ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 2 - 2026

Research on the Generation of Personalized Talent Training Program based on Cognitive diagnosis Model and generative adversarial network

Tao Wang[1,*], *Wensheng Liu*[1] *and Qiulin Hu*[1]

[1] School of Marxism of Yantai Nanshan University, Longkou 265713, Shandong, China

SUMMARY: *With the development of artificial intelligence and big data technology in education, the generation of talent training programs needs to shift from experience design to diagnosis-driven intelligent modeling. In this paper, a collaborative framework of cognitive diagnosis model and generative adversarial network is constructed, and the texts of learning behavior log, course evaluation, practical task and post ability are preprocessed and feature encoded. The condition vector is formed by knowledge state modeling, ability portrait, knowledge defect priority sorting and course goal-post ability graph fusion. And the conditional generative adversarial network is used to generate course modules, practical training, ability improvement paths and evaluation arrangements. The experiment was based on 3268 groups of student samples, 12 courses, 86 knowledge attributes and 42 types of post ability points. The results show that the knowledge defect location rate of the complete model is 93.8%, the effective rate of the generation scheme is 95.2%, the coverage rate of ability compensation is 92.4%, the matching degree of post ability is 91.7%, and the average generation time is 1.63 s. The research provides technical support for the intelligent generation and dynamic optimization of talent training programs in colleges and universities.*

KEYWORDS: *Cognitive diagnosis model; Generative adversarial networks; Personalized talent training; Portrait of competence*

1 Introduction

With the development of artificial intelligence, learning analysis and educational big data technology, the formulation of talent training programs is gradually shifting from experience-driven to data-driven and intelligent generation [1]. Traditional talent training programs usually rely on the manual design of professional leaders, curriculum teams and industry experts, which can reflect the basic relationship between training objectives, curriculum system and job requirements. However, it is often difficult to complete personalized adjustment in time in the face of differences in students 'cognitive basis, changes in learning behavior and rapid update of industrial ability requirements. Different students have obvious differences in knowledge mastery degree, ability shortcoming, curriculum adaptability and career development tendency, and a single training path is difficult to meet the needs of fine training [2, 3]. Therefore, how to use computer models to diagnose students' knowledge states and automatically generate talent training programs with pertinence, interpretability and dynamic adaptability has become an important issue in the research of intelligent education and educational informatization [4].

The cognitive diagnosis model can identify students 'mastery status on different

---

[*] wt199219wt@163.com

https://doi.org/10.65102/is2026906

knowledge attributes based on students 'answer records, course scores, learning process data and knowledge point relationships, which is suitable for constructing learners 'ability portraits. Compared with general performance evaluation methods, the cognitive diagnosis model can further reveal the weak links in the process of students 'ability formation and provide a fine-grained basis for the generation of personalized training programs [5]. However, the cognitive diagnosis model alone can only complete the ability identification and state assessment, and it is difficult to directly generate a complete talent training program [6]. The training program not only includes course recommendation, but also involves complex content such as training target matching, course module combination, practical task configuration, ability improvement path and post ability adaptation, which requires strong structure generation ability and multi-constraint optimization ability [7, 8].

Generative adversarial networks have strong expressive power in tasks such as data generation, structure completion and sample augmentation. By introducing it into the generation scenario of talent training programs, the generator can be used to learn the structure distribution of high-quality training programs, and the discriminator can be used to feedback constraints on the rationality of programs, curriculum cohesion and ability coverage, so as to improve the integrity and adaptability of the generated results [9, 10]. Based on this, this paper constructs a talent training program generation method based on the collaboration of cognitive diagnosis model and generative adversarial network: The cognitive diagnosis model is used to extract students 'knowledge mastery status, ability defects and learning behavior characteristics, and the course objectives and post ability map are combined to form structured inputs. Then the conditional generative adversarial network is used to generate personalized training programs, and the diagnosis feedback mechanism is introduced to dynamically optimize the generated results.

The research focus of this paper is to realize the collaborative modeling among students 'ability diagnosis, training element coding, scheme automatic generation and feedback correction. This method can improve the matching degree between curriculum configuration, ability objectives and students' individual needs on the basis of maintaining the structural integrity of training programs, and provide a computable, optimized and scalable technical path for the intelligent design of talent training programs in colleges and universities.

2 Related work

After artificial intelligence technology enters the field of education, learning data modeling, personalized recommendation, intelligent feedback and teaching decision support have gradually become the research focus. Tahiru systematically composes AI education research and points out that machine learning, intelligent tutoring system, learning analysis and automatic evaluation can improve the datatization degree of teaching process, but there are also problems such as insufficient model interpretation, weak ethical governance and difficult scene transfer [11]. In terms of learning process modeling, Carlon and Cross applied knowledge tracing to the metacognitive tutoring system, and predicted subsequent performance by continuously recording learners' answer behavior and cognitive changes, which provided a basis for adaptive learning path adjustment [12]. Ghiasabadi Farahani et al. proposed a personalized recommendation method combining learning automata and item clustering from the perspective of recommendation system, emphasizing that learning resource matching should consider learners' state, resource similarity and dynamic feedback at the same time [13]. Celik et al. pay attention to the supporting role of artificial intelligence in teachers' work, and believe that AI can assist teachers to complete learning diagnosis, resource screening and classroom intervention, but teachers' understanding, trust and ability to use algorithm results still affect the implementation effect of the system [14]. da Silva et al. reviewed the research on education recommendation system and pointed out that the existing methods have been extended from single content recommendation to learning path, teaching activity and evaluation resource recommendation, but there are still problems such as insufficient learning objective modeling and limited cold start processing ability [15]. Gligorea et al. analyzed AI-driven adaptive learning, and believed that learner profiling, behavior prediction, content adaptation and real-time feedback were the key links to build a personalized learning system [16]. Crompton and Burke further summarized the research status of AI application in higher education, pointing out that intelligent writing, learning analytics, automatic grading and personalized guidance are changing the way of course organization, but high-quality data, model transparency and teaching value verification still need to be strengthened [17]. Shum et al. introduced context adaptation and gating mechanism in serious game scenarios, indicating that personalized learning not only depends on static ability level, but also should be dynamically adjusted by combining task context, behavior trajectory and learning rhythm [18]. In terms of generative models, Farhood et al. used generative adversarial networks for student outcome prediction, which verified the potential of GAN in small sample enhancement, feature generation and prediction performance improvement [19]. Megahed and Mohammed reviewed the principles, applications and challenges of GAN, and pointed out that adversarial training can improve the authenticity of generated data, but training stability, mode collapse and evaluation criteria are still important limitations [20]. Starting from ChatGPT's educational influence, Rejeb et al. showed that generative AI can participate in content generation, learning support and teaching management, but its output needs to be combined with domain constraints and manual review [21]. Siafis et al. focused on teacher recommendation systems and proposed that teacher-oriented resource recommendation should take into account curriculum objectives, teaching contexts and students' needs [22]. Amin et al. constructed a Top-N course recommendation framework for online learning and proved that personalized course matching could improve the accuracy of learning path planning [23]. Lagos-Castillo et al. research on smart classroom shows that personalized learning schemes are moving from resource recommendation to environment perception, learning profiling and intelligent decision fusion [24]. It can be seen that the existing research provides a basis for learning state recognition, personalized recommendation and generative education support, but the research on the collaborative generation of training programs between cognitive diagnosis results and generative adversarial networks is still limited, especially the lack of a technical path to model knowledge defects, course objectives, post ability map and generation result constraints.

