

Exploring Cultural Identity: Perspectives from Social Sciences and Humanities

Exploring the role of machine learning in forecasting student performance in education: An in-depth review of literature

Bavani Raja Pandian^a✉ | Azlinda Abdul Aziz^b | Hema Subramaniam^c | Haslinda Sutan Ahmad Nawi^d^aTunku Abdul Rahman University of Management and Technology (TAR UMT), Jalan Genting Kelang, Setapak, 53300 Kuala Lumpur, Malaysia.^bUniversiti Selangor, Jalan Timur Tambahan, 45600 Batang Berjuntai, Selangor, Malaysia.^cUniversiti Malaya, 50603 Kuala Lumpur, Wilayah Persekutuan Kuala Lumpur, Malaysia.^dInformation Science & Computing Department, Faculty of Information Sciences and Engineering, Management and Science University, Malaysia.

Abstract The literature on the role of machine learning in forecasting student performance in education lacks comprehensive data quality, equitable and interpretable models, consideration of contextual and causal factors, and the integration of human expertise. This review will explore machine learning types, algorithms, predictive performance, and impact on student performance in education. A systematic literature review based on articles published in the past 2019–2023 period. Methods: The IEEE Xplore database was searched by using keywords such as "Educational Data Mining," "Student Performance Prediction," "Evaluations of Students," "Performance Analysis of Students," and "Learning Curve Prediction." were employed and 50 papers were selected. Results: The analysis of the results highlighted prominent patterns. Half of the studies favored supervised learning methods, with decision trees leading (19 instances), followed by Long Short-Term Memory and Random Forest (16 each), and K-Nearest Neighbor and Naive Bayes (12 and 11 times). Support Vector Machine and Logistic Regression were noted 10 and 9 times, respectively. Noteworthy were ANN, CNN, and Xgboost. Positive impacts were evident in 36 cases; only one showed negative effects, while 13 indicated intricate relationships. This study helped in understanding the prevalent machine learning methods used for predicting student performance, provides a benchmark for assessing the effectiveness of new or alternative techniques. Conclusion: This review highlights the varied machine learning uses in predicting student performance in education, emphasizing supervised methods, diverse algorithms, and complex intervention impacts.

Keywords: EduData Mining, predictive analytics, student performance, ML in education, learning analytics, academic forecasting

1. Introduction

The convergence of data science and pedagogy in the digital age is profoundly changing the landscape of education. As traditional classrooms evolve into dynamic virtual environments and educational interactions become increasingly digitized, the field of education finds itself at the nexus of innovation and analytics. The incorporation of machine learning, a formidable branch of artificial intelligence that has the potential to revolutionize how we comprehend, forecast, and improve student performance, is at the core of this paradigm shift. There are many ways that machine learning is being applied in education, and these applications are revolutionizing the subject by giving the ability to completely change how we perceive, anticipate, and improve student performance. Educators use machine learning to identify children who are experiencing difficulty and to take steps to enhance students' academic performance and likelihood of staying in school (Sanusi et al., 2023).

Personalized learning techniques enabled by machine learning give instructors the ability to adjust learning routes to the specific needs of individual pupils. Improving the whole campus experience: The campus experience is being improved by institutions through machine learning, and self-service capabilities are also being enabled (Pinto et al., 2023). Machine learning is assisting in expanding the reach and impact of online educational information in various ways, including localization, transcription, text-to-speech conversion, and personalization. Additionally, it has the ability to swiftly and effectively detect the obstacles faced by learners, forecast the consequences of future learning, and tailor instruction (Hilbert et al., 2021). To assess and boost student happiness, machine learning might be used to determine which interventions and support should be made available to various registered student archetypes. In general, machine learning is changing how educational institutions



monitor student performance and identify problems. It is anticipated that over the next few years, machine learning will play a greater role in the education sector (Webb et al., 2021).

The expense that may be incurred to bring machine learning technology into the classroom may vary. To adequately support deployment, considerable financial investment is necessary in terms of both hardware and software (Ara Shaikh et al., 2022). The implementation of new technologies, such as machine learning, can lead to resistance from educators, administrators, and other stakeholders. Overcoming resistance to change and fostering a culture of innovation and collaboration are essential for successful implementation (Haleem et al., 2022). Due to a general lack of professional expertise, educational institutions may face challenges in adopting machine learning. Implementing machine learning requires skilled data scientists and educators who can effectively utilize the technology. To provide reliable forecasts, algorithms for machine learning require access to vast volumes of data. However, educational institutions may face challenges in accessing and collecting relevant data. In addition, the quality of the data as well as its consistency might influence the efficiency of machine learning models (Schiff, 2021).

It also involves handling sensitive student data. Ensuring the privacy and security of these data is crucial. Educational institutions need to have robust data protection measures in place to safeguard student information. It can inadvertently perpetuate biases and inequalities present in the education system. It is important to address ethical considerations and ensure that machine learning models are fair, transparent, and unbiased (Kim et al., 2022). Despite these obstacles, there is substantial room for improvement in educational outcomes that might be achieved with the application of machine learning. By addressing these challenges and leveraging the opportunities provided by machine learning, educational institutions can enhance teaching and learning processes, personalize education, and improve student outcomes.

It is vital to evaluate the predictive performance of machine learning algorithms to determine the practical value of these tools. These models are used to anticipate the outcomes of students. This systematic review seeks to provide thorough knowledge of the advantages and disadvantages of various models by evaluating their accuracy, precision, recall, and other pertinent metrics. Furthermore, examining the impact of implementing predictive models on educational practices and student outcomes will provide insights into whether these technologies contribute positively to educational outcomes, retention rates, and personalized learning experiences.

In conclusion, by synthesizing and assessing the literature on machine learning applications for forecasting student performance in education, this systematic review aims to add to the existing body of knowledge. This review aims to prove the potential benefits of integrating machine learning into educational practices, inform educators, researchers, and policymakers about the state of the field, and guide the selection of appropriate algorithms and data sources by achieving the outlined objectives.

1.1. Research Gap Identification

Several systematic literature reviews on machine learning and predicting student performance have been carried out. The evaluations revealed a number of research gaps in the currently available literature, including the following:

1. There is a lack of standardization in the features used to predict student performance (Issah et al., 2023).
2. Predicting student performance using deep learning methods is limited (Sekeroglu et al., 2021).
3. There have been few studies on how data preparation methods affect prediction models for student performance (Albreiki et al., 2021).
4. The ethical issues of utilizing machine learning to predict student achievement have received little attention (Alalawi et al., 2023).

These research gaps suggest that there is a need for more standardized approaches to feature selection and data preprocessing in machine learning models for predicting student performance. In addition, there is a need for more studies on the ethical repercussions that may arise from the use of machine learning in educational settings. Finally, there is a need for more studies on the use of deep learning strategies in the process of forecasting student success.

The comprehensive literature on the use of machine learning in forecasting student success in education aims to achieve the following objectives: explore various types of machine learning, delve into key machine learning algorithms, analyze machine learning evaluation metrics, and assess predictive performance along with its impact.

There are a number of gaps and difficulties in the existing research on applying machine learning to predict student performance. A complete knowledge of the advantages and disadvantages of these methods is hampered by the paucity of related research. The formation of best practices and outcome comparisons are hampered by the lack of standardized methodologies. Education data quality problems make it difficult to construct reliable models, and ethical issues related to privacy and prejudice must be carefully addressed. Despite encouraging research, closing these gaps and overcoming these obstacles is essential if machine learning in education forecasting is to reach its full potential.

