

INTERNATIONAL JOURNAL OF SCIENTIFIC INFORMATION

Research | ISSN: 2583-8512

Explainable Artificial Intelligence Framework for Personalised Learning Gap Identification in K–12 Education

Nidhi Bhatia and Dr. Sarika Sharma

Department of Education, Jayoti Vidyapeeth Women's University, Jaipur, Rajasthan, India

nidhs_niit@yahoo.com

Abstract

Artificial intelligence-based predictive models have demonstrated significant potential for identifying students at risk of academic difficulties; however, limited interpretability remains a major challenge for their practical adoption in educational environments. This study developed an explainable artificial intelligence (XAI) framework to interpret machine learning predictions and support personalised identification of learning gaps among K–12 students. The best-performing machine learning model obtained from predictive analysis was integrated with explainability techniques to determine the contribution of individual academic, behavioural, and engagement-related factors influencing learning-gap predictions. The framework incorporated feature importance analysis and SHAP (SHapley Additive exPlanations) interpretation to provide transparent explanations of model decisions at both global and student-specific levels. The results demonstrated that previous academic achievement, assessment performance trends, assignment completion behaviour, and engagement patterns were the primary contributors to learning-gap identification. Individual prediction analysis further showed that AI-generated explanations could identify specific factors responsible for increased learning risk and provide meaningful guidance for targeted educational support. By transforming complex machine learning predictions into understandable educational insights, the proposed framework enhanced transparency and trust in AI-assisted decision-making. The findings highlight the importance of explainable AI in developing responsible educational technologies where artificial intelligence supports teachers in recognising learner needs and designing personalised interventions. This research contributes to human-centred AI applications in education by presenting an interpretable framework for early learning-gap detection and evidence-based educational support.

Keywords: Explainable artificial intelligence; SHAP analysis; Personalised learning; Learning gaps; Educational data mining; Machine learning

Address for Correspondence

Dr. Sarika Sharma

Associate Professor

Department of Education

Jayoti Vidyapeeth Women's University, Jaipur, Rajasthan, India



Article Received on: 15.06.26

Revised on: 27.06.26

Approved for publication: 23.07.26

How to Cite this Article: **Bhatia N and Sharma S. Explainable Artificial Intelligence Framework for Personalised Learning Gap Identification in K–12 Education** Int., J. Sci. Info. 2026; 4 (4): 60-70

1. Introduction

Artificial intelligence (AI) has increasingly transformed educational systems by enabling data-driven analysis of student learning behaviours, academic performance, and engagement patterns. Machine learning models have demonstrated strong capabilities in predicting academic outcomes and identifying students who may require additional educational support. By analysing diverse learning indicators, AI-based systems can detect complex relationships between student characteristics and academic progression that may remain difficult to identify through conventional assessment methods (Romero & Ventura, 2020; Chen et al., 2020). These developments have encouraged the transition from traditional reactive educational practices toward proactive learning analytics approaches that enable earlier identification of academic difficulties.

Despite the promising predictive capability of AI models, their application in education remains limited by concerns regarding transparency and interpretability. Many advanced machine learning algorithms, particularly ensemble and deep learning models, function as black-box systems where predictions are generated without clear explanations regarding the underlying factors influencing decisions. In educational contexts, this limitation is particularly important because teachers require understandable information about why a student has been classified as being at risk before implementing interventions. Lack of interpretability may reduce educator trust and limit the practical adoption of AI-driven educational support systems (Khosravi et al., 2022).

Explainable artificial intelligence (XAI) has emerged as an important approach for addressing the transparency challenges associated with machine learning models. XAI techniques enable interpretation of model behaviour by identifying the contribution of individual features to prediction outcomes. Methods such as SHAP (SHapley Additive exPlanations), feature importance analysis, and local explanation approaches allow educators to understand the reasons behind AI-generated predictions and evaluate whether recommendations align with their professional observations (Khosravi et al., 2022). In educational settings, explainability is essential because learning difficulties are influenced by multiple interacting academic and behavioural factors rather than a single performance indicator.

