Review

Not peer-reviewed version

Deep Learning and Reinforcement Learning for Assessing and Enhancing Academic Performance in University Students. A Scoping Review

Fabrizio Stasolla * , Antonio Zullo , Roberto Maniglio , Anna Passaro , Mariacarla Di Gioia , Enza Curcio , Elvira Martini

Posted Date: 24 December 2024

doi: 10.20944/preprints202412.2037.v1

Keywords: University Students; Artificial Intelligence; Deep Learning; Reinforcement Learning; Academic Performance

Review

Deep Learning and Reinforcement Learning for Assessing and Enhancing Academic Performance in University Students. A Scoping Review

Fabrizio Stasolla [1,*], Antonio Zullo [2], Roberto Maniglio [2], Anna Passaro [1], Mariacarla Di Gioia [2], Enza Curcio [1] and Elvira Martini [1]

University Giustino Fortunato of Benevento (Italy)
[2] Universitas Mercatorum of Rome (Italy)
* Correspondence: f.stasolla@unifortunato.eu

Abstract: Background: University students may encounter challenges throughout their academic careers, with an increased risk of academic drop-out.Among new technologies, those based on Artificial Intelligence (AI), such as Deep Learning (DL) and Reinforcement Learning (RL) strategies, have emerged as valuable resources to assess and enhance academic performance. **Objectives**: To evaluate the effects DL and RL AI-based strategies for assessing and enhancing academic performance in university students. **Method**: A scoping review on the use of DL and RL AI-based strategies for assessing and enhancing academic performance was carried out. A literature overview on the newest empirical studies was conducted. **Results**: Recent studies were reviewed. Data were encouraging and promising. Strengths and weaknesses of DL and RL AI-based strategies in were critically discussed. **Conclusions**: DL and RL AI-based strategies are effective for assessing and improving academic performance. The limitations of the analyzed studies were examined, and some useful insights for future research were highlighted.

Keywords: university students; artificial intelligence; deep learning; reinforcement learning; academic performance

1. Introduction

University students face intricate and multifaceted challenges throughout their academic path, with learning difficulties constituting one of the most significant obstacles to success. A primary issue lies in the increasing cognitive demands of higher education, which necessitate advanced critical thinking, analytical reasoning, and problem-solving skills. Unlike the structured learning environment of secondary education, university settings demand a high degree of independence and self-regulated learning (SRL), requiring students to autonomously manage their academic progression. However, a significant portion of students exhibit limited proficiency in SRL, which hinders their ability to meet academic expectations (Zimmerman & Moylan, 2009). The ineffective application of SRL strategies impacts essential features such as time management, task prioritization, and adaptability in learning techniques, all of which are crucial for mastering the complexities of higher education (Chitra et al., 2022).

This gap leaves many students struggling to effectively manage their studies, particularly in disciplines that require independent learning and self-assessment. Commonly reported difficulties include organizing information, understanding complex academic material, and translating theoretical concepts into practical applications (Mikroyannidis et al., 2014). These issues are exacerbated by a lack of metacognitive awareness, which limits students' capacity to identify effective study strategies and make necessary adjustments to their learning approaches (Renner et al., 2020).

A critical consequence of insufficient SRL skills is the inability to efficiently manage time and tasks. Time management is a cornerstone of academic success, yet many students exhibit repetitive

 doi:10.20944/preprints202412.2037.v1

behaviors that exacerbate stress and diminish productivity. Procrastination is often associated with low self-discipline and volition, both of which are integral to SRL. Students who fail to establish structured schedules or allocate sufficient time for study tend to adopt surface learning strategies, relying on rote memorization rather than engaging in deeper cognitive processing. Such approaches not only compromise academic performance but also hinder the acquisition of transferable skills essential for professional growth(Valenzuela et al., 2020).

Cognitive difficulties are particularly pronounced in fields that require the integration of theoretical and practical knowledge. For instance, students of the faculties of Human Sciences often struggle to synthesize diverse perspectives and construct coherent arguments. These challenges are exacerbated when students lack the skills to decompose complex tasks into manageable steps, recognize gaps in their understanding, and seek appropriate resources for improvement (Ismayilli Karakoc et al., 2022).

Motivational factors play a pivotal role in academic performance. Students with low intrinsic motivation or poor self-efficacy are less likely to engage deeply with their studies, often perceiving academic tasks as insurmountable rather than as opportunities for growth. This mindset, coupled with external pressures such as competitive academic environments or unrealistic familial expectations, further undermines academic success. Additionally, the absence of well-defined goals frequently results in a lack of direction and persistence(Travers et al., 2015).

Given the multifaceted challenges faced, such as limited SRL, difficulty in time management, and motivational deficits, the integration of emerging technological solutions, including Deep Learning (DL) and Reinforcement Learning (RL), might support the academic performance of university students. These AI-driven systems could personalize learning experiences, foster meta-cognitive development, and enhance task prioritization through adaptive feedback and real-time support (Alnasyan et al., 2024;Nie et al., 2023).

2. A Theoretical Framework of RL and DL Based Solution

DL, a transformative branch of Machine Learning (ML), operates through artificial neural networks inspired by the human brain to process vast amounts of data and learn hierarchical representations. This approach allows DL to excel in tasks such as image recognition, Natural Language Processing (NLP), and predictive modeling, where it identifies and refines meaningful patterns from raw inputs. The versatility of DL applications spans across diverse domains, including healthcare, agriculture, and education. By leveraging its capability to identify trends and extrapolate insights, DL has proven to be an invaluable asset for addressing complex challenges in these fields (LeCun et al., 2015; Khalid et al., 2024)

A pivotal application of DL resides in its potential to personalize learning experiences, particularly within higher education. DL algorithms analyze data streams from varied sources—such as student interactions, performance metrics, and individual preferences—to construct adaptive models tailored to specific learning needs. These models dynamically recommend resources, assessments, or activities, aligning with the learner's pace, strengths, and areas requiring improvement. As a result, these AI-driven systems optimize educational outcomes by customizing the learning trajectory (Dong et al., 2021).

In higher education, adaptive learning systems, increasingly powered by DL, utilize predictive analytics to guide students toward success. By analyzing historical performance data, these systems anticipate areas of difficulty, proactively delivering supplementary materials or alternative instructional strategies. This preemptive approach mitigates potential frustration and disengagement, ensuring a more seamless and effective learning process (Janiesch et al., 2021).

Furthermore, DL-driven platforms often incorporate NLP tools to enhance interaction and engagement. Chatbots, underpinned by sophisticated neural networks, provide real-time assistance by addressing inquiries, guiding students through complex content, and elucidating concepts. Such advancements democratize education, offering ubiquitous support and facilitating accessibility irrespective of geographical or temporal constraints (Mageira et al., 2022).

In university contexts, where theoretical knowledge is commonly integrated with practical applications, DL technologies offer interactive and engaging learning opportunities. Virtual laboratories and simulation-based environments, powered by AI, replicate real-world scenarios, enabling students to conduct experiments in risk-free settings. For instance, medical students can simulate surgical procedures with a complexity akin to real-life operations, fostering confidence and refining decision-making capabilities(Tran et al., 2021).

The integration of DL in academic settings also addresses scalability challenges inherent in traditional education systems. Institutions accommodating large student populations often struggle to provide personalized attention. DL tools bridge this gap by automating routine processes such as grading and scheduling, thus allowing educators to concentrate on mentorship and critical teaching interactions. For example, intelligent grading systems assess extensive volumes of submissions with precision and consistency, alleviating the administrative burden on instructors while maintaining academic standards (Kamilaris & Prenafeta-Boldú, 2018).

RL, another paradigm of ML, optimizes agent behaviors through iterative interactions within defined environments. Guided by feedback mechanisms in the form of rewards or penalties, RL agents develop adaptive strategies to achieve specific objectives: it enables adaptive learning and dynamically adjusts task difficulty based on user performance (Stasolla et al., 2023;2024). This framework, based on Markov Decision Processes (MDPs), systematically model states, actions, and transitions, enabling dynamic decision-making in evolving contexts. Unlike other ML approaches, RL emphasizes sequential learning, rendering it particularly suited to domains requiring continuous adaptation and feedback-based optimization, such as education (Ertmer & Newby, 1993; Meyn, 2022).

The application of RL in education has facilitated notable advancements, especially in the realm of Intelligent Tutoring Systems (ITS). These systems employ RL algorithms to tailor instructional content to individual learner profiles, optimizing engagement and efficacy (Iglesias et al., 2009). Bellotti et al. (2009)demonstrated the relevance of RL in game-based learning environments, where adaptive engines dynamically adjust task difficulty and feedback, thus promoting active and personalized learning experiences. This alignment with constructivist pedagogical models underscores RL's relevance in contemporary educational frameworks (Narvekar et al., 2020).

