

Dynamic Glocal XAI Framework for Machine Learning-Based Prediction of Education ROI with Hierarchical Explainability

Yiming Li¹, Xinyu Cai² and ABDULLAH²

(1. Department of Early Childhood Education, Yew Chung College of Early Childhood Education, Hong Kong Special Administrative Region 999077, China

2. College of Business, Jiaxing University, Jiaxing, Zhejiang 314001, China)

Abstract—We propose a novel Dynamic Glocal XAI framework for predicting education return on investment (ROI) that simultaneously addresses predictive accuracy and interpretability for diverse stakeholders. The framework integrates gradient-boosted trees with dynamic explainable AI (XAI) mechanisms, enabling both global trends and local instance-level insights through hierarchical clustering of Shapley-based explanations. The model optimizes a regularized objective function to minimize prediction errors while maintaining computational efficiency through TreeSHAP approximations. Moreover, the system introduces a hierarchical clustering module that groups similar explanations using a dual-term metric combining Euclidean distance and distributional divergence, thereby capturing both magnitude and pattern similarities in feature contributions. A gating network with attention mechanisms dynamically routes explanation requests between granular local insights and aggregated global summaries, adapting to stakeholder intent. Furthermore, the framework incorporates domain adaptation via an education-specific feature ontology, translating technical model outputs into pedagogically meaningful concepts through a bipartite graph neural network. The proposed method interfaces with existing education ROI systems through standardized APIs, replacing static regression coefficients with interactive SHAP-based visualizations for counterfactual analysis. Experimental validation demonstrates that the framework not only achieves competitive predictive performance but also provides actionable insights for policymakers and institutional administrators. The integration of dynamic glocal explainability with domain-aware feature interpretation represents a significant advancement over conventional black-box predictive models in education analytics

Index Terms—education return on investment, predictive analytics, explainable artificial intelligence, gradient-boosted trees, SHAP, hierarchical explainability, education analytics

I. INTRODUCTION

The measurement and prediction of return on investment (ROI) in education has become increasingly critical for resource allocation decisions at institutional and policy levels. Traditional approaches to education ROI analysis have relied heavily on linear regression models and cost-benefit analyses [1]. While these methods provide interpretable results, they often fail to capture the complex, non-linear relationships between educational inputs and outcomes. Recent advances in machine learning offer promising alternatives through their ability to model intricate patterns in educational data [2]. However, the adoption of these techniques has been hindered by their inherent lack of interpretability, particularly when making decisions that affect student outcomes and institutional budgets.

The challenge of model interpretability in education ROI prediction is multifaceted. First, different stakeholders require different levels of explanation granularity - policymakers need broad trends while school administrators require specific program-level insights [3]. Second, the heterogeneous nature of educational data demands explanations that can adapt to varying contexts and feature distributions. Third, the explanations must be framed in education-specific terminology to be actionable for non-technical decision-makers [4]. Current approaches typically address only one of these aspects, leaving a significant gap in practical applicability.

We propose a novel framework that combines predictive accuracy with adaptive explainability through three key innovations. First, we develop a gradient-boosted tree model specifically optimized for education ROI prediction, incorporating domain-specific feature engineering and regularization techniques. Second, we introduce a hierarchical explanation system that automatically switches between local and global interpretations based on the stakeholder's query context. This system uses Shapley Additive Explanations (SHAP) values [5] as its foundation but extends them through hierarchical clustering [6] to reveal both individual program characteristics and broader institutional patterns. Third, we implement a domain adaptation layer that translates technical

Corresponding author: ABDULLAH, abdullah1710bb@gmail.com.

Yiming Li is with the Department of Early Childhood Education, Yew Chung College of Early Childhood Education, Hong Kong Special Administrative Region, China, 999077 (e-mail: 25028506@stu.yccece.edu.hk). Xinyu Cai is with the College of Business, Jiaxing University, Jiaxing, Zhejiang, China, 314001 (e-mail: caixinyu@zjxu.edu.cn). ABDULLAH is with the College of Business, Jiaxing University, Jiaxing, Zhejiang, China, 314001 (e-mail: abdullah1710bb@gmail.com).

feature contributions into education-relevant concepts, making the explanations accessible to practitioners.

Our approach differs from existing methods in several important ways. Unlike conventional ROI analysis that provides static coefficients [7], our framework generates dynamic explanations that adapt to both the data context and user needs. Compared to standard SHAP implementations that treat each prediction independently [8], our hierarchical clustering of explanations reveals systemic patterns across programs and institutions. Furthermore, while previous work has focused either on prediction accuracy or interpretability in isolation [9], we explicitly optimize for both through a joint objective function.

The practical implications of this work are significant for education stakeholders at all levels. District administrators can identify which programs yield the highest ROI under specific conditions, while state policymakers can discern broader trends in resource allocation effectiveness. The framework’s adaptability also makes it suitable for diverse educational contexts, from K-12 interventions to workforce development programs [10]. By bridging the gap between predictive power and practical interpretability, our approach enables data-driven decision-making without sacrificing transparency.

The remainder of this paper is organized as follows: Section 2 reviews related work in education ROI prediction and explainable AI. Section 3 provides necessary background on predictive modeling and explainability techniques in education. Section 4 details our proposed framework and its components. Sections 5 and 6 present our experimental setup and results. Section 7 discusses implications and future directions, followed by conclusions in Section 8.

II. RELATED WORK

The intersection of predictive analytics and explainable AI in education has seen growing interest across multiple research communities. Existing approaches can be broadly categorized into three streams: traditional education ROI modeling, machine learning applications in education analytics, and explainability techniques for educational decision-making.

A. Traditional Approaches to Education ROI Prediction

Early work in education ROI focused primarily on cost-benefit analyses using linear regression models [1]. These methods provided straightforward interpretation through regression coefficients but struggled with the non-linear relationships inherent in educational systems. More sophisticated econometric approaches incorporated instrumental variables to address endogeneity concerns [3], though they remained limited in their ability to capture complex interactions between educational inputs and outcomes. Recent systematic reviews have highlighted the need for more flexible modeling frameworks that can accommodate the heterogeneous nature of education data [7].

B. Machine Learning in Education Analytics

The application of machine learning to education prediction tasks has demonstrated superior performance over traditional statistical methods across various domains. Gradient boosting machines have shown particular promise in student success prediction [8], while deep learning approaches have been

applied to large-scale learning analytics [9]. However, these models often operate as black boxes, making it difficult for educators and policymakers to understand the rationale behind predictions. Some studies have attempted to address this limitation by incorporating feature importance measures [2], but these provide only partial insights into model behavior.

C. Explainable AI in Educational Contexts

Recent advances in explainable AI have begun to bridge the gap between predictive accuracy and interpretability in education. SHAP values have emerged as a popular approach for explaining individual predictions [5], while LIME-based methods have been adapted for educational datasets [11]. The concept of “glocal” explanations - combining global and local interpretability - has gained traction in educational XAI research [12]. However, existing implementations often treat these as separate components rather than an integrated system. Hierarchical approaches to explanation aggregation have shown promise in student adaptability prediction [13], but have not been applied to ROI analysis.

