JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(3), SEPTEMBER, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

# From Features to Fairness: A Thematic Review of Optimized Machine Learning for Student Placement

[1]Ridwan Kamaldeen, [2]Badamasi Imam Ya'u, [3]Kabiru Musa Ibrahim
[1]*Department of Computer Science, Federal Polytechnic Nasarawa, Nigeria*
[2&3]*Department of Computer Science, Abubakar Tafawa Balewa University, Bauchi, Nigeria*

**ABSTRACT**
*Machine learning (ML) is increasingly deployed in academic placement and student transition prediction, yet adoption remains constrained by persistent challenges: high-dimensional and redundant feature spaces, opaque model behaviour, and the frequent absence of disciplined baseline-versus-optimized comparisons. This review synthesizes theoretical foundations and empirical evidence across ML-based educational prediction, with focused attention on feature selection strategies, explainable artificial intelligence (XAI), and the inherent trade-off between predictive accuracy and interpretability. The paper further identifies critical open challenges, including limited dataset generalizability, class imbalance, algorithmic fairness, and the absence of standardized evaluation frameworks, and concludes by proposing actionable directions for future inquiry.*

**ARTICLE INFO**
*Article History*
Received: April, 2026
Received in revised form: May, 2026
Accepted: August, 2026
Published online: September, 2026

**KEYWORDS**
Machine Learning, Feature Selection, Explainable AI, Academic Placement, Educational Data Mining

## INTRODUCTION

Educational institutions are increasingly turning to data-driven decision-making to guide student placement, progression, and stream allocation. Decisions that were once based largely on examination scores and administrative judgement are now frequently informed, or in some cases automated, by machine learning (ML) models trained on cognitive, affective, psychomotor, and demographic indicators (Zawacki-Richter et al., 2022). These models promise greater consistency and predictive power than traditional statistical approaches, and a substantial body of educational data mining (EDM) research has demonstrated their potential to forecast student performance, identify at-risk learners, and support timely intervention (Yağcı, 2022; Hashim et al., 2020).

However, the adoption of ML in high-stakes educational contexts such as academic placement raises two persistent concerns. The first is the problem of high-dimensionality and feature redundancy: educational datasets often contain numerous correlated or irrelevant variables that inflate computational cost, increase the risk of overfitting, and can inadvertently encode bias against particular groups of students (Mehrabi et al., 2021; Chandrashekar & Sahin, 2014). The second is the problem of interpretability. Many of the algorithms that achieve the strongest predictive accuracy, including ensemble methods and neural networks, operate as "black boxes" whose internal reasoning is opaque to the teachers, administrators, and policymakers who must act on their recommendations (Molnar, 2022; Rudin, 2019). In a domain where a single prediction can influence a child's academic trajectory, this lack of transparency is not a peripheral concern but a central barrier to responsible deployment.

This review addresses that gap by synthesizing the theoretical foundations and empirical literature on machine learning applications in academic placement. The specific objectives of the review are to: (i) outline the theoretical foundations underpinning ML-based educational prediction, including the bias-variance trade-off, explainable AI theory, and algorithmic fairness theory; (ii) synthesize, by theme, the empirical literature on classification algorithms, feature selection strategies, and XAI methods applied to student performance and placement

*Corresponding author: Ridwan Kamaldeen*
*Department of Computer Science, Federal Polytechnic Nasarawa, Nigeria.*

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(3), SEPTEMBER, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

prediction; (iii) critically analyze the strengths, weaknesses, and points of convergence or contradiction across the reviewed studies, grounded in an illustrative empirical case; and (iv) identify the persistent challenges and open issues that continue to constrain the field.

## METHODOLOGY

This paper adopts a narrative, thematic literature review methodology, which is appropriate for synthesizing a heterogeneous body of work spanning machine learning algorithms, feature selection techniques, and explainable AI, and for situating these strands within a coherent conceptual framework rather than producing a purely statistical meta-analysis.

