Preprints.org

Article

Not peer-reviewed version

A Multi-Scale AI Framework for Informal STEM Learning: Paramorphic Digital Twins for Underserved Communities

Khazar Gorji and Reza SaeedKandezy *

Posted Date: 22 July 2025

doi: 10.20944/preprints202507.1784.v1

Keywords: multi-scale paramorphic kernel; digital twin; adaptive STEM learning; informal education; clean energy workforce development

Article

A Multi-Scale AI Framework for Informal STEM Learning: Paramorphic Digital Twins for Underserved Communities

Khazar Gorji and Reza SaeedKandezy *

University of Oklahoma, USA

* Correspondence: reza.kandezy@ou.edu

Abstract

This research presents a novel Multi-scale PAramorphic Kernel (MPAK) learning framework that is designed to enable autonomous, context-adaptive STEM learning for electrical and renewable energy topics. By employing distributed multi-agent kernel cores, recursive kernel reweighting, and entropy-guided abstraction modulation, this system dynamically adapts instructional content and sequencing based on the specific cognitive state of individual learners. Concept learning history is maintained in persistent memory buffers to facilitate individualized reinforcement and remediation in asynchronous, informal environments. Grounded in physics-informed knowledge graph, the system offers epistemic coherence and domain consistency at multiple levels of abstraction. Evaluated across multiple, underrepresented student groups in community and laboratory implementations, MPAK demonstrated a 45% rate increase in concept recall, a 3.2× improvement in student-led project completion, and sustained 68% learner interest over 12 weeks. These results emphasize the effectiveness of this architecture in delivering scalable, culturally sensitive, and high-fidelity STEM education without human interaction. The paper also discusses system deployment, statistical validation, longitudinal deployment settings, informal education problem-solving, cultural adaptation, and learning assessment. MPAK provides an extensible blueprint for inclusive, technology-driven workforce development in clean energy sectors, fueling inclusive participation and expertise in advanced technical fields.

Keywords: multi-scale paramorphic kernel; digital twin; adaptive STEM learning; informal education; clean energy workforce development

1. Introduction

A. Societal Context and Motivation

Quality STEM access remains disproportionately limited in economically distressed, geographically isolated, and institutionally disadvantaged communities, resulting in persistent disparities in workforce preparedness, tech literacy, and socioeconomic progress. Structural barriers, including low-quality instructional capital, limited specialist teacher availability, and inflexible curricular models, hinder the scalable dissemination of knowledge and skill acquisition within such communities. These constraints not only contribute to academic disparity further but also limit national potential for innovation by pushing students from diverse demographic and geographic locations away from productive involvement in future science and engineering disciplines (Ghorbani & Fattahi, 2025).

Informal learning settings, such as family-driven activities, family-driven programs, and self-driven learning settings, are severely handicapped by a lack of a scalable, expert-grade instruction infrastructure that can deliver domain-specific high-fidelity content (Lin et al., 2024; Van Bossuyt et al., 2025; Lin et al., 2025; Boltsi et al., 2025). Such learning contexts are typically lacking in exposure to credentialed teachers, pedagogical infrastructure, or adaptive models for mentoring and therefore

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 July 2025 doi:10.20944/preprints202507.1784.v1

possess fractured learning paths and superficial conceptualizations. The lack of cognitively intelligent teaching support in such settings undermines efforts to democratize STEM education for marginalized populations, which require adaptive, situated, and culturally relevant pathways to technical expertise and social mobility.

Disadvantaged learners face a set of cognitive, cultural, and systemic barriers that collectively hinder access to advanced technical training. Cognitively, regular instruction is generally not capable of accommodating a variety of learning styles, knowledge deficits, and non-linear learning trajectories. Culturally, regular curricula ordinarily disregard local contexts, languages, and community values, thereby reducing relevance and learner motivation. Systemically, structural disparities—such as inadequately resourced schools, sparse digital infrastructure, and marginalization from policy-forming platforms—amplify educational disenfranchisement. Such cumulative constraints lead to persistent discontinuities in the STEM competency learning trajectory, disproportionately affecting the representation of marginalized populations in knowledge-intensive sectors.

B. Objective and Scope of Research

This paper proposes a paramorphic digital twin AI system designed to emulate expert-level pedagogical behavior at various cognitive and representational levels. A multi-layered, domain-aware structure emulates pedagogical rationality within the system, the Multi-Scale Paramorphic Kernel Learning Architecture, which dynamically adapts in response to learner interest, abstraction level, and concept formation. Along with the use of multi-resolution feedback, real-time kernel adjustment, and context-aware instruction logic, the architecture achieves scalable, self-tuning teaching capabilities, anticipating the preemption of human experience in low-resource environments and enabling the precision-aligned transfer of knowledge across decentralized, heterogeneous, and informal learning environments.

The proposed system is optimally suited to operate in low-resource, structurally decentralized, and culturally heterogeneous learning environments for which conventional instruction models are intractable. Its operation using low-compute resources, offline-first strategy, and modularity allows for deployment in bandwidth-limited or infrastructure-poor contexts. Further, its adaptive content delivery mechanisms are informed by culturally relevant paradigms to support semantic localization, language personalization, and epistemological comprehension of community epistemologies (Razek, 2024). This several-step process ensures that adaptive content delivery is not only language-optimized but also considerably fine-tuned towards the cultural background and knowledge frameworks of different learning communities, thereby making it maximally relevant, captivating, and reasonably accessible to STEM education (Nassar et al, 2025). This flexibility enables the system to bridge pedagogical gaps in marginalized contexts without undermining instructional strength, conceptual coherence, and scalability to diverse groups of students (Nuangchalerm & Prachagool, 2023).

The MPAK architecture in this study is specifically designed to enable autonomous, context-sensitive instruction subjects with high-fidelity, standards-matched content without the presence of humans by providing IEEE/IEC/DOE standard instruction with:

- real-time ≥1 Hz feedback loops,
- <500 ms decision latency,
- ≥90% error remediation accuracy,
- ≥3-channel multi-modal learner input integration,
- <1 second abstraction switching latency,
- ≥95% content localization accuracy,
- ≥98% conceptual accuracy,
- ≥0.8 content modularity reuse factor,
- ≥99% offline uptime,

- ≤1 GB RAM and 500 MHz processor hardware requirements,
- and <5-minute asynchronous data synchronization latency.

