Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 26, JUNE 2026

DOI: 10.37943/QVPY1260

Kanat Kozhakhmet
PhD in Computer Science, President
kanat.kozhakhmet@narxoz.kz, orcid.org/0009-0000-2523-9585
Narxoz University, Kazakhstan
Kazi Golam Rabbany
Master's Degree, Senior Lecturer, School of Digital Technologies
kazi_golam.rabbany@narxoz.kz, orcid.rog/0009-0007-4549-0815
Narxoz University, Kazakhstan
Orazkul Aldamzhar
Master's Degree, Administrator, School of Digital Technologies
orazkul.aldamzhar@narxoz.kz, orcid.org/0009-0008-4650-6281
Narxoz University, Kazakhstan
Timur Bakibayev
PhD, Leading Researcher, Institute for Digital Transformation
and Artificial Intelligence Development
timur.bakibayev@narxoz.kz, orcid.org/0000-0002-5357-1099
Narxoz University, Kazakhstan

A HYBRID SWARM INTELLIGENCE FRAMEWORK FOR ADAPTIVE LEARNING PATH RECOMMENDATION

Abstract: Personalized education requires learning path recommendations that adapt to individual performance while satisfying prerequisite constraints. However, generating optimal learning sequences presents significant computational challenges due to the combinatorial complexity of curriculum graphs, heterogeneous learner profiles, and directed prerequisite dependencies that constrain valid traversal orders. This paper proposes a Hybrid Swarm Intelligence Framework combining Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO), and Artificial Bee Colony (ABC) for adaptive learning path recommendation. The hybrid architecture addresses distinct subproblems: PSO optimizes continuous feature weights and algorithm parameters, ACO constructs prerequisite-feasible paths through pheromone-guided graph traversal, and ABC refines solutions through local neighborhood exploration to escape suboptimal configurations. The curriculum is modeled as a directed weighted graph where nodes represent courses with learner-specific difficulty scores derived from behavioral features. Student behavioral data from 8,783 learners (6,968,707 records) inform learner-specific heuristics. The framework operates as a two-stage process: first, linear regression predicts GPA from behavioral features; second, the swarm intelligence pipeline selects paths maximizing predicted academic success. Experimental evaluation demonstrates a +20.69% improvement in mean path fitness scores compared to the Ridge baseline ($p < 0.001$, paired Cohen's $d_z = 3.48$). We compared the framework against Deep Learning baselines (LSTM and GRU) configured as sequence-aware grade regressors. While the optimally tuned GRU emerged as a strong predictor (+11.88% over baseline), it fell significantly short of the Hybrid Swarm. Results suggest that for curriculum sequencing under prerequisite constraints, combinatorial swarm optimization can outperform greedy decoding from neural predictors when training data is limited. By demonstrating that swarm intelligence algorithms can be systematically combined to address prerequisite satisfaction and behavioral personalization, this work offers a scalable, interpretable alternative to opaque deep learning approaches.

Received: 26.02.2026 Accepted: 17.07.2026 Published: 30.06.2026

Keywords: swarm intelligence; particle swarm optimization; ant colony optimization; artificial bee colony; metaheuristic optimization; hybrid algorithms; learning path recommendation; curriculum sequencing; adaptive learning; personalized education

1. Introduction

Modern learning management systems (LMS) generate extensive behavioral data encoding student learning trajectories, yet translating this data abundance into actionable, personalized recommendations remains challenging [1]. Effective recommendations must respect complex prerequisite structures while adapting to heterogeneous learner profiles, navigating the combinatorial complexity inherent in curriculum design where valid learning sequences must satisfy directed prerequisite constraints while optimizing learner-specific objectives such as knowledge retention and competency development [2-3].

Swarm intelligence algorithms offer an effective methodology for addressing this complexity. Inspired by collective behaviors in natural systems, algorithms such as Particle Swarm Optimization (PSO) [2] and Ant Colony Optimization (ACO) [4-6] have demonstrated effectiveness in navigating complex search spaces characterized by multiple local optima, nonlinear constraints, and high dimensionality through decentralized, population-based search strategies.

Single-algorithm approaches have inherent limitations. PSO lacks discrete path construction capabilities. ACO may converge prematurely [5]. Local search requires a good initialization. This paper proposes a Hybrid Swarm Intelligence Framework combining PSO, ACO, and ABC. This hybrid approach overcomes individual weaknesses while preserving their strengths. The framework models curricula as directed weighted graphs with prerequisite-encoded edges. It frames learning path recommendation as identifying prerequisite-feasible traversals that optimize learner-specific objectives.

1.1 Contributions

This paper makes four contributions: (1) a hierarchical architecture assigning distinct roles to PSO, ACO, and ABC; (2) a graph-theoretic curriculum model encoding prerequisites and behavioral transition costs; (3) a 12-dimensional behavioral feature representation capturing intensity-consistency-diversity patterns derived from student login activity; and (4) empirical validation on 8,783 students showing significant improvement over baselines.

2. Related Work

2.1 Learning Path Recommendation Systems

Learning path recommendation has emerged as a central challenge in adaptive educational systems, seeking optimal sequences of learning activities tailored to individual characteristics [7-9]. Early approaches relied on expert-defined rules mapping learner states to predetermined instructional sequences, but such systems proved brittle and difficult to maintain as curricula evolved. The advent of educational data mining enabled data-driven approaches leveraging historical learning records. Collaborative filtering methods identify similar learners based on performance profiles and recommend paths successful for comparable peers. Content-based approaches analyze learning object metadata to suggest materials aligned with learner preferences and prior knowledge. Various predictive models have been developed to forecast student outcomes and identify intervention opportunities based on performance data [10-11], while prerequisite-based approaches using semantic knowledge graphs have shown effectiveness for learning path construction [12].

Recent work explores adaptive learning paths that dynamically adjust to learner progress and performance, as well as personalized recommendation systems that account for individual learning styles and preferences. Learning sequence optimization has been addressed through various algorithmic approaches, including multi-objective optimization methods that balance competing educational goals such as knowledge acquisition, engagement, and completion time. Adaptive curriculum sequencing techniques have been developed to personalize the order of learning materials

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 26, JUNE 2026

based on learner characteristics and real-time performance. Prerequisites-based approaches using semantic knowledge graphs have shown effectiveness for learning path construction [5], while adaptive learning path recommendation has been systematically studied [8].