### *Search Strategy and Sources*

Relevant literature was identified through a combination of academic databases and indexing platforms, including Scopus, IEEE Xplore, ScienceDirect, SpringerLink, and Google Scholar. Search terms were combined using Boolean operators around three conceptual clusters: (a) machine learning and student performance/placement prediction (e.g., "machine learning", "educational data mining", "student placement", "academic prediction"); (b) feature selection (e.g., "feature selection", "filter method", "wrapper method", "embedded method", "recursive feature elimination", "hybrid feature selection"); and (c) interpretability and fairness (e.g., "explainable AI", "XAI", "SHAP", "LIME", "algorithmic fairness", "bias in machine learning").

### *Inclusion and Exclusion Criteria*

Studies were included if they: (i) applied supervised machine learning to a student-related prediction or placement task; (ii) addressed feature selection, model interpretability, or both; (iii) were published, in the majority, between 2014 and 2025, to capture both foundational theory and recent empirical advances; and (iv) were available in English with sufficient methodological detail to support thematic extraction. Studies were excluded where they addressed unrelated domains, offered no empirical or theoretical contribution to feature optimization or interpretability, or could not be verified for methodological soundness.

### *Thematic Synthesis and Illustrative Case*

Selected works were screened by title and abstract, then read in full and coded manually into recurring themes: classification algorithms, feature selection strategies, explainable AI, and fairness/ethics.

### Theoretical and Foundational Background

The reviewed literature converges on three theoretical pillars that together frame the design and evaluation of ML-based academic placement systems: the bias-variance trade-off, explainable AI theory, and algorithmic fairness theory.

### *The Bias-Variance Trade-off*

The bias-variance trade-off describes the balance between model complexity and generalisation performance. Excessive features or overly flexible models increase variance and the risk of overfitting to training data, while an impoverished feature set increases bias and reduces predictive power (Molnar, 2022). Feature selection is theoretically positioned as a mechanism for finding the optimal point on this trade-off, and the Structural Risk Minimisation (SRM) principle that underlies Support Vector Machines formalises this by simultaneously maximising the margin between classes and penalising model complexity through a regularisation parameter.

### *Explainable AI Theory*

Explainable AI (XAI) theory holds that automated systems operating in high-stakes domains, including education, must provide understandable justifications for their decisions (Rudin, 2019). Rudin (2019) argues that reliance on post-hoc explanation tools for inherently opaque models may be insufficient in high-stakes settings, and that inherently interpretable models should be preferred wherever feasible. Molnar (2022) complements this position by detailing model-agnostic, post-hoc explanation techniques, notably SHapley Additive exPlanations (SHAP)

*Corresponding author: Ridwan Kamaldeen*
✉
*Department of Computer Science, Federal Polytechnic Nasarawa, Nigeria.*

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(3), SEPTEMBER, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

and Local Interpretable Model-agnostic Explanations (LIME), which can be layered onto complex models to approximate feature-level justification even when full interpretability is not built into the model architecture.

***Algorithmic Fairness Theory***

Algorithmic fairness theory addresses the risk that automated decision systems reproduce or amplify discriminatory outcomes, particularly against specific demographic groups (Mehrabi et al., 2021). In educational placement, this theory underscores the importance of scrutinising which features are permitted to influence predictions, since the inclusion of sensitive or proxy variables can distort outcomes even when the model appears statistically accurate. Fairness-aware feature selection and evaluation metrics such as demographic parity and equal opportunity are proposed as practical mechanisms for operationalising this theory.