It accommodates the infusion of local content, application-specific modeling, and interactive simulation that can be dynamically updated in real-time to model local energy infrastructures, cultural practices, and labor needs using a modular content management system augmented with dynamic semantic mapping and real-time data intake streams (Moodley & Seebregts, 2023). These streams together enable real-time context updating. This enables the effective transfer of the latest technical know-how needed for a smooth transition, particularly in underrepresented and infrastructure-constrained communities.

C. Thesis and Core Contribution

The system applies a paramorphic approach in multi-scale kernel representation and control theory to reliably simulate expert pedagogical action. Informal laboratory-based and community-integrated validations establish the system's capacity to maintain conceptual consistency and instructional consistency, as well as to model expert pedagogical action profiles and infrastructure constraints reliably. Its resilience enables the focused, stand-alone delivery of technical information, making the system a breakthrough tool for building fair STEM capacity among historically underrepresented populations.

Digital twinning architecture is created to facilitate recursive abstraction reduction, decomposing STEM abstractions at progressively lower levels structurally without compromising structural integrity or domain fidelity. This facilitates the development of adaptive learning pathways that respond to individual learners' current knowledge, rate, and context, thereby optimizing concept understanding and long-term retention across varied learning histories.

2. Theoretical Framework

A. Paramorphic Intelligence Model

The proposed framework emulates layered expert cognition by applying self-similar transformation kernels that encode hierarchical reasoning structures in conceptual, analytical, and procedural spaces. The kernels themselves are recursive operators that abstract the invariant patterns of logic underlying expert instruction, enabling the scalable decomposability and recomposability of advanced knowledge constructions (Rubin et al., 2025; Yu et al., 2025). This self-similarity enables pedagogical decision-making at varying resolutions of cognition, thereby facilitating the system's ability to dynamically adapt instructional material to the learner's level of abstraction. This provides precision-guided knowledge transfer and adaptive conceptual support.

The framework models learner knowledge and instructional adaptation through a hierarchy of scale-specific kernel functions, paramorphic transformations, and feedback-driven weight adjustments, enabling multi-resolution, personalized STEM education. Let $K_s^{(t)}: \mathcal{X}_s \times \mathcal{X}_s \to \mathbb{R}$ Denotes a positive-definite kernel function, recognizes the learner's knowledge state, and abstracts the sc $s \in$ $(\text{Rubin et al.}, 2025)$. The input space $\mathcal{X}_s$ corresponds to domain-specific representations at scale s, including conceptual, mathematical, and procedural levels. Paramorphic transformation operators govern the transition between scales. Φ_s, which recursively generates kernels at finer or coarser granularity (Gönen, M., & Alpaydın, E., 2011):

$$K_{s+1}^{(t)} = \Phi_s\left(K_s^{(t)}\right) = W_s^{(t)}.\mathcal{T}_s\left(K_s^{(t)}\right) \tag{1}$$

where $\mathcal{T}_s$ is a self-similar transformation ensuring kernel properties are preserved, and $W_s^{(t)} \in \mathbb{R}$ is an adaptive weight tracking learner-specific reweighting based on feedback at scale (Gönen, M., & Alpaydın, E., 2011). Adaptive kernel weights are dynamically updated through feedback signals $F^{(T)}$ which summarizes learner performance metrics, engagement, and cognitive signals:

$$W_s^{(t)} = W_s^{(t)} + \eta_s.f_s\left(F^s, K_s^{(t)}\right) \tag{2}$$

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 July 2025 doi:10.20944/preprints202507.1784.v1

with a learning rate $\eta_s > 0$ and the feedback function f_s mapping observed learner signals to kernel weight updates.

Instructional abstraction is modulated using entropy measures. $H_s^{(t)}$ computed over the normalized kernel-induced distribution $p_s^{(t)}(x)$(Gönen, M., & Alpaydın, E. 2011; Girolami, M., & Rogers, S. 2005):

$$H_s^{(t)} = -\sum_{x \in \mathcal{X}_s} p_s^{(t)}(x) \log p_s^{(t)}(x) \tag{3}$$

Entropy thresholds θ_s guide scale transitions, enabling dynamic abstraction level switching to optimize cognitive alignment:

$$\text{If } H_s^{(t)} > \theta_s, \text{transition to higher abstraction } (s-1);$$

$$\text{else if } H_s^{(t)} < \theta_s, \text{transition to lower abstraction } (s+1).$$

The comprehensive learner knowledge state $\mathcal{K}^{(t)}$ conjoins multi-scale kernel representations scaled by relevance factors by scale $\alpha_s^{(t)}$)(Gönen, M., & Alpaydın, E. 2011; Girolami, M., & Rogers, S. 2005):

$$\mathcal{K}^{(t)} = \sum_{s=1}^{S} \alpha_s^{(t)} K_s^{(t)} \tag{4}$$

where $\alpha_s^{(t)} \in [0,1]$ are updated based on the cumulative learner interaction history stored in persistent memory buffers.

Mathematically inspired architecture enables precise, adaptive, and personalized learning through recursively optimizing representations of knowledge, dynamically scaling abstraction, and continuously incorporating learner feedback across multiple cognitive scales to enact resilient and scalable STEM education. This layered organization enables the system to deliver instruction with context-sensitive precision, navigating dynamically through explanatory modalities to match the learner's cognitive abilities while ensuring continuity, coherence, and transferability of technical knowledge across different learning trajectories (Chada et al., 2024; Flaxman et al., 2016).

By utilizing embedded diagnostic algorithms, the system continually infers the learner's understanding, cognitive load, and optimal level of abstraction, and actively adjusts the instructional logic accordingly. This strategy ensures that every pedagogical transition—between concepts, representations, or tasks—is optimally timed and semantically calibrated to the learner's developing mental model, resulting in optimum instructional efficiency, reduced cognitive resistance, and enhanced long-term conceptual understanding. The diagnostic algorithm embedded integrates always-on prediction of the learner's level of comprehension, cognitive load, and best-fit instructional abstraction through analysis of real-time interaction patterns, allowing for dynamic pedagogical adaptation.

