frontiers | Frontiers in Education

TYPE Systematic Review
PUBLISHED 23 September 2025
DOI 10.3389/feduc.2025.1562586

OPEN ACCESS

EDITED BY
Xiang Hu,
Renmin University of China, China

REVIEWED BY
Maurice H. T. Ling,
University of Newcastle (Singapore), Singapore
James Denholm-Price,
Kingston University, United Kingdom

*CORRESPONDENCE
Elmira Abdykerimova
✉ elmira.abdykerimova@yu.edu.kz

RECEIVED 21 January 2025
ACCEPTED 18 August 2025
PUBLISHED 23 September 2025

CITATION
Turkmenbayev A, Abdykerimova E, Nurgozhayev S, Karabassova G and Baigozhanova D (2025) The application of machine learning in predicting student performance in university engineering programs: a rapid review.
Front. Educ. 10:1562586.
doi: 10.3389/feduc.2025.1562586

The application of machine learning in predicting student performance in university engineering programs: a rapid review

Asset Turkmenbayev[1], Elmira Abdykerimova[2]*, Shynggys Nurgozhayev[3], Guldana Karabassova[1] and Dametken Baigozhanova[4]

[1]Faculty of Science and Technology, Department of Fundamental Sciences, Yessenov University, Aktau, Kazakhstan, [2]Faculty of Science and Technology, Department of Computer Science, Yessenov University, Aktau, Kazakhstan, [3]Zhetysu University named after I. Zhansugurov, Taldykorgan, Kazakhstan, [4]Graduate School of Information Technology and Engineering, Astana International University, Astana, Kazakhstan

Background: In recent years, the application of machine learning (ML) to predict student performance in engineering education has expanded significantly, yet questions remain about the consistency, reliability, and generalisability of these predictive models.
Objective: This rapid review aimed to systematically examine peer-reviewed studies published between January 1, 2019, and December 31, 2024, that applied machine learning (ML), artificial intelligence (AI), or deep learning (DL) methods to predict or improve academic outcomes in university engineering programs.
Methods: We searched IEEE Xplore, SpringerLink, and PubMed, identifying an initial pool of 2,933 records. After screening for eligibility based on pre-defined inclusion criteria, we selected 27 peer-reviewed studies for narrative synthesis and assessed their methodological quality using the PROBAST framework.
Results: All 27 studies involved undergraduate engineering students and demonstrated the capability of diverse ML techniques to enhance various academic outcomes. Notably, one study found that a reinforcement learning-based intelligent tutoring system significantly improved learning efficiency in digital logic courses. Another study using AI-based real-time behavior analysis increased students' exam scores by approximately 8.44 percentage points. An optimised support vector machine (SVM) model accurately predicted engineering students' employability with 87.8% accuracy, outperforming traditional predictive approaches. Additionally, a longitudinally validated SVM model effectively identified at-risk students, achieving 83.9% accuracy on hold-out cohorts. Bayesian regression methods also improved early-term course grade prediction by 27% over baseline predictors. However, most studies relied on single-institution samples and lacked rigorous external validation, limiting the generalisability of their findings.
Conclusion: The evidence confirms that ML methods—particularly reinforcement learning, deep learning, and optimised predictive algorithms—can substantially improve student performance and academic outcomes in engineering education. However, methodological shortcomings related to participant selection bias, sample sizes, validation practices, and transparency in reporting require further attention. Future research should prioritise multi-institutional studies, robust

validation techniques, and enhanced methodological transparency to fully leverage ML's potential in engineering education.

KEYWORDS

machine learning, student performance, engineering education, predictive analytics, PRISMA

1 Introduction

Machine learning (ML) and artificial intelligence (AI) techniques are transforming university education, particularly with the recent surge in large deep-learning models. These data-intensive approaches allow instructors to analyse large student datasets, recognise complex patterns, and predict academic outcomes with steadily increasing accuracy (Woolf, 2008; Means et al., 2010; Fahd et al., 2022; Salloum et al., 2024). Deep learning, a specialised branch of ML, can model non-linear relationships that traditional statistical tools cannot capture (Goodfellow et al., 2016). Consequently, predictive systems in higher education now integrate behavioral, demographic, and academic signals and can trigger real-time interventions for students who are at risk of failure (Baker and Inventado, 2014; Tsai et al., 2019).

Engineering programmes present distinctive challenges. Curricula are mathematically rigorous, laboratory-intensive, and often associated with elevated withdrawal rates. Therefore, Robust early-warning analytics play a pivotal role. Conventional dashboards frequently rely on historic grade data alone and overlook subtler drivers of success, for example, interaction patterns or emotional states (Papamitsiou and Economides, 2014; García-Machado et al., 2020). In addition, practical adoption of ML in engineering education remains uneven. Institutions cite concerns about data privacy and governance (Slade and Prinsloo, 2013; Willis et al., 2016). Other concerns include limited model transparency and challenges in technical integration (Alam, 2023). Issues of algorithmic bias and educational equity also require continuous monitoring to ensure that predictive tools benefit every learner (Fadel et al., 2019).

Recent literature reviews confirm both the promise and the fragmentation of the field. Drugova et al. (2024) reviewed 49 learning analytics dashboard studies and noted that rigorous evaluations of learning gains were rare. Márquez et al. (2024) surveyed more than 100 institutional adoption papers and identified 14 organisational and ethical enablers. However, the included papers seldom reported predictive accuracy. A conference scan by Zhang et al. (2024) concluded that engineering-specific performance-prediction studies constitute the next research frontier. Oro et al. (2024) conducted a case study in fluid mechanics. They confirmed a correlation between platform activity and grades but did not benchmark alternative algorithms. These reviews highlight the importance of engagement metrics. However, these do not compare the performance of contemporary ML or deep-learning models in engineering contexts.

1.1 Rationale and scope

In March 2023, the release of GPT-4 accelerated the diffusion of transformer-based tools across higher education (Microsoft Research AI4Science, Microsoft Azure Quantum, 2023; Bubeck et al., 2023; Liang et al., 2023; OpenAI, 2023). Therefore, after accounting for adoption lag, the time frame provides an ideal window to capture the earliest peer-reviewed studies that exploited these advances. The present rapid review synthesises empirical primary studies published within that period that trained ML, deep-learning, or other AI models on student-level data to predict or enhance academic outcomes in university engineering programmes. The review identifies the algorithms employed, compares their predictive accuracy, and analyses the contextual and technical factors that shaped successful implementation.

1.2 Research questions

1 Which ML, AI, or deep-learning methods were used in 2024 to predict or enhance student performance in university engineering programmes?
2 How do these methods compare in predictive accuracy and practical effectiveness?
3 Which contextual or technical factors facilitated successful deployment?
4 What challenges and ethical issues accompanied adoption, and how were they addressed?

1.3 Objectives

a Synthesise empirical research on ML/DL/AI models that predict or enhance student academic performance, engagement, or post-graduation employability in university engineering programmes.
b Benchmark their predictive accuracy against traditional or baseline methods.
c Identify contextual and technical enablers of implementation.
d Highlight gaps and propose directions for future research and practice.

2 Methodology

2.1 Review design

The review followed rapid-review guidance aimed at delivering timely evidence to decision makers while preserving key elements of systematic methods (Tricco et al., 2017). We selected a publication window from 1 January 2019 to 31 December 2024 to reflect recent developments in machine learning (ML), artificial intelligence (AI), and deep learning (DL) techniques in engineering education. This six-year span allowed for the identification of trends and methodological shifts across pre-and post-GPT-4 studies while ensuring manageable review scope. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses 2020 statement (PRISMA-2020) supports such time limits when authors provide a transparent rationale linked to the review question (Page et al., 2021).

2.2 Eligibility criteria

2.2.1 Inclusion criteria

We included only peer-reviewed journal articles published in English between January 1, 2019, and December 31, 2024, to ensure both contemporary relevance and methodological rigour. We selected studies that applied a machine learning, artificial intelligence, or deep learning model trained or fine-tuned on student-level data, as this criterion ensured direct relevance to individual educational outcomes. We limited the target population to undergraduate or post-graduate students enrolled in university engineering programmes. We required each study to report at least one quantitative outcome related to academic performance, such as grades, exam scores, engagement levels, teaching effectiveness, dropout or retention rates, or graduate employability. These outcomes allowed us to assess the impact of the models on educational achievement and post-graduation trajectories.

