This article has been accepted for publication in a future proceedings of this conference, but has not been fully edited. Content may change prior to final publication. Citation information: DOI: 10.1109/ICCIAA68481.2026.11544019, 2026 2nd International Conference on Computational Intelligence Approaches and Applications (ICCIAA)

AI-Driven Student Performance Prediction in Higher Education Using Multi-Source Educational Data

Mohammad AlShaikh-Hasan [12]
[1] *Computer Science Department*
University of Petra
Amman, Jordan
MAlSheikh@uop.edu.jo
[2] *Computer Science Department*
Brunel University London
London, UK
Mohammad.AlShaikhHasan@brunel.ac.uk

Alessandro Pandini
Computer Science Department
Brunel University London
London, UK
Alessandro.Pandini@brunel.ac.uk

XiaoHui Liu
Computer Science Department
Brunel University London
London, UK
XiaoHui.Liu@brunel.ac.uk

Gheorghita Ghinea
Computer Science Department
Brunel University London
London, UK
George.Ghinea@brunel.ac.uk

***Abstract*— Predicting how students will perform remains a central challenge in educational data mining and learning analytics. Anticipating difficulties early allows instructors to refine feedback and plan interventions to keep learners on track. Our study describes an AI driven framework for predicting academic success using a multi layered dataset collected at the University of Petra's IT College in the 2022 2023 year. The data contain 952 anonymised records and integrate five complementary data layers: registration details, aggregated course-level intended learning outcomes (CILOs), Learning Management System (LMS) behaviours data, placement test scores and program-level intended learning outcomes (PILOs). We outline a rigorous preprocessing and feature engineering pipeline and benchmark eight traditional machine learning classifiers against two neural network models. Class imbalance is mitigated by random oversampling, and hyper parameters are tuned systematically. Gradient boosted tree ensembles deliver the strongest results, achieving around 85 % accuracy and an area under the curve of approximately 0.96, while long short term memory and convolutional models perform slightly worse. TreeSHAP analyses reveal that course outcomes and placement tests are the most influential predictors. We situate these results within the literature and discuss implications for timely interventions and fair, interpretable analytics. The study's contributions include (A) the integration of five complementary educational data layers and (B) the explicit modelling of CILOs and PILOs in student-performance prediction.**

Keywords*— *Student Performance Prediction, Machine Learning, Deep Learning, Long Short-Term Memory (LSTM), Convolutional Neural Networks (CNN), XGBoost, Logistic Regression, Random Forest, Support Vector Machines (SVM), Educational Data Mining, Intended Learning Outcomes (ILOs).

I. INTRODUCTION

Improving student success in higher education is a priority shared by educators, administrators and policymakers worldwide. Retention rates, graduation rates and overall academic achievement influence institutional reputation, accreditation outcomes and societal prosperity. Yet reliably forecasting which students will excel, struggle or withdraw remains challenging because academic trajectories are shaped by complex interactions among demographic background, prior knowledge, motivation, learning behaviour and curricular design [1]. Advances in educational data mining (EDM) and learning analytics promise to illuminate these interactions by converting raw digital traces into actionable insights. EDM methods analyse a variety of data sources—demographics, grades, click logs, forum posts, learning-management-system events and survey responses—to uncover patterns that would otherwise remain hidden [1]. A recent systematic literature review of 62 studies reported that regression and other supervised machine-learning models are the most commonly used approaches for student-performance prediction and that online activity, term grades and student emotions are among the dominant predictors [2]. Beyond the scope of individual predictive tasks, EDM draws on a wide range of analytic techniques—including classification, regression, clustering, association rule mining and natural language processing—to extract knowledge from large educational datasets [3]. A more recent review emphasised that modern machine-learning algorithms generally outperform traditional statistical methods when forecasting student performance, yet they also introduce challenges related to bias, transparency and generalisability [4].