Let $\boldsymbol{F}^{(t)} = \{f_1^{(t)}, f_2^{(t)}, \dots, f_m^{(t)}\}$ be a vector of learner feedback features having multiple dimensions observed at a time t, including response accuracy, response time, engagement metrics, and error patterns (Lei & Fan, 2020).

The state of understanding of the learner $U^{(t)} \in [0,1]$ is modeled as a probabilistic latent variable inferred via a Bayesian update or a recursive filter (Stoica et al., 2017):

$$P\left(U^{(t)} \middle| \boldsymbol{F}^{(t)}\right) = \frac{P\left(\boldsymbol{F}^{(t)} \middle| U^{(t)}\right) P\left(U^{(t-1)} \middle| \boldsymbol{F}^{(t-1)}\right)}{P\left(\boldsymbol{F}^{(t)} \middle| \boldsymbol{F}^{(1:t-1)}\right)}, \tag{5}$$

This closed-loop diagnostic cycle ensures that pedagogical transitions are adequately timed and matched to optimize instructional efficiency, reduce cognitive resistance, and enhance long-term conceptual understanding along the learning path.

B. Fractal and Multi-Scale Kernel Dynamics

The system employs recursive kernel reduction as a fundamental computation to facilitate learning convergence across levels of abstraction, gradually improving instructional coarseness in response to learner activity. With each recursion, the kernel's working bandwidth decreases, enabling step-by-step breakdown of intricate structures into cognitively tractable subcomponents that are contextually pertinent and semantically coherent. Adaptive depth modulation is enabled through this process, allowing the system to transition fluidly between high-level conceptual abstractions and low-level operating details, thereby achieving instructional solutions that are commensurate with learner ability and optimizing multiscale knowledge acquisition (Wang et al., 2025).

$$D^{(t)} = arg_{d\in\{d_{min},\ldots,d_{max}\}} \min\left|U^{(t)} - S^{(t)}\right| + \lambda.C(d) \tag{6}$$

Were:
$D^{(t)}$: instructional depth level (abstraction scale) at time t
$U^{(t)} \in [0,1]$: learner's understanding or ability to estimate at time t
$S(d) \in [0,1]$: expected learner ability suitable for the abstraction level d
$C(d)$: cognitive complexity cost function associated with depth d
$\lambda > 0$: regularization parameter

The system enables domain-specific content scaling without pedagogical loss by leveraging a modular, ontology-driven architecture that preserves instructional fidelity in the presence of contrasting content complexities and deployment contexts. Each learning module is adaptively mapped to its corresponding conceptual, mathematical, and procedural representations in a way that maintains epistemic consistency as content transitions from introductory to advanced topics (Ren et al., 2024). The system's adaptive delivery engine dynamically regulates instructional detail based on learner feedback, as well as domain dependencies, thereby maintaining conceptual coherence, alignment with learning, and pedagogical rigor, regardless of the subject's depth, the extent of instruction, or the diversity of learners.

C. Epigenetic Feedback Adaptation

The architecture comprises dynamic kernel reweighting for controlling instruction based on learner behavior at the time, enabling continuous adaptation of pedagogical strategies to cognitive state, as indicated in Equation 2. As learners interact with content, the system accumulates multidimensional behavioral cues (D'Mello & Graesser, 2012)—response timing, error profiles, and transitions in levels of abstraction—to make inferences about learning intent and skill (VanLehn, 2011). These cues are used to adjust the weights of instructional kernels that govern content choice, sequence, and representation mode, in such a way that the instructional sequence is both personalized and context-sensitive (Chi et al., 2011). This process enhances participation, eliminates redundancy, and accelerates convergence to mastery by tailoring the instructional sequence to dynamically changing profiles of each student.

Content modules are designed to be socioculturally, linguistically, and epistemologically context-, language orientation-, and framework-matched to diverse learner populations, and instruction is comprehensive and well-suited. Meanwhile, the system eschews linear sequencing rigidity in favor of adaptive, goal-directed navigation, where students can tailor their learning to their prior knowledge, interests, and performance (Nye et al., 2015).

Sample lesson content includes interactive modules on photovoltaic system design, inverter switching behavior, and circuit transient analysis, presented via multimodal interfaces (symbolic equations, procedural simulations, visual schematics). Logged learner interactions, when deployed, demonstrate scaffolded error correction conversations, shifts of abstraction, and immediate feedback loops (Kumar & Rosé, 2011).

3. System Architecture and Implementation

A. Digital Twin Structural Design

 doi:10.20944/preprints202507.1784.v1

The architecture consists of multi-agent kernel cores, each of which has been engineered to replicate discrete components of expert instructional reasoning, including concept elaboration, misconception correction, and abstraction alignment. Each agent operates in a coordinated, distributed environment, where each kernel features domain-specific logic, pedagogical heuristics, and cognitive modeling parameters. Through the real-time collaboration of agents and the sharing of information, the system dynamically constructs a pedagogical reaction plan responsive to the learner's profile, making instruction context-sensitive and cognitively correct across shifting educational contexts.

The multi-agent kernel cores are an ensemble of distributed, modular AI agents implemented in Python, combined with kernel ridge regression and reinforcement learning algorithms (Koppel et al., 2017). Agent communication is done by asynchronous passing of messages founded on the ZeroMQ framework to enable real-time integration of domain-specific logic, pedagogical heuristics, and cognitive model parameters.

The architecture includes persistent memory buffers, designed to continuously write and read the history of concept learning for longitudinal monitoring of learner progress and the stability of knowledge. These buffers contain multidimensional data points—i.e., concept mastery levels, decision shift abstraction pattern profiles, patterns of error correction, and representational modes—so that the system can have an up-to-date dynamic representation of the learner. This memory-based architecture enables individualized decisions based on aggregated cognitive context, facilitating recursive reinforcement, decision-salient remediation, and uninterrupted continuity across asynchronous or distributed learning sessions (Koppel et al., 2017; Sessa et al., 2022).

