Predicting Online Learners' Performance Through Ontologies: A Systematic Literature Review

Safa Ridha Albo Abdullah* and Ahmed Al-Azawei
*Department of Software, College of Information Technology, University of Babylon, Babel, Iraq; *Corresponding author*

Abstract

This systematic review sheds light on the role of ontologies in predicting achievement among online learners, in order to promote their academic success. In particular, it looks at the available literature on predicting online learners' performance through ontological machine-learning techniques and, using a systematic approach, identifies the existing methodologies and tools used to forecast students' performance. In addition, the environment for generating ontologies, as considered by academics in the field, is likewise identified. Based on the inclusion criteria and by adopting PRISMA as a research methodology, seven studies and two systematic reviews were selected. The findings reveal a scarcity of research devoted to ontologies in the prediction of learners' achievement. However, the research outcomes suggest that building an ontological model to harness machine-learning capabilities could help accurately predict students' academic performance. The results of this systematic review are useful for higher education institutes and curriculum planners. This is especially pertinent in online learning settings to avoid dropout or failure. Also highlighted in this study are numerous possible directions for future research.

Keywords: data mining, decision tree, education, ontology, Semantic Web, classification algorithm

Introduction

Online learning has been established as an alternative to traditional face-to-face learning, representing a close fusion of contemporary education and information technology (IT). Moreover, it is regarded by some scholars as crucial to the advancement of educational equity (Hussain et al., 2018). It should be made clear that online learning, open and distributed learning (ODL), and distributed education (DE) are not synonyms. Although online learning is concerned with the delivery of education via digital means, open and distributed learning (ODL) refers to educational approaches that enable access to learning materials and opportunities independent of a geographical location. Distributed education (DE), on the other hand, is an educational method that emphasizes offering learning experiences and materials across places and platforms, frequently using technology to support learning that is not limited to traditional classroom environments (Prinsloo et al., 2022).

Online learning is extensively employed worldwide due to its high degree of flexibility in terms of time and place, low bar for knowledge acquisition, and abundance of learning resources (Qiu et al., 2022). Platforms enable students to access a variety of learning activities, including reading, downloading, submitting papers, uploading content, and designing and delivering presentations, all achievable from any location at any time (Chweya et al., 2020). The adoption of online learning has therefore been growing, due to its benefits in comparison with conventional learning systems. Without being required to attend a class, students can access learning materials via online learning systems (El Aissaoui & Oughdir, 2020).

Nevertheless, this does not mean that online learning is without limitations. The lack of direct interaction has been indicated as a factor of academic failure or unsatisfactory grades among significant numbers of students in online learning environments (Mogus et al., 2012). Moreover, the student dropout rate is higher in online learning courses than in more traditional settings. This is especially evident in developing countries, since there are still many obstacles to the successful implementation of online learning in the developing world, such as low student motivation, difficulties with direct teacher-student connection, and poor access to the online learning environment (Al-Azawei, 2017; Al-Azawei et al., 2017).

Consequently, there is a need for efficient approaches to reduce the risk of failure by predicting learners' performance at an early stage (Aslam et al., 2021). Early prediction is a recently recognised phenomenon that involves forecasting outcomes to help educational institutions understand how to help students complete their online courses successfully. However, an accurate prediction of students' academic success necessitates a thorough comprehension of the variables and characteristics that could affect their achievement (Yağcı, 2022). Moreover, according to the literature (Sultana et al., 2019), predicting students' success is a complicated task because of the massive amounts of data kept in contexts such as educational and learning management databases. Earlier research, therefore, attempted to identify features that could lead to enhancing the prediction of online learners' performance (Abdullah & Al-Azawei, 2024; Al-Masoudy & Al-Azawei, 2023). Online learning performance refers to the efficacy in which students attain educational goals in a digital learning environment, so it includes a variety of measures, including engagement, completion rates, exam scores, and overall course satisfaction (Kara, 2020). However, the emergence of the Semantic Web has facilitated this process.

The Semantic Web is an extension of the traditional Web, serving as an invaluable tool for expanding and deepening the comprehension of information between people and computers (Pelap et al., 2018), consequently revolutionizing online and distance learning. To enhance the quality of educational content and provide learning activities that are tailored to the needs of each student, education systems employ Semantic Web technologies such as ontologies and semantic rules. The objective of an ontology in this instance is to help learners achieve their learning objectives by enabling them to transfer their learner profiles between the components of an e-learning system (Bolock et al., 2021).

In the education context, ontologies have been developed to gather data and categorise learning content, thereby facilitating human-machine communication (Zeebaree et al., 2019). Thus, there are numerous advantages to be gained by implementing ontologies in online or e-learning, such as enhancing students' retention, implementing timely interventions to assist students at risk of failure, determining the factors that influence students' academic performance, and improving the quality of education in practice (Al-Yahya et al., 2015). In another study (Icoz et al., 2015), ontologies were proposed to create conceptual maps to expose students to different experiences. However, there is still a dearth of systematic reviews concerning the use of ontologies to predict academic achievement among online learners, especially within the context of ODL. Ontological approaches can provide novel solutions to improve the ODL experience. By organizing educational data, ontologies enable the design of individualized and adaptable learning routes, based on each learner's particular pace and demands. Ontologies can also increase the accessibility of educational materials by ensuring that learning content is structured in a way that is easy to locate, regardless of learner's individual background, guaranteeing that each learner can succeed in ODL contexts (Wang & Wang, 2021). Hence, the present study is one of the first to undertake a systematic review of the available literature on the use of ontologies to predict students' academic performance. Eight questions were formulated to address this research gap:

RQ1. How can an ontology model be built to predict students' academic performance?

RQ2. Which techniques and learning platforms enable the prediction of students' academic performance based on ontologies?

RQ3. What are the primary research objectives of the chosen studies?

RQ4. What features are used, and what are the contexts of these studies?

RQ5. Which evaluation techniques are used?

RQ6. What are the key benefits of implementing ontologies to predict online learners' academic performance?

RQ7. In which countries were the experiments in the included papers conducted?

RQ8. When were the previous experiments on ontologies conducted?

Specifically, this paper looks at the ontologies used across a number of fields, together with previous systematic reviews on the implementation of ontologies in e-learning. It includes an outline and

justification of the methodology adopted in this systematic review and reports the research findings. The key outcomes of this study are then discussed, with a concluding section that highlights the main concepts extracted and the research limitations.