Over the last decade, machine-learning models have been deployed to predict student outcomes ranging from course dropout to final grades. Prior studies have applied traditional machine learning models for student performance prediction based on demographics, academic, and behavioural attributes, often achieving reasonable accuracy on small datasets [5]. More recent work harnesses behavioural indicators based on the current and previous clickstreams to early improve predictive power [6], and deep-learning architectures such as long short-term memory networks (LSTMs) and convolutional neural networks (CNNs) are beginning to appear in the literature [7], [8]. Despite these advances, two limitations persist: first, many studies treat learning outcomes as black boxes rather than explicitly incorporating course- or program-level objectives; and second, there is little consensus on whether sophisticated deep models genuinely outperform well-tuned, interpretable tree-based learners. Our research addresses both issues by integrating CILOs and PILOs as explicit predictors and by systematically comparing ML and DL models under identical experimental conditions.

Another impetus for this work is the need for early warning. Several large-scale analyses have demonstrated that student engagement patterns in the early weeks of a course are strongly correlated with final outcomes [6], [9], [10]. Models

that can identify at-risk students by mid-term enable instructors to offer timely support, potentially reducing dropout rates. However, early-prediction models often face class imbalance because the number of high-achieving students far exceeds those at risk. Failure to address this imbalance can bias models toward the majority class and obscure important signals. In our framework, we apply random oversampling to the minority class to ensure fair learning. Moreover, we adopt evaluation metrics beyond accuracy—including precision, recall, F1-score and AUC—to provide a holistic picture of model performance. Fairness and interpretability are also emphasised; algorithms used in education must respect equity and transparency [11]. We therefore examine model output through confusion matrices and SHAP (SHapley Additive exPlanations) values to understand which features drive predictions and to detect any inadvertent bias.

The remainder of this paper is organised as follows. Section II reviews related work on student-performance prediction and highlights gaps that our study addresses. Section III describes the multi-layer dataset, feature construction and target definition. Section IV outlines the modelling methodology, including preprocessing, oversampling, model development and hyper-parameter tuning. Section V presents the experimental results and comparative analysis. Section VI discusses the findings in light of existing literature, interprets model behaviour through feature importance. Section VII concludes with key contributions and future research directions.

II. Related Works

The earliest predictive models for student success relied on relatively simple statistical techniques. Logistic regression has been widely used because of its interpretability and solid theoretical grounding. For example, a comparative analysis of machine-learning models for online programming courses reported that logistic regression achieved a high area under the curve (AUC $\approx$ 0.935) when predicting pass/fail outcomes [1]. Another study on MOOC learners found that logistic regression achieved 72.1 % accuracy for a four-class problem and 92.4 % accuracy for a binary pass/fail classification, outperforming several baselines [12]. Yet logistic regression assumes linear relationships between predictors and outcomes and struggles with complex feature interactions. Random forests and other ensemble methods offer non-linear decision boundaries and can automatically capture interactions; another study using random forest on learning-behaviour data obtained 98.15 % accuracy in classifying student performance [13]. Similarly, research on data-mining techniques for predicting exam performance indicated that random forest and decision tree classifiers correctly classified 250 of 253 test instances [14]. Tree-based models have therefore emerged as strong contenders due to their predictive accuracy and interpretability via feature importance.

Feature selection and engineering play a vital role in traditional ML. A study employing Moodle LMS activity features and Zoom lecture-attendance indicators achieved 78% accuracy with a random forest model when predicting final course grades, underscoring the value of behavioural engagement signals for early identification of at-risk students [15]. Another dataset of 830 undergraduate students revealed that including mid-term, practical exam and final assessment scores led to improved predictions [14]. However, these studies often used a limited set of features such as demographics or grades and rarely incorporated higher-level learning outcomes like CILOs or PILOs.