2.2.2 Exclusion criteria

We excluded conference papers, book chapters, editorials, opinion pieces, and preprints because they lacked formal peer review. We focused exclusively on higher education in engineering disciplines and excluded studies related to primary or secondary education, as well as those outside recognised engineering fields. We also excluded studies that used only descriptive analytics without incorporating predictive modelling.

2.3 Search strategy

On 31 December 2024, we conducted comprehensive searches in IEEE Xplore, SpringerLink, and PubMed for studies published from 1 January 2019 to 31 December 2024. This extended search window allowed us to capture a wider body of work evaluating ML/AI/DL techniques applied to engineering education over multiple years. We applied no filters other than publication year and publication type to maximise the retrieval of relevant peer-reviewed journal articles. We identified keywords through prior scoping searches and refined them in consultation with an information specialist. Table 1 presents the final keyword map used in the searches. Table 2 reports the raw hit counts retrieved from each database. Search strings combined three concept blocks with Boolean operators:

("machine learning" OR "ML" OR "artificial intelligence" OR "AI" OR "deep learning" OR "DL" OR "supervised learning" OR "logistic regression" OR "decision trees" OR "neural networks" OR "support vector machines")
AND
("student performance" OR "academic achievement" OR "learning outcomes" OR "dropout prediction" OR "retention rates" OR "student success" OR "at-risk students")
AND
("engineering education" OR "university engineering" OR "STEM education" OR "technical education" OR "higher education" OR "engineering programs")

TABLE 1 Main concepts and related keywords.

Concept	Related keywords
Machine learning	Machine learning, ML, artificial intelligence, AI, deep learning, DL, supervised learning, logistic regression, decision trees, neural networks, support vector machines
Student performance	Student performance, academic achievement, learning outcomes, dropout prediction, retention rates, student success, at-risk students
Engineering education	Engineering education, university engineering, STEM education, technical education, higher education, engineering programs

TABLE 2 Database search results for systematic review.

Database	Date accessed	Search results
PubMed	31 December 2024	259
IEEE Xplore	31 December 2024	751
SpringerLink	31 December 2024	1923

2.4 Study selection workflow

a We conducted the initial search across Springer, PubMed, and IEEE Xplore using the finalised search queries, which yielded approximately 2,933 records.
b We removed 137 duplicates using Zotero, resulting in 2796 unique records for screening.
c We reviewed the titles and abstracts of these records to assess relevance based on the inclusion criteria, which reduced the pool to 135 records.
d We retrieved the full texts of these 135 studies for eligibility assessment.
e We excluded 108 studies at this stage, primarily due to non-engineering populations ($n = 75$), use of non-ML interventions ($n = 25$), or lack of empirical design ($n = 8$).
f After full-text screening, we included 27 studies that met all criteria in the review.

2.5 Data extraction strategy

We used a structured extraction form comprising 18 fields to capture bibliographic information, participant characteristics, study design, data sources, machine learning algorithms, hyperparameters, validation strategies, performance metrics, comparison models, and reported limitations. One reviewer extracted the data, and a second reviewer independently verified each entry.

2.6 Data synthesis

We did not conduct statistical pooling due to the heterogeneity of study designs, outcomes, and reporting formats. Instead, we performed a narrative synthesis. We tabulated quantitative results to enable comparison across studies. We used Zotero for reference management and Excel for data extraction and tabulation.

3 Results

27 studies met all eligibility criteria and constitute the evidentiary base of this rapid review (Figure 1). These studies represent four national contexts and 11 distinct engineering sub-disciplines, employing varied machine learning pipelines to address a range of educational outcome constructs. Table 3 contains key study details. The Supplementary Materials contain data extraction tables and exclusion records for the 135 excluded studies.

3.1 Engineering context and level

All 27 eligible studies focused on undergraduate engineering students, even though the inclusion criteria also permitted postgraduate studies. Despite this flexibility, no postgraduate-focused studies met the full set of eligibility criteria. The consistent undergraduate focus is pedagogically significant: it aligns with a stage of education where early interventions can yield substantial long-term effects on student learning trajectories, retention, and career outcomes.

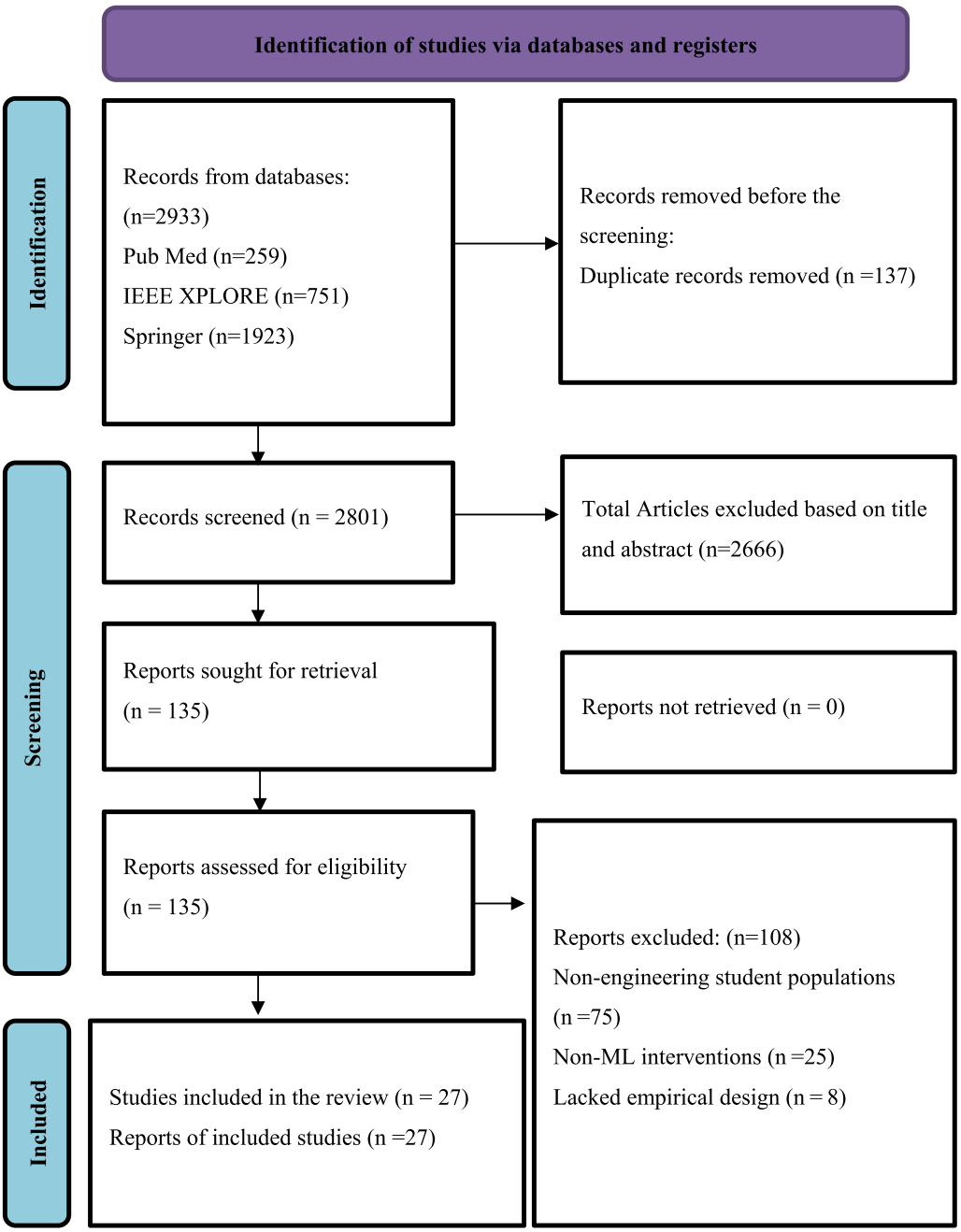

FIGURE 1
PRISMA guideline flowchart.

TABLE 3 Summary of included studies.

