Int. J. of AI and Education Technology

♦ ISSN: 2835-2432 Vol: 05, No. 01, 2026 ♦ PP. 01–07

Article DOI: https://doi.org/10.54216/IJAIET.050101



FULL LENGTH ARTICLE

From Signals to Action: Explainable AI for Engagement-Responsive Instructional Support in Digital Higher Education

Andino Maseleno[1,*] **Meinhaj Hussain**[2] **Aygul Z. Ibatova**[3]

[1] Institut Bakti Nusantara, Lampung, Indonesia

[2] Rennier University, Ireland

[3] Tyumen Industrial University, Russia

Emails: andino.maseleno@ibnus.ac.id; meinhaj@rennier.online; aigoul@rambler.ru

Received: November 02, 2025 Revised: December 25, 2025 Accepted: January 03, 2026 ⋆ Corresponding author

ABSTRACT

Artificial intelligence (AI) is increasingly used to monitor learning processes in higher education, but many analytics pipelines stop at prediction and provide limited support for instructional action. This study develops an explainable AI framework that uses behavioral and contextual data from digital learning environments to generate tailored instructional support. Using the 480-record xAPI-Edu-Data dataset, the framework constructs an engagement index and support profiles, predicts multiclass performance, and allocates actions through rules. Three classifiers are compared using stratified cross-validation; Random Forest performs best overall. On the hold-out sample, it achieves 0.8021 accuracy, 0.8204 macro-precision, 0.8010 macro-recall, 0.8084 macro-F1, and 0.9140 macro-AUC. Student absence, the composite engagement index, gender, student–guardian relationship, support profile, and digital-resource access are the most influential predictors. The decision layer assigns students to attendance-first intervention, adaptive engagement support, family-engagement reinforcement, structured progression coaching, or challenge-and-extend pathways. The resulting workflow connects explainable AI with actionable instructional decision support for digital higher education.

Keywords: Artificial intelligence in education ▪ Learning analytics ▪ Higher education ▪ Explainable AI ▪ Educational technology ▪ Student support

1. INTRODUCTION

Learning analytics and artificial intelligence (AI) underpin many higher-education platforms, supporting monitoring, feedback, personalisation, and performance analysis. Recent research describes AI-enabled predictive modelling, chatbots, personalised support, and decision-focused dashboards [1, 2, 3, 4]. Yet as these systems grow more technically complex, the instructional pathways that follow their outputs are not always well developed. Analytics often stop at visualisations, performance scores, or risk flags, offering little guidance for implementing support decisions.

The pedagogical potential of AI depends on whether its outputs can inform interventions. Reviews of learning analytics and student engagement report continued interest in performance prediction and engagement monitoring, alongside uneven translation of analytics into pedagogical practice [5, 6]. Research on human-centred AI likewise stresses that tools should be transparent, explainable, and accessible to educators rather than detached technical layers [7].

The next wave of AI adoption in higher education amplifies

© 2026 The Author(s). Published by ASPG. This article is licensed under CC BY 4.0.

this need. Research on generative-AI governance, accessibility, and learning support highlights both innovation and concerns involving ethics, institutional agency, assessment integrity, and inclusivity [8, 9, 10]. AI in education therefore needs to be pragmatically meaningful as well as statistically effective.

This study proposes an explainable AI approach to engagement-responsive instructional support. For student-performance prediction, the framework integrates four processes: constructing an engagement index, profiling support needs, classifying performance outcomes in a multiclass setting, and allocating actions. Using a public dataset supports transparency and replicability. The study asks: how can traces from a learning environment be translated into actionable, interpretable support plans for educators and student-success teams?

2. LITERATURE REVIEW

2.1 AI, Learning Analytics, and Student Engagement

AI in education has developed from basic automation toward broader feedback, personalisation, and educational-support systems. A systematic review by Chiu et al. [1] identifies opportunities for adaptive learning and performance support alongside methodological and ethical challenges. Studies of higher education similarly describe growing use of generative AI in teaching and learning, as well as concerns around assessment, quality, and governance [2, 4, 8, 11].