With the rising popularity of deep learning, researchers have begun exploring neural architectures for student-performance prediction. Hybrid models combining convolutional and recurrent layers can capture sequential patterns in clickstream data and identify local patterns in learning behaviours. For example, a hybrid CNN–LSTM model achieved a very high accuracies on two publicly available datasets, outperforming all traditional ML baselines [7]. Another study proposed an ANN–LSTM model to predict MOOC outcomes and reported that it outperformed recurrent neural networks and gated recurrent units by 6–14 % [6]. A Bi-LSTM framework with SHAP-based interpretability achieved average accuracy of 88.23 % and statistically outperformed CatBoost, XGBoost and other baselines [16]. These results demonstrate the potential of deep architectures; however, such models require large training datasets to avoid overfitting and often lack interpretability. On our medium-sized dataset of 952 students, we find that deep models provide competitive recall but generally lower precision than ensemble trees, suggesting that dataset size and feature diversity influence whether deep models are advantageous.

Ensemble and hybrid models combine the strengths of various algorithms. A multi-output hybrid integrated model using XGBoost and gradient boosting predicted mid-term and final grades with an accuracy of 78.37 %, outperforming individual models by 3–8 % [17]. Graph neural networks have also been applied; a recent study introduced a GNN–Transformer–InceptionNet (GNN-TINet) and reported 98.5 % accuracy and 97.8 % F1-score on multi-label student-performance prediction [18]. In our work, we evaluate both deep and tree-based ensembles to assess their relative merits on a medium-sized dataset. Our study adopts SHAP values to visualise feature contributions and discusses fairness in the context of educational interventions..

III. Data Description

Our dataset comprises 952 anonymised student records gathered from the IT College at the University of Petra for the 2022 2023 academic year. Each record is linked through an anonymised identifier enabling us to merge information from multiple systems. **Table I** summarises the five data layers and the variables drawn from each. The registration layer includes socio demographic attributes (e.g., major, nationality, gender), high school background and enrolment details. The course level intended learning outcomes (CILOs) layer aggregates attainment measures for each student across courses taken in the semester (e.g., ILO_mean, ILO_sum, ILO_max, ILO_min). The learning management system (LMS) behavioural layer captures engagement and assessment signals such as assignment scores, quiz statistics and attendance percentages. Placement tests provide entry level English and computing grades, and the programme level PILOs layer aggregates attainment indicators across knowledge, intellectual, practical and transferable skills.

TABLE I. THE FIVE LAYERS AND THEIR ASSOCIATED VARIABLES

Layer	Included Variables
1) Registration records (REG)	Major, nationality, gender, high-school type/country/score/year, enrolment year/semester, age, credit hours registered/completed (and GPA target).
2) Course-level ILOs (CILOs)	Student-level aggregates of course-level attainment values (e.g., ILO_mean, ILO_sum, ILO_max, ILO_min).
3) LMS behavioural indicators (Student activities)	Aggregated engagement and assessment signals (e.g., assign_mean, quiz_mean, quiz_count, present_percentage, absent_percentage).
4) Placement tests	Entry-level grades from English and computer placement tests (EGrade, CGrade).
5) Program-level PILOs (PILOs)	Program-wide attainment indicators (K/I/P/T attainment averages).

A wide range of machine-learning methods has been applied to analyse these complex, multi-layer educational datasets [19]. **Figure 1** illustrates how these layers interconnect. We aggregated micro level events at the student level using mean, sum and count operators, then merged the tables using the anonymised student identifier. Continuous missing values were imputed with the median, whereas categorical gaps were encoded as a distinct category and then one hot encoded. Preliminary analysis indicated an imbalanced target variable (final GPA category), with roughly 61 % of students in the "high" class, 28 % in the "medium" class and 11 % in the "low" class, necessitating special handling during modelling.