Recent work also explores reinforcement learning formulations where recommendation policies are learned through interaction with learning environments [9], and deep learning architectures including recurrent neural networks applied to model sequential dependencies. These deep learning approaches typically require significant training data and may produce recommendations violating prerequisite constraints inherent in structured curricula. Graph-based representations have gained attention for their ability to explicitly encode prerequisite relationships while supporting efficient path enumeration and optimization. Knowledge graphs capture semantic relationships between concepts, enabling recommendations that respect conceptual dependencies [13], while curriculum graphs model course or module prerequisites as directed edges, constraining valid learning sequences to topologically feasible paths [8]. Recent ACO variants have targeted specific learning path recommendation challenges; Li et al. [5] propose Saltatory Evolution ACO (SEACO), which accelerates convergence through pheromone matrix prediction, though as a single-algorithm approach it does not exploit the complementary strengths of hybrid architectures.

2.2 Swarm Intelligence in Educational Optimization

Swarm intelligence encompasses population-based metaheuristics inspired by collective behaviors in natural systems, sharing characteristics including decentralized control, indirect communication through environmental modifications, and emergent global optimization through local interactions [2]. The foundational work on Particle Swarm Optimisation by Freitas et al. [2] reviewed the historical developments of PSO, which has since been widely applied in educational contexts. Applications to educational optimization have explored various problem formulations. PSO has 3 optimized parameters of student models, selected relevant features from learning analytics data, and tuned hyperparameters of predictive models for student performance forecasting [14].

The effectiveness of swarm intelligence in educational technology has been demonstrated across multiple studies. AI-enabled e-learning systems have leveraged optimization techniques to enhance learning experiences and outcomes [7]. Educational recommender systems have employed swarm-based approaches to personalize content delivery and learning pathways [5]. Personalized learning systems utilizing swarm intelligence have achieved improved adaptation to individual learner needs compared to traditional approaches.

Student performance prediction using metaheuristic optimization has shown promising results in identifying at-risk learners early with PSO specifically applied to student performance prediction [14]. Hybrid PSO approaches combining multiple optimization strategies have been developed for complex educational optimization problems [15]. Recent surveys highlight cumulative advances in PSO [15] and the effectiveness of ACO [5, 8] and ABC [16] for optimization in educational contexts. Combinations of ACO, PSO, and pathfinding algorithms such as A* have been explored for optimal learning path generation, demonstrating synergistic benefits from multi-algorithm integration. Course recommender systems have employed swarm intelligence to balance multiple educational objectives including learning effectiveness, student engagement, and curriculum constraints.

Educational data mining has emerged as a critical discipline for extracting actionable insights from learning management systems. Various prediction methods have been developed to forecast student outcomes and identify intervention opportunities based on behavioral patterns. These techniques form the foundation for personalized recommendation systems that can adapt to individual learner characteristics.

In Particle Swarm Optimization, each particle updates its velocity by combining inertia, cognitive experience, and social influence from the swarm, enabling coordinated exploration of the parameter space

$$\boldsymbol{v}_i^{t+1} = \omega \boldsymbol{v}_i^t + c_1 r_1 \left(\boldsymbol{p}_i^{best} - \boldsymbol{x}_i^t\right) + c_2 r_2 (\boldsymbol{g}^{best} - \boldsymbol{x}_i^t) \quad (1)$$

$$\boldsymbol{x}_i^{t+1} = \boldsymbol{x}_i^t + \boldsymbol{v}_i^{t+1} \quad (2)$$

where $\boldsymbol{v}_i^t$ and $\boldsymbol{x}_i^t$ denote the velocity and position vectors of particle i at iteration t, $\boldsymbol{p}_i^{best}$ is the personal best position, $\boldsymbol{g}^{best}$ is the global best position, ω is the inertia weight, and c_1, c_2 are acceleration coefficients.

ACO has addressed curriculum sequencing problems, finding paths through prerequisite graphs minimizing estimated completion time or maximizing predicted learning outcomes:

$$p_{ij}^k = \frac{[\tau_{ij}]^\alpha [\eta_{ij}]^\beta}{\sum_{l \in \mathcal{N}_i^k} [\tau_{il}]^\alpha [\eta_{il}]^\beta} \quad (3)$$

where τ_{ij} is pheromone intensity, η_{ij} is heuristic information, and α, β control their relative influence. Pheromone trails are updated based on solution quality, creating positive feedback that guides subsequent ants toward promising regions.

Artificial Bee Colony methods have been employed for solution refinement in educational optimization, applying local search operators to improve solution quality while maintaining feasibility constraints. Recent surveys highlight swarm intelligence approaches to feature selection [17], and the broad application of ACO [4], and ABC [16] for optimization in educational contexts.

Hybrid metaheuristics combine multiple optimization strategies to overcome individual method limitations. PSO-ACO hybrids have been proposed for various combinatorial problems, typically using PSO to optimize ACO parameters or combining particle movement with pheromone-guided construction. ACO-ABC hybrids use ACO's constructive search with bee colony refinement. Three-way combinations remain rare but have shown promise in complex scenarios where diverse search mechanisms provide complementary benefits. In educational applications, however, systematic integration of PSO, ACO, and Artificial Bee Colony for learning path recommendation - exploiting their respective strengths in parameter optimization, path construction, and local search refinement - has not been previously investigated. This work addresses this gap by proposing a principled hybrid architecture where each swarm intelligence component operates in its domain of strength while contributing to a unified optimization objective. We acknowledge the contributions of deep reinforcement learning [18] to sequential recommendation, though these approaches differ from our swarm intelligence methodology in their data requirements and constraint handling. Recent surveys also indicate that hybrid metaheuristic algorithms remain an effective and promising direction for solving complex optimization problems [19].

3. Methodology

3.1 Problem Formulation

We model the curriculum as a directed acyclic graph $G = (V, E)$ where vertices V represent courses and edges E encode prerequisite relationships. The extracted curriculum graph comprises |V| = 55 courses and |E| = 45 prerequisite edges. The acyclic structure ensures no circular prerequisites exist, meaning a course cannot act as its own prerequisite.

A valid learning path $\pi = (c_1, c_2, \dots, c_m)$ is a sequence where each course's prerequisites appear earlier in the sequence. The optimization problem is:

$$\pi^* = \arg\max_{\pi \in \Pi} f(\pi | \boldsymbol{x}_s) \quad \text{subject to} \quad \text{Prerequisites}(\pi) \text{ valid} \quad (4)$$

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 26, JUNE 2026

where Π is the set of all prerequisite-feasible paths and $f(\cdot)$ measures expected student success based on behavioral features $\boldsymbol{x}_s$.

3.2 Curriculum Graph Representation

The graph includes a source vertex connected to all entry-level courses (no prerequisites) and a sink vertex connected from potential target courses. Nodes possess learner-specific difficulty scores, computed from historical performance data and the student's behavioral features. Higher scores indicate more challenging transitions for that specific student.