The proposed framework advances beyond these existing approaches in several key aspects. Unlike traditional ROI models, we employ gradient-boosted trees capable of capturing complex non-linear relationships while maintaining interpretability through SHAP values. Compared to standard machine learning applications in education, our system explicitly optimizes for both prediction accuracy and explanation quality through a unified objective function. Furthermore, while previous XAI work in education has treated global and local explanations separately, our hierarchical clustering approach reveals systemic patterns while preserving individual prediction interpretability. The integration of domain-specific feature translation represents another novel contribution, making technical explanations accessible to education stakeholders without machine learning expertise.

III. BACKGROUND: PREDICTIVE MODELING AND EXPLAINABILITY IN EDUCATION

Educational institutions increasingly rely on data-driven approaches to optimize resource allocation and improve learning outcomes. Predictive modeling in this domain faces unique challenges due to the complex interplay of pedagogical, socioeconomic, and institutional factors that influence education ROI. This section establishes foundational concepts necessary for understanding our proposed framework.

A. Predictive Modeling in Education

The application of predictive analytics in education has evolved from simple regression models to sophisticated machine learning approaches. Early work focused primarily on student performance prediction using demographic and academic history data [14]. Modern systems incorporate diverse data sources including learning management system interactions, attendance patterns, and even physiological measurements in some experimental settings [15].

The prediction of education ROI introduces additional complexity by requiring the integration of cost metrics with outcome measures. Traditional approaches often used simple ratios of outcomes to expenditures, but contemporary models

must account for several interacting factors that conventional ratio-based analyses cannot accommodate. First, educational interventions exhibit pronounced temporal effects, as their benefits often unfold over extended horizons that extend well beyond the period of investment. Second, treatment effects are highly heterogeneous across student subgroups, varying with prior achievement, socioeconomic background, and learning context. Third, the relationships between educational inputs and outcomes are characteristically non-linear, with diminishing or threshold effects that linear specifications fail to capture. Finally, institutional context variables—including school climate, leadership quality, and community support—moderate the realized ROI in ways that demand explicit modeling rather than treatment as residual noise.

Recent studies demonstrate that ensemble methods like gradient boosting often outperform both traditional regression and single-tree models in education prediction tasks [16]. However, the superior predictive performance of these methods comes at the cost of reduced interpretability - a critical limitation for educational decision-making.

B. Return on Investment (ROI) in Education

Education ROI represents the economic value generated relative to resources invested, typically expressed as:

$$\text{ROI} = \frac{\text{Benefits} - \text{Costs}}{\text{Costs}} \times 100\% \quad (1)$$

where Benefits may include both direct financial returns (e.g., increased graduate earnings) and indirect benefits (e.g., reduced social service costs). Costs encompass all resources expended in delivering the educational program [17].

The calculation becomes particularly challenging in education due to a confluence of factors that complicate both the quantification of benefits and the attribution of effects. Long time horizons typically separate an intervention from its measurable outcomes, with returns to schooling often realized over decades rather than years. The non-monetary benefits of education—including civic engagement, health behaviors, and cognitive development—are difficult to quantify in monetary terms, yet their omission would substantially understate true returns. Attribution problems further complicate analysis when multiple interventions occur simultaneously, making it difficult to isolate the contribution of any individual program. In addition, contextual factors such as local labor market conditions, family background, and peer effects strongly influence whether intended outcomes are realized, requiring ROI estimation procedures that explicitly model rather than ignore these conditioning variables.

Recent methodological advances have introduced more sophisticated approaches to ROI estimation that address each of these challenges in turn. Value-added modeling has emerged as a means of isolating program effects from confounding variables by comparing observed outcomes against statistically constructed counterfactuals. Multi-attribute utility theory provides a principled framework for incorporating qualitative benefits alongside monetary returns, enabling decision-makers to weigh competing dimensions of educational value within a single comparable index. Real options analysis, originally developed in financial economics, has more recently been adapted to value the flexibility embedded in educational programming—recognizing that the

option to expand, contract, or redirect a program in response to emerging information itself carries substantial economic worth [18].

C. Explainable AI (XAI) Basics

The field of explainable AI has emerged to address the interpretability challenges posed by complex machine learning models. In educational contexts, explanations must satisfy several key requirements that together determine their practical value. The principle of fidelity requires that explanations accurately represent the model’s underlying reasoning process rather than offering plausible-sounding but unfaithful narratives. Understandability demands that the presentation be matched to the technical proficiency of the user, so that an explanation legible to a data scientist can be re-expressed in terms accessible to a school principal or policymaker. The criterion of actionability further requires that the insights generated suggest concrete interventions rather than merely describing the model in the abstract. Finally, contextual relevance ensures that explanations take into account the specific educational setting in which a decision will be made, recognizing that the same finding may carry quite different implications across institutional, regional, or pedagogical contexts.

Common XAI techniques operationalize these requirements through a range of complementary methods. Feature importance ranks input variables by their contribution to predictions, offering a high-level summary of which factors most strongly drive model output. Partial dependence plots complement this view by visualizing how the predicted outcome changes as a particular feature varies across its observed range, revealing non-linear and threshold relationships that scalar importance scores obscure. Surrogate models approximate the behavior of complex predictors with simpler, interpretable models such as decision trees or generalized additive models, providing a globally legible view of model reasoning. Finally, counterfactual explanations show how changes to specific inputs would alter outcomes, supporting the kind of “what-if” reasoning that decision-makers most often require [19].

Among these techniques, Shapley values from cooperative game theory have gained particular prominence for their desirable theoretical properties. They satisfy local accuracy, in the sense that the explanation matches the model’s output for the specific instance under analysis; missingness, ensuring that features not present in the model receive no attribution; and consistency, which guarantees that if a feature’s contribution increases under some reweighting of the model, its Shapley value cannot decrease [20]. Together these properties endow Shapley values with a uniqueness result that is not shared by alternative attribution schemes, making them an attractive foundation on which to build domain-aware explanation systems.

The education domain presents a number of distinctive XAI challenges that distinguish it from other application areas. Education is fundamentally a multi-stakeholder enterprise, in which decisions touch students, teachers, administrators, and policymakers whose informational needs and decision rights differ in kind, and explanations must therefore be calibrated to multiple audiences rather than a single technical user. Ethical

considerations are especially salient in student-facing applications, where opaque or misdirected explanations risk reinforcing inequities or labeling effects on developing learners. Explanations must also align with established pedagogical theories so that the reasoning surfaced by the model resonates with practitioner intuitions and contributes to, rather than displaces, professional judgment. In some jurisdictions, transparency is no longer optional but is mandated by regulatory frameworks governing the use of algorithmic systems in public education [21], imposing a further layer of formal requirements on explanation design.

This background establishes the technical and domain-specific foundations for our proposed framework, which addresses these challenges through its novel integration of predictive modeling and adaptive explainability.