***Feature Selection Taxonomy***

The literature consistently classifies feature selection methods into three families (Chandrashekar & Sahin, 2014). Filter methods, such as the Pearson correlation coefficient and mutual information, evaluate feature relevance independently of any learning algorithm and are computationally inexpensive but ignore feature interactions. Wrapper methods, such as Recursive Feature Elimination (RFE), iteratively train and evaluate a model to identify the optimal feature subset, capturing interactions at greater computational cost. Embedded methods, such as Lasso (L1-regularised) regression, integrate feature selection into the model training process itself, offering a compromise between the two. Hybrid frameworks that combine filter, wrapper, and embedded stages, as employed in the illustrative case reviewed here, are increasingly favoured as a way of balancing computational efficiency against predictive robustness.

**Thematic Synthesis**

***Classification Algorithms in Educational Prediction***

A wide range of supervised learning algorithms has been applied to student performance and placement prediction. Decision Trees remain popular for their interpretability and their tree-like, rule-based structure that is comprehensible to non-technical stakeholders (Mustapha et al., 2023). Artificial Neural Networks (ANNs) are valued for their capacity to model complex, non-linear interactions among predictors, though this flexibility comes at the cost of transparency (Olabanjo & Wusu, 2022). Logistic Regression continues to be widely used for binary placement outcomes because of its simplicity and interpretable coefficients (Hashim et al., 2020). Naïve Bayes classifiers, despite their simplifying independence assumption, often perform surprisingly well on educational classification tasks and scale efficiently to large datasets (Beckham et al., 2022).

Support Vector Machines (SVMs), particularly the linear variant, are consistently reported as strong performers on structured, moderately correlated educational data because they explicitly construct a maximum-margin hyperplane rather than relying on axis-aligned partitions. Random Forests and other ensemble techniques, bagging, boosting, voting, and stacking, combine multiple weak learners to achieve higher accuracy than any single constituent model (Al Husaini et al., 2022; Bujang et al., 2021), and have been widely applied to student performance prediction (Çetinkaya et al., 2023; Ghorbani & Ghousi, 2020). K-Nearest Neighbours (KNN), by contrast, is reported to be highly sensitive to noisy or poorly scaled features, a limitation that becomes particularly important once feature selection is considered (Hashim et al., 2020).

A recurring finding across this literature is that no single algorithm is universally superior; rather, performance is contingent on the structure of the underlying data. Yağcı (2022), for example, achieved accuracy ranging from 50% to 81% across Random Forest, Logistic Regression, KNN, and SVM depending on demographic and behavioural feature composition, while Farhood et al. (2024) found Gradient Boosting Machines most

*Corresponding author: Ridwan Kamaldeen*

*Department of Computer Science, Federal Polytechnic Nasarawa, Nigeria.*

effective when prior academic performance and attendance were the dominant predictors.

***Feature Selection Strategies in Educational ML Pipelines***

Feature selection emerges across the literature as a critical, though frequently under-optimised, stage of the ML pipeline. Ahmed and Khan (2023) combined filter and wrapper techniques to improve classification accuracy and reduce computational time relative to traditional approaches, although their study offered limited analysis of how the selected features influenced downstream decision-making. Zhou et al. (2024) applied dimensionality reduction techniques, including Principal Component Analysis (PCA) and RFE, and found that reduced dimensionality minimised overfitting and improved generalisation, but noted that PCA's transformed features undermined interpretability, a direct illustration of the tension between the bias-variance and explainability goals discussed in Section 3.

In the illustrative case reviewed here, a three-stage hybrid pipeline was constructed: filter-based screening using Pearson correlation and mutual information to remove weakly relevant or multicollinear features; wrapper-based Recursive Feature Elimination tied directly to the target LinearSVM classifier to capture feature interactions; and embedded Lasso (L1) regularization during model training to enforce sparsity. This layered design reflects the broader consensus in the literature that hybrid approaches, rather than any single method in isolation, tend to yield the best balance of efficiency and predictive robustness (Chandrashekar & Sahin, 2014; Ahmed et al., 2024).

