

Explainable At-Risk Student Prediction with SHAP-Guided Risk Segmentation and Rule-Based Intervention Recommendations

¹P. L. Barilee, ^{*2}D. Matthias, ^{*3}V. I. E. Anireh

^{1,2,3}Department of Computer Science, Rivers State University, Port-Harcourt, Nigeria

E-mail: ¹patience.barilee@rsu.edu.ng, ²matthias.daniel@ust.edu.ng, ³anireh.ike@ust.edu.ng

Abstract: Early identification of students at risk of failure or withdrawal can support timely academic assistance, but many predictive systems provide risk labels without transparent reasons or guidance on suitable support actions. An explainable at-risk student prediction workflow combined Random Forest and XGBoost classification, SHAP-based explanation, k-means risk segmentation, and rule-based intervention recommendations. The primary evaluation used Open University Learning Analytics Dataset records observed at day 60, with the 2014J presentation retained as a locked temporal holdout; Kalboard 360 provided a secondary stratified-cross-validation robustness check. XGBoost was selected during development and achieved accuracy of 0.679, precision of 0.554, recall of 0.802, F1-score of 0.656, ROC-AUC of 0.798, and PR-AUC of 0.747 on the OULAD temporal holdout. The recall-oriented threshold identified 2,695 of 3,359 at-risk students, while producing 2,167 false-positive alerts requiring staff review. SHAP identified assessment completion ratio, mean assessment score, available weighted score, and last active day as the leading global predictors. Clustering predicted at-risk students into three groups produced a silhouette score of 0.409 and stability ARI of 1.000; the resulting profiles were mapped to tutoring, study-skills support, engagement monitoring, and academic-adviser review. The findings show that ensemble prediction can be connected to transparent risk profiles and practical support recommendations. However, the workflow remains a decision-support prototype requiring local validation and human oversight.

Keywords: at-risk student prediction, educational data mining, explainable artificial intelligence, SHAP, intervention recommendation.

I. INTRODUCTION

Failure, withdrawal, and persistent underperformance are continuing concerns in higher education, particularly where large cohorts and digital learning environments make individual monitoring difficult. Educational data generated from assessments, prior academic history, and learning-platform activity can support earlier identification of students whose observed progress is consistent with academic risk. An early-risk investigation used engagement, demographic, and performance data to evaluate supervised machine-learning models for identifying at-risk students [1]. A deep ensemble approach evaluated student-performance prediction using a continuous final-score target [2]. Genetic optimisation of Random Forest components has been evaluated for multi-class student-performance prediction [3]. A hybrid CNN-LSTM model has also been applied to academic-performance prediction [4]. These results support predictive modelling in educational settings, but their reported values are not directly interchangeable because

targets, datasets, feature availability, and validation designs differ substantially.

Prediction alone is insufficient for educational decision-making. A risk label or probability does not tell educators which observed factors contributed to a flag, whether the signal is educationally plausible, or which support action should be considered. This limitation is particularly important where automated outputs inform decisions about student outreach and allocation of scarce academic-support resources. Explainable artificial intelligence addresses part of this problem by exposing feature contributions for individual predictions and across a population. For tree-based models, SHAP provides an efficient route from local explanations to global feature-importance analysis [5]. Educational applications have used SHAP and related methods to make student predictions more inspectable [6], but explanation quality can be affected by model characteristics and correlated features, requiring cautious interpretation rather than causal claims [7].

Existing work therefore establishes three related needs: accurate identification of students likely to be at risk, transparent interpretation of model outputs, and a defensible connection between risk information and support planning. Deep-learning models with SHAP have reported strong classification results, but their outputs do not necessarily include a mechanism for grouping comparable risk patterns or assigning support actions [8]. A useful decision-support workflow should preserve the distinction between model prediction and professional judgement while making its evidence and recommended next steps visible to staff.

