STATE-CONDITIONED PEDAGOGICAL ADAPTIVE ROUTING FOR EXERCISE RECOMMENDATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Exercise recommendation is inherently non-stationary: learner representations are unreliable under sparse interaction histories, the set of pedagogically appropriate exercises changes with knowledge mastery, and the desired balance among consolidation, exploration, and challenge evolves throughout the learning process. However, existing methods often optimize multiple objectives using globally fixed trade-offs and handle representation learning, candidate retrieval, and ranking as separate components, limiting their ability to adapt to both cold-start uncertainty and changing cognitive states. To address these limitations, we propose **SPAR**, a state-conditioned pedagogical adaptive routing framework that treats exercise recommendation as three interdependent decisions. First, a prototype-guided alignment module regularizes sparse learner representations using behavioral priors derived from similar active learners. Second, a ZPD (Zone of Proximal Development)-inspired retrieval module adapts the target challenge level according to recent performance and constructs a personalized candidate set within an appropriate difficulty range. Third, a state-conditioned dynamic routing network dynamically coordinates relevance, diversity, and difficulty within the retrieved candidate space according to the learner's current cognitive state. Experiments on three public datasets show that SPAR consistently improves ranking performance over strong exercise-recommendation baselines, with particularly pronounced gains for learners with sparse interaction histories.

1 INTRODUCTION

Online educational platforms have accumulated large-scale records of learners' exercise-taking behaviors, enabling data-driven Exercise Recommendation (ER) to provide individualized learning support. Traditional ER methods primarily rely on knowledge tracing and cognitive diagnosis to estimate learners' knowledge mastery and recommend exercises they are likely to answer correctly Su et al. (2018); Wang et al. (2020). Recent studies have further considered pedagogical factors such as relevance, difficulty, diversity, and personalized learning progression Zhou et al. (2026). These developments reflect an important shift in ER: effective recommendations should not only predict learner–exercise interactions, but also support knowledge consolidation, appropriate challenge, and concept exploration.Liu et al. (2025)

Despite this progress, existing ER methods remain limited in adapting to learners' evolving states and providing individualized instruction. First, educational interaction data typically exhibit a long-tailed distribution, leaving many learners with only sparse histories Fan et al. (2021); Yang et al. (2023). Under insufficient behavioral evidence, their representations may become dominated by population-level patterns, making distinct learners less distinguishable and resulting in generic recommendations Yun et al. (2022); Qiu et al. (2022), as illustrated in Figure 1(a). Second, candidate exercises are often retrieved using globally fixed difficulty thresholds. The range of pedagogically appropriate exercises should vary with learners' current mastery and recent performance. This is consistent with the Zone of Proximal Development (ZPD), which emphasizes tasks that are challenging yet achievable Shabani et al. (2010). Third, existing multi-objective methods commonly adopt fixed scalarization weights or shared trade-offs Lin et al. (2019); Huang et al. (2019), overlooking that the relative importance of relevance, difficulty, and diversity varies both across learners and throughout the learning process.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

As shown in Figure 1(b), struggling learners may require focused consolidation, whereas exploratory learners may benefit from broader coverage and increased difficulty. These limitations are closely coupled: unreliable representations impair mastery estimation, inaccurate mastery estimates lead to inappropriate candidate sets, and downstream re-ranking cannot recover suitable exercises that have already been excluded.

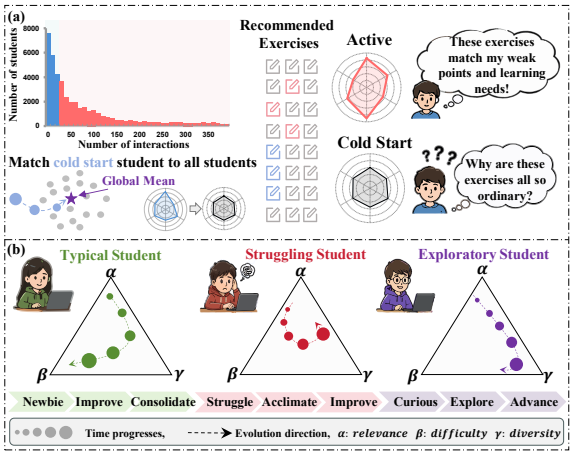


Figure 1: Motivation of the proposed framework. (a) The long-tailed distribution of learner interactions and representation homogenization under sparse histories, which may lead to generic recommendations. (b) Learner-specific trajectories of pedagogical priorities over relevance, difficulty, and diversity as learning progresses.

To address these challenges, we propose **SPAR**, a State-conditioned Pedagogical Adaptive Routing framework for exercise recommendation. SPAR formulates ER as a hierarchical state-conditioned decision process consisting of three complementary stages. At the macro level, a prototype-guided alignment module regularizes sparse learner representations using local behavioral priors derived from similar active learners, rather than indiscriminately transferring information from the entire population. At the meso level, a ZPD-inspired retrieval module dynamically adapts the target challenge level according to the learner's concept mastery and recent performance, thereby constructing a state-appropriate candidate space. At the micro level, a state-conditioned dynamic routing network coordinates relevance, difficulty, and diversity within the retrieved candidate set according to the learner's current cognitive state. Through this hierarchy, SPAR progressively integrates the three stages together and determines *who the learner is under limited observations*, *which exercises are currently appropriate*, and *which pedagogical objective should be prioritized*.

- We formulate exercise recommendation as a unified hierarchical, state-conditioned decision process that treats learner representation under sparse histories, pedagogically appropriate candidate construction, and adaptive multi-objective coordination as interdependent decisions.
- We develop a prototype-guided alignment mechanism for sparse-history learners and a ZPD-inspired retrieval strategy that dynamically adapts the target challenge level to construct personalized candidate sets.
- We design a state-conditioned multi-objective re-ranking mechanism that dynamically coordinates relevance, difficulty, and diversity according to learners' evolving cognitive states, achieving consistent improvements across three public datasets.

2 Related Work

Exercise Recommendation (ER). Exercise recommendation aims to select exercises that match learners' knowledge states and support their learning progress Cheng et al. (2026). Early methods mainly incorporated knowledge tracing and cognitive diagnosis to estimate learner mastery and improve recommendation accuracy. For example, LSTMCQP Huo et al. (2020) combines contextual representations with LSTM-based knowledge tracing, while ExRec Özyurt et al. (2025) introduces semantically grounded knowledge tracking for personalized recommendation. Beyond mastery estimation, subsequent studies have considered broader pedagogical factors, including exercise difficulty, concept coverage, and personalized learning preferences, as explored by KCPER Wu et al. (2020) and Wang et al. Wang et al. (2025). Similarly, DLPR Zhang et al. (2024) explicitly models item difficulty for adaptive educational sequencing. Recent work has investigated explainability and sparse-history recommendation. KG4EER Guan et al. (2025) and KG4Ex Guan et al. (2023) leverage knowledge graphs to model learner–concept–exercise relations, while MMER Liu et al. (2023a) uses meta-learning to improve recommendation. However, they remain limited in handling representation uncertainty and changing learning needs within a unified framework.

