{TP} + \text{FP})}$

Recall. It provides the proportion of genuine cases that the model accurately identified as true in relation to all of the true examples that ought to have been suggested. The learner in the actual system finds these to be intriguing calculated by:

Recall = $\frac{\text{TP}}{(\text{TP} + \text{FN})}$.

F1-Score. When there is an imbalance between classes (for example, a small number of useless cycles versus a large number of helpful cycles), tis a statistic used to assess how well a classification model is performing. It balances recall and precision by combining both into a single number.

F1-Score =2* $\frac{(\text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})}$.

Accuracy. is a performance metric used to evaluate the correctness of a classification or recommendation system. It measures the proportion of correctly predicted instances (both positive and negative) out of the total number of predictions made.

Accuracy = $\frac{\text{TP} + \text{TN}}{(\text{Total Number of Prediction})}$

3.5.2 Experiences of Those Who Have Tested This System

Test A: the system requested 50 recommended courses. The evaluation metrics used are precision, recall, F1 score, and accuracy. The recommendation system generated 50 course recommendations. Of these courses:

49 courses were truly relevant (high quality with a review).

One course was irrelevant.

Thus, we can define:

y_true = ground truth (the actual relevance of courses): 49 courses are relevant; one course is irrelevant.

y_pred = system predictions: all 50 courses predicted as relevant.

- TP: True Positives (correctly recommended relevant courses) = 49
- FP: False Positives (incorrectly recommended irrelevant course) = 1
- FN: False Negatives = 0
- TN: Not applicable in this case, as only relevant recommendations were considered

Precision= TP / (TP + FP) = 49 / 50

Recall= TP / (TP + FN) = 49 / 49

F1 Score= 2 × (Precision × Recall) / (Precision + Recall)

Accuracy= (TP + TN) / Total = 49 / 50

Table 3. *Analyses Results.*

Metric	**Value**
Precision	0.98
Recall	1.00
F1 Score	≈0.99
Accuracy	0.98

The great effectiveness of the proposed semantic-based recommendation system is evident from the visual representation of the system's performance metrics, which include Accuracy, Precision, Recall, and F1 Score. According to the four separate bar charts in figure. These outcomes show how dependable the suggested technique is at recommending pertinent courses with little false positives. Technical and non-technical stakeholders can more easily understand the evaluation results thanks to the use of distinct bar charts, which are unambiguous visual indicators of the system's performance.

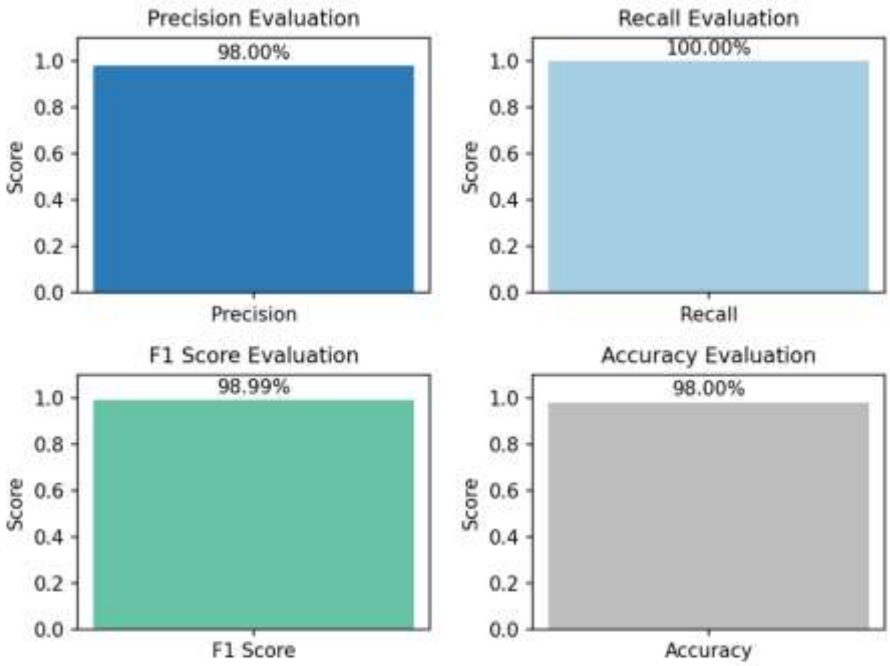

Figure 7. *System Performance Results for Generating 50 Training Course Recommendations*

Test B. We expanded the scope of testing to verify the robustness and accuracy of the recommendation system, going beyond basic precision and recall metrics. A multi-level evaluation strategy was applied. The evaluation scope was expanded to include 500 randomly selected Coursera courses. Of these courses:

480 were rated as relevant to user preferences.