This paper evaluates an explainable at-risk student prediction workflow based on Random Forest and XGBoost classification, SHAP explanation, k-means segmentation of predicted at-risk students, and rule-based intervention recommendations. The principal experiment uses the Open University Learning Analytics Dataset (OULAD), with a day-60 observation window and a locked 2014J temporal holdout. A Kalboard 360 experiment is reported as a secondary robustness check under stratified cross-validation, not as a directly comparable temporal validation. The paper first examines relevant work in prediction, explainability, and risk segmentation; it then describes the evaluation workflow, reports classification, explanation, clustering, and recommendation results, and concludes with the practical scope and limitations of the evidence.

II. RELATED WORKS

Educational data mining research has applied statistical, machine-learning, and deep-learning methods to predict student performance and identify students who may require additional support. Early-risk work demonstrates that student records and learning activity can be used to distinguish students with elevated risk, while also showing that performance depends on the institutional setting, observation period, and target definition [1]. Ensemble methods remain relevant because they model nonlinear relationships across heterogeneous academic and engagement features without requiring a single parametric decision rule. A deep ensemble study reported that combining predictive models can improve student-performance classification [2], while genetic optimisation of Random Forest components has been used to improve performance on a multi-class student dataset [3]. Hybrid CNN-LSTM work similarly shows that temporal and behavioural patterns can support high predictive performance [4]. These results motivate comparative evaluation of ensemble methods, but high scores from random

splits or different performance targets cannot substitute for a future-presentation test of at-risk classification.

The growing use of complex predictive models has increased the need for explanations that educators can inspect. Tree-based SHAP decomposes a prediction into feature contributions and supports both local explanations for a flagged student and global rankings across a cohort [5]. In educational machine learning, explainability has been applied to student-adaptability prediction with Random Forest and multiple post-hoc techniques, demonstrating the practical value of feature-level evidence for interpreting classifications [6]. However, explanations should not be treated as causal accounts. Comparative analysis of SHAP and LIME in educational settings identifies instability risks associated with feature correlation, model choice, and local perturbation procedures [7]. An explainable model therefore requires both appropriate visualisation and disciplined interpretation of feature importance.

Several studies combine predictive models with post-hoc explainability. A Bi-LSTM study of student academic performance combined SHAP with statistical validation and reported both deep-learning and XGBoost baselines [8]. Explainable models have also been used to estimate overall course scores, where the evaluation highlights the trade-off between the fidelity of complex models and the natural transparency of simpler decision structures [9]. Other dropout-risk work combines a hybrid classifier with SHAP and LIME, demonstrating that explanations can expose risk factors after prediction [10]. Ensemble regression with SHAP has likewise been used to identify influential prior-performance features for institutional decision-making [11]. These approaches advance interpretability, but they primarily remain prediction-and-explanation systems: they do not establish a repeatable path from patterns shared by flagged students to differentiated support actions.

Clustering offers a complementary population-level view by grouping students with similar academic and engagement profiles. A systematic review of clustering in student-risk prediction reports that k-means and related techniques can help identify performance groups, but it also identifies inconsistent validation and limited integration with explainability as persistent issues [12]. Clustering-based preprocessing has been used to represent course difficulty and improve classification inputs [13], yet preprocessing-oriented clustering does not itself provide an intervention policy. Qualitative work on explainable AI in education further stresses the trade-off between predictive

complexity, explanation fidelity, and educator understanding [14]. Broader reviews of deep learning in educational data mining also identify performance prediction and personalised support as active but distinct research areas, with data sparsity and computational demands limiting straightforward deployment [15]. Feature-selection research reinforces the importance of examining feature relevance and redundancy before interpreting predictive outputs [16].

The unresolved problem is therefore not only to predict risk or display feature importance. It is to evaluate a transparent workflow under a realistic temporal holdout, then link predicted at-risk students to validated, interpretable profiles and explicit support rules. The workflow addresses that boundary by combining ensemble classification with SHAP explanations, k-means segmentation of predicted at-risk students, and cluster-level recommendations. The recommendation rules are intended to make support planning auditable; they do not claim that any intervention causes improved student outcomes.