In online learning platforms, RL advances methodologies such as adaptive experimentation and instructional sequencing. For instance, multi-armed bandit algorithms optimize resource allocation by adapting learning materials to individual needs, while task sequencing maximizes learning efficiency (Doroudi et al., 2019).

Deep Reinforcement Learning (DRL), an advanced integration of RL with neural networks, has broadened educational applications by addressing the complexities of high-dimensional data. DRL facilitates dynamic difficulty adjustment mechanisms within e-learning platforms, tailoring challenges based on real-time learner assessments. Furthermore, incorporating Explainable AI (XAI) within RL ensures transparency and equity, addressing ethical concerns in AI-driven education technologies (Wells & Bednarz, 2021).

Simulation-based learning environments, driven by RL, effectively bridge theoretical constructs and practical applications. In fields such as engineering and medicine, RL-powered simulations offer risk-free procedural training, enhancing both competence and confidence. Similarly, RL algorithms refine Massive Open Online Courses (MOOCs) by personalizing course recommendations and adapting learning pathways, thereby promoting inclusivity and accommodating learner variability (Y. Li et al., 2023).

Current studies on DL and RL in education face several limitations, including narrow datasets, restricted capacity of abstraction, and an over-reliance on controlled environments. Many approaches lack scalability and are often constrained by computational resource demands, limiting their applicability in diverse academic settings. Additionally, the absence of longitudinal research leaves uncertainties about sustained impacts on learning outcomes. This review aims to address these challenges by providing a synthesis of empirical evidence in university contexts. By analyzing studies on DL and RL to assess and improve academic performance of university students, the review might identify trends, gaps, and best practices. This approach can provide concrete insights for educators and policy makers, promoting the development of optimized AI-based solutions in higher education (Dong et al., 2021; Sutton & Barto, 2018).

In line with the above, a scoping review was conducted to map the research carried out in this area, as well as to identify potential applications and current limitations in using RL- and DP-based solutions to support academic performance in university students. The research question of the scoping review aims to explore the main RL- and DP-based solutions developed to evaluate and improve academic performance. Furthermore, it aims to investigate the possible practical implications, the limitations of these solutions and the perspectives for future studies.

3. Method

The standard guidelines adopted in this review were in line with PRISMA statement (Moher et al., 2009), according to the protocol available at https://osf.io/5ausc/?view_only=2093e5f348d94abca8883bcf87d76b56

As for the eligibility criteria, studies were included if the following parameters were observed:

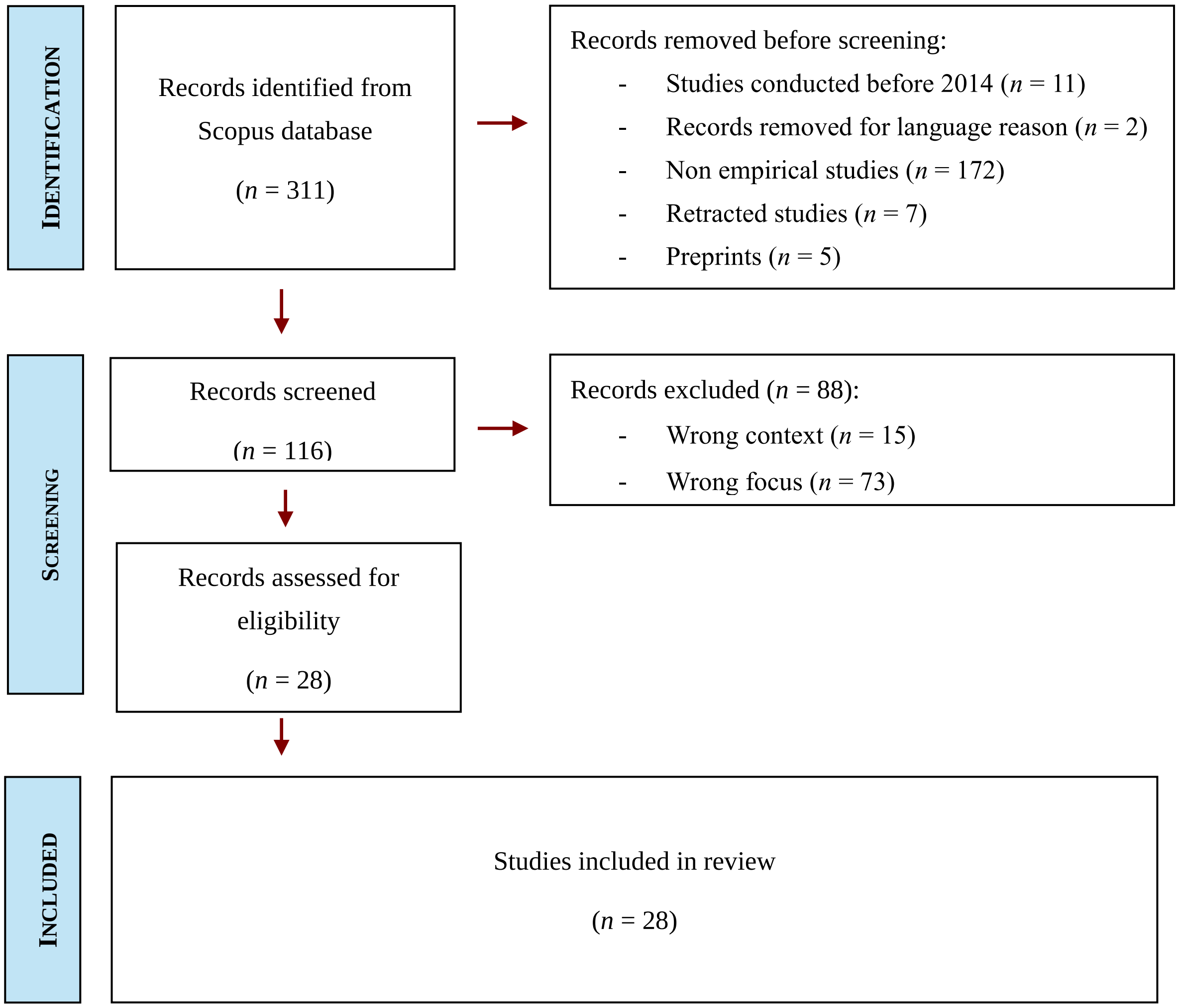
- Keyword: "University Students", "Artificial Intelligence", "Reinforcement Learning" or "Deep Learning";
- Studies published from 2014 to 2024;
- Empirical studies;
- Language: English;
- Pertinence to the research question (DL and RL for assessing and enhancing academic performance in university students);
- Participants: university students.
- The exclusion criteria were:
- reviews and conference papers;
- retracted papers;
- preprints;
- DL and RL for other purpose.
- participants: other students.

An initial search was conducted on Scopus, entering the search keywords: "University Students" and "Artificial Intelligence" and "Deep Learning" or "Reinforcement Learning: the search produced 311 results.

Including all studies conducted between 2014 and 2024, 300 documents were identified. From these, only empirical studies wrote on English were included. The search led to 128documents. Of these documents, 7 were discarded because they were retracted, while another 5 were not considered because they were still in press and not yet published.The search led to 116 documents.From these documents, researches not relating to the use of DL and RL for assessing and enhancing academic performance in university students were excluded. The final outcomes led to 28 results.

The research phases were replicated by two independent judges, with a 13% disagreement. The disagreement was resolved by a third independent judge.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 December 2024 doi:10.20944/preprints202412.2037.v1



 doi:10.20944/preprints202412.2037.v1

Figure 1. PRISMA flowchart of the study selection process.

4. Results

4.1. DL and RL for Assessing Academic Performance

This category of studies explores the application of DL- and RL- based solutions for evaluating academic performance and predicting student outcomes. Eight studies are described, collectively involving data from 49,210students across various university educational settings.

Kadhim & Hassan (2020) aimed to enhance e-learning systems by integrating Recurrent Neural Networks (RNN) optimized with the Adaptive Momentum (ADAM) algorithm DL-based to predict students' learning continuity. The experiment involved 1000 rows collected from students enrolled in virtual classroomsfrom students of the University of Technology, Baghdad, with 80% used for training and 20% for testing. The system analyzed student grades and behaviors to classify continuation or dropout risks. Results indicated that the RNN-ADAM model achieved a 99.1% accuracy, outperforming other methods, including Multi-Layer Perceptron (97.99%), decision tree (64.78%), and random forest (77.96%). This hybrid approach demonstrated its capability in enhancing predictive accuracy for educational performance monitoring.

Liu, Wang, and Yuan (2022) aimed to predict academic performance using a feed forward spiking neural network (SNN). The model processed data from 55 students over six semesters, encompassing 62 courses. The SNN integrated AI and DL by encoding input data into spiking sequences, adapting synaptic weights and time delays, and decoding outputs to classify grades into three levels: high, medium, and low. Results demonstrated that the proposed SNN outperformed traditional methods, achieving the highest accuracy and significantly improving prediction precision for low-grade categories.