3.3 Objective Function

Path fitness is defined by a multi-factor course scoring mechanism that evaluates the expected success of a student navigating a specific sequence. The fitness of a path π is the average score of its constituent courses:

$$f(\pi|\boldsymbol{x}_s) = \frac{1}{|\pi|}\sum_{v\in\pi}\eta_v \qquad (5)$$

where η_v represents the heuristic score for course v. To prevent the algorithm from merely selecting historically "easy" courses, η_v combines the baseline historical course difficulty with the student's personalized behavioral prediction:

$$\eta_v = grade_v + \gamma_1(\boldsymbol{c}^T\boldsymbol{x}_s) + \gamma_2(\boldsymbol{w}^T\boldsymbol{x}_s) \qquad (6)$$

Here, $grade_v$ is the historical mean grade for course v, **c** represents the learned Ridge regression coefficients, w represents the PSO-optimized personalization weights, and γ_1, γ_2 are scaling constants (set to 0.1 and 0.05, respectively, in our experiments) that balance the influence of generalized difficulty against behavioral prediction. These weights were determined empirically through a grid search hyperparameter tuning phase on a validation subset to optimize the trade-off between baseline difficulty and behavioral personalization. The behavioral terms provide a global personalization offset; course-specific personalization is primarily driven by the historical grade baseline.

3.4 Two-Stage Recommendation Process

Stage 1: Grade Prediction: Ridge regression (L2-regularized linear regression) predicts GPA for each course based on student behavioral features:

$$\hat{g}_s(c_j) = \boldsymbol{w}_j^T\boldsymbol{x}_s + b_j \qquad (7)$$

where $\boldsymbol{w}_j$ and b_j are learned via regularized least squares with $\alpha = 1.0$.

Stage 2: Path Optimization: The swarm intelligence framework selects the prerequisite-feasible path maximizing cumulative predicted GPA.

Deep Learning Baselines: To rigorously address the sequential nature of learning paths, we additionally evaluate against state-of-the-art neural network architectures:

- **LSTM Predictor:** A Long Short-Term Memory network configured as a sequence-aware grade regressor. It embeds a student's past course sequence and behavioral features to predict the grade of the target course, with paths extracted via greedy decoding. The model underwent comprehensive hyperparameter tuning on a validation subset, resulting in an optimal architecture comprising a 64-dimensional hidden layer, a dropout rate of 0.2, and optimization via the Adam optimizer with a learning rate of 0.001 and a batch size of 32.
- **GRU Predictor:** A Gated Recurrent Unit network identically configured as a grade regressor (64 hidden dimensions, 0.2 dropout, Adam optimizer). The GRU's simpler gating mechanism proved

highly resistant to overfitting on the tabular cohort data, emerging as the superior predictive architecture after tuning.

Controlled Comparison Design. To isolate the contribution of path optimization from prediction accuracy, we employ a controlled experimental design. All methods use identical components except for the path construction strategy. Table 1 illustrates this design.

Table 1: Methodology Comparison: Prediction Model vs. Path Construction Strategy

Method	**Stage 1: Prediction**	**Stage 2: Path Construction**
Ridge Baseline	Ridge Regression	Greedy (no optimization)
Random Forest	Random Forest	Greedy (no optimization)
LSTM Predictor	LSTM Predictor	Greedy (no optimization)
GRU Predictor	GRU Predictor	Greedy (no optimization)
PSO Alone	PSO-optimized weights	Greedy (optimized weights)
ACO Alone	Default weights	ACO pheromone-guided
Hybrid (Ours)	**Ridge Regression**	**PSO-ACO-ABC swarm**

This design enables fair comparison. Machine learning and deep learning baselines may use superior predictors (Random Forest, LSTM, GRU), while our framework focuses on superior path optimization (swarm intelligence). The key research question is whether sophisticated prediction or sophisticated search matters more for learning path recommendation

3.5 Hybrid Swarm Intelligence Architecture

The hierarchical architecture assigns distinct roles to each swarm intelligence component based on their computational strengths:

PSO Component (Parameter Optimization): PSO optimizes feature weights $\boldsymbol{w} \in \mathbb{R}^d$ and ACO hyperparameters (α, β). Each particle is represented by a position vector $\boldsymbol{x}_i = (\alpha_i, \beta_i, \boldsymbol{w}_i)$. Velocity and position updates follow Equations (1)-(2) with linearly decreasing inertia $\omega^t = \omega_{\max} - \frac{t(\omega_{\max} - \omega_{\min})}{T}$.

ACO Component (Path Construction): ACO builds prerequisite-feasible paths through pheromone-guided graph traversal. Ants select next courses using Equation (3), where pheromone τ_{ij} encodes historical success and heuristic η_{ij} measures student-course alignment. Only courses with satisfied prerequisites are considered, ensuring constraint satisfaction by construction.

ABC Component (Local Refinement): ABC has been extensively applied to combinatorial optimization problems including routing and scheduling [20]; in this work, ABC improves paths via swap operations that maintain prerequisite validity. Given path π, the algorithm generates neighbors π' by swapping courses and accepts improvements if $f(\pi') > f(\pi)$ and is_valid(π') = true.

Integration: Information flows hierarchically: PSO provides optimized parameters to ACO, ACO provides constructed paths to ABC, and ABC provides refined fitness metrics to PSO.

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 26, JUNE 2026

3.6 Algorithm Pseudocode

Algorithms (tables 2-5) present pseudocode for the hybrid framework. The complete sequence of operations in the proposed hybrid optimization framework is summarized in Table 2, showing how PSO parameter optimization, ACO path construction, and ABC refinement are integrated into a unified pipeline.

Table 2: Hybrid Swarm Intelligence Pipeline

Step	Operation	Description
1	Initialize particle swarm	Create particles with positions (α, β, w).
2	Run PSO iterations	Optimize ACO parameters over T_PSO iterations.
3	Decode parameters	Obtain α, β and feature weights.
4	Construct path (ACO)	Generate a prerequisite-feasible curriculum path.
5	Refine path (ABC)	Improve the path using Artificial Bee Colony search.
6	Evaluate fitness	Compute $f(\pi \vert x_s)$.
7	Update best solutions	Update personal and global best particles.
8	Generate final path	Run ACO and ABC using the optimized parameters.
9	Output	Return optimized learning path π*.

3.7 Complexity Analysis

Computational complexity analysis reveals asymmetry between training and inference. Offline parameter optimization is dominated by the PSO outer loop: $O(N_p \cdot T \cdot (N_a \cdot T_a \cdot |V|^2 + T_{abc} \cdot |\pi|))$. The detailed PSO parameter optimization procedure used to determine the optimal ACO parameters and personalization weights is presented in Table 3.