III. METHODOLOGY

The evaluated workflow connects data preparation, ensemble classification, SHAP explanation, segmentation of predicted at-risk students, and rule-based intervention mapping. Figure 1 shows these components and their sequence. Student records first undergo cleaning, missing-value handling, categorical encoding, and numerical preparation. Random Forest and XGBoost then generate at-risk probabilities and classifications. The selected classifier is interpreted with TreeSHAP, which supports global feature rankings and case-level feature contributions [5]. Students predicted to be at risk are grouped using k-means, and their cluster profiles are mapped to transparent support rules. The resulting output combines a risk decision, explanatory evidence, risk-group assignment, and recommended staff actions.

The primary evaluation used the Open University Learning Analytics Dataset. The analysis retained registered and active students at a day-60 observation cut-off and defined the 2014J course presentation as a locked temporal test set. This design tests whether a model developed from earlier presentations can identify risk in a later presentation without student overlap between development and test data. After cohort construction, the OULAD analysis comprised 26,346 rows: 16,332 development rows and 8,815 temporal-holdout rows.

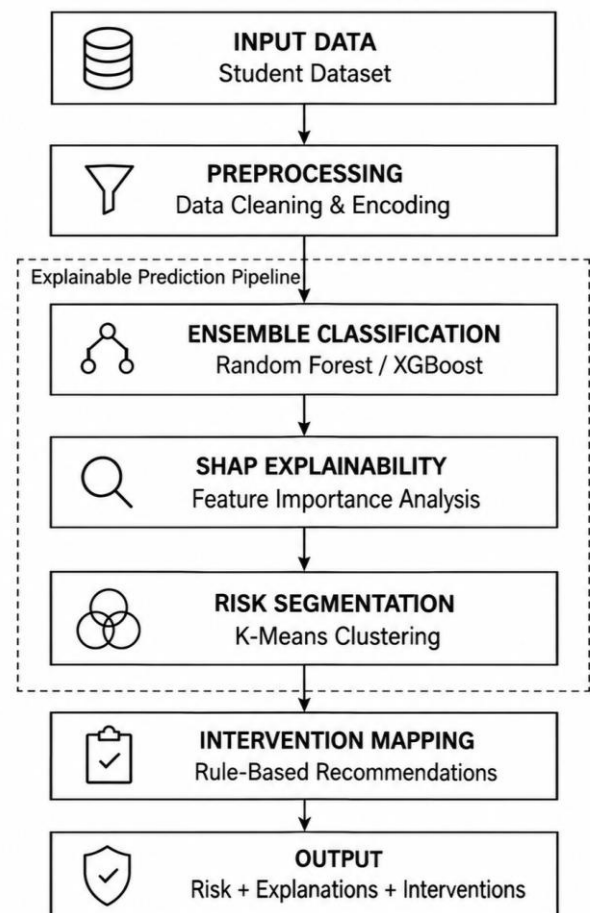


Figure 1: Architecture of the explainable at-risk student prediction and intervention workflow

The binary outcome classified Fail and Withdrawn as at-risk and Pass and Distinction as not at-risk. Demographic, prior-study, assessment, and virtual-learning-environment features were prepared for classification. Protected attributes were excluded from the predictor set and retained only for post-hoc subgroup evaluation.

Kalboard 360 was used as a secondary robustness dataset. It contains 480 records and 16 demographic, academic-context, behavioural, and parental-involvement features, with Low, Medium, and High performance categories. OULAD and Kalboard 360 serve different purposes in the evaluation: OULAD supplies the primary temporal-generalisation evidence, whereas the smaller Kalboard dataset is assessed with stratified cross-validation. Table 1 summarises the two evaluation settings.