What Is an Ontology?

An ontology is a modeling tool that deploys a standard vocabulary to define and represent domains in a formal manner (Rami et al., 2018). Ontologies and the Semantic Web have been deployed in online learning in wide-ranging ways, including to convey domain knowledge, offer metadata for significant ideas and entities, promote richer definition and retrieval of educational content, enable the definition, sharing, and exchange of learning content, develop curricula, and measure learning quality (Al-Yahya et al., 2015). In short, ontologies are essential Semantic Web technologies that are regarded as the backbone of the Web (Raad & Cruz, 2015). Furthermore, they constitute one of the most successful means of organizing a body of knowledge (Al-Chalabi & Hussein, 2020).

Ontologies have been employed in numerous applications and disciplines and can be used to organize content in any field (Zeebaree et al., 2019). They are used, for example, to store the data from learning objects, actions, behaviors, feelings, and models created by students. Ontology models are composed of many different classes or sub-ontologies derived from a range of data sources (Rahayu et al., 2022). Within online learning, ontologies are frequently used to describe its services and components for collaboration between diverse systems (Al-Yahya et al., 2015). However, building an ontology is an expensive, time-consuming, and error-prone task, especially in an online learning environment (Al-Chalabi & Hussein, 2020). This is because the process of creating an ontology requires extensive skills and human experience in knowledge engineering. Hence, the acquisition of ontological knowledge in the e-learning context requires an expert in the field (George & Lal, 2019).

Previous Systematic Reviews on the Implementation of Ontologies in Online Learning

Earlier literature on the research topic did not adopt taxonomies for the educational objective of assessing how well students met their educational goals or achieved high academic performance. Moreover, while this review identified a number of related articles, it also detected a research gap, with few articles found to cover the combined use of computational ontologies and learning analytics. For example, Costa et al. (2018) investigated 21 journal articles published between 2010 and 2018, with a focus on the ways in which education analytics and computational ontologies led by taxonomy and learning goals can help evaluate academic performance.

Furthermore, Costa et al. (2020) reviewed 31 journal articles published between 2010 and 2019. These studies were analyzed from two perspectives, specifically in terms of learning analytics, computational ontologies, taxonomies of educational objectives, and the relationship between these components and academic performance. However, in both Costa et al. (2018) and Costa et al. (2020), a research gap may be

noted with regard to computational ontologies and learning analytics in the online learning context. Conversely, unlike these systematic reviews and others in the research context, the present study explores the contribution of ontological techniques to the prediction of online learners' performance.

The rationale for this study was to build an ontological model that would leverage machine-learning capabilities and help accurately predict academic achievement, while also recommending a more general approach based on concepts rather than specific features of a particular course. Therefore, this systematic literature review is expected to fill the identified research gap and inform future trends in the use of ontological techniques to predict the academic achievement of online students.

Methodology

For this study, the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram was selected to find, select, evaluate, and synthesize research in the field. The PRISMA flow diagram has been widely used in several fields since its initial publication in 2009. The rationale behind the adoption of PRISMA over other protocols includes two reasons. First, it is a recognized approach in several academic disciplines because of its comprehensiveness and acceptance in the academic community (Liberati et al., 2009; Zainuddin et al., 2024). Second, PRISMA can help researchers ensure that their literature evaluation is conducted in a systematic manner (Liu et al., 2024). The current study adopts the PRISMA checklists technique to highlight the application of ontologies in e-learning and learning technologies within the education sector (Liberati et al., 2009).

The most recent iteration of PRISMA (the 2020 statement) was adopted in this study as a guide for setting the criteria (Page et al., 2021). Several checklists are included in PRISMA, encompassing eligibility criteria (exclusion and inclusion), search strategy, data-gathering procedures, method of selecting studies, synthesis approach, and synthesis results. Figure 1 presents the key steps of the PRISMA research methodology.

Figure 1

Flow of Studies From Identification to Inclusion
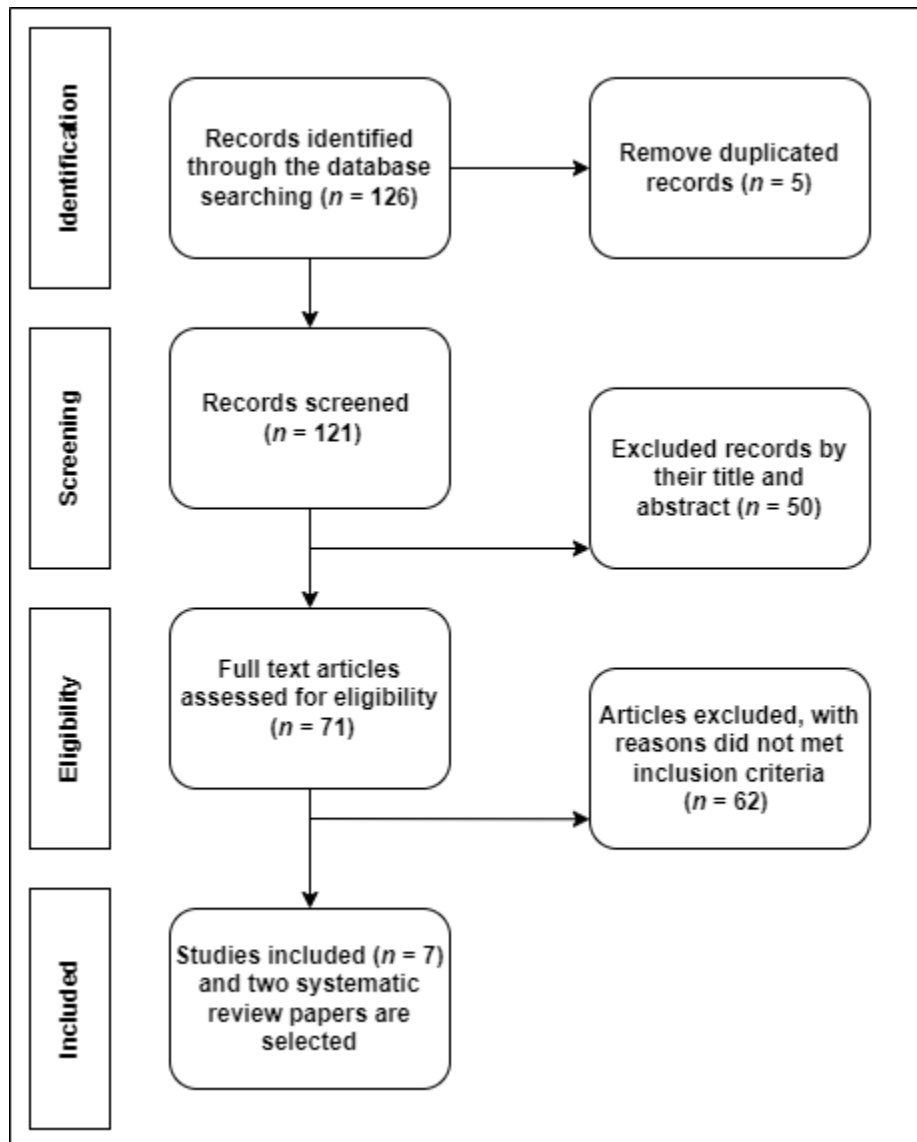
Note. Adapted from "The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews," by M. J. Page, J. E. McKenzie, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, C. D. Mulrow, L. Shamseer, J. M. Tetzlaff, E. A. Akl, S. E. Brennan, R. Chou, J. Glanville, J. M. Grimshaw, A. Hróbjartsson, M. M. Lalu, T. Li, E. W. Loder, E. Mayo-Wilson, S. McDonald, ... D. Moher, 2021, *BMJ*, *372*(71), 5. (https://doi.org/10.1136/bmj.n71). CC BY 4.0.