3 The generation model design of talent training program based on the collaboration of cognitive diagnosis model and generative adversarial network

3.1 The knowledge state modeling framework of personalized talent training needs

The premise of the generation of personalized talent training program is to transform students 'learning foundation, knowledge mastery level, ability shortboard and job adaptation needs into computable knowledge state representation. Based on the cognitive diagnosis model, this paper constructs a knowledge state modeling framework for talent training needs, and maps learning behavior logs, course assessment records, course scores, practical task completion and post ability requirements into students 'knowledge state space. The system first cleans the multi-source educational data, aligns the time series and normalizes the characteristics, and then constructs the knowledge attribute set according to the course knowledge points, ability goals and job skills requirements, so that the mastery probability of students on different knowledge attributes can be quantified. Different from relying solely on performance evaluation, this framework pays more attention to the structural differences between students 'course modules, practical ability and professional ability, so as to provide stable conditional input for subsequent generative adversarial networks.

Let the learning behavior characteristics, assessment performance characteristics, course score characteristics and post ability matching characteristics of student i be x_i^b, x_i^a, x_i^c and x_i^j respectively, then the student knowledge state vector can be expressed as follows:

$$z_i = \sigma\left(W_b x_i^b + W_a x_i^a + W_c x_i^c + W_j x_i^j + b\right) \tag{1}$$

Among them, z_i represents the knowledge state vector of student i, which reflects her mastery degree in multiple knowledge attributes and ability dimensions. W_b, W_a, W_c, W_j represent the weight matrices of different feature sources respectively. b is the bias term; $\sigma(\cdot)$ is a nonlinear activation function, which is used to map multi-source features into a unified knowledge state space. The formula realizes the joint coding of learning behavior, assessment results, course performance and post ability requirements, and provides basic input for the generation of subsequent personalized training programs.

In the process of knowledge state modeling, it is also necessary to incorporate the mapping relationship between course knowledge points and ability goals. Let the set of knowledge attributes be $K = \{k_1, k_2, \ldots, k_m\}$, the set of course objectives is $C = \{c_1, c_2, \cdots, c_n\}$,and the incidence matrix between them can be defined as follows:

$$R_{mn} = [r_{uv}], r_{uv} = \begin{cases} 1, k_u supports c_v \\ 0, k_u doesnotsupport c_v \end{cases} \tag{2}$$

Here,R_{mn} represents the mapping matrix between knowledge attributes and course objectives. r_{uv} indicates whether the u-th knowledge attribute supports the v-th curriculum goal. Through this matrix, the system can judge which course objectives will be affected by students 'short board of knowledge, and further determine the curriculum modules, practical tasks and ability training content that need to be strengthened in the training program. This modeling method transforms the personalized training needs from fuzzy description to structured calculation results. Figure 1 shows the generation framework of talent training program for the collaboration of cognitive diagnosis and generative adversarial network.

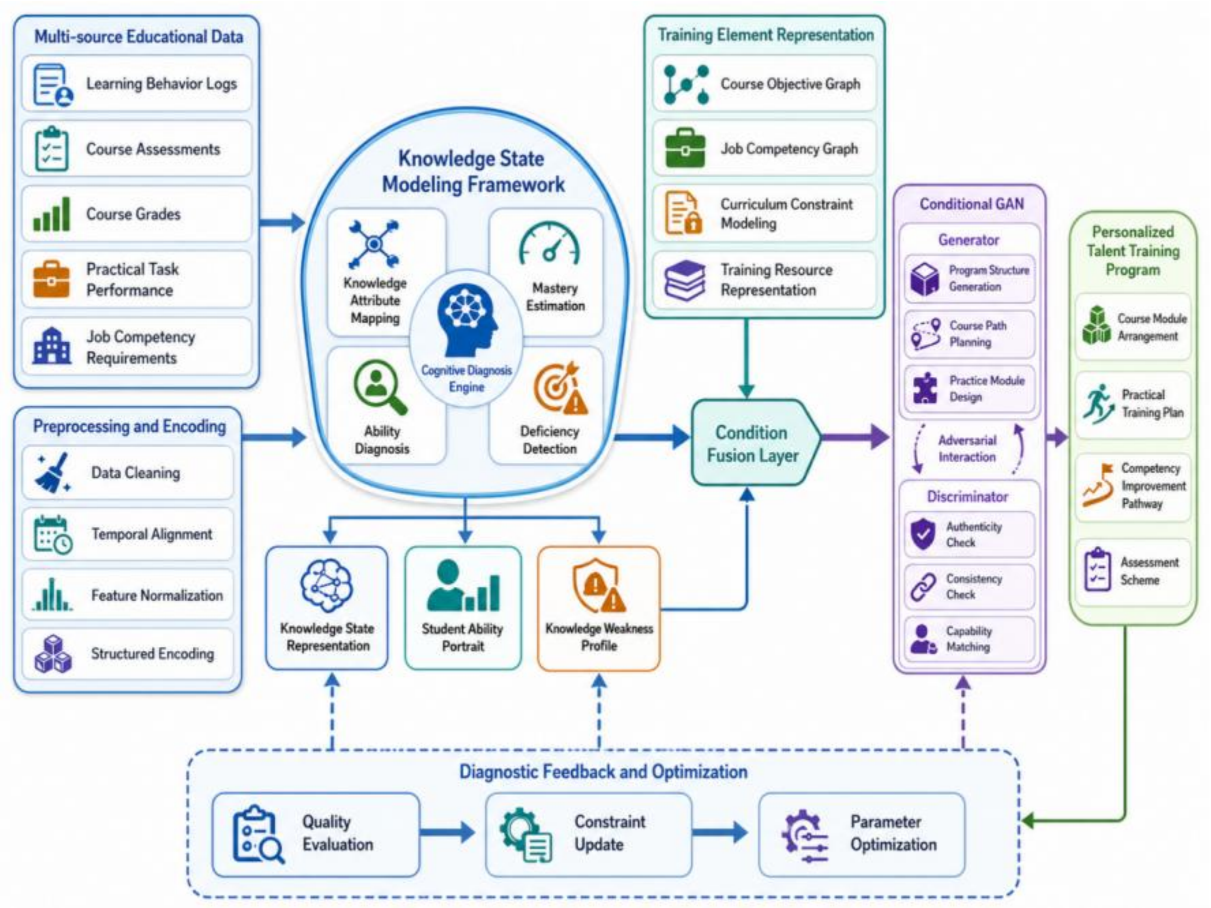

Figure 1: Framework diagram of talent training program generation based on cognitive diagnosis and generative adversarial network collaboration

As shown in Figure 1, the knowledge state modeling framework constructed in this paper includes multi-source educational data input, preprocessing and feature coding, cognitive diagnosis module, training element representation, conditional generative adversarial network, and personalized training scheme output. The knowledge state vector and student ability portrait output by the cognitive diagnosis module enter the condition fusion layer together with the course goal map and the post ability map to provide constraints for the generator and discriminator. Through the above framework, students' individual differences can be transformed into inputable, trainable and feedback-able computational features, which not only ensures the fine-grained expression of cognitive diagnosis results, but also lays a model foundation for generative adversarial networks to generate talent training programs that meet students' development needs.

3.2 Students' ability portrait and knowledge defect identification method based on cognitive diagnosis model

Students' ability profiling and knowledge defect identification are the key links in the generation of personalized talent training programs, whose core is to transform students' scattered performance in course learning, online behavior, stage evaluation and practical tasks into computable ability states. Based on the cognitive diagnosis model, this paper takes students 'answer results, course knowledge attributes, learning behavior sequences and post ability requirements into the diagnosis process. Through the neural network feature coding and knowledge attribute mapping, students 'mastery level on different knowledge points and ability units is identified. Compared with the single performance evaluation, this method can refine to the knowledge attribute level, determine which course modules students have weak links, and further analyze the impact of these weak links on the formation of post ability and the achievement of training objectives.

Suppose that the multi-source learning features of student i are encoded to obtain the hidden layer representation h_i. The mastery probability of student I on the KTH knowledge attribute can be expressed as follows:

$$p_{ik} = \frac{1}{1 + \exp[-(w_k^T h_i + \eta_k)]} \tag{3}$$

Here, p_{ik} represents the mastery probability of student i to the KTH knowledge attribute. h_i represents the hidden layer features of students 'learning behavior, curriculum evaluation, practical performance and other data encoded. wk is the trainable weight parameter corresponding to the knowledge attribute; η_k is the attribute bias term.

