Acta Scientific COMPUTER SCIENCES

Volume 6 Issue 7 July 2024

Research Article

Predictive Power of Machine Learning Models on Degree Completion Among Adult Learners

Emily Barnes[1], James Hutson[2]*, Karriem Perry[3]

[1]*Capitol Technology University, Maryland*

[2]*Lindenwood University, United States*

[3]*Capitol Technology University, Maryland*

***Corresponding Author:** James Hutson, Lindenwood University, United States.

Received: June 03, 2024

Published: June 24, 2024

Abstract

The integration of machine learning (ML) into higher education has been recognized as a transformative force for adult learners, a growing demographic facing unique educational challenges. This study evaluates the predictive power of three ML models—Random Forest, Gradient-Boosting Machine, and Decision Trees—in forecasting degree completion among this group. Utilizing a dataset from the academic years 2013-14 to 2021-22, which includes demographic and academic performance metrics, the study employs accuracy, precision, recall, and F1 score to assess the efficacy of these models. The results indicate that the Gradient-Boosting Machine model outperforms others in predicting degree completion, suggesting that ML can significantly enhance data-driven decision-making in educational settings. By highlighting the factors influencing adult learners' educational success, such as age and socioeconomic status, this research supports the strategic implementation of tailored educational policies and interventions, aimed at improving the retention and graduation rates of adult learners in higher education institutions.

Keywords: Machine Learning; Adult Learners; Educational Outcomes; Predictive Analytics; Degree Completion

Introduction

The incorporation of machine learning into higher education is increasingly acknowledged as a transformative force, adeptly responding to the changing needs of varied student populations, especially those of adult learners [1]. Studies have traced the historical evolution of higher education, noting a significant shift towards adult learners, who face unique challenges like balancing education with other commitments and adapting to new learning environments [1-3]. The function of ML in higher education surpasses conventional educational frameworks and the analysis of student performance data and the forecasting of educational results. Various ML algorithms, including Random Forests, Decision Trees, and neural networks, have demonstrated effectiveness in forecasting student achievement, retention, and graduation likelihoods [4,5]. ML's capability in enhancing data-driven decision-making in higher education and aids in the development of inclusive and flexible educational models that cater to the needs of adult learners. This shift towards ML-driven analytics represents a significant advancement in predicting and improving educational strategies and student support mechanisms [1-2,4-5].

Buenaño-Fernández., *et al*. [6] and Korkmaz and Correia [7] identified the capability of ML to refine personalized learning and enhance student progress monitoring. Oliveira., *et al*. [8] and Kannan [9] advocated for additional research on feature selection and class balancing, along with exploring human-centric methodologies in education. Hilbert., *et al*. [10] and Son., *et al*. [11] investigated the potential and challenges of ML in educational sciences, considering the prospects of automated evaluations and individualized feedback. Pinto., *et al*. [12] and Hessen [13] examined the possible ML applications in higher education, including its part in predicting academic success and employability, as well as its relevance in various industry-related contexts.

The potential of ML in higher education is vast, with several promising directions for future research and application [14]. Personalized learning experiences are anticipated to become more refined through advanced ML algorithms, enabling education systems to adapt to individual learning styles and preferences. The expansion of predictive analytics applications is expected to encompass a broader range of contexts, including emotional well-being and career readiness [11].

Statement of Problem

Higher education institutions are increasingly confronted with the challenge of integrating data-driven decision-making processes, hampered by rapidly advancing technologies, antiquated information systems, limited resources, and entrenched practices [15]. The incorporation of advanced methodologies, such as machine learning (ML), informs policies, strategies, and tasks related to early intervention and determining student success factors [16,17]. Despite its potential, the use of such technologies to enhance educational outcomes, particularly for adult learners, remains insufficiently explored. This study aims to evaluate effective ML algorithms that can predict graduation rates among adult learners and elucidate the factors that affect their educational progress. Furthermore, as Hussain., *et al.* [18] have indicated, machine learning can shape inclusive educational practices and policies within higher education frameworks. Consequently, this research focuses on examining the application of machine learning in higher education settings, specifically analyzing its impact on the success of adult learners.

Research Questions and Hypotheses

- **Research Question 1:** Which machine learning model among Random Forest, Gradient-Boosting Machine, and Decision Trees most accurately predicts degree completion?
- **Hypothesis 1:** The Gradient-Boosting Machine model will exhibit superior accuracy in predicting degree completion compared to Random Forest and Decision Trees, positioning it as the most effective analytical tool for data-informed decision-making.

Term	Definition
Artificial Intelligence (AI)	Replication of cognitive functions associated with human intelligence by machines, particularly computer systems. Includes capabilities such as learning, reasoning, and self-improvement
Adult Learner	Non-traditional student who engages in the pursuit of higher education or lifelong learning beyond the conventional college-age bracket of 18 to 24 years
Cross-Validation Score (CV Score)	Performance measure that calculates the predictive performance of a ML model
Decision Tree (DT)	Model that employs a hierarchical decision structure resembling a tree to navigate through decisions and their potential outcomes, commonly utilized for tasks involving classification and regression
Estimators	Number of trees in the forest. In a RF model, multiple Decision Trees are created, and their predictions are aggregated to give the final output
F1 score	Statistical measure used to evaluate the performance of a binary classification model, particularly useful when the class distribution is imbalanced
Gradient-Boosting Machine (GBM)	Ensemble method that populates trees in a consecutive manner, where each tree trains in a sequential manner to improve accuracy of outcomes
Hyperparameters	Configuration settings used to structure a machine-learning model. Set prior to the training process and dictate the behavior of the learning algorithm. Directly influence the performance of the model
Max_Depth	Maximum depth of each tree in the forest. The depth of a tree is the longest path from the root node down to the farthest leaf node. Measure of how many splits a tree can make before reaching a prediction
Machine Learning (ML)	Segment of artificial intelligence that focuses on the development of algorithms to perform tasks without explicit instructions
Precision	Ratio of true positive predictions to the total number of positive predictions made, reflecting the accuracy of the model in labeling an instance as positive
Random Forest (RF)	Classification and regressive ensemble learning model that constructs Decision Trees trained on different segments of the same training set
Recall	Ratio of true positive predictions to the total number of actual positive instances, reflecting the model's ability to identify all relevant instances
Standard Deviation	A statistical measure that quantifies the amount of variation or dispersion in a set of values. It is used to assess the variability of a model's prediction errors or performance metrics across different runs or folds of cross-validation.

Table 1: Definitions and Abbreviations.

Literature

This literature review establishes a comprehensive understanding of the evolving dynamics within higher education, especially focusing on the integration of machine learning (ML) and its impact on adult learning. By situating the study within the realms of ML application in higher education, the review highlights the current academic landscape and identifies gaps that this research aims to address. It sets the stage for exploring the effectiveness of sophisticated algorithms in predicting academic outcomes for adult learners, aiming to enhance data-informed educational practices and policy development.

Historical context and current trends

Higher education has undergone significant evolution, from its origins in religious institutions to becoming inclusive settings that cater to a diverse student body, adapting continually to societal and technological changes. The expansion of access and curricula, alongside the integration of technology, marks key developments that have shaped its current form [1,2]. Recent shifts focus particularly on adult learners who form a substantial segment of the student population with some college experience but no credentials. This demographic shift requires educational systems to adapt to the needs of adult learners, driven by job market demands and the value of lifelong learning [1,3,19].

Adapting to adult learners

The trend towards accommodating adult learners represents a strategic transformation in higher education to address broader community needs. This shift is supported by an increasing preference for flexible learning options such as online courses and part-time programs, which are crucial for adults balancing education with other responsibilities [16,17]. However, adult learners face unique barriers such as accessibility, affordability, and aligning educational outcomes with workplace demands. Addressing these challenges is essential for developing inclusive and flexible educational models that support lifelong learning and cater to the evolving needs of the adult learner population [15,20,21].