Eligibility Criteria

The papers included in this systematic review were evaluated and selected based on precise inclusion and exclusion criteria. These inclusion and exclusion criteria are set out in Table 1.

Table 1

Inclusion and Exclusion Criteria

Inclusion criteria	Exclusion criteria
Research involving ontology and the prediction of online learners' performance.	Research including an ontology that was not used to predict online learners' performance.
Research studies written in English.	Research studies written in a language other than English.
Research with empirical findings, regardless of geographical boundaries.	Research involving an ontology used in circumstances related to learning as a whole, but not for predicting the achievement of online learners.
Research published between 2010 and 2023.	Research published before 2010.
Research published in peer-reviewed journals.	Research published in journals, etc., with no peer-review process.

Search Strategy

This study involved a systematic review of literature in seven leading databases and search engines, namely IEEE, Scopus, ACM, ERIC, Science Direct, Springer, and Google Scholar. The review was conducted between January and March 2023. However, the papers sourced were published between 2010 and 2023, as illustrated in Table 2. The keywords used for the database search were "e-learning," "Semantic Web," "ontology," "prediction of online learners' performance," and "learners' performance."

Table 2

Number of Studies by Database or Search Engine and Year of Publication

Database or Search Engine	n		
	2010–2016	2017–2023	Total
Scopus	5	11	16
IEEE	17	8	25
Science Direct	11	10	21
Google Scholar	15	8	23
Springer	10	11	21
ACM	7	11	18
ERIC	6	5	11
Total	71	55	126

Selection of Related Articles

After deleting duplicates ($n = 5$), 121 peer-reviewed papers were initially obtained. The researcher carefully analysed the abstracts and conclusions of these papers and scanned the content to determine which studies satisfied the inclusion criteria. In the screening stage, fifty articles were excluded based on title/abstract. This led to reviewing with a total of 71 papers. In the eligibility step, 62 articles were removed because they did not meet the inclusion criteria. Such articles did not include empirical findings or merely offered a framework for integrating ontologies. Finally, only seven papers and two systematic reviews were included.

Quality Assessment

Along with inclusion and exclusion criteria, another element that may be employed in the selection of studies for a systematic review is the quality rating of an article. The quality evaluation checklist used in this study defined nine criteria for evaluating the quality of the research retained for further analysis. For each of the nine criteria, a study received one point for "yes" and zero points for "no." A total score ranging from 0 to 9 was then awarded to each study, with a higher score indicating that the study was of good quality. The nine criteria include the following:

Q1. Have the goals of the study been outlined?

Q2. Were the research goals described?

Q3. Is the methodology understandable?

Q4. Was the research empirically tested?

Q5. Have all techniques and resources been well defined?

Q6. Were the findings satisfying?

Q7. Does the study increase researchers' comprehension or knowledge?

Q8. Have the data collection techniques been sufficiently described?

Q9. Do the researchers explain the problems?

Table 3 displays the results of the quality assessment checklist for the selected studies. Based on these results, these seven studies are of a high enough quality that would make them suitable for future investigation.

Table 3

Results of the Quality Assessment on the Selected Studies

Study	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Total
El-Rady (2020)	1	1	1	1	1	1	1	1	1	9
Costa et al. (2021)	1	0	1	0	1	0	0	1	1	5
El-Rady et al. (2017)	1	1	1	1	1	1	1	1	1	9
Boufardea & Garofalakis (2012)	1	1	1	1	1	0	0	1	1	7
Grivokostopoulou et al. (2014)	1	1	1	1	1	1	1	1	1	9
Hamim et al. (2021)	1	1	1	1	1	1	0	1	1	8
López-Zambrano et al. (2022)	1	1	1	1	1	1	1	1	1	9

Results and Discussion

The results presented in this systematic review are based on the nine research objectives extracted from the selected studies, published between 2010 and 2023.

Building an Ontology Model

The first research question (RQ1) asked: How can an ontology model be built to predict students' academic performance? This study found that the prediction model in the primary studies used an ontology as a learning model. Moreover, an ontological model of prediction was designed or implemented in many tools and settings (Grivokostopoulou et al., 2014). The studies identified the Protégé tool as a popular environment for creating and developing ontologies, as shown in Table 4. The result demonstrates widespread adoption of ontology language standards.

Although the World Wide Web Consortium (W3C) provides many language standards, the Web Ontology Language (OWL) was employed in nearly half the primary research selected for this study. Moreover, Table 4 reveals that ontology languages were used in the prediction models (El-Rady, 2020; Costa et al., 2021; El-Rady et al., 2017; Boufardea & Garofalakis, 2012; Grivokostopoulou et al., 2014; Hamim et al., 2021; López-Zambrano et al., 2022). In El-Rady et al. (2017), Grivokostopoulou et al. (2014), and Hamim et al. (2021), the Protégé tool was used to build each rule in the Semantic Web Rule Language (SWRL), based on a decision tree. The rules were then converted to SWRL and added to the ontology. In other experiments,

ontologies were combined with an intelligent recommendation system. López-Zambrano et al. (2022) presented a new technique using Semantic Web resources. This technique creates an ontology based on an activity taxonomy to capture how students engage with the Moodle learning management system. Another study found that an ontology learning goal was used to generate knowledge corresponding to the cognitive level of Bloom's revised learning objective taxonomy (Costa et al., 2021). Because it typically takes a long time to build an ontology, it was proposed that METHONTOLOGY could be used to build a learner ontology model (El-Rady, 2020), but the environment for creating ontologies was not specified (Costa et al., 2020).