The objective of this study is to conduct an exhaustive review of previous research in the form of studies, articles, and papers that examine the application of machine learning algorithms to the task of predicting the academic performance of students in various educational environments. The following components are included in the scope of this project; however, this list is not exhaustive:

1. Examining various machine learning techniques (such as decision trees, neural networks, support vector machines, etc.) used to forecast student performance, along with their relative merits and drawbacks
2. Investigation of the features and variables used as inputs for machine learning models, including academic, demographic, and socioeconomic attributes, as well as potential feature engineering techniques to enhance prediction accuracy
3. The data sources used in these studies included academic records, attendance records, demographic information, student behavior data, etc.
4. Analysis of the assessment metrics and procedures used to rate the performance of the generated prediction models in terms of accuracy, precision, recall, F1-score, and other criteria
5. Identification of factors that have been identified as significantly impacting student performance prediction, such as course difficulty, engagement, study habits, and socioeconomic background
6. Exploration of how machine learning models vary in effectiveness across different educational levels (primary, secondary, and tertiary) and subject domains (mathematics, sciences, humanities, etc.)
7. Discussion of privacy, prejudice, fairness, and transparency as ethical issues when using machine learning algorithms to forecast student performance
8. Examination of how predictive insights are generated by machine learning can inform educational strategies, interventions, and policies to enhance student outcomes.

2. Research Methodology

Recent literature assessments undertaken by eminent standard publishers, including IEEE Xplore, have emphasized the dynamic research environment from 2019 to 2023 in the field of educational data mining and student performance prediction. These reviews delve into various methodologies and techniques employed for enhancing the accuracy and reliability of predictions regarding student performance in education. Noteworthy keywords such as "Educational Data Mining," "Student Performance Prediction," "Evaluations of Students," "Performance Analysis of Students," and "Learning Curve Prediction" have been central to the studies explored. These reviews reveal a burgeoning interest in the utilization of advanced machine learning algorithms, including but not limited to neural networks, decision trees, and ensemble methods, to uncover insightful patterns from educational data. Moreover, a shift toward incorporating diverse data sources, such as socioeconomic factors, behavioral attributes, and learning styles, has been observed, contributing to more holistic predictive models. This review underscores the potential of these predictive models to facilitate early interventions and personalized learning strategies, fostering improved educational outcomes and instructional efficacy.

2.1. Search Strategy

Utilizing the IEEE Xplore database to search for studies published between 2019 and 2023, inclusive, helped to achieve the review objectives. The databases were searched, and articles were selected using keywords such as "Educational Data Mining," "Student Performance Prediction," "Evaluations of Students," "Performance Analysis of Students," and "Learning Curve Prediction." The total number of publications identified through the literature search was 25,470. Records were excluded after applying filters for the years 2019–2023 ($n = 16,527$), open access ($n = 8,202$), early access articles ($n = 26$), and conferences ($n = 6$). The number of records assessed for eligibility was 709. After screening, highly relevant studies were identified ($n = 50$) (Adnan et al., 2021, 2022; Alamri & Alharbi, 2021; Alhazmi & Sheneamer, 2023; Aljaloud et al., 2022; Alshanqiti & Namoun, 2020; Asthana et al., 2023; Bujang et al., 2021; Chan et al., 2023; Chui et al., 2020; Feng et al., 2022; Gao et al., 2019; Ghorbani & Ghousi, 2020; Hassan et al., 2022; Hussain et al., 2021; Kastrati et al., 2020; Kusumawardani & Alfarozi, 2023; Latif et al., 2023; Liu et al., 2020; Mengash, 2020; Motz et al., 2021; Munshi & Alhindi, 2021; Nabil et al., 2021; Nguyen-Huy et al., 2022; Okoye et al., 2023; Orlando et al., 2020; Pek et al., 2023; Prabowo et al., 2021; Rafique et al., 2021; Rahman et al., 2021).

2.2. Inclusion/Exclusion Criteria

The literature review focused on studies published between 2019 and 2023, written in English, and primarily concerned the application of machine learning (ML) in the context of student academic performance. The review excluded studies published outside this timeframe, those not written in English, nonprimary studies, and works that did not pertain to the specified academic themes. Additionally, the scope encompassed journal articles and fully accessible and comprehensive full-text materials, incomplete or inaccessible sources were excluded, and any duplicates were eliminated.

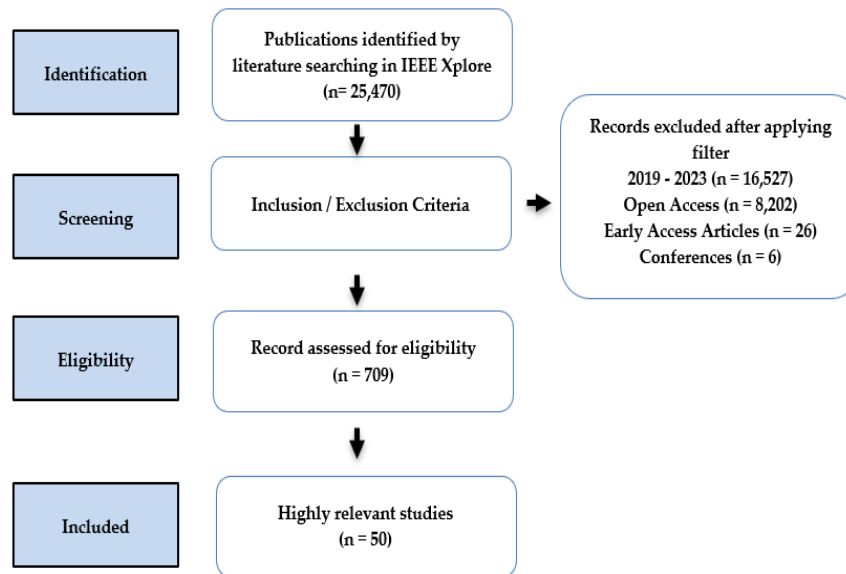


Figure 1 PRISMA diagram.

2.3. Quality Assessment

In our systematic literature review, we incorporated quality assessment as a vital criterion to ascertain whether the chosen articles adhered to specific quality standards. These criteria were used to evaluate the appropriateness of each selected article for inclusion in our review. For each quality assessment, respondents had three possible choices: "Yes" if the paper fully met the quality criteria, "Partly" if it met them partially, and "No" if it did not meet them at all. Table 1 outlines the five quality assessment criteria that were employed in this systematic literature review.

Table 1 Quality Assessment of Selected Studies.

CRITERIA	YES	NO	PARTLY
Is the Research Question Clearly, Stated?	49	0	1
Is the Methodology Appropriate and Transparent?	40	0	10
Is the Sample or Data Collection Method Adequate?	45	0	5
Are Valid and Reliable Measures Used?	40	0	10
Are Limitations and Confounding Factors Addressed?	39	0	11

2.4. Block diagram of the framework

As illustrated in Figure 2, there are many interconnected stages in the block diagram of the framework for machine learning applications to predict student progress in education. Data collection began with student demographics, academic records, and socioeconomic indicators. Missing values, outliers, and feature engineering are preprocessed to improve the data. After the data have been analyzed, they are segmented into training, validation, and testing sets so that the model can be trained, adjusted, and evaluated. Machine learning algorithms, from regression and classification to neural networks and ensemble models, form the foundation. Iterative training and optimization utilizing training data fine-tune these algorithms' hyperparameters through cross-validation. The trained models are tested on the validation set to make modifications. The optimum model is evaluated on the unseen testing set to replicate real-world prediction skills. This comprehensive approach provides insights into student performance to help instructors make educated decisions to improve learning outcomes.

3. Main finding of existing papers

3.1. Top 10 Most Cited Articles

Table 2 shows the top 10 influential studies across multiple years. The top of the list is Wei et al. (2020), with 90 citations, followed by Mengash et al. (2020) and Ghorbani et al. (2020), with 81 and 74 citations, respectively. These studies likely contribute to a significant field, as demonstrated by their citation counts. Kastrati et al. (2020) and Sindhu et al. (2019) also feature notable citations. The list showcases a mix of research spanning the years 2019 to 2021, indicating ongoing relevance. Noteworthy findings might emerge from Adnan et al. (2021), Chui et al. (2020), and others. Bujang et al. (2021) and Alshantqi et al. (2020) removed the top 10 with substantial citations, signifying their impact.