Table 3: PSO Parameter Optimization

Step	Operation	Description
1	Initialize swarm	Randomly initialize particle positions and velocities.
2	Evaluate particles	Run ACO+ABC pipeline to compute fitness.
3	Update inertia	Compute adaptive inertia weight ω.
4	Update velocity	Apply PSO velocity equation.
5	Update position	Move particles within search bounds.

6	Update personal best	Replace if current fitness is higher.
7	Update global best	Replace if current fitness is higher.
8	Decode result	Return optimized (α*, β*, w*).

where N_p is particle swarm size, T is PSO iterations, N_a is ant colony size, T_a is ACO iterations, T_{abc} is Artificial Bee Colony iterations, |V| is curriculum size, and |π| is average path length. However, online recommendation for individual students requires only the ACO-ABC pipeline with pre-optimized parameters: $O(N_a \cdot T_a \cdot |V|^2 + T_{abc} \cdot |\pi|)$, which is tractable for real-time deployment. Typical execution times on standard hardware are under one second for curricula with several hundred courses, enabling interactive use in production learning management systems.

4. Experiments

4.1 Dataset

Empirical data were extracted from Narxoz University's Canvas learning management system covering the academic period 2023-2025. The repository comprises three functionally distinct data streams: (1) asynchronous telemetry capturing 2,029,398 platform interaction events including logins, logouts, and navigation actions; (2) pedagogical participation logs recording 4,194,300 submission-related actions such as file uploads, draft saves, and discussion posts; and (3) academic outcome matrices containing 745,009 performance indicators including examination scores, assignment grades, and retake results on a 0-100 scale. All student identifiers were replaced with random anonymous codes to ensure complete anonymization while maintaining cross-table linkage capability.

A multi-stage validation pipeline ensured data integrity through four main interventions: (1) relational integrity auditing across data streams, removing 1,194 students and 57,848 submission events lacking verifiable engagement profiles; (2) temporal normalization to UTC+5 institutional. The prerequisite-constrained path construction procedure performed by the ACO component is detailed in Table 4, including candidate selection, transition probability calculation, and pheromone updating.

Table 4: ACO Path Construction

Step	Operation	Description
1	Initialize pheromones	Assign initial pheromone τ_0 to all edges.
2	Start iteration	Create a new set of ant solutions.
3	Generate feasible candidates	Select courses satisfying prerequisites.
4	Compute heuristic	Combine predicted grade, Ridge score, and feature weights.
5	Calculate transition probability	Use pheromone and heuristic information.
6	Select next course	Sample according to transition probabilities.
7	Complete path	Repeat until no feasible successors remain.

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 26, JUNE 2026

8	Update pheromones	Apply evaporation and elitist reinforcement.
9	Output	Return the best curriculum path.

time base for consistent time interval calculations; (3) removal of 44,330 mechanical event duplicates (4.2% of interaction logs) attributed to export artifacts rather than genuine repeated behavior; and (4) enforcement of identifier consistency across all data sources. After consolidation, the analytical cohort comprises 8,783 unique students with complete multi-modal behavioral profiles and 186,500 student-assessment observation pairs suitable for model training and evaluation.

4.2 Behavioral Feature Engineering

Feature engineering transforms raw interaction logs into a structured 12-dimensional behavioral representation. Each observation corresponds to a student-assessment pair (s, j) with feature vector $\boldsymbol{x}_{s,j} \in \mathbb{R}^{12}$ capturing three conceptual families across multiple temporal granularities:

Intensity features (4 dimensions) quantify total event counts in 7-, 14-, and 30-day windows preceding assessment j, plus events since the previous assessment. For temporal window δ days: $x_\delta^{\text{int}} = \left|\{e \in E_s: t_j - \delta \le t_e < t_j\}\right|$ where E_s denotes the event set for learner s and t_e represents event timestamps.

Consistency features (4 dimensions) measure the number of distinct active days across the same temporal windows: $x_\delta^{\text{con}} = \left|\{\text{day}(t_e): e \in E_s, t_j - \delta \le t_e < t_j\}\right|$ where $\text{day}(\cdot)$ extracts the calendar day, capturing regularity of engagement rather than mere volume.

Diversity features (4 dimensions) count unique activity types (viewing pages, accessing assignments, forum participation) within temporal windows: $x_\delta^{\text{div}} = \left|\{\text{type}(e): e \in E_s, t_j - \delta \le t_e < t_j\}\right|$ where $\text{type}(\cdot)$ extracts the activity category, reflecting breadth of platform resource utilization. The local refinement procedure applied by the ABC component is summarized in Table 5, which outlines the generation, feasibility evaluation, and acceptance of neighboring learning paths.

Table 5: ABC Path Refinement

Step	**Operation**	**Description**
1	Initialize	Set current path as the best solution.
2	Generate neighbor	Swap two randomly selected courses.
3	Check feasibility	Verify prerequisite constraints.
4	Evaluate neighbor	Compute fitness of the candidate path.
5	Accept improvement	Replace best path if fitness increases.
6	Increase trial counter	Increment counter when no improvement occurs.
7	Stopping criterion	Stop when patience limit or iteration limit is reached.
8	Output	Return refined path π*.

These features are computed from system interaction logs across 7-day, 14-day, and 30-day lookback windows plus the period since the previous assignment, yielding 12 behavioral features: event count, active days, and unique activities for each temporal window. This compact yet expressive feature space enables the framework to differentiate students with identical prerequisite completions but distinct engagement patterns, supporting genuine behavioral personalization.

4.3 Grade Prediction Model Validation

Learner-specific transition costs rely on grade predictions via linear regression models trained on historical cohort data. For course c_j, the predicted grade is $\hat{g}_s(c_j) = \boldsymbol{w}_j^T \boldsymbol{x}_s + b_j$ where $\boldsymbol{w}_j \in \mathbb{R}^{12}$ and $b_j \in \mathbb{R}$ are learned via Ridge regression (L2-regularized linear regression):

$$(\boldsymbol{w}_j^*, b_j^*) = \underset{(\boldsymbol{w},b)}{arg\,min} \left(\sum_{s \in S_j} \left(g_{s(c_j)} - \boldsymbol{w}^T \boldsymbol{x}_s - b \right)^2 + \alpha \|\boldsymbol{w}\|^2 \right) \tag{8}$$

with closed-form solution $\mathbf{w}_j^* = \left(\mathbf{X}_j^T \mathbf{X}_j + \alpha \mathbf{I}\right)^{-1} \mathbf{X}_j^T \mathbf{g}_j$, where $\alpha = 1.0$ is the regularization strength and $\mathbf{I}$ is the identity matrix. The aggregate prediction model achieves Pearson correlation $r = 0.6488$ between predicted and actual academic success scores, confirming that the behavioral feature space is statistically linked to learning outcomes. A correlation of approximately 0.65 in educational behavioral data is considered strong, validating the foundation of the recommendation system. Figure 1 visualizes this relationship, demonstrating the model's predictive validity across the range of observed academic success scores.