Multi-objective Recommendation. Multi-objective recommendation commonly balances competing goals through weighted scalarization Zheng & Wang (2021), multi-task learning, or Pareto-efficient optimizationLin et al. (2022). Representative methods include MMoE Ma et al. (2018), PE-LTR Lin et al. (2019), and PAPERec Xie et al. (2021). These methods provide general mechanisms for coordinating multiple objectives, but they are not specifically designed to model learners' evolving cognitive states or pedagogically feasible exercise spaces. Multi-objective optimization has also been introduced into exercise recommendationVernerey et al. (2026). Sun et al. Sun et al. (2024b) use an evolutionary framework to assemble exercise groups according to learners' knowledge proficiency. Huang et al. Huang et al. (2019) optimize multiple objectives such as concept exploration, engagement, and difficulty smoothness through scalarization. NR4DER Cheng et al. (2025) combines sequence enhancement with neural re-ranking to balance difficulty matching and recommendation diversity. However, these methods generally rely on fixed objective combinations or globally shared optimization policies, and thus cannot explicitly adapt pedagogical priorities to each learner's changing state.

More importantly, existing studies treat learner representation under sparse histories, candidate construction, and re-ranking as separate problems. Consequently, they overlook the cascading dependencies among these decisions: representation uncertainty affects the estimation of a learner's cognitive boundary, candidate construction determines the feasible space for downstream ranking, and pedagogical objectives must be coordinated within that space according to the learner's state.

In contrast, SPAR formulates exercise recommendation as a hierarchical state-conditioned decision process. It first improves sparse learner representations through prototype-guided alignment, then constructs a personalized candidate space through ZPD-inspired challenge adaptation, and finally coordinates relevance, difficulty, and diversity through state-conditioned routing. This design jointly models representation, retrieval, and re-ranking as coupled decisions rather than treating them as independent optimization stages.

3 METHOD

3.1 PROBLEM DEFINITION

ER is inherently state-dependent, as the desired balance among pedagogical objectives varies with the learner's evolving cognitive state. We therefore formulate ER as a hierarchical state-conditioned decision problem:

$$\pi^*(u,t) = \arg\max_{\pi} \mathcal{F}\left(\mathrm{Rel}, \mathrm{Diff}, \mathrm{Div} \mid \mathbf{s}_u^{(t)}\right), \tag{1}$$

where $\mathbf{s}_u^{(t)}$ denotes learner u's cognitive state at time step t, and Rel, Diff, and Div represent relevance, difficulty, and diversity, respectively. Let $\mathcal{U}$, $\mathcal{E}$, and $\mathcal{C}$ denote the sets of learners, exercises, and knowledge concepts, respectively. The interaction history of learner u up to time t is represented as $\mathcal{H}_u^{(t)} = \{(e_i, a_i, \mathcal{Q}(e_i), \tau_i)\}_{i=1}^{t}$, where $e_i \in \mathcal{E}$ is the exercised item, $a_i \in \{0,1\}$ indicates whether the response is correct, $\mathcal{Q}(e_i) \subseteq \mathcal{C}$ denotes the concepts associated with e_i, and τ_i is the interaction timestamp. Given $\mathcal{H}_u^{(t)}$, the goal is to construct a personalized candidate set and rank its exercises for the next interaction according to the learner's state.

3.2 OVERVIEW

Figure 2 illustrates the overall architecture of SPAR, which organizes exercise recommendation into a hierarchical Macro–Meso–Micro decision process. Given a learner's historical interaction sequence, the macro-level prototype alignment module first incorporates behavioral priors from active learners to improve sparse learner representations (Stage 1). Based on the resulting representation and mastery state, the meso-level ZPD-inspired retrieval module dynamically estimates a personalized target challenge level and constructs a state-appropriate candidate set (Stage 2). Finally, the micro-level state-conditioned re-ranking module dynamically coordinates relevance, difficulty, and diversity according to the learner's current cognitive state, producing the final Top-K recommendation list (Stage 3). Together, these stages progressively reduce representation uncertainty, constrain the pedagogically feasible exercise space, and adapt the ranking policy to the learner's evolving state.

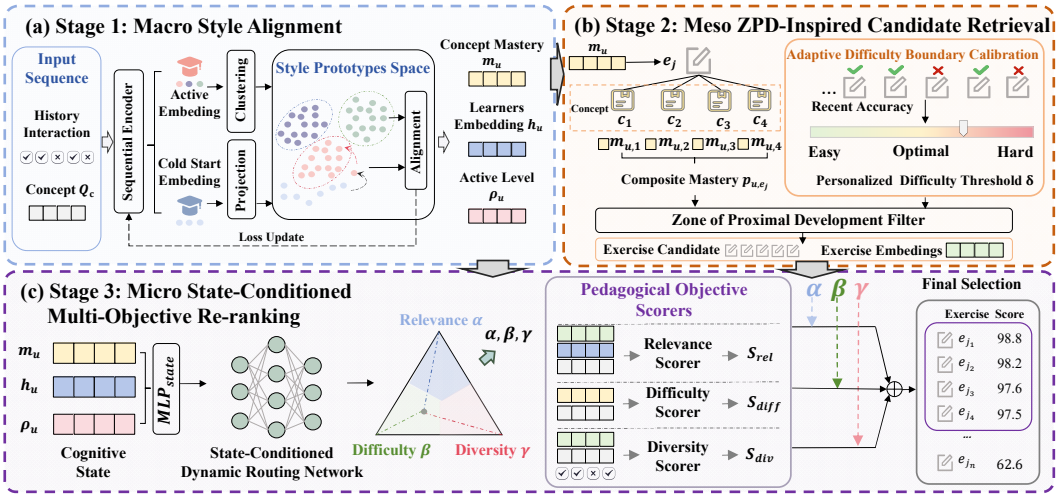


Figure 2: Framework and workflow of our SPAR method.

3.3 Macro Style Alignment for Learner Representation

To effectively model learners with sparse interaction histories, we propose a macro style alignment module that leverages style prototype clustering and dynamic style prototype alignment to enhance cold-start representations while preserving their personalized intrinsic styles.

Style Prototype Clustering. To capture the diverse macro learning styles, we first encode learners' historical interaction sequences via a sequential encoder, specifically implemented as a mLSTM. We regard learners with sequence lengths $|X_u| \geq \tau$ as active learners, whose encoded representations h_u are robust and semantically rich. Building upon this, style prototypes are updated online via mini-batch K-means clustering in the active representation space H_{active}. During each forward pass, prototypes are recomputed from the detached representations of active learners. This process identifies K distinct style prototypes, segmenting learners into different cognitive archetypes Chen et al. (2022). The objective is to discover a set of cluster centroids $Z = \{z_1, z_2, \ldots, z_K\}$ that minimizes the intra-cluster variance.