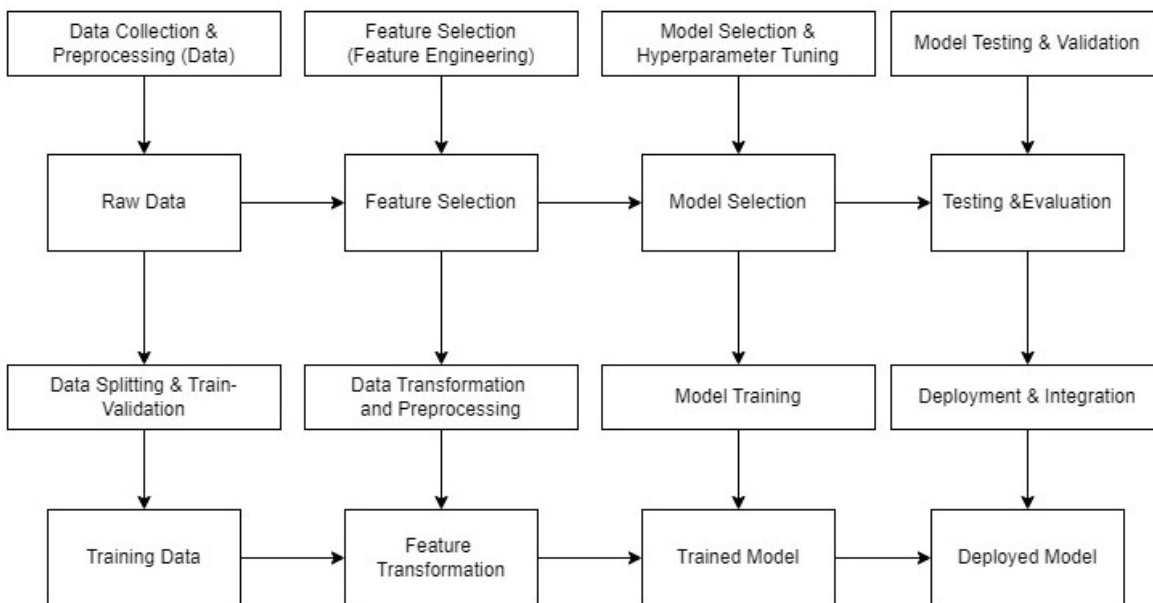


Figure 2 Machine Learning Framework for Educational Prediction of Student Performance.

Table 2 Top 10 Most Cited Articles.

RANK	STUDY	PUBLICATION YEAR	CITATION COUNT
1	(Wei et al., 2020)	2020	90
2	(Mengash, 2020)	2020	81
3	(Ghorbani & Ghousi, 2020)	2020	74
4	(Kastrati et al., 2020)	2020	66
5	(Sindhu et al., 2019)	2019	59
6	(Adnan et al., 2021)	2021	49
7	(Chui et al., 2020)	2020	46
8	(Samin & Azim, 2019)	2019	35
9	(Bujang et al., 2021)	2021	34
10	(Alshanjiti & Namoun, 2020)	2020	32

3.2. Trends of Publication and Citations of Studies

The publication output displayed a gradual trend over the years, with a peak in 2020 (198 publications) and a subsequent decline. Citations followed a similar pattern, with the highest occurring in 2020 (3478 citations). However, as shown in Figure 3, both publications (105) and citations (39) significantly decreased by 2023 in comparison to earlier years.

3.3. Top Keyword Occurrences

The results of this SLR showed the ranking of keywords by their occurrences in a dataset. "Humans" tops with 20 mentions, followed by "machine learning" (9) and "algorithms" (8). The pandemic terms "COVID-19" and "deep learning" each have 6 mentions. Technical terms such as "EEG methods" and "EHRs" appear 4 times. Specific AI concepts are also present, such as CNNs (3). The gender categories "female" and "male" also appeared three times, as shown in Table 3.

3.4. Distribution of Studies by Dataset Size

The distribution of studies based on dataset size is as follows: 8 studies employed datasets with fewer than 100 instances; 12 studies used datasets containing 101 to 500 instances; and 8 studies utilized datasets ranging from 501 to 1000 instances. Additionally, 11 studies employed datasets sized between 1001 and 5000 instances, 2 studies used datasets containing 5001 to 10,000 instances, and 9 studies worked with datasets exceeding 10,000 instances, as shown in Figure 4.



The distribution of studies based on dataset size can impact the accuracy and generalizability of student performance prediction models. Studies that use datasets with fewer than 100 instances may have limited data to train and test their models, which can result in overfitting and poor generalizability to new data. Studies that use datasets containing 101 to 500 and 501 to 1000 instances may have a better balance between dataset size and complexity, which can lead to more accurate and generalizable models. Studies that use datasets exceeding 10,000 instances may use larger and more diverse datasets to train and test their models, which can lead to more accurate and generalizable models. The variability in dataset size across studies can make it difficult to compare and generalize results across studies. Researchers should carefully consider the dataset size and its potential impact on their research question and modeling approach.

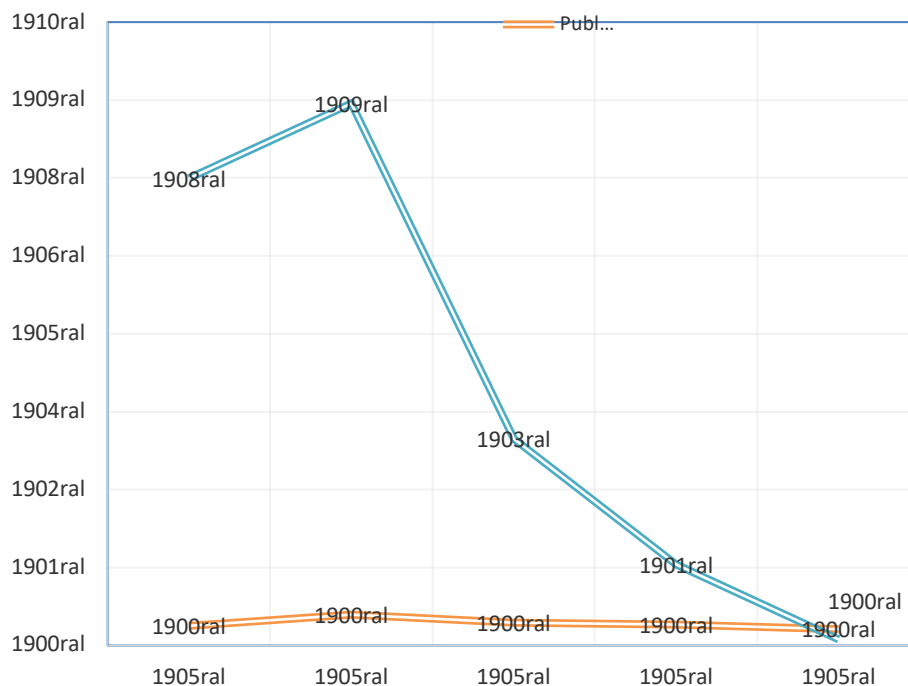


Figure 3 Publication with citations of studies by year.

Table 3 Occurrences of the Top 10 Keywords.

RANK	KEYWORD	OCCURRENCES	TLS
1	humans	20	114
2	machine learning	9	54
3	algorithms	8	57
4	covid-19	6	32
5	deep learning	6	32
6	electroencephalography/methods	4	16
7	electronic health records	4	25
8	convolutional neural network	3	16
9	Female	3	32
10	Male	3	32

4. Analysis of the limitations of the Methodology of existing papers and algorithms

4.1. Limitations of the Methodology of Existing Papers

Figure 5 presents the methodological limitations and their respective percentage distributions. The largest portion of the included studies in this systematic review, 41%, pertains to limited data and its applicability to broader contexts. Focus and scope limitations followed at 14%. Assumption and approach constraints contribute 11%, as do data quality and preprocessing issues. Algorithm and model restrictions account for 9%, while methodology itself represents 8%. Finally, the evaluation metric limitations are 7%.