Jing et al. (2022) aimed to enhance oral English assessment using a hybrid model combining fuzzy logic with a neural network. Ten participants (5 men, 5 women) with comparable oral literacy were recruited from diverse regions. The model employed advanced speech recognition algorithms to evaluate pronunciation quality, fluency, and accuracy. The intervention involved testing 10 participants. Results demonstrated the model’s superior accuracy (86.16%) and lower error rates compared to traditional methods to standardize oral English evaluation and provide consistent, data-driven feedback for language learning improvement.

Yuhua (2024) developed a Back Propagation Neural Network (BPNN)-based model to assess English Language Teaching (ELT) within a digital cultural framework. The study relied on classified sample data to train and validate the BPNN model, ensuring the inclusion of diverse instructional assessment scenarios. However, specific details about participants were not provided. This DL model analyzed instructional levels, addressing complex, nonlinear relationships between teaching factors and outcomes. By leveraging adaptive learning rate adjustments and a multi-layer neural network, the model provided comprehensive insights into teaching quality. Results demonstrated a 22.64% improvement in assessment accuracy compared to traditional systems, emphasizing the model’s potential to enhance ELT by providing reliable, data-driven evaluations that support instructional refinement and better decision-making in educational environments.

Li and Mohamad (2023) developed the Latent Dirichlet Integrated DL (LDiDL) framework for assessing English oral proficiency. This model combined Latent Dirichlet Allocation (LDA) for uncovering latent topics and a DL model to analyze acoustic and linguistic features. Using a dataset of 500 spoken English samples across varying proficiency levels, the LDiDL system categorized proficiency and provided personalized feedback on grammar, fluency, pronunciation, and vocabulary. Results showed an accuracy of 99% in proficiency assessment, outperforming traditional NLP methods.

Tsai et al. (2020) aimed to predict university student dropouts between their second and fourth years based on data from their first year. The researchers employed a multilayer perception model, a DL-based architecture, which analyzed eight variables, including academic performance, student loan applications, absences, and alerted subjects. Using a dataset of 3,552 students, the model achieved 77% accuracy and high specificity, showing efficiency in identifying dropout at-risk students.

Sayed (2024) developed a dropout prediction model using a DL-based Convolutional Neural Network (CNN). The model leveraged data from the Arab Open University's Learning Management System (LMS) and Student Information System (SIS) encompassing 12,000 student records. Behavioral data were transformed into matrices for CNN analysis, which integrated convolutional and pooling layers with dropout mechanisms to enhance feature extraction and reduce overfitting. The model predicted dropout probabilities and identified critical risk factors influencing student retention. Experimental results demonstrated the CNN's superior performance, achieving an accuracy of 98.6%, surpassing traditional methods and offering robust early intervention capabilities.

Ujkani et al. (2024) utilized Explainable AI (XAI), incorporating SHapley Additive exPlanations (SHAP), to investigate student success in online education. The objectives were to predict academic outcomes, identify at-risk students, and uncover critical success factors. Data from 32,593 students in the Open University Learning Analytics Dataset (OULAD) provided insights into engagement, registration timelines, and course interactions. Using the Open University Learning Analytics Dataset (OULAD), models like Random Forest, Gradient Boosting, and custom neural networks were employed. DL models were refined through hyper-parameter tuning, and pre-processing emphasized dataset integrity. SHAP provided transparency into predictions, revealing student engagement and registration timelines as pivotal for success. Custom neural networks excelled with 94% accuracy in predictions.

4.2. DL and RL for Enhancing Academic Performance

This category of studies investigates the application of AI and DL technologies, including neural networks, RL, and NLP, to optimize teaching strategies, personalize learning experiences, and improve educational outcomes. A total of 20 studies are described, collectively involving 4,545students across various university educational settings.

Xu and Yu (2024) designed an AI-driven online learning platform utilizing a DL model. Tool integrated an advanced resource scheduling technology to enhance teaching efficacy and learning engagement. A total of 1,212 students participated over six months, engaging in technical, language, entertainment, and sports activities. The intervention employed a blockchain-supported decision tree algorithm and fuzzy convolutional neural networks, leveraging the DL model to optimize resource allocation and personalize learning paths. Results showed improved average scores across all subject categories, enhanced student interaction levels, and a 30% increase in resource utilization efficiency compared to traditional methods.

Wang, Zou, and Xue (2023) investigated the use of EAP TALK, an AI-powered system integrating DL, big data, and speech recognition, to enhance oral English proficiency. The model processed student speech, evaluating pronunciation, fluency, and comprehension through real-time scoring. The research included 110 non-English major university students, utilizing questionnaires and semi-structured interviews. Results revealed that EAP TALK improved students' oral fluency and pronunciation, with 65.5% of participants expressing satisfaction with its speech recognition capabilities.

Naseer et al. (2024) aimed to evaluate the integration of AI-driven adaptive learning platforms in higher education to create personalized learning pathways. It employed advanced DL techniques, specifically Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), to analyze student data and dynamically adapt learning content. The study involved 300 students, divided into control and experimental groups. The experimental group used the AI platform, while the control group followed traditional instruction. Results demonstrated a 25% improvement in grades, educational outcomes, engagement and test scores in the experimental group, with a significant p-value of 0.00045.

Jia and Zhang (2021) investigated the integration of AI into the teaching modes of psychology and pedagogy in universities. The researchers used a weighted evaluation algorithm to analyze cognitive abilities, combining AI and DL to optimize teaching methods. Data were collected from 290

 doi:10.20944/preprints202412.2037.v1

participants through surveys and interviews. Results showed that over 90% of participants believed AI improved teaching quality and learning outcomes.

Francisco and Silva (2022) developed an Intelligent Tutoring System (ITS) to enhance the teaching of Software Maintenance (SM) through personalized content recommendations. The model employed Q-Learning, a RL algorithm, to adaptively recommend educational activities based on students' learning status. The research involved extensive analysis of curricula and simulation-based experiments with 10 virtual students and 110 learning activities. Results demonstrated the model's efficiency, with the Q-Learning-based system achieving optimal recommendations after fewer iterations compared to traditional methods, highlighting its potential to improve SM education by reducing learning time and increasing accuracy in activity selection.

Yin, Peng, and Liu (2024) developed and validated the E-GPPE-C smart teaching model to enhance personalized learning and classroom engagement. The model integrated AI and DL technologies, including Convolutional Neural Networks (CNNs) and subject knowledge mapping, to analyze learning data, construct learner profiles, and recommend personalized learning paths. The research involved 103 college students in smart teaching environments and utilized structural equation modeling for validation. Results demonstrated positive effects on learning strategies (β = 0.286), classroom engagement (β = 0.211), and participation (β = 0.20), supporting the model's effectiveness in fostering active and personalized educational experiences.

Liu, Chen, and Yao (2022) investigated the integration of AI technology into teaching and learning processes in universities to promote learning among students. The researchers utilized a DL-based assessment framework employing the YOLOv3 convolutional neural network to analyze classroom behaviors and optimize learning pathways. The study involved three rounds of action research with first-year university students (N= 40), comparing an experimental class using an intelligent teaching model against a control group. Results showed significant improvements in learning, knowledge mastery, ability development, and emotional engagement in the experimental group, highlighting the potential of AI to enhance educational efficiency and foster personalized, in-depth learning experiences.

Liu (2024)aimed to enhance Civics and Political Science education using an AI-driven intelligent teaching system. The model incorporated RL with a Q-learning algorithm to optimize teaching interventions and collaborative filtering to recommend personalized course resources. The research involved 50 first-year students majoring in ideology and politics, with a semester-long application of the system. Results showed an increase in overall assessment scores from 82.51 to 88.76, significant improvements in online learning and practical achievements, and a high satisfaction rate (4.625–4.647 on a 5-point scale). The system demonstrated effectiveness in fostering engagement and improving learning outcomes.

Li, Wang, and Wang (2024) designed and implemented a personalized teaching system tailored for local comprehensive universities using AI technology. The system incorporated face recognition, NLP, and virtualization technologies to optimize teaching and analyze classroom behavior. Testing involved two classes with 44 students each; one used the AI-based system, while the other followed traditional methods. Results demonstrated a significant increase in student participation (45.78%) and attention levels (0.6–0.9) in the AI-supported class. This approach highlights the effectiveness of AI-driven personalized education in enhancing interactivity and teaching outcomes in higher education.