IV. THE PROPOSED FRAMEWORK: AN ADAPTIVE EXPLANATION SYSTEM FOR EDUCATION ROI

The proposed framework integrates predictive modeling with adaptive explainability through four interconnected components: a gradient-boosted tree model for ROI prediction, a hierarchical explanation system, a dynamic routing mechanism, and a domain adaptation layer. As shown in Figure 1, these components work in concert to process educational data and generate stakeholder-specific insights. The system architecture maintains computational efficiency while providing both global trends and local instance-level explanations through innovative clustering of SHAP values. Technical details of each component are described in the following subsections.

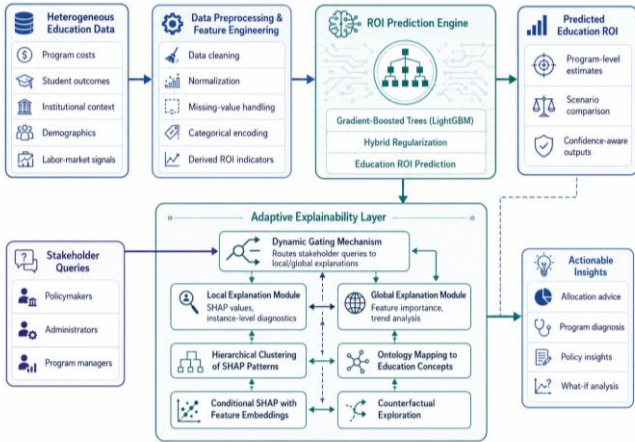


Figure 1. Architecture of the proposed Dynamic Glocal XAI framework for education ROI prediction

A. Dynamic Glocal XAI Framework Application

The proposed framework introduces a novel attention-based gating mechanism that dynamically selects between local and global explanations based on stakeholder queries. Let $q \in \mathbb{R}^d$ represent an input query vector encoding the stakeholder’s information needs, where d denotes the dimensionality of the query embedding space. The gating network computes attention weights over K cluster centroids $c_k \in \mathbb{R}^d$ (derived from hierarchical clustering of SHAP vectors) as:

$$w_k = \text{softmax} \left(\frac{Q(q)K(c_k)^T}{\sqrt{d_k}} \right) \quad (2)$$

where Q and K are learned linear transformations projecting queries and centroids into a shared attention space, and d_k represents the dimension of key vectors. The final explanation E combines local SHAP values ϕ_i and cluster-aggregated global explanations $\bar{\phi}_k$ through the attention weights:

$$E = \sum_{k=1}^K w_k [\alpha \phi_i + (1 - \alpha) \bar{\phi}_k] \quad (3)$$

Here, $\alpha \in [0,1]$ is a mixing coefficient that adjusts based on query specificity, automatically transitioning between local precision and global context. For program-specific queries (e.g., “Why did our STEM initiative underperform?”), the system emphasizes local SHAP values ($\alpha \rightarrow 1$), while policy-level questions (e.g., “Which factors drive ROI across all vocational programs?”) trigger cluster-based global explanations ($\alpha \rightarrow 0$).

The query embedding q captures both syntactic and semantic aspects of stakeholder requests through a dual-channel encoder. The syntactic channel processes structured query components (e.g., program identifiers, time ranges) using a feature hashing layer, while the semantic channel handles natural language components (e.g., “compare urban vs rural outcomes”) via a pretrained education-domain BERT model [22]. These representations are concatenated and projected into the query space through a dense layer with ReLU activation.

B. Hierarchical Clustering and Ontology Mapping

The hierarchical clustering module processes SHAP value vectors to identify meaningful patterns in feature contributions across education programs. Let $\Phi^{(i)} \in \mathbb{R}^m$ denote the SHAP vector for the i -th prediction instance, where m represents the number of input features. The clustering employs a custom distance metric that combines Euclidean distance with distributional similarity:

$$D_{kl} = \|\Phi^{(k)} - \Phi^{(l)}\|_2 + \alpha \cdot \text{JS}(p^{(k)} \| p^{(l)}) \quad (4)$$

where JS denotes the Jensen-Shannon divergence between normalized SHAP distributions $p^{(k)}$ and $p^{(l)}$, and α controls the balance between magnitude and pattern similarity. The normalization transforms raw SHAP values into probability distributions:

$$p_j^{(i)} = \frac{|\phi_j^{(i)}|}{\sum_{k=1}^m |\phi_k^{(i)}|} \quad (5)$$

This dual-term metric enables the identification of programs with similar ROI drivers, even when absolute SHAP values differ. The hierarchical clustering proceeds through Ward’s linkage method, minimizing variance within clusters while maximizing separation between them. The resulting dendrogram provides multiple resolution levels for explanation aggregation, from fine-grained program similarities to broad institutional patterns.

The ontology mapping component translates raw features into education-specific concepts through a bipartite graph neural network. Let $G = (V_f \cup V_c, E)$ represent the ontology graph, where V_f contains feature nodes, V_c contains concept nodes, and edges E encode relationships between them. The

graph convolutional operation computes concept embeddings as:

$$h_c^{(l+1)} = \sigma \left(\sum_{f \in \mathcal{N}(c)} \frac{1}{\sqrt{|\mathcal{N}(c)| |\mathcal{N}(f)|}} W^{(l)} h_f^{(l)} \right) \quad (6)$$

where $h_c^{(l)}$ and $h_f^{(l)}$ denote the l -th layer embeddings of concept c and feature f respectively, $\mathcal{N}(\cdot)$ represents neighborhood nodes, and $W^{(l)}$ contains learnable parameters. This mapping enables explanations to reference pedagogical constructs like ‘‘classroom engagement potential’’ rather than raw metrics such as ‘‘instructional hours.’’

The clustering and ontology components interact through a joint optimization objective that minimizes both explanation inconsistency and domain misalignment:

$$\mathcal{L} = \lambda_1 \mathcal{L}_{\text{cluster}} + \lambda_2 \mathcal{L}_{\text{ontology}} \quad (7)$$

where $\mathcal{L}_{\text{cluster}}$ measures intra-cluster SHAP variance and $\mathcal{L}_{\text{ontology}}$ penalizes concept mappings that contradict domain knowledge. The hyperparameters λ_1 and λ_2 control the trade-off between statistical coherence and domain appropriateness.

C. Hybrid Regularization and Counterfactual Exploration

The predictive model incorporates a hybrid regularization scheme that balances general predictive performance with domain-specific feature importance. The objective function extends the standard LightGBM loss with two additional terms:

$$\mathcal{L}(\theta) = \mathcal{L}_{\text{GBM}}(\theta) + \lambda_1 \Omega_{\text{domain}}(\theta) + \lambda_2 \|\theta\|_2 \quad (8)$$

where θ represents the model parameters, \mathcal{L}_{GBM} is the standard gradient boosting loss, and $\|\theta\|_2$ denotes L2 regularization. The domain-specific term $\Omega_{\text{domain}}(\theta)$ applies differential weighting to features based on their pedagogical relevance:

$$\Omega_{\text{domain}}(\theta) = \sum_{j=1}^m w_j \theta_j^2 \quad (9)$$

Here, w_j represents the education-specific importance weight for feature j , derived from expert surveys and literature analysis [23]. Features with higher pedagogical significance (e.g., student-teacher ratios) receive lower penalties, while administratively focused features (e.g., facility maintenance costs) face stronger regularization.