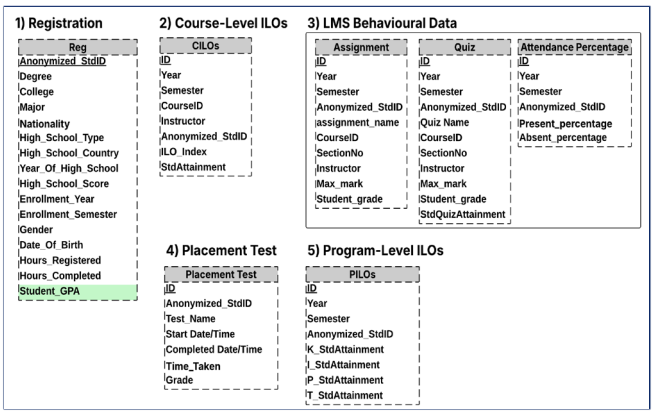

Fig. 1. Five Layers Schema

The predictive task is formulated as a three-class classification problem. Each student's cumulative GPA (scaled on a 4.0 system) is discretised into "low", "medium" and "high" categories based on thresholds determined by academic regulations: 0–1.99, 2.0–2.99 and 3.0–4.0, respectively. These intervals correspond to institutional standards for academic probation, satisfactory progress and honours. The aim is to predict the GPA category at the end of the academic year based on information available up to mid-term. Early prediction facilitates timely interventions and supports advising strategies [6].

IV. METHODOLOGY

A. Overall framework

Figure 2 illustrates the proposed modelling framework. Data from the five layers are loaded and preprocessed to produce a clean, consistent, machine-readable dataset; rows with missing GPA are removed, categorical attributes are encoded [20]. After which feature selection is performed to remove redundant or low-variance features. The dataset is split into 80 % training and 20 % testing subsets stratified by the GPA category. To address the class imbalance, we apply random oversampling on the training data, replicating minority-class instances until the class frequencies are balanced. Hyper-parameter tuning is then conducted on the oversampled training set using grid search for traditional ML models and Hyperband for deep models.

The modelling stage encompasses two branches: (1) traditional ML classifiers—logistic regression, decision tree, k-nearest neighbours (KNN), support vector machine (SVM), random forest, AdaBoost, multi-layer perceptron (MLP) and XGBoost; and (2) deep models—LSTM and CNN. Each model is trained using 10-fold cross-validation on the oversampled training set. The best hyper-parameters are selected based on the mean F1-score. The final models are then evaluated on the hold-out test set, where no oversampling is applied. Evaluation metrics include accuracy, recall, precision, F1-score and AUC. Our study also adopts SHAP values to visualise feature contributions.

B. Model Suite

To assess predictive performance on the proposed multi-layer dataset, we compare a range of models, covering linear classifiers, instance-based approaches, margin methods, ensemble trees, and deep architectures.

1) Logistic regression: A linear classifier with multinomial loss, offering interpretability and serving as a transparent baseline; the regularisation coefficient was tuned via cross validation.

2) Decision Tree: A single classification tree using entropy based splits, with maximum depth and minimum leaf size adjusted to prevent overfitting.

3) K-Nearest Neighbours (KNN): A distance based method that classifies each instance by majority vote among its nearest neighbours; different values of k and weighting schemes were explored.

4) Support Vector Machine (SVM): A kernel based classifier that maximises the margin between classes; both linear and radial basis function kernels were tested with varied penalty parameters.

5) Random Forest: An ensemble of hundreds of decision trees trained on bootstrap samples and random feature subsets to reduce variance and improve generalisation.

6) AdaBoost: A boosting algorithm that sequentially combines weak learners (decision stumps), reweighting observations to focus on misclassified cases; the number of estimators and learning rate were tuned.

7) Multi Layer Perceptron (MLP): A feed forward neural network with one hidden layer using rectified linear units and softmax output; network size, activation function and learning rate were optimised.

8) XGBoost: An efficient gradient boosted tree ensemble that incorporates regularisation and handles missing values; the number of estimators, tree depth and learning rate were tuned. In our experiments, it achieves the highest accuracy and AUC, corroborating findings from prior research that gradient-boosted ensembles can outperform other models on educational data *[8]*, *[17]*, *[21]*.