Ou (2024) aimed to enhance English learning outcomes through a blended teaching model integrating the BOPPPS framework with intelligent algorithms, including Bayesian Knowledge Tracing and RL. Conducted with 105 university students, the research compared an experimental group utilizing the model to a control group with traditional teaching. Results showed significant improvements in English skills, with the experimental group outperforming the control group across metrics like listening, reading, and speaking. The BOPPPS-based approach fostered better cognitive, skill-based, and affective learning outcomes, demonstrating its effectiveness in advancing personalized and participatory education in higher learning environments.

 doi:10.20944/preprints202412.2037.v1

Qiao and Fu (2022) implemented Outcome-Based Education (OBE) approach for university mathematics courses through the integration of AI technologies, including the Ant Colony Optimization (ACO) algorithm and Wireless Sensor Networks (WSN). The study aimed to enhance learning outcomes by optimizing teaching methods and personalizing learning experiences. The study does not explicitly mention the number of participants but analyzes data from online self-paced mathematics microcourses conducted between 2011 and 2020, encompassing multiple student interactions over a decade. The AI-based model employed Apriori algorithms for frequent pattern analysis and ACO for decision-making and resource allocation. The experimental phase included various microcourses offered to students via online platforms. The results demonstrated a significant improvement in learning outcomes, with accuracy rates reaching up to 98.87% for specific optimization tasks.

Liu and Ren (2022) investigated the application of AI in an English education platform under the “Internet+” framework to improve teaching quality and learning outcomes. A total of 1,046 students participated, utilizing data from the UCI repository.The model employed a genetic algorithm integrated with wireless sensor networks to personalize learning and automate test paper generation. Using data from the UCI repository, the system optimized course management and student interaction. Experimental results involving 1,046 students demonstrated significant improvements in learning efficiency, student engagement, and teaching effectiveness.

Li and Wu (2023) aimed to enhance university teaching methods by developing an embedded voice teaching system based on cloud computing and DL. The system integrated a hybrid HMM-LSTM model for voice recognition, leveraging HMM's time-processing capabilities and LSTM's characterization and generalization features. It optimized teaching interactions by providing real-time feedback and personalized learning experiences. Tested with 100 university students, the system achieved a high recognition accuracy of 96.25%, robust noise immunity, and stability across functional modules. Results demonstrated improved student engagement and learning efficiency, underscoring the system's potential to support digital transformation in higher education.

Riedel et al. (2023)aimed to evaluateChatGPT's performance in medicaleducation, specifically in obstetrics and gynecology (OB/GYN) examinations. The research utilized two datasets: OB/GYN course exams at a German university and state medical licensing exams. The study involved a total of 786 participants across five OB/GYN course exams conducted between 2021 and 2023. The participant distribution included 154 in February 2023, 125 in July 2022, 185 in February 2022, 149 in July 2021, and 173 in February 2021. These exams were mandatory for the OB/GYN course and aimed to test theoretical and clinical knowledge. ChatGPT, a DL-based large language model, processed natural language questions to generate responses. It demonstrated comparable performance to medical students, answering 85.6% and 70.4% of questions correctly in the respective datasets. Additionally, qualitative assessments highlighted ChatGPT's potential to enhance autonomous medical learning, providing accurate, concise, and relevant insights.

He and Wang (2020) explored the application of a blended teaching model enhanced by AI technologies in public administration courses. The study does not specify the exact number of participants. It focuses on applying a blended teaching model in public administration courses, leveraging AI and big data to enhance teaching and learning processes.The model integrated big data analysis and AI-supported learning platforms to personalize educational content, adapt teaching strategies, and assess learning outcomes. The experimental design involved both online and offline teaching components, utilizing AI to analyze student knowledge, track progress, and suggest tailored learning paths. Results indicated improved teaching quality, greater student engagement, and enhanced learning efficiency.

Wang Y. et al. (2024) aimed to evaluate the impact of AI-powered tools on vocabulary acquisition in English as a Foreign Language (EFL) instruction. Utilizing AI systems such as the UNIPUS AIGC platform and iTEST mobile applications, the authors analyzed data from 110 university students across four majors. The Apriori algorithm from data mining was used to identify associations between learning practices and outcomes. Results highlighted the significant role of AI in enhancing vocabulary learning efficiency, particularly through personalized feedback and interactive learning environments. The findings underscore the potential of integrating AI technologies to optimize language learning strategies and outcomes in EFL education.

Wang and Zheng (2022) aimed to enhance English communication skills among students using a cognitive psychology-based framework supported by DL models. Sixty students were divided into experimental and control groups. Over two months, the experimental group underwent targeted training in reading, vocabulary memorization, and situational dialogue. Results showed significant improvements in the experimental group, with an average score increase of 17.75%, highlighting the model's efficacy in fostering English proficiency and communication skills.

Koć-Januchta et al. (2022) investigated the impact of an AI-enriched biology textbook on university students' learning experiences and outcomes. The objectives were to explore relationships between cognitive load types (intrinsic, germane, and extraneous), usability, self-regulation, and learning gains. The AI-textbook incorporated a 5000-concept knowledge base, NLP for answering questions, and features supporting schema construction.The research involved 42 participants using the textbook during an introductory biology course. Findings revealed that germane cognitive load dominated, indicating DL. Students achieved significant learning gains and valued interactive elements like pop-up definitions. However, usability challenges, such as non-intuitive question formats, increased extraneous load. The results underscore the need for refined design to optimize cognitive engagement.

Chen, Yu, and Wu (2024) developed a DL-based English Vocabulary Teaching Assistance System for college students to improve vocabulary acquisition efficiency. A total of 81 students from two fifth-grade classes in School Y, City W, participated. Class A (40 students) utilized the deep learning system, while Class B (41 students) followed traditional teaching methods. The system employed neural networks and personalized recommendation algorithms to analyze students' learning behaviors and provide tailored vocabulary learning suggestions. Experimental data were gathered from two fifth-grade classes, one utilizing the system and another relying on traditional methods. Results demonstrated a significant improvement in vocabulary test scores for the experimental group, with averages increasing from 75 to 93.

Shao et al. (2022) developed and evaluated an AI-based Arabic Language and Speech Tutor (AI-ALST) focused on teaching Moroccan Arabic pronunciation. Twelve participants engaged with the tutor system, which analyzed 3,851 audio recordings of 17 commonly used words. The system utilized a DL model based on Mel Frequency Cepstrum Coefficients (MFCC), bidirectional Long Short-Term Memory (BiLSTM), and an attention mechanism. Experimental research involved analyzing pronunciation errors using audio recordings from Moroccan Arabic classes. The AI-ALST demonstrated robust performance, achieving high accuracy and F1-scores in pronunciation error detection. It offered real-time feedback and adaptive learning paths, enabling students to improve their speaking proficiency effectively. The results validate the AI-ALST as an innovative tool for second-language learning.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 December 2024 doi:10.20944/preprints202412.2037.v1

Figure 2. Synoptic table of reviewed studies.

Authors	Objective	Method	Participants	Results
Kadhim & Hassan (2020)	Enhance e-learning systems by predicting students' learning continuity.	Utilized a Recurrent Neural Network (RNN) optimized with the Adaptive Momentum (ADAM) algorithm to process grades and behavioral data, achieving superior accuracy in predictions.	1000 rows collected from students enrolled in virtual classrooms	RNN-ADAM achieved 99.1% predictive accuracy.
Liu, Wang, and Yuan (2022)	Predictacademic performance.	Designed a feedforward spiking neural network (SNN) that encoded input data into spiking sequences, adapted synaptic weights, and decoded outputs to classify grades across three categories (high, medium, low).	55 students over sixsemesters	Achieved highest accuracy; improved low-grade prediction.
Jing et al. (2022)	Enhance oral English assessment.	Combined fuzzy logic and neural networks with advanced speech recognition algorithms to analyze pronunciation quality, fluency, and emotional expression in oral English assessments.	10 participants	Model achieved 86.16% accuracy, standardizing oral evaluation.
Yuhua (2024)	Assess English Language Teaching.	Designed a Back Propagation Neural Network (BPNN) to analyze nonlinear relationships between instructional factors and outcomes, using multi-layer neural networks with adaptive learning rates.	Classified sample data. Number of participants not specified	The BPNN-based model improved assessment accuracy by 22.64% compared to traditional instructional systems.
Li and Mohamad (2023)	Assess English oral proficiency using.	Combined Latent Dirichlet Allocation (LDA) for topic modeling with a Deep Learning (DL) framework to analyze linguistic and acoustic features for proficiency categorization.	500 spoken English samples	Achieved 99% accuracy in proficiency assessment.
Tsai et al. (2020)	Predict university student dropouts using a multilayer perceptron model.	A logistic regression model (statistical learning) and a deep learning model using a multilayer perceptron algorithm trained with the TensorFlow framework to predict dropout probabilities.	3552 university students in Taiwan (2093 females, 1459 males) with data from their first academic year.	The DL model achieved a 77% accuracy rate (and higher specificity), while the logistic regression achieved 68% accuracy.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 December 2024 doi:10.20944/preprints202412.2037.v1