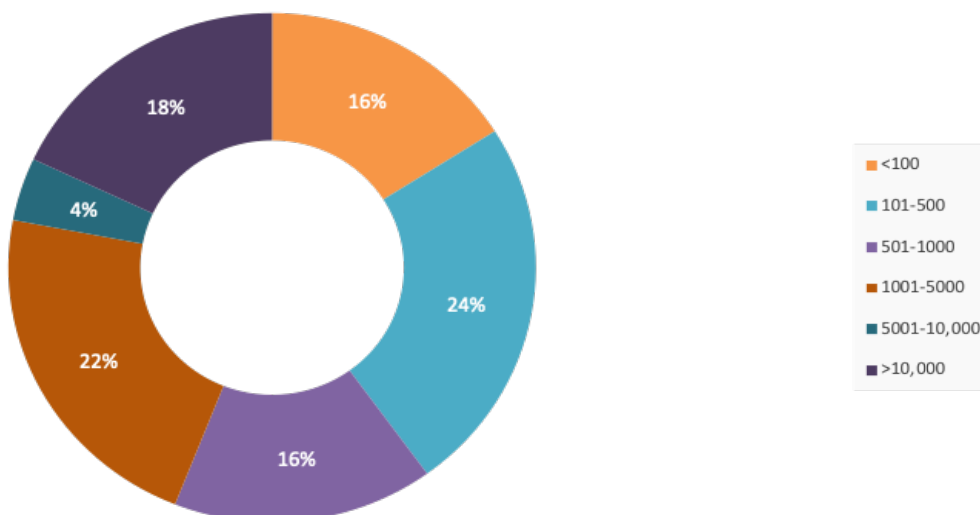


Figure 4 Distribution of Studies Based on the Size of the Selected Studies' Datasets.

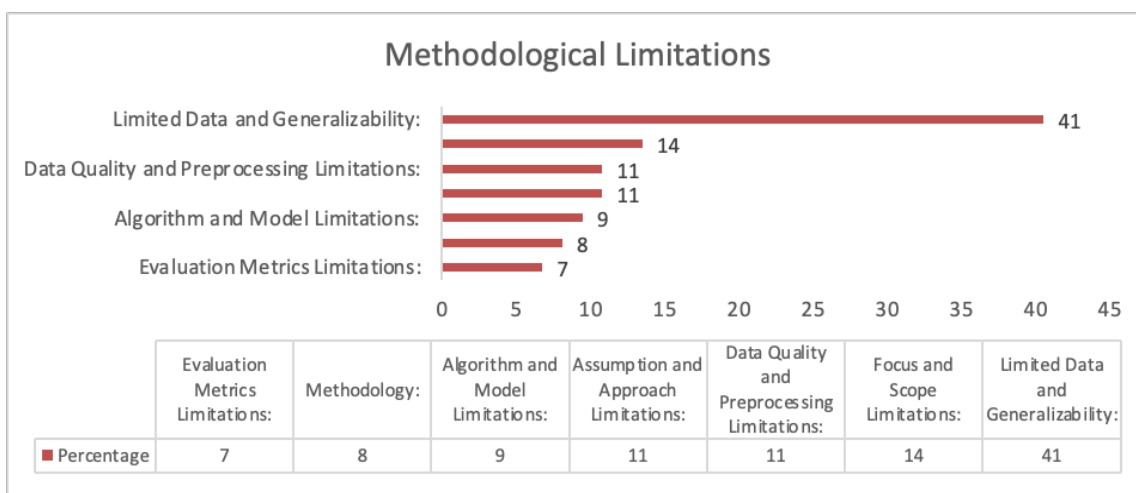


Figure 5 Analysis of the Limitations of the Methodology of Existing Papers.

4.2. Limitations of the Existing Published Algorithms

Figure 6, analyzing the limitations of existing algorithm papers, reveals a multifaceted landscape. Approximately 9% of concerns revolve around architecture and complexity, emphasizing scalability. Data-related issues, constituting 13%, highlight data quality and quantity challenges. Model performance (10%) raises questions about efficacy, while interpretability (6%) and other factors (5%) collectively underscore the need for comprehensive algorithm assessment and development.

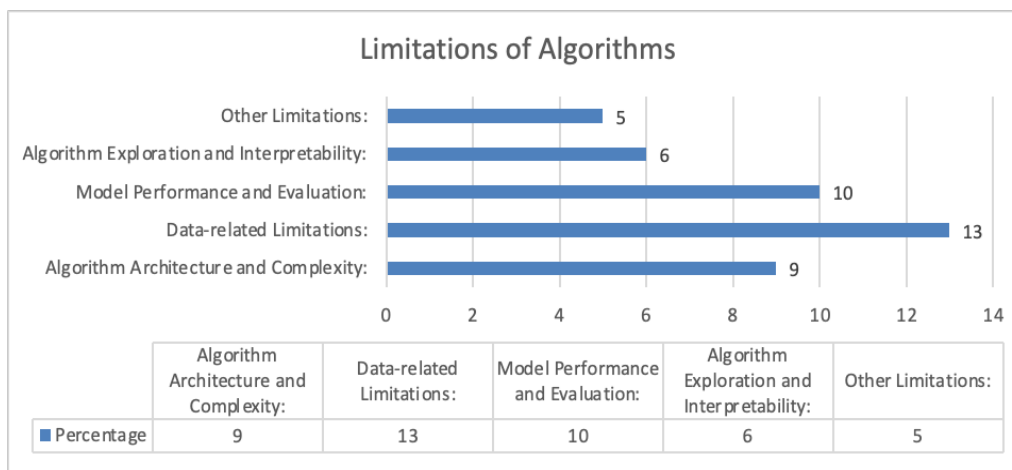


Figure 6 Analysis of the Limitations of the Existing Papers Algorithms.



4.3. Advantages and Disadvantages of Existing Research with Desired Outputs

This systematic review revealed multiple advantages and disadvantages of educational technology in enhancing the academic performance of students in educational institutes.

Advantages:

1. *Efficient Feedback and Personalization*
 - Efficient qualitative feedback processing.
 - Dynamic modeling of student effort and ability.
 - Efficient course allocation and recommendations.
 - Real-time engagement analysis and feedback.
 - Informed admissions decisions.
2. *Accurate Performance Prediction:*
 - Accurate student performance prediction.
 - Improved prediction through a hybrid model.
 - GPA prediction with a dual-input model.
 - Effective entrepreneurial intention prediction.
 - High GPA prediction and model efficiency.
3. *Insights for Education Enhancement:*
 - Insights into teaching effectiveness and attitudes.
 - Study on imbalanced dataset effects.
 - Factors impacting student success.
 - Educator insights and student support.
 - Insights into emotions, nuanced representation.
4. *Early Identification and Support:*
 - Identifying at-risk kids early.
 - A predictive model for pupils who are in danger.
 - Early intervention and personalized support.
 - Reduced missed assignments; immediate notifications.
 - Early learning effect prediction has high accuracy.
5. *Data-Driven Analysis and Insights*
 - Interformable study algorithms.
 - Streamlined data collection and program evaluation.
 - Improved information literacy and interactive learning.
 - Feature selection and early interventions.
 - Early prediction requires a data-driven approach.
6. *Enhanced Visualization and Decision Making:*
 - Informative dashboard for teachers.
 - Real-time feedback and visualized insights.
 - Structured pricing framework and flexibility.
7. *Performance enhancement and efficiency:*
 - Insights into evolving programming skills.
 - Comprehensive understanding of 3D tool use.
 - Insights into gamification effects.
 - Improved accuracy and time-series analysis.
 - Improved accuracy and insights into factors.
 - High model accuracy and IoT-based insights.

Disadvantages:

1. *Model Complexity and Implementation*
 - Complex model implementation.
 - Technical complexity is resource intensive.
 - A longer runtime and greater clustering effectiveness.
2. *Prediction Inaccuracies*
 - Potential inaccuracies in predictions.

- Inaccurate recommendations.
 - The model accuracy is limited.
 - Inaccurate predictions and student pressure.
3. *Bias and Interpretation:*
 - Mismatched recommendations and biases.
 - Overlooking noncognitive skills and biases.
 - Bias in expert evaluations.
 - Simplification of complex relationships.
 - Bias in behavior measurement.
 4. *Data and Scope Limitations:*
 - Limited dataset scope.
 - Exclusion of external factors.
 - Limited by data quality and implementation information.
 - Model limitations and narrow scope.
 - Oversimplification and neglect of other factors.
 5. *Privacy and ethical concerns:*
 - Privacy concerns and potential negative impact.
 - Privacy concerns, stigmatization.
 - Ethics, privacy concerns, stigmatization.
 6. *Effectiveness and Applicability*
 - Possible misinterpretation of engagement signals.
 - Resistance to change.
 - The generalizability and short-term focus of this study are limited.
 - Limited reach, muted notifications.
 - There is limited applicability and factor coverage.
 - Model limitations and applicability concerns.
 - The accuracy and individual differences are limited.
 - There is a lack of studies that address accuracy and explainability.
 - Customization is needed for data quality.
 - Evolving dynamics, oversimplification.
 - There is a lack of depth and contextual limitations.
 7. *Resource and Training Requirements:*
 - Resource and training requirements.
 - Technological infrastructure, resistance.
 - Variable accuracy and external factors.