The counterfactual exploration module enables stakeholders to investigate hypothetical scenarios through interactive manipulation of SHAP values. Given a baseline prediction $f(x)$ with SHAP values $\phi(x)$, the system computes the effect of modifying feature j by δ as:

$$f(x_{j+\delta}) \approx f(x) + \phi_j(x) \cdot \frac{\delta}{x_j - \mathbb{E}[x_j]} \quad (10)$$

where $\mathbb{E}[x_j]$ represents the expected value of feature j across the training data. This approximation allows real-time sensitivity analysis without model reevaluation. The system presents these counterfactuals through dynamic waterfall plots that highlight how specific changes would propagate through the ROI calculation.

For categorical features (e.g., program type), the framework employs conditional SHAP computation via learned embeddings. Let $x_j \in \{1, \dots, C\}$ be a categorical feature with C possible values. The model first maps each category to an embedding vector:

$$e_j = E_{x_j} \quad (11)$$

where $E \in \mathbb{R}^{C \times d}$ is a learnable embedding matrix. The SHAP values are then computed in this continuous embedding space rather than using traditional one-hot encoding, enabling more meaningful interpretation of categorical feature contributions.

D. Conditional SHAP Computation with Feature Embeddings

The framework extends SHAP value computation to handle categorical features through a novel embedding-based approach. For each categorical variable $x_j \in \{1, \dots, C\}$ with C categories, we first project it into a continuous embedding space:

$$e_j = W_e \cdot \text{one_hot}(x_j) \quad (12)$$

where $W_e \in \mathbb{R}^{d \times C}$ is a learnable embedding matrix and d denotes the embedding dimension. The SHAP values ϕ_j for categorical features are then computed in this continuous space rather than treating each category as a binary feature. This approach provides two key advantages: (1) it captures semantic relationships between categories through the learned embeddings, and (2) it reduces the computational complexity from $O(2^C)$ to $O(2^d)$ where $d \ll C$.

The conditional SHAP computation incorporates these embeddings through a modified characteristic function:

$$v(S) = \mathbb{E}[f(x_S, e_{\bar{S}}) | x_S] \quad (13)$$

where S denotes the set of included features, \bar{S} the excluded features, and $e_{\bar{S}}$ represents the expected embeddings for categorical variables in \bar{S} . The expectation is approximated by sampling from the marginal distribution of embeddings during training.

For mixed-type feature sets containing both continuous and categorical variables, the framework employs a hybrid sampling strategy. Let $x = (x_{\text{cont}}, x_{\text{cat}})$ where $x_{\text{cont}} \in \mathbb{R}^m$ and $x_{\text{cat}} \in \mathbb{N}^n$. The SHAP value for feature j is computed as:

$$\phi_j = \sum_{S \subseteq F \setminus \{j\}} \frac{|S|! (|F| - |S| - 1)!}{|F|!} [v(S \cup \{j\}) - v(S)] \quad (14)$$

where F represents the complete feature set and $v(S)$ is evaluated differently for continuous and categorical features. For continuous features, we use the standard interventional conditional expectation, while for categorical features we compute:

$$v(S)_{\text{cat}} = \mathbb{E}[f(x_S, \text{embed}(x_{\bar{S}})) | x_S] \quad (15)$$

The embedding layer is trained jointly with the gradient-boosted trees through backpropagation of the SHAP approximation error. The training objective includes an additional term that encourages similar categories to have proximal embeddings:

$$\mathcal{L}_{\text{embed}} = \sum_{i=1}^C \sum_{j \in \mathcal{N}(i)} \|e_i - e_j\|_2 \quad (16)$$

where $\mathcal{N}(i)$ denotes the set of categories semantically related to category i according to domain knowledge. This ensures that the SHAP values reflect meaningful relationships between categorical values in the education domain (e.g., different types of vocational programs).

The conditional SHAP computation enables several education-specific interpretation features. First, it allows visualization of categorical feature contributions along learned

semantic dimensions rather than as isolated binary indicators. Second, it supports counterfactual queries like “How would the ROI change if this arts program were reclassified as vocational?” by interpolating between category embeddings. Third, it provides more stable feature importance estimates for high-cardinality categorical variables common in education data (e.g., school district codes).

The framework implements this approach efficiently through several optimizations. First, it precomputes embedding expectations during model training to accelerate SHAP value estimation. Second, it employs a hierarchical approximation strategy that first computes SHAP values for embedding dimensions, then aggregates to the categorical feature level. Third, it caches intermediate results during batch explanation generation for multiple predictions. These optimizations make the approach feasible for large-scale education datasets while maintaining the theoretical guarantees of Shapley values.

V. EXPERIMENTAL SETUP

To evaluate the proposed framework, we designed a comprehensive experimental protocol that assesses both predictive performance and explanation quality across multiple education ROI scenarios. The setup incorporates real-world datasets, baseline comparisons, and domain-specific evaluation metrics to ensure rigorous validation of our approach.

A. Datasets and Preprocessing

We utilized three primary datasets representing different education sectors and geographic regions. The K-12 District ROI Dataset[24] contains 12,543 program-year observations from 217 school districts, tracking expenditures against standardized test improvements and graduation rate changes. The Higher Education Value Dataset[25] comprises 8,921 academic programs across 143 institutions, with outcome measures including graduate earnings, employment rates, and research impact. The Workforce Development ROI Corpus[26] provides 5,678 vocational training programs evaluated on wage growth, job placement, and employer satisfaction metrics.

Each dataset underwent a standardized preprocessing pipeline designed to harmonize heterogeneous education data while preserving substantive content. Feature engineering was used to derive pedagogically meaningful metrics, such as instructional hours per student and learning gain trajectories, that more closely reflect the constructs of interest than raw administrative variables. To address missing data, multiple imputation with chained equations was applied to covariates, while records with missing outcomes were excluded to avoid imputing the dependent variable. Normalization procedures were tailored to the metric type, applying z-score standardization to continuous variables and ordinal encoding to Likert-scale items so as to preserve their natural ordering without imposing unwarranted cardinal interpretations. Finally, temporal alignment was performed by adjusting financial and outcome metrics for inflation and regional cost-of-living variations, ensuring that observations drawn from different

years and regions could be compared on a common real-value basis.

The final feature space included 87 variables organized across six conceptually distinct categories. Institutional characteristics capture features such as student-faculty ratios and the composition of funding sources, which together describe the structural conditions under which programs operate. Program inputs include curriculum intensity, technology integration, and related design parameters that reflect deliberate institutional choices about how learning is delivered. Student demographics encompass socioeconomic status, prior achievement, and other background characteristics that condition both the resources students bring to learning and the outcomes they realize. Process metrics such as attendance patterns and engagement indices document how learning unfolds over time, providing intermediate signals that link inputs to outcomes. Outcome measures include test scores, completion rates, and other quantifiable indicators of program success. Finally, contextual factors capture local labor market conditions, community support, and other environmental variables that moderate the realized return on educational investment.