Sayed (2024)	Develop dropout prediction model.	Convolutional Neural Network (CNN) with a feature-weighting method, Nadam optimizer, and pooling layers for dropout prediction using AOU-LMS and AOU-SIS datasets.	12,000 students from the Arab Open University, diverse in age (below 20 to 29+), with data on demographics, GPA, and blended learning engagements.	Achieved 98.6% prediction accuracy for dropout rates.
Ujkani et al. (2024)	Predict course success and early identification of at-risk students	ML models (Random Forest, Gradient Boosting, k-NN, Neural Networks) and SHAP for explanation.	32,593 students (Open University Learning Analytics Dataset), diverse demographics and academic backgrounds.	94% prediction accuracy; engagement identified as a critical factor.
Xu and Yu (2024)	Enhance online learning platform.	Integrated a DL model with blockchain-supported decision tree algorithms and fuzzy convolutional neural networks (CNNs) to optimize resource scheduling and adapt learning paths in real-time.	1,212 students over sixmonths	Improved scores, interaction levels, and resource efficiency by 30%.
Wang, Zou, and Xue (2023)	Enhanceoral English proficiency.	Created the EAP TALK system, integrating AI, DL, and big data to evaluate pronunciation, fluency, and comprehension using real-time scoring with speech recognition algorithms.	110 university students in China (27 males, 83 females) aged 17-29, mostly freshmen and sophomores, with middle-level English entrance scores.	Improved fluency and pronunciation; 65.5% satisfaction rate.
Naseer et al. (2024)	Evaluate AI-driven adaptive learning platforms for personalized pathways.	Leveraged Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) to analyze student data and dynamically adapt learning materials and assessments based on performance patterns.	300 students (control + experimental)	25% improvement in grades and engagement; significant p-value.
Jia and Zhang (2021)	Integrate AI into psychology and pedagogy teaching modes.	Weighted Evaluation Algorithm for cognitive ability, combining AI with traditional teaching methods.	290 teachers and students (age not specified).	90% reported improvement in teaching quality and learning outcomes.
Francisco and Silva (2022)	Enhance Software Maintenanceteaching.	Employed a Q-Learning algorithm, a type of Reinforcement Learning (RL), to model states and actions in educational tasks, refining content recommendations dynamically based on prior outcomes.	10 virtualstudents	Optimal recommendations with fewer iterations; improved activity selection.

 doi:10.20944/preprints202412.2037.v1

Yin, Peng, and Liu (2024)	Enhancepersonalized learning.	Developed the E-GPPE-C model using CNNs and subject knowledge mapping to construct learner profiles and recommend personalized learning paths based on engagement and performance data.	103 college students in smart teaching classes, tracked over one semester.	The E-GPPE-C model showed significant improvement in learning engagement (β=0.286), participation (β=0.203), and creativity (β=0.424). AI-driven tools effectively promoted personalized learning and collaboration, enhancing the overall smart learning environment.
Liu, Chen, and Yao (2022)	Promote learning.	Applied the YOLOv3 Convolutional Neural Network (CNN) to analyze classroom behavior data, providing real-time feedback to inform teaching strategies and address learning gaps.	40 first-yearuniversitystudents	Significant improvements in learning and emotional engagement.
Liu (2024)	Optimize Civics teaching using adaptive systems.	Utilized RL (Q-Learning) to tailor teaching interventions and resource recommendations, dynamically adjusting strategies based on student feedback and performance.	50 first-year university students majoring in Ideology and Politics at University D.	Post-intervention scores improved by an average of 6.25 points; personalized recommendation accuracy exceeded 92.5%, and satisfaction ratings for system functionality averaged 4.63/5. The system enhanced learning engagement, outcomes, and adaptability across student groups.
Li, Wang, and Wang (2024)	Implement AI-driven personalized teaching system for local universities.	Face recognition, NLP, and virtualization technologies.	88 students from two classes, ages not specified, divided into experimental and control groups.	Increased student participation (45.78%) and attention levels (0.6–0.9). Enhanced interaction and teaching outcomes in the AI-supported class.
Ou (2024)	Enhance English learning outcomes.	Integrated Bayesian Knowledge Tracing and RL to track learning progress and adapt teaching strategies dynamically, optimizing comprehension and addressing gaps.	105 students from School Z (52 in experimental group, 53 in control group).	Significant improvements in English proficiency (listening, reading, writing, translation, speaking) and interest in English for the experimental group. Enhancedeffectiveness of AI-supported blended teaching.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 December 2024 doi:10.20944/preprints202412.2037.v1

Qiao and Fu (2022)	Enhanceuniversitymathematics learning.	Applied Ant Colony Optimization (ACO) and Apriori algorithms to allocate resources and sequence tasks dynamically based on real-time learner data.	The study does not explicitly mention the number of participants but analyzes data from online self-paced mathematics microcourses conducted between 2011 and 2020, encompassing multiple student interactions over a decade	Proposed algorithms achieved 98.87% accuracy, enhancing microcourse customization and efficiency in meeting learning outcomes.
Liu and Ren (2022)	Improve English learning efficiency using AI and genetic algorithm.	Used a genetic algorithm-based framework analyzing English course performance using UCI repository data	1046 students; highereducation	Significant improvement in learning and engagement.
Li and Wu (2023)	Develop embedded voice teaching system.	Hybrid HMM-LSTM model for voice recognition integrated with a cloud computing platform	100 universitystudents	High voice recognition rate (96.25%) with robust noise immunity; improved learning engagement and satisfaction; efficient feedback on exercises; enriched course content fostering independent study skills.
Riedel et al. (2023)	Evaluate ChatGPT's performance in medical education exams.	Leveraged ChatGPT, a large language model based on DL, to process natural language questions and provide responses for medical education assessments.	The study involved a total of 786 participants across five OB/GYN course exams conducted between 2021 and 2023. The participant distribution included 154 in February 2023, 125 in July 2022, 185 in February 2022, 149 in July 2021, and 173 in February 2021.	ChatGPT achieved 85.6% accuracy on university questions and 70.4% on state exam questions. It performed comparably to medical students, especially in general knowledge.
He and Wang (2020)	Enhance public administration teaching using AI and blended learning.	Blended learning model integrating AI technologies for personalized learning pathways, knowledge tracking, and assessment.	Students enrolled in Public Administration courses at a Chinese university (number unspecified).	The AI-enhanced blended model improved learning outcomes, student engagement, and teaching effectiveness. Personalized pathways and automated assessments were pivotal in addressing diverse learner needs.
Wang Y. et al. (2024)	Evaluate AI-powered tools for vocabulary acquisition in EFL.	Apriori algorithm for analyzing survey data on AI-driven language platforms	110 second-year university students from diverse majors (ages 19–21).	Enhanced vocabulary learning by identifying effective strategies;

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 December 2024 doi:10.20944/preprints202412.2037.v1

				personalized learning experiences led to better language acquisition.
Wang and Zheng (2022)	Enhance English communication.	Deep Neural Network (DNN) incorporating cognitive psychology principles for English grammar detection and communication training.	60 university students from North China University of Water Resources and Electric Power, divided into experimental and control groups.	Experimental group improved significantly: reading comprehension (+13.33%), question answering (+15.19%), situational dialogue (+17.39%), topic description (+28.3%). Overall, class A's average score rose by 17.75% compared to a minimal 3.25% improvement in the control group.
Koć-Januchta et al. (2022)	Investigate AI-enriched biology textbooks' impact on learning.	AI-enriched textbook leveraging NLP and a knowledge base.	42 university students (69% female, ages 17-44, M = 26.28).	Germane cognitive load significantly higher than intrinsic and extraneous loads, indicating meaningful engagement and learning.
Chen, Yu, and Wu (2024)	Improve English vocabulary acquisition using DL-based system.	The study used DL neural network models for student behavior detection, facial orientation recognition, and personalized recommendation.	81 university students in School Y, divided into two classes (A and B) with similar initial abilities. Ages and detailedcharacteristicsnotspecified.	Class A (using the system) showed significant improvement in vocabulary test scores (average increase from 75 to 93) compared to Class B (traditional teaching). Improved engagement and efficiency.
Shao et al. (2022)	Develop AI-based Arabic tutor for pronunciation teaching.	The AI-ALST system used Mel-Frequency Cepstrum Coefficient (MFCC) for feature extraction and an attention BiLSTM model to process audio data and detect pronunciation errors. A cost-based learning strategy addressed class imbalance.	12 participants learning Moroccan Arabic at the University of Arizona	The system achieved high accuracy in detecting mispronunciations. Attention BiLSTM outperformed BiLSTM for precision, recall, and F1-score across most words. It successfully provided tailored feedback, enabling learners to improve pronunciation effectively.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 December 2024 doi:10.20944/preprints202412.2037.v1

5. Discussion

The reviewed studies demonstrate that DL and RL provide significant opportunities for advancing academic evaluation and student support. These technologies could be useful in identifying students at risk of academic failure and distinguishing them from those likely to experience regular and successful academic progress.