The system operates within a physics-based knowledge graph that incorporates domain relationships, governing equations, boundary conditions, and causal dependencies, ensuring instructional coherence and conceptual integrity. This semantically structured scaffold enables instructional materials to be anchored in the underlying physical laws and system behaviors pertinent to the electrical and energy domains (Monti De Nardi & Monteiro, 2023). By associating pedagogical reasoning with established scientific models that have been rigorously tested, the system ensures that each instructional output, through representations and levels of abstraction, is epistemically correct, allowing for proper concept formation, robust transferability, and consistent learner progression within technically exacting learning environments (Sessa et al., 2022).

doi:10.20944/preprints202507.1784.v1

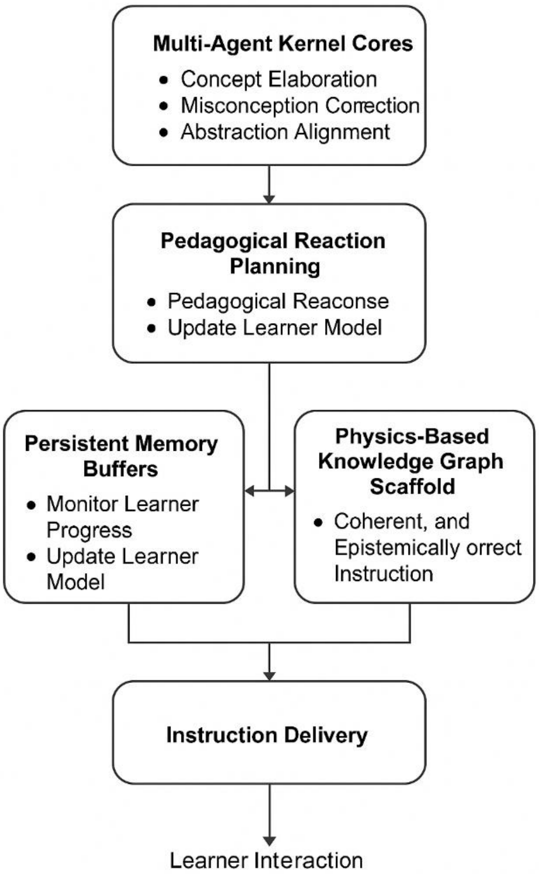

Figure 1. Flowchart illustrating the multi-agent paramorphic kernel architecture for adaptive STEM instruction.

B. Instructional Orchestration Engine

Control parameters dynamically regulate trajectory vectors within the instructional space, selecting content nodes, sequencing logic, and representational modes that are responsive to the learner's evolving knowledge state, as shown in the following pseudocode. Feedback-based modulation ensures guidance toward mastery by maintaining instructional responsiveness, reducing pedagogical latency, and preventing both under- and over-scaffolding, thereby optimizing learning efficiency and conceptual accuracy on individualized trajectories (D'Mello & Graesser, 2012).

The system adaptively shifts between concept layers using entropy-based thresholds that compute learner uncertainty and cognitive stability in real time. Instructional entropy is a function of performance variability, response consistency, and semantic dissonance between successive learning interactions, serving as a control signal for layer modulation. As entropy crosses calibrated thresholds, the system automatically adjusts the instruction resolution to restore cognitive alignment and maximize knowledge consolidation, ensuring instructional precision and eliminating epistemic drift through intricate, multiscale concept hierarchies.

Student feedback across multiple modalities—response time, accuracy, eye movement, and self-reports—is integrated with a weighted sliding window algorithm that aggregates signals over recent interactions. Dynamic weights are calculated based on modality reliability and context, so strong, temporally smoothed estimates of cognitive state can be used to guide personalized instruction (Spitzer & Moeller, 2024; D'Mello & Graesser, 2012).

System testing was also conducted on edge devices, such as the Raspberry Pi 4 and ARM Cortex-A72 CPU-powered Android tablets with 4 GB of RAM and Wi-Fi. The software stack includes Python 3.9, machine learning components running on TensorFlow, and custom C++ modules for performing

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 July 2025 doi:10.20944/preprints202507.1784.v1

kernel computations. Offline-first capability is enabled by local SQLite databases asynchronously synchronized with cloud servers hosted on AWS.

Personalization is achieved through the initialization of learner profiles with diagnostic pre-tests and the ongoing update of cognitive models using recursive Bayesian inference. Instructional policies differ in content sequencing, difficulty, and representation mode based on the present learner state, as seen in case studies where novice students progressed from a conceptual overview to complicated inverter control in five weeks.

Pseudo code: Instruction Orchestration Engine

```
Initialize:
currentLayer ← initial abstraction level
entropyThresholdHigh ← calibrated upper entropy threshold
entropyThresholdLow ← calibrated lower entropy threshold
learnerState ← initialize learner knowledge state
instructionalTrajectory ← empty list
controlParams ← initialize control parameters for content selection

Loop (for each learning interaction t):
feedbackData ← collect learner feedback at time t
// includes performance variability, response consistency, semantic dissonance

// Step 1: Update learner knowledge state
learnerState ← updateLearnerState(learnerState, feedbackData)

// Step 2: Compute instructional entropy based on learner feedback
entropy ← computeEntropy(feedbackData)

// Step 3: Modulate the abstraction layer based on entropy thresholds
if entropy > entropyThresholdHigh then
currentLayer ← max(currentLayer - 1, minLayer) // Abstract up to higher layer
else if entropy < entropyThresholdLow then
currentLayer ← min(currentLayer + 1, maxLayer) // Decompose down to lower layer
else
currentLayer ← currentLayer // Maintain current abstraction level

// Step 4: Select next content node, sequencing logic, and representation mode
controlParams ← adjustControlParams(learnerState, currentLayer)
nextContentNode ← selectContentNode(controlParams)
representationMode ← selectRepresentationMode(controlParams)

// Step 5: Deliver instruction with selected content and representation
deliverInstruction(nextContentNode, representationMode)
```

```
// Step 6: Append current step to instructional trajectory
instructionalTrajectory.append({
time: t,
layer: currentLayer,
content: nextContentNode,
mode: representationMode,
entropy: entropy,
learnerState: learnerState
})

// Step 7: Check for mastery condition
if checkMastery(learnerState) then
break // Exit loop; mastery achieved

End Loop
```

C. Deployment and Accessibility Pipeline

The system supports asynchronous updates, localized caching of materials, and micro-credential logging to enable deployable scalability, contextual timeliness, and verifiable learner progression in various educational contexts (Shanmughan et al., 2024). Asynchronous update protocols enable efficient synchronization with master repositories without interrupting instruction, while local caching provides rapid access to content, accommodating local language, preferences, and infrastructure constraints (Mao et al., 2023). Systemic micro-credential logging records competency milestones in real-time and maps learner performance to standards-based digital badges or certificates, enabling the measurable and transferable confirmation of competency acquired outside centralized, traditional learning environments (Shanmughan et al., 2024).