20 course was irrelevant.

And these were our results.

Table 4. *Extended Metrics*.

Metric	Average Value
Precision	0.959
Recall	1.0
F1-Score	0.979
accuracy	95.99%

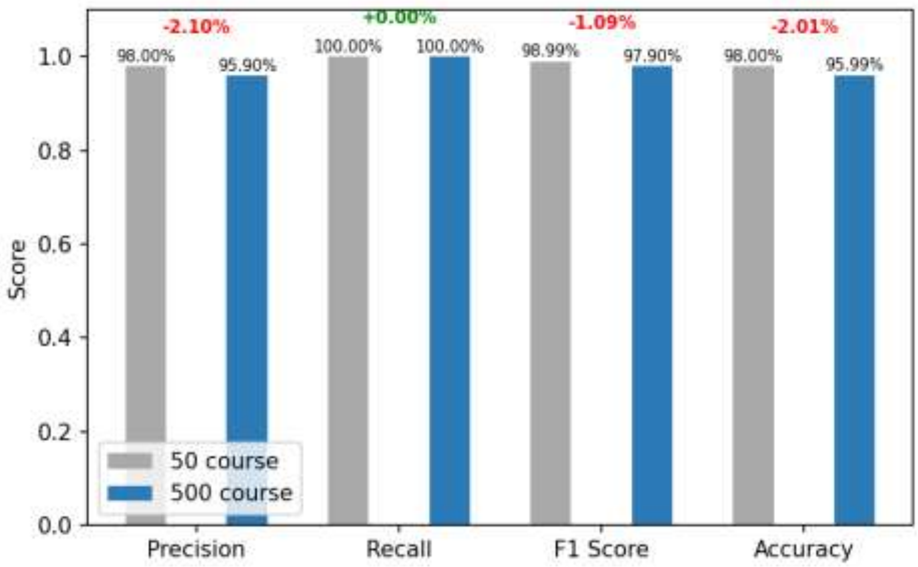

Figure 8. *Comparison of Evaluation Metrics: 50 Course vs 500 Course*

Figure 8 compares the system's performance using two different learning courses: 50 and 500 courses. Both sets achieve high performance on metrics like precision, recall, F1 score, and accuracy. The 50-cycle set shows slightly higher precision and F1 score, but these differences are small and don't significantly impact recommendation quality. The system maintains high efficiency as data volume increases, enhancing reliability and generalizability to larger learning environments.

4. Comparison and Discussion

Table 1, summarizes recent advancements in recommendation systems for online learning platforms, reflecting a diverse range of methodologies, datasets, and application contexts—spanning Coursera, Udemy, NPTEL, and social e-learning environments. A key differentiation lies in the incorporation of ontologies and semantic technologies. While most studies (e.g.,[9],[10],[11]) rely on traditional machine learning or deep learning techniques without semantic structuring, other works such as [13]. and [15] illustrate the benefits of ontology-based systems, achieving notably high accuracy rates of 97% and 97.68%, respectively. Similarly, the proposed model, **Ontology Course Recommender**, utilizes a content-based recommendation strategy enriched with semantic features and applied to the Coursera platform. Experimental results affirm the effectiveness of this approach: when generating 50 recommendations, the model achieved a Precision of 0.98, Recall of 1.00, F1-Score of approximately 0.99, and Accuracy of 0.98. When the recommendations-generation is scaled to 500 recommendations, it maintained strong performance, with an average Precision of 0.959, Recall of 1.0, F1-Score of 0.979, and Accuracy of 95.99%. These results further substantiate the efficacy of ontology-driven personalization in enhancing the relevance and precision of recommendations in e-learning environments.

Traditional approaches like content-based filtering and collaborative filtering remain robust. For instance, [10] employed Random Forest and SVD models, achieving a precision of 86.21%, while [11]used LSTM networks, reporting an 84.53% accuracy and a 77.31% F1 score. More recent trends explore the potential of deep learning and large language models (LLMs); [17], for example, utilized Mistral-7B-Instruct with over 78% accuracy, and [12] applied ensemble learning with promising results.