Table 1: Datasets and validation protocols

Dataset	Evaluation role	Records used	Validation protocol	Outcome setting
OULAD	Primary evaluation	26,346 cohort rows	Development on earlier presentations; locked 2014J temporal holdout	Binary at-risk versus not-at-risk
Kalboard 360	Secondary robustness check	480 records	Stratified cross-validation	Three performance levels: Low, Medium, and High

During OULAD development, three classifiers were compared: a dummy baseline, Random Forest, and XGBoost. The dummy baseline establishes the performance expected without learning meaningful predictive structure. Random Forest used 100 estimators with unrestricted maximum depth, while XGBoost used 100 estimators, a learning rate of 0.1, and maximum depth of 6. All stochastic stages used *random_state=42*. Model selection considered development-fold

F1-score, ROC-AUC, PR-AUC, Brier score, and the practical requirement to identify at-risk students. Candidate probability thresholds were tuned during development rather than fixed at 0.5. The final temporal-holdout evaluation used the selected XGBoost classifier and a threshold of 0.315, reflecting the recall-oriented screening objective. Table 2 records the principal experimental configuration.

Table 2: Main experimental configuration

Component	Configuration
Dummy classifier	Development baseline
Random Forest	100 estimators; max_depth=None; random_state=42
XGBoost	100 estimators; learning rate 0.1; maximum depth 6; random_state=42
Final OULAD classifier	XGBoost with decision threshold 0.315
SHAP	TreeExplainer for the selected tree-based classifier
Clustering	k-means with k in {3, 4, 5}; n_init=10; random_state=42

Classification performance was assessed with accuracy, precision, recall, F1-score, ROC-AUC, PR-AUC, and Brier score. Accuracy provides overall classification correctness; precision quantifies the proportion of alerts that correspond to actual at-risk students; recall measures the proportion of actual at-risk students identified; and F1-score balances precision and recall. ROC-AUC and PR-AUC evaluate probability ranking across thresholds, while the Brier score measures the quality of predicted probabilities. The OULAD temporal holdout was kept separate from all development-stage comparison and threshold-selection activities. Kalboard 360 results were calculated with stratified cross-validation because the dataset does not provide an equivalent later-presentation split.

TreeSHAP was applied after model selection to calculate feature contributions for the selected XGBoost classifier. Global importance was computed by averaging the absolute SHAP value of each feature across the evaluated records, while local explanations decomposed an individual risk probability into its contributing features. These explanations describe associations learned by the model rather than causal determinants of student

outcome. This distinction is important for educational use because correlated variables and modelling choices can change feature attributions [7].

Student segmentation was applied only to students predicted to be at risk. Academic and engagement profiles were standardised before k-means clustering, and solutions with three, four, and five clusters were compared. Selection considered silhouette score, stability measured with the adjusted Rand index, minimum cluster share, and the interpretability of cluster profiles. A systematic review identifies the need to evaluate clustering as a meaningful risk-grouping step rather than treating it as an unvalidated analytical addition [12]. Clustering-based preprocessing research further shows that clustering can be used to reduce ambiguity and support downstream dropout classification [13]. Cluster centroids were examined for deficits in assessment completion, assessment performance, prior attempts, and engagement measures.

The intervention layer used explicit rules rather than a learned treatment model. A cluster with completion and engagement deficits was mapped to study-skills and assessment-

planning support together with engagement outreach and activity monitoring. A cluster characterised primarily by assessment deficit was mapped to academic tutoring and prerequisite review. A cluster characterised by prior-attempt risk and engagement deficit was mapped to academic-adviser progress review together with engagement outreach and monitoring. The rules provide consistent, inspectable support suggestions for staff review. They do not estimate intervention effects, replace educator judgement, or authorise automatic action.

IV. RESULTS AND DISCUSSION

The primary OULAD execution used the day-60 cohort, a locked 2014J temporal holdout, the selected XGBoost threshold of 0.315, and random_state=42 for stochastic stages. The full pipeline completed in 85.08 seconds, including preprocessing, model execution, explanation, clustering, and recommendation generation. Final XGBoost fitting required 1.38 seconds. These timings describe the recorded pipeline execution rather than a general deployment-latency guarantee.