***Explainable AI and Interpretability in Educational Systems***

A second major theme concerns the growing integration of XAI techniques into educational prediction systems. García et al. (2024) incorporated SHAP-based explanations into predictive models for learning analytics and reported improved transparency and stakeholder trust, though without a systematic feature optimisation process. Nakamura et al. (2024) integrated SHAP explanations into a Gradient Boosting model for university admissions prediction, improving stakeholder trust but without a structured comparison between baseline and optimised feature sets. Okafor and Adeyemi (2025) developed a neural network-based academic track allocation system that achieved strong predictive performance but was explicitly acknowledged by its authors as lacking inherent interpretability, illustrating the accuracy-transparency trade-off that recurs throughout the field.

XAI methods are typically classified along three axes: whether they are model-agnostic or model-specific, whether they offer local (single-prediction) or global (whole-model) interpretability, and whether interpretability is achieved intrinsically (through model design, e.g., decision trees) or post-hoc (through external tools such as SHAP or LIME applied after training) (Molnar, 2022). The literature suggests that intrinsically interpretable models, such as Linear SVM and Logistic Regression, offer a practical middle ground: they retain competitive predictive accuracy on structured educational data while remaining more transparent than ensemble or deep learning alternatives, reducing (though not eliminating) the need for supplementary post-hoc explanation tools.

***Fairness, Bias, and Ethical Considerations***

Mehrabi et al. (2021) provide a comprehensive account of how improper feature selection and biased training data can produce discriminatory outcomes in automated decision systems, a concern that is particularly acute in educational placement, where biased predictions can disadvantage specific demographic groups. Ibrahim and Yusuf (2025) applied fairness metrics such as demographic parity and equal opportunity to student placement models and found that feature selection choices materially influenced fairness outcomes, while Kumar and Patel (2024) demonstrated that removing sensitive features and applying fairness constraints improved equity, albeit with limited accompanying interpretability

*Corresponding author: Ridwan Kamaldeen*

*Department of Computer Science, Federal Polytechnic Nasarawa, Nigeria.*

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(3), SEPTEMBER, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

analysis. Holmes and Tuomi (2022) situate these concerns within a broader ethical framework, arguing that AI systems deployed in education must be designed with accountability and fairness as first-order requirements rather than after-the-fact considerations.

## DISCUSSION OF FINDINGS

First, the evidence strongly suggests that model optimization, rather than algorithmic sophistication in isolation, is frequently the decisive factor in predictive performance. The gap between the conventional ML models and the optimized counterparts corroborates the argument advanced by Probst et al. (2019) and Bergstra and Bengio (2012) that proper hyperparameter tuning can transform a mediocre model into a high-performing one. This finding complicates a common pattern in the reviewed literature, in which studies report the performance of a single, tuned model without a disciplined comparison against an untuned baseline (Li et al., 2023; Brown et al., 2025), making it difficult to attribute reported gains to feature selection, algorithm choice, or optimization effort.

Second, Al Braiki et al. (2023) observed that well-tuned linear classifiers often outperform more complex ensembles when features exhibit multicollinearity, since trees must construct unnecessarily deep, piecewise-constant partitions to approximate an underlying linear decision boundary, increasing variance and overfitting risk. This contrasts with findings such as Olabanjo and Wusu (2022) and Farhood et al. (2024), who reported ensemble superiority on datasets with heterogeneous, high-dimensional, or non-linear features. The apparent contradiction is resolved once dataset structure is taken into account: algorithm superiority in educational ML is data-dependent rather than universal, reinforcing calls in the literature for correlation and separability analysis to precede model selection (Zhou et al., 2024).

Third, the interpretability-performance trade-off identified by Rudin (2019) and Molnar (2022) is visible throughout the thematic synthesis but is only partially resolved by the reviewed studies. Where PCA-based dimensionality reduction improved performance at the cost of interpretability (Zhou et al., 2024), and neural network models achieved strong accuracy while remaining acknowledged black boxes (Okafor & Adeyemi, 2025), the illustrative case suggests that this trade-off is not absolute: an intrinsically interpretable linear model, when properly optimized and paired with a hybrid feature selection pipeline, can achieve competitive accuracy without sacrificing transparency. This supports Rudin's (2019) position that inherently interpretable models deserve greater consideration before defaulting to post-hoc explanation of black-box systems, though it should be noted that the illustrative case used a compact, three-feature input space; the generalisability of this finding to higher-dimensional, more heterogeneous educational datasets remains an open question.