Table 4

Construction of Ontologies in the Selected Studies

Study	Ontology language	Tool
El-Rady (2020)	OWL	Protégé
Costa et al. (2021)	RDF/XML	Protégé
El-Rady et al. (2017)	OWL	Protégé
Boufardea & Garofalakis (2012)	RDF/XML	Protégé
Grivokostopoulou et al. (2014)	-	Protégé
Hamim et al. (2021)	OWL\DL	Protégé
López-Zambrano et al. (2022)	-	Protégé

Note. OWL = Web Ontology Language; RDF = Resource Description Framework; XML = Extensible Markup Language; DL = Description Logic.

Effective Techniques and Learning Platforms

The second research question (RQ2) asked: Which techniques and learning platforms enable the prediction of students' academic performance based on ontologies? Table 5 tabulates the machine-learning techniques used with the ontologies. The review found that various machine-learning techniques were implemented alongside ontologies to predict academic performance. This indicates that the studies selected either used a decision tree algorithm alone or in combination with other machine-learning techniques.

Three studies solely employed the decision tree technique (Boufardea & Garofalakis, 2012; Grivokostopoulou et al., 2014; López-Zambrano et al., 2022), three studies used different machine-learning techniques (El-Rady, 2020; Costa et al., 2021; El-Rady et al., 2017), and one study implemented linear and logistic regression algorithms (Costa et al., 2021). Regarding the type of platform used, six studies (Costa et al., 2021; El-Rady et al., 2017; Boufardea & Garofalakis, 2012; Grivokostopoulou et al., 2014; Hamim et al., 2021; López-Zambrano et al., 2022) did not specify the type of platform, whereas one study clearly mentioned the use of Moodle (El-Rady, 2020).

Table 5

Machine-Learning Algorithms Used in the Selected Studies

Study	Technique
El-Rady (2020); El-Rady et al. (2017)	Bayes Net, Naive Bayes, random forest, AdTree, J48, IB1, KSTAR, JRip, OneR, SMO, SimpleLogistics, LWL, simple CART
Costa et al. (2021)	Linear and logistic regression algorithms
Boufardea & Garofalakis (2012)	J48, C4.5
Grivokostopoulou et al. (2014)	J48 and the CART (decision tree)
Hamim et al. (2021)	C5.0
López-Zambrano et al. (2022)	J48 (decision tree)

Table 5 presents various machine learning algorithms that were used in earlier research for the classification tasks. This includes, but are not limited to, decision trees, Bayesian networks, and ensemble methods. For further details, researchers can refer to a research study conducted by Tan, Steinbach, and Kumar (2016).

Primary Research Objectives

RQ3 asked: What are the primary research objectives of the chosen studies? In response, each paper was classified into one of four categories: (a) predicting final performance to help tutors obtain deeper insights; (b) tracing the achievement of students who are underperforming or in danger of failing; (c) presenting apposite references and recommendations to each student to promote their success on a course and drive broader pedagogical improvement, and (d) improving prediction model portability in terms of predictive accuracy.

Table 6 shows that the majority of studies selected traced underperforming students ($n = 6$). The second-highest number ($n = 4$) offered sound recommendations and appropriate advice. One study measured students' understanding and improved the portability of prediction models, whereas another evaluated students' academic performance, and a further study monitored the status of students' knowledge.

Table 6

Objectives of the Selected Studies

Objective	Studies
Predicting final performance	El-Rady, (2020); Costa et al. (2021); El-Rady et al. (2017); Boufardea & Garofalakis(2012); Grivokostopoulou et al. (2014); Hamim et al. (2021); López-Zambrano et al. (2022)

Tracing students' achievement	El-Rady, (2020); Costa et al. (2021); El-Rady et al. (2017); Grivokostopoulou et al. (2014); Hamim et al. (2021)
Presenting apposite references and recommendations	Costa et al. (2021); Boufardea & Garofalakis(2012); Grivokostopoulou et al. (2014)
Improving the portability of the prediction model	López-Zambrano et al. (2022)

Features and Context

The fourth research question (RQ4) was: What are the features used and what are the contexts of these studies? Table 7 shows the features used in all the articles collected, revealing that they were conducted using educational datasets or learning environments based on online learning or e-learning modes (El-Rady, 2020; Costa et al., 2021; El-Rady et al., 2017; Boufardea & Garofalakis, 2012; Grivokostopoulou et al., 2014; Hamim et al., 2021; López-Zambrano et al., 2022).

In some educational domains, while the datasets are too small (El-Rady, 2020; Costa et al., 2021; El-Rady et al., 2017; Boufardea & Garofalakis, 2012; Grivokostopoulou et al., 2014; Hamim et al., 2021; López-Zambrano et al., 2022), in others, they are usually large enough (El-Rady, 2020; Hamim et al., 2021; López-Zambrano et al., 2022). Moreover, only one of the studies reviewed failed to mention the features of the dataset used (Costa et al., 2021).

Table 7

Contexts and Features Used in the Selected Studies

Study	Features used to build ontologies	Context
El-Rady, (2020)	Average number of comments, posts and likes submitted by the learner on Facebook groups for the course, learner's age, address, gender, learner's activities (related or unrelated to the curriculum), number of sessions attended, grades for exercises, mid-term grades, family members, average time spent on learning, number of previous failures, final grade	E
Costa et al. (2021)	Not mentioned	E
El-Rady et al. (2017)	Number of sessions attended, grades for exercises, mid-term grades, learner's activities, learner's age, learner's address, average number of comments, gender, family members, average time spent learning, number of previous failures, learner's activities, posts and likes submitted by learner on Facebook groups for the course	O

Boufardea & Garofalakis(2012)	Grades for projects, grades for exams, the type of exam, student's gender, student's age, student's marital status, and educational background	O
Grivokostopoulou et al. (2014)	Marks in tests, student's gender, academic year, marks in final exams for the course	O
Hamim et al. (2021)	Personal identity, social identity, digital identity, family background, personality, professional experience, physical limitations, knowledge profile, learning profile, academic background, cognitive profile	O
López-Zambrano et al. (2022)	Questionnaires, quizzes, surveys, forums, glossaries, assignments, databases, chats, choice, lessons, workshops, scorepackages, wikis	O

Note. E = e-learning; O = online learning.