Machine learning in educational research

Machine learning (ML) is significantly shaping educational research, particularly in analyzing and predicting student performance data. Research highlights the effectiveness of artificial neural networks (ANNs) and other ML algorithms in early prediction and intervention strategies for student performance, with notable focus on STEM education and e-learning systems [22,23]. ML applications are not only enhancing predictive accuracy but also tailoring educational content to meet individual student needs through adaptive learning systems [24].

Efficacy of different ML algorithms

Decision Trees: Studies have proven Decision Trees effective in forecasting the academic outcomes of adult learners, employing methodologies like the C4.5 algorithm for nuanced predictions in various educational contexts, from graduation rates to licensure examinations [25,26].

Gradient Boosting Machines: These have been applied to predict diverse educational outcomes, including student employability and lifelong education participation, illustrating their utility in strategic educational decisions [27,28].

Random Forests: Known for their robustness, Random Forest algorithms have been effective in predicting academic performance and identifying dropout risks, with studies reporting high accuracy rates [29,30].

Comparative analysis across models

Research across various ML models demonstrates their capacity to address different aspects of academic administration and student support. Whether assessing student mental health or academic success, models like XGBoost, ANNs, and Random Forests have shown superior efficacy, underscoring the need for tailored approaches based on specific educational goals and contexts [31].

Feature importance in machine learning and educational research

Feature importance in machine learning is crucial for enhancing model interpretability and identifying the factors influencing predictions. This concept has become fundamental in educational research, particularly in projects that aim to forecast student performance and success [31,32]. Studies have demonstrated that variables such as demographics and academic performance are pivotal predictors. Advanced machine learning techniques, including vector machines and neural networks, have proven effective in identifying key predictors and refining educational models [33,34].

Integration of ML in education data

The integration of machine learning in educational data mining has significantly advanced, addressing challenges such as student dropout rates and enhancing the adaptability of learning systems to meet individual needs [19]. The adoption of feature selection methodologies, like the fast correlation-based filter (FCBF) and correlation-based feature selection (CFS), has improved the precision of predictive models [27].

Research also highlights the challenges of employing machine learning in education, including issues with data quality, model bias, and the complexity of interpreting opaque models [35-36]. Studies emphasize the necessity of integrating various feature importance measures and using sophisticated explanation techniques to enhance understanding and application of these models in educational settings [37,38].

Machine learning and student retention in higher education

Machine learning (ML) has proven effective in predicting student enrollment, performance, and retention with high accuracy, addressing significant concerns for educational institutions that view retention as a revenue stream [39]. Studies by Cardona., *et al*. [33] (2020) and Palacios., *et al*. [40] have demonstrated the capability of ML models to predict student retention accurately, identifying key factors such as secondary educational scores and socio-economic indicators. This research supports the use of ML to develop early intervention strategies and tailored student support systems.

Further advancements in ML applications, such as ML-based recommendation systems by Arqawi., *et al*. [41] have shown that these systems can enhance student retention effectively. Their work illustrates how ML not only predicts outcomes but also supports the creation of personalized educational interventions, evolving from analytical tools to integral components of educational strategy and policy development. This shift represents ML's transformational ability for educational approaches to student retention, emphasizing the need for a holistic strategy that incorporates data-driven insights to optimize educational outcomes and support mechanisms effectively.

Enhancing machine learning models with socio-economic and behavioral data in education

Incorporating socio-economic and behavioral data into machine learning (ML) models significantly enhances educational research, particularly in predicting student performance. Researchers like Brdesee., *et al*. [42] and Alsariera., *et al*. [31] have shown that using advanced algorithms such as Long Short-Term Memory (LSTM) networks and Artificial Neural Networks (ANNs) can improve the accuracy of models in predicting student outcomes. These studies emphasize the importance of integrating a broad range of data, including demographic, academic, and behavioral factors, to gain a comprehensive understanding of the influences on student success.

The application of Automated Machine Learning (AutoML) and data mining algorithms, as discussed by Zeineddine [43] and Chen., *et al*. [20], further demonstrates the adaptability of ML models to handle complex datasets and enhance prediction capabilities. Research by Cardona., *et al*. [33] has explored the impact of socio-economic backgrounds and behavioral aspects using models like Support Vector Machines and Multilayer Perceptrons, focusing on how non-academic factors contribute to educational outcomes. This holistic approach acknowledges the significant influence of external circumstances on student performance.

Studies by Gaftandzhieva., *et al*. [44] and Palacios., *et al*. [40] have also incorporated variables such as educational backgrounds and community poverty indices into their models, using Random Forest algorithms to analyze the intricate relationship between various factors and student success. The ongoing refinement of these ML models reveal their potential to inform effective, targeted interventions that cater to the diverse needs of students, leading to more personalized and equitable educational strategies. The integration of comprehensive socio-economic and behavioral data into ML models promises to enhance the personalization of learning experiences, making educational interventions more responsive and effective for an increasingly diverse student body.

Machine learning's role in evolving educational landscapes

Machine learning (ML) has become integral to advancing post-secondary education by enhancing student retention, predicting academic success, and supporting educational decisions. Researchers like Oqaidi., *et al*. [45] have highlighted ML's capability in managing educational data and improving student retention. Studies by Pinto., *et al.* [12] and Jusslin., *et al*. [46] further explore ML's potential in forecasting academic outcomes and integrating into curriculum development to bolster digital education. These investigations collectively illustrates ML's part in enhancing various aspects of educational administration and support.

The adaptability of ML to the dynamic educational landscape-marked by shifting policies and diverse student needs-is crucial for its effective integration. Researchers like Nauman., *et al*. [47] focus on developing customizable ML models that cater to individual learner preferences, reflecting an understanding that diverse student backgrounds require sophisticated, personalized educational approaches. The studies by Allen., *et al.* [48] and Nieto., *et al*. [49] on refining ML model accuracy and adapting them to e-learning environments further demonstrate how ML can be tailored to meet the challenges of modern educational settings.

This body of work suggests that continuous research and development, along with a careful selection of ML algorithms, are essential for leveraging ML to create responsive, personalized educational experiences that adapt to new educational paradigms. The collaboration between educational researchers and data scientists is pivotal, fostering a holistic approach that combines technical ML expertise with deep educational insights to revolutionize educational practices effectively. This interdisciplinary effort is key to developing ML applications that enhance teaching and learning across diverse educational landscapes.

Effective machine learning implementations in education

Machine learning (ML) has proven instrumental in various educational contexts by facilitating the early identification of at-risk students and enhancing personalized learning strategies. Studies like those by Tarmizi., *et al.* [50] and Yağcı [51] highlight successful ML applications, despite challenges such as data privacy concerns and the adaptation to online learning platforms during the COVID-19 pandemic. These implementations have significantly improved student retention and engagement by enabling timely interventions and sustaining educational continuity.

For instance, Martins., *et al*. [52] integrated active, problem-based learning to teach ML concepts in high school, enhancing student engagement and understanding of complex topics. Similarly, Jun-on., *et al*. [53] used ML algorithms to analyze student performance in English courses, identifying patterns that predict academic outcomes. These studies encountered challenges like the steep learning curve of ML and the need for comprehensive data but ultimately demonstrated the potential of ML to predict and improve student performance.

The integration of ML with pedagogical strategies, particularly for adult learners, is emerging as a promising approach to enhance educational outcomes. Research by Oliveira., *et al.* [8] and Davari., *et al*. [54] has shown how AI and ML can analyze learning analytics to customize learning interventions, making education more responsive to the diverse needs of adult learners. Moreover, strategies that align ML with clear educational objectives, such as those advocated by Liu [55] and Martins [52], encourage the practical application of ML in educational settings, ensuring that technological implementations are grounded in solid pedagogical principles.