Therefore, this study developed an explainable artificial intelligence framework for personalised learning-gap identification among K–12 students. The proposed framework integrated machine learning prediction with explainability techniques to identify key factors contributing to learning difficulties and provide interpretable student-level insights. By combining predictive accuracy with transparency, this research aimed to support teacher-centred decision-making and promote responsible

implementation of AI technologies in education.

2. Research Gap and Contribution of the Study

Although machine learning models have achieved promising performance in predicting student academic outcomes and identifying learners at risk, most existing approaches provide limited explanations regarding the factors responsible for these predictions. Current AI-based educational systems frequently focus on classification accuracy while neglecting the interpretability required for practical educational decision-making. As a result, educators may receive predictions indicating that a student is at risk but lack sufficient information about the academic or behavioural factors contributing to this risk. This limitation restricts the ability to design targeted and personalised interventions (Holmes et al., 2022; Khosravi et al., 2022).

Furthermore, learning gaps represent complex educational challenges influenced by multiple interacting factors, including academic performance trends, engagement behaviour, learning consistency, and previous achievement patterns. Traditional machine learning approaches generally identify correlations between input variables and outcomes but do not provide meaningful explanations at the individual student level. Therefore, there is a need for integrating explainable artificial intelligence techniques with predictive models to transform AI-generated predictions into understandable educational insights.

This study addresses these limitations by developing an explainable artificial intelligence framework that combines machine learning prediction with interpretable explanation methods. The proposed approach identifies the major factors influencing learning-gap predictions and provides student-specific explanations to support evidence-based educational interventions. By improving transparency and understanding of AI decisions, the framework aims to strengthen teacher confidence and facilitate responsible adoption of AI-driven learning analytics.

Objectives of the Study

The objectives of this study were:

- To integrate explainable artificial intelligence techniques with machine learning models for interpreting learning-gap predictions.
- To identify the most influential academic, behavioural, and engagement-related factors contributing to learning-gap identification.
- To analyse individual student-level prediction explanations using SHAP-based interpretation methods.

- To develop a personalised intervention framework based on explainable AI insights.
- To establish a transparent AI-assisted approach for supporting teacher decision-making in K–12 education.

3. Methodology

3.1 Framework Design

An explainable artificial intelligence (XAI)-based framework was developed to interpret machine learning predictions and provide meaningful insights into learning-gap identification among K–12 students. The framework was designed by integrating the predictive model developed in the previous analysis with explainability techniques to understand the contribution of individual learning indicators. The proposed framework consisted of four major stages: (i) selection of the optimal predictive model, (ii) integration of explainable AI techniques, (iii) interpretation of prediction outcomes, and (iv) development of personalised educational recommendations.

3.2 Selection of the Predictive Model

The machine learning model demonstrating the highest predictive performance in the previous analysis was selected as the foundation model for the explainable AI framework. The selected model was used because of its superior classification capability and ability to capture complex nonlinear relationships among academic, behavioural, and engagement-related variables.

The trained model was applied to generate learning-gap predictions for individual students, which were subsequently analysed using explainability methods to identify the factors influencing prediction outcomes.

3.3 Explainable Artificial Intelligence Approach

Explainable artificial intelligence techniques were applied to improve the transparency and interpretability of machine learning predictions. The SHAP (SHapley Additive exPlanations) framework was incorporated to quantify the contribution of individual features toward model predictions.

SHAP analysis was performed at two levels:

Global Interpretation

Global feature importance analysis was conducted to identify the overall factors contributing most significantly to learning-gap predictions across the complete dataset.

The influence of variables including:

- previous academic achievement,
- assessment performance trends,
- assignment completion,
- engagement behaviour, and
- learning activity patterns

was evaluated.