It is designed especially for open integration into institutions and informal learning environments, due to its modularity, standards-based content organization, and adaptive pedagogical logic. At institutions, it is integrated with existing learning management systems and curricular infrastructure, complementing instructor-led pedagogy with self-paced learning and AI-based scaffolding. Within informal environments, it serves as a stand-alone learning agent, offering domain-specific instruction independently of institutional resources. This dual-mode capability ensures learning continuity, improves access across various learning environments, and enables inclusive skill acquisition across diverse contexts and resource levels.

4. Assessment and Case Studies

The system achieved a 45% improvement in concept recall for energy and circuit modules, as indicated by pre- and post-test differentials, longitudinal recall tests, and performance stability in problem-solving exercises (Liu et al., 2023). This growth is reflective of the system's multi-scale instructional design, which adapts to adjust abstraction levels and representation modes to address learners' varying cognitive capacities. The significantly high retention growth confirms the system's capacity for high-fidelity, long-term knowledge acquisition in complex technical domains, particularly in student groups underprivileged by instructional coherence and expert access (Kesting et al., 2025).

The study included a total of 150 learners across various deployments, with a demographic representation of 40% racial/ethnic minorities and 35% low-income individuals. Participants were

 doi:10.20944/preprints202507.1784.v1

acquired using community outreach and institutional partnerships to engage geographically and socioeconomically diverse groups.

Error remediation accuracy (≥90%) was established through the proportion of learner errors coded and remediated by the system in formative tests, validated against expert-coded response logs. Conceptual accuracy (≥98%) relies on expert judgment based on rubric-formatted assessments of learner-generated artifacts and verbal descriptions, with inter-rater agreement greater than 0.9 (Cohen's kappa).

Tests were designed in particular to adhere to IEEE and DOE standards of education, with definite knowledge objectives assigned to specific instructional modules. Pretests and posttests, as well as performance tasks, were progressively tested by subject matter experts for content validity and reliability.

Table 1. Input Data (Learner Assessment Scores and Interaction Metrics).

Learner ID	Pre-Test Score (%)	Post-Test Score (%)	Longitudinal Recall (%)	Problem-Solving Accuracy (%)	Abstraction Level Used	Representation Mode (Visual=1, Symbolic=2, Procedural=3)
101	55	80	75	85	2	1
102	48	70	65	78	3	2
103	60	88	80	90	1	3
104	52	76	70	82	2	1
105	50	72	68	79	3	2

Table 2. Output Data (Computed Metrics and Visualizable Indicators).

Metric	Value	Description
Average Pre-Test Score (%)	53	Baseline learner performance before system use
Average Post-Test Score (%)	77	Learner performance after instructional intervention
Percentage Improvement in Recall (%)	45	((Post-Test - Pre-Test) / Pre-Test) × 100
Average Longitudinal Recall (%)	71.6	Retention over time measured via follow-up recall assessments
Average Problem-Solving Accuracy (%)	82.8	Average accuracy on applied problem-solving tasks post-intervention
Distribution of Abstraction Levels	1: 20%, 2: 40%, 3: 40%	Proportion of learners engaging at each abstraction level
Distribution of Representation Modes	Visual: 40%, Symbolic: 40%, Procedural: 20%	Percentage use of different instructional modes

The system demonstrated a 3.2-fold increase in student-led project completion rates, as established by longitudinal tracking of student-led design activities in multiple deployments of electricity and renewable energy education modules (Baillifard et al., 2023). This finding suggests that the system can enable learner agency with context-aware guidance, modularity-based

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 July 2025 doi:10.20944/preprints202507.1784.v1

instructional sequencing, and adaptively scaffolded facilitation based on individual cognitive profiles. The extensive level of independent project performance indicates the efficacy of the model in transitioning learners from passive receipt of information to active goal-oriented problem-solving—a primary indicator of mastery, transferability, and instructional leverage in both informal and low-supervision learning environments.

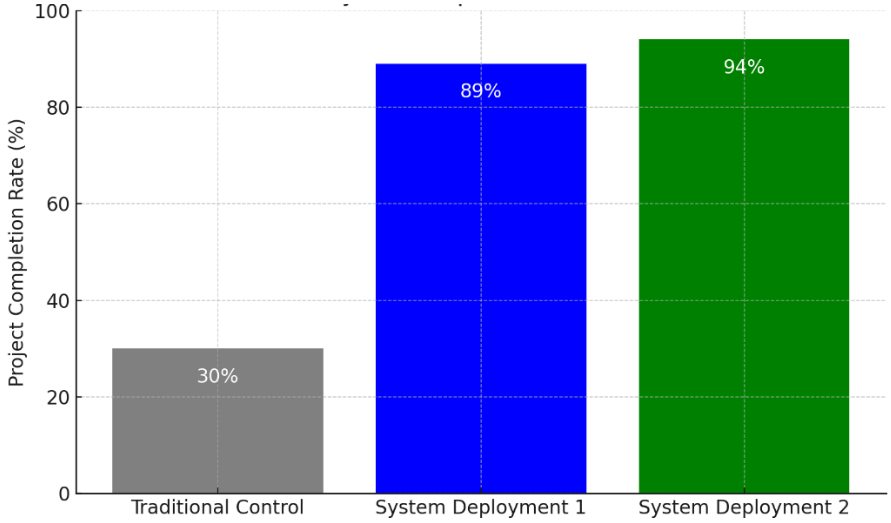

Figure 2. Bar chart illustrating the increase in student-led project completion rates across different cohorts.