Overall, this systematic review presents a balanced view of the potential benefits and challenges of educational technology. While it has the potential to enhance teaching and learning, careful interpretation, customization, and ethical considerations are necessary to ensure its effective use.

5. Result analysis with existing research papers

5.1. Types of machine learning

Figure 5 shows the findings of this comprehensive literature analysis on machine learning applications for forecasting student performance in education, which revealed that supervised learning techniques dominated the field and accounted for 50% of the research examined. Classification, a subset of supervised learning, was explored in 5% of the studies. Deep learning, known for its advanced neural network techniques, was employed in 2% of cases. Data mining and logistic regression each had a 3% representation, while clustering was used in 2% of the studies. A noteworthy 36% of the research fell under the category of "other" methods, suggesting a diverse array of machine learning approaches for predicting student performance.

5.2. Types of machine learning algorithms

Figure 6 illustrates the widespread use of a variety of machine learning algorithms in the area of education for the purpose of forecasting student performance. Decision trees emerged as the most frequently utilized algorithm (19 occurrences), followed closely by long short-term memory and random forest (both with 16 occurrences). K-nearest neighbor and naive Bayes also demonstrated substantial presence, appearing in 12 and 11 instances, respectively. The study also highlights the significance of support vector machine and logistic regression, which are observed in 10 and 9 instances,

respectively. Additionally, artificial neural networks, convolutional neural networks, and Xgboost exhibit notable representation, underscoring their relevance in predicting student performance.

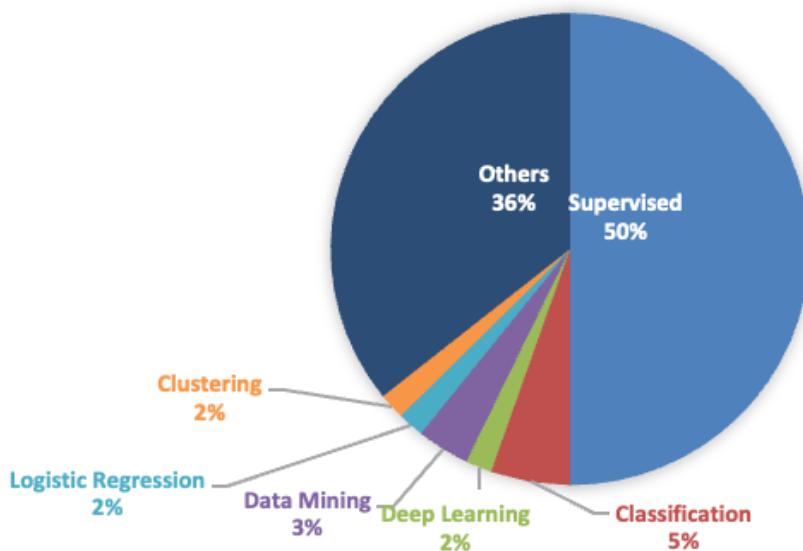


Figure 5 Division of Studies Based on Various Machine Learning Techniques.

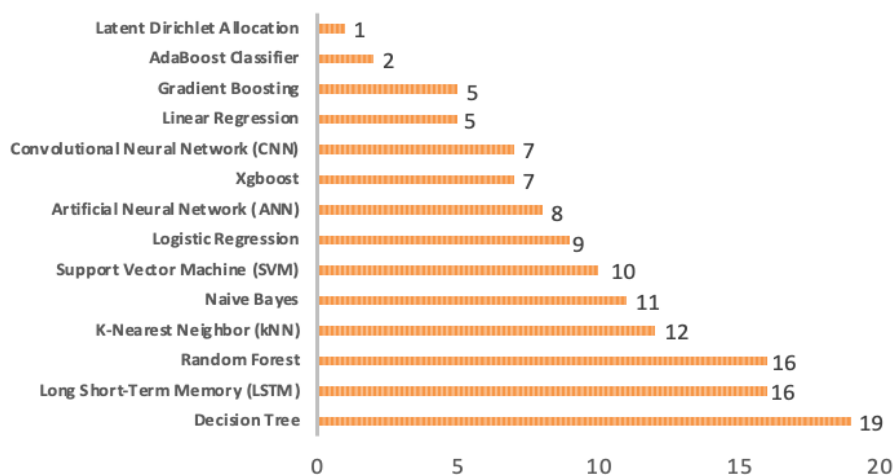


Figure 6 Recurrence of different ML algorithms.

5.3. Machine learning algorithm categorization

Figure 7 shows the machine learning algorithm categories. Machine learning algorithms are categorized into different types based on their underlying tasks and characteristics. These algorithms assign input data to predefined classes or categories. Classification models make up approximately 77% of machine learning algorithms across categories. These algorithms mostly categorize incoming data into specified classifications. Regression algorithms, at 12%, predict continuous numerical values. Clustering algorithms, 10%, aggregate related data elements to reveal patterns. Dimensionality reduction approaches make up just 1% of machine learning, highlighting their relative insignificance. This distribution shows that classification, regression, and clustering are the most commonly used approaches, whereas dimensionality reduction is rare.

5.4. Performance Analysis with Existing Works

5.4.1. Pearson's chi-squared test

Using R Studio software, the chi-square test was conducted to determine whether there was a significant association between the variables' positive impact and negative impact within the data in Table 4. The test resulted in a chi-square statistic of 21.863 with 9 degrees of freedom and a p value of approximately 0.0093. The p value is below the commonly used significance level of 0.05. This suggests that there is evidence of a statistically significant relationship between the positive impact and negative impact variables in the data.



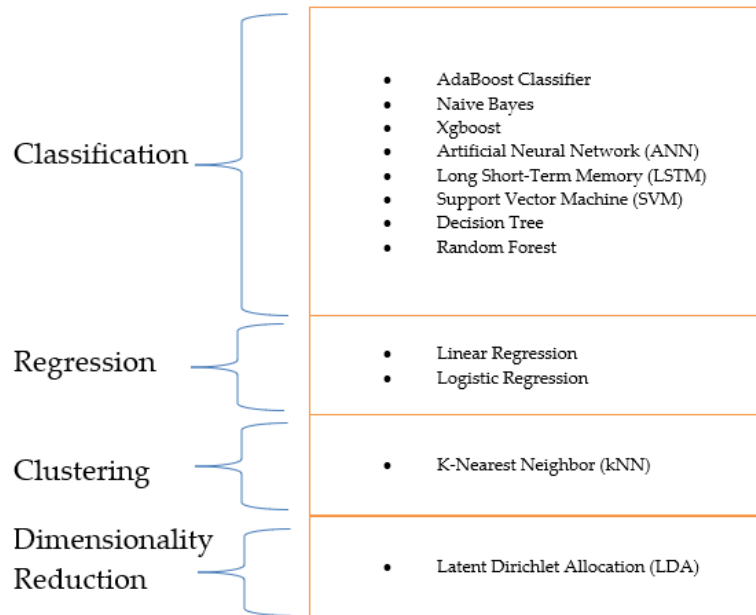


Figure 7 Categorizing different ML algorithms.

Table 4 Performance analysis with ML types in the selected studies.