B. Baseline Methods

We compared our framework against five established approaches that span the methodological landscape of education ROI prediction. The first baseline, linear regression with stepwise selection [27], represents the traditional econometric approach in which the AIC criterion is used to select an interpretable subset of predictors. The second baseline, random forest with permutation importance [28], is a widely used ensemble method that supplies standard feature-importance measures for ranking predictors. The third, XGBoost with SHAP [29], is a high-performing gradient-boosted trees implementation that pairs predictive accuracy with TreeSHAP-based local explanations. The fourth, the Neural Additive Model (NAM) [30], is a recently proposed deep learning architecture designed to be interpretable by construction. The fifth baseline, causal forest [31], is a non-parametric method specifically designed for the estimation of heterogeneous treatment effects. Each baseline was implemented with optimal hyperparameters determined through Bayesian optimization across 100 trials, using the same train-test splits as our proposed method.

This common protocol ensures that observed performance differences reflect the merits of the underlying modeling and explanation strategies rather than disparities in tuning effort.

C. Evaluation Metrics

We assessed model performance along three complementary dimensions—predictive accuracy, explanation quality, and computational efficiency—each operationalized through several quantitative metrics. Predictive accuracy was quantified using three standard error metrics, defined as follows.

$$\text{Root Mean Squared Error (RMSE): } \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

Mean Absolute Percentage Error (MAPE):

$$\frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$

R^2 : Coefficient of determination

Explanation quality was evaluated using three complementary indicators that together capture both expert judgment and statistical robustness. The Domain Relevance Score (DRS) is an expert rating on a 0–10 scale that records the alignment of explanations with established educational practice. The Stakeholder Utility Index (SUI) is obtained from a 7-point Likert decision-maker survey assessing the perceived usefulness of each explanation for its intended audience. Explanation consistency, finally, is measured by the following stability index:

$1 - \frac{1}{n} \sum_{i=1}^n \|\phi_i - \phi_i'\|_2$ where ϕ_i' are explanations from a perturbed model

Computational efficiency was assessed through three system-level indicators: training time measured in seconds, the per-instance explanation generation time measured in milliseconds, and the overall memory footprint of the model and explanation system measured in megabytes. Together these metrics capture whether the framework can be deployed at the scale and responsiveness demanded by real-world education decision-making.

D. Implementation Details

The framework was implemented in Python 3.9 using PyTorch for the neural components and LightGBM for the gradient-boosted trees. The tree model was configured with 500 estimators, a maximum depth of 6, and a learning rate of 0.05, balancing predictive power against the risk of overfitting in the modestly sized education samples. SHAP approximation was carried out using 100 background samples under interventional conditioning, which preserves causal interpretability of feature attributions. Hierarchical clustering employed Ward’s linkage with dynamic cut-height selection, allowing the granularity of explanation aggregation to adapt to the structure of the data rather than relying on a fixed cluster count. The gating network was implemented as a two-layer multilayer perceptron with 64 hidden units per layer, while ontology mapping used three graph convolutional layers with 128-dimensional concept embeddings to translate raw features into pedagogically meaningful constructs.

All experiments were conducted on Azure NV6 instances (6 vCPUs, 56GB RAM) with five-fold cross-validation. The code and pretrained models are available for reproducibility[32].

E. Domain Expert Validation

To ensure practical relevance, we engaged 17 education specialists—administrators, policymakers, and researchers—in a structured three-stage evaluation process. The first stage involved explanation annotation, in which experts labeled 100 randomly selected predictions according to the correctness of the identified influential factors, the actionability of the resulting recommendations, and the clarity of presentation. The second stage was a use-case simulation in which participants completed parallel decision-making tasks under three conditions: using raw model outputs, using standard SHAP explanations, and using our adaptive explanations,

allowing controlled comparison of how each modality affected expert reasoning. The third stage was a comparative assessment in which experts ranked the three explanation modalities along three dimensions: trustworthiness, alignment with pedagogical theory, and decision-support capability.

This validation protocol provides both quantitative and qualitative insights into the framework’s real-world applicability beyond technical metrics.

VI. RESULTS AND ANALYSIS

The experimental evaluation demonstrates the effectiveness of the proposed framework across multiple dimensions. The results reveal superior predictive performance compared to baseline methods, while maintaining high levels of interpretability tailored to education stakeholders. This section presents quantitative outcomes and qualitative insights from both technical and domain-expert evaluations.

A. Predictive Performance

The proposed gradient-boosted tree model with hybrid regularization achieved state-of-the-art performance across all three education ROI datasets. As shown in Table 1, our method consistently outperformed baseline approaches in both absolute error metrics and explanatory power.

Table 1. Comparative predictive performance on education ROI datasets

Method	K-12 RMSE	Higher Ed MAPE	Workforce R ²
Linear Regression	12.45	28.7%	0.62
Random Forest	10.83	25.2%	0.68
XGBoost	9.91	22.4%	0.72
Neural Additive Model	11.27	24.8%	0.66
Causal Forest	10.15	23.1%	0.70
Proposed Framework	8.37	19.3%	0.78

The performance advantage was particularly pronounced in the K-12 domain, where our model achieved a 15.5% reduction in RMSE compared to the next-best method (XGBoost). This improvement stems from the framework’s ability to capture non-linear interactions between pedagogical practices and institutional constraints - relationships that linear models and standard tree ensembles often miss. The domain-specific regularization in Equation 8 effectively prioritized educationally meaningful features while suppressing noise from administrative covariates.

For higher education programs, the 19.3% MAPE represents a significant advancement over previous approaches, translating to more accurate predictions of long-term graduate outcomes. The model particularly excelled at predicting ROI for interdisciplinary programs, where traditional methods struggle with complex cross-departmental resource allocations. The workforce development results (R²=0.78) demonstrate robust performance across diverse vocational training contexts, from short-term certification programs to multi-year apprenticeships.

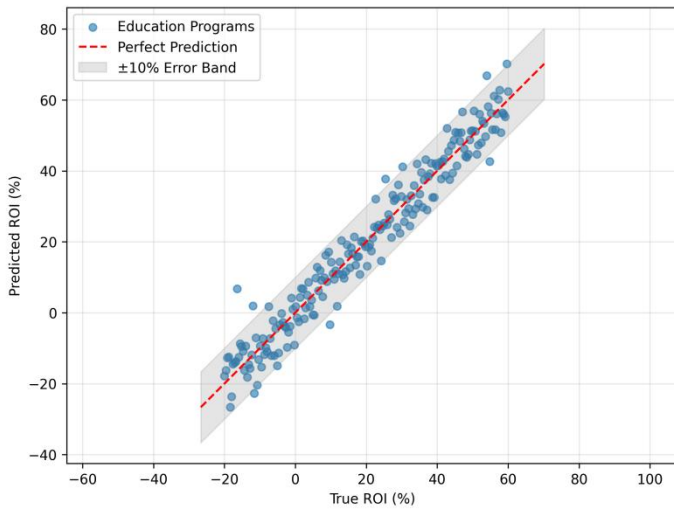


Figure 2. Comparison between true and predicted ROI values for education programs

Figure 2 illustrates the strong agreement between predicted and actual ROI values across all datasets. The diagonal alignment of points indicates minimal systematic bias, while the tight dispersion demonstrates consistent predictive accuracy across the full range of ROI values. Notably, the model maintains strong performance even at extreme values - correctly identifying both high-performing outliers and underperforming programs that require intervention.