Intra-cluster Style Alignment. For a cold-start learner u with a sparse interaction history ($|X_u| < \tau$), we first map this sparse representation into the active semantic space via a learnable linear transformation matrix W_A, yielding $\tilde{h}_u = W_A h_u$. Instead of blindly pulling $\tilde{h}_u$ toward a generic global mean, we match the cold-start learner to their nearest style prototype by calculating the Euclidean distance in the semantic space: $k^* = \arg\min_{k \in \{1,2,\ldots,K\}} \|\tilde{h}_j - z_k\|$. Once the most analogous learning style, k^*, is identified, we perform dynamic alignment of style prototypes. We use the local semantic structure formed by active learners as a style-aware prior to regularize sparse representations. Specifically, the cold-start learner's representation is encouraged to move toward the semantic region occupied by active learners within the matched cluster Guo et al. (2021). The alignment loss is formulated as the weighted mean squared error:

$$\mathcal{L}_{align} = \frac{1}{M} \sum_{u=1}^{M} w_j \cdot \|\tilde{h}_u - \mu_{k^*}\|^2, \tag{2}$$

where M is the number of cold-start learners in the current batch, and μ_{k^*} is the localized average representation of the active learners assigned to the matched cluster k^*. To ensure the alignment process is adaptive to both the training stage and the specific sparsity level of the learner, we introduce a dynamic scaling weight w_j:

$$w_j = \frac{epoch}{E_{max}} \left(1 - \frac{|X_u|}{L_{max}}\right), \tag{3}$$

where $epoch$ is the current training epoch, E_{max} is the maximum number of epochs, and L_{max} is the predefined maximum sequence length. After optimization, the model generates aligned learner representations h_u. Based on the output state of the base encoder, the concept mastery vector $\mathbf{m}_u$,

where $m_{u,c}$ denotes the mastery of concept c, is obtained through a linear projection followed by a Sigmoid activation.

3.4 MESO ZPD-INSPIRED CANDIDATE RETRIEVAL

We design a meso ZPD-inspired candidate retrieval module to efficiently downsample the item space into a coarse-grained candidate pool. Crucially, rather than employing a static filtering threshold, this module dynamically calibrates the retrieval boundary based on ZPD theory, ensuring the retrieved exercises are neither overly easy nor difficult.

To accurately evaluate a learner's current learning status, we continuously track their cognitive boundary by evaluating their recent performance in interactions. Given the learner's response sequence, we calculate the recent accuracy $acc_{recent}^{(u)}$ over a sliding window of size W.

Adaptive Difficulty Boundary Calibration. According to the ZPD theory, optimal learning occurs when tasks are slightly beyond a learner's current level of independent mastery but are achievable with guidance. To approximate each learner's feasible cognitive boundary under the ZPD principle, we estimate a personalized target difficulty region based on recent learning performance:

$$\delta_u = \text{Clip}\Big(\delta_{base} - \alpha \cdot acc_{recent}^{(u)},\ \delta_{min},\ \delta_{max}\Big), \tag{4}$$

where δ_{base} is a base threshold, α is a scaling factor adjusting the sensitivity to recent performance, and $[\delta_{min}, \delta_{max}]$ is the allowable threshold range. If a learner demonstrates high recent accuracy, δ_u decreases. A lower target mastery level implies selecting more challenging exercises, pushing the learner out of their comfort zone to explore new knowledge.

ZPD-Driven Candidate Filtering. With the personalized threshold δ_u established, the module scans the exercise pool to retrieve exercises that match this cognitive target. For a candidate exercise e_j, the learner's composite mastery p_{u,e_j} on this exercise is approximated by the product of their mastery over the required concepts Junker & Sijtsma (2001):

$$p_{u,e_j} = \exp\left(\frac{1}{|\mathcal{Q}(e_j)|}\sum_{c\in\mathcal{Q}(e_j)}\log(m_{u,c}+\epsilon)\right), \tag{5}$$

where $\mathrm{m}_{u,c}$ is the learner's estimated mastery state for concept c, derived from the base sequence encoder, ϵ is a sufficiently small constant added to prevent mathematical instability, and $\exp(\cdot)$ denotes the exponential function. We then compute the absolute distance between the learner's predicted proficiency and their personalized ZPD threshold:

$$r_{u,e_j} = \left|\delta_u - p_{u,e_j}\right|. \tag{6}$$

Finally, we select the top N exercises with the minimal r_{u,e_j} to construct the candidate pool E_c.

3.5 MICRO STATE-CONDITIONED MULTI-OBJECTIVE RE-RANKING

To balance the demands of multiple teaching objectives, we propose a micro state-conditioned multi-objective re-ranking that gauges the learner u's cognitive state and uses a dynamic routing network to adaptively harmonize relevance, diversity, and difficulty for each candidate exercise.

Dynamic Cognitive State Modeling. A learner's learning state is highly dynamic and multi-faceted Ying et al. (2018). To comprehensively characterize the learner's current learning state, we construct a unified state representation s_u by integrating three complementary signals: knowledge mastery, recent learning dynamics, and activity level. Specifically, we obtain the concept mastery m_u and the learner embedding h_u from stage 1. To make the model aware of the sparsity of the learner's interactions, we compute the active level $\rho_u = \frac{|X_u|}{L_{max}}$ Zhou et al. (2020). The composite cognitive state vector s_u is generated via a multi-layer perceptron:

$$s_u = \text{MLP}_{state}([m_u \parallel h_u \parallel \rho_u]), \tag{7}$$

where $\parallel$ denotes concatenation. This decomposition allows the state encoder to jointly capture stable knowledge proficiency and transient learning dynamics within a unified cognitive state space.

Pedagogical Objective Scorers. With the candidate pool E_c retrieved, we evaluate each candidate exercise $e_j \in E_c$ using three parallel objective scorers, including relevance scorer (S_{rel}), diversity scorer (S_{div}), and difficulty scorer (S_{diff}) . The relevance scorer aims to consolidate previously learned concepts. We concatenate the learner embedding h_u, the candidate exercise embedding $\mathbf{e}_j$, and the exercise's concept multi-hot encoding q_j:

$$S_{rel}(u, e_j) = \text{MLP}_{rel}(h_u \parallel \mathbf{e}_j \parallel q_j), \tag{8}$$

To prevent cognitive fixation and encourage the exploration of new knowledge boundaries, we model the personalized diversity gain. Specifically, we first organize and encode the learners' historical exercises into concept-wise interaction sequences. A multi-head self-attention mechanism is then applied to capture the learner's personalized concept preferences, yielding a concept attention vector $\theta \in \mathbb{R}^{|\mathcal{Q}|}$.To quantify the diversity contribution of an exercise e_j, we compute its marginal gain in knowledge coverage. Let R denote the set of concepts covered by the learner's historical exercises. The concept coverage is defined as Cheng et al. (2025):

$$Coverage(R) = 1 - \prod_{e_j \in R} (1 - q_j), \tag{9}$$

The diversity gain of a exercise e_j is the incremental coverage after adding it to the historical set:

$$\Delta R_j = Coverage(R \cup \{e_j\}) - Coverage(R). \tag{10}$$

The attention weights and the coverage gain are combined to generate the diversity score:

$$S_{div}(u, e_j) = \text{MLP}_{div}(\theta \odot \Delta R_j), \tag{11}$$

where $\odot$ denotes element-wise multiplication. To ensure the exercise meticulously aligns with the learner's micro-level ZPD, we introduce a dedicated difficulty scorer. For a candidate e_j containing a set of concepts $\mathcal{Q}(e_j)$, we extract the learner's average mastery over these specific concepts: $\bar{m}_{u,j} = \frac{\sum_{c \in \mathcal{Q}(e_j)} m_{u,c}}{|\mathcal{Q}(e_j)|}$ from Stage 2. We then concatenate the q_j with this localized mastery scalar:

$$S_{diff}(u, e_j) = \text{MLP}_{diff}([q_j \parallel \bar{m}_{u,j}]), \tag{12}$$

where $\text{MLP}_{rel}, \text{MLP}_{div}, \text{MLP}_{diff}$ are uniformly implemented as two-layer feed-forward networks with ReLU activation and a Sigmoid.