Student engagement is a prominent construct in learning analytics. In higher education, engagement is multifaceted and not always clearly defined [5]. Distance-education studies likewise treat engagement and performance prediction as central analytical targets, while noting that subsequent intervention actions are often underspecified [6]. This is consequential in higher education: behavioral traces are readily collected, but support strategies are difficult to standardise.

2.2 Explainability and Intervention-Oriented Analytics

Research on explainability focuses on transparency and human-centred AI. Alfredo et al. [7] argue that learning analytics and AI should support human agency and make their processes comprehensible to teachers. Explainable machine-learning approaches can also support educational prediction while increasing interpretability [12]. Research on self-regulated and socially shared learning examines how analytics and AI can characterise learners' regulation processes and collaborative activity [10, 13].

Intervention-aware analytics is also emerging. Reviews of learning-analytics dashboards emphasize designs that connect data displays with student learning and action [3], while explainable learning-analytics research examines the stability and interpretability of student-success predictions [14]. Together, these studies indicate that prediction is more educationally useful when analytical models are integrated with an action system.

2.3 Research Gap

The literature provides substantial evidence on AI adoption, learning analytics, and explainability in higher education. Fewer studies, however, integrate behavioral-signal engineering, interpretable segmentation, multiclass prediction, and practical action allocation in a single reproducible workflow. This study addresses that gap by using digital trace data to produce differentiated instructional support plans rather than risk labels alone. The selected studies are summarized in Table 1.

3. CONCEPTUAL FOUNDATION AND ANALYTICAL FRAMEWORK

3.1 Analytical Premise

The proposed framework assumes that observable digital behavior contains signals that can be aggregated, interpreted, and translated into differentiated instructional action. In digital higher education, engagement is not directly observable as a latent construct; instead, it is approximated through trace indicators such as interaction frequency, resource access, and participation. Let $\mathbf{x}_i \in \mathbb{R}^p$ denote the predictor vector for student i, and let $y_i \in \{H, M, L\}$ denote the observed performance class, where H, M, and L represent high, medium, and low performance, respectively.

To provide a concise behavioral representation, the study constructs a composite engagement index from four numerical indicators: raised hands, visited resources, viewed announcements, and discussion activity. For student i, the index is

$$\mathrm{EI}_i = \frac{1}{4}\sum_{k=1}^{4} z_{ik}, \tag{1}$$

where z_{ik} is the standardized value of behavioral variable k for student i. The index does not replace the original variables in all analyses; rather, it provides an interpretable summary for segmentation and model interpretation.

The framework segments students into support profiles by combining EI_i with attendance and contextual indicators. The segmentation distinguishes digitally active students, low-touch absentees, moderate-progress students, and support-sensitive students. This intermediate layer reduces the distance between raw behavioral signals and educational interpretation and supplies an additional explanatory feature for the predictive model.

3.2 Mathematical Specification of the Predictive Layer

The classification problem is treated as a multiclass learning task. For multinomial logistic regression, the conditional probability of class c for student i is

$$P(y_i = c \mid \mathbf{x}_i) = \frac{\exp(\beta_c^\top \mathbf{x}_i)}{\sum_{j\in\{H,M,L\}} \exp(\beta_j^\top \mathbf{x}_i)}. \tag{2}$$

The predicted class is

$$\hat{y}_i = \arg\max_{c\in\{H,M,L\}} P(y_i = c \mid \mathbf{x}_i). \tag{3}$$

For ensemble tree models, class assignment follows majority voting across T trees:

$$\hat{y}_i = \arg\max_{c\in\{H,M,L\}} \frac{1}{T}\sum_{t=1}^{T} \mathbb{I}(h_t(\mathbf{x}_i) = c), \tag{4}$$

Table 1. Selected post-2022 studies informing the analytical framing.