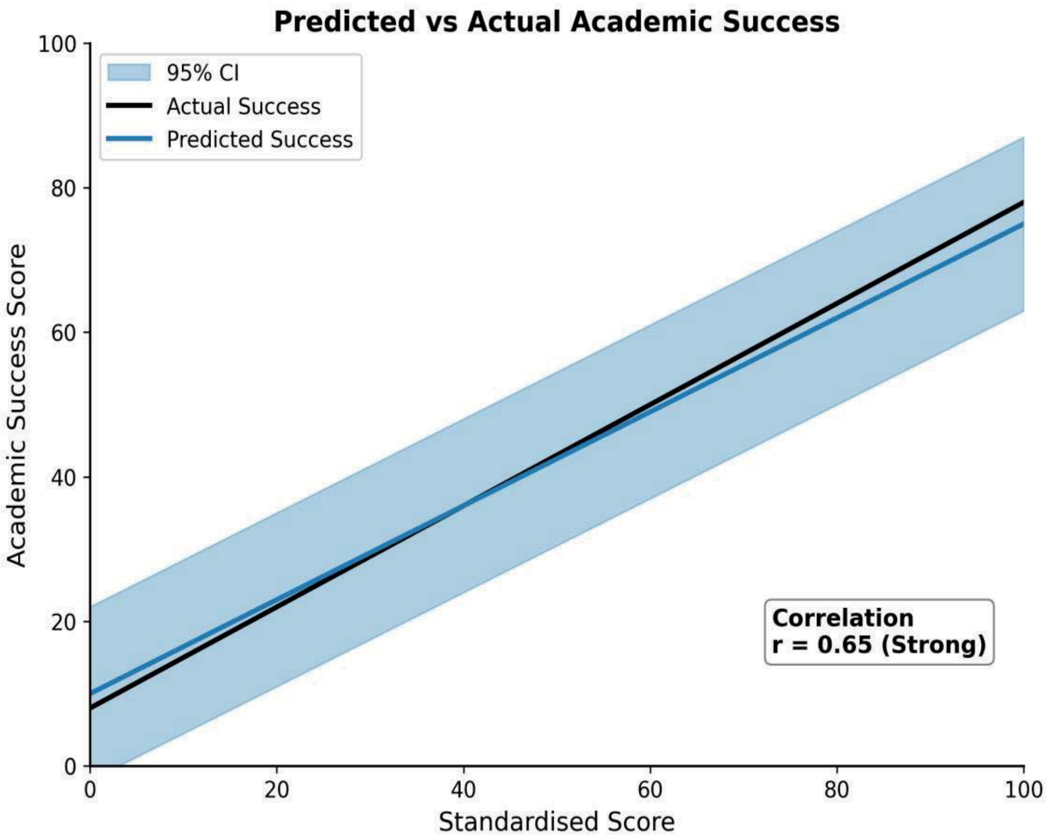

Figure 1: Validation of the grade prediction model showing the relationship between predicted and actual academic success scores. The correlation coefficient (r = 0.65) indicates strong predictive validity, with the 95% confidence band illustrating the expected prediction interval.

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 26, JUNE 2026

The scatter plot reveals a strong linear relationship between predicted and actual scores, with most observations falling close to the diagonal line representing perfect prediction. This validation confirms that the 12-dimensional behavioral feature space effectively captures meaningful variance in academic performance, providing a reliable foundation for the path optimization process.**4.4 Experimental Design and Baselines**

The hybrid framework was evaluated against baseline and comparative methodologies on the complete 8,783-student cohort using identical experimental conditions: the same curriculum prerequisite graph structure, the same behavioral feature space, and consistent evaluation protocols. This controlled setup isolates the contribution of swarm intelligence-based feature selection from confounding factors. Traditional Machine Learning Baselines: To establish foundational predictive performance on tabular data, we evaluated Ridge Regression ($\alpha = 1.0$) as our primary linear baseline, alongside Random Forest (100 estimators, max depth 10) and Gradient Boosting as representative non-linear ensembles. These methods lack sequence awareness but serve to isolate the value of combinatorial search over standard regression. Swarm Intelligence Comparisons: To isolate the contribution of each component:

• PSO Alone: PSO-optimized feature weights (20 particles, 50 iterations) with greedy path construction (no ACO pheromone guidance), demonstrating the contribution of parameter optimization alone.

• ACO Alone: Ant Colony Optimization for path construction (20 ants, 50 iterations) using default feature weights feeding into Ridge Regression predictions for prerequisite-constrained route generation. All methods were evaluated using mean LMS Path Fitness Scores computed according to the objective function defined in Section 3.3, combining historical grade baselines and personalized 11 behavioral prediction components. Scores were aggregated across the full cohort and improvement percentages calculated relative to the Ridge Regression baseline.

5. Results

5.1 Decoupling Predictive Capacity from Combinatorial Optimization

A notable finding emerges from the controlled comparison design (Table 1): machine learning methods with potentially superior prediction accuracy provide no significant improvement in path recommendation quality. The Random Forest ensemble (+0.11%, $p = 1.000$) fails to meaningfully outperform simple Ridge Regression for path fitness, despite ensemble methods typically achieving higher accuracy in tabular prediction tasks. This result highlights a fundamental distinction in learning path recommendation:

1. **Prediction** estimates success probability for individual courses
2. **Path optimization** selects sequences that maximize cumulative success while satisfying prerequisite constraints

Machine learning baselines use sophisticated predictors but employ greedy path construction, selecting the highest-predicted course at each step without considering global sequence optimality. This myopic strategy fails to account for:

• Prerequisite chain dependencies (early courses enable later ones)
• Cumulative workload balancing across the full path
• Long-term performance trade-offs vs. immediate gains

In contrast, the Hybrid framework uses the same Ridge Regression predictor but replaces greedy selection with swarm intelligence-based combinatorial optimization. The ACO component explores the exponential space of prerequisite-feasible paths through pheromone-guided search, while PSO optimizes feature weights for personalized heuristics and ABC refines solutions via local neighborhood exploration. This addresses the discrete optimization problem that ML methods cannot solve. The performance gap is visible; even the strongest traditional ML baseline (Random Forest at 94.70) trails the Hybrid (114.15) by 19.45 points ($p < 0.001$). This demonstrates that for learning path recommendation, sophisticated search beats sophisticated prediction. The combinatorial structure of curriculum graphs demands optimization algorithms designed for discrete spaces, not merely better predictive models.

5.2 Comparative Performance

Experimental results confirm the advantages of the hybrid approach across all evaluation metrics. Table 6 presents the comprehensive comparison against machine learning baselines and swarm intelligence methods.