After the mastery probability estimation is completed, the system needs to further judge the severity of the student's knowledge deficiency. In this paper, considering the lack of knowledge, the importance of knowledge attributes and the learning risk status of students, the strength of knowledge defects is defined as follows:

$$d_{ik} = (1 - p_{ik}) \cdot \omega_k \cdot \rho_i \tag{4}$$

Here, d_{ik} denotes the defect strength of student i on the KTH knowledge attribute; p_{ik} is the mastery probability of the knowledge attribute; ω_k represents the important weight of knowledge attributes in the course objectives and post ability structure. ρ_i represents the learning risk coefficient of students, which can be comprehensively obtained by the recent score fluctuation, task submission delay, learning activity decline and assessment error rate. The formula can distinguish between common weaknesses and key ability gaps, so that the system can give priority to the core issues that affect the achievement of training objectives when generating subsequent programs.

In order to form a complete student ability portrait, this paper fuses knowledge mastery probability, learning behavior coding and practical task performance to construct a unified ability portrait vector:

$$a_i = \mathrm{ReLU}\big(W_p[p_i \| s_i \| r_i] + b_p\big) \tag{5}$$

where a_i represents the ability profile vector of student i; p_i represents the mastery probability set of students on all knowledge attributes. S_i represents learning behavior sequence coding results; r_i represents the practice task performance coding results; $\|$ denote feature splicing operation; W_p is the mapping matrix of capability portrait; b_p is the bias term; $\mathrm{ReLU}(\cdot)$ is the activation function. The formula compressed the knowledge state, behavior process and practical ability into the same feature space, so that the student portrait had the unified dimension and expression ability required for model input.

In order to make the knowledge defect identification results serve for the generation of personalized training programs, this paper further constructs a defect priority sorting function, which integrates knowledge defects, course prerequisite relationships and post ability requirements into the judgment process:

$$q_{ik} = \frac{\exp(d_{ik} + \lambda_1 u_k + \lambda_2 v_k)}{\sum_{k=1}^{m} \exp(d_{ik} + \lambda_1 u_k + \lambda_2 v_k)} \tag{6}$$

Here, q_{ik} represents the defect priority of student i on the KTH knowledge attribute. d_{ik} represents the strength of knowledge deficiency; u_k represents the importance of the prior knowledge attribute corresponding to the course module; v_k represents the matching weight between the knowledge attribute and the post ability requirement. λ_1 and λ_2 are the adjustment coefficients; m denotes the total number of knowledge attributes. The formula sorted different knowledge defects through the normalization mechanism, so that the training program generation model could give priority to the knowledge units with stronger relationships with curriculum connection, ability improvement and job adaptation.

In the overall calculation process, the system first encodes students 'multi-source data, and then outputs the knowledge mastery probability by the cognitive diagnosis model, and then generates the ability portrait according to the defect intensity and priority sorting. The portrait contains four types of information, including knowledge status, ability level, weak knowledge points and reinforcement direction, which can be used as the input constraints of the conditional generative adversarial network to guide the generator to output a talent training program that is more in line with the individual needs of students. In this way, the analysis of students 'ability is transformed from static evaluation to dynamic calculation process, which provides stable data support for subsequent course module configuration, practical task recommendation and training path optimization.

3.3 Structural representation of training elements integrating course objectives and post ability map

After the formation of students 'ability portraits, the generation of talent training programs also needs to transform curriculum objectives, post capabilities, curriculum resources and practical tasks into a unified structured representation. Based on the course goal graph and the post ability graph, this paper constructs the course modules, knowledge attributes, professional ability points and practical projects as heterogeneous graph nodes, and describes the supporting relationship between training elements through the graph correlation weights, so that the generative adversarial network can generate training programs under explicit constraints.

Let the matching strength between the course target node and the post ability node be:

$$s_{uv} = \sigma\left(c_u^T W_g g_v + r_{uv}\right) \tag{7}$$

where, s_{uv} represents the matching strength between the u-th course goal and the v-th post ability. cu represents the embedding vector of course objectives. g_v represents the position ability embedding vector; W_g is the graph matching weight matrix; r_{uv} represents the explicit correlation characteristics of the two in course standards and job descriptions. This formula is used to measure the support degree of course objectives to post ability and provide a basis for the selection of training elements.

Training elements such as courses, practical projects and evaluation tasks need to be further encoded into a unified vector, which is expressed as follows:

$$e_j = \mathrm{ReLU}\left(W_e[o_j||m_j||l_j] + b_e\right) \tag{8}$$

where e_j denotes the encoding vector of the JTH cultivation element; o_j represents the characteristics of the course objectives to which the elements belong; m_j represents the feature of the corresponding capability module. l_j represents the constraint features such as learning load, difficulty and prior knowledge relationship. W_e and b_e are the mapping matrix and the bias term, respectively. The formula can transform different types of cultivation elements into the same feature space, which is convenient for subsequent condition fusion.

In the input stage of the generative model, the system fuses the student ability portrait, the course goal map and the post ability map to form a condition vector:

$$f_i = \mathrm{LayerNorm}(W_f[a_i||\bar{e}||\bar{g}] + b_f) \tag{9}$$

where f_i represents the generation condition vector for student i; a_i represents a student's ability profile. $\bar{e}$ represents the aggregation representation of cultivation elements; $\bar{g}$ represents the aggregation representation of post ability map; W_f and b_f are the fusion parameters. LayerNorm(·) is used to stabilize the feature distribution. This vector connects the individual needs of students with the constraints of training objectives, and is an important input of conditional generative adversarial networks.Table 1 shows the characteristic variables and coding methods for the generation of personalized talent training program.

Table 1: Characteristic variables and coding method table of personalized talent training program generation

Feature category	Specific variables	Data source	Encoding method	Technical function
Student knowledge state features	Knowledge point mastery probability, knowledge deficiency intensity, deficiency priority	Course assessments, response records, cognitive diagnosis results	Probability vector encoding	Represents students' current knowledge foundation and weak points
Learning behavior features	Login frequency, learning duration, task submission delay, resource access path	Online learning platform logs	Temporal feature encoding	Captures students' learning engagement and behavioral change trends
Course objective features	Course objective hierarchy, ability objective description, course module relationship	Training program texts, course syllabi	Graph node embedding encoding	Describes supporting relationships among course objectives
Job competency features	Professional skill requirements, job task requirements, occupational competency levels	Job standards, recruitment texts, industry competency frameworks	Semantic embedding and graph encoding	Establishes the connection between training objectives and occupational demands
Practical task features	Experimental projects, practical training tasks, project difficulty, task outcome requirements	Practical teaching plans, project task library	Category encoding and difficulty scalar encoding	Supports practical module generation and competency transfer training
Training constraint features	Credit limits, prerequisite relationships, course cycle, learning load	Teaching plans, course management system	Rule constraint encoding	Controls the executability and course coherence of the generated program

Through the above structured representation, the student states, course objectives, post capabilities and teaching constraints required for the generation of training programs are unified as trainable features, which provide stable input conditions for subsequent conditional generative adversarial networks.

3.4 Design of Conditional generative adversarial Networks for Training scheme Generation

After the extraction of students 'ability portraits and the structured representation of training elements, this paper constructs a conditional generative adversarial network for the generation of personalized talent training programs. The model takes students' ability portrait, training element coding and post ability map aggregation results as conditional inputs, learns the structure distribution of training programs through the generator, and then jointly discriminates the authenticity, curriculum cohesion and ability matching of programs by the discriminator. Different from general text generation or recommendation models, conditional generative adversarial networks are more suitable for dealing with multi-element collaborative generation tasks such as course module combination, practical project arrangement and evaluation scheme configuration, which can enhance the feasibility and constraint consistency of output results while maintaining the diversity of schemes.