Despite these advances, scalability remains a key challenge, primarily due to the limited diversity of datasets and the inflexibility of many current system architectures. Few studies—such as [14] and [15]have investigated real-time adaptive mechanisms through clustering and semantic frameworks. Additionally, psychological and pedagogical aspects are often overlooked in user modeling. Although [9] attempted to integrate educational psychology, this area remains underexplored.

In summary, the integration of Semantic Web technologies, deep learning methods, and educational theories represents a promising direction for developing personalized, interpretable, and adaptive learning systems in real time. This integration can contribute to building more accurate, effective, and learner-centered learning environments by delivering personalized educational content tailored to the needs and learning styles of different individuals. Recent research indicates that large models, such as Large Language Models (LLMs) and Ensemble Models, have proven effective in personalization due to their ability to understand context and interact naturally with users. However, traditional recommendation techniques remain effective in some contexts, but their integration with Semantic Web and Deep Learning approaches offers greater potential for adaptability and personalization.

5. Conclusion

We presented a content-based recommendation system for Coursera leveraging web scraping, data preprocessing, ontology modeling, and TF-IDF-based similarity. Online learning is convenient, but it also presents a number of new issues that we must research and resolve. Resources for online learning are expanding geometrically in tandem with the market's growing demand for online education. Convenient learning resources are one of its benefits, but it also makes it difficult for users to locate and match their personal learning resources among the many available. Users are often distracted by too much information and are unable to select the initial target. Integrating learner profile with natural language processing (NLP) techniques to extract implicit preferences and learning qualities from user-generated text, such as learning objectives, course feedback, or browsing behavior, is one promising avenue. Transformer-based models, for example, can be utilized to more accurately link pertinent course content with semantic patterns in user input. By offering dynamic suggestions based on intricate learner attributes like learning history, course completion behavior, or inferred interests, rule-based reasoning via Semantic Web Rule Language could also improve the personalization layer.

Language-independent ontologies and multilingual NLP models can be used to increase the system's adaptability in multicultural and multilingual environments, which is another area for advancement. In order to increase the accuracy and diversity of course recommendations, the recommendation engine may also profit from hybrid approaches that combine content-based filtering with collaborative filtering or ensemble learning strategies.

Lastly, to guarantee generalizability and prevent overfitting, more thorough assessment techniques like k-fold cross-validation could be used, particularly as the system grows to accommodate bigger datasets and a wider range of user demographics. The existing framework can develop into a more responsive, inclusive, and intelligent recommendation system that satisfies the changing demands of a global learner base by investigating these improvements.

6. Future Work

By adding knowledge graphs and domain-specific data to augment the core ontology, we aim to improve the semantic representation of course material. More accurate semantic reasoning will be enabled, facilitating the matching of advanced courses based on student goals, skill development, and prior knowledge. The accuracy of recommendations can be improved and dynamically adapt to evolving learner behavior by incorporating real-time user interaction data, such as clickstreams, time spent on course pages, and user comments. To facilitate compatibility with other educational platforms and standards, we also intend to explore strategies for ontology alignment and integration. This concept may extend beyond Coursera and facilitate cross-platform recommendations.

Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.
- **Conflict of interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper
- **Acknowledgement:** The authors declare that they have nobody or no-company to acknowledge.
- **Author contributions:** The authors declare that they have equal right on this paper.
- **Funding information:** The authors declare that there is no funding to be acknowledged.
- **Data availability statement:** The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

[1] Jena, K. K., et al. (2023). E-learning course recommender system using collaborative filtering models. *Electronics (Switzerland), 12*(1). https://doi.org/10.3390/electronics12010157

[2] Ma, F. (2025). Learning behavior analysis and personalized recommendation system of online education platform based on machine learning. *Computers and Education: Artificial Intelligence, 8*. https://doi.org/10.1016/j.caeai.2025.100408

[3] Amanoul, S. V., Abdulrahman, L. M., Abdullah, R. M., & Qashi, R. (2023). Orchestrating distributed computing and web technology with semantic web and big data. *Journal of Smart Internet of Things, 2023*(2), 174–192. https://doi.org/10.2478/jsiot-2023-0019