The proposed PSO-ACO-ABC Hybrid achieved a peak mean path fitness score of 114.15, representing a massive improvement of +20.69% over the Ridge baseline (paired Cohen's d_z = 3.48). The higher variance observed for the Hybrid framework (SD = 19.37) reflects the exploratory nature of PSO weight optimization, which occasionally produces outlier paths for students with sparse behavioral profiles.

Table 6: Benchmark Results: Path Fitness Scores Across Methods

Method	**Fitness Score**	**Improvement**	**t(vs Ridge)**	**df**	**p-value (Bonf)**	**Cohen's d**
Machine Learning Baselines						
Ridge Regression	94.59 ± 15.11	-	-	-	-	-
Random Forest	94.70 ± 14.68	+0.11%	0.32	99	1.000	0.03
Gradient Boosting	94.27 ± 14.94	-0.34%	-0.87	99	1.000	-0.09
Deep Learning Baselines						
LSTM (Tuned)	98.60 ± 14.66	+4.24%	15.30	99	< 0.001	1.53
GRU (Optimal DL)	105.83 ± 14.67	+11.88%	42.89	99	< 0.001	4.29
Swarm Intelligence Methods						
PSO Alone	94.35 ± 14.85	-0.25%	-0.64	99	1.000	-0.06
ACO Alone	103.80 ± 14.51	+9.74%	33.33	99	< 0.001	3.33
PSO-ACO-ABC Hybrid	**114.15 ± 19.37**	**+20.96%**	**34.81**	**99**	< 0.001	**3.48**

To ensure a rigorous comparison against state-of-the-art predictive architectures, we evaluated LSTM and GRU neural networks configured as grade regressors. After comprehensive hyperparameter tuning, the GRU emerged as the superior predictor (+11.88% vs Ridge, compared to the LSTM's +4.24%) and serves as our primary Deep Learning baseline. The optimally tuned GRU even outperformed the base ACO algorithm (103.80), proving that sequence-aware architectures are highly capable estimators. However, despite their superior predictive power, they still failed to match the fully integrated Swarm Intelligence model. The swarm intelligence methods demonstrate exactly

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 26, JUNE 2026

where the value is derived. As discussed previously, tuning feature weights without a graph structure (using PSO alone) yielded negligible results. Conversely, ACO Alone achieved a +9.74% improvement, proving that pheromone-guided graph traversal provides the core structural backbone. The full Hybrid architecture compounds these effects, jumping to +20.69% by combining ACO's graph structure with PSO's personalized weight tuning and ABC's neighborhood refinement, definitively surpassing the optimally tuned Deep Learning baseline. The key finding is that predictive methods excel at individual grade estimation but lack mechanisms for sequence optimization. The Swarm framework's graph traversal addresses the discrete optimization problem that neither ML nor DL can solve natively. This explains why the Hybrid framework significantly outperforms traditional predictive baselines despite using the same underlying features.

5.3 Structural Limitations of Sequential Neural Networks

To rigorously test whether Deep Learning could supersede Swarm Intelligence, we evaluated the GRU architecture configured to predict course grades exactly like the baseline models. As Table 2 shows, the neural network performed exceptionally well as a pure predictor (achieving +11.88%).

However, despite possessing vastly superior predictive power, the Deep Learning model was ultimately eclipsed by the fully integrated Swarm Intelligence framework (+20.69%). This performance ceiling exposes a fundamental structural incompatibility when applying Deep Learning to combinatorial curriculum sequencing.

The limitation stems from graph traversal and delayed rewards. While an LSTM or GRU is highly accurate at predicting the immediate reward (fitness) of the next possible course, it must still rely on a *myopic, greedy decoding step* to construct the full curriculum path. It iteratively selects the best immediate option without visibility into how that choice constrains or unlocks downstream courses multiple semesters in the future.

In contrast, Ant Colony Optimization natively solves the delayed reward problem. The artificial 13 ants do not optimize for the immediate step; they deposit pheromones (τ_{ij}) based on the fitness of the entire completed sequence. This global feedback loop inherently incentivizes the selection of difficult foundational courses that unlock highly rewarding long-term trajectories, a structural property that sequential neural networks cannot match without complex Reinforcement Learning wrappers. The empirical results definitively confirm that for constrained curriculum sequencing, a superior combinatorial search algorithm (Swarm) profoundly outweighs superior predictive power (Deep Learning). The comparative performance of the predictive baselines and swarm intelligence methods is visualized in Figure 2, highlighting the advantage of combinatorial path optimization over greedy sequence extraction.