State-Conditioned Dynamic Routing Network. Since a learner's optimal learning strategy shifts over time, we use a state-conditioned dynamic routing network to address this. Taking the real-time cognitive state s_u as input, the network dynamically allocates trade-off weights $[\alpha, \beta, \gamma]$ corresponding to relevance, difficulty, and diversity, respectively:

$$[\alpha, \beta, \gamma] = \text{Softmax}(\text{MLP}_{gate}(s_u)), \tag{13}$$

where MLP_{gate} is a non-linear multi-layer perceptron mapping the d_{state} dimensional cognitive state into the 3-dimensional routing space. The Softmax activation ensures that $\alpha + \beta + \gamma = 1$, enabling an adaptive routing mechanism that dynamically balances multiple educational objectives conditioned on the learner's cognitive state. Finally, we perform a dynamic scalarization to fuse the multi-objective scores. The ranking score for a candidate exercise e_j is the state-weighted linear combination of the three objectives:

$$Score(u, e_j) = \alpha \cdot S_{rel}(u, e_j) + \beta \cdot S_{diff}(u, e_j) \quad + \gamma \cdot S_{div}(u, e_j). \tag{14}$$

The candidate pool E_c is sorted in descending order based on this score to produce the Top-N recommendation list.

3.6 Optimization & Training

We formulate the primary recommendation task as a binary classification problem. We adopt the binary cross-entropy loss $\mathcal{L}_{BCE}$ to penalize discrepancies between the predicted final score and the ground-truth learner response labels Piech et al. (2015). To prevent objective dominance and encourage a balanced, diverse weight distribution, we introduce an entropy regularization term $\mathcal{H}_r$:

$$\mathcal{H}_r = -\frac{1}{B} \sum_{i=1}^{B} \left(\alpha_i \log \alpha_i + \beta_i \log \beta_i + \gamma_i \log \gamma_i \right), \tag{15}$$

where B is the batch size. The joint loss function for the overall framework is thus defined as:

$$\mathcal{L}_{total} = \mathcal{L}_{BCE} - \lambda_{ent}\mathcal{H}_r, \tag{16}$$

where λ_{ent} is a tunable hyperparameter controlling the strength of the entropy regularization. Maximizing this entropy ensures the dynamic weights remain non-degenerate throughout the training process Shazeer et al. (2017). $\mathcal{L}_{align}$ is exclusively used to optimize the prototype alignment.

4 EXPERIMENT

To evaluate the performance of our proposed framework, we conduct extensive experiments to address the following Research Questions (RQs):

RQ1 (Superiority) Can SPAR achieve better prediction performance than state-of-the-art ER models?

RQ2 (Effectiveness) How does each core component of SPAR contribute to its overall effectiveness?

RQ3 (Robustness) Can SPAR effectively alleviate the cold-start problem for learners with sparse interaction histories?

RQ4 (Adaptability) Does SPAR learn meaningful pedagogical trajectories dynamically conditioned on learners' cognitive states?

4.1 EXPERIMENTAL SETUP

Datasets and Baselines. We investigated three public datasets, including Assist2009, Assist2012 Feng et al. (2009), and XES3G5M Liu et al. (2023d), where the statistics of datasets are in Table 1. To evaluate our model's performance, we compared it against 8 state-of-the-art methods, including SimpleKTRec Liu et al. (2023c), MIKTRec Sun et al. (2024a), KSAP Wang et al. (2025), KCPER Wu et al. (2020), MMER Liu et al. (2023a), KG4Ex Guan et al. (2023), FAKTER Ji et al. (2025), and NR4DER Cheng et al. (2025).

Table 1: Dataset statistics.

Dataset	learner	Question	Concept	Interaction
XES3G5M	18,066	7,652	865	5,549,184
Assist2009	4,217	17,737	123	374,422
Assist2012	29,018	53,091	265	2,711,813

Evaluation Metrics. To evaluate the performance of the recommendations, We adopt four metrics in our experiments: normalized discounted cumulative gain (NDCG), Recall, list-coverage F1 (LC-F1), and diversity (DIV), where LC-F1 evaluates the coverage of historically weak concepts by the recommended concept list after removing duplicates Liu et al. (2023b). Higher values of these metrics indicate better recommendation quality. In addition, we compute Metrics@K (K = 1, 3, 5, 10) based on the lengths of the exercise lists to assess the effectiveness of recommendations across different list sizes.

Implementation Details. Hyperparameters are empirically selected based on validation performance: $\delta_{base} = 0.6$, $\alpha = 0.4$, $\delta_{min} = 0.2$, $\delta_{max} = 0.8$, window size $W = 20$, $K = 5$, and $\lambda_{ent} = 0.01$. To ensure a fair comparison, the same number of candidate exercises ($N = 150$) were evaluated for methods that used retrieval and re-ranking. The baseline models construct their candidate pools using their respective original retrieval strategies, whereas SPAR uses our own meso ZPD-inspired retrieval strategy before re-ranking. We train the ranking score using weak concept-level supervision. A candidate is labeled positive if it covers at least one previously incorrect concept, and negative if all its required concepts were answered correctly. Candidates involving unseen concepts are excluded from the BCE loss.

4.2 PERFORMANCE COMPARISON (RQ1)

To assess the effectiveness of our proposed framework, we conduct a thorough comparative analysis across the three datasets. Table 2 presents the experimental results of our model compared with eight state-of-the-art baselines across various top-K recommendation metrics. Our model substantially outperforms the baseline methods on most evaluation metrics. In particular, with respect to ranking quality, our approach achieves notable gains in both NDCG and Recall. Nonetheless, it attains comparatively lower performance improvement on the LC-F1 metric. This is because LC-F1 primarily

Table 2: Performance comparison of different models on three datasets. Best results are in bold. All performance gains of SPAR over the strongest baseline are statistically significant under a two-sided paired t-test ($p < 0.05$).