[4] Gan, B., Zhang, C., Dong, Q., & Sun, W. (2021). Design of online course knowledge recommendation system based on improved learning diagnosis model. In *Journal of Physics: Conference Series*. IOP Publishing. https://doi.org/10.1088/1742-6596/1861/1/012052

[5] Abdulazeez, D., & Salah, R. (2020). Developing an ontology for retrieving massive open online courses (MOOCs) information in Coursera platform. *The Journal of The University of Duhok*, *23*(1), 103–114. https://doi.org/10.26682/sjuod.2020.23.1.11

[6] Hazar, M. J., Zrigui, M., & Maraoui, M. (2022). Learner comments-based recommendation system. In *Procedia Computer Science*, 2000–2012. https://doi.org/10.1016/j.procs.2022.09.259

[7] Islam, N., Khan, A. R., Ahmed, U., Jaffer, A., & Akhtar, S. (2022). Course recommendation, exploratory data analysis and visualizations of massive open online courses (MOOCs). *Journal of Independent Studies and Research: Computing, 20*(1). https://doi.org/10.31645/JISRC.22.20.1.9

[8] Ben Hadj Boubaker, N., Kodia, Z., & Yacoubi Ayadi, N. (2024). Learning knowledge graph-based recommender system using ensemble attention networks. In *16th International Conference on Management of Digital Ecosystems*. https://hal.science/hal-04892534v1

[9] Wei, X., Sun, S., Wu, D., & Zhou, L. (2021). Personalized online learning resource recommendation based on artificial intelligence and educational psychology. *Frontiers in Psychology, 12*. https://doi.org/10.3389/fpsyg.2021.767837

[10] Amin, S., Uddin, M. I., Mashwani, W. K., Alarood, A. A., Alzahrani, A., & Alzahrani, A. O. (2023). Developing a personalized e-learning and MOOC recommender system in IoT-enabled smart education. *IEEE Access, 11*, 136437–136455. https://doi.org/10.1109/ACCESS.2023.3336676

[11] Ezaldeen, H., Bisoy, S. K., Misra, R., & Alatrash, R. (2023). Semantics aware intelligent framework for content-based e-learning recommendation. *Natural Language Processing Journal, 3*, 100008. https://doi.org/10.1016/j.nlp.2023.100008

[12] Putra, S. A., et al. (2024). The development of intelligent web for rural social e-learning. *Jurnal Nasional Teknik Elektro dan Teknologi Informasi, 13*(3), 112–121. https://doi.org/10.22146/jnteti.v13i3.9872

[13] Bousalem, S., Benchikha, F., & Marir, N. (2025). Personalized learning through MBTI prediction: A deep learning approach integrated with learner profile ontology. *IEEE Access.* https://doi.org/10.1109/ACCESS.2025.3542701

[14] Ma, F. (2025). Learning behavior analysis and personalized recommendation system of online education platform based on machine learning. *Computers and Education: Artificial Intelligence, 8*. https://doi.org/10.1016/j.caeai.2025.100408

[15] Wang, L., Han, W., & Xi, Z. (2025). Exploring semantic web tools in education to boost learning and improve organizational efficiency. *International Journal on Semantic Web and Information Systems, 21*(1). https://doi.org/10.4018/IJSWIS.370315

[16] Hazar, M. J., Jaballi, S., Maraoui, M., Zrigui, M., & Nicolas, H. (2025). A hybrid e-learning recommendation system incorporating user reviews and ratings for enhanced course selection. https://doi.org/10.21203/rs.3.rs-5729775/v1

[17] Dutta, S., Beier, F., & Werth, D. (2025). AI-based personalized multilingual course recommender system using large language models. In *International Conference on Agents and Artificial Intelligence* (pp. 1069–1076). Science and Technology Publications. https://doi.org/10.5220/0013260100003890

[18] Cheok, S. M., Hoi, L. M., Tang, S. K., & Tse, R. (2021). Crawling parallel data for bilingual corpus using hybrid crawling architecture. In *Procedia Computer Science* (pp. 122–127). https://doi.org/10.1016/j.procs.2021.12.218

[19] Vo, N. N. Y., Vu, Q. T., Vu, N. H., Vu, T. A., Mach, B. D., & Xu, G. (2022). Domain-specific NLP system to support learning path and curriculum design at tech universities. *Computers and Education: Artificial Intelligence, 3*. https://doi.org/10.1016/j.caeai.2021.100042