The evaluation compares a dummy baseline, Random Forest, and XGBoost during development, then reports the selected XGBoost model on the temporal holdout. It also evaluates global feature explanations, cluster separation and stability, subgroup recall variation, and a Kalboard 360 robustness check. Table 3 presents the development cross-validation results, while Table 4 presents the primary temporal-holdout evidence.

4.1 Classification Performance Under Temporal Holdout

Table 3 shows that both ensemble classifiers substantially exceeded the dummy baseline across all predictive measures. Random Forest produced marginally higher development accuracy than XGBoost, by 0.5 percentage points. XGBoost, however, produced the highest F1-score, ROC-AUC, and PR-AUC, as well as the lowest Brier score. The paired development-fold comparison for F1-score gave a t-statistic of -3.813 and a p-value of 0.019 for Random Forest relative to XGBoost. XGBoost was consequently selected for the temporal-holdout evaluation.

Table 3: Development cross-validation performance on OULAD

Model	Accuracy	Precision	Recall	F1-score	ROC-AUC	PR-AUC	Brier
Dummy	0.567 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.500 ± 0.000	0.433 ± 0.000	0.245 ± 0.000
Random Forest	0.736 ± 0.006	0.667 ± 0.007	0.781 ± 0.009	0.720 ± 0.006	0.828 ± 0.006	0.818 ± 0.006	0.163 ± 0.003
XGBoost	0.731 ± 0.006	0.651 ± 0.006	0.817 ± 0.014	0.724 ± 0.008	0.837 ± 0.007	0.827 ± 0.008	0.158 ± 0.004

On the locked 2014J presentation, XGBoost achieved a recall of 0.802 and an F1-score of 0.656, as shown in Table 4. The classifier identified 2,695 of 3,359 actual at-risk students from 8,815 test records. Its ROC-AUC of 0.798 and PR-AUC of 0.747 show useful ranking of student risk probabilities under the

temporal test. The selected threshold also produced 2,167 false positives, which reduced precision to 0.554. The result is therefore suited to recall-oriented staff screening, where missing an at-risk student is costly, but it should not be used to make automatic intervention decisions.

Table 4: XGBoost performance on the OULAD 2014J temporal holdout

Test records	Actual at-risk	Threshold	Accuracy	Precision	Recall	F1-score	ROC-AUC	PR-AUC	Brier
8,815	3,359	0.315	0.679	0.554	0.802	0.656	0.798	0.747	0.175

Figure 2 displays the temporal-holdout confusion matrix. The 664 false negatives show that not all at-risk students were identified at the selected threshold, while the 2,167 false positives establish the workload that a staff-review process must be able to absorb. The predicted at-risk rate was 0.552, above the observed prevalence of 0.381, which is consistent with the recall-

oriented operating threshold.

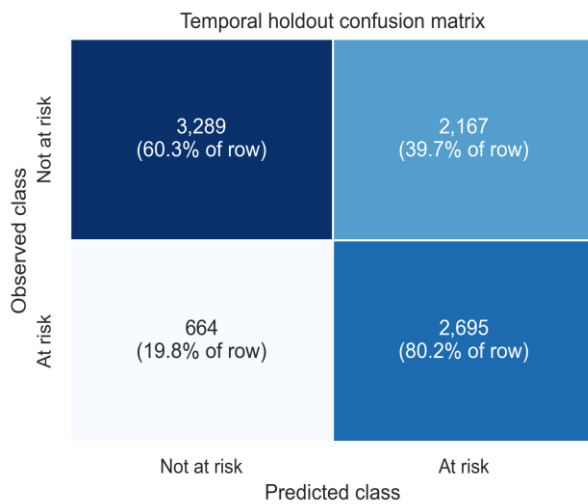


Figure 2: Confusion matrix for XGBoost on the OULAD 2014J temporal holdout

4.2 Explainability Evidence

The selected XGBoost model was interpreted using global SHAP values. Table 5 shows that assessment-completion ratio was the highest-ranked feature, followed by mean assessment score, available weighted score, and last active day. These results indicate that the classifier relied most strongly on observed assessment progression, assessed performance, and recent learning activity when ranking risk. Educational background, completion weight, recent active days, module indicator, activity trend, and clicks in days 31–60 also contributed, but with lower mean absolute SHAP values.