Study	Focus, design, and main contribution
Chiu et al. (2023)	AI in education; systematic review. Synthesizes opportunities, challenges, and future research directions.
Bergdahl et al. (2024)	Student engagement in learning analytics; systematic review. Clarifies the multidimensional nature of engagement and actionability challenges.
Palanci et al. (2024)	Learning analytics in distance education; systematic review. Documents the growing use of analytics for engagement, prediction, and instructional redesign.
Alfredo et al. (2024)	Human-centred learning analytics; systematic review. Emphasizes interpretability, human oversight, and responsible human–AI collaboration.
Ghimire et al. (2024)	Explainable AI for student-score estimation; empirical modelling. Demonstrates the value of explainable machine-learning models in educational prediction.
AlGhamdi (2024)	AI-generated feedback; blinded empirical study. Shows that AI feedback can improve technical-writing performance in computing education.
Lee et al. (2024)	Generative AI in higher education; educator survey. Examines opportunities and risks from teaching staff perspectives.
Paulsen and Lindsay (2024)	Learning analytics dashboards; systematic review. Synthesizes design approaches that connect analytics to student learning.
Tiukhova et al. (2024)	Explainable learning analytics; empirical study of model stability. Examines the stability and interpretability of student-success predictions.
Alhazbi et al. (2024)	Learning analytics and self-regulated learning; systematic review. Reviews empirical measures of self-regulated learning in higher education.
Dang et al. (2024)	AI-supported analysis of socially shared regulation; exploratory study. Examines collaborative learning interactions through AI-enabled analytics.
Dabis and Csáki (2024)	Generative-AI policy and ethics in higher education; institutional policy analysis. Reviews early policy responses across universities.
Pierrès et al. (2024)	AI and students with disabilities in higher education; scoping review. Identifies accessibility and inclusion considerations.
Kallunki et al. (2024)	Generative AI in higher education; faculty and staff perspectives. Examines perceived effects and institutional support needs.
Ogunleye et al. (2024)	Generative AI for teaching and learning; systematic review. Synthesizes research trends and pedagogical implications.

where $h_t(\mathbf{x}_i)$ is the prediction of tree t and $\mathbb{I}(\cdot)$ is the indicator function. Logistic Regression, Random Forest, and Extra Trees use the same train–test split and preprocessing logic to ensure internally consistent model comparisons.

Model performance is assessed using accuracy, macro-precision, macro-recall, macro-F1, and macro-AUC. For class c,

$$\text{Precision}_c = \frac{TP_c}{TP_c + FP_c}, \qquad \text{Recall}_c = \frac{TP_c}{TP_c + FN_c}. \quad (5)$$

The macro-F1 score is

$$F_1^{\text{macro}} = \frac{1}{C}\sum_{c=1}^{C} \frac{2\,\text{Precision}_c\,\text{Recall}_c}{\text{Precision}_c + \text{Recall}_c}, \quad (6)$$

where $C = 3$.

3.3 Action Translation and Support-Plan Allocation

To move from prediction to action, the framework applies a priority score that ranks students by intervention intensity:

$$\text{PS}_i = 0.45\,\hat{p}_{iL} + 0.20A_i + 0.20(1 - g_{\text{VR},i}) + 0.15(1 - g_{\text{RH},i}), \quad (7)$$

where $\hat{p}_{iL}$ is the predicted probability of class L, A_i is the absence indicator, $g_{\text{VR},i}$ is normalized resource visitation, and $g_{\text{RH},i}$ is normalized hand-raising behavior. The score does not replace class prediction; it ranks students within and across predicted classes for support prioritization.

The decision layer allocates cases to a small portfolio of interpretable instructional plans. Students predicted as low-performing with high absence receive attendance-first intervention; low-performing students without high absence receive structured progression coaching; medium-performing students with weak parental-engagement signals receive family-engagement reinforcement; high-performing students receive a challenge-and-extend pathway; and remaining cases are routed to adaptive engagement support. Thus, the final analytical output is an actionable support plan.

4. MATERIALS AND METHODS

4.1 Dataset and Variables

The analysis uses the public xAPI-Edu-Data dataset, which contains 480 student observations and demographic, contextual, and behavioral characteristics derived from a digital learning environment. The target variable is a three-class performance label (H, M, L). Numerical behavioral indicators include raised hands, visited resources, announcement views, and discussion participation. Contextual indicators include educational stage, grade level, topic, section, semester, student–guardian relation, parental-survey participation, parental-school satisfaction, and student absence days.