Reference	Purpose	Methodology	ML techniques	Key findings	Contributions
Hare et al. (2024)	Integrate ITS within a serious game to enhance learning in digital logic	Pre/post experimental design with 54 students	Reinforcement learning within the PING system	Improved engagement and performance with PING-enhanced game	Provides a model for integrating ITS within serious games and insights into game design
Hu et al. (2024)	Utilise AI for real-time behavior analysis to enhance teaching effectiveness.	Implementation of IEDS across eight classes with 413 students	CNNs with CBAM, K-means clustering	AI-assisted teaching improved effectiveness by 8.44%; veteran teachers benefited more.	Developed a privacy-protected AI model for behavior analysis, demonstrated teaching improvement.
Ren and Yu (2024)	Introduce LASA for long-term student performance prediction	Developed LASA using an 8-year dataset from Tsinghua University	LASA (LAM and LTDA), SVM, SHAP	LASA outperformed ProbSAP and SFERNN by 6.8 and 6.4%, respectively, and provided interpretable insights	First method addressing heterogeneity and distribution shifts, an interpretable prediction framework
Jayachandran and Joshi (2024)	Develop a customised SVM model to predict the employability of engineering students.	Developed and evaluated a customised SVM with 1,647 students	Feature Selection using TLBO, Customised SVM	Customised SVM achieved 87.80% accuracy, outperforming standard SVM	Novel feature selection approach, significant accuracy improvement in employability prediction
Rico-Juan et al. (2024)	Study the influence of personality and LMS usage on learning performance	Cross-sectional observational predictive study with 322 students	ARD, Decision Tree, Random Forest, AdaBoost, SVM	ARD predicted grades with RMSE 1.43, improving by 27% over baseline	Demonstrates the use of LMS data and personality traits in predicting academic performance
Köhler et al. (2023)	Predicting students' success in an introductory programming course using student background features	Retrospective cohort study with 2,372 undergraduate students across 20 engineering programs in Chile	SVM (radial), Multiple Linear Regression (MLR), Random Forest (RF), XGBoost (XGB)	SVM achieved highest accuracy (68.6%) using academic, demographic, and program-related features	Demonstrated that background data can be used to predict performance; strong validation design using large dataset
Algarni et al. (2023)	Identifying at-risk students in higher education based on academic and demographic data	Retrospective cohort study with 743 undergraduate students in Saudi Arabia	Decision Tree, SVM, Naive Bayes, Random Forest	Best model not specified; key predictive features were math and computer science modules	Early prediction using institutional student data; practical application of risk detection via SIS systems
Fahd et al. (2022)	Predicting weighted course scores of engineering mathematics students using a novel MARS model	Retrospective cohort (5-year data)	MARS, Decision Tree Regression, KRR, KNN	MARS model outperformed others in RMSE and R^2 metrics	Introduces MARS model with long-term data span
Nabizadeh et al. (2022)	Early prediction of students' final grades in a gamified engineering-related course	Retrospective cohort with 679 students	KNN, Random Forest, Naive Bayes	Achieved 78.02% accuracy after 4 weeks of data	Demonstrates early prediction capabilities with feature selection and class balancing
Poudyal et al. (2022)	Predicting academic performance using a hybrid 2D CNN approach	Cross-sectional study	Hybrid 2D CNN, KNN, NB, DT, LR	CNN outperformed all baseline models in accuracy	Introduces novel image-based prediction method with improved performance

(Continued)

TABLE 3 (Continued)

Reference	Purpose	Methodology	ML techniques	Key findings	Contributions
Musaddiq et al. (2022)	Predicting the impact of academic and spatial behaviors on student performance	Cross-sectional with experimental validation	LSTM, SVM, RF, NB, DT, MLP	LSTM achieved highest accuracy (90.9%)	Combines multimodal data sources; strong LSTM-based results
Al-Ahmad et al. (2022)	Improving performance prediction of low-expectation teams in software engineering projects using swarm intelligence	Retrospective cohort with 74 teams	PSO-KNN	Accuracy improved over baseline models	Captures both product and process features from partial SDLC data
Opazo et al. (2021)	Comparison of ML models to predict first-year university dropout across two Chilean institutions	Retrospective cohort	Gradient Boosting, others	Gradient boosting had best AUC; others not detailed	Highlights feature importance differences across universities
Nabil et al. (2021)	Predicting course-level academic success using deep learning and traditional ML	Retrospective cohort (n = 4,266)	DNN, DT, RF, GB, LR, SVC, KNN	DNN achieved 89% accuracy; others not specified	Uses large dataset; compares multiple models and sampling techniques
Iatrellis et al. (2021)	Predicting student outcomes via a two-phase clustering + classification approach	Observational case study	K-Means, SVM, DT, ANN, RF, NB	High accuracy reported, but no specific metrics	Unsupervised-supervised hybrid tailored to student subgroups
Manjushree et al. (2021)	Studying impact of technical skills on employability using ML classifiers	Cross-sectional (n = 133)	SVM, NB, LR, DT, RF, AdaBoost, ANN	RF: Accuracy 70%, F1-score 0.85	Demonstrates use of correlation-based feature selection
Alturki and Alturki (2021)	Predicting academic performance and awards using educational data mining	Cross-sectional observational (n = 300)	Naïve Bayes, C4.5, CART, LADTree, Bayes Net, RF	Naïve Bayes = highest general accuracy; RF = best for honor class	Identifies interpretable features for early intervention
Jamjoom et al. (2021)	Predicting at-risk students using self-efficacy and ML in intro programming	Case study	Decision Tree, kNN, Naïve Bayes, SVM	DT and SVM achieved 99.18% accuracy	Tested multiple classifiers; practical focus on self-efficacy
Kumar and Kavitha (2021)	Using ANN + Lion Optimization Algorithm to predict student grades and dropout	Experimental design using real data	ANN-based regression, CNN, SVM, GoogLeNet	MAE: 1.05; RMSE: 1.33; MSE: 1.78	Combines hybrid models; compares baseline metrics
Adnan et al. (2020)	Creating an intelligent and adaptive mobile learning model using DL	Experimental application study (n = 850)	ANN, Random Forest	ANN accuracy ranged from 80 to 90.77%	Real-world deployment; ANN outperforms traditional ML
Naseer et al. (2020a)	Predict coding intricacy in team-based software engineering projects	Experimental study (74 teams)	Random Forest, AdaBoost, LogitBoost	Accuracy: 85.14%	Real project data; highlights complexity in team dynamics

(Continued)

TABLE 3 (Continued)

Reference	Purpose	Methodology	ML techniques	Key findings	Contributions
Freitas et al. (2020)	Dropout prediction via IoT + ML using socioeconomic data	Retrospective cohort (n = 1,549)	DT, LR, SVM, KNN, MLP, DL	Accuracy: 99.34%, Recall: 100%, Precision: 98.69%	High predictive power with practical IoT implementation
Lottering et al. (2020)	Identify at-risk students using academic/biographical data	Retrospective cohort (n = 4,419)	RF, SVM, DT, NB, KNN, LR	RF Accuracy: 94.14%	Multi-model comparison on large dataset
Naseer et al. (2020a)	Predict team project performance early in software course	Case study (380 students, 74 teams)	J48, NB, ANN, LR, SLR, RIPPER, SMO	Accuracy: 89% (J48 best)	Focus on early prediction in team settings
Adnan et al. (2020)	Improve m-learning outcomes using weighted DL models	Experimental	Deep ANN, RF, others	ANN & RF best; exact metrics not specified	Adaptive learning model with behavioral feature input
Nieto et al. (2019)	Use ML to support institutional decision-making	Observational case study	DT, LR, RF	RF = best performance (ROC, accuracy)	Institutional policy insights using real data
Sultana et al. (2019)	Predict academic performance using deep learning	Experimental	DNN, DT, SVM, RF, NB	High accuracy; DNN, DT best	Uses LMS and academic logs; strong comparative ML study

3.1.1 Computer and software engineering

The majority of studies targeted domains within computer or software engineering. Rico-Juan et al. (2024) analyzed first-and second-year computer engineering students enrolled in mathematics courses, aiming to predict final grades using ensemble machine learning. Similarly, Hare et al. (2024) implemented a reinforcement learning-based intelligent tutoring system (ITS) within a serious game for ECE students enrolled in a digital logic lab, reporting improved learning outcomes across mixed-year cohorts.