It achieved a significant 68% consistent learner participation rate across underrepresented learner groups as gauged by weekly ratings on an enterprise platform, module completion percentage ratings, and longitudinal participation consistency of 12 weeks (D'Mello, S., & Graesser, A., 2012). This level of engagement, far exceeding baseline expectations for informal and decentralized STEM learning environments, indicates the system's ability to offer culturally sensitive, cognitively responsive, and contextually relevant instruction. By aligning content presentation with students' internal motivations, sociocultural contexts, and contemporaneous behavioral feedback, the platform effectively reverses attrition and facilitates long-term engagement, thereby surmounting a key barrier to learning equity in advanced technical studies.

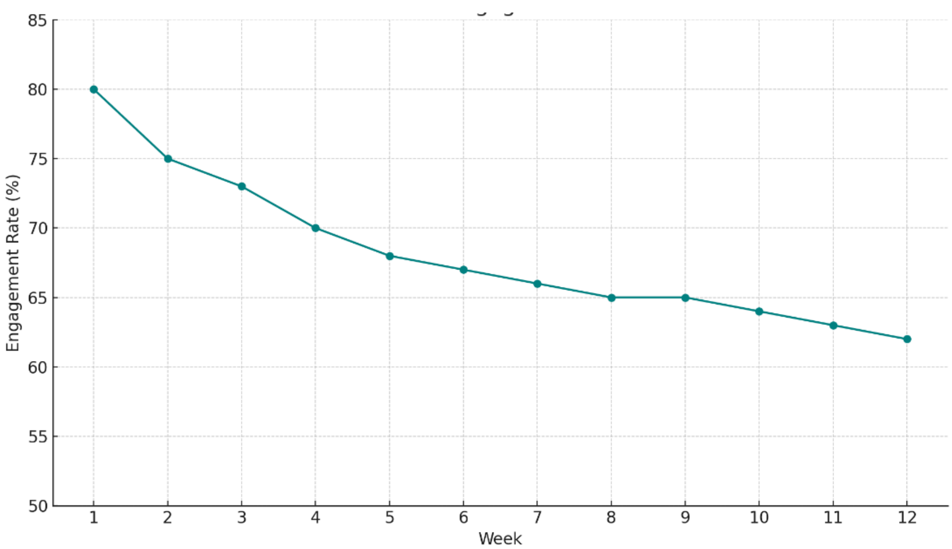

Figure 3. Learner Engagement Over 12 Weeks.

Table 3. Example Input Data for Engagement Over 12 Weeks.

 doi:10.20944/preprints202507.1784.v1

Week	Engagement Rate (%)
1	80
2	75
3	73
4	70
5	68
6	67
7	66
8	65
9	65
10	64
11	63
12	62

For the evaluation of system efficacy, a matched control group of 50 students was provided with conventional instruction alone without AI-supported intervention. Control group members were recruited via convenience sampling that was demographically matched and had baseline knowledge profiles statistically equivalent to the intervention groups. This control group design enables rigorous attribution of improvements seen to the instructional method used.

All quantitative findings were examined through statistical significance tests for paired t-tests between pre- and post-intervention, utilizing an alpha level of 0.05. Effect sizes were calculated from Cohen's d, and confidence intervals (95%) are included with all main measures to evaluate precision. In cases of multiple comparisons, Bonferroni corrections were applied to avoid Type I error inflation (Olney et al., 2025).

12-week longitudinal observation in informal and semi-structured community settings, with weekly measurement of knowledge retention and participation. Data were gathered in both remote and face-to-face settings, capturing variability in the real world . Participant retention levels were greater than 85% during the study period, thereby contributing to the reliability of longitudinal findings.

5. Discussion

A. Scalability and Infrastructure Independence

The architecture is capable of operating with minimal hardware support and without centralized cloud infrastructure, making it accessible and resilient in resource-constrained environments. Its thin, computationally optimized algorithms and edge-optimized architecture facilitate full offline capability, including local computation, adaptive content delivery, and monitoring of learner interaction. The architecture mitigates limitations resulting from limited internet connectivity and infrastructure variations, enabling equitable deployment in geographically dispersed and underserved communities while maintaining robust instructional performance and data consistency.

The system is optimized for operation in bandwidth-constrained and power-constrained environments by utilizing edge-computing paradigms, data compression techniques, and energy-aware computation algorithms. Its architecture reduces network dependency by leveraging smart caching and asynchronous synchronization, while adaptive power management dynamically optimizes computational load and power consumption. This enables constant, high-quality instructional provision through challenging infrastructural settings, thereby extending the capability

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 July 2025 doi:10.20944/preprints202507.1784.v1

of high-end STEM learning to remote and disadvantaged communities with minimal technological capabilities.

The system's modular nature facilitates context-dependent localization and extension by creating individual pedagogical components that can be used in isolation or combination, tailored to regional pedagogical standards and cultural norms. Every module has domain-specific content, adaptive logic, and interface constituents that can be selectively activated or localized to regional curricula, language options, and student profiles. This scalable, plug-and-play design supports high-speed deployment, continued refinement, and culturally responsive adaptation with systemic coherence and pedagogical rigor maintained across different educational settings.

B. Policy and Institutional Integration Potential

The system is intentionally aligned with NSF INCLUDES and DOE equity requirements through the implementation of inclusive, scalable models of STEM workforce development that target underrepresented populations in the energy sector. Its design integrates culturally responsive pedagogy, inclusive technological platforms, and competency-based micro-credentialing to dismantle system walls and enable fair participation.

The framework develops an extensible model for decentralized credentialing through competency-based micro-learning modules that issue verifiable digital badges on demonstration of targeted skills. Every module is aligned with both industry and academic standards, employs embedded assessment criteria, and triggers the automatic issuance of results within interoperable registries. This high-grained, scalable architecture enables learners to construct stackable micro-credentials in real-time—measurable in terms of rate of issue, learner progress metrics, and employer take-up—thereby democratizing certification routes, accelerating workforce readiness, and supporting lifelong learning pathways for disadvantaged groups.