4.2 Preprocessing and Experimental Design

The analysis proceeds in four stages. First, the dataset is cleaned and recoded, and numerical behavioral variables are standardized to construct the engagement index. Second, support profiles are derived from engagement, attendance, and support-related variables. Third, multiclass prediction models are estimated. Categorical variables are one-hot encoded, numerical variables are standardized where needed, and the sample is divided into training and hold-out subsets using stratified sampling. Fourth, the selected model is interpreted through permutation-importance analysis and translated into support plans through the rule-based action layer.

Model comparison uses stratified cross-validation. The final model is selected according to macro-F1 because a three-class outcome requires balanced treatment of all classes rather than accuracy alone. The analyses are reproducible and are included in the project package.

5. RESULTS

5.1 Descriptive Characteristics of the Sample

The sample comprises 142 students in class H, 211 in class M, and 127 in class L. Support-profile segmentation yields four groups: 189 digitally active students, 114 low-touch absentees, 100 moderate-progress students, and 77 support-sensitive students. The descriptive evidence shows marked behavioral separation across classes. Students in class H display higher levels of hand-raising, resource visits, announcement views, and discussion activity than students in class L, while class M occupies an intermediate position. Absence patterns follow the opposite direction, with a larger proportion of students in the low-performance group having more than seven absence days. Figure 1 summarizes the class and absence patterns, while Figure 2 compares the behavioral signatures of the four support profiles.

5.2 Comparative Model Performance

Table 2 reports cross-validated performance for the candidate models. Random Forest provides the strongest overall macro-F1 score (0.8048), followed by Extra Trees (0.8033) and Logistic Regression (0.7757). The two ensemble models perform similarly, but Random Forest has the highest macro-precision and is selected for final hold-out evaluation. Figure 3 visualizes the cross-validation comparison.

Table 2. Cross-validated model performance.

Model	Accuracy	Macro-Precision	Macro-Recall	Macro-F1
Random Forest	0.8000	0.8161	0.8004	0.8048
Extra Trees	0.7979	0.8039	0.8064	0.8033
Logistic Regression	0.7688	0.7835	0.7726	0.7757

5.3 Hold-Out Performance of the Selected Model

On the hold-out sample, the selected Random Forest model achieves accuracy of 0.8021, macro-precision of 0.8204, macro-recall of 0.8010, macro-F1 of 0.8084, and macro-AUC of 0.9140 (Table 3). Class-specific metrics are reported in Table 4. Figure 4 shows the confusion matrix and overall hold-out metrics. The strongest separation is for class L, while residual confusion remains between classes H and M, which plausibly share some behavioral characteristics.

Table 3. Hold-out performance of the selected Random Forest model.

Metric	Value
Accuracy	0.8021
Macro-Precision	0.8204
Macro-Recall	0.8010
Macro-F1	0.8084
Macro-AUC	0.9140

Table 4. Class-specific hold-out performance.

Class	Precision	Recall	F_1
H	0.7442	0.7879	0.7654
L	0.9375	0.8824	0.9091
M	0.7797	0.7321	0.7552

5.4 Predictor Importance and Support-Plan Portfolio

Permutation-importance analysis identifies StudentAbsenceDays as the most influential predictor, followed by EngagementIndex, gender, Relation, SupportProfile, VisITedResources, and raisedhands. The ranking suggests that attendance and sustained participation provide the strongest signals for performance differentiation in this dataset.

The action-allocation layer translates predictions into a small portfolio of support plans. Within the hold-out sample, the largest assigned category is adaptive engagement support (31 students), followed by the challenge-and-extend pathway (25), attendance-first intervention (23), family-engagement reinforcement (16), and structured progression coaching (1). Actionable analytics should therefore distinguish not only severe risk but also cases requiring moderate support, attendance stabilisation, enrichment, or family-linked engagement strategies. Figure 5 presents the predictor-importance ranking and the distribution of assigned support plans.