Let the condition vector of student i be f_i and the noise vector be z_i, then the personalized training scheme output by the generator is expressed as follows:

$$\hat{y}_i = G(z_i, f_i; \theta_G) \tag{10}$$

where $\hat{y}_i$ represents the training scheme vector output by the generator, which includes curriculum module arrangement, practical training plan, ability improvement path and evaluation scheme. G(·) denotes the generator mapping function; Let θ_G be the generator parameter. The discriminator receives the real culture scheme y_i and the generated scheme $\hat{y}_i$, and combines the condition vector to determine whether the scheme meets the requirements of high-quality culture design. Its adversarial objective function is defined as follows:

$$\mathcal{L}_{adv} = \mathbb{E}_{y_i \sim p_{data}}[\log D(y_i, f_i)] + \mathbb{E}_{z_i \sim p_z}[\log(1 - D(G(z_i, f_i), f_i))] \tag{11}$$

where D(·) represents the discriminator; p_{data} represents the distribution of true training schemes; p_z denotes the noise distribution. The first term is used to improve the discriminator's ability to recognize the true scheme, and the second term is used to compress the difference between the generated scheme and the true scheme. With the help of this adversarial mechanism, the generator could continue to approach the structural characteristics of high-quality training programs, and the discriminator continuously strengthened its ability to distinguish curriculum logic, ability coverage and configuration rationality.

Only relying on adversarial loss is difficult to ensure that the scheme fully matches the student's ability shortboard. Therefore, this paper adds the ability matching loss and the structure constraint loss in the generation stage to construct the joint optimization objective:

$$\mathcal{L} = \mathcal{L}_{adv} + \alpha\mathcal{L}_{match} + \beta\mathcal{L}_{struct} \tag{12}$$

where $\mathcal{L}$ is the total loss function; $\mathcal{L}_{match}$ represents the matching loss between the student's ability profile and the output of the training program. $\mathcal{L}_{struct}$ represents the loss of structural constraints such as course prerequisite relationship, credit limit, practice load and post ability coverage. α and β are the weight coefficients. This objective function enables the model to promote the generation of training schemes from "looking reasonable" to "structurally executable and ability adaptable" while maintaining the ability of adversarial learning. The conditional generative adversarial network structure is shown in Figure 2. Conditional generative adversarial networks consist of a conditional input layer, a generator, a discriminator, and a scheme output layer. The conditional input layer receives students' ability portrait, knowledge defect priority, course goal map and post ability map. The generator is responsible for outputting candidate training schemes. The discriminator discriminates from three levels: authenticity, consistency and ability matching. The final output includes course module combination, practical training arrangement, ability improvement path and evaluation scheme. The structure can explicitly embed the cognitive diagnosis results and training element constraints into the generation process, which provides a stable basis for subsequent feedback optimization.

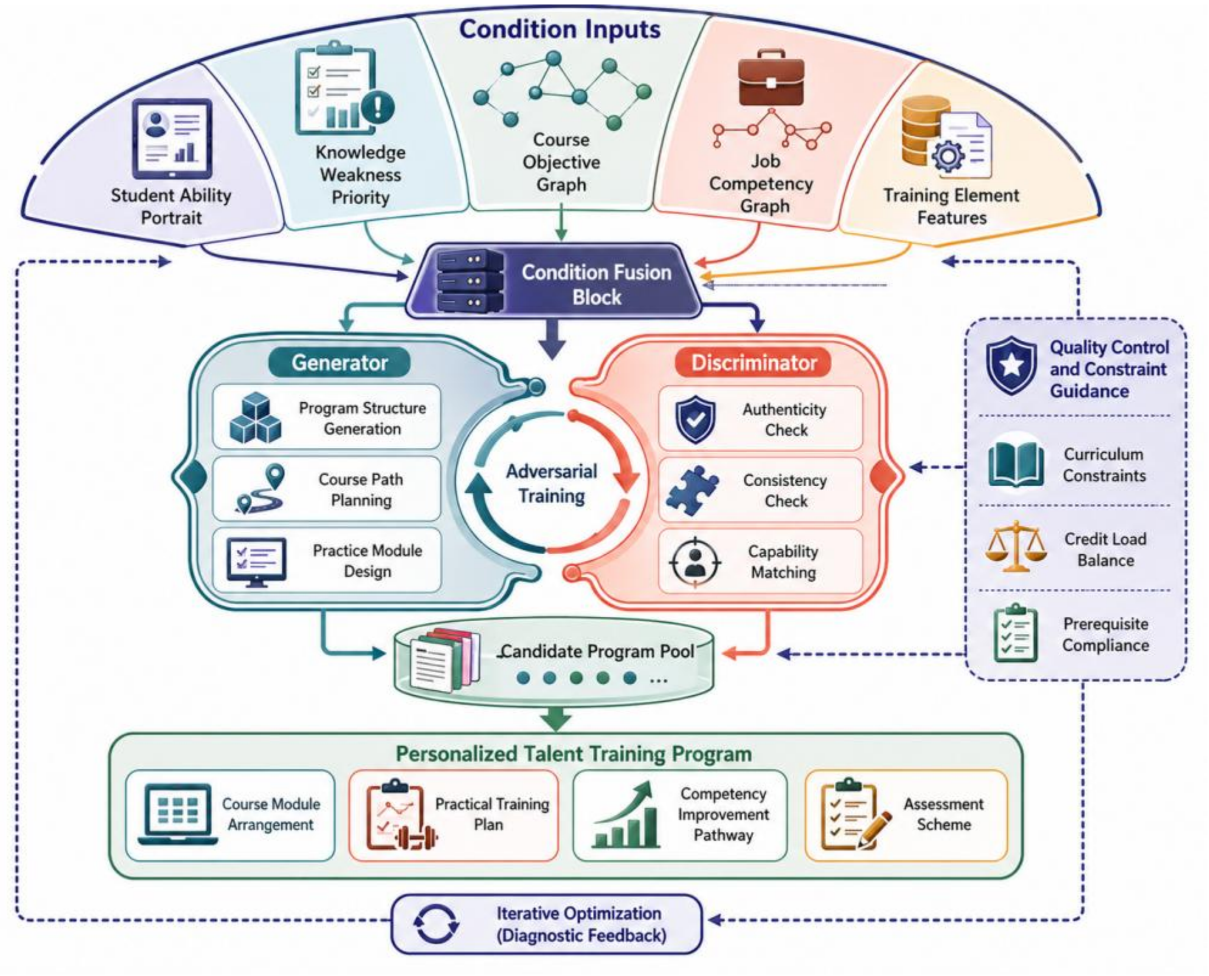

Figure 2: Structure diagram of conditional generative adversarial networks

Through the above design, the generation of personalized talent training program changes from static rule matching to condition-driven adversarial learning process, which not only improves the flexibility of training program generation, but also enhances the correspondence between curriculum configuration and students 'development needs, and lays a model foundation for the subsequent dynamic optimization mechanism driven by cognitive diagnosis feedback.

3.5 Cognitive diagnosis feedback-driven generation result constraint and dynamic optimization mechanism

Conditional generative adversarial networks can output candidate talent training programs, but the generated results still need to further meet multiple constraints such as curriculum structure, post ability, learning load and students 'knowledge defects. In order to improve the executability and individual adaptability of the scheme, this paper constructs a dynamic optimization mechanism of generation results driven by cognitive diagnosis feedback. Taking the candidate scheme output by the generator as the initial result, the students 'knowledge defects, ability gap, course prerequisite relationship, credit load and practical task matching degree into the joint constraints, and the quality evaluation results were used to reverse modify the generation parameters, so that the training scheme was gradually close to the optimal state in multiple rounds of iterations.