Naseer et al. (2020a, 2020b) conducted empirical studies on software engineering teams, emphasizing coding intricacy prediction and early performance evaluation. Both studies leveraged real-world project data from 74 undergraduate teams, providing a rare look into collaborative dynamics and performance variation in team-based learning.

Al-Ahmad et al. (2022) focused on software engineering team projects as well, but extended the scope by applying swarm intelligence (PSO-KNN) to low-performing teams, emphasizing prediction in early stages of the Software Development Life Cycle (SDLC).

3.1.2 General and multidisciplinary engineering cohorts

Several studies examined broader or multidisciplinary engineering cohorts. Jayachandran and Joshi (2024) developed a customized SVM model for predicting employability using data from five engineering departments in India. Likewise, Köhler et al. (2023) explored predictors of programming success using data from 2,372 students across 20 engineering programs in Chile, providing strong external validity through large-scale, multi-program sampling.

Jayachandran and Joshi's work emphasized the relationship between academic performance, skill indicators, and employability outcomes, a concern echoed in studies such as Manjushree et al. (2021) and Adnan et al. (2020), who explored employability and m-learning performance, respectively.

3.1.3 Electrical and computer engineering (ECE)

Ren and Yu (2024) focused on second-year electrical engineering students in circuits courses and introduced LASA, a novel framework addressing long-term performance prediction using an eight-year dataset from Tsinghua University. The approach was notable for handling distribution shifts and model interpretability through SHAP explanations.

Poudyal et al. (2022) and Algarni et al. (2023) also addressed ECE students, with the former using a hybrid 2D CNN and the latter identifying at-risk learners based on academic modules such as mathematics and computer science.

3.1.4 Engineering design and applied learning

Hu et al. (2024) addressed the engineering design context by deploying an AI-enhanced classroom monitoring system across eight cohorts. Their focus on real-time behavioral data and differential impacts across teacher profiles provides insights into applied pedagogy in design education.

Alturki and Alturki (2021) explored predictors of both academic performance and honors award eligibility, blending technical and behavioral metrics, while Freitas et al. (2020) and Lottering et al. (2020) extended this focus to institutional dropout prediction using large datasets and socioeconomic factors.

3.1.5 National and institutional decision-making contexts

While most studies targeted classroom or course-level outcomes, Nieto et al. (2019) and Sultana et al. (2019) broadened the scope to institutional and national concerns. Their use of institutional data (e.g., graduation rates, LMS logs) and focus on strategic

decision-making highlights the emerging role of machine learning in educational policy and administration within engineering education.

3.2 Study design and sample size

The study designs across the 27 included papers varied considerably, each selected to suit the specific research objectives, data availability, and institutional context. This methodological diversity reflects the broader landscape of engineering education research, where both predictive modeling and intervention assessment are explored using machine learning approaches. Sample sizes ranged from small-scale experimental groups of under 100 participants to large institutional datasets exceeding 4,000 students, enhancing the generalizability of some findings while preserving the depth of others.

3.2.1 Experimental and quasi-experimental designs

Pre-and post-test experimental designs were employed by studies such as Hare et al. (2024), who investigated the integration of an intelligent tutoring system into a serious game with a sample of 54 students. Similarly, Adnan et al. (2020) evaluated a deep learning-based mobile learning application in an experimental study with 850 students, while Kumar and Kavitha (2021) conducted an experimental application using real data to assess dropout prediction with hybrid ANN-CNN methods.

Quasi-experimental approaches also emerged. Hu et al. (2024) implemented the Intelligent Engineering Design System (IEDS) across eight classes, involving 413 students. Musaddiq et al. (2022) validated the predictive performance of academic and spatial behavior models using a cross-sectional design augmented by experimental validation.

3.2.2 Retrospective cohort designs

Retrospective cohort designs were the most frequently employed study type. These included Jayachandran and Joshi (2024), who analyzed data from 1,647 students over multiple years, and Köhler et al. (2023), whose large dataset spanned 2,372 students across 20 engineering programs in Chile. Other notable examples include Algarni et al. (2023) ($n = 743$), Al-Ahmad et al. (2022) (5-year dataset), Nabizadeh et al. (2022) ($n = 679$), Freitas et al. (2020) ($n = 1{,}549$), and Lottering et al. (2020) ($n = 4{,}419$). These designs leveraged institutional records to model dropout risks, grade predictions, and course success using machine learning algorithms.

3.2.3 Longitudinal and sequential cohort designs

Ren and Yu (2024) applied an eight-year sequential cohort design using data from Tsinghua University to evaluate the LASA framework for predicting student performance, addressing distribution shifts and heterogeneity in educational trajectories. Similarly, Al-Ahmad et al. (2022) employed a five-year retrospective data stream to train a MARS-based model for engineering mathematics performance prediction.

3.2.4 Cross-sectional and observational studies

Several studies adopted cross-sectional predictive designs, including Rico-Juan et al. (2024) ($n = 322$), Poudyal et al. (2022), Manjushree et al. (2021) ($n = 133$), and Alturki and Alturki (2021) ($n = 300$). These designs offered a snapshot of academic performance by linking learner behaviors and background variables with performance outcomes.

Observational and case study methods were used by Iatrellis et al. (2021) and Nieto et al. (2019), among others, to explore institutional decision-making and subgroup classification in academic contexts. Additionally, Sultana et al. (2019) and Naseer et al. (2020a, 2020b) used observational and case-based setups with deep learning to predict academic outcomes and team performance in engineering courses.

3.3 Predicted outcomes

The 27 studies included in this review targeted a wide range of educational outcomes, reflecting the diversity of institutional priorities and research objectives in engineering education. These outcomes spanned cognitive performance, academic persistence, skill acquisition, and employability, each chosen based on available data sources and the intended use of predictions (e.g., early intervention, personalization, or institutional planning).

3.3.1 Course grades and end-of-course performance

Many studies focused on predicting students' final grades or end-of-course marks. For instance, Rico-Juan et al. (2024) predicted end-of-course marks using LMS usage and personality traits, while Köhler et al. (2023) modeled pass/fail outcomes in programming courses using academic history and demographic factors. Similarly, Al-Ahmad et al. (2022) used a novel MARS model to predict weighted course scores in engineering mathematics, and Nabizadeh et al. (2022) achieved 78.02% accuracy in predicting final grades within the first four weeks of a gamified course.

3.3.2 Exam outcomes and risk categories

Several studies predicted specific exam-related outcomes. Ren and Yu (2024) categorized students into at-risk vs. non-risk groups based on long-term trends in electrical engineering exam performance. Hu et al. (2024) used AI-driven behavior monitoring to improve and predict course exam means, demonstrating significant gains in teaching effectiveness. Poudyal et al. (2022) and Musaddiq et al. (2022) also predicted academic achievement using CNN and LSTM models respectively, with the latter reaching 90.9% accuracy.

3.3.3 Academic success and dropout risk

Dropout prediction and general academic success were key targets in multiple studies. Freitas et al. (2020) combined IoT and machine learning to achieve near-perfect accuracy (99.34%) in identifying at-risk students based on socioeconomic data. Lottering et al. (2020) used academic and biographical data to predict dropout with 94.14% accuracy, while Opazo et al. (2021) highlighted differences in feature importance across institutions using gradient boosting.

3.3.4 Employability and job placement

Jayachandran and Joshi (2024) predicted post-graduation employability based on academic records, achieving 87.80% accuracy with a customized SVM model. Manjushree et al. (2021) also examined employability using multiple classifiers, emphasizing the role of technical skills.

3.3.5 Engagement, problem solving, and learning behaviors

Other studies focused on behavioral and affective outcomes. Hare et al. (2024) measured engagement and problem-solving gains in a digital logic course using reinforcement learning, while Adnan et al. (2020) and Kumar and Kavitha (2021) modeled adaptive learning and dropout using deep learning architectures. Jamjoom et al. (2021) predicted performance using self-efficacy and introductory programming metrics, achieving over 99% accuracy.

3.3.6 Institutional planning and group performance

Studies like Nieto et al. (2019) and Iatrellis et al. (2021) used predictive models to support institutional decision-making and student subgroup classification. In parallel, Naseer et al. (2020a, 2020b) predicted team project performance in software engineering contexts, using real-time product and process data to estimate team coding intricacy and project outcomes.