MACHINE LEARNING TYPE	POSITIVE IMPACT	NEGATIVE IMPACT	TOTAL
Supervised Machine Learning	28	1	29
Probabilistic Topic Modeling	2	1	3
Additive Factor Model (AFM)	2	1	3
Dual-Input Deep Learning Model	2	1	3
Data Mining and Association Rules	1	1	2
Explainable Machine Learning Models	0	1	1
Unsupervised Clustering	0	1	1
Mixed Methodology	0	1	1
Classification	2	2	4
Unspecified	1	2	3

X-squared = 21.863, df = 9, p value = 0.009323

5.4.2. Pairwise Comparison of Proportions

After identifying a significant association between the positive impact and negative impact variables, a pairwise comparison of proportions for each combination of machine learning types using R Studio software was conducted to determine whether there were statistically significant differences in the proportions of positive impacts among the different machine learning types. The results are presented in Table 5, where each cell represents the p value of a pairwise comparison between two machine learning types. The p values were adjusted using the Holm method to control for multiple comparisons.

Overall, the analyses indicate a significant association between positive impact and negative impact and provide insights into how machine learning types compare in terms of their positive impact proportions.

6. Summary of the review

The systematic exploration of the rapidly evolving field stemming from a systematic literature review on machine learning applications for predicting student success in education has yielded a profound comprehension of its landscape. The most influential studies, spanning the years 2019 to 2021, have highlighted an amalgamation of research efforts that continuously contributed to the field's pertinence. Notably, trends in publications and citations showed a pinnacle in 2020, followed by a gradual decrease in 2023.



A meticulous examination of keyword occurrences underscored the significance of terms such as "humans," "machine learning," and "algorithms". Notably, even terms linked to the pandemic, such as "COVID-19" and "deep learning," were prominently featured.

The distribution of studies based on dataset size revealed a spectrum of sample sizes, with larger datasets gaining prominence over time.

Diverse methodological constraints have come to light, spanning issues regarding data applicability, scope, assumptions, data quality, algorithms, and evaluation metrics. Similarly, existing algorithmic studies have encountered concerns related to the scalability of architectures, data quality, model performance, and interpretability.

This review revealed the advantages of educational technology, including efficacious feedback mechanisms, personalized learning, and precise performance prognosis. However, challenges include complexities in model design, inherent biases, and apprehensions surrounding privacy.

The landscape was predominantly influenced by supervised learning techniques, with decision trees, LSTM, and random forest emerging as frequently employed algorithms for predicting student performance.

Using R Studio, an exploration using Pearson's chi-squared test revealed a noteworthy association ($p \approx 0.0093$) between variables signifying positive and negative impacts. Subsequent pairwise comparisons of diverse machine learning types, adjusted using the Holm method, underscored noteworthy differences in proportions of positive impacts, thereby providing valuable insights into their relative efficacy.

Systematic scrutiny presented a comprehensive overview of the field's advancement, delineating the wide array of methodologies, challenges, and benefits that compose it. This thorough review provides invaluable guidance for researchers, educators, and policymakers, who are steering the formulation and application of effective machine learning strategies in the domain of education.

Table 5 Pairwise comparison of proportions.

PAIRWISE COMPARISONS USING PAIRWISE COMPARISON OF PROPORTIONS									
DATA: DATA\$POSITIVEIMPACT OUT OF DATA\$TOTAL									
	1	2	3	4	5	6	7	8	9
2	1.00	-	-	-	-	-	-	-	-
3	1.00	1.00	-	-	-	-	-	-	-
4	1.00	1.00	1.00	-	-	-	-	-	-
5	1.00	1.00	1.00	1.00	-	-	-	-	-
6	1.00	1.00	1.00	1.00	1.00	-	-	-	-
7	1.00	1.00	1.00	1.00	1.00	-	-	-	-
8	1.00	1.00	1.00	1.00	1.00	-	-	-	-
9	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	-
10	0.47	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

7. Final considerations

In conclusion, a thorough study of machine learning applications for forecasting student performance in the area of education was offered in this systematic literature review. This review revealed a dynamic landscape in which influential studies from 2019 to 2021 significantly contributed to the field, as evidenced by their citation counts. The publication trend peaked in 2020, reflecting the heightened interest in the subject, with subsequent years showing a decline. Notably, the review highlighted the dominance of supervised learning methods, with classification, deep learning, data mining, and logistic regression making up specific methodological approaches.

The distribution of studies based on dataset size showcased a diverse range, indicating the diversity of data availability and its implications for predictive modeling. Methodological limitations were identified, with data-related challenges, model performance concerns, and broader applicability issues among the key factors hindering the development and deployment of predictive algorithms.

It was also shown how often machine learning algorithms are used to predict student performance, with decision trees, long short-term memory (LSTM), random forests, and K-nearest neighbors (kNNs) emerging as popular options. The positive impact of interventions on student performance outcomes was evident across multiple instances, while the complexity of relationships between interventions and outcomes was also acknowledged.

8. Future directions



Future directions in this field should focus on addressing the identified limitations. Mitigating data quality and quantity issues, refining algorithm architectures for scalability, and enhancing interpretability and model performance are key challenges. Additionally, there is a need for continuous efforts to ensure ethical considerations in education technology, especially when dealing with personal student data. Researchers should explore hybrid approaches that combine the strengths of different machine learning techniques to achieve more accurate and interpretable predictions. The integration of natural language processing and sentiment analysis could further enrich predictive models by incorporating textual feedback and emotional indicators. Collaborative efforts between educators, researchers, and policymakers are essential to leveraging the potential of educational technology while responsibly navigating its challenges. In conclusion, although the systematic review provides insightful information about the state of machine learning applications for predicting student performance, the fact that the environment is constantly changing makes it necessary for ongoing research and innovation to fully realize the potential of these techniques for improving educational outcomes.

Acknowledgments

The authors is grateful to the Tunku Abdul Rahman University of Management and Technology (TAR UMT), Jalan Genting Kelang, Setapak, 53300 Kuala Lumpur, Malaysia

Ethical Considerations

This research is a form of consideration in making policies and research on MSMEs.

Conflict of Interest

The authors declare no conflict of interest.

Funding

Research is a purely private fund as a form of advanced research foundation.

References

- Adnan, M., Habib, A., Ashraf, J., Mussadiq, S., Raza, A. A., Abid, M., Bashir, M., & Khan, S. U. (2021). Predicting at-Risk Students at Different Percentages of Course Length for Early Intervention Using Machine Learning Models. *IEEE Access*, *9*, 7519–7539. <https://doi.org/10.1109/access.2021.3049446>
- Adnan, M., Uddin, M. I., Khan, E., Alharithi, F. S., Amin, S., & Alzahrani, A. A. (2022). Earliest Possible Global and Local Interpretation of Students' Performance in Virtual Learning Environment by Leveraging Explainable AI. *IEEE Access*, *10*, 129843–129864. <https://doi.org/10.1109/access.2022.3227072>
- Alalawi, K., Athauda, R., & Chiong, R. (2023). Contextualizing the current state of research on the use of machine learning for student performance prediction: A systematic literature review. *Engineering Reports*, e12699. <https://doi.org/10.1002/ENG2.12699>
- Alamri, R., & Alharbi, B. (2021). Explainable Student Performance Prediction Models: A Systematic Review. *IEEE Access*, *9*, 33132–33143. <https://doi.org/10.1109/access.2021.3061368>
- Albreiki, B., Zaki, N., & Alashwal, H. (2021). A Systematic Literature Review of Student' Performance Prediction Using Machine Learning Techniques. *Education Sciences 2021*, Vol. 11, Page 552, 11(9), 552. <https://doi.org/10.3390/EDUCSCI11090552>
- Alhazmi, E., & Sheneamer, A. (2023). Early Predicting of Students Performance in Higher Education. *IEEE Access*, *11*, 27579–27589. <https://doi.org/10.1109/access.2023.3250702>
- Aljaloud, A. S., Uliyan, D. M., Alkhalil, A., Elrhman, M. A., Alogali, A. F. M., Altameemi, Y. M., Altamimi, M., & Kwan, P. (2022). A Deep Learning Model to Predict Student Learning Outcomes in LMS Using CNN and LSTM. *IEEE Access*, *10*, 85255–85265. <https://doi.org/10.1109/access.2022.3196784>
- Alshanjiti, A., & Namoun, A. (2020). Predicting Student Performance and Its Influential Factors Using Hybrid Regression and Multi-Label Classification. *IEEE Access*, *8*, 203827–203844. <https://doi.org/10.1109/access.2020.3036572>
- Ara Shaikh, A., Kumar, A., Jani, K., Mitra, S., García-Tadeo, D. A., & Devarajan, A. (2022). The Role of Machine Learning and Artificial Intelligence for making a Digital Classroom and its sustainable Impact on Education during Covid-19. *Materials Today: Proceedings*, *56*, 3211–3215. <https://doi.org/10.1016/j.MATPR.2021.09.368>
- Asthana, P., Mishra, S., Gupta, N., Derawi, M., & Kumar, A. (2023). Prediction of Student's Performance With Learning Coefficients Using Regression Based Machine Learning Models. *IEEE Access*, *11*. <https://doi.org/10.1109/access.2023.3294700>
- Bujang, S. D. A., Selamat, A., Ibrahim, R., Krejcar, O., Herrera-Viedma, E., Fujita, H., & Ghani, N. A. M. (2021). Multiclass Prediction Model for Student Grade Prediction Using Machine Learning. *IEEE Access*, *9*, 95608–95621. <https://doi.org/10.1109/access.2021.3093563>
- Chan, R. Y.-Y., Wong, C. M. V., & Yum, Y. N. (2023). Predicting Behavior Change in Students With Special Education Needs Using Multimodal Learning Analytics. *IEEE Access*, *11*, 63238–63251. <https://doi.org/10.1109/access.2023.3288695>
- Chui, K. T., Liu, R. W., Zhao, M., & Ordóñez de Pablos, P. (2020). Predicting Students' Performance With School and Family Tutoring Using Generative Adversarial Network-Based Deep Support Vector Machine. *IEEE Access*, *8*, 86745–86752. <https://doi.org/10.1109/access.2020.2992869>
- Feng, G., Fan, M., & Chen, Y. (2022). Analysis and Prediction of Students' Academic Performance Based on Educational Data Mining. *IEEE Access*, *10*, 19558–19571. <https://doi.org/10.1109/access.2022.3151652>
- Gao, L., Zhao, Z., Qi, L., Liang, Y., & Du, J. (2019). Modeling the Effort and Learning Ability of Students in MOOCs. *IEEE Access*, *7*, 128035–128042. <https://doi.org/10.1109/access.2019.2937985>
- Ghorbani, R., & Ghousi, R. (2020). Comparing Different Resampling Methods in Predicting Students' Performance Using Machine Learning Techniques. *IEEE Access*, *8*, 67899–67911. <https://doi.org/10.1109/access.2020.2986809>