6. DISCUSSION

The empirical results support three main observations. First, simple digital-trace variables contain substantial information for analysing student performance. Attendance, resource access, and participation behavior provide the strongest predictive signals, consistent with research identifying engagement behavior as central to higher-education learning analytics [5, 6].

Second, the explanatory value of AI improves when prediction is embedded within intermediate analytical layers. The engagement index and support-profile segmentation provide more interpretable representations of student behavior than a single predictive probability. This is consistent with human-centred learning-analytics research, which argues that AI systems should make their logic intelligible to educators while preserving space for pedagogical judgment [7]. It also accords with dashboard and explainable-learning-analytics research emphasizing the connection between interpretable analysis and student learning support [3, 14].

Third, the action-allocation layer illustrates how AI can sup-



Figure 1. Descriptive diagnostics of performance, support profiles, engagement, and absence.

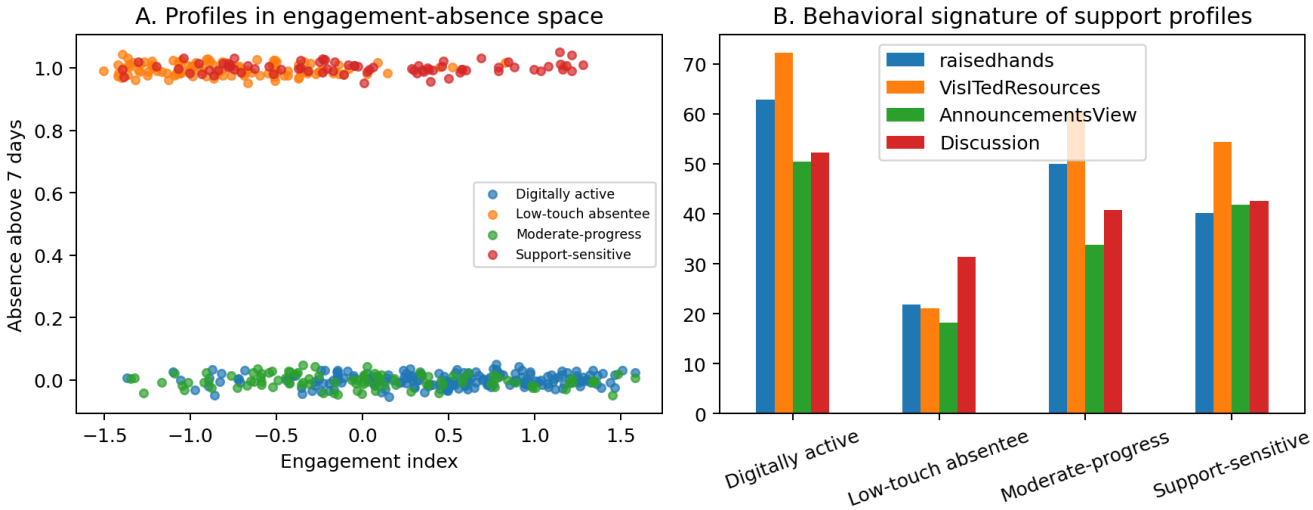


Figure 2. Behavioral signatures of the support profiles.

port differentiated instructional planning. The resulting portfolio is heterogeneous: students may need attendance stabilisation, adaptive engagement support, family-linked reinforcement, or extension-oriented pathways. Educational response is rarely one-dimensional; in practice, undifferentiated alerts may overstate some cases and underspecify others. A portfolio-based approach offers a more usable basis for instructional coordination.

Several limitations should be acknowledged. The dataset is moderate in size and represents one specific educational context. The support plans are analytically inferred rather than validated through field implementation. In addition, the action-allocation rules are deliberately interpretable and there-fore simpler than optimization-based assignment procedures. Future research may extend the framework to larger contem-porary datasets, compare alternative intervention-allocation mechanisms, and examine how generative-AI tools may be integrated into support-plan execution.