Table 5: Ten highest-ranked global SHAP features for the selected XGBoost model

Rank	Feature	Mean absolute SHAP
1	assessment_completion_ratio	0.567
2	mean_assessment_score	0.430
3	weighted_score_available	0.264
4	last_active_day	0.214
5	highest_education_Lower Than A Level	0.195
6	assessment_weight_completed	0.168
7	active_days_last_14	0.139
8	code_module_FFF	0.118
9	activity_trend	0.109
10	clicks_days_31_60	0.096

Figure 3 visualises the global SHAP ranking. The concentration of the leading contributions in assessment and activity variables provides a transparent basis for discussing an alert with staff, but it does not establish that changing a single feature would cause a different student outcome. The explanations should therefore support case review alongside educator knowledge and student context.

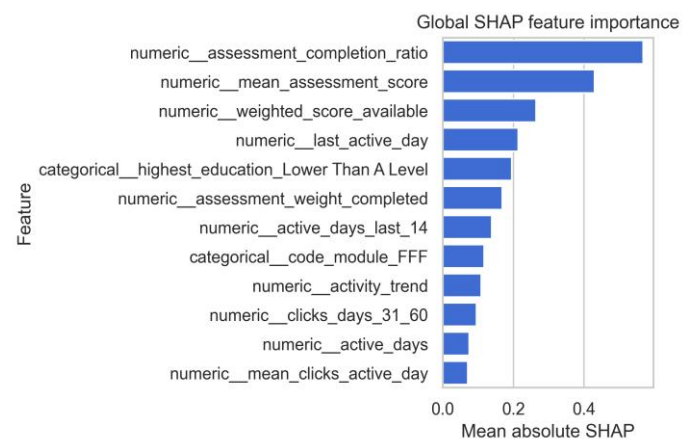


Figure 3: Global SHAP feature importance for the selected XGBoost classifier

4.3 Risk Segmentation and Intervention Mapping

K-means clustering was evaluated for three, four, and five clusters among students predicted to be at risk. Table 6 shows that the three-cluster solution had the highest silhouette score (0.409), perfect stability ARI (1.000), and a minimum cluster share of 0.153. The four-cluster solution had a similar silhouette score but did not improve separation, while the five-cluster solution reduced the silhouette score to 0.347. The three-cluster configuration was selected because it balanced measured separation, repeatability, and interpretability.

Table 6: Cluster-selection results for predicted at-risk students

k	Silhouette	Stability ARI	Minimum cluster share	Inertia
3	0.409	1.000	0.153	10,128.090
4	0.407	1.000	0.142	7,499.867
5	0.347	0.994	0.140	6,542.442

Figure 4 presents the standardised cluster profiles. Cluster 0 showed a pronounced completion deficit and was assigned study-skills and assessment-planning support with engagement outreach. Cluster 1 was dominated by assessment deficit and was mapped to academic tutoring and prerequisite review. Cluster 2

showed the strongest prior-attempt risk together with an engagement deficit and was mapped to academic-adviser review, outreach, and monitoring. Table 7 records the resulting cluster-level recommendations.

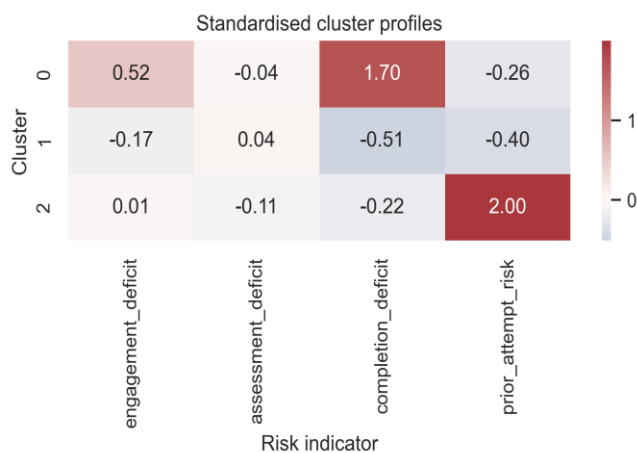


Figure 4: Standardised risk-indicator profiles for the selected three-cluster solution