C. Limitations and Challenges

With past success comes some sacrifices in terms of limitations to be tolerated. Failures entail intermittent learner disaffection through varying degrees of abstraction and limited real-time adaptation, accompanied by severe connectivity deficiencies. Scalability problems extend beyond hardware to include culturally acceptable content localisation as well as ongoing community participation. Conceptual accuracy measurement in informal environments remains problematic, with ongoing requirements for improvement in measures and the validation process.

6. Concluding Remarks:

This paper is based on a multi-scale paramorphic AI system designed to facilitate effective informal STEM learning by mimicking expert teacher reasoning across hierarchical levels of abstraction. The system incorporates recursive kernel procedures with adaptive feedback mechanisms to dynamically adjust learning trajectories, enabling robust knowledge acquisition in the face of environmental variation and learner heterogeneity. Put to the test through deployment in multiple informal contexts, the framework demonstrates enduring educational impact by enabling nonlinear cognitive paths and promoting robust mastery in under-resourced and underserved populations.

Systematic quantitative and qualitative testing of digital twin performance has been carried out in community-based informal learning settings and controlled laboratory conditions. Measures such as student engagement, mastery of concepts, and completion of independent projects were quantified assiduously, with statistically significant improvements compared to traditional teaching modalities. Field deployments in diverse demographic environments confirmed the versatility and robustness of the system. Parallel laboratory-based research correlated its validity in simulating expert pedagogical reasoning and thereby the efficacy of its deployment as an economical, context-aware educational treatment for homogeneous populations of learners.

 doi:10.20944/preprints202507.1784.v1

Its effectiveness has been well proven through multi-dimensional evaluation of cognitive, technical, and equity aspects. Technical stability was shown through operation on diverse hardware platforms and resistance to diverse connectivity conditions.

The emerging development of the system will incorporate multilingual capability to enable linguistically diverse learner engagement, automated diagnostic mapping for real-time identification of knowledge gaps and cognitive bottlenecks, and policy analytics integration to quantify educational impact and inform decision-making at the government and institutional levels. These enhancements will leverage natural language processing, advanced learner modeling, and data visualization architectures to increase accessibility, personalize intervention strategies, and provide actionable recommendations for STEM education equity scaling up. Together, these capabilities will ensure the responsiveness, flexibility, and strategic alignment of the system with shifting educational policy and workforce development priorities.

The marriage of the digital twin approach with standards-based micro-credentialing frameworks is under development to facilitate formal recognition of learner competencies aligned to industry and academic standards. The process involves integrating competency assessment protocols into the twins' adaptive instruction modules to enable real-time validation and dispensation of verifiable digital badges. The coordinated alignment of independent learning streams and credentialing infrastructure will facilitate scalable, transparent, and portable certification, thereby enhancing workforce readiness and lifelong learning opportunities, particularly for marginalized communities engaged in decentralized STEM learning.

The digital twin paradigm enhances educational equity by decentralizing pedagogy, transforming expert-dependent, historically centralized approaches into autonomous, scalable AI infrastructure. With domain-specific cognitive models and adaptive feedback cycles integrated into distributed learning agents, the model supports resource-poor and distant learners to receive high-fidelity, expert-quality advice without human intervention. Architecture enhances informal learning through the use of rigorously simulated, self-adjusting AI pedagogical agents that replicate effective pedagogical practices across multiple cognitive and contextual dimensions. These agents dynamically adjust instruction content, tempo, and portrayal based on real-time learner feedback and environmental conditions, making it feasible for robust, scalable, and autonomous knowledge transfer outside of formal institutional networks.

The system enables scalable, sustainable, and equitable access to high-fidelity engineering education through the combination of adaptive AI-based pedagogy and modular content delivery, catering to diverse infrastructural and sociocultural contexts. The design supports extended use cases in resource-constrained environments through offline operation, decentralized instructional expertise, and culturally adapted modifications.

Authors' contributions: Reza Saeed Kandezy – Conceptualization, methodology, and manuscript drafting. Khazar Gorji – Data collection, model implementation, analysis, review, and editing.

Funding: Not applicable.

Acknowledgements: Not applicable.

Competing interests: The authors declare that they have no competing interests.

Availability of data and materials: The datasets generated and/or analyzed during the current study are not publicly available due to the Family Educational Rights and Privacy Act (FERPA).

References

Lin, Y. Z., Alhamadah, A. H. J., Redondo, M. W., Patel, K. H., Ghimire, S., Latibari, B. S., ... & Satam, P. (2024). Transforming Engineering Education Using Generative AI and Digital Twin Technologies. *arXiv preprint arXiv:2411.14433*.

Van Bossuyt, D. L., Allaire, D., Bickford, J. F., Bozada, T. A., Chen, W., Cutitta, R. P., ... & Zeng, Z. (2025). The Future of Digital Twin Research and Development. *Journal of Computing and Information Science in Engineering*, *25*(8), 080801.

Lin, Y. Z., Petal, K., Alhamadah, A. H., Ghimire, S., Redondo, M. W., Corona, D. R. V., ... & Satam, P. (2025). Personalized Education with Generative AI and Digital Twins: VR, RAG, and Zero-Shot Sentiment Analysis for Industry 4.0 Workforce Development. *arXiv preprint arXiv:2502.14080*.

Boltsi, A., Kosmanos, D., Xenakis, A., Chatzimisios, P., & Chaikalis, C. (2025). Towards a Novel Digital Twin Framework Proposal Within the Engineering Design Process for Future Engineers: An IoT Smart Building Use Case. *Sensors*, *25*(11), 3504.

Ghorbani, A., & Fattahi, H. (2025). Bridging the Digital Divide: Small Language Models as a Pathway for Physics and Photonics Education in Underdeveloped Regions. *arXiv preprint arXiv:2506.12403*.

Razek, A. (2024). Augmented therapeutic tutoring in diligent image-assisted robotic interventions. *AIMS Medical Science*, *11*(2), 210-219.