Suppose that the course constraint, capacity constraint, load constraint and practice task constraint of the generated scheme are C_{cur}, C_{com}, C_{load} and C_{pra} respectively, then the multi-constraint penalty term is defined as follows:

$$\mathcal{L}_{con} = \mu_1 \mathcal{C}_{cur} + \mu_2 \mathcal{C}_{com} + \mu_3 \mathcal{C}_{load} + \mu_4 \mathcal{C}_{pra} \tag{13}$$

Here, $\mathcal{L}_{con}$ represents the comprehensive constraint loss of the generation scheme. μ_1, μ_2, μ_3, μ_4 are the weight coefficients of different constraint terms, respectively. C_{cur} is used to describe the course prerequisite relationship and the course order conflict. The curriculum constraints can be further refined into curriculum structure constraints and prerequisite relationship constraints in the optimization process, so as to improve the logical consistency and implementation rationality of the training program. C_{com} is used to measure the insufficient coverage of the program to the post ability point; C_{load} is used to limit too high or too low study load during the semester. C_{pra} is used to assess the degree of match between practice tasks and competency training objectives. This formulation transforms the decentralized rule constraints into an optimizable loss, so that the generated results can accept a unified computational constraint.

In the cognitive diagnosis feedback layer, the system modifies the generation scheme according to the students 'latest knowledge mastery state. Let the student defect priority vector be q_i, and the reinforcement weight of the corresponding training element in the generation scheme be r_i. Then the diagnostic feedback correction can be expressed as follows:

$$\Delta_\mathrm{i} = \mathrm{Softmax}(\mathrm{W_d}[\mathrm{q_i}||\mathrm{r_i}] + \mathrm{b_d}) \tag{14}$$

Here, Δ_i represents the scheme correction weight for student i; q_i represents the priority of knowledge deficiency obtained from the cognitive diagnosis model. r_i represents the strengthening weights of course modules, practice tasks and evaluation links in the generation scheme. W_d and b_d are feedback mapping parameters; $||$ denotes feature concatenation. The correction amount can be used to adjust the proportion of course modules, the intensity of practical training and the ability improvement path, so that the program is more suitable for students' current ability shortcomings.

The generated results also need to be comprehensively scored by the quality evaluation module. In this paper, the quality scoring function is constructed from four aspects: authenticity, consistency, ability matching and execution feasibility:

$$\mathrm{S_i} = \tau_1 \mathrm{A_i} + \tau_2 \mathrm{K_i} + \tau_3 \mathrm{M_i} + \tau_4 \mathrm{E_i} \tag{15}$$

where S_i represents the comprehensive quality score of student i corresponding to the generated scheme. A_i represents the degree of similarity between the distribution of the program and the real high-quality training program; K_i indicates the consistency of course structure and prerequisite relationship. M_i represents the matching degree between the scheme and the students' ability defects and the post ability requirements. E_i represents the implementation feasibility of the scheme under the conditions of credits, cycles and teaching resources. τ_1, τ_2, τ_3, and τ_4 are the scoring weights. The score can provide quantitative basis for candidate selection and dynamic iteration.

In the model optimization stage, the system combines adversarial loss, multi-constraint loss and feedback correction results to update the generator parameters, and the update rules are as follows:

$$\theta_G^{t+1} = \theta_G^t - \eta \nabla_{\theta_G} (\mathcal{L}_{adv} + \lambda \mathcal{L}_{con} - \gamma S_i) \tag{16}$$

Here, θ_G^t and θ_G^{t+1} denote the generator parameters in round t and round t+1, respectively. η is the learning rate; ∇_{θ_G} denotes the gradient of the generator parameters; Let λ denote the constraint loss weight; Let γ denote the quality score incentive weight. This update method makes the model reduce the generation error and constraint conflict, and improve the generation probability of high-quality training programs.

Figure 3 shows the multi-constraint joint optimization mechanism of the generated results.

Figure 3: Diagram of the resulting multi-constraint joint optimization mechanism

Through this mechanism, the generation process of training scheme forms a closed-loop optimization process of "generation-evaluation-feedback-correction-regeneration". The results of cognitive diagnosis provide an individualized basis for the adjustment of the program, the multi-constraint calculation ensures the reasonable structure of the program, and the dynamic parameter update improves the continuous optimization ability of the model, so as to enhance the adaptability and enforceability of the personalized talent training program in real teaching scenarios.

4 Results

4.1 Dataset construction and description of learning behavior characteristics

In order to verify the effectiveness of cognitive diagnosis model and generative adversarial network to collaboratively generate personalized talent training program, this paper constructs a multi-source education dataset including learning behavior, curriculum evaluation, practical tasks and post ability requirements. Data sources included online learning platform logs, course stage test results, course performance records, practical project evaluation forms and post ability description texts. A total of 3268 groups of student samples were collected, covering 12 professional courses, 86 knowledge attributes, 42 types of post ability points and 31 types of practical tasks. On average, each group of samples contained 34 dimensions of learning behavior features, 18 dimensions of course performance features, 86 dimensions of knowledge mastery features and 42 dimensions of post ability matching features.

Learning behavior characteristics mainly included login frequency, learning duration, resource access path, video completion rate, task submission delay and assessment error rate, which were used to reflect students 'learning engagement and process changes. The course evaluation data is used to construct the knowledge attribute mastery matrix, the practical task data is used to describe the quality of students 'project completion and ability transfer performance, and the post ability text is transformed into the ability graph node through semantic embedding. The dataset was divided into training set, validation set and test set according to 7:1.5:1.5, and the proportion of students with different ability levels was basically the same. The dataset can support subsequent experiments such as knowledge diagnosis, condition generation and scheme evaluation.

4.2 Multi-source data preprocessing and capability feature coding process

There are some problems in the collection process of multi-source educational data, such as inconsistent time granularity, missing fields, different evaluation criteria and differences in text expression, so it is necessary to complete unified preprocessing before modeling. Firstly, this paper uses student number, course number and timestamp as the key to associate and match the learning log, course evaluation, course performance, practical task evaluation and post ability text. Then, abnormal login time, duplicate submission records and invalid evaluation samples were removed. For continuous features with missing rate less than 20%, the sample mean of the same course and ability level was used to complete. The samples with higher missing rate are directly deleted to reduce the interference of noise on cognitive diagnosis results.

In order to eliminate the differences of different feature dimensions, this paper normalizes the continuous learning behavior features:

$$\tilde{x}_{ij} = \frac{x_{ij} - \min(x_j)}{\max(x_j) - \min(x_j) + \varepsilon} \quad (17)$$

Here, $\tilde{x}_{ij}$ denotes the normalized result of student i on the JTH feature; x_{ij} is the original feature value; $\min(x_j)$ and $\max(x_j)$ are the minimum and maximum values of the feature in the sample set, respectively. Let ε be the smoothing term that prevents the denominator from being zero. The formula can map the data of different scales such as learning time, login frequency and task delay into a unified interval to ensure the stability of subsequent model training.

In the ability feature coding stage, the learning behavior sequence uses the sliding window method to extract dynamic change features, and the encoding form is as follows:

$$b_i = \frac{1}{T}\sum_{t=1}^{T} \phi\left(l_{i,t}\right) \quad (18)$$

Here, b_i represents the learning behavior encoding vector of student i; T denotes the number of time steps in the sliding window; $l_{i,t}$ denotes the learning log features of the student at the TTH time step; Let $\phi(\cdot)$ denote the nonlinear feature mapping function. The formula can compress the continuous learning behavior into a stable behavior representation, and reflect the change trend of students 'learning engagement, resource access and task completion process.

The course evaluation results were converted into knowledge mastery vectors according to the knowledge attribute matrix, the practical task performance was formed into practical ability characteristics through level coding and difficulty weight, and the post ability text was converted into ability node vectors using the semantic embedding model. Finally, all the features were concatenated into a unified input matrix of "learning behavior - knowledge mastery - practical performance - post ability", which provided a stable data basis for subsequent evaluation index calculation and ablation experiments.