Evaluation Techniques

The next research question (RQ5) was: Which evaluation techniques are used? Table 8 illustrates the techniques used to evaluate the proposed ontologies with machine-learning algorithms. There was no specific evaluation methodology to assess the quality of the ontologies developed. The studies variously applied an evaluation matrix (El-Rady, 2020; El-Rady et al., 2017; Boufardea & Garofalakis, 2012; Grivokostopoulou et al., 2014; Hamim et al., 2021) or other evaluation techniques (López-Zambrano et al., 2022). Only one study implemented interviews and questionnaires to assess the quality of the suggested architecture's viability, taking into account how it would affect the teaching and learning process (Costa et al., 2021).

Table 8

Evaluation Techniques in the Selected Studies

Study	Evaluation technique
El-Rady, (2020); El-Rady et al. (2017); Boufardea & Garofalakis (2012); Grivokostopoulou et al. (2014); Hamim et al. (2021)	Accuracy, precision, recall, accuracy matrix
Costa et al. (2021)	Interviews, questionnaires
López-Zambrano et al. (2022)	Area under AUC loss, ROC curve

Note. AUC = Area Under the Curve; ROC = Receiver Operating Characteristic.

Table 9 illustrates the outcomes of assessing the performance of the various methods adopted in the selected studies. Data-mining techniques, such as the J48 algorithm, yielded better accuracy and performance (87%; Grivokostopoulou et al., 2014). The random forest algorithm was used to predict whether students would pass or fail based on a dataset, attaining 91.36% accuracy (El-Rady et al., 2017). In another study, the C5.0 algorithm achieved 83.6% accuracy (Hamim et al., 2021). Similarly, in Boufardea & Garofalakis (2012), the C5.0 algorithm outperformed other studies (85.2%), and the AUC metric was found to yield 0.63% accuracy

(López-Zambrano et al., 2022). Ultimately, El-Rady, (2020) achieved the best accuracy (95.8%), using the J48 and random forest algorithms.

Table 9

Overall Accuracy of the Algorithms as Reported in the Included Studies

| Study	Techniques	Accuracy, %
El-Rady, (2020)	Random forest and J48 algorithm	95.8
El-Rady et al. (2017)	Random forest algorithm	91.36
Boufardea & Garofalakis (2012)	C5.0 algorithm	85.2
Grivokostopoulou et al. (2014)	J48 algorithm	87
Hamim et al. (2021)	C5.0 algorithm	83.6
López-Zambrano et al. (2022)	J48 algorithm (decision tree)	0.63

Key Benefits

Research question six (RQ6) asked: What are the key benefits of implementing ontologies to predict online learners' academic performance? In response, this research found that ontological machine-learning techniques were successfully implemented in online learning. Each student in the selected studies had received learning activities via ontologies, which also helped improve the quality of the curriculum. Moreover, ontologies were used to represent students' characteristics and the teaching domain (Grivokostopoulou et al., 2014). In El-Rady et al. (2017) and Boufardea & Garofalakis (2012), the aim was to predict learners' progress or status and generate information about students' final performance. The application of ontologies was also aimed at predicting students' failure, success, and dropout rate (Hamim et al., 2021). In terms of predictive accuracy, the use of ontologies appeared to enhance the portability of the prediction models, as ontological models developed on one course could be used for other course goals at different levels of application, without sacrificing their prediction accuracy (López-Zambrano et al., 2022). Additionally, some of the included papers emphasised how the ontology model could help identify students who required extra help and make appropriate recommendations to close the gaps in students' learning and lower the student failure ratio (Grivokostopoulou et al., 2014).

Country Context

The next research question (RQ7) asked: In which countries were the experiments in the included papers conducted? Figure 2 depicts the articles' distribution across countries, indicating that the studies took place in Brazil (n = 3), Greece (n = 2), Egypt (n = 2), Spain (n = 1), and Morocco (n = 1).

Figure 2

Distribution of Articles Across Countries
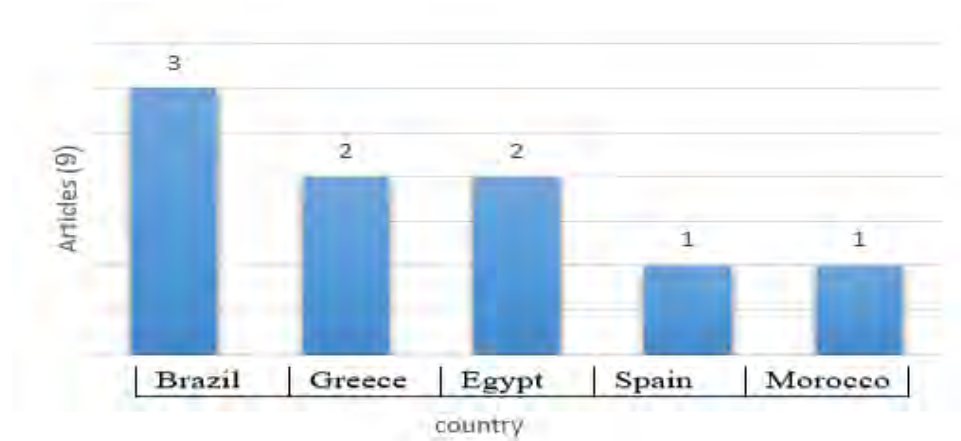
Time of Previous Experiments

The final research question (RQ8) asked: When were the previous experiments on ontologies conducted? Figure 3 presents the selected studies organized by year of publication. This classification served to build a clearer picture of the distribution of publications for the past 13 years and a summary of the research output in this field.