- Haleem, A., Javaid, M., Qadri, M. A., & Suman, R. (2022). Understanding the role of digital technologies in education: A review. *Sustainable Operations and Computers*, 3, 275–285. <https://doi.org/10.1016/J.SUSOC.2022.05.004>
- Hassan, Y. M. I., Elkorany, A., & Wassif, K. (2022). Utilizing Social Clustering-Based Regression Model for Predicting Student's GPA. *IEEE Access*, 10, 48948–48963. <https://doi.org/10.1109/access.2022.3172438>
- Hilbert, S., Coors, S., Kraus, E., Bischl, B., Lindl, A., Frei, M., Wild, J., Krauss, S., Goretzko, D., & Stachl, C. (2021). Machine learning for the educational sciences. *Review of Education*, 9(3), e3310. <https://doi.org/10.1002/REV3.3310>
- Hussain, W., Spady, W. G., Khan, S. Z., Khawaja, B. A., Naqash, T., & Conner, L. (2021). Impact Evaluations of Engineering Programs Using Abet Student Outcomes. *IEEE Access*, 9, 46166–46190. <https://doi.org/10.1109/access.2021.3066921>
- Issah, I., Appiah, O., Appiahene, P., & Inusah, F. (2023). A systematic review of the literature on machine learning application of determining the attributes influencing academic performance. *Decision Analytics Journal*, 7, 100204. <https://doi.org/10.1016/J.DAJOUR.2023.100204>
- Kastrati, Z., Imran, A. S., & Kurti, A. (2020). Weakly Supervised Framework for Aspect-Based Sentiment Analysis on Students' Reviews of MOOCs. *IEEE Access*, 8, 106799–106810. <https://doi.org/10.1109/access.2020.3000739>
- Kim, J., Lee, H., & Cho, Y. H. (2022). Learning design to support student-AI collaboration: perspectives of leading teachers for AI in education. *Education and Information Technologies*, 27(5), 6069–6104. <https://doi.org/10.1007/S10639-021-10831-6/TABLES/1>
- Kusumawardani, S. S., & Alfarozi, S. A. I. (2023). Transformer Encoder Model for Sequential Prediction of Student Performance Based on Their Log Activities. *IEEE Access*, 11, 18960–18971. <https://doi.org/10.1109/access.2023.3246122>
- Latif, G., Abdelhamid, S. E., Fawagreh, K. S., Brahim, G. Ben, & Alghazo, R. (2023). Machine Learning in Higher Education: Students' Performance Assessment Considering Online Activity Logs. *IEEE Access*, 11. <https://doi.org/10.1109/access.2023.3287972>
- Liu, D., Zhang, Y., Zhang, J., Li, Q., Zhang, C., & Yin, Y. (2020). Multiple Features Fusion Attention Mechanism Enhanced Deep Knowledge Tracing for Student Performance Prediction. *IEEE Access*, 8, 194894–194903. <https://doi.org/10.1109/access.2020.3033200>
- Mengash, H. A. (2020). Using Data Mining Techniques to Predict Student Performance to Support Decision Making in University Admission Systems. *IEEE Access*, 8, 55462–55470. <https://doi.org/10.1109/access.2020.2981905>
- Motz, B. A., Mallon, M. G., & Quick, J. D. (2021). Automated Educative Nudges to Reduce Missed Assignments in College. *IEEE Transactions on Learning Technologies*, 14(2), 189–200. <https://doi.org/10.1109/tlt.2021.3064613>
- Munshi, A. A., & Alhindi, A. (2021). Big Data Platform for Educational Analytics. *IEEE Access*, 9, 52883–52890. <https://doi.org/10.1109/access.2021.3070737>
- Nabil, A., Seyam, M., & Abou-Elfetouh, A. (2021). Prediction of Students' Academic Performance Based on Courses' Grades Using Deep Neural Networks. *IEEE Access*, 9, 140731–140746. <https://doi.org/10.1109/access.2021.3119596>
- Nguyen-Huy, T., Deo, R. C., Khan, S., Devi, A., Adeyinka, A. A., Apan, A. A., & Yaseen, Z. M. (2022). Student Performance Predictions for Advanced Engineering Mathematics Course With New Multivariate Copula Models. *IEEE Access*, 10, 45112–45136. <https://doi.org/10.1109/access.2022.3168322>
- Okoye, K., Daruich, S. D. N., De La O, J. F. E., Castano, R., Escamilla, J., & Hosseini, S. (2023). A Text Mining and Statistical Approach for Assessment of Pedagogical Impact of Students' Evaluation of Teaching and Learning Outcome in Education. *IEEE Access*, 11, 9577–9596. <https://doi.org/10.1109/access.2023.3239779>
- Orlando, S., Gaudio, E., & De La Paz, F. (2020). Supporting Teachers to Monitor Student's Learning Progress in an Educational Environment With Robotics Activities. *IEEE Access*, 8, 48620–48631. <https://doi.org/10.1109/access.2020.2978979>
- Pek, R. Z., Ozyer, S. T., Elhage, T., Ozyer, T., & Alhaji, R. (2023). The Role of Machine Learning in Identifying Students At-Risk and Minimizing Failure. *IEEE Access*, 11, 1224–1243. <https://doi.org/10.1109/access.2022.3232984>
- Pinto, A., Abreu, A., Costa, E., Information, J. P.-J. of, & 2023, undefined. (2023). How Machine Learning (ML) is Transforming Higher Education: A Systematic Literature Review. *Comum.Rcaap.PtAs Pinto, A Abreu, E Costa, J PaivaJournal of Information Systems Engineering and Management*, 2023•comum.Rcaap.Pt, 2023(2), 21168. <https://doi.org/10.55267/iadt.07.13227>
- Prabowo, H., Hidayat, A. A., Cenggoro, T. W., Rahutomo, R., Purwandari, K., & Pardamean, B. (2021). Aggregating Time Series and Tabular Data in Deep Learning Model for University Students' GPA Prediction. *IEEE Access*, 9, 87370–87377. <https://doi.org/10.1109/access.2021.3088152>
- Rafique, A., Khan, M. S., Jamal, M. H., Tasadduq, M., Rustam, F., Lee, E., Washington, P. B., & Ashraf, I. (2021). Integrating Learning Analytics and Collaborative Learning for Improving Student's Academic Performance. *IEEE Access*, 9, 167812–167826. <https://doi.org/10.1109/access.2021.3135309>
- Rahman, M. M., Watanobe, Y., Matsumoto, T., Kiran, R. U., & Nakamura, K. (2022). Educational Data Mining to Support Programming Learning Using Problem-Solving Data. *IEEE Access*, 10, 26186–26202. <https://doi.org/10.1109/access.2022.3157288>
- Raza, M., & Faria, J. P. (2020). The ProcessPAIR Method for Automated Software Process Performance Analysis. *IEEE Access*, 8, 141569–141583. <https://doi.org/10.1109/access.2020.3013328>
- Ros, S., Gonzalez, S., Robles, A., Tobarra, L. L., Caminero, A., & Cano, J. (2020). Analyzing Students' Self-Perception of Success and Learning Effectiveness Using Gamification in an Online Cybersecurity Course. *IEEE Access*, 8, 97718–97728. <https://doi.org/10.1109/access.2020.2996361>
- Sahlaoui, H., Alaoui, E. A. A., Nayyar, A., Agoujil, S., & Jaber, M. M. (2021). Predicting and Interpreting Student Performance Using Ensemble Models and Shapley Additive Explanations. *IEEE Access*, 9, 152688–152703. <https://doi.org/10.1109/access.2021.3124270>
- Samin, H., & Azim, T. (2019). Knowledge Based Recommender System for Academia Using Machine Learning: A Case Study on Higher Education Landscape of Pakistan. *IEEE Access*, 7, 67081–67093. <https://doi.org/10.1109/access.2019.2912012>
- Sanusi, I. T., Oyelere, S. S., Vartiainen, H., Suhonen, J., & Tukiainen, M. (2023). A systematic review of teaching and learning machine learning in K-12 education. *Education and Information Technologies*, 28(5), 5967–5997. <https://doi.org/10.1007/S10639-022-11416-7/TABLES/5>
- Schiff, D. (2021). Out of the laboratory and into the classroom: the future of artificial intelligence in education. *Ai & Society*, 36(1), 331. <https://doi.org/10.1007/S00146-020-01033-8>
- Sekeroglu, B., Abiyev, R., Ilhan, A., Arslan, M., & Idoko, J. B. (2021). Systematic Literature Review on Machine Learning and Student Performance Prediction: Critical Gaps and Possible Remedies. *Applied Sciences* 2021, Vol. 11, Page 10907, 11(22), 10907. <https://doi.org/10.3390/AP112210907>
- Setiawan, R., Arif, F. A. S., Putro, J. O., Princes, E., Silalahi, F. T. R., Geraldina, I., Julianti, E., & Safitri, J. (2023). E-Learning Pricing Model Policy for Higher Education. *IEEE Access*, 11, 38370–38384. <https://doi.org/10.1109/access.2023.3266954>
- Shao, Z., Sun, H., Wang, X., & Sun, Z. (2020). An Optimized Mining Algorithm for Analyzing Students' Learning Degree Based on Dynamic Data. *IEEE Access*, 8, 113543–113556. <https://doi.org/10.1109/access.2020.3001749>