9) Long Short Term Memory (LSTM): A recurrent neural architecture adapted here to tabular data; a single LSTM layer feeds into dense and dropout layers, with units and dropout rates selected through Hyperband.

10) Convolutional Neural Network (CNN): A one dimensional CNN with two convolutional layers, batch normalisation and pooling, followed by dense layers; kernel sizes, filter counts and dropout rates were tuned via Hyperband.

C. Evaluation metrics and interpretability

To assess model performance comprehensively, we report accuracy (proportion of correct predictions), recall (sensitivity to the minority class), precision (purity of positive predictions), F1-score (harmonic mean of precision and recall) and area under the receiver-operating characteristic curve (AUC) [22]. AUC is particularly important because it measures the trade-off between true positives and false positives across different thresholds and is robust to class imbalance. In addition, confusion matrices provide class-specific insights into misclassification patterns. We also compute SHAP values for the best model to quantify feature contributions for each prediction and generate global feature importance rankings. SHAP, grounded in cooperative game theory, attributes a consistent value to each feature by considering all possible feature subsets. This approach aligns with calls for more transparent AI in education and allows stakeholders to understand why a model made a particular prediction.

D. Proposed methodology diagram

The overall workflow is summarised in **Figure 2**. From left to right, the multi-layer dataset is loaded, preprocessed and subject to feature selection. The data are split into training and testing subsets, and random oversampling balances the training set. Hyper-parameter tuning then feeds into model development for both traditional and deep learners. Finally, the trained models are evaluated and interpreted using a battery of metrics and visualisations.

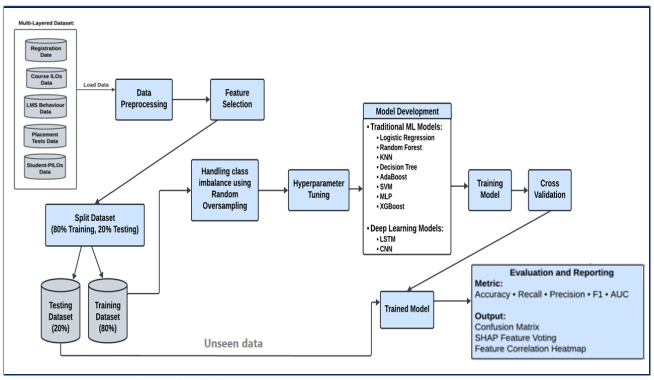

Fig. 2. Proposed methodology

V. Results

After training and tuning the models, we evaluate them on the 20 % hold-out test set. **Table II** summarises the performance metrics for all ten models. The XGBoost classifier yields the highest accuracy (0.852) and AUC (0.955), closely followed by random forest and SVM. Deep models (LSTM and CNN) deliver respectable recall but slightly lower precision and AUC. Logistic regression and decision tree provide interpretable baselines but lag behind ensemble methods. These results are consistent with prior findings that gradient-boosted trees excel on structured tabular data [1], [8], [21].

TABLE II. Test-set performance of ML and DL models

Classifier	Accuracy	Recall	Precision	F1 Score	AUC
Logistic Regression	0.754	0.784	0.748	0.757	0.915
Random Forest	0.832	0.837	0.833	0.831	0.946
KNN	0.820	0.821	0.817	0.815	0.892
Decision Tree	0.754	0.747	0.762	0.751	0.807
AdaBoost	0.781	0.798	0.774	0.783	0.792
SVM	0.820	0.824	0.817	0.818	0.933
MLP	0.770	0.799	0.758	0.771	0.921
XGBoost	0.852	0.850	0.861	0.852	0.955
LSTM	0.805	0.822	0.793	0.804	0.921
CNN	0.809	0.816	0.802	0.807	0.931

The confusion matrix for the XGBoost model (**Figure 3**) reveals that most misclassifications occur between neighbouring GPA categories: a few low-GPA students are predicted as medium, and some medium-GPA students are predicted as high. No low-GPA student is erroneously classified as high, indicating that the model rarely overlooks at-risk learners.