Dataset	Method	NDCG				LC-F1				Recall			
		@1	@3	@5	@10	@1	@3	@5	@10	@1	@3	@5	@10
XES3G5M	SimpleKTRec	58.41	61.29	63.39	64.69	74.49	71.30	62.08	50.02	41.19	57.20	62.93	73.55
	MIKTRec	57.33	62.35	63.38	68.52	75.52	72.29	62.87	51.05	42.28	56.17	62.31	73.44
	KSAP	49.11	48.25	60.03	61.18	72.61	67.36	61.55	47.57	33.61	51.84	55.95	67.31
	KCPer	33.89	44.58	50.92	54.33	31.83	45.86	52.19	44.21	36.83	41.16	44.98	47.59
	KG4Ex	48.63	69.36	73.10	73.24	66.87	65.93	60.28	49.93	42.01	53.79	49.79	51.53
	FAKTER	61.93	69.63	74.03	74.55	68.78	68.93	61.67	50.98	42.24	52.39	60.38	68.92
	MMER	47.23	63.37	66.61	68.62	48.96	65.98	65.42	51.02	44.25	46.47	45.79	47.48
	NR4DER	75.51	78.92	82.45	83.41	79.50	71.39	62.56	52.81	45.76	56.82	63.87	75.21
	SPAR	**78.89**	**84.92**	**86.27**	**88.29**	**80.12**	**72.53**	**66.08**	**54.00**	**45.98**	**57.91**	**66.32**	**78.89**
Assist2009	SimpleKTRec	54.39	66.67	70.20	71.34	54.21	46.98	43.56	38.55	31.19	36.74	41.11	54.32
	MIKTRec	56.33	67.98	72.92	73.16	56.28	47.81	45.57	40.54	36.32	37.15	45.33	56.11
	KSAP	53.28	66.98	70.11	71.51	53.22	46.53	42.81	37.42	32.13	35.79	41.12	53.21
	KCPer	21.61	29.12	31.13	29.71	21.60	28.29	30.45	32.04	21.63	24.15	25.13	26.10
	KG4Ex	25.09	34.91	37.41	32.42	25.19	29.55	30.06	30.42	24.94	25.01	24.04	25.16
	FAKTER	55.57	69.31	73.29	74.65	54.13	49.36	47.09	40.23	30.96	36.46	41.24	56.63
	MMER	32.73	45.92	50.23	52.88	32.72	50.09	49.93	49.22	32.72	33.24	33.34	34.68
	NR4DER	56.98	71.13	75.13	77.54	56.84	48.03	47.32	37.14	30.25	36.94	42.53	56.93
	SPAR	**77.59**	**85.00**	**85.07**	**86.53**	**56.98**	**61.23**	**59.83**	**54.13**	**42.66**	**69.44**	**77.17**	**80.94**
Assist2012	SimpleKTRec	38.55	52.10	54.26	55.14	38.34	39.60	35.39	26.35	17.36	25.44	30.95	38.56
	MIKTRec	39.67	53.19	55.67	56.53	39.41	39.98	36.91	27.85	19.63	29.87	33.94	40.55
	KSAP	33.64	49.71	51.96	53.17	33.97	37.65	33.17	24.21	16.31	23.65	28.47	33.65
	KCPer	14.36	28.47	34.61	41.15	14.23	26.53	30.83	33.42	14.32	21.22	23.93	25.46
	KG4Ex	35.59	37.52	48.19	59.24	35.95	36.88	37.98	41.53	19.74	26.55	32.39	35.53
	FAKTER	36.35	48.65	53.81	54.78	36.23	39.13	39.86	34.92	27.10	34.47	41.98	45.33
	MMER	31.58	50.66	56.24	60.33	31.35	39.93	41.79	43.44	30.74	31.24	31.47	31.56
	NR4DER	49.15	64.93	72.15	80.64	49.12	50.35	50.01	46.36	33.01	38.95	41.58	49.16
	SPAR	**78.10**	**79.59**	**80.80**	**83.04**	**49.29**	**53.11**	**50.63**	**47.27**	**38.75**	**57.55**	**62.61**	**65.32**

Table 3: Ablation study on three datasets.

Dataset	Method	NDCG				LC-F1				Recall			
		@1	@3	@5	@10	@1	@3	@5	@10	@1	@3	@5	@10
XES3G5M	w/o macro	76.98	83.39	85.10	87.09	79.23	71.52	64.98	52.98	44.36	56.71	64.87	77.36
	w/o meso	74.83	81.67	82.39	86.12	78.30	68.96	62.23	50.41	40.69	54.84	62.55	75.16
	w/o micro	76.34	83.09	84.81	86.62	78.91	71.13	64.55	52.34	43.98	56.19	64.26	76.89
Assist2009	w/o macro	76.66	84.28	83.69	83.24	50.86	60.65	59.62	54.06	43.02	68.94	76.87	80.87
	w/o meso	74.80	81.37	79.65	78.52	49.30	57.97	57.27	53.48	41.64	65.94	73.65	79.96
	w/o micro	68.89	67.65	77.40	75.02	45.97	56.29	57.22	53.73	39.03	64.50	74.13	80.47
Assist2012	w/o macro	77.48	79.22	80.27	82.32	47.40	52.65	49.31	42.21	38.35	57.12	62.32	65.26
	w/o meso	75.86	78.36	78.23	81.03	45.87	49.95	47.25	40.35	37.39	55.32	60.98	64.12
	w/o micro	77.06	78.56	79.76	81.98	47.84	52.06	49.60	42.27	37.73	56.50	61.59	64.32

emphasizes classification accuracy instead of ranking quality, whereas our framework is explicitly tailored to optimize list-wise ranking performance, pedagogical diversity, and difficulty adaptation. The sensitivity analysis identifies five style prototypes, $\lambda_{\text{ent}} = 0.01$, and a candidate pool size of 150 as the most effective settings, and SPAR maintains a competitive diversity score across all datasets, which are provided in the Appendix.

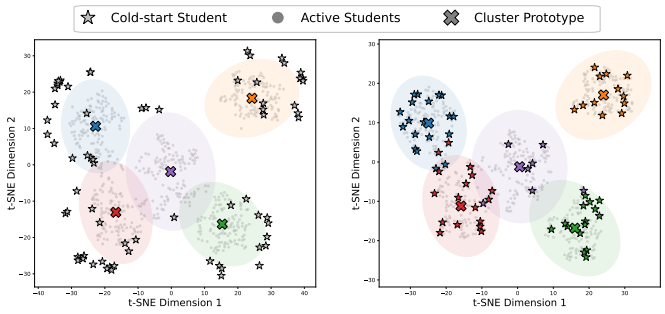
4.3 Ablation Study (RQ2)

To assess the contribution of each component, we construct three variants: w/o macro, which removes prototype-based style alignment; w/o meso, which replaces ZPD-inspired retrieval with global exercise retrieval; and **w/o micro**, which replaces state-conditioned routing with uniform weights ($\alpha = \beta = \gamma = \frac{1}{3}$). As shown in Table 3, removing any component degrades performance, demonstrating the complementary contributions of the three hierarchical stages. We also conduct fairness evaluation under different retrieval and re-ranking settings to prove our method's effectiveness in the Appendix. The fairness evaluation shows that both the ZPD-inspired retrieval and state-

conditioned re-ranking modules independently improve existing two-stage methods, while their combination achieves the best performance.

4.4 Cold-Start Analysis (RQ3)

We assess the effectiveness of macro style alignment for cold-start learners by dividing learners into cold-start (1–30 interactions) and active (>30 interactions) cohorts. The zero-interaction case is excluded because no behavioral interaction is available for modeling learner-specific preferences or states. Figure 3 visualizes their Stage-1 embeddings before and after alignment, with active learners providing the semantic background and cold-start learners highlighted. After alignment, cold-start learners move toward learning-style prototypes and form more coherent semantic regions, while the overall structure of active learners remains largely preserved. Quantitative results are reported in the Appendix, showing that SPAR maintains strong performance across sparsity levels, with particularly notable gains for extremely cold-start learners (interaction ≤ 5).