4.3 Evaluation index setting

In order to avoid evaluating the effect of the model only by the classification accuracy, the evaluation index is adjusted to a comprehensive index system oriented to "diagnosis-generation-match-deployment". At the knowledge diagnosis level, the recognition accuracy rate of knowledge status and the localization rate of knowledge defects are used to measure the recognition ability of the cognitive diagnosis model to students 'mastery status and weak knowledge points. At the scheme generation level, the efficiency rate of the generation scheme and the compliance rate of the course structure are set. The former indicates whether the generated results contain complete course modules, practical tasks and evaluation arrangements, and the latter is used to test whether the course prerequisite relationship, credit allocation and training cycle conform to the teaching rules.

At the personalized adaptation level, the ability compensation coverage rate and the post ability matching degree were used to reflect the strengthening degree of the generation scheme to students 'short board ability, and the corresponding relationship between the course content and the post ability map. At the generation quality level, the diversity index and repetition rate of the scheme are added to judge whether the model can generate differentiated cultivation paths and avoid output template schemes. At the system efficiency level, the average generation time, single-sample inference delay and video memory occupation were used to evaluate the running feasibility of the model under large-scale student samples. Through the above indicators, the comprehensive performance of the model in knowledge diagnosis, scheme generation, personalized adaptation and engineering deployment can be more comprehensively reflected.

4.4 Ablation experiments

In order to verify the contribution of each core module to the generation effect of personalized talent training program, this paper sets up an ablation experiment. Under the same data set, training rounds and evaluation indicators, the cognitive diagnosis module, post ability map constraints, conditional generation confrontation structure and feedback optimization mechanism are removed respectively, and compared with the complete model. The experiment focuses on the changes of indicators such as knowledge defect location, scheme generation effectiveness, ability compensation coverage and post ability matching, so as to judge the role of different modules in the process of diagnosis, generation and optimization. Table 2 shows the results of ablation experiments with different model configurations.

Table 2: Table of ablation experiment results for different modules

Model configuration	Knowledge deficiency localization rate/%	Generated program effectiveness rate/%	Capability compensation coverage rate/%	Job competency matching degree/%	Average generation time/s
Without cognitive diagnosis module	82.6	87.4	80.9	84.7	1.48
Without job competency graph constraint	89.1	90.2	86.5	81.8	1.56
Without conditional generative adversarial structure	88.4	84.9	85.7	86.3	1.31
Without feedback optimization mechanism	90.3	91.6	87.8	88.1	1.42
Complete model	93.8	95.2	92.4	91.7	1.63

As can be seen from Table 2, the full model performs optimally in all core indicators. After removing the cognitive diagnosis module, the knowledge defect location rate and ability compensation coverage rate decreased most obviously, indicating that the cognitive diagnosis result was the key basis for personalized generation. After removing the constraints of the post ability map, the matching degree of the post ability was reduced to 81.8%, indicating that the ability map could effectively enhance the connection between the training program and the occupational needs.

5 Discussion

5.1 Performance comparison analysis with the existing talent training program generation method

In order to verify the comprehensive advantages of the proposed method in the generation task of personalized talent training program, this paper compares the rule recommendation, collaborative filtering recommendation, knowledge tracking recommendation, and GAN generation model with the model in this paper. The comparison indicators were carried out around the four links of knowledge diagnosis, scheme generation, ability matching and feedback optimization, which were used to observe the performance differences of different methods in the utilization of diagnosis results, the organization of training schemes, the adaptation of post capabilities and dynamic correction. Rule recommendation and collaborative filtering recommendation mainly rely on static rules or similar user relationships, which are weak in response to students 'knowledge defects and post ability changes. Knowledge tracking recommendation can improve the recognition of knowledge status, but it is insufficient to support the generation of complete training program structure. The single GAN model has strong generation ability, but it lacks cognitive diagnosis constraints, which is prone to the problem of insufficient scheme personalization. The performance gain of different methods in the collaborative process of diagnosis generation is shown in Figure 4.

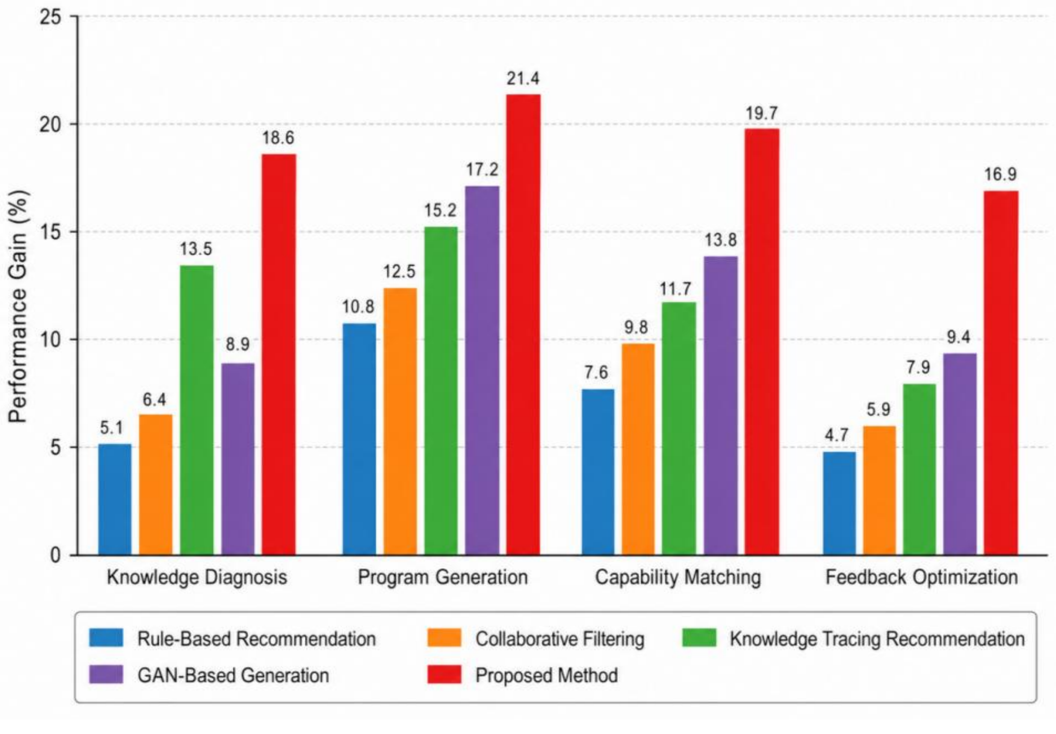

Figure 4: Bar chart of performance gain comparison of different methods during diagnosis generation collaboration

It can be seen from Figure 4 that the proposed model achieves the highest gain in all four indicators. The rate of knowledge diagnosis was 18.6%, which was higher than 13.5% of knowledge tracking recommendation. The scheme generation process reaches 21.4%, which is 4.2 percentage points higher than 17.2% of the single GAN model. Capability matching and feedback optimization reach 19.7% and 16.9%, respectively, which are significantly higher than other methods. The results show that the cooperation of cognitive diagnosis, conditional generative adversarial network and feedback optimization mechanism can improve the quality of scheme generation and the adaptation effect of individual ability at the same time.

5.2 Analysis of the influence of cognitive diagnosis results on the personalization ability of the generative model

The cognitive diagnosis results directly affect the input quality of conditional generative adversarial networks. In order to analyze its effect on the ability of personalization generation, this paper divided the student samples into four categories according to the degree of knowledge deficiency: low deficiency, medium deficiency, high deficiency and compound deficiency, and compared the personalization matching degree of the training program under the three conditions of no cognitive diagnosis input, only knowledge mastery input and complete cognitive diagnosis input. The matching degree mainly reflects the degree of adaptation between the generation scheme and students 'knowledge shortboards, ability portraits, course objectives and post ability requirements.Figure 5 shows the distribution of personalized matching degree of training scheme under different knowledge defect levels.