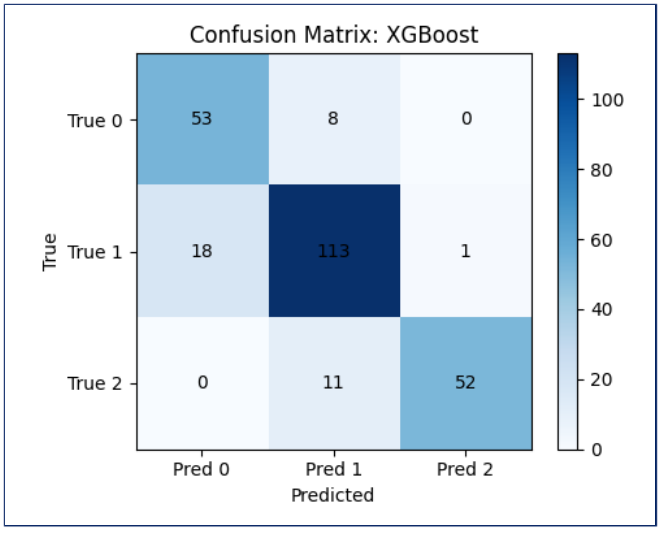

Fig. 3. Confusion matrix for the XGBoost model

We next inspect feature importance using SHAP values for XGBoost (**Figure 4**). The mean CILO attainment emerges as the most influential predictor, corroborating our hypothesis that incorporating learning-outcome data enhances predictive

power. Placement-test scores rank next, reflecting the importance of prior knowledge. Behavioural indicators (e.g., number of quiz attempts and time spent in Moodle) also contribute significantly, highlighting the value of fine-grained engagement data [23]. Demographic variables such as nationality and gender have comparatively lower importance, suggesting that academic and behavioural factors carry more weight in the model's decisions. This pattern is consistent with evidence that behavioural indicators and course assessments are among the most informative predictors of learning outcomes, supporting early identification and intervention [2], [8], [12].

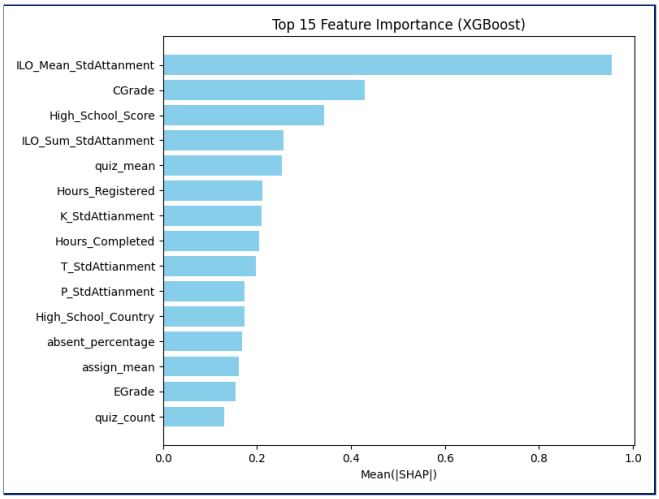

Fig. 4. SHAP feature importance for XGBoost model

VI. Discussion

The results demonstrate that gradient-boosted trees provide the best trade-off between accuracy and interpretability on our medium-sized, multi-layered dataset. XGBoost outperforms deep models despite their capacity to learn complex patterns. This finding resonates with the broader "no free lunch" principle in prediction modelling: algorithmic superiority is highly context-dependent and influenced by dataset size, feature richness and class imbalance. The limited gain from deep models may stem from the modest sample size (952 students) and the absence of temporal sequences; aggregated features summarise behaviour and outcomes at a single time point, thereby limiting the advantage of recurrent architectures.