Figure 3

Distribution of Selected Articles According to Year of Publication
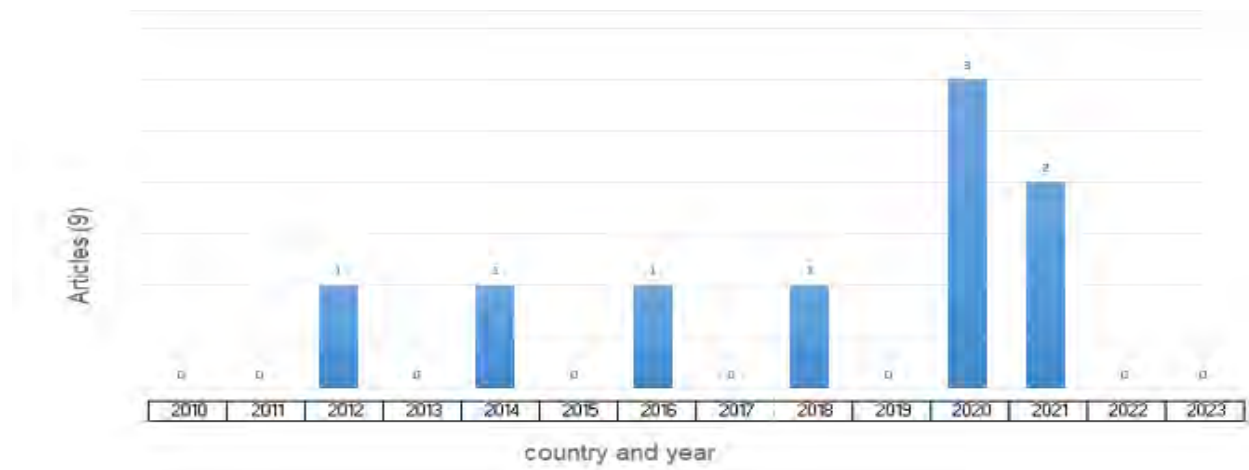
It is clear from Figure 3 that from 2010 to 2023, there was no substantial increase in the number of publications on ontological techniques for forecasting the success of online learners. In eight of those years, no such papers were published. In 2012, 2014, 2016, and 2018, one journal article on this topic was published for each year, whereas in 2020, there were three. This increase may be attributed to the evolution

in associated areas such as Web 3.0 and e-learning, or the widespread use of online learning during the COVID-19 pandemic.

The use of ontological approaches to predict online academic success provides a novel way to understand and promote students' progress. This approach can open the door for more flexible, responsive, and successful educational systems as it can help define learners' characteristics and then improve customization of the learning experience.

This research highlights how most of the ontological methods reported in the literature employed Protégé software to create ontologies, but two studies failed to define the environment for ontology creation. The OWL representation language was identified as a common tool for encoding ontologies, whereas the RDF representation language was less frequently used; three studies used an OWL representation and two used an RDF representation language, as shown in Table 4. Conversely, other selected studies did not mention the representation language used.

Among the machine-learning techniques mentioned in the selected studies, the decision tree technique was found to be the most commonly deployed. Meanwhile, different machine-learning techniques were applied in three of the studies, but two studies omitted to specify the technique used to forecast students' academic performance, as shown in Table 5. The findings show there was no substantial increase in the number of journal articles published between 2010 and 2023 on the topic of ontology-based techniques for predicting academic performance. This means that related research may still be facing some issues in implementing ontological prediction techniques. The purpose of most of the studies shown in Table 6 was to investigate the accuracy of using ontologies to predict students' performance. However, some studies were designed to serve more than one purpose.

To evaluate the research findings, several techniques were employed: a convolution matrix, AUC loss, area under the ROC curve, interviews in five of the studies reviewed, and a questionnaire in one study. In contrast, three studies omitted to mention any evaluation technique, as illustrated in Table 8. Among the machine-learning techniques used in the selected studies, the highest accuracy (95.8%) was achieved with the random forest and J48 algorithms, as displayed in Table 9.

Regarding the location of these studies, three were carried out in Brazil, two in Greece, two in Egypt, one in Spain, and one in Morocco, as shown in Figure 2. Furthermore, there was no substantial rise in the number of publications on ontology-based techniques for predicting academic performance between 2010 and 2023. See Figure 3.

Despite the advantages of implementing ontologies to predict students' academic achievement, there remain many research gaps. First, experiments and prediction models should be trained on large datasets. In addition, artificial datasets might not reflect students' actual performance or behaviour (Boufardea & Garofalakis, 2012). Second, most of the included studies employed machine-learning algorithms (López-Zambrano et al., 2022). However, the use of deep learning algorithms with an ontology could help improve the prediction of academic performance. Third, a common factor among the reviewed studies was that they all relied solely on the original features of the datasets, whereas generating new features could lead to better prediction accuracy (Al-Azawei & Al-Masoudy, 2020). Fourth, some researchers have mainly used Protégé

with a Pellet reasoner, but there are many other reasoners that could be used. Reasoners that also support Protégé are Snorocket, RACER, FACT++, HermiT, CEL, ELK SWRL-IQ, and TrOWL (Nafea et al., 2016). Furthermore, although this systematic review was conducted across several databases, only a few studies were found to have implemented ontological techniques to predict students' performance. This could signify that the research area requires further empirical investigation to identify its key advantages in predicting academic performance.

Finally, evaluating the performance of ontological approaches was clearly lacking from this body of literature, which could be attributed to the absence of relevant evaluation standards. In nearly half the reviewed studies, no clear evaluation techniques appear to have been implemented, which could affect the overall reliability of those ontologies. Khalilian (2019) suggested that evaluating the quality of an ontology is crucial to determine the most suitable ontology for a specific purpose.

Conclusion

This article presents a systematic review, conducted to explore the role of ontologies in predicting learners' academic achievement. The PRISMA methodology was used to collect possible research papers. Seven studies and two systematic reviews published between 2010 and 2023 were selected. This research therefore provides an inclusive review of ontological methods of predicting online students' achievement.

The reviewed papers were classified according to the techniques implemented, ontology language represented, learning platform used, classification applied, location of affiliations, year of publication, datasets implemented, number of participants, and evaluation methods deployed. This systematic review provided a clear analysis of the research gaps and limitations, in order to inform possible future trends. It is envisaged that this review will broaden the boundaries of knowledge and provide relevant literature for scholars who are interested in furthering the topic of inquiry. It also highlighted the influence of ontologies on DE and ODL theories. This lies in the fact that the use of ontology can provide tools for organizing, integrating, and analyzing educational information and processes. It can also improve the efficacy and theoretical underpinning of educational models by contributing to curriculum design, interoperability, scalability, and flexibility. Hence, such integration can lead to promoting innovation and increasing outcomes of both learners and educational institutions.