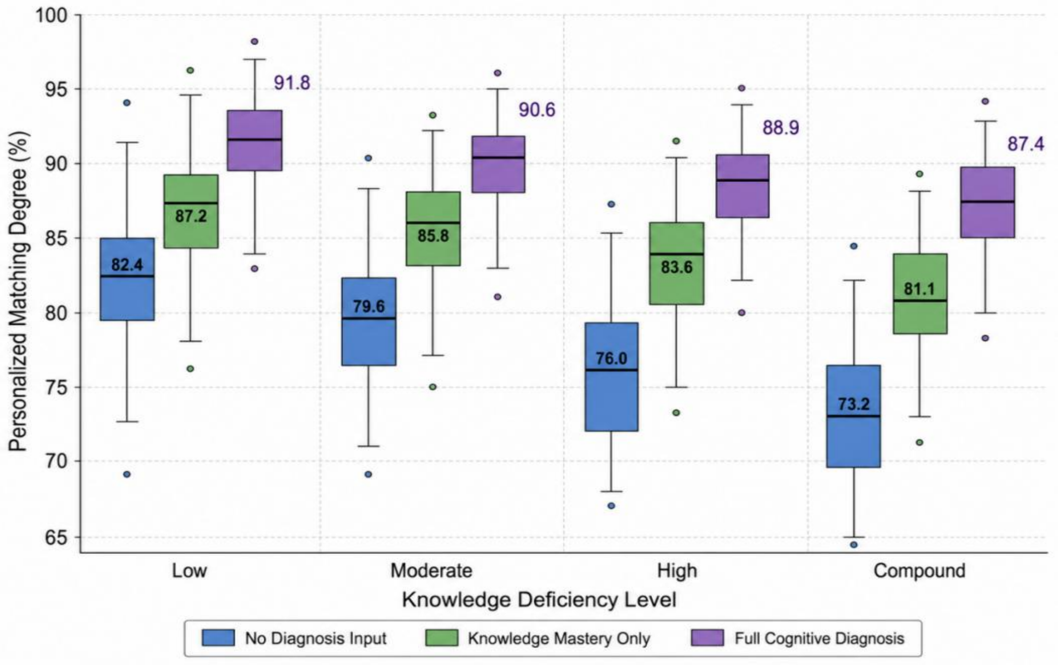

Figure 5: Box plots for comparison of personalized matching degree of training schemes under different levels of knowledge deficiency

It can be seen from Figure 5 that the complete cognitive diagnosis input performs best at all defect levels, and the median personalized matching degree of low defect, medium defect, high defect and composite defect samples reaches 91.8%, 90.6%, 88.9% and 87.4%, respectively. Compared with no cognitive diagnosis input, the composite defect sample has the most obvious improvement, and the matching degree is increased from 73.2% to 87.4%, an increase of 14.2 percentage points. The results show that cognitive diagnosis can not only improve the recognition ability of the generative model for general learning differences, but also enhance the effect of the model in generating personalized plans under the condition of complex knowledge shortcomings.

5.3 Model stability and adaptability verification under complex learning behavior data

The problems of missing, noise, delay and performance fluctuation in complex learning behavior data will weaken the stability of cognitive diagnosis results and further affect the quality of training program generation. In order to verify the adaptability of the proposed model under complex data conditions, this paper sets up five scenarios: no disturbance, behavior loss, evaluation noise, task delay and performance fluctuation, and compares them with the single GAN generation model. Figure 6 shows the quality change of the generation scheme under the learning behavior perturbation condition.

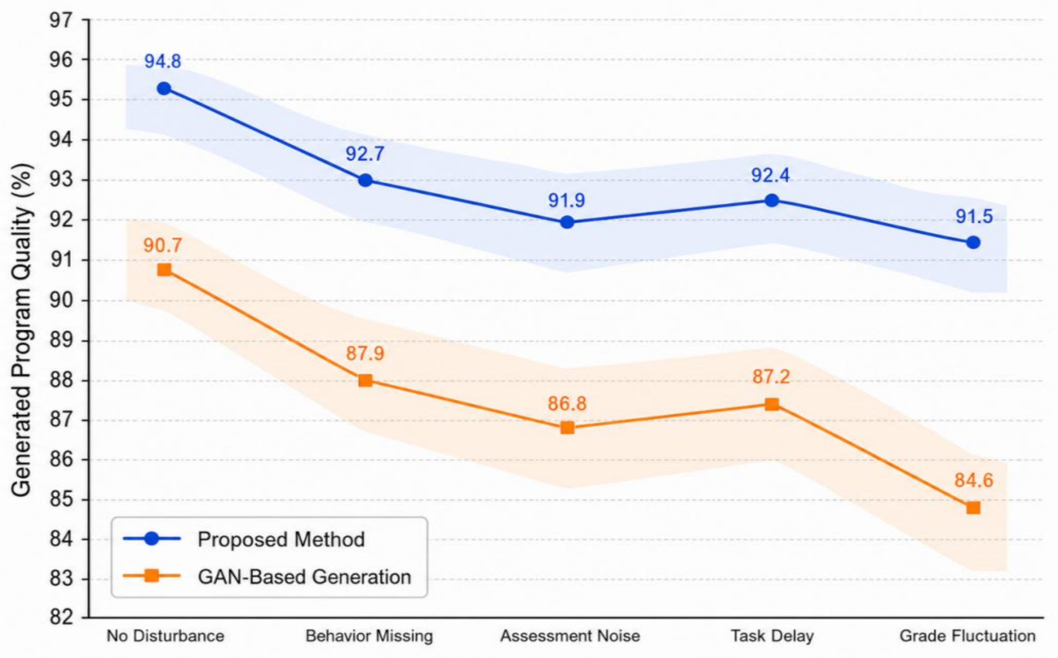

Figure 6: Line chart of quality change of generation scheme under learning behavior perturbation

It can be seen from Figure 6 that the proposed model maintains high generation quality in all kinds of disturbance scenarios. Under the condition of no disturbance, the quality of the generated scheme is 94.8%. In the scenarios of behavior loss, evaluation noise, task delay and performance fluctuation, the quality is 92.7%, 91.9%, 92.4% and 91.5%, respectively, and the maximum decline is 3.3 percentage points. In contrast, the single GAN model is 90.7%, 87.9%, 86.8%, 87.2%, and 84.6% in the same scenario, respectively, with the most obvious decrease in the performance fluctuation scenario. The results show that the cognitive diagnosis feedback and multi-constraint optimization can enhance the robustness of the model to complex learning behavior data.

5.4 Computing Resource consumption and feasibility analysis of large-scale deployment

In the generation task of personalized talent training scheme, the model should not only ensure the generation quality, but also take into account the computational overhead and deployment efficiency. In order to verify the engineering feasibility of the proposed model under the condition of large-scale student samples, this paper selects rule recommendation, knowledge tracking recommendation, GAN generation models and the proposed model for comparison. The average generation time and video memory occupation are used to characterize the computational overhead, and the efficiency of generation scheme is used to characterize the output quality. The trade-off between the computational overhead and the quality of scheme generation for different models is shown in Figure 7.