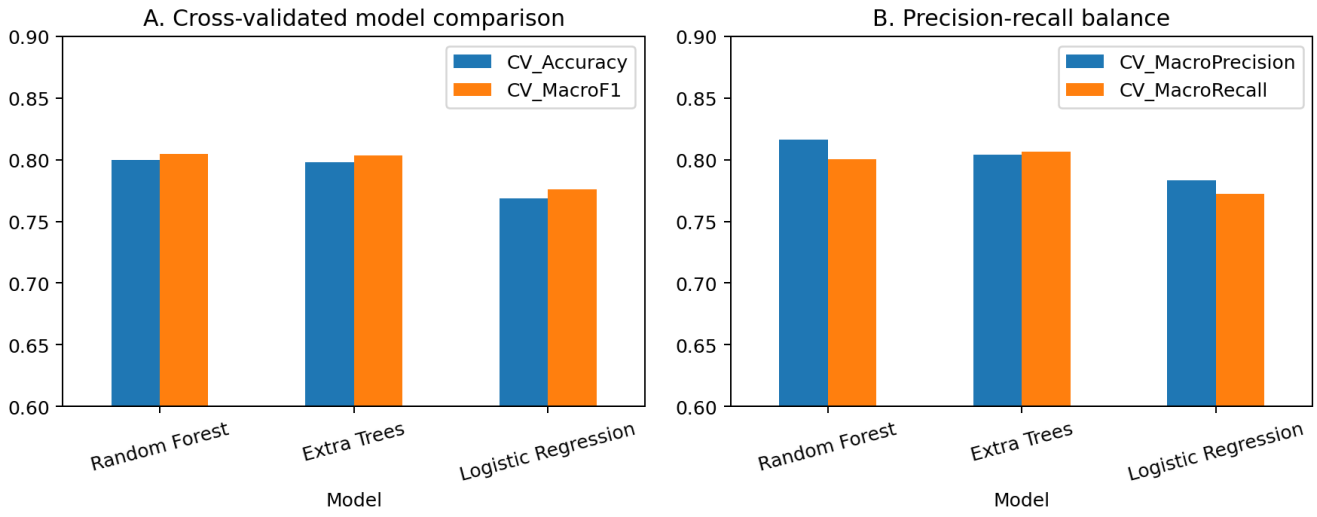


Figure 3. Cross-validated benchmarking of candidate models.

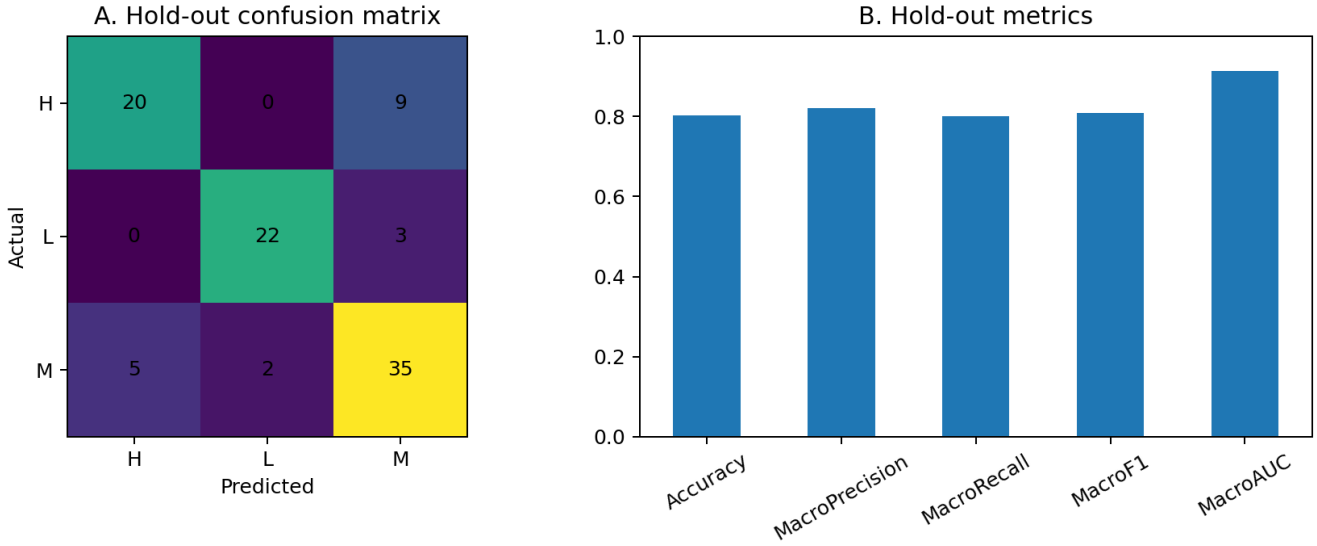


Figure 4. Hold-out performance of the selected model.