Regardless of the importance of this work, a number of research limitations should be mentioned. The first is that no repeat searches for pertinent research were conducted after a set length of time, which could mean that some papers were disregarded or released later. Second, only specific databases were used, so papers published elsewhere might not have been retrieved. Third, the results did not include research written in other languages. Fourth, the only keywords used to identify relevant studies were "online learning," "Semantic Web," "ontology," and "learners' performance." Employing additional keywords could have yielded more accurate and comprehensive results. Finally, the included studies did not explore the influence of ontologies on personalizing online learning platforms according to learners' needs and characteristics. As such, further research is invited to investigate the relationship between these two concepts.

References

Abdullah, S. R. A., & Al-Azawei, A. (2024). Enhancing the early prediction of learners' performance in a virtual learning environment. In A. M. Al-Bakry, M. A. Sahib, S. O. Al-Mamory, J. A. Aldhaibani, A. N. Al-Shuwaili, H. S. Hasan, R. A. Hamid, & A. K. Idrees (Eds.), *New trends in information and communications technology applications: 7th national conference, NTICT 2023, proceedings* (pp. 252–266). Springer. https://doi.org/10.1007/978-3-031-62814-6_18

Al-Azawei, A., & Al-Masoudy, M. A. A. (2020). Predicting learners' performance in virtual learning environment (VLE) based on demographic, behavioral and engagement antecedents. *International Journal of Emerging Technologies in Learning (IJET), 15*(9), 60–75. https://doi.org/10.3991/ijet.v15i09.12691

Al-Azawei, A., Parslow, P., & Lundqvist, K. (2017). The effect of universal design for learning (UDL) application on e-learning acceptance: A structural equation model. *The International Review of Research in Open and Distributed Learning, 18*(6), 54–87. https://doi.org/10.19173/irrodl.v18i6.2880

Al-Azawei, A. H. S. (2017). *Modelling e-learning adoption: The influence of learning style and universal learning theories* [Doctoral dissertation, University of Reading]. CentAUR: Central Archive at the University of Reading. https://centaur.reading.ac.uk/77921/

Al-Chalabi, H. K. M., & Hussein, A. M. A. (2020, June). Ontology applications in e-learning systems. In *Proceedings of the 12th International Conference on Electronics, Computers and Artificial Intelligence—ECAI—2020* (pp. 1–6). IEEE. https://doi.org/10.1109/ECAI50035.2020.9223135

Al-Masoudy, M. A. A., & Al-Azawei, A. (2023). Proposing a feature selection approach to predict learners' performance in virtual learning environments (VLES). *International Journal of Emerging Technologies in Learning (iJET), 18*(11), 110–131. https://doi.org/10.3991/ijet.v18i11.35405

Al-Yahya, M., George, R., & Alfaries, A. (2015). Ontologies in e-learning: Review of the literature. *International Journal of Software Engineering and Its Applications, 9*(2), 67–84. https://www.earticle.net/Article/A242009

Aslam, N., Khan, I. U., Alamri, L. H., & Almuslim, R. S. (2021). An improved early student's academic performance prediction using deep learning. *International Journal of Emerging Technologies in Learning (iJET), 16*(12), 108–122. https://doi.org/10.3991/ijet.v16i12.20699

Boufardea, E., & Garofalakis, J. (2012). A predictive system for distance learning based on ontologies and data mining. In T. Bossomaier & S. Nolfi (Eds.), *COGNITIVE: Proceedings of 4th International Conference on Advanced Cognitive Technologies and Applications* (pp. 151–158). International Academy, Research and Industry Association. https://personales.upv.es/thinkmind/dl/conferences/cognitive/cognitive_2012/cognitive_2012_7_30_40125.pdf

Chweya, R., Shamsuddin, S. M., Ajibade, S. S., & Moveh, S. (2020). A literature review of student performance prediction in E-learning environment. *Journal of Science, Engineering, Technology and Management*, *1*(1), 22–36.

Costa, L. A., Nascimento Salvador, L. D., & Amorim, R. R. (2018). Evaluation of academic performance based on learning analytics and ontology: A systematic mapping study. In J. Rhee (Chair), *2018 IEEE Frontiers in Education Conference* (pp. 1–5). IEEE. https://doi.org/10.1109/FIE.2018.8658936

Costa, L. A., Pereira Sanches, L. M., Rocha Amorim, R. J., Nascimento Salvador, L. D., & Santos Souza, M. V. D. (2020). Monitoring academic performance based on learning analytics and ontology: A systematic review. *Informatics in Education*, *19*(3), 361–397. https://doi.org/10.15388/infedu.2020.17

Costa, L. A., Souza, M., Salvador, L. N., Silveira, A. C., & Saibel, C. A. (2021). Students' perceptions of academic performance in distance education evaluated by learning analytics and ontologies. In *Anais do XXXII Simpósio Brasileiro de Informática na Educação—SBIE 2021* (pp. 91–102). https://doi.org/10.5753/sbie.2021.218423

El Bolock, A., Abdennadher, S., & Herbert, C. (2021). An ontology-based framework for psychological monitoring in education during the COVID-19 pandemic. *Frontiers in Psychology*, *12*, Article 673586. https://doi.org/10.3389/fpsyg.2021.673586

El Aissaoui, O., & Oughdir, L. (2020). A learning style-based ontology matching to enhance learning resources recommendation. In B. Benhala, K. Mansouri, A. Raihani, M. Qbadou, & N. El Makhfi (Eds.), *2020 1st International Conference on Innovative Research in Applied Science, Engineering and Technology—IRASET 2020* (pp. 1–7). IEEE. https://doi.org/10.1109/IRASET48871.2020.9092142

El-Rady, A. A. (2020). An ontological model to predict dropout students using machine learning techniques. In *3rd ICCAIS 2020 International Conference on Computer Applications & Information Security* (pp. 1–5). IEEE. https://doi.org/10.1109/ICCAIS48893.2020.9096743

El-Rady, A. A., Shehab, M., & El Fakharany, E. (2017). Predicting learner performance using data-mining techniques and ontology. In A. E. Hassanien, K. Shaalan, T. Gaber, A. T. Azar, & M.F. Tolba (Eds.), *Proceedings of the International Conference on Advanced Intelligent Systems and Informatics 2016* (pp. 660–669). Springer. https://doi.org/10.1007/978-3-319-48308-5_63