Challenges and limitations of machine learning in education

ML has significant potential to transform educational practices by enhancing personalized education and improving academic outcomes, as highlighted by Hilbert., *et al.* [10]. However, this potential comes with challenges, including the need for new approaches to model evaluation and a deeper understanding of ML's limitations. Allen., *et al*. [48] and Pinto., *et al*. [12] emphasized the importance of developing best practices for teaching AI and ML, particularly focusing on boosting student confidence and accommodating diverse educational backgrounds. Nieto., *et al*. [49] explored practical applications of ML, such as dropout prediction and support for strategic decision-making in educational settings. Korkmaz and Correia [7] reviewed current trends in ML research within educational technology, underlining the need for rigorous development of best practices to fully leverage ML in enhancing postsecondary education. These studies collectively underscore the importance of addressing the challenges associated with ML implementation in education to realize its full potential in improving learning outcomes and educational quality.

The adoption of ML in higher education raises significant ethical challenges and concerns, particularly as institutions increasingly rely on data-driven decision-making [5] Saltz., *et al*. [56] and Salihoun [5] have highlighted the importance of integrating ethical considerations into ML education and developing robust ethical frameworks to guide ML practices. Issues such as data privacy, the potential for bias, and the ethical use of algorithms are paramount as educators and technologists navigate the complexities of applying ML in academic settings. Toms and Whitworth [57] and Musso., *et al*. [58] further discussed the ethical dilemmas posed by ML in educational research and its applications, emphasizing the need for ethical guidelines. An issue identified is the quality of data used in ML models. Educational datasets often suffer from inconsistencies and gaps that can significantly impact the performance and reliability of ML models, making the pursuit of high-quality, accurate data a priority in educational ML applications [5,12]. These concerns underscore the need for careful consideration of ethical and data quality issues to ensure that ML technologies are used responsibly and effectively in educational contexts.

The influence of machine learning on the future of higher education

ML is increasingly recognized as a transformative force in higher education, particularly in enhancing the learning experiences of adult learners and improving educational outcomes across diverse student populations. ML extends beyond traditional educational methods, effectively analyzing student performance data and predicting educational outcomes with advanced algorithms such as Random Forests, Decision Trees, and neural networks [1,4,5]. This shift towards data-driven strategies facilitates the development of inclusive and adaptable educational models, tailored to meet the specific needs of adult learners.

The future of ML in education promises further advancements, with studies suggesting numerous potential applications to enhance personalized learning and improve student monitoring [6,8]. Research continues to explore the integration of ML with pedagogical strategies, aiming to align educational practices with the nuanced needs of students, particularly in adapting to individual learning styles and preferences. The advancement of NLP and ethical AI also indicates a broadening of ML's applications in education, addressing challenges such as bias, privacy, and the ethical deployment of technology in educational settings.

Longitudinal studies on ML in education

Longitudinal studies in higher education have shed light on the enduring impacts of various educational strategies on student learning and retention, revealing the complex dynamics of educational interventions over time. Research by Pande., *et al*. [59] indicates that immersive learning environments, like virtual reality simulations, sustain their benefits over extended periods, underscoring the importance of incorporating advanced technologies in education for long-term learning enhancement. Similarly, Maravé-Vivas., *et al*. [60] found that while service-learning positively affects student outcomes initially, its effectiveness tends to wane, suggesting the necessity for continuous engagement and iterative adjustments in educational practices.

Further, Holenstein., *et al*. [61] demonstrated the essentialness of mathematical literacy in sustaining academic success across various educational domains, highlighting the importance of foundational skills that support long-term academic achievements. Carless [62] advocated for evolving feedback processes that adapt over time, emphasizing the dynamic nature of learning and the potential for ongoing improvements in feedback effectiveness. In addition, active student involvement and repeated exposure are crucial for enhancing learning retention, stressing the need for educational environments that promote consistent student engagement and participation. These insights from longitudinal studies emphasize adaptability, continuous interaction, and the integration of innovative methods in fostering sustained educational success.

Conclusions from Research

This literature review has examined the integration of ML within higher education, focusing particularly on adult learning. It highlights the evolving place of ML in shaping educational outcomes for adult learners, noting significant historical shifts towards inclusivity and the increasing adoption of flexible, technology-enhanced learning models. Reports from the National Student Clearinghouse Research Center [3] underline a rise in adult learners returning to education, underscoring a need for educational strategies that accommodate their distinct needs and lifestyles.

Longitudinal studies reveal how ML applications can potentially enhance student success metrics, though they also expose challenges such as gaps in technical expertise, resource limitations, and the need for robust ethical frameworks within educational institutions. These challenges require strategic responses to effectively integrate ML technologies and leverage their capabilities.

The review concludes that ML holds transformative potential for higher education. By addressing specific challenges and enhancing institutional capacities for ML integration, higher education can better serve adult learners, ultimately enriching their educational experiences and contributing positively to the broader educational landscape.

Methods

In this study, we quantitatively assessed three ML models-Random Forest, Gradient-Boosting Machine (GBM), and CART Decision Tree-focusing on their accuracy in predicting degree completion rates among adult learners. We used key statistical measures such as accuracy, precision, recall, and F1 score to evaluate the reliability of these models. The analysis included a comprehensive dataset from the academic period 2013-14 to 2021-22, encompassing variables like age, ethnicity, gender, Pell Grant eligibility, and academic performance metrics.

Preprocessing

The study's methodology involves meticulous data collection and preprocessing to maintain data integrity and confidentiality. Data is collected securely and processed to correct inaccuracies and integrate multiple sources. This ensures that the data is suitable for ML analysis, which could profoundly influence educational strategies and policies.

Model building and evaluation procedures

The research methodology includes segmenting the dataset into training (80%) and testing (20%) subsets to evaluate the models' accuracy and generalization capabilities. This segmentation ensures a balanced approach to model training and validation, allowing comprehensive learning from the training data and effective performance assessment on unseen testing data.

For this specific study, we adopted an 80/20 split ratio, partitioning the dataset into two segments: 7,999 entries for training and 2,000 for testing. This widely accepted ratio within the machine learning and data science communities ensures a balanced approach between training and validation. The substantial training subset is used to finetune the models' parameters, allowing them to learn from a broad spectrum of data points and scenarios. This comprehensive training is crucial for the models to identify and learn the underlying patterns in the data, enhancing their predictive power.

The smaller testing subset evaluates the model's performance by applying it to new, unseen data. This phase is essential for assessing the effectiveness of the training and the model's ability to generalize its learning to different situations. It is key to ensuring the model's consistent and accurate performance across varied datasets. Bottom of Form.

Initial parameter selection and iterative refinement

The initial parameter settings for the RF, GBM, and CART models are selected based on a combination of best practices and empirical research. For instance, the RF model's parameters, such as the number of trees and their maximum depth, are chosen to balance accuracy with computational efficiency. This setup facilitates iterative refinement and testing, enhancing the models' predictive precision.

Model	Strengths	Usage
Random Forest Classifier	Robust, handles large data volumes with high accuracy and minimal computation, combines multiple Decision Trees to improve accuracy and reduce overfitting	Ideal for educational data analysis, adaptability and feature selection capabilities facilitate the identification of predictors for graduation rates, aiding the development of targeted educational interventions.
Gradient-Boosting Machine (GBM)	Strong predictive performance, refines accuracy by iteratively correcting previous prediction errors	Suitable for complex educational data that may initially obscure underlying patterns, particularly in forecasting educational outcomes like graduation rates.
CART Decision Tree	Excels in classifying and predicting student performance through a clear, interpretable approach	Uses historical data to predict future educational outcomes, improvements in Decision Tree algorithms, such as the C4.5, have significantly boosted their predictive accuracy, enhancing their utility in educational settings.

Table 2: Model Selection.

Model	Max_Depth	Estimators
RF1	3	100
RF2	5	200
RF3	10	50

Table 3: Random forest model parameters.

Note: Table 3 provides an overview of the parameter configurations and the resulting accuracy for three iterations of the RF Classifier model: RF1, RF2, and RF3.