Table 7: Cluster-level intervention recommendations

Cluster	Students	Dominant risk indicators	Rule IDs	Recommended support
0	1,027	Completion deficit; engagement deficit	R3; R1	Study-skills and assessment-planning support; engagement outreach and activity monitoring
1	3,089	Assessment deficit	R2	Academic tutoring and prerequisite review
2	746	Prior-attempt risk; engagement deficit	R4; R1	Progress review with an academic adviser; engagement outreach and activity monitoring

The clusters turn a common alert list into differentiated profiles rather than treating all predicted at-risk students as a single group. The recommendations are deliberately rule based: they can be inspected and revised by educators, but their

effectiveness was not experimentally evaluated in the reported evaluation.

4.4 Subgroup Audit and Robustness Check

Protected attributes were excluded from the prediction inputs and used for post-hoc performance auditing. Recall ranged by 0.022 across gender groups and 0.031 across disability groups. The largest range was across age bands (0.146); however, the 55-and-above group contained only 81 test records, making its recall estimate less stable. The audit identifies a priority for additional evaluation rather than demonstrating fairness or bias conclusively.

The secondary Kalboard 360 evaluation produced higher cross-validation scores than the OULAD temporal holdout. Random Forest achieved accuracy of 0.910 ± 0.035 , precision of 0.821 ± 0.083 , recall of 0.849 ± 0.078 , and F1-score of 0.833 ± 0.068 . XGBoost achieved accuracy of 0.908 ± 0.045 , precision of 0.846 ± 0.071 , recall of 0.794 ± 0.131 , and F1-score of 0.817 ± 0.095 . These results provide secondary evidence that the modelling approach performs effectively on a smaller behavioural dataset, but they are not directly comparable with the OULAD figures because the datasets, target structures, feature sets, sample sizes, and validation protocols differ.

The evidence is limited to public benchmark data, a day-60 observation window, and rule-based recommendations whose intervention outcomes were not measured. The false-positive burden on OULAD also means that institutional deployment would require local validation, threshold selection matched to support capacity, practitioner review of the rules, and ongoing monitoring of subgroup performance.

V. CONCLUSION

The evaluation assessed an explainable at-risk student prediction workflow that combines ensemble classification, SHAP explanation, k-means risk segmentation, and rule-based intervention recommendations. On the locked OULAD 2014J temporal holdout, the selected XGBoost classifier achieved accuracy of 0.679, precision of 0.554, recall of 0.802, F1-score of 0.656, ROC-AUC of 0.798, and PR-AUC of 0.747. Assessment completion ratio, mean assessment score, available weighted score, and last active day were the strongest global SHAP predictors. The selected three-cluster solution achieved a silhouette score of 0.409 and stability ARI of 1.000, enabling differentiation among completion-and-engagement, assessment,

and prior-attempt-and-engagement risk profiles.

The workflow provides a transparent decision-support basis for prioritising staff review and matching risk profiles to tutoring, study-skills support, engagement monitoring, and adviser review. Its principal limitation is the recall-precision trade-off: the selected threshold identified most actual at-risk students but produced false-positive alerts that require human judgement and available support capacity. The evaluation is also limited to public benchmark datasets and does not measure the effects of recommended interventions. Future work should validate the workflow with institutional data, assess intervention uptake and outcomes, refine thresholds with staff, and monitor subgroup performance over larger and more representative cohorts.