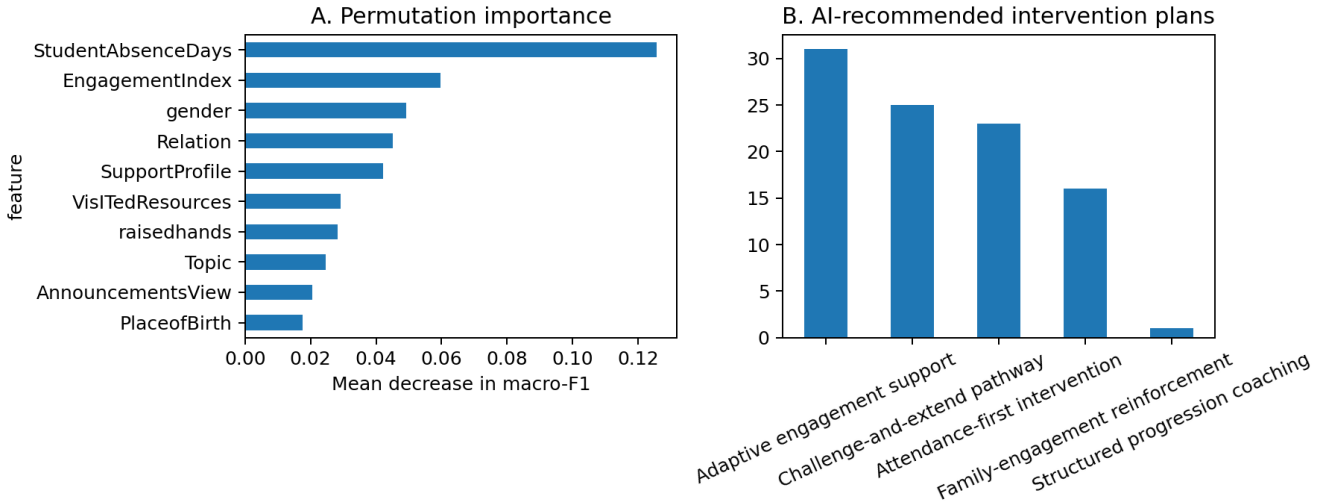


Figure 5. Predictor importance and distribution of AI-recommended support plans.

7. PRACTICAL IMPLICATIONS

The results have direct implications for AI-enabled education-technology management in higher education. First, institutions should distinguish clearly between predictive modelling and support design: a model that performs well statistically but does not lead to a credible response mechanism has limited practical value. Second, analytical systems should com-

bine attendance, behavioral, and support-related signals rather than depend on a single engagement indicator. Third, instructional support should be managed as a portfolio. Advising units, instructors, and student-success teams need multiple response modes, including stabilisation, coaching, reinforcement, and enrichment pathways.

These implications are especially important for institutions seeking to operate AI-enabled student-support systems through transparent and ethical governance. From this perspective, explainability is a technical attribute that helps educators make informed decisions based on educational data.

8. CONCLUSION

This research designed and evaluated an explainable AI approach to engagement-responsive support in digital higher education. It integrated engagement-index construction, support-profile segmentation, multiclass performance prediction, and support-action allocation in a replicable workflow using the public xAPI-Edu-Data dataset. The Random Forest model achieved good hold-out results, while the intermediate profile and action layers supported interpretability and action relevance.

The results demonstrate how AI in education technology can be structured around decision support rather than prediction alone. Associating digital traces with actionable support profiles provides an explainable foundation for higher-education intervention.