(a) Before macro alignment (b) After macro alignment

Figure 3: An example of the style alignment module's efficacy

4.5 Pedagogical Policy Evolution Analysis (RQ4)

To validate that the routing mechanism learns meaningful pedagogical adaptation, we visualize the temporal evolution of the routing weights $[\alpha, \beta, \gamma]$ on a ternary simplex with point colors from dark to bright indicating the progression of time. As illustrated in Figure 4, different learner types exhibit different pedagogical routing trajectories. For example, the typical learner's trajectory initially focuses on relevance, gradually shifts toward difficulty, and ultimately explores diversity. This continuous migration confirms that SPAR dynamically tailors pedagogical priorities to learners' evolving states rather than relying on static trade-offs.

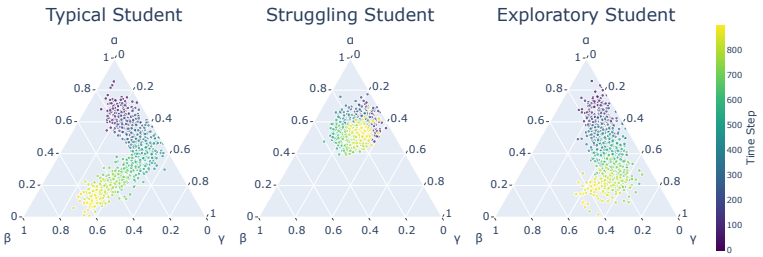


Figure 4: State-conditioned pedagogical policy trajectories for different learner archetypes.

5 Conclusion

In this paper, we propose SPAR, a state-conditioned adaptive routing framework for exercise recommendation. SPAR integrates prototype-guided macro alignment to alleviate cold-start sparsity, ZPD-inspired meso retrieval to calibrate difficulty boundaries, and micro-level dynamic routing to balance relevance, diversity, and difficulty conditioned on cognitive states. Extensive experiments across three public datasets demonstrate its consistent superiority over state-of-the-art baselines, especially under sparse interactions. Future work will explore online feedback modeling, explainability, and educational fairness.

REFERENCES

Yongjun Chen, Zhiwei Liu, Jia Li, Julian McAuley, and Caiming Xiong. Intent contrastive learning for sequential recommendation. In *Proceedings of the ACM web conference (WWW)*, pp. 2172–2182, 2022.

Xinghe Cheng, Xufang Zhou, Liangda Fang, Chaobo He, Yuyu Zhou, Weiqi Luo, Zhiguo Gong, and Quanlong Guan. Nr4der: Neural re-ranking for diversified exercise recommendation. In *Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, pp. 1738–1747, 2025.

Xinghe Cheng, Guiyong Zhuang, Yusheng Xie, Jiapu Wang, Yixin Liu, Quanlong Guan, Liangda Fang, and Shirui Pan. Unier: A unified benchmark for item-level and path-level exercise recommendation. *arXiv preprint arXiv:2605.16750*, 2026.

Ziwei Fan, Zhiwei Liu, Shen Wang, Lei Zheng, and Philip S Yu. Modeling sequences as distributions with uncertainty for sequential recommendation. In *Proceedings of the ACM international conference on information & knowledge management (CIKM)*, pp. 3019–3023, 2021.

M. Feng, N. T. Heffernan, and K. R. Koedinger. Addressing the assessment challenge in an intelligent tutoring system that tutors as it assesses. *User Modeling and User-Adapted Interaction*, 19: 243–266, 2009.

Neil D Fleming and Colleen Mills. Not another inventory, rather a catalyst for reflection. *To improve the academy*, 11(1):137–155, 1992.

Quanlong Guan, Fang Xiao, Xinghe Cheng, Liangda Fang, Ziliang Chen, Guanliang Chen, and Weiqi Luo. Kg4ex: An explainable knowledge graph-based approach for exercise recommendation. In *Proceedings of the ACM international conference on information and knowledge management (CIKM)*, pp. 597–607, 2023.

Quanlong Guan, Xinghe Cheng, Fang Xiao, Zhuzhou Li, Chaobo He, Liangda Fang, Guanliang Chen, Zhiguo Gong, and Weiqi Luo. Explainable exercise recommendation with knowledge graph. *Neural Networks*, 183:106954, 2025.

Dandan Guo, Long Tian, Minghe Zhang, Mingyuan Zhou, and Hongyuan Zha. Learning prototype-oriented set representations for meta-learning. *arXiv preprint arXiv:2110.09140*, 2021.

Zhenya Huang, Qi Liu, Chengxiang Zhai, Yu Yin, Enhong Chen, Weibo Gao, and Guoping Hu. Exploring multi-objective exercise recommendations in online education systems. In *Proceedings of the ACM International Conference on Information and Knowledge Management (CIKM)*, pp. 1261–1270, 2019.

Yujia Huo, Derek F Wong, Lionel M Ni, Lidia S Chao, and Jing Zhang. Knowledge modeling via contextualized representations for lstm-based personalized exercise recommendation. *Information Sciences*, 523:266–278, 2020.

Wei Ji, Zhiyu Chen, and Jing Xiao. Exercise recommendation based on feature-aligned knowledge tracing. In *Proceedings of the International Conference on Database Systems for Advanced Applications (DASFAA)*, pp. 168–184. Springer, 2025.

Brian W Junker and Klaas Sijtsma. Cognitive assessment models with few assumptions, and connections with nonparametric item response theory. *Applied Psychological Measurement*, 25(3): 258–272, 2001.

David A Kolb. *Experiential learning: Experience as the source of learning and development*. FT press, 2014.

Xi Lin, Zhiyuan Yang, Xiaoyuan Zhang, and Qingfu Zhang. Pareto set learning for expensive multi-objective optimization. *Advances in neural information processing systems*, 35:19231–19247, 2022.

Xiao Lin, Hongjie Chen, Changhua Pei, Fei Sun, Xuanji Xiao, Hanxiao Sun, Yongfeng Zhang, Wenwu Ou, and Peng Jiang. A pareto-efficient algorithm for multiple objective optimization in e-commerce recommendation. In *Proceedings of the ACM Conference on recommender systems (RecSys)*, pp. 20–28, 2019.

Fei Liu, Xuegang Hu, Shuochen Liu, Chenyang Bu, and Le Wu. Meta multi-agent exercise recommendation: A game application perspective. In *Proceedings of ACM SIGKDD conference on Knowledge Discovery and Data Mining (KDD)*, pp. 1441–1452, 2023a.

Guowei Liu, Meirui Ren, Longjiang Guo, Jin Li, and Miao Ma. Comprehensive exercise recommendation with practicality, generalizability, and versatility in ai-driven education. *Information Processing & Management*, 62(3):104051, 2025.

Weiwen Liu, Yunjia Xi, Jiarui Qin, Xinyi Dai, Ruiming Tang, Shuai Li, Weinan Zhang, and Rui Zhang. Personalized diversification for neural re-ranking in recommendation. In *Proceedings of the IEEE International Conference on Data Engineering (ICDE)*, pp. 802–815. IEEE, 2023b.

Zitao Liu, Qiongqiong Liu, Jiahao Chen, Shuyan Huang, and Weiqi Luo. simplekt: a simple but tough-to-beat baseline for knowledge tracing. *arXiv preprint arXiv:2302.06881*, 2023c.

Zitao Liu, Qiongqiong Liu, Teng Guo, Jiahao Chen, Shuyan Huang, Xiangyu Zhao, Jiliang Tang, Weiqi Luo, and Jian Weng. Xes3g5m: A knowledge tracing benchmark dataset with auxiliary information. In *Proceedings of the Advances in Neural Information Processing Systems: Annual Conference on Neural Information Processing Systems (NeurIPS)*, pp. 32958–32970, 2023d.