Fourth, while several reviewed studies address fairness (Mehrabi et al., 2021; Ibrahim & Yusuf, 2025; Kumar & Patel, 2024) and several address interpretabilities (García et al., 2024; Nakamura et al., 2024), few studies integrate feature optimization, interpretability, and fairness evaluation within a single, structured comparative framework. This fragmentation constitutes the central research gap identified across the synthesized.

### Challenges and Open Issues

Several challenges recur across the reviewed literature and remain only partially resolved by current empirical work.

#### *Dataset generalizability*

Most studies, including the illustrative case reviewed here, rely on a single institutional or platform-sourced dataset, limiting confidence that reported performance gains will transfer to other schools, regions, or educational systems with different student populations and assessment frameworks.

#### *Class imbalance*

Educational outcome variables, such as transition or dropout status, are frequently imbalanced in real-world settings. While

*Corresponding author: Ridwan Kamaldeen*
*Department of Computer Science, Federal Polytechnic Nasarawa, Nigeria.*

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(3), SEPTEMBER, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

techniques such as class weighting, SMOTE, and cost-sensitive learning are proposed as mitigations, their effectiveness varies across studies and datasets, and imbalance continues to distort model evaluation when not explicitly addressed.

***Lack of ML model optimization***

As noted in Section 5, many studies report only a final, tuned model without a documented baseline, making it difficult to isolate the contribution of feature selection from that of hyperparameter optimization or algorithm choice.

***Interpretability at scale***

While intrinsically interpretable models perform well on compact feature sets, it remains unclear how well this advantage persists as feature spaces grow to include socioeconomic, behavioural, and institutional variables, which several studies recommend incorporating for more comprehensive prediction.

***Fairness evaluation is inconsistently applied***

Formal fairness metrics such as demographic parity and equal opportunity are used in only a minority of the reviewed studies, and rarely alongside feature selection and interpretability analysis in the same framework.

***Longitudinal and external validation is largely absent***

Nearly all reviewed studies, including the illustrative case, evaluate models on a single, static dataset snapshot rather than tracking performance across multiple academic sessions or validating on independently collected data, leaving open the question of model drift over time.

***Reproducibility and computational cost***

Wrapper methods such as RFE and repeated cross-validation regimes offer robustness but at meaningfully higher computational cost than filter methods, a trade-off that is not always transparently reported alongside performance gains.

**CONCLUSION**

This review synthesizes the theoretical and empirical literature on machine learning applications in academic placement prediction, with a focus on feature selection, explainable AI, and fairness. The review identifies a persistent gap: relatively few studies integrate feature optimization, interpretability, and fairness evaluation within a single framework. Thus, external and longitudinal validation across diverse educational settings remains rare.

Future research should prioritize on the development of a unified evaluation framework that report predictive performance, feature-selection rationale, interpretability outcomes (via SHAP, LIME, or comparable tools), and fairness metrics; the extension of feature spaces to include socioeconomic, behavioral, and institutional variables without eroding interpretability; and multi-institutional, longitudinal validation to establish the generalizability and durability of predictive models before they are deployed in real-world academic placement decisions.

**REFERENCES**

Ahmed, R., & Khan, S. (2023). A hybrid feature selection model combining filter and wrapper techniques for educational data classification. Journal of Educational Data Mining, 15(2), 45–62.

Ahmed, T., Bello, A., & Yusuf, K. (2024). Combining genetic algorithms with recursive feature elimination for educational dataset optimisation. International Journal of Machine Learning Applications, 9(1), 22–38.