The prominence of CILO and PILO features confirms our contributions. Many prior works neglected explicit learning outcomes and instead relied on grades or generic LMS metrics. By aggregating course- and program-level outcomes, we provide a more fine-grained measure of learning quality, which emerges as the strongest predictor of final GPA. These findings suggest that educational institutions should invest in robust outcome-assessment systems and integrate them into learning analytics platforms [24]. The high importance of placement-test scores underscores the predictive value of pre-admission proficiency and supports the use of diagnostic exams for placement and early advising [25].

Our results align with several recent studies that favour ensemble methods over deep networks on small/medium tabular datasets. For instance, a hybrid approach reported that XGBoost achieved 78.37 % accuracy, outperforming other models by 3–8 % [17]. Another study reported that random forest performs strongly for student performance prediction on behavioural data [13], while logistic regression achieved competitive AUC for predicting at risk students [1]. In contrast**,** one recent study reported that a Bi-LSTM model slightly outperformed XGBoost and was statistically superior on their dataset [16].

Although our study integrates multiple data sources, several limitations remain. First, the dataset comes from a single institution and a single academic year, which may limit generalisability. Future work could combine data from multiple universities to assess cross-institutional robustness. Second, we discretise GPA into categories; regression models predicting continuous GPA could provide more nuanced insights. Third, we summarise behavioural data at the student level; time-series modelling of weekly engagement patterns may yield additional predictive signals, allowing LSTM and CNN architectures to exploit temporal dynamics. Finally, we only address class imbalance through oversampling; exploring alternative techniques such as SMOTE, class-weighted loss functions and ensemble strategies could further improve minority-class performance.

VII. Conclusion

This study proposed an AI-driven framework for predicting university student performance using a multi-layer dataset that integrates registration variables, CILOs, LMS behavioural indicators, placement-test scores and PILO aggregates. By comparing a diverse set of machine-learning and deep-learning models, we demonstrated that gradient-boosted tree ensembles achieve the highest predictive accuracy and offer interpretable insights through SHAP values. The explicit inclusion of course- and program-level learning outcomes is shown to be a key factor in improving predictions, underscoring the importance of outcome-based assessment in learning analytics. Our findings suggest that modestly sized educational datasets may benefit more from well-tuned ensemble models than from deep architectures. Future research should explore transfer learning across institutions, temporal modelling of engagement patterns and fairness-aware algorithm design to ensure that predictive analytics supports equitable and effective educational interventions.

Acknowledgment

The authors thank the University of Petra, Amman, Jordan, for its support of this research.

References

[1] F. E. Arévalo-Cordovilla and M. Peña, "Comparative analysis of machine learning models for predicting student success in online programming courses: a study based on LMS data and external factors," *Mathematics,* vol. 12, *(20),* pp. 3272, 2024. .

[2] A. Namoun and A. Alshanqiti, "Predicting student performance using data mining and learning analytics techniques: A systematic literature review," *Applied Sciences,* vol. 11, *(1),* pp. 237, 2020. .

[3] I. Papadogiannis, M. Wallace and G. Karountzou, "Educational data mining: A foundational overview," *Encyclopedia,* vol. 4, *(4),* pp. 1644–1664, 2024. .

[4] A. Almalawi *et al*, "Predictive models for educational purposes: A systematic review," *Big Data and Cognitive Computing,* vol. 8, *(12),* pp. 187, 2024. .

[5] K. O. Adefemi, M. B. Mutanga and V. Jugoo, "Hybrid Deep Learning Models for Predicting Student Academic Performance," *Mathematical and Computational Applications,* vol. 30, *(3),* pp. 59, 2025. .

[6] F. A. Al-Azazi and M. Ghurab, "ANN-LSTM: A deep learning model for early student performance prediction in MOOC," *Heliyon,* vol. 9, *(4),* 2023. .

[7] K. O. Adefemi and M. B. Mutanga, "A Robust Hybrid CNN–LSTM Model for Predicting Student Academic Performance," *Digital,* vol. 5, *(2),* pp. 16, 2025. .