In Table 3, the learning rate is a parameter affecting the progression of learning in GBM models. Friedman's [63] research emphasizes the importance of a learning rate that prevents overfitting by moderating the influence of each successive tree. GBM models were set with learning rates of 0.01 for GBM1 and GBM3, and 0.05 for GBM2. Additionally, the assigned number of trees and their depths were set with careful consideration to achieve an equilibrium between a model's complexity and its ability to generalize. Table 3 lists each GBM iteration and the different levels of depth and estimators.

When setting the parameters for the Decision-Tree model, the max depth and min samples split were chosen with an eye towards creating a model that was neither overly simplistic nor excessively complex, drawing from Quinlan's [64] insights into the optimal construction of Decision Trees. As seen in Table 4, the iterations DT1, DT2, and DT3 with max depths of 5, 3, and 2, alongside min samples splits of 2 and 3, were designed to test various levels of model simplicity and complexity. This approach was aimed at fostering models that are well-calibrated to discern the relevant patterns in the data without succumbing to the pitfalls of overfitting or underfitting [65,66].

Model	Max_Depth	Estimators	Learning Rate
GBM1	3	100	.01
GBM2	4	150	.05
GBM3	5	200	.01

Table 4: Gradient boosting machine model parameters.

Note: Table 4 outlines the configuration parameters and accuracy outcomes for three iterations of the Gradient-Boosting machine model, labeled GBM1, GBM2, and GBM3.

Model	Max_Depth	Min_Samples_Split
DT1	5	2
DT2	3	3
DT3	2	3

Table 5: Decision tree model parameters.

Note: Table 5 outlines the parameters and accuracy for three separate iterations of the Decision Tree model -DT1, DT2, and DT3.

Tables 3, 4, and 5 serve as references for understanding how the initial parameters were set and adjusted through subsequent iterations to enhance each model's predictive performance. They reflect the empirical nature of model development, where theoretical knowledge and industry standards intersect with practical experimentation and outcome analysis. The iterative training process, coupled with these informed parameter settings, aims to produce models that not only perform well on known data but also maintain their accuracy when faced with new, unseen data, ultimately leading to more reliable and robust predictions.

Integrating feature importance into the evaluation process involves leveraging methods such as Gini importance for RF models, which measures the frequency and depth with which features contribute to the Decision Trees within the model [42,67]. For Gradient-Boosting Machines and CART Decision Trees, techniques such as permutation importance and SHAP (SHapley Additive exPlanations) values are employed. A technique Huynh-Cam., *et al.* [65] similarly performed to quantify the impact of each feature on the prediction outcome, thereby identifying not just how well a model predicts, but why it predicts in a certain way.

The analysis of feature importance serves multiple purposes within the context of evaluating ML models for educational data. It aids in identifying key predictors of academic success, such as socioeconomic status, attendance patterns, or GPA, thereby allowing educators and policymakers to focus on the most impactful areas for intervention. Furthermore, it assists in model refinement by highlighting features that contribute little to prediction accuracy, which can be candidates for removal in the interest of model simplicity and efficiency [42,65-67].

Incorporating this dimension into the model evaluation framework aligns with the study's goals of enhancing educational outcomes and informing policy. By understanding the drivers behind model predictions, the study goes beyond merely identifying the most effective model to providing actionable insights into the factors influencing adult learners' success. This approach not only ensures a holistic assessment of model performance but also contributes to the development of targeted, data-driven strategies designed to support adult learners more effectively. Consequently, the incorporation of feature importance analysis into the comprehensive model evaluation and optimal selection phase represents a step towards achieving more inclusive, adaptable, and effective educational interventions tailored to the diverse needs of adult learners in the modern educational landscape.

Findings

This section of the research presents a detailed analysis of the outcomes derived from implementing various machine learning (ML) models within the study. It specifically examines the performance of three distinct models-Random Forest (RF), Gradient Boosting Machine (GBM), and Decision Tree-across multiple iterations. These analyses provide deep insights into each model's predictive capabilities and their effectiveness in interpreting the academic success of adult learners at private, non-profit educational institutions.

The performance of each model was rigorously evaluated over three iterative rounds using several key metrics, including mean accuracy, standard deviation, cross-validation score, precision, recall, and F1 score. These metrics highlight the efficacy, reliability, and suitability of each model in predicting student performance within educational datasets. The iterative process of training and evaluation was in refining the models, thereby ensuring they effectively addressed the core research questions. These included identifying the most accurate model for predicting degree completion, assessing the reliability and consistency of the models via their cross-validation scores, and understanding how the models' feature importance correlates with actual determinants of academic success.

Random forest classifier results

The evaluation of the Random Forest (RF) classifier across three iterations-RF1, RF2, and RF3-reveals significant insights into each model's performance, guided by a comprehensive array of metrics such as mean accuracy, standard deviation, cross-validation score, precision, recall, and F1 score. These metrics, as outlined in table 5 and detailed further in table 6, help assess the effectiveness, reliability, and suitability of each model in predicting student outcomes within educational datasets.

Model	Accuracy	Avg CV Score	Mean Accuracy	Standard Deviation
RF1	0.8335	0.74918	0.74918	0.15
RF2	0.8350	0.73708	0.73708	0.13
RF3	0.8185	0.57968	0.57968	0.11

Table 6: Random forest performance metrics.

Note: Number of records in training set = 7999, Number of records in testing set = 2000.

Table 6 illustrates that RF2 not only displayed the highest accuracy but also excelled in precision, indicating its robustness in minimizing false positives and enhancing predictive reliability. The standard deviations for each model (0.15 for RF1, 0.13 for RF2, and 0.11 for RF3) highlight a moderate variability in performance, which shows the model consistency and reliability in predicting educational outcomes.

Model	Accuracy	Precision	Recall	F1-Score
RF1	0.8335	0.824251	0.941634	0.879041
RF2	0.8350	0.840828	0.916732	0.877141
RF3	0.8185	0.823282	0.913619	0.866101

Table 7: Forest model effectiveness and reliability scores.

Note: Table 7 presents a detailed comparative analysis of the effectiveness and reliability of three RF Classifier models-RF1, RF2, and RF3-through the lens of four performance metrics: accuracy, precision, recall, and F1-score. Number of records in training set = 7999, Number of records in testing set = 2000.

Each iteration underwent extensive training and testing, using 7,999 records for training and 2,000 for testing, to ensure that the models could generalize effectively across different datasets and prevent overfitting. The performance evaluation of these models contributes significantly to the field of predictive analytics in higher education, offering a robust framework for future research and policy-making aimed at enhancing educational experiences for adult learners.

Gradient-Boosting machine results

The Gradient-Boosting Machine models, specifically GBM1, GBM2, and GBM3, have been rigorously tested across a dataset involving 7,999 training instances and 2,000 testing instances. Their performance metrics are captured comprehensively, focusing on accuracy, mean cross-validation (CV) scores, standard deviation, precision, recall, and F1 scores.

Model	Accuracy	Mean CV Score	Standard Deviation
GBM1	0.8335	0.65038	0.114630
GBM2	0.8350	0.58188	0.130518
GBM3	0.8185	0.61498	0.145319

Table 8: Gradient boosting machine performance metrics.

Note: Table 8 provides a synopsis of the performance metrics for GBM1, GBM2, and GBM3. The performance of these models is assessed based on three key metrics: accuracy, mean cross-validation (CV) score, and standard deviation.

GBM1 showed a precision of 81.76% and a recall of 92.60%, leading to an F1-score of 87.99%. GBM2's figures were slightly lower, with an F1-score of 87.17%. GBM3, however, while having lower precision at 80.33%, had the highest recall at 97.89%, resulting in an F1-score of 88.24%. These scores reflect GBM3's ability to identify true positives effectively, though with a slightly lower precision.