REFERENCES

[1] T. K. F. Chiu, Q. Xia, X. Zhou, C. S. Chai, and M. Cheng, "Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education," *Computers and Education: Artificial Intelligence*, vol. 4, p. 100118, 2023.

[2] D. Lee, M. Arnold, A. Srivastava, K. Plastow, P. Strelan, F. Ploeckl, D. Lekkas, and E. Palmer, "The impact of generative AI on higher education learning and teaching: A study of educators' perspectives," *Computers and Education: Artificial Intelligence*, vol. 6, p. 100221, 2024.

[3] L. Paulsen and E. Lindsay, "Learning analytics dashboards are increasingly becoming about learning and not just analytics: A systematic review," *Education and Information Technologies*, vol. 29, pp. 14 279–14 308, 2024.

[4] V. Kallunki, P. Kinnunen, E. Pyörälä, A. Haarala-Muhonen, N. Katajavuori, and L. Myyry, "Navigating the evolving landscape of teaching and learning: University faculty and staff perceptions of the artificial intelligence-altered terrain," *Education Sciences*, vol. 14, no. 7, p. 727, 2024.

[5] N. Bergdahl, M. Bond, J. Sjöberg, M. Dougherty, and E. Oxley, "Unpacking student engagement in higher education learning analytics: A systematic review," *International Journal of Educational Technology in Higher Education*, vol. 21, p. 63, 2024.

[6] A. Palanci, R. M. Yı lmaz, and Z. Turan, "Learning analytics in distance education: A systematic review study," *Education and Information Technologies*, vol. 29, pp. 22 629–22 650, 2024.

[7] R. Alfredo, V. Echeverria, Y. Jin, L. Yan, Z. Swiecki, D. Gašević, and R. Martinez-Maldonado, "Human-centred learning analytics and AI in education: A systematic literature review," *Computers and Education: Artificial Intelligence*, vol. 6, p. 100215, 2024.

[8] A. Dabis and C. Csáki, "AI and ethics: Investigating the first policy responses of higher education institutions to the challenge of generative AI," *Humanities and Social Sciences Communications*, vol. 11, p. 1006, 2024.

[9] O. Pierrès, M. Christen, F. M. Schmitt-Koopmann, and A. Darvishy, "Could the use of AI in higher education hinder students with disabilities? a scoping review," *IEEE Access*, vol. 12, pp. 27 810–27 828, 2024.

[10] S. Alhazbi, A. Al-ali, A. Tabassum, A. Al-Ali, A. Al-Emadi, T. Khattab, and M. A. Hasan, "Using learning analytics to measure self-regulated learning: A systematic review of empirical studies in higher education," *Journal of Computer Assisted Learning*, vol. 40, no. 4, pp. 1658–1674, 2024.

[11] B. Ogunleye, K. I. Zakariyyah, O. Ajao, O. Olayinka, and H. Sharma, "A systematic review of generative AI for teaching and learning practice," *Education Sciences*, vol. 14, no. 6, p. 636, 2024.

[12] S. Ghimire, S. Abdulla, L. P. Joseph, S. Prasad, A. Murphy, A. Devi, P. D. Barua, R. C. Deo, R. Acharya, and Z. M. Yaseen, "Explainable artificial intelligence–machine learning models to estimate overall scores in tertiary preparatory general science course," *Computers and Education: Artificial Intelligence*, vol. 7, p. 100331, 2024.

[13] B. Dang, A. Nguyen, and S. Järvelä, "Deliberative interactions for socially shared regulation in collaborative learning: An AI-driven learning analytics study," *Journal of Learning Analytics*, vol. 11, no. 3, pp. 192–209, 2024.

[14] E. Tiukhova, P. Vemuri, N. G. López Flores, A. S. Islind, M. Óskarsdóttir, S. Poelmans, B. Baesens, and M. Snoeck, "Explainable learning analytics: Assessing the stability of student success prediction models by means of explainable AI," *Decision Support Systems*, vol. 182, p. 114229, 2024.

[15] R. AlGhamdi, "Exploring the impact of ChatGPT-generated feedback on technical writing skills of computing students: A blinded study," *Education and Information Technologies*, vol. 29, pp. 18 901–18 926, 2024.