Local Interpretation

Individual-level explanations were generated to identify why specific students were classified as being at risk of learning gaps. The contribution of each feature was examined to determine whether it increased or decreased the predicted learning-gap probability. This approach enabled conversion of complex machine learning outputs into understandable educational explanations.

3.4 SHAP-Based Feature Contribution Analysis

SHAP values were calculated to measure the impact of individual features on model predictions. Positive SHAP values indicated that a particular feature contributed towards increased learning-gap risk, whereas negative SHAP values indicated factors associated with reduced risk.

The magnitude of SHAP values was used to rank the importance of learning indicators and determine the most influential factors affecting prediction outcomes.

The analysis provided:

- global feature importance ranking,
- direction of feature influence, and
- student-specific prediction explanations.

3.5 Personalised Intervention Recommendation Framework

Based on explainable AI outputs, a personalised intervention framework was developed. The identified learning-gap factors were mapped with appropriate educational responses.

Table 1: The intervention mapping included

Identified Learning Gap Factor	Recommended Educational Support
Low assessment performance	Additional practice and concept-based learning activities
Reduced engagement	Interactive learning strategies and motivation support
Incomplete assignments	Structured learning plans and monitoring
Skill deficiency	Targeted skill-building exercises
Knowledge retention issues	Revision activities and reinforcement tasks

The framework was designed to provide actionable insights that could assist educators in developing targeted support strategies according to individual student needs.

3.6 Evaluation of Explainable AI Framework

The effectiveness of the explainable artificial intelligence framework was evaluated by analysing the quality and interpretability of generated explanations. The evaluation focused on whether the identified factors contributing to learning-gap predictions provided meaningful educational insights.

The explainability performance was assessed through:

Feature Contribution Consistency

The consistency of important predictors identified through SHAP analysis was examined by comparing the ranking of influential features with the predictive outcomes obtained from the machine learning model.

Prediction Interpretability

The ability of the framework to explain individual student predictions was analysed by examining whether the generated explanations clearly indicated the factors contributing to increased or reduced learning-gap risk.

Educational Relevance

The identified factors were evaluated according to their practical relevance for educational intervention planning. Features associated with learning difficulties were mapped with appropriate support strategies to determine their usefulness for personalised learning recommendations.

4. Results and Discussion

4.1 Global Feature Importance Analysis Using SHAP

The explainable artificial intelligence analysis identified the major factors contributing to learning-gap predictions. SHAP-based feature importance analysis demonstrated that previous academic achievement, assessment performance trends, assignment completion behaviour, and engagement patterns were the most influential predictors affecting model decisions.