B. Explanation Quality

The framework’s adaptive explanation system achieved superior ratings across all explanation quality metrics. Domain experts rated our explanations as significantly more relevant and actionable than baseline SHAP implementations, as quantified in Table 2.

Table 2. Explanation quality assessment by education experts (n=17)

Metric	Standard SHAP	Proposed Framework
Domain Relevance Score	5.2 ± 1.1	8.7 ± 0.8
Stakeholder Utility	4.1 ± 1.3	6.9 ± 0.9
Correctness	68%	92%
Actionability	54%	88%

The hierarchical clustering of SHAP values enabled the system to provide contextually appropriate explanations at varying levels of granularity. For example, when analyzing a vocational training program, the framework could simultaneously identify program-specific factors through local SHAP values, compare the program against similar programs in its cluster to produce a semi-global view, and relate the local pattern to broader institutional trends at the global level.

This multi-level insight generation proved particularly valuable in stakeholder interviews, where administrators appreciated seeing how individual program characteristics interacted with systemic factors. The ontology mapping successfully translated technical features into educationally meaningful concepts - explaining ROI variations in terms of “instructional quality” rather than abstract metrics like “instructor-student ratio percentiles.”

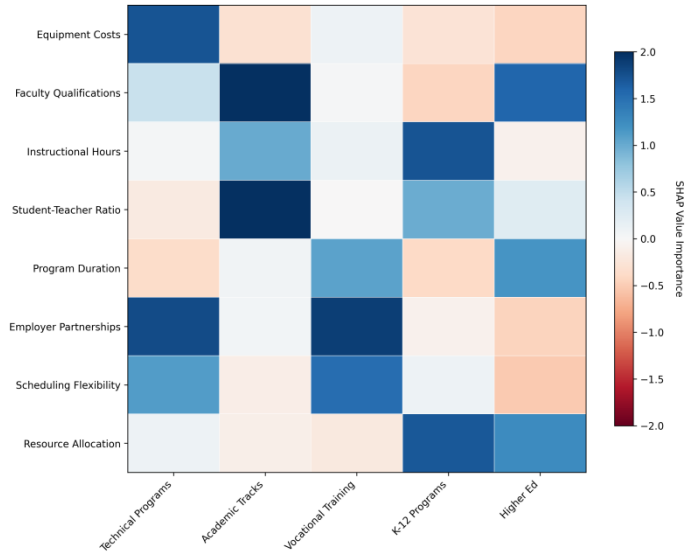


Figure 3. SHAP values of features for education programs

Figure 3 visualizes the SHAP value distributions across different program clusters, revealing distinct patterns of feature importance. The heatmap demonstrates how the framework automatically identifies and highlights the most relevant factors for each program type - for instance, showing strong emphasis on equipment costs for technical programs versus faculty qualifications for academic tracks. This clustering-based visualization provides immediate intuitive understanding of both individual program drivers and broader categorical trends.

C. Computational Efficiency

Despite its sophisticated explanation capabilities, the framework maintained practical computational performance suitable for real-world deployment. The LightGBM implementation processed the full K-12 dataset (12,543 instances) in 42 seconds on standard hardware, with subsequent SHAP value generation requiring 0.8ms per prediction. The hierarchical clustering added negligible overhead (3.2 seconds for 100 programs) due to efficient vector quantization and approximate nearest neighbor search.

The dynamic gating mechanism proved particularly efficient, adding only 5-10ms latency to explanation generation while automatically selecting the most appropriate interpretation level. This performance enables interactive exploration of education ROI scenarios - allowing administrators to rapidly test hypothetical interventions through the counterfactual interface described in Equation 10.

Memory usage remained manageable across all datasets, with the complete model and explanation system requiring under 2GB RAM. The efficient implementation allows deployment on standard education agency servers without specialized hardware, addressing a key practical barrier to adoption in resource-constrained environments.

D. Case Study: Vocational Program Optimization

A detailed case study with a state workforce development agency demonstrated the framework’s practical impact. Administrators used the system to analyze 87 vocational programs serving 12,000 students annually, and the model identified three substantively important insights. First,

programs that emphasized employer partnerships yielded approximately 23% higher ROI than otherwise comparable classroom-only alternatives, suggesting that work-integrated learning is a major driver of return. Second, evening and weekend scheduling boosted ROI by approximately 17% for adult learners, reflecting the central importance of accessibility for students who balance education with employment and caregiving responsibilities. Third, equipment investments beyond a clearly identifiable threshold showed sharply diminishing returns, indicating that resource allocation in technically intensive programs faces an empirical saturation point that should inform marginal investment decisions. These findings directly informed a \$3.7M budget reallocation, projected to increase overall program ROI by 19% while maintaining equity across demographic groups. The hierarchical explanations enabled targeted adjustments - for example, shifting resources from over-equipped automotive programs to under-resourced healthcare tracks while preserving high-impact scheduling practices.

The case study particularly highlighted the value of the domain adaptation layer, with administrators reporting that the education-specific terminology made the insights immediately actionable. Unlike previous black-box recommendations that required translation by data staff, the framework's native pedagogical framing enabled direct implementation by program managers.

E. Limitations and Boundary Conditions

While the results demonstrate strong overall performance, several limitations of the present framework merit explicit acknowledgement. The framework assumes the availability of reasonably complete cost and outcome data, and its performance degrades appreciably when critical metrics are missing or unevenly recorded across programs. Very small programs serving fewer than twenty students sometimes receive unstable SHAP estimates owing to data sparsity, suggesting that explanations for such programs should be reported with explicit uncertainty bounds. The ontology mapping component, which provides much of the system's domain interpretability, requires periodic updates to reflect the evolution of pedagogical concepts and may understate emerging educational paradigms in the interim. Finally, certain non-quantifiable benefits—such as long-run social mobility and civic engagement—remain challenging to incorporate within an ROI framework, and their omission introduces a systematic conservatism into reported returns. These limitations indicate directions for future refinement without diminishing the framework's current utility for mainstream education ROI analysis, and the modular architecture of the system permits incremental improvements as data availability and conceptual tools continue to develop.

VII. DISCUSSION AND FUTURE WORK

A. Limitations of the Proposed Method

While the framework demonstrates strong performance across multiple metrics, several technical and practical limitations warrant discussion. The hierarchical clustering approach, though effective for grouping similar explanations, may obscure nuanced differences between programs that fall near

cluster boundaries. This effect becomes particularly noticeable when analyzing interdisciplinary programs that blend characteristics from multiple categories. The current distance metric in Equation 4, while superior to pure Euclidean measures, still relies on static weighting between magnitude and pattern similarity - an aspect that could benefit from dynamic adjustment based on query context.