An expansion into the realm of model effectiveness and reliability is provided in Table 8, where precision, recall, and F1-scores are considered. GBM1 offers a precision of 81.76 percent and a recall of 92.60 percent, culminating in an F1-score of 87.99 percent. GBM2, while exhibiting comparable precision and recall to GBM1, sees a slight reduction in its F1-score to 87.17 percent. GBM3's precision drops to 80.33 percent, but its recall ascends to a noteworthy

Model	Accuracy	Precision	Recall	F1-Score
GBM1	0.8335	0.817635	0.926070	0.879942
GBM2	0.8350	0.823529	0.926070	0.871795
GBM3	0.8185	0.803321	0.978988	0.882497

Table 9: Gradient boosting machine model effectiveness and reliability scores.

Note: Table 9 provides performance metrics for GBM1, GBM2, and GBM3. The performance of these models is assessed based on three key metrics: accuracy, precision, recall, and F1-Score.

97.89 percent, yielding an F1-score of 88.24 percent. These scores highlight GBM3's particular adeptness at identifying true positives, although this is counterbalanced by a marginally lower precision.

The GBM models have showcased a potent capacity to analyze complex data relationships within educational settings, with each model presenting unique strengths. GBM1 and GBM2 offer higher accuracy, making them suitable for scenarios where general prediction accuracy is important. In contrast, GBM3, with its high recall rate, might be preferable in situations where missing a true positive has significant consequences.

CART decision tree results

The Classification and Regression Trees (CART), specifically models DT1, DT2, and DT3, have undergone comprehensive evaluations to assess their effectiveness in predicting educational outcomes. The analysis involves 7,999 records for training and 2,000 for testing, focusing on metrics such as accuracy, mean cross-validation (CV) score, standard deviation, precision, recall, and F1 score.

DT1 and DT2 both achieved an accuracy rate of 83.15%, indicating robust predictive capabilities. DT2 is noted for its superior mean CV score of 70.12%, suggesting greater generalizability across various data subsets. It also shows slightly higher variability with a standard deviation of 0.145, compared to DT1.

DT3 registered a lower accuracy of 80.85% but still offers reliable predictions with a mean CV score of 65.16%. However, it has the highest standard deviation at 0.162, indicating the most significant variability among the three.

Model	Accuracy	Mean CV Score	Standard Deviation
DT1	0.8315	0.60008	0.143494
DT2	0.8315	0.70128	0.145324
DT3	0.8085	0.65158	0.162175

Table 10: Decision tree performance metrics.

Model	Accuracy	Precision	Recall	F1-Score
DT1	0.8315	0.815160	0.954086	0.879168
DT2	0.8315	0.795143	0.993774	0.883431
DT3	0.8085	0.897707	0.792218	0.841670

Table 11: Decision tree model effectiveness and reliability scores.

DT1 presented a balanced precision of 81.52% and a high recall of 95.41%, leading to an F1 score of 87.92%. DT2, while having slightly lower precision at 79.51%, showed an exceptional recall of 99.38%, resulting in an F1 score of 88.34%. DT3 displayed the highest precision at 89.77% but lower recall at 79.22%, culminating in an F1 score of 84.17%.

This matrix can be dissected into four quadrants: the top left quadrant represents the True Positive (TP) count, which signifies the instances where the models accurately identified 1226 cases as positive. Conversely, the top right quadrant details the False Positive (FP) count, indicating 59 instances where the models incorrectly labeled negative cases as positive. The bottom left quadrant corresponds to the False Negative (FN) count, where the models misclassified 278 positive instances as negative. The bottom right

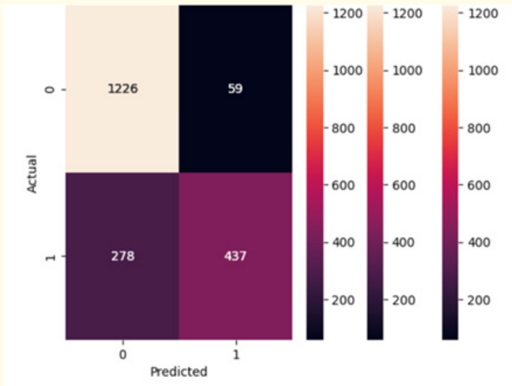

Figure 1: Decision tree accuracy-confusion matrix.
Note: The confusion matrix provides a visualization of the actual versus predicted values of DT1, DT2, and DT3.

quadrant shows the True Negative (TN) count, with 437 instances correctly identified as negative. The models DT1 and DT2 demonstrate robust predictive capabilities with an accuracy level of 83.15 percent, calculated as the sum of true positives and negatives (1226 + 437 = 1663) divided by the total predictions (1226 + 437 + 59 + 278 = 2000).

The findings articulate the efficacy of CART models in educational settings, with DT2 emerging as particularly powerful due to its high accuracy and remarkable ability to identify relevant cases. The models' distinct performance characteristics suggest tailored applications based on specific predictive needs. This rigorous evaluation supports the continuous refinement and application of CART models in enhancing data-driven decision-making within educational research and policy frameworks.

Analysis

The study systematically gathered data, employing designated collection tools while recognizing the foundational assumptions. This thorough methodology lays a groundwork for investigating the predetermined research questions and hypotheses, providing insights into the contribution of each model to the realm of data-driven educational approaches.

The Random Forest Classifier models, denoted as RF1, RF2, and RF3, have undergone extensive evaluations to ascertain their effectiveness in classifying educational outcomes. This assessment was performed using a dataset split into 7,999 training instances and 2,000 testing instances, focusing on various performance metrics including accuracy, mean cross-validation (CV) score, and standard deviation.

RF2's balanced configuration provided the best performance in terms of accuracy, suggesting that moderate complexity in tree depth and a higher number of estimators are beneficial for capturing nuanced patterns in educational data. The variation in the mean CV scores, with RF2 scoring the highest at 70.12%, highlights its superior generalization across different subsets of the data. Conversely, the higher standard deviation in RF3 suggests more variability in its performance, indicating potential overfitting. All models showed considerable effectiveness with F1 scores reflecting high precision and recall, particularly in RF2, which balanced these aspects most effectively.

Feature importance

Age was consistently deemed the most significant predictor across all models. Attendance and Pell Grant eligibility also emerged as important factors but with varying degrees of influence across the models.

Model	Age	Attendance	Pell	Entry GPA	Ethnicity	Generation
RF1	0.40	0.33	0.15	0.01	.01	.01
RF2	0.46	0.29	0.08	0.02	.01	.01
RF3	0.50	0.22	0.07	0.04	.03	.01

Table 12: Random forest feature importance.

Note: Table 12 provides a quantitative analysis of the feature importance as assessed by three different RF models: RF1, RF2, and RF3.

The analysis confirms the robustness of the Random Forest Classifier in handling educational datasets, with RF2 identified as the most effective model due to its optimal complexity and strong generalizability.

Insights into feature importance underscore the relevance of age, attendance, and financial aid in predicting educational success, providing valuable information for educational institutions to tailor their strategies and interventions.

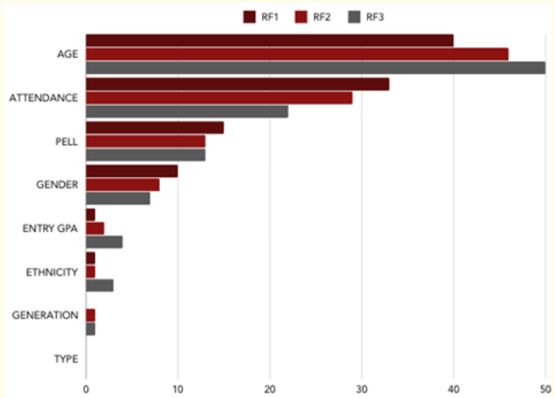

Figure 2: Random forest feature importance comparison.

Gradient-boosting machine analysis

The Gradient-Boosting Machine models, labeled as GBM1, GBM2, and GBM3, were scrutinized to assess their performance based on different configurations of Estimators, Learning Rate, and Max_Depth. GBM2's configuration indicates that a moderate increase in depth and estimators, coupled with a higher learning rate, can slightly enhance model accuracy, suggesting an optimal balance of complexity and learning pace. Despite its complexity, GBM3 exhibited the highest recall at 97.9%, showing its effectiveness in identifying true positives but with a decrease in precision, indicating potential overpredictions.