Through advanced algorithms, DL models analyze vast datasets encompassing grades, attendance, participation, behavioral patterns, and even speech characteristics. For instance, Kadhim & Hassan (2020) showed that Recurrent Neural Networks (RNNs) optimized with the Adaptive Momentum (ADAM) algorithm achieved a 99.1% accuracy in predicting learning continuity. This surpasses traditional methods, emphasizing the robustness of DL in managing large datasets and extracting meaningful insights to identify at-risk students. Similarly, Tsai et al. (2020) used a multilayer perceptron model to analyze academic and personal data, achieving 77% accuracy in predicting university dropouts. These findings underline the predictive power of DL, offering educators valuable tools to identify and support struggling students early in their academic journey.

The application of DL extends beyond risk identification to include subject-specific improvements. The model developed by Jing et al. (2022) achieved 86.16% accuracy in pronunciation assessments, outperforming traditional evaluation methods. By incorporating advanced speech recognition algorithms, the system provided consistent, data-driven feedback, enhancing language learning for diverse participants. Similarly, Li and Mohamad (2023) utilized Latent Dirichlet Integrated DL to assess English oral proficiency, achieving 99% accuracy while providing personalized feedback on fluency, grammar, and vocabulary. These studies highlighted how DL technologies contribute to subject-specific learning by offering precise evaluations and tailored feedback, ensuring that students improve in targeted areas while reducing errors in assessment.

RL adds another dimension to educational innovation by tailoring interventions to individual students' needs. Models based on RL algorithms, such as Q-learning, adaptively recommend educational activities, enabling struggling students to recover their performance and maintain motivation. Francisco and Silva (2022) demonstrated the effectiveness of RL in enhancing software maintenance education, achieving optimal recommendations for educational activities with reduced learning time. Similarly, Liu (2024) employed Q-learning to optimize teaching strategies in Civics and Political Science education, leading to significant improvements in assessment scores and engagement levels. By dynamically adjusting to the unique learning trajectories of students, RL ensures that interventions are personalized, targeted, and effective in addressing specific gaps.

One of the main advantages of combining DL and RL is their ability to work synergistically to improve academic outcomes. While DL models are highly effective in analyzing vast datasets and identifying patterns of risk or success, RL offers a pathway to implement personalized solutions based on those insights. For example, Naseer et al. (2024) combined DL techniques, including CNNs and RNNs, with adaptive learning platforms to create personalized pathways for students. Their study showed a 25% improvement in grades and engagement for students in the experimental group compared to traditional teaching methods. This integration of DL and RL allows educators to predict risks and actively intervene to mitigate them, creating a comprehensive framework for academic support.

Practical applications of these technologies are far-reaching. AI-driven systems powered by DL and RL can significantly enhance teaching efficacy and learning experiences. Xu and Yu (2024) demonstrated that an AI-driven online learning platform incorporating DL and blockchain-supported decision trees increased resource utilization efficiency by 30%, while also improving engagement and learning outcomes. These results showed the capacity of AI technologies to optimize resource allocation, streamline instructional processes, and foster an engaging learning environment. Similarly, the study by Liu and Ren (2022) highlighted how a genetic algorithm integrated with wireless sensor networks optimized English course management and student interaction, leading to improved learning efficiency and teaching effectiveness.

Beyond academic performance, these technologies address critical issues such as student motivation and self-regulated learning. For instance, Yin, Peng, and Liu (2024) validated the E-GPPE-C smart teaching model, which integrates DL technologies and subject knowledge mapping to construct personalized learning paths. This approach positively impacted classroom engagement, participation, and learning strategies, demonstrating the potential to foster active and self-directed learning. By enabling students to take ownership of their learning process, these models contribute to long-term academic success and personal development.

Another significant implication of these technologies is their ability to support diverse learners and adapt to various educational contexts. The studies reviewed encompass a wide range of applications, from oral language proficiency assessments to complex decision-making in course management. For instance, Shao et al. (2022) developed an AI-based Arabic Language and Speech Tutor (AI-ALST) focused on Moroccan Arabic pronunciation. This system provided real-time feedback and adaptive learning paths, achieving high accuracy in pronunciation error detection. Such applications demonstrate the versatility of DL and RL in addressing the unique needs of students across disciplines, languages, and learning environments.

Moreover, the scalability of DL and RL models ensures that their benefits extend beyond individual classrooms. Large-scale datasets, such as the Open University Learning Analytics Dataset (OULAD) used by Ujkani et al. (2024), enable the development of predictive models that identify critical success factors for thousands of students. Their use of SHapley Additive exPlanations (SHAP) provided transparency into predictions, revealing key indicators such as student engagement and registration timelines. These insights not only aid in individual student support but also inform institutional strategies for enhancing overall academic performance.

However, the implementation of DL and RL in education also presents challenges. Ethical concerns, such as data privacy and algorithmic transparency, must be addressed to ensure that these technologies are deployed responsibly. Furthermore, educators and institutions require training to effectively integrate AI-driven systems into their teaching practices. Despite these challenges, DL and RL based solution can analyze complex datasets, provide personalized interventions, and enhance engagement and learning outcomes, positioning themselves as indispensable tools for the future of higher education.

5.1. Limitations and Implications for Future Research

A common limitation across many studies, such as those by Xu and Yu (2024), Kadhim and Hassan (2020), and He and Wang (2020), was their reliance on data from single institutions or specific academic settings, restricting the generalizability of their findings. To address this, future research could validate these models across diverse institutions, disciplines, and populations, incorporating multi-institutional datasets and more heterogeneous samples. Expanding the contexts in which these models are tested could refine their scalability and ensure their adaptability across broader educational landscapes.

Sample size constraints also posed challenges, as seen in studies like Jing et al. (2022) and Wang and Zheng (2022), which used small participant pools of 10 and 60 students, respectively. Such limited sample sizes reduce the robustness of findings and their applicability to diverse student populations. Future research could engage larger, more varied participant groups to capture linguistic, cultural, and educational diversity. For example, incorporating non-native speakers with diverse accents and dialects into language-focused models, such as those by Jing et al. (2022) and Li and Mohamad (2023), could enhance their linguistic versatility and broaden their utility.

Several studies, including those by Liu (2024), Francisco and Silva (2022), and Qiao and Fu (2022), concentrated on specific academic disciplines, such as civics, software maintenance, and mathematics. This narrow focus limits the applicability of their models to other fields. Expanding these systems to cross-disciplinary applications could increase their relevance and impact in varied educational contexts. For instance, models tailored for specific domains, such as LDiDL for English proficiency or BPNN for language teaching (Li & Mohamad, 2023; Yuhua, 2024), could be adapted to address challenges in STEM or social sciences.

Computational complexity and resource dependency also emerged as barriers to scalability in studies such as those by Naseer et al. (2024), Liu and Ren (2022), and Wang, Zou, and Xue (2023). The reliance on advanced technologies like CNNs, RNNs, and genetic algorithms creates challenges for implementation in resource-constrained institutions. Future research could prioritize the development of lightweight, efficient algorithms and explore cost-effective technological solutions to ensure broader accessibility. For example, integrating these systems with open-source platforms or optimizing their architecture for low-resource settings could enhance their feasibility in diverse academic environments.

A notable limitation in studies like those by Liu, Wang, and Yuan (2022) and Yin, Peng, and Liu (2024) was the exclusion of multimodal and socio-emotional data, which are critical to understanding the complexity of learning processes. Current models often focus narrowly on cognitive metrics, ignoring behavioral, emotional, or contextual factors that significantly influence educational outcomes. Future research could integrate multimodal data streams, including speech patterns, facial expressions, and engagement metrics, to offer a more holistic evaluation of student performance. Such advancements could be particularly beneficial for systems like EAP TALK and the ITS developed by Francisco and Silva (2022), which would gain from incorporating additional layers of interaction and contextual analysis.

Another limitation was the lack of interpretability in several AI models, such as the RNN-ADAM model (Kadhim & Hassan, 2020) and the SNN model (Liu, Wang, & Yuan, 2022). These black-box systems hinder educators' ability to understand and trust their outputs. While some studies, like those by Ujkani et al. (2024), utilized explainable AI techniques, the computational demands of methods like SHAP limited their scalability. Future research could prioritize the development of transparent, user-friendly explainability tools to enhance trust and usability. This could involve creating intuitive visualizations or interfaces that allow educators to explore how predictions are generated, thereby supporting informed decision-making.

Technological integration and user experience issues were also evident in models like EAP TALK (Wang, Zou, & Xue, 2023) and the personalized teaching system by Li, Wang, and Wang (2024). Participants in these studies reported discomfort with prolonged use and challenges in navigating complex interfaces. Future research could focus on improving usability and reducing cognitive load by designing intuitive, user-centered interfaces. Incorporating user feedback during development could ensure these systems align with the needs of both students and educators, enhancing adoption and effectiveness.