[8] M. AlShaikh-Hasan and G. Ghinea, "Predicting student performance using deep learning and machine learning techniques," in *2025 1st International Conference on Computational Intelligence Approaches and Applications (ICCIAA),* 2025, .

[9] W. Hadi and M. N. AlShaikh-Hasan, "AI-driven curriculum transformation and faculty development in developing universities," in *Revolutionizing Urban Development and Governance with Emerging Technologies*Anonymous 2025, .

[10] S. AbuAisheh, Y. W. Teh and L. Shuib, "Monitoring university students' focus during online lectures using artificial intelligence tools," in *2025 1st International Conference on Computational Intelligence Approaches and Applications (ICCIAA),* 2025, .

[11] W. Kim and H. Kim, "Counterfactual fairness evaluation of machine learning models on educational datasets," in *International Conference on Intelligent Tutoring Systems,* 2025, .

[12] H. A. Althibyani, "Predicting student success in MOOCs: a comprehensive analysis using machine learning models," *PeerJ Computer Science,* vol. 10, pp. e2221, 2024.

[13] J. Chen *et al*, "Evaluation of student performance based on learning behavior with random forest model," in *2024 13th International Conference on Educational and Information Technology (ICEIT),* 2024, .

[14] D. Khairy *et al*, "Prediction of student exam performance using data mining classification algorithms," *Education and Information Technologies,* vol. 29, *(16),* pp. 21621–21645, 2024. .

[15] S. Gaftandzhieva *et al*, "Exploring online activities to predict the final grade of student," *Mathematics,* vol. 10, *(20),* pp. 3758, 2022. .

[16] E. Kalita *et al*, "Predicting student academic performance using bi-LSTM: A deep learning framework with SHAP-based interpretability and statistical validation," in *Frontiers in Education,* 2025, .

[17] H. Xue and Y. Niu, "Multi-output based hybrid integrated models for student performance prediction," *Applied Sciences,* vol. 13, *(9),* pp. 5384, 2023. .

[18] X. Zhang *et al*, "Optimizing multi label student performance prediction with GNN-TINet: A contextual multidimensional deep learning framework," *PloS One,* vol. 20, *(1),* pp. e0314823, 2025. .

[19] S. Ajili *et al*, "Predicting digital literacy gains in rural contexts using multidimensional data: A machine learning approach," in *2025 IEEE/ACS 22nd International Conference on Computer Systems and Applications (AICCSA),* 2025, . DOI: 10.1109/AICCSA66935.2025.11315320.

[20] A. Altarawneh *et al*, "ETM-F: an enriched topic modeling and filtration framework integrating ontologies and deep learning for biomedical trend analysis," *Knowledge and Information Systems,* vol. 68, *(1),* pp. 25, 2026. .

[21] M. AlShaikh-Hasan and G. Ghinea, "Evaluating the impact of multi-layer data on machine learning classifiers for predicting student academic performance," in *2024 25th International Arab Conference on Information Technology (ACIT),* 2024, . DOI: 10.1109/ACIT62805.2024.10877023.

[22] M. Baklizi *et al*, "A powerful machine learning method for detecting phishing threats," *Bulletin of Electrical Engineering and Informatics,* vol. 14, *(6),* pp. 4626–4640, 2025. .

[23] A. Shtayyat and A. Gawanmeh, "Gamification and AI for enhanced student engagement and automated assessment in higher education moodle E-learning," in *2025 16th International Conference on Information and Communication Systems (ICICS),* 2025, .

[24] M. A. Arqoub *et al*, "Extending learning management system for learning analytics," in *2022 International Conference on Business Analytics for Technology and Security (ICBATS),* 2022, .

[25] P. Everaert, E. Opdecam and H. Van Der Heijden, "Predicting first-year university progression using early warning signals from accounting education: A machine learning approach," *Accounting Education,* vol. 33, *(1),* pp. 1–26, 2024. .