REFERENCES

- [1] Jimenez Martinez, A. L., Sood, K., & Mahto, R. (2024). Early detection of at-risk students using machine learning. *arXiv*, 2412.09483.
- [2] Tang, B., Li, S., & Zhao, C. (2024). Predicting the performance of students using deep ensemble learning. *Journal of Intelligence*, 12(12), 124.
- [3] Chen, M., & Liu, Z. (2024). Predicting performance of students by optimizing tree components of random forest using genetic algorithm. *Heliyon*, 10(13), e32570.
- [4] Adefemi, K. O., & Mutanga, M. B. (2025). A robust hybrid CNN–LSTM model for predicting student academic performance. *Digital*, 5(2), 16.
- [5] Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., & Lee, S.-I. (2020). From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence*, 2(1), 56–67.
- [6] Nnadi, L. C., Watanobe, Y., Rahman, M. M., & John-Otumu, A. M. (2024). Prediction of students' adaptability using explainable AI in educational machine learning models. *Applied Sciences*, 14(12), 5141.
- [7] Hooshyar, D., & Yang, Y. (2024). Problems with SHAP and LIME in interpretable AI for education: A comparative study of post-hoc explanations and neural-symbolic rule extraction. *IEEE Access*, 12, 137472–137490.
- [8] Kalita, E., Alfarwan, A. M., El Aouifi, H., Kukkar, A., Hussain, S., Ali, T., & Gaftandzhieva, S. (2025). Predicting student academic performance using Bi-LSTM: A deep learning framework with SHAP-based interpretability and statistical validation. *Frontiers in Education*, 10, 1581247.
- [9] Ghimire, S., Abdulla, S., Joseph, L. P., Prasad, S., Murphy, A., Devi, A., Barua, P. D., Deo, R. C., Acharya, R., & Yaseen, Z. M. (2024). Explainable artificial intelligence-machine learning models to estimate overall scores in tertiary preparatory general science course. *Computers and Education: Artificial Intelligence*, 7, 100292.
- [10] Mustofa, S., Emon, Y. R., Mamun, S. B., Akhy, S. A., & Ahad, M. T. (2024). A novel AI-driven model for student dropout risk analysis with explainable AI insights. *Computers and Education: Artificial Intelligence*, 8, 100352.
- [11] Ahmed, W., Wani, M. A., Pławiak, P., Meshoul, S., Mahmoud, A., & Hammad, M. (2025). Machine learning-based academic performance prediction with explainability for enhanced decision-making in educational institutions. *Scientific Reports*, 15, 26879.
- [12] Lu, Y., Yeom, S., Maktoubian, J., Rahman, M. M., & Kim, S.-H. (2025). Improve student risk prediction with clustering techniques: A systematic review in education data mining. *Education Sciences*, 15(12), 1695.
- [13] Rani, N., Pachigolla, V., & Kumar, A. (2025). Clustering-based preprocessing for feature reduction and robust student dropout classification. *International Journal of Information Technology*, 17, 2001–2013.
- [14] Gunasekara, S., & Saarela, M. (2025). Explainable AI in education: Techniques and qualitative assessment. *Applied Sciences*, 15(3), 1239.
- [15] Lin, Y., Chen, H., Xia, W., Lin, F., Wang, Z., & Liu, Y. (2025). A comprehensive survey on deep learning techniques in educational data mining. *Data Science and Engineering*, 10(4), 564–590.
- [16] Malik, S., Patro, S. G. K., Mahanty, C., Hegde, R., Naveed, Q. N., Lasisi, A., Buradi, A., & Emma, A. F. (2025). Advancing educational data mining for enhanced student performance prediction: A fusion of feature selection algorithms and classification techniques with dynamic feature ensemble evolution. *Scientific Reports*, 15, 8738.

Citation of this Article:

P. L. Barilee, D. Matthias, & V. I. E. Anireh. (2026). Explainable At-Risk Student Prediction with SHAP-Guided Risk Segmentation and Rule-Based Intervention Recommendations. *Current Journal of Engineering and Science Research*. 3(10), 10-18. Article DOI: <https://doi.org/10.47001/CJESR/2026.310002>

***** End of the Article *****