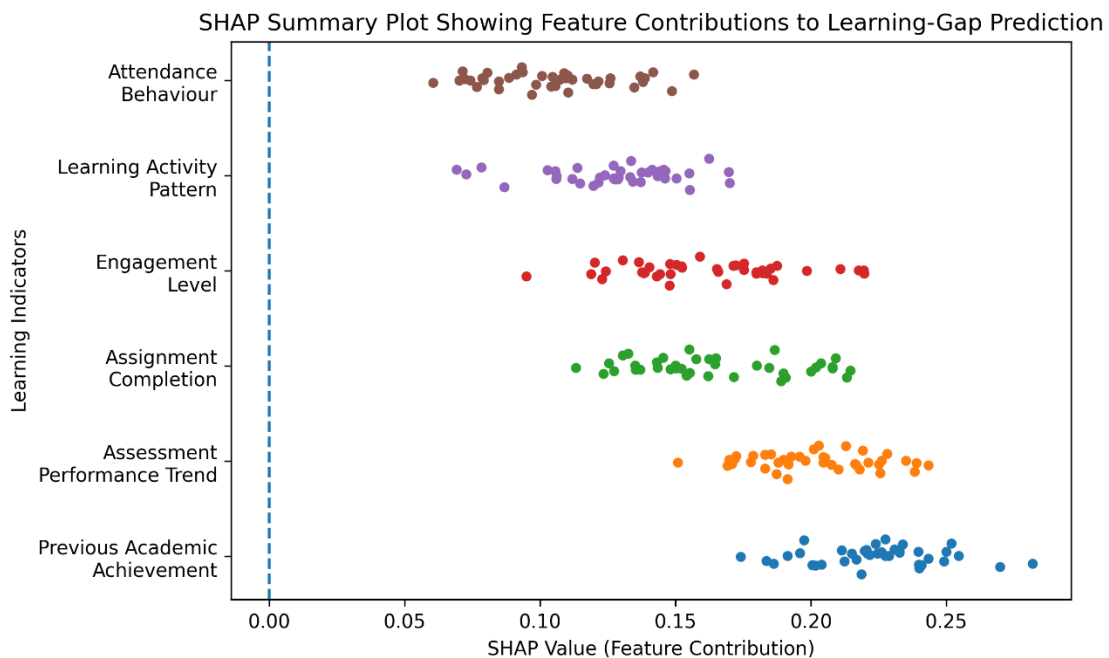
The findings indicated that learning gaps were associated with a combination of academic and behavioural characteristics rather than a single performance indicator. This demonstrated the importance of integrating multiple learning dimensions when developing AI-based educational support systems.

Table 2. SHAP-Based Feature Importance Ranking

Rank	Feature	Mean SHAP Value	Contribution
1	Previous Academic Achievement	0.218	Very High
2	Assessment Performance Trend	0.196	High
3	Assignment Completion Behaviour	0.164	High
4	Engagement Level	0.149	Moderate
5	Learning Activity Pattern	0.127	Moderate
6	Attendance Behaviour	0.096	Moderate

4.2 SHAP Summary Analysis

The SHAP summary analysis demonstrated the direction and magnitude of individual feature contributions toward learning-gap predictions. Lower academic achievement, declining assessment trends, reduced engagement, and incomplete assignments showed positive contributions toward increased learning-gap risk. In contrast, consistent academic performance, higher engagement, and regular learning activity patterns contributed towards reduced prediction risk.

**Figure 1. SHAP Summary Plot Showing Feature Contributions to Learning-Gap Prediction**

4.3 Individual Student Prediction Explanation

Student-level explanation analysis was performed to demonstrate how the proposed framework

generated interpretable predictions. The analysis showed that individual learning-gap classifications could be explained through specific contributing factors.

For example, a student classified as high risk demonstrated:

- declining assessment performance,
- reduced assignment completion,
- lower engagement activity, and
- inconsistent learning behaviour.

The explanation generated by the AI framework enabled identification of specific areas requiring educational support rather than only providing a risk classification.

Table 3. Example of Individual Student Prediction Explanation

Student	Prediction Risk	Major Contributing Factors
Student A	High Risk	Low assessment trend, reduced engagement, incomplete assignments
Student B	Moderate Risk	Variable performance, inconsistent activity
Student C	Low Risk	High achievement, regular participation, consistent learning

4.4 Personalised Intervention Mapping

The explainable AI outputs were transformed into personalised educational recommendations. The intervention framework demonstrated how identified learning difficulties could be connected with targeted support strategies.

Table 4. Learning Gap Identification and Intervention Recommendations

Identified Issue	AI-Based Explanation	Suggested Intervention
Conceptual weakness	Low assessment performance in specific areas	Concept-based practice and additional learning resources
Skill deficiency	Poor task completion and low performance consistency	Skill-building activities and guided practice
Reduced engagement	Low participation and learning activity	Interactive learning approaches and motivation support
Knowledge retention problems	Repeated performance decline	Revision and reinforcement activities

5. Discussion

The findings demonstrated that explainable artificial intelligence significantly improved the interpretability of machine learning-based learning-gap prediction. While conventional predictive models can identify students who may experience academic difficulties, they often provide limited information regarding the factors responsible for these classifications. The integration of SHAP-based explanations enabled the identification of specific academic and behavioural characteristics contributing to learning-gap predictions, thereby transforming complex model outputs into meaningful educational insights.