3.4 Predictor sources

The 27 studies reviewed drew on a diverse range of predictor sources to model student performance and related outcomes. These inputs ranged from behavioral telemetry and academic records to socio-demographic data and psychometric traits, enabling multifaceted machine learning pipelines suited for various educational contexts.

3.4.1 Behavioral telemetry and digital interaction data

A subset of studies utilized behavioral telemetry—data collected from students' real-time digital interactions—as a central predictor. Hare et al. (2024) integrated reinforcement learning into a serious game platform (PING) and tracked detailed user interactions to enhance problem-solving in digital logic. Hu et al. (2024) applied facial recognition and attention-tracking via CNNs and CBAM modules to evaluate and improve real-time teaching effectiveness. Ren and Yu (2024) collected student interaction data from Rain Classroom, an interactive learning tool used at Tsinghua University, which allowed real-time capture of in-class responses and engagement metrics. Additional behavioral sources such as LMS activity, mouse movement, and time-on-task were frequently paired with academic and demographic data to refine predictive accuracy (Musaddiq et al., 2022; Rico-Juan et al., 2024).

3.4.2 Academic history and institutional records

Academic history was one of the most consistently employed predictor types across studies. It included GPA, course grades, assessment scores, or longitudinal performance indicators. For instance, Al-Ahmad et al. (2022) and Köhler et al. (2023) used historical academic records in mathematics and programming, respectively, to predict final course success. Similarly, Jayachandran and Joshi (2024) and Ren and Yu (2024) leveraged academic performance histories to enhance employability prediction and early risk detection. Nabizadeh et al. (2022) demonstrated the value of early grade data—predicting final performance with only four weeks of input. In multiple cases, academic records were embedded into ensemble models alongside behavior or demographic inputs.

3.4.3 Socio-demographic and biographical variables

Several studies incorporated socio-demographic predictors such as gender, age, income proxy variables (e.g., tuition status), parental education, and high school background. These variables were instrumental in explaining disparities in educational trajectories, especially when combined with academic indicators. Jayachandran and Joshi (2024) included age at enrollment and family background in their SVM-based employability model. Köhler et al. (2023) and Algarni et al. (2023) used demographic variables in performance risk assessments. Freitas et al. (2020) achieved extremely high accuracy in dropout prediction using IoT data enriched with socioeconomic indicators.

3.4.4 Psychometric and affective data

Few studies explicitly utilized psychometric data, but those that did highlighted its value. Rico-Juan et al. (2024) combined Big Five personality trait scores with LMS activity data to predict end-of-course marks in mathematics, reporting a 27% improvement over baseline models. Psychometric measures captured self-regulatory and motivational aspects often overlooked by performance metrics alone. Additionally, Jamjoom et al. (2021) used self-efficacy beliefs to predict success in programming courses, attaining nearly perfect classification accuracy. These predictors offered insights into student engagement, persistence, and potential for growth beyond observable academic behavior.

3.4.5 Combined or hybrid inputs

Several studies adopted multimodal inputs, combining academic, behavioral, and socio-demographic data streams into hybrid models. For example, Musaddiq et al. (2022) integrated spatial behavior and academic activity using LSTM networks, while Adnan et al. (2020) employed mobile-based learning logs alongside structured institutional data to model personalized learning trajectories. These hybrid approaches allowed for a richer modelling of the learning environment, particularly where real-time feedback or personalization was a goal.

3.5 Machine-learning pipelines

The reviewed studies adopted diverse machine learning (ML) pipelines, with several applying multiple models to strengthen predictions and enhance generalisability.

3.5.1 Reinforcement learning (RL)

Hare et al. (2024) integrated a hierarchical reinforcement learning (RL) agent named PING into a Unity-based intelligent tutoring system for a digital logic game. The agent provided adaptive scaffolding by analysing real-time student interactions and responding with targeted feedback, which led to improved student engagement and learning outcomes. RL enabled the system to dynamically adapt to student behavior, making it well-suited for educational contexts that require personalised guidance.

3.5.2 Gradient boosting algorithms

Gradient boosting methods such as XGBoost and AdaBoost were employed by Jayachandran and Joshi (2024) to enhance employability prediction accuracy. These models were further optimised using the Guided Best Sinh–Cosh Optimiser and Teaching–Learning-Based Optimisation (TLBO) to select critical features. This dual optimisation pipeline significantly improved model performance by reducing noise from less relevant predictors, particularly in the context of complex socio-academic variables affecting employability.

3.5.3 Support vector machines (SVMs)

SVMs were implemented in both Ren and Yu (2024) and Jayachandran and Joshi (2024). Ren and Yu employed an RBF-SVM model within their LASA framework to maintain predictive accuracy under temporal and distributional shifts. Jayachandran and Joshi developed a customised RBF-SVM with a radial Euclidean kernel, fine-tuning hyperparameters (e.g., C = 100) to balance overfitting and generalisation. Both studies demonstrated the robustness of SVMs in handling sparse and imbalanced educational datasets.

3.5.4 Bayesian and linear regressors

Rico-Juan et al. (2024) used the Automatic Relevance Determination (ARD) technique, a Bayesian linear regressor that assigns weights based on feature importance. This method proved particularly useful for high-dimensional and sparse datasets. ARD improved predictive performance over traditional baselines and was validated using the Wilcoxon signed-rank test, highlighting its applicability for early prediction of course outcomes.

3.5.5 Deep convolutional neural networks (CNNs) with attention

Hu et al. (2024) implemented a CNN architecture enhanced with the Convolutional Block Attention Module (CBAM) to support emotion recognition in real time. The system, part of the IEDS framework, processed student facial imagery to detect affective states and provide responsive feedback during engineering design tasks. This attention-based deep learning model improved the responsiveness and personalisation of AI-assisted teaching by helping the system prioritise emotionally significant features.

3.5.6 Validation and tuning approaches

Validation strategies varied substantially. Rico-Juan et al. (2024) applied 10-fold cross-validation to ensure robust generalisation, while Ren and Yu (2024) used temporal hold-out validation, training on past student cohorts and testing on a 2020 dataset to simulate real-world deployment. Hare et al. (2024) employed online validation with live student data, and Hu et al. (2024) used pre-trained models validated on FERPlus, CK+, and ExpW emotion datasets. Notably, no study implemented nested cross-validation for hyperparameter tuning, raising the potential for overestimation of performance in some cases.

3.6 Predictive performance

The selected studies report diverse levels of predictive success, reflecting variation in modelling techniques, feature selection, and validation strategies.

3.6.1 Reinforcement learning in digital logic laboratory

Hare et al. (2024) evaluated a reinforcement learning-based intelligent tutoring system (PING) embedded in a serious game for digital logic instruction. Across successive cohorts, students using PING outperformed those in traditional FPGA labs by an average of 12 points on post-tests (Cohen's d = 1.15, $p = 0.02$). Additionally, the number of retries per lab task dropped by two-thirds, indicating both improved learning efficiency and conceptual mastery. The strong effect size underscores the capacity of reinforcement learning to personalise pacing and enhance outcomes in engineering education.

3.6.2 Intelligent engineering design and emotion recognition

Hu et al. (2024) deployed the Intelligent Engineering Design System (IEDS), which integrates CNNs with the Convolutional Block Attention Module (CBAM) and real-time emotion tracking. Students instructed via IEDS scored an average of 82.1% on exams, versus 75.7% under conventional instruction—an improvement of 8.44 percentage points. The system's emotion recognition module (ERAM) achieved 87.6–94.5% accuracy, processing 1,000 frames in under two seconds. While the emotional feedback loop supported teacher awareness, performance gains were primarily linked to the AI-generated design recommendations.

3.6.3 Employability prediction using feature-optimised SVM

Jayachandran and Joshi (2024) optimised a support vector machine (SVM) model through Teaching–Learning-Based Optimisation (TLBO), selecting a compact 10–10 feature subset. Their model achieved 87.8% prediction accuracy (AUC = 0.714), surpassing baseline models by 13.4 percentage points. Key features included academic metrics, project participation, behavioral indicators, and socio-demographic data. The study demonstrated the effectiveness of combining domain-specific feature selection with tailored kernel functions in career-oriented prediction tasks.