Consistent across all models, Attendance and Age were identified as the most influential features, highlighting the importance of regular engagement and the maturity of students in predicting educational outcomes. Pell Grant eligibility also emerged as a significant factor in influencing student success. Entry GPA and Gender were deemed less influential in determining graduation outcomes. Provide a clear depiction of feature importance across models, illustrating the weighted significance of each factor in influencing educational success.

Model	Attendance	Age	Pell	Entry GPA	Gender
GBM1	0.38	0.35	0.23	0.01	.01
GBM2	0.38	0.35	0.23	0.01	.01
GBM3	0.38	0.35	0.23	0.01	.03

Table 13: Gradient boosting machine feature importance.

The analysis of the Gradient-Boosting Machine's robustness in managing complex data relations, with each model configuration revealing unique strengths that can be leveraged based on specific educational predictive needs. Insights into feature importance inform institutions about the factors affecting student success, advocating for targeted strategies that emphasize regular attendance and support for older or financially challenged students. This detailed exploration not only validates the applicability of GBM models in educational settings highlights key predictors of student success, thereby providing a foundation for informed educational strategies and policy development.

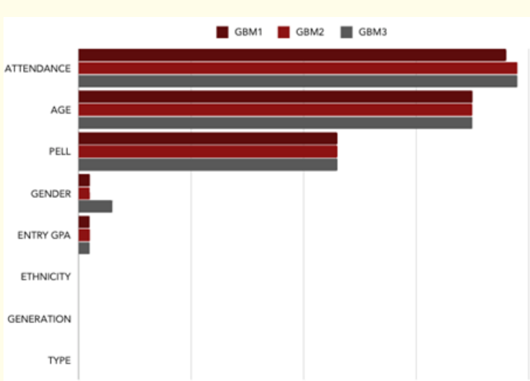

Figure 3: Gradient boosting machine feature importance comparison.

Decision tree model analysis

This analysis examines the Decision Tree (DT) models—DT1, DT2, and DT3—and their effectiveness in predicting educational outcomes. DT2 stands out with the highest mean CV score, indicating superior generalizability across different data segments. DT3, despite its lower accuracy, offers a dependable prediction capability, suggesting that minimal depth with moderate splitting criteria can still yield reliable outcomes, albeit with possible variations in performance that could benefit from further tuning.

Feature importance

The analysis reveals insights into the factors most impactful in predicting academic success:

Attendance: Emerges as a significant predictor across all models, with its importance peaking in DT3 at 53%. Age, follows closely, particularly in DT3 where it accounts for 45% of the model's predictive power, highlighting the influence of students' life stages on their educational trajectories. Consistently, pell grant eligibility is recognized across models as a key factor, reflecting the importance of financial support in achieving academic success.

Attendance and Age consistently appear as top factors, with Pell Grant eligibility also important, but to a lesser extent. The consis-

Model	Attendance	Age	Pell	Entry GPA	Generation	Gender	Ethnicity
DT1	0.39	0.34	0.24	0.03	.02	.02	.01
DT2	0.40	0.36	0.25	0.01	.01	.00	.00
DT3	0.53	0.45	0.01	0.00	.00	.00	.00

Table 14: Decision tree feature importance.

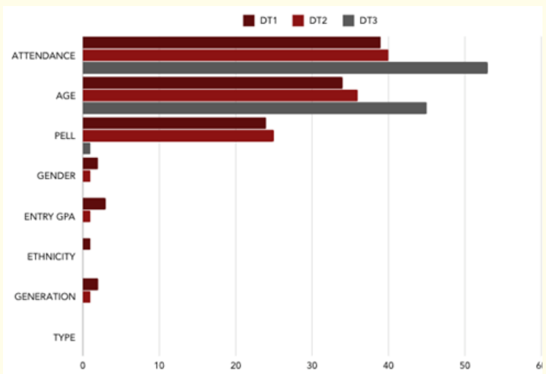

Figure 4: Decision tree machine feature importance comparison.

tency of certain features as key predictors across models suggests areas where educational interventions might be most effective, particularly in enhancing student engagement and support mechanisms.

The nuanced results of Age and Financial aid in the models indicate that tailored approaches, which consider the diversity of student backgrounds and needs, could improve educational outcomes.

Despite their lower influence, the recurring presence of Entry GPA and Gender points to their subtle yet significant part in comprehensive educational assessments.

The Decision Tree analysis reinforces the utility of these models in educational settings, providing a robust tool for predicting student outcomes. The insights gained underscore the importance of addressing both academic engagement and the broader socio-economic factors affecting student success. Institutions can leverage these findings to implement targeted strategies that enhance support systems and improve graduation rates, ultimately fostering an environment where all students can succeed.

Comparison analysis

The comparative analysis of the Random Forest (RF), Gradient-Boosting Machine (GBM), and Decision Tree (DT) models provides a rich understanding of how different machine learning strategies and configurations can impact the prediction of student outcomes in educational settings.

- Random Forest Classifier Analysis: The RF models, particularly RF2, demonstrated high accuracy, achieving 83.5% with a configuration of Max_Depth of 5 and 200 Estimators. This model strikes an optimal balance, effectively capturing generalizable patterns without succumbing to overfitting. The analysis consistently identified Age and Attendance as factors influencing academic success, with Age gaining increased importance in more complex models like RF3. This suggests that both demographic factors and student engagement impact educational outcomes, alongside financial aid considerations such as Pell Grant eligibility.
- Gradient-Boosting Machine Analysis: GBM models, especially GBM2, exhibited slight enhancements in performance due to balanced adjustments in Estimators, Learning Rate, and Max_Depth. GBM2's accuracy of 83.5%, supported by a strategic configuration, reveals its capability to manage intricate data relationships effectively. Across all GBM iterations, Attendance, Age, and Pell Grant eligibility were consistently highlighted as significant predictors, emphasizing the importance of engagement and socioeconomic factors in influencing student success.
- Decision Tree Model Analysis: DT models responded sensitively to parameter modifications, with DT2 showing a balanced recall capability despite variations in precision. This model's ability to identify true positive outcomes while managing false positives shows the importance of parameter tuning. DT3, while placing greater emphasis on Attendance and Age, introduced Gender as a minor but noteworthy predictor, opening a dialogue about the influence of demographic factors on academic outcomes.

The comparative analysis of Random Forest (RF), Gradient-Boosting Machine (GBM), and Decision Tree (DT) models provides valuable insights into the complexities of predicting academic success among adult learners. A key finding from this analysis is the delicate balance needed between model complexity and predictive accuracy. Overly complex models may overfit and lose their ability to generalize to new data, whereas models that are too simple might fail to capture the detailed patterns necessary for accurate forecasting.

Across all models, Age, Attendance, and Pell Grant eligibility consistently emerge as predictors of academic success, underscoring their importance in any educational predictive model. While there is agreement on the significance of these features, variations in their relative importance across different models offer insights into how each model weights these predictors. The introduction of Gender as a predictive factor in the DT models indicates a growing recognition of the impact demographic factors have on educational trajectories, adding complexity to the predictions of academic outcomes.

The recall scores for DT1, DT2, GBM3, and RF1, all above 94 percent, highlight these models' effectiveness in identifying true positive outcomes. GBM1 exhibits the lowest standard deviation (11), indicating it as the most stable and accurate model in this analysis. These findings underscore the robust capabilities of these models to handle diverse data and reflect the nuanced dynamics of adult education, informing targeted interventions and strategies that can enhance academic success in higher education settings.