George, G., & Lal, A. M. (2019). Review of ontology-based recommender systems in e-learning. *Computers & Education*, *142*, Article 103642. https://doi.org/10.1016/j.compedu.2019.103642

Grivokostopoulou, F., Perikos, I., & Hatzilygeroudis, I. (2014, December). Utilizing semantic web technologies and data mining techniques to analyze students learning and predict final performance. In D. Carnegie (Chair), *TALE 2014—Proceedings of IEEE International Conference on Teaching, Assessment and Learning for Engineering* (pp. 488–494). IEEE. https://doi.org/10.1109/TALE.2014.7062571

Hamim, T., Benabbou, F., & Sael, N. (2021). An ontology-based decision support system for multi-objective prediction tasks. *International Journal of Advanced Computer Science and Applications*, *12*(12), 183–191. https://doi.org/10.14569/IJACSA.2021.0121224

Hussain, M., Zhu, W., Zhang, W., & Abidi, S. M. R. (2018). Student engagement predictions in an e-learning system and their impact on student course assessment scores. *Computational Intelligence & Neuroscience, 2018,* Article 6347186. https://doi.org/10.1155/2018/6347186

Icoz, K., Sanalan, V. A., Cakar, M. A., Ozdemir, E. B., & Kaya, S. (2015). Using students' performance to improve ontologies for intelligent e-learning system. *Educational Sciences: Theory and Practice*, *15*(4), 1039–1049. https://jestp.com/menuscript/index.php/estp/article/view/661/598

Kara, M. (2020, May 21). Influential readers [Review of the book *Distance education: A systems view of online learning*, by M. G. Moore & G. Kearsley]. *Educational Review*, *72*(6), 800. https://doi.org/10.1080/00131911.2020.1766204

Khalilian, S. (2019). A survey on ontology evaluation methods. *Digital and Smart Libraries Researches*, *6*(2), 25–34. https://doi.org/10.30473/mrs.2020.48615.1402

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *PLoS Med*, *6*(7), Article e1000100. https://doi.org/10.1371/journal.pmed.1000100

Liu, F., Ziden, A. A., & Liu, B. (2024). Emotional support in online teaching and learning environment: A Systematic literature Review (2014–2023). *Journal of Curriculum and Teaching*, *13*(4), 209–218. https://doi.org/10.5430/jct.v13n4p209

López-Zambrano, J., Lara, J. A., & Romero, C. (2022). Improving the portability of predicting students' performance models by using ontologies. *Journal of Computing in Higher Education*, *34*, 1–19. https://doi.org/10.1007/s12528-021-09273-3

Mogus, A. M., Djurdjevic, I., & Suvak, N. (2012). The impact of student activity in a virtual learning environment on their final mark. *Active Learning in Higher Education*, *13*(3), 177–189. https://doi.org/10.1177/1469787412452985

Nafea, S., Maglaras, L. A., Siewe, F., Smith, R., & Janicke, H. (2016). Personalized students' profile based on ontology and rule-based reasoning. *EAI Endorsed Transactions on E-Learning, 3*(12), Article e6. https://doi.org/10.4108/eai.2-12-2016.151720

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021, March 29). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*(71). https://doi.org/10.1136/bmj.n71

Pelap, G. F., Zucker, C. F., & Gandon, F. (2018). Semantic models in Web-based educational system integration. In M. J. Escalona, F. D. Mayo, T. Majchrzak, & V. Monfort (Eds.), *Proceedings of the 14th International Conference on Web Information Systems and Technologies—WEBIST* (pp. 78–89). SciTePress. https://doi.org/10.5220/0006940000780089

Prinsloo, P., Slade, S., & Khalil, M. (2022). Introduction: Learning analytics in open and distributed learning—Potential and challenges. In P. Prinsloo, S. Slade, & M. Khalil (Eds.), *Learning analytics in open and distributed learning* (pp. 1–13). Springer. https://doi.org/10.1007/978-981-19-0786-9_1

Qiu, F., Zhang, G., Sheng, X., Jiang, L., Zhu, L., Xiang, Q., Jiang, B., & Chen, P.-K. (2022). Predicting students' performance in e-learning using learning process and behaviour data. *Nature: Scientific Reports, 12*, Article 453. https://doi.org/10.1038/s41598-021-03867-8

Raad, J., & Cruz, C. (2015). A survey on ontology evaluation methods. In A. Fred, J. Dietz, D. Aveiro, K. Liu, & J. Filipe (Eds.), *Proceedings of the 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management* (vol. 2, pp. 179–186). SciTePress. https://doi.org/10.5220/0005591001790186

Rahayu, N. W., Ferdiana, R., & Kusumawardani, S. S. (2022). A systematic review of ontology use in e-learning recommender system. *Computers and Education: Artificial Intelligence, 3*, Article 100047. https://doi.org/10.1016/j.caeai.2022.100047

Rami, S., Bennani, S., & Idrissi, M. K. (2018). A novel ontology-based automatic method to predict learning style using Felder-Silverman model. In *ITHET 2018—17th International Conference on Information Technology Based Higher Education and Training* (pp. 1–5). IEEE. https://doi.org/10.1109/ITHET.2018.8424774

Sultana, J., Rani, M. U., & Farquad, M. A. H. (2019). Student's performance prediction using deep learning and data mining methods. *International Journal of Recent Technology and Engineering (IJRTE), 8*(1S4), 1018–1021.

Wang, Y., & Wang, Y. (2021). A survey of ontologies and their applications in e-learning environments. *Journal of Web Engineering, 20*(6), 1675–1720. https://doi.org/10.13052/jwe1540-9589.2061

Yağcı, M. (2022). Educational data mining: Prediction of students' academic performance using machine learning algorithms. *Smart Learning Environments, 9,* Article 11. https://doi.org/10.1186/s40561-022-00192-z

Zainuddin, N. M., Mohamad, M., Zakaria, N. Y., Sulaiman, N. A., Jalaludin, N. A., & Omar, H. (2024). Trends and benefits of online distance learning in the English as a second language context: A systematic literature review. *Arab World English Journal, 10,* 284–301. https://doi.org/10.24093/awej/call10.18

Zeebaree, S. R. M., Al-Zebari, A., Jacksi, K., & Selamat, A. (2019). Designing an ontology of e-learning system for Duhok Polytechnic University using Protégé OWL tool. *Journal of Advanced Research in Dynamical and Control Systems, 11*(5), 24–37. https://jardcs.org/abstract.php?id=984