Al Braiki, B., Rustamov, Z., Rustamov, J., & Zaki, N. (2021). Enhancing predictive performance in identifying at-risk students: Integration of topological features and node embeddings in machine learning models.

Al Husaini, Y., Syufiza, N., Shukor, A., Said, Y. N., & Husaini, A. (2022). Factors affecting students' academic performance: A review (Vol. 12). https://www.researchgate.net/publication/367360842

Corresponding author: Ridwan Kamaldeen
Department of Computer Science, Federal Polytechnic Nasarawa, Nigeria.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 14(3), SEPTEMBER, 2026
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbufstejoste.com

Albreiki, B., Zaki, N., & Alashwal, H. (2021). A systematic literature review of student performance prediction using machine learning techniques. Education Sciences, 11(9). https://doi.org/10.3390/educsci11090552

Almeida, F., & Costa, R. (2023). Comparing logistic regression with ensemble methods in placement predictions. Journal of Applied Educational Analytics, 7(3), 112–128.

Beckham, N. R., Akeh, L. J., Mitaart, G. N. P., & Moniaga, J. V. (2022). Determining factors that affect student performance using various machine learning methods. Procedia Computer Science, 216, 597–603. https://doi.org/10.1016/j.procs.2022.12.174

Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. Journal of Machine Learning Research, 13, 281–305.

Brown, S., Ahmed, L., & Wei, T. (2025). A comparative study of machine learning models in education. Journal of Educational Technology Systems, 53(1), 61–79.

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

Çetinkaya, A., Baykan, Ö. K., & Kırgız, H. (2023). Analysis of machine learning classification approaches for predicting students' programming aptitude. Sustainability (Switzerland), 15(17). https://doi.org/10.3390/su151712917

Chandrashekar, G., & Sahin, F. (2014). A survey on feature selection methods. Computers & Electrical Engineering, 40, 16–28. https://doi.org/10.1016/j.compeleceng.2013.11.024

Chen, Y., Li, M., & Zhao, Q. (2024). Recursive feature elimination combined with SHAP explanations for educational prediction. Expert Systems with Applications, 238, 121–134.

Davis, K., Mensah, P., & Okoro, J. (2025). A comparative framework evaluating baseline and optimised machine learning models in education. Journal of Learning Analytics, 12(1), 88–104.

Farhood, H., Joudah, I., Beheshti, A., & Muller, S. (2024). Evaluating and enhancing artificial intelligence models for predicting student learning outcomes. Informatics, 11(3), 46. https://doi.org/10.3390/informatics11030046

García, M., López, R., & Fernández, A. (2024). Explainable AI methods in learning analytics systems. Computers & Education: Artificial Intelligence, 6, 100–112.

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

Gupta, V., Singhal, P., & Khattri, V. (2024). Analysis of student academic performance using machine learning algorithms: A study. https://doi.org/10.5281/zenodo.10851854

Hashim, A. S., Awadh, W. A., & Hamoud, A. K. (2020). Student performance prediction model based on supervised machine learning algorithms. IOP Conference Series: Materials Science and Engineering, 928(3). https://doi.org/10.1088/1757-899X/928/3/032019

Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. European Journal of Education, 57, 542–570. https://doi.org/10.1111/ejed.12533

Ibrahim, A., & Yusuf, K. (2025). Applying fairness metrics to student placement models. Journal of Fair and Ethical AI, 3(1), 15–29.

Idris, R., Govindasamy, P., Nachiappan, S., & Bacotang, J. (2023). Exploring the impact of cognitive factors on learning, motivation and career in Malaysia's STEM education. International Journal of Academic Research in Business and Social Sciences, 13(6). https://doi.org/10.6007/ijarbss/v13-i6/17227

Kumar, R., & Patel, S. (2024). Fairness-aware machine learning techniques in student selection systems. Journal of Educational AI Ethics, 2(2), 55–70.