Conclusions

The Gradient-Boosting Machine (GBM) models demonstrated a high accuracy of 83.5% in predicting degree completion among adult learners, supporting the hypothesis that GBM models are exceptionally effective for educational data analysis. This finding is crucial as it addresses the primary research question of identifying the most accurate machine learning model for educational data analysis. Further analysis confirmed that features such as attendance, age, and Pell Grant eligibility are highly predictive of academic success, aligning with real-world data and supporting the third research question. This alignment not only validates the effectiveness of GBM models in highlighting factors influencing degree completion but also states the importance of these factors in regular data collection by educational institutions.

Limitations

The study's limitations stem primarily from the constraints associated with its higher education context and the age of the data analyzed, which spans a decade. Over such a period, numerous factors may compromise data integrity and relevance. For instance, the definition of an "adult student" may have evolved, reflecting changes in demographic trends, educational policy, or institutional criteria. Additionally, ten years of data accumulation inherently raises concerns regarding human error and inconsistent data management practices, which could include improper categorization or incomplete records. These factors collectively pose significant challenges to the accuracy and applicability of the findings, potentially limiting the generalizability of the study's conclusions to current educational contexts.

Significance of Key Findings

The study's analysis delineates age as a primary determinant of student degree completion, markedly surpassing conventional metrics like entry GPA, modality, and attendance status. This revelation prompts a reconsideration of how predictors of educational success are traditionally viewed, challenging entrenched assumptions that prioritize factors like ethnicity and gender. Such insights necessitate a profound reevaluation of entry requirements and admission policies within higher education institutions. The data-driven approach illuminated by this study advocates for more inclusive and equitable institutional practices. Such practices promise to enhance access to higher education for underrepresented populations, thereby improving enrollment and retention rates. This approach also fosters a conducive learning environment that acknowledges and leverages the life experiences and maturity levels of older students, viewing these attributes not merely as background characteristics but as substantial assets within the educational ecosystem.

The pronounced influence of being a non-traditional, older student underscores the necessity for a strategic reorientation in how educational institutions approach student recruitment and support. This paradigm shift indicates a crucial need for institutions to expand their focus beyond traditional student demographics and develop initiatives tailored to meet the diverse needs of students across varying age ranges. Institutions are encouraged to view the varied life experiences and maturity levels of older students as valuable contributions that enhance the educational ecosystem, rather than mere contextual variables. This perspective shift could lead to the adoption of more holistic educational practices that are sensitive to the unique challenges and strengths of non-traditional students, ultimately fostering a more inclusive and supportive learning environment.

Future Research

Future research should delve deeper into the factors impacting adult student success, particularly focusing on age, attendance, and Pell Grant eligibility. This exploration could extend to qualitative studies interviewing adult learners about their experiences, learning habits, and the cultural and historical influences that shape their educational journey. Such studies could provide richer, more nuanced insights into the variables that impact adult learners differently across various age ranges.

Additionally, there's a compelling opportunity to reassess and potentially challenge previous studies that positioned GPA as a crucial determinant of student success. By applying modern data sets to old methodologies, researchers could explore whether GPA should still be considered a primary factor in predicting student success. This could involve comparing the academic outcomes of traditional versus adult students to assess the relevance of GPA across different student demographics.

A targeted study might explore how GPA impacts the likelihood of college completion on first versus subsequent attempts, particularly for adult learners who have returned to education after significant gaps. This investigation could reveal whether GPA's influence diminishes over time as adult learners accumulate life experiences that may compensate for earlier academic performance.

Expanding the scope of research to include broader demographic factors—such as socioeconomic status, employment, and family responsibilities—combined with age, could offer a holistic view of the student experience. Both quantitative and qualitative methods could be employed to capture the complex interplay of these factors on educational success. Direct engagement with students to capture their perspectives on what influences their academic success would enrich the data.

Furthermore, examining the perceptions of faculty and policymakers could shed light on the receptiveness to using machine learning and data-driven decision-making in educational settings, especially when such approaches challenge established norms. Understanding these perspectives could help gauge the readiness and potential resistance within institutions toward adopting data-driven frameworks.

Outside of feature importance, the rapid technological advancements present an opportunity to evaluate the readiness of higher education institutions to implement and benefit from machine learning. Research could assess the variability in institutional capabilities for data collection and management, identifying gaps that may hinder the effective use of data-driven approaches.

The premise that age significantly determines academic success could transform higher education practices, making them more inclusive and adaptive to the needs of diverse student populations. Future research in this area could lead to a refined understanding of how different life stages affect educational outcomes, supporting the development of policies and practices that better cater to adult learners. This shift toward a more nuanced, data-informed approach to educational decision-making has the potential to profoundly influence policy formulation and institutional operations, enhancing the overall educational landscape for adult students.

Bibliography

1. Rueda E and Swift C. "Academic belonging in higher education". Routledge (2023).
2. Ryan R and Deci E. "Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being". *American Psychologist* 55.1 (2000): 68-78.
3. National Student Clearinghouse Research Center. "Some college, no credential student Outcomes: Annual progress report-academic year-2021/22 (2023).
4. Ivanov A. "Decision trees for evaluation of mathematical competencies in higher education: A case study". *Mathematics* 8.5 (2020): Article 748.
5. Salihoun M. "State of art data mining and learning analytics tools in higher education". *International Journal of Emerging Technologies in Learning* 15.21 (2020): 58.
6. Buenaño-Fernández D., *et al.* "Application of machine learning in predicting performance for computer engineering students: A case study". *Sustainability* 11.10 (2019): Article 2833.
7. Korkmaz C and Correia A. "A review of research on machine learning in educational technology". *Educational Media International* 56.3 (2019): 250-267.
8. Oliveira A., *et al.* "Emerging technologies as pedagogical tools for teaching and learning science: A literature review". *Human Behavior and Emerging Technologies* 1.2 (2019): 149-160.
9. Kannan R. "Advancements in machine learning techniques for educational data mining: An overview of perspectives and trends". *International Journal of Membrane Science and Technology* 10.3 (2023): 1820-1839.
10. Hilbert S., *et al.* "Machine learning for the educational sciences". *Review of Education* 9.3 (2021).
11. Son N., *et al.* "The applications of machine learning in education science research. *VNU Journal of Science: Education Research* 37.4 (2021).
12. Pinto A., *et al.* "How machine learning (ML) is transforming higher education: A systematic literature review". *Journal of Information Systems Engineering and Management* 8.2 (2023): Article 21168.
13. Hessen S., *et al.* "Developing multiagent e-learning system-based machine learning and feature selection techniques". *Computational Intelligence and Neuroscience* Article 2941840 (2022): 1-8.

14. Zhang Y and Teng Z. "Natural language processing: A machine learning perspective". *Computational Linguistics* 48.1 (2022): 233-235.

15. Teng Y., *et al.* "Data-driven decision-making model based on artificial intelligence in higher education system of colleges and universities". *Expert Systems* (2022): 40.

16. Ashaari M., *et al.* "Big data analytics technology capability and data-driven decision making in Malaysian higher education institutions: A conceptual framework". *IOP Conference Series Materials Science and Engineering* 874.1 (2020): Article 012021.

17. Feng L. "Research on higher education evaluation and decision-making based on data mining". *Scientific Programming* (2021): 1-9.

18. Hussain M., *et al.* "Using machine learning to predict student difficulties from learning session data". *The Artificial Intelligence Review* 52.1 (2019): 381-407.

19. Albreiki B., *et al.* "A systematic literature review of student' performance prediction using machine learning techniques". *Education Sciences* 11.9 (2021): Article 552.

20. Chen F and Cui Y. "Utilizing student time series behaviour in learning management systems for early prediction of course performance". *Journal of Learning Analytics* 7.2 (2020): 1-17.

21. Dake D and Buabeng-Andoh C. "Using machine learning techniques to predict learner drop-out rate in higher educational institutions. *Mobile Information Systems* (2022): Article 2670562.

22. Mandinach E and Schildkamp K. "Misconceptions about data-based decision making in education: An exploration of literature. *Studies in Educational Evaluation* 69 (2020): Article 100842.