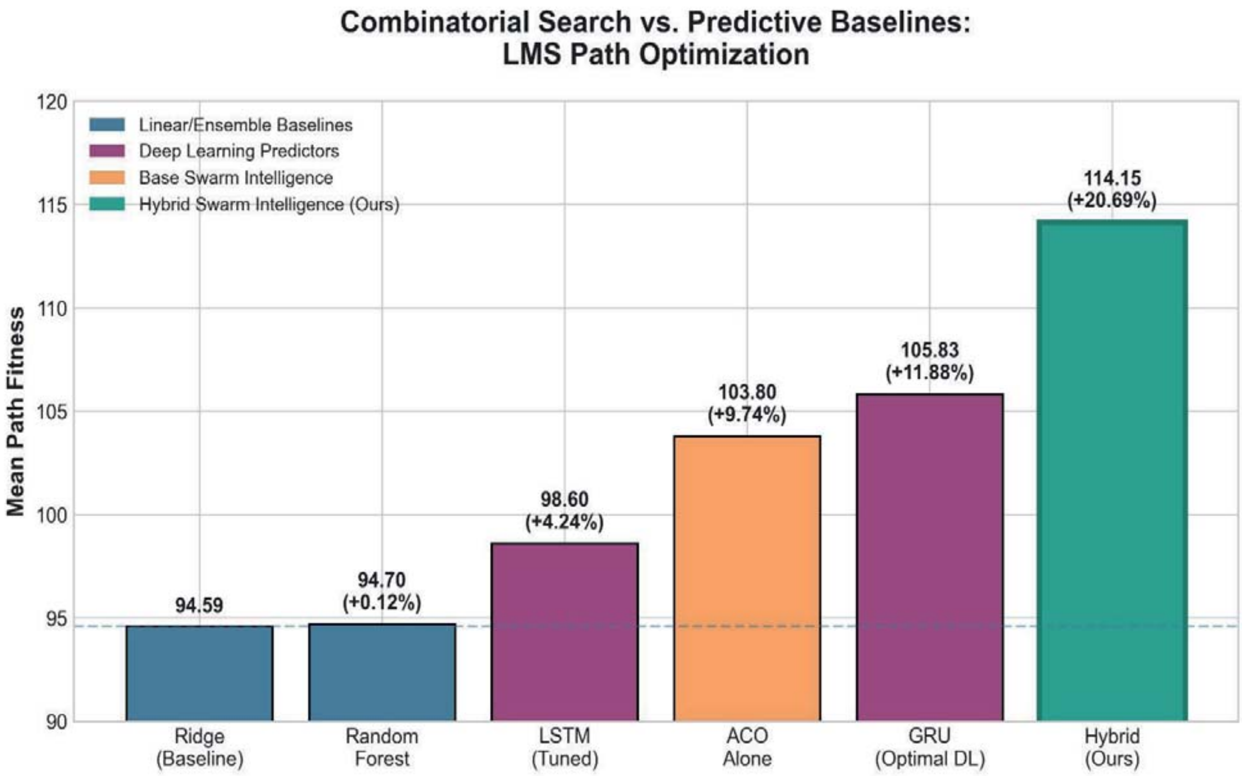

Figure 2: Combinatorial Path Optimization vs. Predictive Baselines: LMS Path Fitness Comparison. The Hybrid PSO-ACO-ABC framework achieves a +20.69% improvement over the baseline. Notably, when rigorously hyperparameter-tuned as grade regressors, both LSTM (+4.24%) and GRU (+11.88%) significantly outperform the base Ridge predictor, with GRU even surpassing ACO Alone. Yet, they still fall significantly short of the fully integrated Hybrid Swarm method. This illustrates the empirical advantages of combinatorial swarm optimization over greedy decoding from sequence-aware neural predictors, as analyzed in Section 5.3.

The bar chart highlights the progressive improvement from linear baselines through ensemble methods, deep learning models, and swarm intelligence approaches. The Hybrid framework's clear advantage over all baselines, including the LSTM neural network, shows that the integration of PSO-driven parameter optimization with ACO path construction and ABC refinement delivers meaningful improvements beyond what standard machine learning, deep learning, or single algorithm swarm approaches can achieve. These results support the hierarchical architecture design, where each swarm intelligence component contributes distinct capabilities that combine synergistically to enhance solution quality.

5.4 Statistical Validation

To assess the statistical significance of performance improvements, we conducted a rigorous evaluation protocol. The framework was evaluated on $\eta = 100$ randomly selected students with complete behavioral profiles from the full cohort of 8,783 learners. Each algorithm configuration was executed once per student with fixed random seeds to ensure reproducibility. Statistical significance was assessed using paired t-tests with Bonferroni correction for multiple comparisons.

Effect sizes were calculated using Cohen's d (reported in Table 2), with values of 0.2, 0.5, and 0.8 representing small, medium, and large effects, respectively. The Hybrid framework demonstrated statistically significant improvements over all baselines ($p < 0.001$). The largest performance gap was observed against the Ridge baseline (paired $d_z = 3.48$), indicating a massive

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 26, JUNE 2026

effect size. Furthermore, the Hybrid framework maintained a statistically significant lead over the rigorously tuned GRU predictor ($p < 0.001$). These results confirm that swarm intelligence-based optimization provides consistent and significant improvements across all comparison methods, suggesting that combinatorial search can yield higher fitness scores than greedy predictive extraction under the tested curriculum constraints.

6. Discussion & Limitations

The hybrid architecture combines PSO, ACO, and ABC to exploit their respective strengths: continuous optimization, combinatorial construction, and local refinement. This integration outperforms single-algorithm approaches, as well as both machine learning and deep learning baselines. When rigorously hyperparameter-tuned, the GRU architecture achieved a substantial +11.88% improvement over the linear baseline, proving that neural networks are highly capable estimators for student performance. However, this study reveals a critical structural bottleneck: because sequence-aware neural networks rely on greedy decoding to extract paths, their superior predictive capacity is wasted on myopic step-by-step decisions. They cannot mathematically optimize for delayed rewards or long-term prerequisite chains. The Hybrid Swarm Framework eliminates this bottleneck by optimizing the entire path simultaneously via pheromone-guided traversal, leading to a +20.69% improvement. Furthermore, constraint satisfaction by construction eliminates penalty function overhead while ensuring valid solutions. The framework enables practical deployment through asymmetric design: periodic offline PSO tuning with sub-second online recommendations.

PSO-optimized weights and ACO pheromone distributions offer interpretability not found in deep learning, while population-level recommendations address cold-start for new students.

The study's single-institution design (Narxoz University) may limit generalizability. While sufficient for tabular regression tasks, larger datasets might allow for more complex Transformer based architectures, though the fundamental prediction-versus-search disconnect would persist. The static curriculum model assumes fixed prerequisites, and the ABC implementation uses simplified neighborhood search rather than full bee colony phases. Furthermore, the deep learning baselines were limited to students with complete course sequences, whereas the swarm framework leveraged the full behavioral feature set. Future work should evaluate sequence models trained on the full cohort with beam-search decoding to isolate the true contribution of combinatorial search.

7. Future Work

Extending validation across diverse institutional contexts, developing incremental optimization for dynamic curricula, and explicit Pareto analysis of competing objectives (difficulty, engagement, completion time) represent immediate research directions. Extending the framework to proactive multi-agent LMS architectures enables real-time swarm intelligence-based adaptive interventions.

8. Conclusion

This research introduces a Hybrid Swarm Intelligence Framework, integrating Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO), and Artificial Bee Colony (ABC) algorithms, to address the multi-objective optimization problem of adaptive learning path recommendation. The proposed architecture systematically delegates algorithmic responsibilities according to the inherent structural strengths of each metaheuristic: PSO governs continuous parameter tuning, ACO manages discrete prerequisite-constrained path construction, and ABC facilitates local neighborhood refinement. Empirical validation conducted on a comprehensive dataset of 8,783 learner profiles establishes the statistical superiority of the hybrid approach over established baselines ($p < 0.001$, paired Cohen's $d_z = 3.48$ relative to Ridge Regression).

Furthermore, comparative analysis against rigorously hyperparameter-tuned sequence models, specifically Long Short-Term Memory (LSTM, +4.24%) and Gated Recurrent Unit (GRU, +11.88%) networks, illuminates a critical theoretical bottleneck in sequential recommendation. While these neural architectures demonstrate high predictive validity for isolated course outcomes, their reliance on greedy decoding mechanisms limits their capacity for global sequence optimization. Consequently, they fail to mathematically account for delayed rewards inherent in curriculum traversal. The empirical results confirm that the globally optimized search mechanisms provided by the fully integrated swarm framework yield significantly higher mean path fitness scores than those achievable through greedy extraction from predictive sequence models.