Li, X., Wang, H., & Zhang, F. (2023). Student performance prediction using Random Forest and Support Vector Machines. Journal of Educational Computing Research, 61(4), 789–806.

*Corresponding author: Ridwan Kamaldeen*

*Department of Computer Science, Federal Polytechnic Nasarawa, Nigeria.*

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. ACM Computing Surveys, 54, 1–35. https://doi.org/10.1145/3457607

Molnar, C. (2022). Interpretable machine learning: A guide for making black box models explainable (Self-published). https://christophm.github.io/interpretable-ml-book/

Mustapha, R., Uju, A. S., Haruna, K., Lawal, I. A., Abubakar, R., & Jibril, A. U. (2023). Placement model for students into appropriate academic class using machine learning. Journal of Science & Technology, 7(1), 170–183. https://doi.org/10.56471/slujst.v7i

Nakamura, T., Sato, K., & Yamamoto, R. (2024). Integrating SHAP explanations into a gradient boosting model for university admissions prediction. Journal of AI in Higher Education, 4(1), 33–47.

Ofori, F., Maina, E., Gitonga, R., & Elizaphan Maina, D. (2024). Using machine learning algorithms to predict students' performance and improve learning outcome: A literature-based review. Stratford Peer Reviewed Journals and Book Publishing Journal of Information and Technology, 4, 2616–3573.

Ojo, T., & Mensah, K. (2025). The role of data preprocessing in improving academic prediction models. African Journal of Educational Technology, 6(1), 41–56.

Okafor, C., & Adeyemi, B. (2025). A neural network-based system for academic track allocation in tertiary institutions. Journal of Computing in Higher Education, 37(1), 101–119.

Olabanjo, O., & Wusu, A. (2022). A machine learning prediction of academic performance of secondary school students using radial basis function neural network. https://doi.org/10.20944/preprints202207.0088.v1

Oliveira, T., Costa, M., & Ferreira, P. (2023). The impact of correlated features on educational predictions. Journal of Data Science in Education, 5(2), 77–91.

Probst, P., Boulesteix, A.-L., & Bischl, B. (2019). Tunability: Importance of hyperparameters of machine learning algorithms. Journal of Machine Learning Research, 20, 1934–1965.

Rahman, M., Islam, S., & Chowdhury, N. (2024). Comparing filter-based and wrapper-based feature selection methods in academic datasets. Journal of Educational Informatics, 8(2), 60–75.

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence, 1, 206–215. https://doi.org/10.1038/s42256-019-0048-x

Singh, A., & Verma, R. (2024). Application of gradient boosting algorithms for student placement prediction. Journal of Machine Learning in Education, 3(1), 18–32.

Smith, J., & Taylor, L. (2025). Evaluating interpretable boosting machines in academic classification tasks. Journal of Explainable AI Applications, 2(1), 5–19.

Thomas, R., & Clark, D. (2023). Overfitting challenges in educational datasets. Educational Data Science Review, 4(2), 91–105.

Wang, L., Chen, X., & Liu, Y. (2024). Model compression techniques to simplify complex predictive systems. Journal of Applied Machine Learning, 6(1), 24–39.

Yağcı, M. (2022). Educational data mining: Prediction of students' academic performance using machine learning algorithms. Smart Learning Environments, 9(1). https://doi.org/10.1186/s40561-022-00192-z

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2022). Systematic review of research on artificial intelligence applications in higher education – Where are the educators? International Journal of Educational Technology in Higher Education, 19(1), 1–27. https://doi.org/10.1186/s41239-021-00249-8

Zhou, H., Fang, J., & Xu, K. (2024). Dimensionality reduction techniques in educational datasets: PCA and recursive feature elimination. Journal of Learning Analytics and Data Science, 5(1), 29–46.

*Corresponding author: Ridwan Kamaldeen*

*Department of Computer Science, Federal Polytechnic Nasarawa, Nigeria.*