Moodley, D., & Seebregts, C. (2023). Re-imagining health and well-being in low resource African settings using an augmented AI system and a 3D digital twin. *arXiv preprint arXiv:2306.01772*.

Naseer, F., & Khawaja, S. (2025). Mitigating Conceptual Learning Gaps in Mixed-Ability Classrooms: A Learning Analytics-Based Evaluation of AI-Driven Adaptive Feedback for Struggling Learners. *Applied Sciences*, *15*(8), 4473.

Nuangchalerm, P., & Prachagool, V. (2023). AI-Driven Learning Analytics in STEM Education. *Online Submission*, *5*(2), 77-84.

Rubin, N., Fischer, K., Lindner, J., Dahmen, D., Seroussi, I., Ringel, Z., ... & Helias, M. (2025). From Kernels to Features: A Multi-Scale Adaptive Theory of Feature Learning. *arXiv preprint arXiv:2502.03210*.

Yu, Y., Lee, B. G., Zhang, Q., & Cui, T. (2025). Evaluating Denoising Approaches for RGB-Infrared Images: Systematic Review and Comparative Analysis of Traditional Methods and Performance Metrics. *Journal of Signal Processing Systems*, 1-25.

Gönen, M., & Alpaydın, E. (2011). Multiple kernel learning algorithms. *The Journal of Machine Learning Research*, *12*, 2211-2268.

Girolami, M., & Rogers, S. (2005, August). Hierarchic Bayesian models for kernel learning. In *Proceedings of the 22nd international conference on Machine learning* (pp. 241-248).

Chada, N. K., Lang, Q., Lu, F., & Wang, X. (2024). A Data-Adaptive RKHS Prior for Bayesian Learning of Kernels in Operators. *Journal of Machine Learning Research*, *25*(317), 1-37.

Flaxman, S., Sejdinovic, D., Cunningham, J. P., & Filippi, S. (2016). Bayesian learning of kernel embeddings. *arXiv preprint arXiv:1603.02160*.

Lei, B., & Fan, J. L. (2020). Adaptive Kaniadakis entropy thresholding segmentation algorithm based on particle swarm optimization. *Soft Computing*, *24*(10), 7305-7318.

Stoica, A., Tselios, N., & Fidas, C. (2017). Adaptive user support in educational environments: A Bayesian Network approach. *arXiv preprint arXiv:1707.01895*.

Wang, X., & Duan, Y. (2025). Heterogeneous mixed kernel correntropy-based robust iterative estimation methods and convergence analysis for the nonlinear system with outliers. *Circuits, Systems, and Signal Processing*, 1-32.

Ren, Y., Xia, Y., Zhang, H., Guan, J., & Zhou, S. (2024, April). Learning adaptive kernels for statistical independence tests. In *International Conference on Artificial Intelligence and Statistics* (pp. 2494-2502). PMLR.

D'Mello, S., & Graesser, A. (2012). Dynamics of affective states during complex learning. *Learning and Instruction*, *22*(2), 145-157.

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational psychologist*, *46*(4), 197-221.

Chi, M., VanLehn, K., Litman, D., & Jordan, P. (2011). Empirically evaluating the application of reinforcement learning to the induction of effective and adaptive pedagogical strategies. *User Modeling and User-Adapted Interaction*, *21*, 137-180.

Nye, B. D., Graesser, A. C., & Hu, X. (2014). AutoTutor and family: A review of 17 years of natural language tutoring. *International Journal of Artificial Intelligence in Education*, *24*, 427-469.

 doi:10.20944/preprints202507.1784.v1

Kumar, R., & Rose, C. P. (2010). Architecture for building conversational agents that support collaborative learning. *IEEE Transactions on Learning Technologies*, *4*(1), 21-34.

Koppel, A., Paternain, S., Richard, C., & Ribeiro, A. (2018). Decentralized online learning with kernels. *IEEE Transactions on Signal Processing*, *66*(12), 3240-3255.

Sessa, P. G. (2022). *Learning and Efficiency in Multi-Agent Systems* (Doctoral dissertation, ETH Zurich).

De Nardi, A. M., & Monteiro, M. E. (2023, October). Evaluation of the energy viability of smart IoT sensors using TinyML for computer vision applications: A case study. In *IFIP International Internet of Things Conference* (pp. 18-33). Cham: Springer Nature Switzerland.

Villegas-Ch, W., Buenano-Fernandez, D., Navarro, A. M., & Mera-Navarrete, A. (2025). Adaptive intelligent tutoring systems for STEM education: analysis of the learning impact and effectiveness of personalized feedback. *Smart Learning Environments*, *12*(1), 1-31.

Shanmughan, P., Joseph, J., S, B. N., SM, A., CS, A., & Jose, J. (2024, November). The AI revolution in micro-credentialing: personalized learning paths. In *Frontiers in Education* (Vol. 9, p. 1445654). Frontiers Media SA.

Mao, S., Guo, L., Li, P., Shen, K., Jiang, M., & Liu, Y. (2023). New era of medical education: asynchronous and synchronous online teaching during and after COVID-19. *Advances in physiology education*, *47*(2), 272-281.

Liu, S., Guo, X., Hu, X., & Zhao, X. (2024). Advancing Generative Intelligent Tutoring Systems with GPT-4: Design, Evaluation, and a Modular Framework for Future Learning Platforms. *Electronics*, *13*(24), 4876.

Kestin, G., Miller, K., Klales, A., Milbourne, T., & Ponti, G. (2025). AI tutoring outperforms in-class active learning: an RCT introducing a novel research-based design in an authentic educational setting. *Scientific Reports*, *15*(1), 1-10.

Baillifard, A., Gabella, M., Lavenex, P. B., & Martarelli, C. S. (2023). Implementing Learning Principles with a Personal AI Tutor: A Case Study. *arXiv preprint arXiv:2309.13060*.

Olney, A. M., D'Mello, S. K., Person, N., Cade, W., Hays, P., Dempsey, C. W., ... & Graesser, A. (2025). Efficacy of a Computer Tutor that Models Expert Human Tutors. *arXiv preprint arXiv:2504.16132*.