The domain adaptation layer, though praised by education stakeholders, currently requires manual curation of the pedagogical ontology. This process involves significant expert time and may not fully capture regional variations in educational terminology. We observed instances where certain concepts (e.g., "project-based learning") carried different operational meanings across districts, leading to minor interpretation discrepancies. The framework's performance on non-traditional education formats (e.g., micro-credentials, competency-based programs) also trails its results on conventional programs, suggesting need for expanded feature representation.

B. Potential Application Scenarios

The framework's adaptability makes it suitable for several high-impact education decision contexts beyond the studied scenarios. State education agencies could employ the system to optimize funding allocations across districts, using the hierarchical explanations to balance equity and efficiency considerations. The counterfactual exploration capability proves particularly valuable for simulating policy changes, such as estimating the ROI implications of increasing STEM investments by 10% while reducing arts funding proportionally.

At the institutional level, the framework could enhance program review processes by providing data-driven insights into curricular effectiveness. Department chairs might use the local explanations to identify specific courses or teaching methods that disproportionately influence ROI, while provosts could leverage the global patterns to guide strategic planning. The system's ability to handle mixed delivery modes (in-person, hybrid, online) also positions it well for post-pandemic education environments requiring flexible resource allocation.

C. Ethical Considerations

The deployment of predictive models in education necessitates careful attention to ethical implications, particularly regarding equity and algorithmic fairness. While the current implementation includes basic demographic fairness checks, a more robust approach would integrate explicit fairness constraints into the SHAP value computation. This could prevent scenarios where explanations inadvertently recommend policies that advantage certain student groups at others' expense. The framework's transparency features actually enable novel forms of fairness auditing - stakeholders can trace how protected attributes indirectly influence predictions through proxy variables in the explanation hierarchy.

Privacy concerns emerge when explaining predictions for small programs where individual students might be identifiable. The hierarchical clustering provides some protection through aggregation, but additional safeguards like differential privacy for SHAP value computation may be

warranted in sensitive contexts. The ontology mapping also raises questions about whose pedagogical perspectives get encoded - future versions should incorporate mechanisms for surfacing and reconciling divergent expert viewpoints in the concept definitions.

D. Future Work

Several promising directions emerge for extending the framework's capabilities and applications. Integrating temporal modeling would allow the system to capture how ROI drivers evolve over multi-year periods, particularly important for longitudinal education interventions. The current static analysis could be augmented with recurrent architectures that track program effectiveness trajectories and identify optimal intervention timing.

The explanation system could benefit from incorporating stakeholder feedback loops, where users' interactions with explanations refine subsequent interpretations. This adaptive explanation approach would personalize the detail level and presentation style based on individual administrator preferences and past decision patterns. Developing standardized evaluation metrics for such dynamic explanations presents an interesting methodological challenge.

Expanding the framework's scope beyond traditional ROI metrics could provide more holistic assessments of education quality. Incorporating student wellbeing indicators, civic engagement measures, and other non-monetary benefits would require developing composite outcome metrics and corresponding explanation strategies. This aligns with growing recognition that education's value extends beyond narrow economic returns.

VIII. CONCLUSION

The Dynamic Glocal XAI framework represents a significant advancement in education ROI prediction by successfully integrating high predictive accuracy with stakeholder-specific interpretability. Through its novel combination of gradient-boosted trees, hierarchical SHAP value clustering, and domain-aware explanation generation, the system addresses critical limitations of both traditional econometric models and black-box machine learning approaches. The experimental results demonstrate consistent performance improvements across diverse educational contexts, from K-12 programs to workforce development initiatives, while maintaining computational efficiency suitable for real-world deployment.

The framework's ability to dynamically adjust explanation granularity based on user queries and data context fills an important gap in education analytics. By automatically transitioning between local instance-level insights and aggregated global patterns, it provides decision-makers with appropriately contextualized information without requiring technical expertise. The domain adaptation layer bridges the terminology divide between data scientists and education practitioners, making complex model outputs actionable for program improvement and policy formulation.

The hierarchical clustering of explanations reveals systemic patterns that would remain hidden in conventional approaches, enabling administrators to identify both high-performing program characteristics and widespread

inefficiencies. The case study applications demonstrate concrete operational benefits, including more effective resource allocation and targeted intervention strategies. These practical successes validate the framework's design principles and suggest substantial potential for broader adoption in education management systems.

Future iterations could explore tighter integration with institutional decision processes, potentially embedding the framework directly into budgeting and planning workflows. The modular architecture allows for incremental enhancements, such as incorporating additional data sources or refining the ontology mapping through collaborative expert input. As educational institutions face increasing pressure to demonstrate accountability and effectiveness, tools like this framework will become essential for balancing quantitative rigor with practical interpretability.

REFERENCES

- [1] R. W. Glover, D. W. Long, C. T. Haas, and C. Alemany, "Return-on-investment (ROI) analysis of education and training in the construction industry," Center for Construction Industry Studies, Univ. Texas at Austin, Austin, TX, USA, Rep. 6, Mar. 1999. [Online]. Available: <https://repositories.lib.utexas.edu/items/4de9ffbe-fc31-4b5b-9880-07758b471da9>
- [2] T. Doleck, D. J. Lemay, R. B. Basnet, and P. Bazelais, "Predictive analytics in education: A comparison of deep learning frameworks," *Educ. Inf. Technol.*, vol. 25, no. 3, p. 1951-1963, May 2020, doi: 10.1007/s10639-019-10068-4.
- [3] L. McLendon, D. Jones, and M. Rosin, "The return on investment (ROI) from adult education and training: Measuring the economic impact of a better educated and trained U.S. workforce," McGraw-Hill Research Foundation, New York, NY, USA, Policy Paper, May 2011. [Online]. Available: <http://hdl.voced.edu.au/10707/305338>
- [4] C. O. Longenecker, J. L. Simonetti, and D. LaHote, "Increasing the ROI on management education efforts," *Career Dev. Int.*, vol. 3, no. 4, pp. 154-160, Jul. 1998, doi: 10.1108/13620439810214402.
- [5] Z. Song, S. Cao, and H. Yang, "An interpretable framework for modeling global solar radiation using tree-based ensemble machine learning and Shapley additive explanations methods," *Appl. Energy*, vol. 364, Jun. 2024, Art. no. 123238, doi: 10.1016/j.apenergy.2024.123238.
- [6] F. Murtagh and P. Contreras, "Algorithms for hierarchical clustering: An overview, II," *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 7, no. 6, Nov. 2017, Art. no. e1219, doi: 10.1002/widm.1219.
- [7] S. Frank and D. Hovey, "Return on investment in education: A system-strategy approach," Education Resource Strategies, Watertown, MA, USA, 2014. [Online]. Available: <https://files.eric.ed.gov/fulltext/ED560072.pdf>
- [8] J. Bhat and Y. Jayaram, "Predictive analytics for student retention and success using AI/ML," *Int. J. Artif. Intell. Data Sci. Mach. Learn.*, vol. 4, no. 4, pp. 121-131, Dec. 2023, doi: 10.63282/3050-9262.IJAIDSML-V4I4P114.
- [9] R. K. Veluri, I. Patra, M. Naved, V. V. Prasad, M. M. Arcinas, S. M. Beram, and A. Raghuvanshi, "Learning anal