The studies also highlighted limitations in linguistic and cultural adaptability, particularly in language-focused models. For instance, Jing et al. (2022) and Shao et al. (2024) noted challenges in handling diverse accents and dialects, while Li and Mohamad (2023) acknowledged the need to extend their model beyond English proficiency to other languages. Future research could address these issues by expanding datasets to include a wider range of linguistic and cultural contexts and integrating multimodal inputs like gestures and facial expressions. Such advancements could enhance the robustness and inclusivity of these models.

Short-term evaluations were another recurring limitation. Studies like those by Liu (2024) and Wang et al. (2024) emphasized immediate outcomes without exploring long-term impacts. Future research could incorporate longitudinal studies to assess the sustained effects of AI-driven systems on learning outcomes and educational practices. For example, evaluating how models influence students' critical thinking, collaboration, or adaptability over time would provide deeper insights into their efficacy.

Several models focused exclusively on academic metrics, neglecting the broader dimensions of learning, such as emotional intelligence, creativity, or collaborative skills. For instance, Jia and Zhang (2021) and Yuhua (2024) limited their assessments to predefined cognitive categories, which may not capture the full spectrum of student development. Future research could expand these frameworks to include holistic metrics, integrating socio-emotional and interpersonal dimensions alongside academic performance.

Finally, the reliance on simulated or pre-structured datasets, as noted in studies like Qiao and Fu (2022) and Riedel et al. (2023), may not fully represent real-world educational variability. Future research could validate these models using live classroom data and real-world scenarios, ensuring their outputs are aligned with the complexities of dynamic educational environments. Moreover, expanding assessment formats to include open-ended or case-based questions, as suggested for ChatGPT (Riedel et al., 2023), could enhance their utility in fostering critical thinking and problem-solving skills.

Addressing these limitations could significantly enhance the effectiveness, scalability, and inclusivity of AI-driven educational tools. By integrating diverse datasets, multimodal inputs, efficient algorithms, and explainable techniques, future research could refine these models to meet the evolving needs of higher education. This could pave the way for transformative advancements in personalized learning, academic performance evaluation, and educational innovation.

Author Contributions: Conceptualization, F.S., A.Z., R.M., A.P., M.D.G., and E.C.; methodology, EM revision and editing. F.S and A.Z., validation, A.P., M.D.G., E.C.; formal analysis, F.S., A.P., M.D.G. E.C.; investigation, F.S. and A.Z; writing - original draft preparation, A.Z.; writing - review and editing, F.S., R.M. and A.Z.; supervision, F.S. and R.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Alnasyan, B., Basheri, M., & Alassafi, M. (2024). The power of Deep Learning techniques for predicting student performance in Virtual Learning Environments: A systematic literature review. In *Computers and Education: Artificial Intelligence* (Vol. 6). https://doi.org/10.1016/j.caeai.2024.100231
2. Bellotti, F., Berta, R., De Gloria, A., & Primavera, L. (2009). Adaptive Experience Engine for Serious Games. *IEEE Transactions on Computational Intelligence and AI in Games*, *1*(4), 264–280. https://doi.org/10.1109/TCIAIG.2009.2035923
3. Chen, L., Yu, F., & Wu, J. (2024). Research on Deep Learning-based English Vocabulary Teaching Aid System in Colleges and Universities. *Applied Mathematics and Nonlinear Sciences*, *9*(1), 20240855. https://doi.org/10.2478/amns-2024-0855
4. Chitra, E., Hidayah, N., Chandratilake, M., & Nadarajah, V. D. (2022). Self-Regulated Learning Practice of Undergraduate Students in Health Professions Programs. *Frontiers in Medicine*, *9*, 803069. https://doi.org/10.3389/fmed.2022.803069
5. Dong, S., Wang, P., & Abbas, K. (2021). A survey on deep learning and its applications. *Computer Science Review*, *40*, 100379. https://doi.org/10.1016/j.cosrev.2021.100379
6. Doroudi, S., Aleven, V., & Brunskill, E. (2019). Where's the Reward?: A Review of Reinforcement Learning for Instructional Sequencing. *International Journal of Artificial Intelligence in Education*, *29*(4), 568–620. https://doi.org/10.1007/s40593-019-00187-x
7. Ertmer, P. A., & Newby, T. J. (1993). Behaviorism, Cognitivism, Constructivism: Comparing Critical Features from an Instructional Design Perspective. *Performance Improvement Quarterly*, *6*(4), 50–72. https://doi.org/10.1111/j.1937-8327.1993.tb00605.x
8. Francisco, R., & Silva, F. (2022). A Recommendation Module based on Reinforcement Learning to an Intelligent Tutoring System for Software Maintenance: *Proceedings of the 14th International Conference on Computer Supported Education*, 322–329. https://doi.org/10.5220/0011083900003182
9. He, H., & Wang, T. (2020). Application of Blended Teaching in Public Administration Courses under the Background of Artificial Intelligence. *Journal of Physics: Conference Series*, *1575*(1), 012174. https://doi.org/10.1088/1742-6596/1575/1/012174

10. Iglesias, A., Martínez, P., Aler, R., & Fernández, F. (2009). Learning teaching strategies in an Adaptive and Intelligent Educational System through Reinforcement Learning. *Applied Intelligence*, *31*(1), 89–106. https://doi.org/10.1007/s10489-008-0115-1
11. Ismayilli Karakoc, A., Ruegg, R., & Gu, P. (2022). First-year writing in the humanities and social sciences: Requirements, expectations, and perceived weaknesses. *Language Teaching Research*, 13621688221130906. https://doi.org/10.1177/13621688221130906
12. Janiesch, C., Zschech, P., & Heinrich, K. (2021). Machine learning and deep learning. *Electronic Markets*, *31*(3), 685–695. https://doi.org/10.1007/s12525-021-00475-2
13. Jia, S., & Zhang, X. (2021). Teaching Mode of Psychology and Pedagogy in Colleges and Universities Based on Artificial Intelligence Technology. *Journal of Physics: Conference Series*, *1852*(3), 032033. https://doi.org/10.1088/1742-6596/1852/3/032033
14. Jing, C., Zhao, X., Ren, H., Chen, X., & Gaowa, N. (2022). An Approach to Oral English Assessment Based on Intelligent Computing Model. *Scientific Programming*, *2022*, 1–8. https://doi.org/10.1155/2022/4663574
15. Kadhim, M. K., & Hassan, A. K. (2020). Towards Intelligent E-Learning Systems: A Hybrid Model for Predicatingthe Learning Continuity in Iraqi Higher Education. *Webology*, *17*(2), 172–188. https://doi.org/10.14704/WEB/V17I2/WEB17023
16. Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*, *147*, 70–90. https://doi.org/10.1016/j.compag.2018.02.016
17. Khalid, U., Naeem, M., Stasolla, F., Syed, M., Abbas, M., & Coronato, A. (2024). Impact of AI-Powered Solutions in Rehabilitation Process: Recent Improvements and Future Trends. *International Journal of General Medicine*, *Volume 17*, 943–969. https://doi.org/10.2147/IJGM.S453903
18. Koć-Januchta, M. M., Schönborn, K. J., Roehrig, C., Chaudhri, V. K., Tibell, L. A. E., & Heller, H. C. (2022). "Connecting concepts helps put main ideas together": Cognitive load and usability in learning biology with an AI-enriched textbook. *International Journal of Educational Technology in Higher Education*, *19*(1), 11. https://doi.org/10.1186/s41239-021-00317-3
19. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539
20. Li, S., Wang, H., & Wang, L. (2024). Research and Practice of Personalized Teaching System of Local Comprehensive University Based on Artificial Intelligence Technology. *Applied Mathematics and Nonlinear Sciences*, *9*(1), 20240253. https://doi.org/10.2478/amns-2024-0253
21. Li, W., & Mohamad, M. (2023). An Efficient Probabilistic Deep Learning Model for the Oral Proficiency Assessment of Student Speech Recognition and Classification. *International Journal on Recent and Innovation Trends in Computing and Communication*, *11*(6), 411–424. https://doi.org/10.17762/ijritcc.v11i6.7734
22. Li, Y., & Wu, F. (2023). Design and Application Research of Embedded Voice Teaching System Based on Cloud Computing. *Wireless Communications and Mobile Computing*, *2023*, 1–10. https://doi.org/10.1155/2023/7873715
23. Li, Y., Xiong, H., Kong, L., Zhang, R., Xu, F., Chen, G., & Li, M. (2023). *MHRR*: MOOCs Recommender Service With Meta Hierarchical Reinforced Ranking. *IEEE Transactions on Services Computing*, *16*(6), 4467–4480. https://doi.org/10.1109/TSC.2023.3325302
24. Liu, C., Wang, H., & Yuan, Z. (2022). A Method for Predicting the Academic Performances of College Students Based on Education System Data. *Mathematics*, *10*(20), 3737. https://doi.org/10.3390/math10203737
25. Liu, Y. (2024). Research on the Design and Effectiveness Analysis of Artificial Intelligence-Driven Intelligent Teaching and Assisting System for Civics and Political Science Courses. In *Applied Mathematics and Nonlinear Sciences* (Vol. 9, Issue 1). https://doi.org/10.2478/amns-2024-2708
26. Liu, Y., Chen, L., & Yao, Z. (2022). The application of artificial intelligence assistant to deep learning in teachers' teaching and students' learning processes. *Frontiers in Psychology*, *13*, 929175. https://doi.org/10.3389/fpsyg.2022.929175
27. Liu, Y., & Ren, L. (2022). The Influence of Artificial Intelligence Technology on Teaching under the Threshold of "Internet+": Based on the Application Example of an English Education Platform. *Wireless Communications and Mobile Computing*, *2022*, 1–9. https://doi.org/10.1155/2022/5728569