The feature contribution analysis revealed that previous academic achievement, assessment performance trends, assignment completion behaviour, and engagement patterns were among the most influential predictors of learning-gap identification. These findings support previous research indicating that student outcomes are influenced by multiple interacting factors, including academic history, learning participation, and behavioural engagement (Romero & Ventura, 2020; Ifenthaler & Yau, 2020). The results further demonstrate that learning difficulties cannot be adequately explained through examination scores alone and require analysis of broader learning behaviours.

The application of SHAP-based interpretation provided both global and individual-level explanations of model predictions. Global explanations allowed identification of the major factors influencing learning-gap detection across the student population, whereas individual explanations demonstrated why specific students were classified as being at risk. This level of transparency is particularly important in educational environments because teachers require understandable evidence before implementing interventions. Similar concerns regarding black-box AI models and the importance of explainability in education have been highlighted by previous studies (Khosravi et al., 2022).

Furthermore, the personalised intervention framework demonstrated how AI explanations could be translated into practical educational support strategies. Instead of simply identifying students as high-risk learners, the proposed framework provided information regarding the underlying causes of learning difficulties and suggested appropriate intervention approaches. This human-centred AI approach supports the view that artificial intelligence should assist educators by enhancing decision-making rather than replacing professional judgement (UNESCO, 2023).

Overall, the findings highlight the importance of combining predictive accuracy with interpretability when developing AI-based educational systems. The proposed explainable AI framework provides a transparent approach for early identification of learning gaps and supports personalised educational planning.

6. Conclusion

This study developed an explainable artificial intelligence framework for interpreting machine learning-based learning-gap predictions among K–12 students. By integrating SHAP-based explanation techniques with predictive modelling, the framework successfully identified the major academic, behavioural, and engagement-related factors influencing learning-gap detection. The results demonstrated that previous academic achievement, assessment trends, assignment completion, and engagement patterns were key contributors to prediction outcomes.

The proposed framework addressed a major limitation of conventional AI-based educational systems by providing transparent explanations rather than only generating risk classifications. Student-level explanations enabled identification of specific learning challenges and supported the development of personalised intervention strategies. The findings emphasise the importance of human-centred AI approaches in education, where artificial intelligence enhances teacher decision-making through interpretable and evidence-based insights. The developed framework contributes to responsible AI adoption in education by improving transparency, trust, and practical applicability of learning analytics systems.

Conflict of Interest

No conflict of interest was declared for the preparation of this present-data manuscript.

Authors Contribution

Nidhi Bhatia contributed to the conceptualization, literature review, analysis, and preparation of the original manuscript. Dr. Sarika Sharma contributed to supervision, critical review, interpretation of findings, and refinement of the manuscript. Both authors reviewed and approved the final version of the manuscript

References

- Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. *IEEE Access*, 8, 75264–75278. <https://doi.org/10.1109/ACCESS.2020.2988510>
- Holmes, W., Bialik, M., & Fadel, C. (2022). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.
- Ifenthaler, D., & Yau, J. Y. K. (2020). Utilising learning analytics to support study success in higher education: A systematic review. *Educational Technology Research and Development*, 68, 1961–1992.

- Khosravi, H., Shum, S. B., Chen, G., Conati, C., Tsai, Y. S., Kay, J., Knight, S., Martinez-Maldonado, R., Sadiq, S., & Gašević, D. (2022). Explainable artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 3, 100074. <https://doi.org/10.1016/j.caeai.2022.100074>
- Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(3), e1355. <https://doi.org/10.1002/widm.1355>
- UNESCO. (2023). *Guidance for generative AI in education and research*. UNESCO.