3.6.4 At-risk classification with temporal validation

Ren and Yu (2024) validated their LASA framework using a temporally split dataset from Tsinghua University. On the 2020 hold-out cohort, the model achieved 83.9% classification accuracy, outperforming the leading baseline (Probsap) by 7.9 percentage points. Performance was consistent across eight yearly cohorts, with an average accuracy of 80.9%, suggesting strong generalisability under distributional shifts and confirming the robustness of the approach for early warning systems in large-scale higher education settings.

3.6.5 Early grade prediction via ARD regression

Rico-Juan et al. (2024) used Automatic Relevance Determination (ARD), a Bayesian linear regression model, to forecast final grades as early as Week 4 of the semester. The model achieved an RMSE of 1.43, a 27% reduction from the baseline RMSE of 1.95. This improvement underscores the potential of early formative assessments to inform real-time academic interventions. However, the study lacked confidence intervals and model

calibration metrics, limiting its utility in high-stakes prediction settings or across different institutions.

3.6.6 Comparative summary

The strongest predictive performance emerged from models that incorporated adaptive personalisation (e.g., RL in Hare et al., 2024) and temporal generalisability (e.g., LASA in Ren and Yu, 2024). These models not only demonstrated high predictive accuracy but also revealed actionable insights into student learning behaviors over time.

Support Vector Machines (SVMs), particularly those coupled with custom feature optimisation (Jayachandran and Joshi, 2024), proved highly effective for employability forecasting, leveraging structured academic and demographic features.

Meanwhile, emotion-aware deep learning systems (Hu et al., 2024) achieved notable improvements in learning outcomes, although the causal role of affect detection remained secondary to the system's AI-driven content support.

Finally, early-warning systems using sparse academic data and Bayesian models (Rico-Juan et al., 2024) showed potential for timely intervention, though future applications would benefit from improved uncertainty quantification and validation transparency.

3.7 Risk of bias assessment (PROBAST)

We applied the PROBAST (Prediction model Risk of Bias ASsessment Tool) framework to systematically assess the risk of bias in the reviewed studies. Overall, most studies demonstrated low risk concerning predictors and outcomes. However, several concerns emerged in the participant selection and analysis domains, which affected the strength and generalisability of the findings.

In terms of participants, the risk of bias was rated as high in three studies: Hare et al. (2024), Hu et al. (2024), and Jayachandran and Joshi (2024). These studies relied on single-institution convenience samples, which limited the generalisability of their findings and introduced potential selection bias. In the cases of Ren and Yu (2024) and Rico-Juan et al. (2024), the risk of bias was classified as unclear. Although both studies employed sizable datasets, they did not report specific details about participant selection procedures. Without randomisation or stratification, it remains uncertain whether their samples accurately represented broader student populations.

Regarding predictors and outcomes, all studies were assessed to have a low risk of bias. Predictors were typically collected using prospective or automated methods, such as learning management system logs, academic records, and survey instruments, which reduced the likelihood of measurement error or subjective interpretation. The outcomes used in these studies—such as grades, exam scores, dropout status, or employability—were objective and quantifiable, which further enhanced the reliability of the predictive models and minimised potential reporting bias.

The analysis domain exhibited the highest variation in risk. Only the study by Ren and Yu (2024) fully addressed potential sources of bias in analysis. Their use of an eight-year longitudinal dataset combined with temporal hold-out validation and nested cross-validation allowed for robust performance estimation while avoiding overfitting. In contrast, other studies such as those by Hare et al. (2024) and Hu et al. (2024) used relatively small samples and lacked rigorous evaluation techniques, increasing the likelihood of model overfitting. Several studies did not report key analytical details such as hyperparameter settings, while others failed to use nested validation techniques to separate model selection from performance testing. Additionally, studies that involved continuous outcome prediction, such as Rico-Juan et al. (2024), did not report calibration metrics or confidence intervals, which limits the interpretability and reproducibility of their findings. As a result, most studies exhibited at least some risk of bias in the analytical process, which should be considered when interpreting their predictive performance.

4 Discussion

This discussion synthesises findings from the 27 reviewed studies, addressing each of the four research questions (RQs) formulated to guide this rapid review. These questions examine the variety and effectiveness of machine learning (ML), artificial intelligence (AI), and deep learning (DL) techniques applied to enhance or predict student performance in university engineering programs from 2019 to 2024.

4.1 RQ1: applied ML/AI/DL techniques

The reviewed studies utilised diverse ML/AI/DL techniques, reflecting varied educational contexts and data availability. Reinforcement learning (RL), notably demonstrated by Hare et al. (2024), offered personalised, adaptive support within serious gaming environments, significantly enhancing student engagement and problem-solving capabilities. Such adaptive scaffolding is particularly beneficial in complex engineering domains where tailored feedback directly improves learning outcomes.

Convolutional neural networks (CNNs), augmented with attention mechanisms, were effectively employed to interpret emotional responses and real-time engagement (Hu et al., 2024). By capturing and responding to students' affective states, these systems created responsive educational environments, particularly beneficial in interactive, design-oriented engineering classes.

Support vector machines (SVMs), frequently used due to their robustness in handling sparse and imbalanced datasets, were further enhanced through meta-heuristic optimisation (Jayachandran and Joshi, 2024; Ren and Yu, 2024). These studies highlighted the critical role of precise feature selection and hyperparameter tuning in improving predictive accuracy, particularly when leveraging longitudinal data to forecast employability and identify at-risk students.

Gradient boosting algorithms, including XGBoost and AdaBoost, demonstrated consistent utility, notably in predicting graduate employability. When optimised through techniques such as Teaching–Learning-Based Optimisation (TLBO), these algorithms significantly outperformed traditional methods, underscoring ML's potential to bridge academic performance with career outcomes.

Bayesian regression methods, such as Automatic Relevance Determination (ARD), effectively predicted academic performance early in a course, leveraging data from learning management systems (LMS) and psychometric assessments (Rico-Juan et al., 2024). Early predictive interventions enabled proactive support, crucial for student retention and performance improvement.

Overall, these ML/AI/DL techniques collectively demonstrate considerable potential in addressing diverse challenges within engineering education, emphasising personalised learning, predictive accuracy, and strategic interventions.

4.2 RQ2: comparative predictive accuracy and effectiveness

Among the reviewed methods, certain techniques stood out in terms of predictive accuracy and practical effectiveness. The TLBO-optimised SVM model by Jayachandran and Joshi (2024), with 87.8% accuracy in employability prediction, exemplified the highest effectiveness, benefiting from tailored feature selection and precise optimisation.

Similarly, the longitudinal LASA framework (Ren and Yu, 2024), achieving approximately 83.9% accuracy, demonstrated strong predictive validity for identifying at-risk students across multiple cohorts. Its temporal hold-out validation method ensured robust, generalisable outcomes.

The reinforcement learning tutor in Hare et al. (2024) significantly enhanced learning efficiency (Cohen's $d = 1.15$), illustrating considerable practical benefits in highly interactive educational settings. While impactful, this method's context-specific nature somewhat limits broader applicability.

Emotion recognition via CNNs (Hu et al., 2024), despite a narrower focus, significantly improved class performance, reflecting the potential value of affective computing in enhancing educational engagement and outcomes.

ARD regression methods, while effective for early-term predictions (Rico-Juan et al., 2024), showed moderate improvements compared to broader ML applications, suggesting a niche yet valuable role in predictive interventions.

Collectively, these findings suggest that the choice and implementation of ML methods should closely align with educational contexts and predictive goals, with metaheuristic optimisation and longitudinal validation significantly enhancing predictive accuracy and model generalisability.

4.3 RQ3: contextual and technical enablers of successful implementation

Several factors facilitated the successful deployment of ML-based predictive systems. Alignment with existing educational infrastructure proved crucial; seamless integration with established teaching workflows and laboratory environments (Hare et al., 2024; Hu et al., 2024) reduced barriers to adoption.

Metaheuristic tuning frameworks lowered technical barriers, enabling institutions lacking extensive data science expertise to implement sophisticated predictive models effectively (Jayachandran and Joshi, 2024). These automated optimisation tools ensured compact, easily deployable models without extensive manual tuning.