Jiaqi Ma, Zhe Zhao, Xinyang Yi, Jilin Chen, Lichan Hong, and Ed H Chi. Modeling task relationships in multi-task learning with multi-gate mixture-of-experts. In *Proceedings of the ACM SIGKDD international conference on knowledge discovery & data mining (KDD)*, pp. 1930–1939, 2018.

Yilmazcan Özyurt, Tunaberk Almaci, Stefan Feuerriegel, and Mrinmaya Sachan. Personalized exercise recommendation with semantically-grounded knowledge tracing. *arXiv preprint arXiv.2507.11060*, 2025.

Chris Piech, Jonathan Bassen, Jonathan Huang, Surya Ganguli, Mehran Sahami, Leonidas J Guibas, and Jascha Sohl-Dickstein. Deep knowledge tracing. *Advances in neural information processing systems*, 28, 2015.

Ruihong Qiu, Zi Huang, Hongzhi Yin, and Zijian Wang. Contrastive learning for representation degeneration problem in sequential recommendation. In *Proceedings of the ACM international conference on Web Search and Data Mining (WSDM)*, pp. 813–823, 2022.

Karim Shabani, Mohamad Khatib, and Saman Ebadi. Vygotsky's zone of proximal development: Instructional implications and teachers' professional development. *English language teaching*, 3 (4):237–248, 2010.

Noam Shazeer, Azalia Mirhoseini, Krzysztof Maziarz, Andy Davis, Quoc Le, Geoffrey Hinton, and Jeff Dean. Outrageously large neural networks: The sparsely-gated mixture-of-experts layer. *arXiv preprint arXiv:1701.06538*, 2017.

Yu Su, Qingwen Liu, Qi Liu, Zhenya Huang, Yu Yin, Enhong Chen, Chris Ding, Si Wei, and Guoping Hu. Exercise-enhanced sequential modeling for student performance prediction. In *Proceedings of the AAAI conference on artificial intelligence (AAAI)*, volume 32, 2018.

Jianwen Sun, Fenghua Yu, Qian Wan, Qing Li, Sannyuya Liu, and Xiaoxuan Shen. Interpretable knowledge tracing with multiscale state representation. In *Proceedings of the ACM Web Conference (WWW)*, pp. 3265–3276, 2024a.

Yifei Sun, Yifei Cao, Ziang Wang, Sicheng Hou, Weifeng Gao, and Zhi-Hui Zhan. Personalized exercise group assembly using a two archive evolutionary algorithm. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 9(3):2572–2583, 2024b.

Charles Vernerey, Samir Loudni, Noureddine Aribi, and Yahia Lebbah. Innovative constraint models for mining frequent and rare association rules using multi-objective optimization. *Artif. Intell.*, 357:104561, 2026.

Fei Wang, Qi Liu, Enhong Chen, Zhenya Huang, Yuying Chen, Yu Yin, Zai Huang, and Shijin Wang. Neural cognitive diagnosis for intelligent education systems. In *Proceedings of the AAAI conference on artificial intelligence (AAAI)*, volume 34, pp. 6153–6161, 2020.

Xin Wang, Xiuli Diao, Qingtian Zeng, Weijian Ni, and Zhengguo Song. Personalized exercise recommendation based on knowledge structure and learners' attempting preferences. *Knowledge and Information Systems*, 67(10):8417–8448, 2025.

Zhengyang Wu, Ming Li, Yong Tang, and Qingyu Liang. Exercise recommendation based on knowledge concept prediction. *Knowledge-Based Systems*, 210:106481, 2020.

Ruobing Xie, Yanlei Liu, Shaoliang Zhang, Rui Wang, Feng Xia, and Leyu Lin. Personalized approximate pareto-efficient recommendation. In *Proceedings of the Web Conference (WWW)*, pp. 3839–3849, 2021.

Heeyoon Yang, YunSeok Choi, Gahyung Kim, and Jee-Hyong Lee. Loam: improving long-tail session-based recommendation via niche walk augmentation and tail session mixup. In *Proceedings of the international ACM SIGIR conference on research and development in information retrieval (SIGIR)*, pp. 527–536, 2023.

Haochao Ying, Fuzhen Zhuang, Fuzheng Zhang, Yanchi Liu, Guandong Xu, Xing Xie, Hui Xiong, and Jian Wu. Sequential recommender system based on hierarchical attention network. In *IJCAI international joint conference on artificial intelligence (IJCAI)*, 2018.

Sukwon Yun, Kibum Kim, Kanghoon Yoon, and Chanyoung Park. Lte4g: Long-tail experts for graph neural networks. In *Proceedings of the ACM International Conference on Information & Knowledge Management (CIKM)*, pp. 2434–2443, 2022.

Haotian Zhang, Shuanghong Shen, Bihan Xu, Zhenya Huang, Jinze Wu, Jing Sha, and Shijin Wang. Item-difficulty-aware learning path recommendation: From a real walking perspective. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, pp. 4167–4178, 2024.

Yong Zheng and David Wang. Multi-objective recommendations. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery & Data Mining (KDD)*, pp. 4098–4099, 2021.

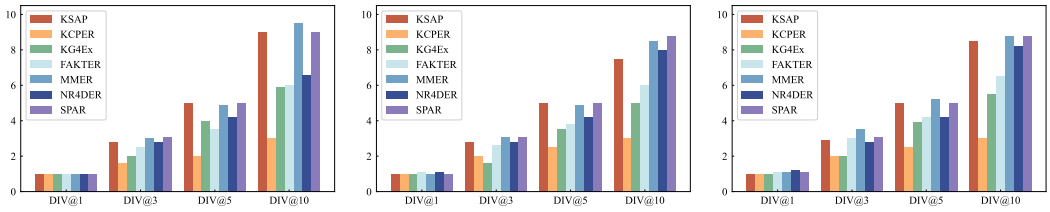
Kun Zhou, Hui Wang, Wayne Xin Zhao, Yutao Zhu, Sirui Wang, Fuzheng Zhang, Zhongyuan Wang, and Ji-Rong Wen. S3-rec: Self-supervised learning for sequential recommendation with mutual information maximization. In *Proceedings of the ACM international conference on information & knowledge management (CIKM)*, pp. 1893–1902, 2020.

Yangtao Zhou, Hua Chu, Ziwen Wang, Jiacheng Liu, Jianan Li, Yueying Feng, Xiangming Li, Zihan Han, Qingshan Li, et al. Knowledge starts with practice: Knowledge-aware exercise generative recommendation with adaptive multi-agent cooperation. *Advances in Neural Information Processing Systems*, 38:92980–93013, 2026.

A Appendix

A.1 Diversity Analysis

To further investigate the effectiveness of our state-conditioned routing mechanism, we present the diversity performance (DIV@K) across three datasets in Figure 5. While SPAR does not always yield the absolute highest diversity score, it consistently maintains a highly competitive level across all settings. This aligns with our pedagogical design philosophy. SPAR does not blindly pursue maximum diversity, as extreme diversity can disrupt pedagogical coherence and overwhelm learners with disjointed concepts. Instead, our micro state-conditioned routing mechanism dynamically calibrates the diversity weight γ according to the learner's real-time cognitive state. Consequently, SPAR achieves a superior pedagogical balance, successfully avoiding recommendation homogenization without sacrificing the efficiency of personalized learning.