28. Mageira, K., Pittou, D., Papasalouros, A., Kotis, K., Zangogianni, P., & Daradoumis, A. (2022). Educational AI Chatbots for Content and Language Integrated Learning. *Applied Sciences*, *12*(7), 3239. https://doi.org/10.3390/app12073239
29. Meyn, S. (2022). *Control Systems and Reinforcement Learning* (1st ed). Cambridge University Press.
30. Mikroyannidis, A., Connolly, T., Law, E. L. C., Schmitz, H. C., Vieritz, H., Nussbaumer, A., Berthold, M., Ullrich, C., & Dhir, A. (2014). Self-regulated learning in formal education: Perceptions, challenges and opportunities. *International Journal of Technology Enhanced Learning*, *6*(2), 145. https://doi.org/10.1504/IJTEL.2014.066860
31. Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. (2009). Research methods and reporting. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *British Medical Journal*, *339*, 332–339. https://doi.org/10.1371/journal.pmed.1000097
32. Narvekar, S., Peng, B., Leonetti, M., Sinapov, J., Taylor, M. E., & Stone, P. (2020). *Curriculum Learning for Reinforcement Learning Domains: A Framework and Survey*. https://doi.org/10.48550/ARXIV.2003.04960
33. Naseer, F., Khan, M. N., Tahir, M., Addas, A., & Aejaz, S. M. H. (2024). Integrating deep learning techniques for personalized learning pathways in higher education. *Heliyon*, *10*(11), e32628. https://doi.org/10.1016/j.heliyon.2024.e32628
34. Nie, A., Reuel, A.-K., & Brunskill, E. (2023). Understanding the Impact of Reinforcement Learning Personalization on Subgroups of Students in Math Tutoring. In *Communications in Computer and Information Science: Vol. 1831 CCIS* (pp. 688–694). https://doi.org/10.1007/978-3-031-36336-8_106
35. Ou, Y. (2024). Strategies for Applying BOPPPS Model Supported by Intelligent Algorithms in Blended Teaching of College English. *Applied Mathematics and Nonlinear Sciences*, *9*. https://doi.org/10.2478/amns-2024-1719
36. Qiao, Y., & Fu, H. (2022). Study on OBE Teaching Concept in the Context of Deep Learning for the Construction of University Mathematics Microcourses. *Computational Intelligence and Neuroscience*, *2022*, 1–10. https://doi.org/10.1155/2022/6860842
37. Renner, B., Wesiak, G., Pammer-Schindler, V., Prilla, M., Müller, L., Morosini, D., Mora, S., Faltin, N., & Cress, U. (2020). Computer-supported reflective learning: How apps can foster reflection at work. *Behaviour & Information Technology*, *39*(2), 167–187. https://doi.org/10.1080/0144929X.2019.1595726
38. Riedel, M., Kaefinger, K., Stuehrenberg, A., Ritter, V., Amann, N., Graf, A., Recker, F., Klein, E., Kiechle, M., Riedel, F., & Meyer, B. (2023). ChatGPT's performance in German OB/GYN exams – paving the way for AI-enhanced medical education and clinical practice. *Frontiers in Medicine*, *10*, 1296615. https://doi.org/10.3389/fmed.2023.1296615
39. Sayed, M. (2024). Student Progression and Dropout Rates Using Convolutional Neural Network: A Case Study of the Arab Open University. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, *28*(3), 668–678. https://doi.org/10.20965/jaciii.2024.p0668
40. Shao, S., Alharir, S., Hariri, S., Satam, P., Shiri, S., & Mbarki, A. (2022). AI-based Arabic Language and Speech Tutor. *2022 IEEE/ACS 19th International Conference on Computer Systems and Applications (AICCSA)*, 1–8. https://doi.org/10.1109/AICCSA56895.2022.10017924
41. Stasolla, F., Akbar, K., Passaro, A., Dragone, M., Di Gioia, M., & Zullo, A. (2024). Integrating reinforcement learning and serious games to support people with rare genetic diseases and neurodevelopmental disorders: Outcomes on parents and caregivers. *Frontiers in Psychology*, *15*, 1372769. https://doi.org/10.3389/fpsyg.2024.1372769
42. Stasolla, F., Passaro, A., Di Gioia, M., Curcio, E., & Zullo, A. (2023). Combined extended reality and reinforcement learning to promote healthcare and reduce social anxiety in fragile X syndrome: A new assessment tool and a rehabilitative strategy. *Frontiers in Psychology*, *14*, 1273117. https://doi.org/10.3389/fpsyg.2023.1273117
43. Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (Second edition). The MIT Press.
44. Tran, K. A., Kondrashova, O., Bradley, A., Williams, E. D., Pearson, J. V., & Waddell, N. (2021). Deep learning in cancer diagnosis, prognosis and treatment selection. *Genome Medicine*, *13*(1), 152. https://doi.org/10.1186/s13073-021-00968-x

45. Travers, C. J., Morisano, D., & Locke, E. A. (2015). Self-reflection, growth goals, and academic outcomes: A qualitative study. *British Journal of Educational Psychology*, *85*(2), 224–241. https://doi.org/10.1111/bjep.12059
46. Tsai, S.-C., Chen, C.-H., Shiao, Y.-T., Ciou, J.-S., & Wu, T.-N. (2020). Precision education with statistical learning and deep learning: A case study in Taiwan. *International Journal of Educational Technology in Higher Education*, *17*(1), 12. https://doi.org/10.1186/s41239-020-00186-2
47. Ujkani, B., Minkovska, D., & Hinov, N. (2024). Course Success Prediction and Early Identification of At-Risk Students Using Explainable Artificial Intelligence. *Electronics*, *13*(21). https://doi.org/10.3390/electronics13214157
48. Valenzuela, R., Codina, N., Castillo, I., & Pestana, J. V. (2020). Young University Students' Academic Self-Regulation Profiles and Their Associated Procrastination: Autonomous Functioning Requires Self-Regulated Operations. *Frontiers in Psychology*, *11*, 354. https://doi.org/10.3389/fpsyg.2020.00354
49. Wang, W., Zou, B., & Xue, S. (2023). AI technology used as a tool for enhancing university students' English speaking skills: Perceptions and practices. In S. Patnaik & T. Shen (Eds.), *Seventh International Conference on Mechatronics and Intelligent Robotics (ICMIR 2023)* (p. 97). SPIE. https://doi.org/10.1117/12.2689728
50. Wang, Y., Wu, J., Chen, F., Wang, Z., Li, J., & Wang, L. (2024). Empirical Assessment of AI-Powered Tools for Vocabulary Acquisition in EFL Instruction. *IEEE Access*, *12*, 131892–131905. https://doi.org/10.1109/ACCESS.2024.3446657
51. Wang, Y., & Zheng, L. (2022). A Novel Deep Framework for English Communication Based on Educational Psychology Perspective. *Frontiers in Public Health*, *10*, 916101. https://doi.org/10.3389/fpubh.2022.916101
52. Wells, L., & Bednarz, T. (2021). Explainable AI and Reinforcement Learning—A Systematic Review of Current Approaches and Trends. *Frontiers in Artificial Intelligence*, *4*, 550030. https://doi.org/10.3389/frai.2021.550030
53. Xu, L., & Yu, W. (2024). Design and Implementation of Artificial Intelligence Online Learning Platform Based on Resource Scheduling Technology: *Journal of Cases on Information Technology*, *26*(1), 1–22. https://doi.org/10.4018/JCIT.349740
54. Yin, Y., Peng, H., & Liu, H. (2024). Research on the Construction of Smart Teaching Mode with Artificial Intelligence Technology Facilitating Education Informatization in Colleges and Universities. In *Applied Mathematics and Nonlinear Sciences* (Vol. 9, Issue 1). https://doi.org/10.2478/amns.2023.2.01409
55. Yuhua, D. (2024). Integrating Deep Learning into English Language Teaching Within the Digital Cultural Framework. *Computer-Aided Design and Applications*, 71–84. https://doi.org/10.14733/cadaps.2024.S16.71-84
56. Zimmerman, B. J., & Moylan, A. (2009). Self-regulation: Where metacognition and motivation intersect. *Handbook of Metacognition in Education*, 299–315.