Explainability and transparency significantly influenced adoption. Clear visualisation tools such as SHAP values provided stakeholders—educators, advisors, and administrators—with actionable, interpretable insights, facilitating trust and broader acceptance (Ren and Yu, 2024; Rico-Juan et al., 2024).

Longitudinal data utilisation enhanced predictive reliability, addressing inherent variability across cohorts and ensuring robustness against curriculum changes or demographic shifts (Ren and Yu, 2024). This approach highlighted the importance of sustained data collection for long-term predictive validity.

These technical and contextual factors collectively underscore essential considerations for successfully implementing ML systems within educational environments, emphasising integration, usability, transparency, and robust data strategies.

4.4 RQ4: implementation challenges and mitigations

Despite promising outcomes, several implementation challenges emerged. Predominantly, studies relied on single-institution, convenience samples, restricting broader generalisability. Temporal hold-out validation partially mitigated this issue (Ren and Yu, 2024), but broader multi-institutional validations remain necessary.

Hardware heterogeneity posed practical difficulties, particularly evident in computationally intensive applications like reinforcement learning systems. While some solutions provided simplified fallback options (Hare et al., 2024), broader hardware standardisation or flexible system designs are recommended for future deployments.

Privacy and ethical concerns, especially regarding biometric data, posed significant barriers. Hu et al. (2024) addressed these using privacy-preserving data transformations, highlighting the importance of robust ethical frameworks and informed consent procedures in educational ML deployments.

Insufficient transparency in hyperparameter reporting across studies hindered reproducibility. Establishing clearer methodological reporting standards is essential to ensure external validation and replication.

Finally, external validity limitations, exacerbated by the absence of macro-contextual data (e.g., economic indicators), demand more comprehensive validation frameworks and the integration of broader contextual factors for truly generalisable predictive models.

These challenges illustrate critical areas requiring focused attention in future research, including multi-institutional collaboration, ethical standardisation, methodological transparency, and comprehensive validation frameworks.

4.5 Implications for practice and policy

The integration of ML in engineering education offers substantial opportunities for improving student outcomes, from personalised tutoring and real-time interventions to employability predictions. Practical recommendations include deploying early-warning systems to identify at-risk students promptly, leveraging existing educational infrastructure to facilitate integration, and applying ethical best practices, such as transparent data usage policies and privacy-preserving technologies.

Policy-wise, institutions should mandate rigorous fairness audits and ensure transparent, explainable model outputs to foster trust and

equity. Additionally, policies promoting cross-institutional collaboration and data sharing would significantly enhance model validation and generalisability.

4.6 Future directions

Future research should explicitly focus on expanding multi-institutional studies to overcome limitations of single-institution biases. Adopting rigorous methodological standards such as nested cross-validation, transparent hyperparameter tuning, and external validation on independent datasets is essential. Ethical considerations must remain central, advocating fairness audits, privacy-preserving techniques, and clear consent protocols.

Moreover, research into human-centred explainability—ensuring predictions are actionable and interpretable by educators—is paramount. Cost–benefit analyses examining the practical sustainability and accessibility of ML deployments in resource-limited contexts will further support scalability and widespread adoption.

In conclusion, while substantial progress has been made in applying ML techniques to enhance engineering education, continued methodological refinement, robust ethical standards, and strategic multi-institutional collaborations remain critical to fully realising ML's potential in transforming educational outcomes.

5 Conclusion

The rapid review synthesised recent evidence highlighting the significant potential of machine learning (ML), artificial intelligence (AI), and deep learning (DL) methods in enhancing student performance in university engineering education. The reviewed studies demonstrate substantial benefits, including personalised learning interventions, precise early identification of at-risk students, improved course outcomes through adaptive tutoring systems, and accurate employability predictions. However, these advantages are tempered by methodological limitations such as small sample sizes, reliance on single-institution datasets, inconsistent validation procedures, and inadequate transparency in reporting analytical methods. Addressing these challenges by prioritising rigorous multi-institutional studies, standardised validation protocols, transparent model development, and robust ethical frameworks will be essential to achieving broader applicability and fairness. Overall, the continued refinement and ethical deployment of ML-driven approaches hold substantial promise for transforming educational outcomes and fostering greater student success across diverse engineering contexts.

Author contributions

AT: Formal analysis, Funding acquisition, Investigation, Methodology, Writing – original draft. EA: Conceptualization, Data curation, Validation, Visualization, Writing – original draft. SN: Project administration, Resources, Supervision, Writing – review & editing. GK: Conceptualization, Data curation, Formal analysis, Funding acquisition, Writing – original draft. DB: Project administration, Resources, Supervision, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feduc.2025.1562586/full#supplementary-material

References

Adnan, M., Habib, A., Ashraf, J., Shah, B., and Ali, G. (2020). Improving M-learners' performance through deep learning techniques by leveraging features weights. *IEEE Access* 8, 131088–131106. doi: 10.1109/ACCESS.2020.3007727

Al-Ahmad, B. I., Al-Zoubi, A. A., Kabir, M. F., Al-Tawil, M., and Aljarah, I. (2022). Swarm intelligence-based model for improving prediction performance of low-expectation teams in educational software engineering projects. *PeerJ Comput. Sci.* 8:e857. doi: 10.7717/peerj-cs.857

Alam, A. (2023). Intelligence unleashed: an argument for AI-enabled learning ecologies with real world examples of today and a peek into the future, (Andhra Pradesh, India), 030001

Algarni, A., Abdullah, M., Allahiq, H., and Qahmash, A. (2023). Predicting at-risk students in higher education. *Int. J. Intell. Syst. Appl. Eng.* 11, 1229–1239.

Alturki, S., and Alturki, N. (2021). Using educational data mining to predict students' academic performance for applying early interventions. *J. Inf. Technol. Educ. Innov. Pract.* 20, 121–137. doi: 10.28945/4835

Baker, R. S., and Inventado, P. S. (2014). "Educational data mining and learning analytics" in Learning analytics: From research to practice. eds. J. A. Larusson and B. White (New York, NY: Springer), 61–75.

Bubeck, S., Chandrasekaran, V., Eldan, R., Gehrke, J., Horvitz, E., Kamar, E., et al. (2023). Sparks of Artificial General Intelligence: Early experiments with GPT-4. arXiv, Ithaca, NY: Cornell University.

Drugova, E., Zhuravleva, I., Zakharova, U., and Latipov, A. (2024). Learning analytics driven improvements in learning design in higher education: a systematic literature review. *J. Comput. Assist. Learn.* 40, 510–524. doi: 10.1111/jcal.12894

Fadel, C., Holmes, W., and Bialik, M. (2019). Artificial intelligence in education: promises and implications for teaching and learning. Boston, MA, USA: Independently published.

Fahd, K., Venkatraman, S., Miah, S. J., and Ahmed, K. (2022). Application of machine learning in higher education to assess student academic performance, at-risk, and attrition: a meta-analysis of literature. *Educ. Inf. Technol.* 27, 3743–3775. doi: 10.1007/s10639-021-10741-7

Freitas, F. A. D. S., Vasconcelos, F. F. X., Peixoto, S. A., Hassan, M. M., Dewan, M. A. A., Albuquerque, V. H. C. D., et al. (2020). IoT system for school dropout prediction using machine learning techniques based on socioeconomic data. *Electronics* 9:1613. doi: 10.3390/electronics9101613

García-Machado, J. J., Figueroa-García, E. C., and Jachowicz, A. (2020). "Sustainable consumption behaviour in Poland through a PLS-SEM model" in Perspectives on consumer behaviour: Theoretical aspects and practical applications. ed. W. Sroka (Cham: Springer International Publishing), 147–172.

Goodfellow, I., Bengio, Y., and Courville, A. (2016). Deep learning. Cambridge, Massachusetts: The MIT Press.