23. Oyedeji A., *et al.* "Analysis and prediction of student academic performance using machine learning". *Journal of Information Technology and Computer Engineering* 4.1 (2020): 10-15.

24. Roberts L., *et al.* "Student attitudes toward learning analytics in higher education: The Fitbit version of the learning world". *Frontiers in Psychology* 7 (2016).

25. Sadiq M and Ahmed N. "Classifying and predicting students' performance using improved Decision Tree C4.5 in higher education institutes". *Journal of Computer Science* 15.9 (2019): 1291-1306.

26. Gotardo M. "Using Decision Tree algorithm to predict student performance". *Indian Journal of Science and Technology* 12 (2019): 1-8.

27. Kim J., *et al.* "Online higher education for nontraditional adult students: Best cases for public universities. In K. Junghwan, H. Shin, K. Smith and J. Hwang (Eds.), *Research Anthology on Developing Effective Online Learning Courses* (2021): 265-286.

28. Saidani O., *et al.* "Predicting student employability through the internship context using gradient boosting models". *IEEE Access* 10 (2022): 46472-46489.

29. Akmeshe O., *et al.* "Use of machine learning techniques for the forecast of student achievement in higher education". *Information Technologies and Learning Tools* 82 (2021): 297-311.

30. Hasan R., *et al.* "Predicting student performance in higher educational institutions using video learning analytics and data mining techniques". *Applied Sciences* 10.11 (2020): Article 3894.

31. Alsariera Y., *et al.* "Assessment and evaluation of different machine learning algorithms for predicting student performance". *Computational Intelligence and Neuroscience* (2020): 1-11.

32. Beaulac C and Rosenthal J. "Predicting university students' academic success and major using Random Forests". *Research in Higher Education* 60.7 (2019): 1048-1064.

33. Cardona T., *et al.* "Data mining and machine learning retention models in higher education". *Journal of College Student Retention: Research, Theory and Practice* 25.1 (2020): 51-75.

34. Kamal P and Ahuja S. "An ensemble-based model for prediction of academic performance of students in undergrad professional course". *Journal of Engineering, Design and Technology* 17.4 (2019): 769-781.

35. Charilaou P and Battat R. "Machine learning models and overfitting considerations". *World Journal of Gastroenterology* 28.5 (2022): 605-607.

36. Jalota C and Agrawal R. "Feature selection algorithms and student academic performance: A study". In D. Gupta, A. Khanna, S. Bhattacharyya, A. Hassanien, S. Anand, & A. Jaiswal, (Eds.), *International Conference on Innovative Computing and Communications: Vol. 1165. Advances in Intelligent Systems and Computing* (2021): 317-328.

37. Islam M., *et al.* "A comprehensive survey on the process, methods, evaluation, and challenges of feature selection". *IEEE Access* 10 (2022): 99595-99632.

38. Kaneko H. "Interpretation of machine learning models for data sets with many features using feature importance". *Acs Omega* 8.25 (2023): 23218-23225.

39. Marcinkowski F., *et al.* "Implications of AI (un)fairness in higher education admissions: The effects of perceived AI (un-) fairness on exit, voice, and organizational reputation". In *ACM Digital Library (Eds.), Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency. Association for Computing Machinery* (2020).

40. Palacios C., *et al.* "Knowledge discovery for higher education student retention based on data mining: Machine learning algorithms and case study in Chile". *Entropy* 23.4 (2021): Article 485.

41. Arqawi S., *et al.* "Predicting university student retention using artificial intelligence". *International Journal of Advanced Computer Science and Applications* 13.9 (2022).

42. Brdesee H., *et al.* "Predictive model using a machine learning approach for enhancing the retention rate of students at-risk". *International Journal of Semantic Web Information Systems* 18 (2022): 1-21.

43. Zeineddine H., *et al.* "Enhancing prediction of student success: Automated machine learning approach". *Computers and Electrical Engineering* 89.1 (2021): Article 106903.

44. Gaftandzhieva S., *et al.* "Exploring online activities to predict the final grade of student". *Mathematics* 10.20 (2022): Article 3758.

45. Oqaidi K., *et al.* "Towards a students' dropout prediction model in higher education institutions using machine learning algorithms". *International Journal of Emerging Technologies in Learning* 17.18 (2022): 103-117.

46. Jusslin S., *et al.* "Post-approaches to education and the arts: putting theories to work in arts educational practices". *Journal for Research in Arts and Sports Education* 6.3 (2022): 1-10.

47. Nauman M., *et al.* "Guaranteeing correctness of machine learning based decision making at higher educational institutions". *IEEE Access* 9 (20221): 92864-92880.

48. Allen L., *et al.* "Natural language processing as a tool for learning analytics-towards a multi-dimensional view of the learning process. In C. Lang, A. Friend Wise, A. Merceron, D. Gaševic, and G. Siemens (Eds.). Handbook of learning analytics. (2nd edition.) (2022).

49. Nieto Y., *et al.* "Supporting academic decision making at higher educational institutions using machine learning-based algorithms". *Soft Computing* 23 (2019): 4145-4153.

50. Tarmizi S., *et al.* "A review on student attrition in higher education using big data analytics and data mining techniques". *International Journal of Modern Education and Computer Science* 11.8 (2019): 1-14.

51. Yağcı M. "Educational data mining: Prediction of students' academic performance using machine learning algorithms". *Smart Learning Environments* 9.1 (2022).

52. Martins M., *et al.* "A data mining approach for predicting academic success: A case study. In A. Rocha, C. Ferrás, M. Paredes (Eds.), Information Technology and Systems: Vol. 918". *Advances in Intelligent Systems and Computing* (2019): 45-56.

53. Jun-on N., *et al.* "Prediction of students' performance in English using machine learning algorithms". *English Language Teaching* 16.4 (2023): 24.

54. Davari M., *et al.* "Predictive analytics of student performance determinants in education". *World Academy of Science, Engineering and Technology, Open Science Index 191, International Journal of Educational and Pedagogical Sciences* 16.11 (2022): 716-721.

55. Liu J., *et al.* "Fair Compass: Operationalising fairness in machine learning. *IEEE Transactions on Artificial Intelligence* 99 (2023): 1-10.

56. Saltz J., *et al.* "Integrating ethics within machine learning courses". *ACM Transactions on Computing Education* 19 (2019): 1-26.

57. Toms A and Whitworth S. "Ethical considerations in the use of machine learning for research and statistics". *International Journal for Population Data Science* 7.3 (2022).

58. Musso M., *et al.* "Predicting key educational outcomes in academic trajectories: A machine-learning approach". *Higher Education* 80.5 (2020): 875-894.

59. Pande P., *et al.* "Long-term effectiveness of immersive VR simulations in undergraduate science learning: Lessons from a media-comparison study". *Research in Learning Technology* 29 (2021): 1-24.

60. Maravé-Vivas M., *et al.* "A longitudinal study of the effects of service-learning on physical education teacher education students". *Frontiers in Psychology* 13 (2022): 787346.

61. Holenstein M., *et al.* "Transfer effects of mathematical literacy: An integrative longitudinal study". *European Journal of Psychology of Education* 36 (2020): 799-825.

62. Carless D. "Longitudinal perspectives on students' experiences of feedback: A need for teacher-student partnerships". *Higher Education Research and Development* 39 (2020): 425-438.

63. Friedman J. "Greedy function approximation: A gradient boosting machine". *The Annals of Statistics* 29.5 (2001): 1189-1232.

64. Quinlan J. "Induction of Decision Trees". *Machine Learning* 1 (1986): 81-106.

65. Huynh-Cam T., *et al.* "Using Decision Trees and Random Forest algorithms to predict and determine factors contributing to first-year university students' learning performance". *Algorithms* 14 (2021): 318.

66. Raschka S and Mirjalli V. "Python machine learning with PyTorch and Scikit-Learn". Packt Publishing (2022).

67. Breiman L. "Random forests". *Machine Learning* 45 (2001): 5-32.