(a) Diversity of XES3G5M (b) Diversity of Assist2009 (c) Diversity of Assist2012

Figure 5: The diversity of three datasets.

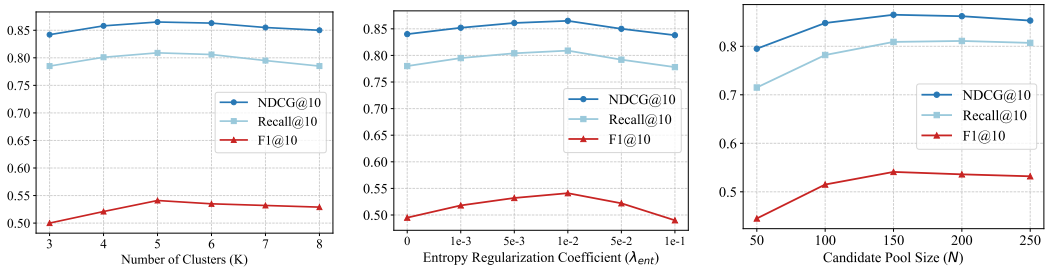
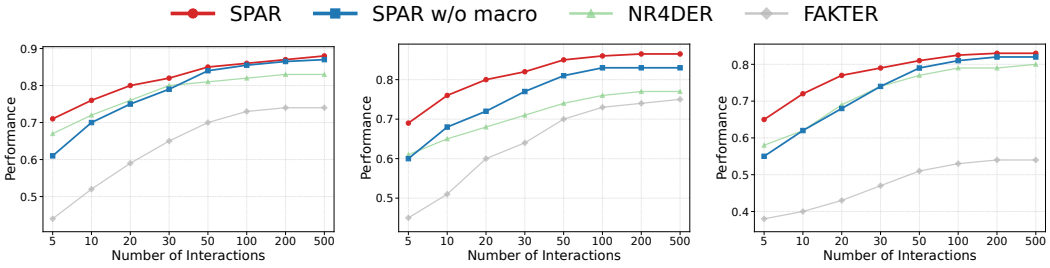


Figure 6: Performance comparison of different parameter settings.



(a) XES3G5M (b) Assist2009 (c) Assist2012

Figure 7: Recommendation performance under cumulative cold-start evaluation.

A.2 Parameter Sensitivity

The number of style prototypes K, the entropy regularization coefficient λ_{ent}, and the candidate pool size N are key parameters because they capture essential signals for personalized cognitive alignment and affect how we model dynamic pedagogical routing. In Figure 6, we examine the sensitivity of model performance to variations in these parameters. For the number of style prototypes K, performance increases as K grows, reaching a maximum at $K = 5$, after which it progressively declines but remains generally stable. Mainstream learning style theories typically categorize learners into 4 to 5 macro-level cognitive archetypes Fleming & Mills (1992); Kolb (2014). Therefore, $K = 5$ provides sufficient granularity to capture diverse intrinsic learning styles and avoid semantic noise, effectively preventing both under-clustering and noise-inducing over-clustering, while perfectly encapsulating a pedagogically meaningful representation of learners' inherent cognitive diversities. For the entropy regularization coefficient λ_{ent}, predictive performance improves steadily as λ_{ent} increases from 0, peaking at 0.01, but then declines once it exceeds this threshold. This performance drop at both extremes occurs because $\lambda_{ent} = 0$ leads to severe weight degeneration, while an excessively large λ_{ent} forces the routing weights to be nearly uniform, stripping the model of its dynamic state-conditioned adaptation capability. For the candidate pool size N, the recommendation quality improves as the pool expands, reaching optimal performance at $N = 150$, and then gradually decreases. This occurs because an overly small pool preemptively filters out cognitively appropriate exercises, whereas a large pool introduces pedagogically irrelevant noise that overwhelms the downstream micro-level re-ranker. Therefore, we set the number of style clusters K to 5, the entropy regularization coefficient λ_{ent} to 0.01, and the candidate pool size N to 150 in our experiments.

A.3 Fairness Evaluation under Different Retrieval and Re-ranking Settings

To further investigate whether the performance gains of SPAR mainly originate from the proposed ZPD-inspired retrieval or the state-conditioned dynamic routing network, we conduct an additional controlled experiment by decoupling the retrieval and re-ranking stages. Unlike most existing exercise recommendation methods that directly rank all exercises end-to-end, both MMER and NR4DER adopt a two-stage recommendation pipeline. Specifically, MMER first retrieves a candidate set using conventional retrieval strategies and then employs a multi-agent reinforcement learning framework to jointly optimize recommendation accuracy and knowledge diversity during re-ranking. NR4DER first enhances inactive learner representations through sequence enhancement and mLSTM-based retrieval, and subsequently applies a neural re-ranking module to generate personalized recommendation lists.

To ensure a fair comparison, we interchange the retrieval and re-ranking modules among MMER, NR4DER, and SPAR while keeping all other experimental settings unchanged, including the candidate pool size ($N = 150$), training protocol, evaluation metrics, and hyperparameter configurations. In this way, the influence of retrieval and ranking can be isolated. Under the SPAR retrieval setting, all re-ranking methods share the same candidate pool, thereby eliminating the influence of candidate generation and enabling a fair comparison of different re-ranking strategies. As shown in Table 4, replacing the original re-ranking module with SPAR consistently improves recommendation performance under both MMER and NR4DER retrieval strategies. When the proposed ZPD-inspired retrieval is combined with existing re-ranking models, both MMER and NR4DER achieve noticeable improvements over their original pipelines.

Retrieval	Re-Ranking	NDCG@10	LC-F1@10	Recall@10
MMER	MMER	68.62	51.02	47.48
MMER	SPAR	73.27	51.91	55.46
SPAR	MMER	81.69	52.13	68.37
NR4DER	NR4DER	83.41	52.81	77.21
NR4DER	SPAR	84.91	53.02	77.38
SPAR	NR4DER	85.33	53.45	78.01
SPAR	SPAR	**88.29**	**54.00**	**78.89**

Table 4: Comparison of different retrieval and ranking methods on XES3G5M.

A.4 Cold-start Quantitative Evaluation

To quantitatively evaluate the effectiveness of the proposed macro style alignment module across different sparsity levels, we conduct a cumulative interaction-threshold evaluation. Specifically, for each threshold T $\in$ 5, 10, 20, 30, 50, 100, 200, 500, we select only learners whose historical interaction count satisfies |Xu| $\leq$ T and evaluate recommendation performance on this subgroup. We compare the full SPAR framework against three representative baselines: (1) SPAR w/o macro, which removes the style alignment module, (2) NR4DER, a method that matches cold-start learners to all learners, and (3) FAKTER, a method that does not perform any processing on cold-start learners. The same evaluation protocol and candidate pool construction are used for all methods. Figure 7 reports the subgroup performance across three datasets.

SPAR consistently achieves the best performance across all interaction thresholds and datasets. The advantage is most pronounced in the extremely sparse regime ($\leq$5 and $\leq$10 interactions), where SPAR substantially outperforms both SPAR w/o macro and the external baselines.