- Shi, Y., Peng, F., & Sun, F. (2022). A Blended Learning Model Based on Smart Learning Environment to Improve College Students' Information Literacy. *IEEE Access*, *10*, 89485–89498. <https://doi.org/10.1109/access.2022.3201105>
- Shi, Y., Sun, F., Zuo, H., & Peng, F. (2023). Analysis of Learning Behavior Characteristics and Prediction of Learning Effect for Improving College Students' Information Literacy Based on Machine Learning. *IEEE Access*, *11*, 50447–50461. <https://doi.org/10.1109/access.2023.3278370>
- Sindhu, I., Muhammad Daudpota, S., Badar, K., Bakhtyar, M., Baber, J., & Nurunnabi, M. (2019). Aspect-Based Opinion Mining on Student's Feedback for Faculty Teaching Performance Evaluation. *IEEE Access*, *7*, 108729–108741. <https://doi.org/10.1109/access.2019.2928872>
- Singelmann, L. N., & Ewert, D. L. (2022). Leveraging the Innovation-Based Learning Framework to Predict and Understand Student Success in Innovation. *IEEE Access*, *10*, 36123–36139. <https://doi.org/10.1109/access.2022.3163744>
- Sun, Q., Wu, J., & Liu, K. (2020). Toward Understanding Students' Learning Performance in an Object-Oriented Programming Course: The Perspective of Program Quality. *IEEE Access*, *8*, 37505–37517. <https://doi.org/10.1109/access.2020.2973470>
- Tao, X., Shannon-Honson, A., Delaney, P., Li, L., Dann, C., Li, Y., & Xie, H. (2022). Data Analytics on Online Student Engagement Data for Academic Performance Modeling. *IEEE Access*, *10*, 103176–103186. <https://doi.org/10.1109/access.2022.3208953>
- Webb, M. E., Fluck, A., Magenheimer, J., Malyn-Smith, J., Waters, J., Deschênes, M., & Zagami, J. (2021). Machine learning for human learners: opportunities, issues, tensions and threats. *Educational Technology Research and Development*, *69*(4), 2109–2130. <https://doi.org/10.1007/S11423-020-09858-2/TABLES/1>
- Wei, Y., Lv, H., Chen, M., Wang, M., Heidari, A. A., Chen, H., & Li, C. (2020). Predicting Entrepreneurial Intention of Students: An Extreme Learning Machine With Gaussian Barebone Harris Hawks Optimizer. *IEEE Access*, *8*, 76841–76855. <https://doi.org/10.1109/access.2020.2982796>
- Wu, Y. C., & Liao, W. H. (2020). Analysis of Learning Patterns and Performance—A Case Study of 3-D Modeling Lessons in the K-12 Classrooms. *IEEE Access*, *8*, 186976–186992. <https://doi.org/10.1109/access.2020.3029947>
- Yanes, N., Mostafa, A. M., Ezz, M., & Almuayqil, S. N. (2020). A Machine Learning-Based Recommender System for Improving Students Learning Experiences. *IEEE Access*, *8*, 201218–201235. <https://doi.org/10.1109/access.2020.3036336>
- Yi, A. C. Y., Ying, T. K., Yee, S. J., Chin, W. M., & Tin, T. T. (2022). InPath Forum: A Real-Time Learning Analytics and Performance Ranking Forum System. *IEEE Access*, *10*, 128536–128542. <https://doi.org/10.1109/access.2022.3227430>
- Yoo, J. E., Rho, M., & Lee, Y. (2022). Online Students' Learning Behaviors and Academic Success: An Analysis of LMS Log Data From Flipped Classrooms via Regularization. *IEEE Access*, *10*, 10740–10753. <https://doi.org/10.1109/access.2022.3144625>
- Yue, J., Tian, F., Chao, K. M., Shah, N., Li, L., Chen, Y., & Zheng, Q. (2019). Learning Unit State Recognition Based on Multi-channel Data Fusion. *IEEE Access*, *7*, 149554–149567. <https://doi.org/10.1109/access.2019.2947091>
- Zhao, L., Chen, K., Song, J., Zhu, X., Sun, J., Caulfield, B., & Namee, B. Mac. (2021). Academic Performance Prediction Based on Multisource, Multifeature Behavioral Data. *IEEE Access*, *9*, 5453–5465. <https://doi.org/10.1109/access.2020.3002791>
- Zhen, Y., Luo, J. Der, & Chen, H. (2023). Prediction of Academic Performance of Students in Online Live Classroom Interactions—An Analysis Using Natural Language Processing and Deep Learning Methods. *Journal of Social Computing*, *4*(1), 12–29. <https://doi.org/10.23919/jsc.2023.0007>