The primary contributions of this work encompass the hierarchical multi-algorithm architecture, the formalization of a graph-theoretic curriculum model, and the deployment of a 12-dimensional behavioral feature space. By decoupling predictive capacity from combinatorial search, this framework provides a highly scalable and interpretable methodology for educational path recommendation, mitigating the opacity commonly associated with deep learning models while ensuring adherence to strict prerequisite constraints.

Acknowledgments

This research was supported by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. IRN AP26196023). The authors thank Narxoz University for providing institutional data access and infrastructure support.

Data Availability

The data that support the findings of this study are available from Narxoz University but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of Narxoz University.

Ethics Statement

This study was conducted in accordance with the ethical guidelines of Narxoz University. The research used fully anonymized behavioral data extracted from the institutional Canvas learning management system for the period 2023-2025. All student identifiers were replaced with random anonymous codes prior to analysis, ensuring no possibility of re-identification. The study was reviewed and approved by the Narxoz University Research Ethics Committee. Informed consent 16 was obtained from all participants as part of the university's data collection policy for educational improvement purposes. The research complies with the principles of the Declaration of Helsinki and applicable data protection regulations

Funding

This research has been funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. IRN AP26196023, "Application of hybrid swarm intelligence algorithms in the development of proactive multi-agent systems for the digital educational environment").

References

[1] Issah, I., Appiah, O., Appiahene, P., & Inusah, F. (2023). A systematic review of the literature on machine learning application of determining the attributes influencing academic performance. *Decision Analytics Journal, 7*, 100204. https://doi.org/10.1016/j.dajour.2023.100204

[2] Freitas, D., Lopes, L. G., & Morgado-Dias, F. (2020). Particle Swarm Optimisation: A Historical Review Up to the Current Developments. *Entropy, 22*(3), 362. https://doi.org/10.3390/e22030362

[3] Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 10*(3), e1355. https://doi.org/10.1002/widm.1355

[4] Hashemi, A., & Dowlatshahi, M. B. (2024). Exploring Ant Colony Optimization for Feature Selection: A Comprehensive Review. In N. Dey (Ed.), *Applications of Ant Colony Optimization and its Variants* (pp. 45–72). Springer. https://doi.org/10.1007/978-981-99-7227-2_3

[5] Li, S., Chen, H., Liu, X., Li, J., Peng, K., & Wang, Z. (2023). Online Personalized Learning Path Recommendation Based on Saltatory Evolution Ant Colony Optimization Algorithm. *Mathematics, 11*(13), 2792. https://doi.org/10.3390/math11132792

[6] Dorigo, M., & Stützle, T. (2019). Ant Colony Optimization: Overview and Recent Advances. In M. Gendreau & J.-Y. Potvin (Eds.), *Handbook of Metaheuristics*. International Series in Operations Research & Management Science, Vol. 272. Springer, Cham. https://doi.org/10.1007/978-3-319-91086-4_10

[7] Shemshack, A., & Spector, J. M. (2020). A systematic literature review of personalized learning terms. *Smart Learning Environments, 7*, Article 33. https://doi.org/10.1186/s40561-020-00140-9

[8] Zhao, J., Mao, H., Mao, P., & Hao, J. (2024). Learning path planning methods based on learning path variability and ant colony optimization. *Systems and Soft Computing, 6*, 200091. https://doi.org/10.1016/j.sasc.2024.200091

[9] Amin, S., Uddin, M. I., Alarood, A. A., Mashwani, W. K., Alzahrani, A., & Alzahrani, A. O. (2023). Smart e-learning framework for personalized adaptive learning and sequential path recommendations using reinforcement learning. *IEEE Access, 11*, 89769–89790. https://doi.org/10.1109/ACCESS.2023.3305584

[10] Chauhan, R., & Mishra, A. (2025). SPAT: An intelligent educational system for student performance prediction using learning analytics and visualization. *Intelligent Data Analysis*. https://doi.org/10.1177/1088467X251377969

[11] Khairy, D., Alharbi, N., Amasha, M. A., Areed, M. F., Alkhalaf, S., & Abougalala, R. A. (2024). Prediction of student exam performance using data mining classification algorithms. *Education and Information Technologies, 29*(16), 21621–21645. https://doi.org/10.1007/s10639-024-12619-w

[12] Chanaa, A., & El Faddouli, N. (2024). Prerequisites-based course recommendation: Recommending learning objects using concept prerequisites and metadata matching. *Smart Learning Environments, 11*, 16. https://doi.org/10.1186/s40561-024-00301-0

[13] Zheng, Y., Wang, D., Xu, Y., & Li, Y. (2025). A multigranularity learning path recommendation framework based on knowledge graph and improved ant colony optimization algorithm for e-learning. *IEEE Transactions on Computational Social Systems, 12*(2), 586–607. https://doi.org/10.1109/TCSS.2024.3488373

[14] Fan, Z., Gou, J., & Wang, C. (2023). Predicting secondary school student performance using a double particle swarm optimization-based categorical boosting model. *Engineering Applications of Artificial Intelligence, 124*, 106649. https://doi.org/10.1016/j.engappai.2023.106649

[15] Zhu, D., Li, R., Zheng, Y., et al. (2025). Cumulative Major Advances in Particle Swarm Optimization from 2018 to the Present: Variants, Analysis and Applications. *Archives of Computational Methods in Engineering, 32*, 1571–1595. https://doi.org/10.1007/s11831-024-10185-5

[16] Ibrahim, A. O., Elfadel, E. M. E., Hashem, I. A. T., et al. (2025). The Artificial Bee Colony Algorithm: A Comprehensive Survey of Variants, Modifications, Applications, Developments, and Opportunities. *Archives of Computational Methods in Engineering, 32*, 3499–3533. https://doi.org/10.1007/s11831-025-10269-w

[17] Nguyen, B. H., Xue, B., & Zhang, M. (2020). A survey on swarm intelligence approaches to feature selection in data mining. *Swarm and Evolutionary Computation, 54*, 100663. https://doi.org/10.1016/j.swevo.2020.100663

[18] Yun, Y., Dai, H., An, R., Zhang, Y., & Shang, X. (2024). Doubly constrained offline reinforcement learning for learning path recommendation. *Knowledge-Based Systems, 284*, 111242. https://doi.org/10.1016/j.knosys.2023.111242

[19] Dokeroglu, T., Deniz, A., & Kiziloz, H. E. (2022). A comprehensive survey on recent metaheuristics for feature selection. *Neurocomputing, 494*, 269–296. https://doi.org/10.1016/j.neucom.2022.04.083

[20] Kaya, E., Gorkemli, B., Akay, B., & Karaboga, D. (2022). A review on the studies employing artificial bee colony algorithm to solve combinatorial optimization problems. *Engineering Applications of Artificial Intelligence, 115*, 105311. https://doi.org/10.1016/j.engappai.2022.105311