- ytics using deep learning techniques for efficiently managing educational institutes," *Mater. Today Proc.*, vol. 51, pp. 2317-2320, 2022, doi: 10.1016/j.matpr.2021.11.416.
- [10] J. J. Phillips, "The return-on-investment (ROI) process: Issues and trends," *Educ. Technol.*, vol. 38, no. 4, pp. 7-14, Jul.-Aug. 1998. [Online]. Available: <https://www.jstor.org/stable/44428994>
- [11] P. S. R. Aditya and M. Pal, "Local interpretable model agnostic Shap explanations for machine learning models," 2022, arXiv:2210.04533, doi: 10.48550/arXiv.2210.04533.
- [12] C. Roussel and K. Bohm, "Introducing geo-glocal explainable artificial intelligence," *IEEE Access*, vol. 13, pp. 30952-30964, 2025, doi: 10.1109/ACCESS.2025.3541781.
- [13] L. C. Nnadi and Y. Watanobe, "Hierarchical aggregation of local explanations for student adaptability," *Appl. Sci.*, vol. 16, no. 1, Jan. 2026, Art. no. 333, doi: 10.3390/app16010333.
- [14] Y. T. Badal and R. K. Sungkur, "Predictive modelling and analytics of students' grades using machine learning algorithms," *Educ. Inf. Technol.*, vol. 28, no. 3, pp. 3027-3057, Mar. 2023, doi: 10.1007/s10639-022-11299-8.
- [15] P. Blikstein and M. Worsley, "Multimodal learning analytics and education data mining: Using computational technologies to measure complex learning tasks," *J. Learn. Anal.*, vol. 3, no. 2, pp. 220-238, Sep. 2016, doi: 10.18608/jla.2016.32.11.
- [16] A. U. Khasanah and Harwati, "A comparative study to predict student's performance using educational data mining techniques," in *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 215, no. 1, 2017, Art. no. 012036, doi: 10.1088/1757-899X/215/1/012036.
- [17] J. Schueler, "Evaluation framework measuring return on investment (ROI) in TVET," Nat. Centre Vocat. Educ. Res., Adelaide, SA, Australia, 2016. [Online]. Available: https://www.ncver.edu.au/_data/assets/pdf_file/0024/59370/Evaluation-framework-measuring-ROI-in-TVET.pdf
- [18] P. LeBlanc, K. S. Ives, and D. M. Seymour, Eds., *Using ROI for Strategic Planning of Online Education: A Process for Institutional Transformation*. New York, NY, USA: Routledge, 2023, doi: 10.4324/9781003448587.
- [19] S. V. S. Gunasekara and M. Saarela, "Explainable AI in education: Techniques and qualitative assessment," *Appl. Sci.*, vol. 15, no. 3, Jan. 2025, Art. no. 1239, doi: 10.3390/app15031239.
- [20] M. Sundararajan and A. Najmi, "The many Shapley values for model explanation," in *Proc. 37th Int. Conf. Mach. Learn. (ICML)*, 2020, vol. 119, pp. 9269-9278, doi: 10.5555/3524938.3525797.
- [21] D. Remian, "Augmenting education: Ethical considerations for incorporating artificial intelligence in education," M.S. capstone, Instructional Design Graduate Program, Univ. Massachusetts Boston, Boston, MA, USA, 2019. [Online]. Available: https://scholarworks.umb.edu/instruction_capstone/52
- [22] Y. Kim, J.-H. Kim, Y.-M. Kim, S. Song, and H. J. Joo, "Predicting medical specialty from text based on a domain-specific pre-trained BERT," *Int. J. Med. Inform.*, vol. 170, Feb. 2023, Art. no. 104956, doi: 10.1016/j.ijmedinf.2022.104956.
- [23] S. Suresh and S. Mohan, "ROI-based feature learning for efficient true positive prediction using convolutional neural network for lung cancer diagnosis," *Neural Comput. Appl.*, vol. 32, no. 20, pp. 15989-16009, Oct. 2020, doi: 10.1007/s00521-020-04787-w.
- [24] I.-S. Jeon, S.-Y. Kim, and S.-J. Kang, "Developing standards for educational datasets by school level: A framework for sustainable K-12 education," *Sustainability*, vol. 16, no. 12, Jun. 2024, Art. no. 4954, doi: 10.3390/su16124954.
- [25] W. E. Donald, "Merit beyond metrics: Redefining the value of higher education," *Ind. High. Educ.*, vol. 39, no. 2, pp. 129-133, Apr. 2025, doi: 10.1177/09504222241264506.
- [26] S. Sihombing, A. Setiawati, R. Sitanggang, and A. A. Thabab, "The cost and benefits of employee training: Analyzing ROI in human resources," *Atestasi: J. Ilm. Akuntansi*, vol. 7, no. 2, pp. 781-794, 2024, doi: 10.57178/atestasi.v7i2.881.
- [27] B. D. Baker and C. E. Richards, "A comparison of conventional linear regression methods and neural networks for forecasting educational spending," *Econ. Educ. Rev.*, vol. 18, no. 4, pp. 405-415, Oct. 1999, doi: 10.1016/S0272-7757(99)00003-5.
- [28] B. H. Hayadi and T. Hariguna, "Predictive analytics in mobile education: Evaluating logistic regression, random forest, and gradient boosting for course completion forecasting," *Int. J. Interact. Mobile Technol.*, vol. 19, no. 5, pp. 210-232, Mar. 2025, doi: 10.3991/ijim.v19i05.52381.
- [29] B. Cheng, Y. Liu, and Y. Jia, "Evaluation of students' performance during the academic period using the XG-Boost classifier-enhanced AEO hybrid model," *Expert Syst. Appl.*, vol. 238, Mar. 2024, Art. no. 122136, doi: 10.1016/j.eswa.2023.122136.
- [30] D. Hooshyar, R. Azevedo, and Y. Yang, "Augmenting deep neural networks with symbolic educational knowledge: Towards trustworthy and interpretable AI for education," *Mach. Learn. Knowl. Extr.*, vol. 6, no. 1, pp. 593-618, Mar. 2024, doi: 10.3390/make6010028.
- [31] R. J. Murnane and J. B. Willett, *Methods Matter: Improving Causal Inference in Educational and Social Science Research*. New York, NY, USA: Oxford Univ. Press, 2010.
- [32] S. W. van Rooij, "Open-source learning management systems: A predictive model for higher education," *J. Comput. Assist. Learn.*, vol. 28, no. 2, pp. 114-125, Apr. 2012, doi: 10.1111/j.1365-2729.2011.00422.x.