Hare, R., Tang, Y., and Ferguson, S. (2024). An intelligent serious game for digital logic education to enhance student learning. *IEEE Trans. Educ.* 67, 387–394. doi: 10.1109/TE.2024.3359001

Hu, J., Huang, Z., Li, J., Xu, L., and Zou, Y. (2024). Real-time classroom behavior analysis for enhanced engineering education: an AI-assisted approach. *Int. J. Comput. Intell. Syst.* 17:167. doi: 10.1007/s44196-024-00572-y

Iatrellis, O., Savvas, I. K., Fitsilis, P., and Gerogiannis, V. C. (2021). A two-phase machine learning approach for predicting student outcomes. *Educ. Inf. Technol.* 26, 69–88. doi: 10.1007/s10639-020-10260-x

Jamjoom, M., Alabdulkreem, E., Hadjouni, M., Karim, F., and Qarh, M. (2021). Early prediction for at-risk students in an introductory programming course based on student self-efficacy. *Informatica* 45, 1–9. doi: 10.31449/inf.v45i6.3528

Jayachandran, S., and Joshi, B. (2024). Customized support vector machine for predicting the employability of students pursuing engineering. *Int. J. Inf. Technol.* 16, 3193–3204. doi: 10.1007/s41870-024-01818-w

Köhler, J., Hidalgo, L., and Jara, J. L. (2023). Predicting students' outcome in an introductory programming course: leveraging the student background. *Appl. Sci.* 13:11994. doi: 10.3390/app132111994

Kumar, K. K., and Kavitha, K. S. (2021). Novel ANN based regression and improved lion optimization algorithm for efficient prediction of student performance. *Int. J. Syst. Assur. Eng. Manag.* 1–15. doi: 10.1007/s13198-021-01259-9

Liang, W., Zhang, Y., Cao, H., Wang, B., Ding, D., Yang, X., et al. (2023). Can large language models provide useful feedback on research papers? A large-scale empirical analysis. *arXiv*. doi: 10.48550/arXiv.2310.01783

Lottering, R., Hans, R., and Lall, M. (2020). A machine learning approach to identifying students at risk of dropout: a case study. *Int. J. Adv. Comput. Sci. Appl.* 11, 417–422. doi: 10.14569/IJACSA.2020.0111052

Manjushree, D. L., Varsha, T. L., Arvind, W. K., and Laxman, D. N. (2021). Performance analysis of the impact of technical skills on employability. *Int. J. Perform. Eng.* 17:371. doi: 10.23940/ijpe.21.04.p5.371378

Márquez, L., Henríquez, V., Chevreux, H., Scheihing, E., and Guerra, J. (2024). Adoption of learning analytics in higher education institutions: a systematic literature review. *Br. J. Educ. Technol.* 55, 439–459. doi: 10.1111/bjet.13385

Means, B., Toyama, Y., Murphy, R., Bakia, M., and Jones, K. (2010). Evaluation of evidence-based practices in online learning: A Meta-analysis and review of online learning studies. Washington, D.C., USA: U.S. Department of Education Office of Planning, Evaluation, and Policy Development Policy and Program Studies Service, 115.

Microsoft Research AI4Science, Microsoft Azure Quantum (2023). The Impact of Large Language Models on Scientific Discovery: a Preliminary Study using GPT-4. *arXiv*. doi: 10.48550/arXiv.2311.07361

Musaddiq, M. H., Sarfraz, M. S., Shafi, N., Maqsood, R., Azam, A., and Ahmad, M. (2022). Predicting the impact of academic key factors and spatial behaviors on students' performance. *Appl. Sci.* 12:10112. doi: 10.3390/app121910112

Nabil, A., Seyam, M., and Abou-Elfetouh, A. (2021). Prediction of students' academic performance based on courses' grades using deep neural networks. *IEEE Access* 9, 140731–140746. doi: 10.1109/ACCESS.2021.3119596

Nabizadeh, A. H., Goncalves, D., Gama, S., and Jorge, J. (2022). Early prediction of students' final grades in a gamified course. *IEEE Trans. Learn. Technol.* 15, 311–325. doi: 10.1109/TLT.2022.3170494

Naseer, M., Zhang, W., and Zhu, W. (2020a). Early prediction of a team performance in the initial assessment phases of a software project for sustainable software engineering education. *Sustainability* 12:4663. doi: 10.3390/su12114663

Naseer, M., Zhang, W., and Zhu, W. (2020b). Prediction of coding intricacy in a software engineering team through machine learning to ensure cooperative learning and sustainable education. *Sustainability* 12:8986. doi: 10.3390/su12218986

Nieto, Y., Gacia-Diaz, V., Montenegro, C., Gonzalez, C. C., and Gonzalez Crespo, R. (2019). Usage of machine learning for strategic decision making at higher educational institutions. *IEEE Access* 7, 75007–75017. doi: 10.1109/ACCESS.2019.2919343

Opazo, D., Moreno, S., Álvarez-Miranda, E., and Pereira, J. (2021). Analysis of first-year university student dropout through machine learning models: a comparison between universities. *Mathematics* 9:2599. doi: 10.3390/math9202599

OpenAI (2023). GPT-4 technical report. OpenAI. Available online at: https://cdn.openai.com/papers/gpt-4.pdf (Accessed March, 15, 2025).

Oro, J. M. F., Regodeseves, P. G., Bertolín, L. S., Pérez, J. G., Barrio-Perotti, R., and Blanco, A. P. (2024). Application of learning analytics for the study of the virtual campus activity in an undergraduate fluid mechanics' course in bachelors of engineering. *Technol. Knowl. Learn.* 30, 1321–1344. doi: 10.1007/s10758-024-09803-9

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The Prisma 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 372:n71. doi: 10.1136/bmj.n71

Papamitsiou, Z., and Economides, A. (2014). Learning analytics and educational data Mining in Practice: a systematic literature review of empirical evidence. *Educ. Technol. Soc.* 17, 49–64.

Poudyal, S., Mohammadi-Aragh, M. J., and Ball, J. E. (2022). Prediction of student academic performance using a hybrid 2D CNN model. *Electronics* 11:1005. doi: 10.3390/electronics11071005

Ren, Y., and Yu, X. (2024). Long-term student performance prediction using learning ability self-adaptive algorithm. *Complex Intell. Syst.* 10, 6379–6408. doi: 10.1007/s40747-024-01476-2

Rico-Juan, J. R., Cachero, C., and Macià, H. (2024). Study regarding the influence of a student's personality and an LMS usage profile on learning performance using machine learning techniques. *Appl. Intell.* 54, 6175–6197. doi: 10.1007/s10489-024-05483-1

Salloum, S. A., Basiouni, A., Alfaisal, R., Salloum, A., and Shaalan, K. (2024). "Predicting student retention in higher education using machine learning., in *breaking barriers with generative intelligence*" in Using GI to improve human education and well-being. eds. A. Basiouni and C. Frasson (Cham: Springer Nature Switzerland), 197–206.

Slade, S., and Prinsloo, P. (2013). Learning analytics: ethical issues and dilemmas. *Am. Behav. Sci.* 57, 1510–1529. doi: 10.1177/0002764213479366

Sultana, J., Rani, M. U., and Farquad, M. A. H. (2019). Student's Performance Prediction using Deep Learning and Data Mining Methods, 8.

Tricco, A. C., Langlois, E. V., and Straus, S. E. (2017). *Rapid reviews to strengthen health policy and systems: a practical guide*. World Health Organization Geneva. Available online at: https://ahpsr.who.int/docs/librariesprovider11/publications/supplementary-material/alliancehpsr_rrguide_trainingslides.pdf (Accessed April 19, 2025).

Tsai, Y.-S., Poquet, O., Gasevic, D., Dawson, S., and Pardo, A. (2019). Complexity leadership in learning analytics: drivers, challenges and opportunities. *Br. J. Educ. Technol.* 50, 2839–2854. doi: 10.1111/bjet.12846

Willis, J. E., Slade, S., and Prinsloo, P. (2016). Ethical oversight of student data in learning analytics: a typology derived from a cross-continental, cross-institutional perspective. *Educ. Technol. Res. Dev.* 64, 881–901. doi: 10.1007/s11423-016-9463-4

Woolf, B. P. (2008). Building intelligent interactive tutors: Student-centered strategies for revolutionizing E-learning. *1st* Edn. Amsterdam, Boston: Morgan Kaufmann.

Zhang, Z., Lu, J., Pei, B., and Huang, Y. (2024). Advancing engineering and computing education through the Lens of learning analytics., in 2024 IEEE Frontiers in education conference (FIE), (IEEE), 1–8. Available online at: https://ieeexplore.ieee.org/abstract/document/10893384/ (Accessed April 19, 2025).