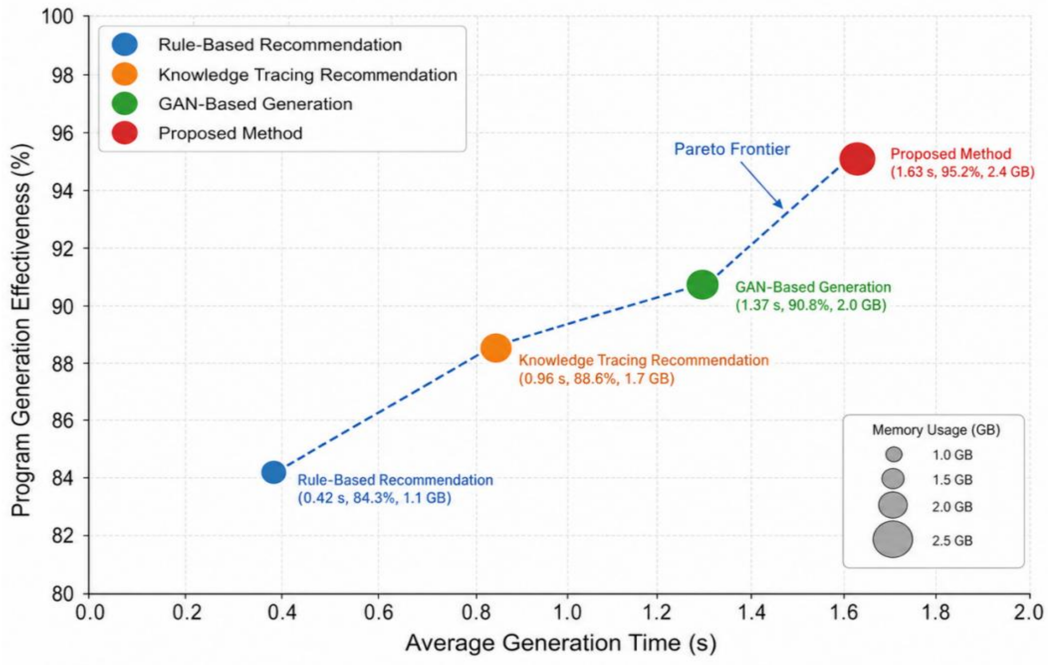

Figure 7: Pareto front plot of model computational overhead versus scheme generation quality

Figure 7 shows that although the average generation time of rule recommendation is only 0.42 s and the video memory occupation is 1.1 GB, the efficiency rate of the generation scheme is only 84.3%. The effective rate of GAN generation model is increased to 90.8%, and the average generation time is increased to 1.37 s. The average generation time of the proposed model is 1.63 s, the video memory occupation is 2.4 GB, and the efficiency of the generation scheme reaches 95.2%, which is located in the Pareto frontier region. The results show that the proposed model achieves higher scheme generation quality under the condition of moderately increasing the computational overhead, and has better large-scale deployment feasibility.

5.5 Application value analysis of model generation results in talent training decision-making

The application value of the generated results of the model is mainly reflected in the transformation of talent training decision-making from experience judgment to data support. The development of traditional training programs relies more on the overall judgment of course leaders and teaching teams, and it is easy to ignore the differences of individual students in knowledge base, learning rhythm, practical ability and career development direction. In this model, cognitive diagnosis results, students 'ability portraits, course objectives and post ability maps are introduced into the generation process, so that the training program is no longer a unified template, but can be dynamically organized around students 'real ability states.

At the level of teaching management, the generated results can help teachers identify the common shortcomings of different student groups, and provide a basis for course module adjustment, teaching content supplement and practical task configuration. At the level of student training, the model can generate differentiated learning paths according to knowledge defects and ability development needs, so that students can clarify the course content, practical projects and ability directions that need to be strengthened in the follow-up. At the professional construction level, the introduction of the post ability map enhances the connection between the training program and the industry needs, and helps to promote the continuous update of the curriculum system.

In general, the personalized talent training scheme generated by the model in this paper not only has the technical automatic generation ability, but also serves the teaching diagnosis, curriculum optimization and training path decision-making. Its value is not limited to replacing manual plans, but lies in providing interpretable, adjustable and sustainable optimization decision-making reference for teachers and managers.

6 Conclusion

Focusing on the problem of personalized talent training program generation, this paper proposes a method of collaboration between cognitive diagnosis model and generative adversarial network. The learning behavior representation was constructed through multi-source educational data cleaning, time alignment and feature normalization, and the cognitive diagnosis model was used to estimate the probability of knowledge mastery, defect intensity and defect priority to form students 'ability portraits. Then, the course goal graph, job ability graph and training constraint characteristics are fused, and the conditional generative adversarial network is used to complete the generation of training scheme, and the executability of the scheme is improved through multi-constraint loss, diagnostic feedback correction and parameter iterative update. The experimental results show that the complete model is superior to other ablation models in the knowledge defect location rate, the effective rate of generation scheme, the coverage rate of ability compensation and the matching degree of post ability, reaching 93.8%, 95.2%, 92.4%and 91.7%respectively. This method can take into account the diagnosis accuracy, generation quality and deployment efficiency, and provide intelligent reference for curriculum optimization, practical task configuration and training path decision-making.

Author's Profile

Wang Tao was born in Yantai, Shandong in November 1998. He obtained his Bachelor's degree in Laws from Shandong Normal University in June 2021 and his Master's degree in Education from Shandong Normal University in June 2023. Currently, he serves as a dedicated lecturer at the School of Marxism, Yantai Nanshan University.

Wensheng Liu was born in Nanyang, Henan in August 1983. He obtained his Bachelor's degree in Management from Zhengzhou University in July 2007 and his Master's degree in Education from Henan University in July 2010. Currently, he serves as a dedicated lecturer at the School of Marxism, Yantai Nanshan University.

Qiulin Hu was born in August 1995 in Yangquan, Shanxi Province. She graduated from Shanxi University in July 2023 with a Master's degree in Philosophy. Currently, he serves as a full-time teacher at the School of Marxism of Yantai Nanshan University.

References

[1] Nájera P, Abad F J, Sorrel M A. Determining the Number of Attributes in Cognitive Diagnosis Modeling[J]. Frontiers in Psychology, 2021, 12: 614470. DOI: 10.3389/fpsyg.2021.614470.

[2] Sorrel M A, Abad F J, Nájera P. Improving Accuracy and Usage by Correctly Selecting: The Effects of Model Selection in Cognitive Diagnosis Computerized Adaptive Testing[J]. Applied Psychological Measurement, 2021, 45(2): 112-129. DOI: 10.1177/0146621620977682.

[3] Yigit H D, Douglas J A. First-Order Learning Models With the GDINA: Estimation With the EM Algorithm and Applications[J]. Applied Psychological Measurement, 2021, 45(3): 143-158. DOI: 10.1177/0146621621990746.

[4] Yamaguchi K, Templin J. A Gibbs Sampling Algorithm with Monotonicity Constraints for Diagnostic Classification Models[J]. Journal of Classification, 2022, 39(1): 24-54. DOI: 10.1007/s00357-021-09392-7.

[5] Helm C, Warwas J, Schirmer H. Cognitive Diagnosis Models of Students’ Skill Profiles as a Basis for Adaptive Teaching: An Example from Introductory Accounting Classes[J]. Empirical Research in Vocational Education and Training, 2022, 14: 9. DOI: 10.1186/s40461-022-00137-3.

[6] Kreitchmann R S, de la Torre J, Sorrel M A, Nájera P, Abad F J. Improving Reliability Estimation in Cognitive Diagnosis Modeling[J]. Behavior Research Methods, 2023, 55(7): 3446-3460. DOI: 10.3758/s13428-022-01967-5.

[7] Olea J, Santos K C. A General Mixture Model for Cognitive Diagnosis[J]. Journal of Educational and Behavioral Statistics, 2024, 49(2): 268-307. DOI: 10.3102/10769986231176012.

[8] Im S. Targeted Assessment of Hypothesis Testing Skills Using Cognitive Diagnostic Models: Implications for Formative Practice[J]. International Journal of Educational Research, 2025, 134: 102801. DOI: 10.1016/j.ijer.2025.102801.

[9] Khamboonruang A. Investigating the Applicability of a Diagnostic Classification Model in a Thai EFL Classroom Writing Assessment: A GDINA Model Study[J]. Language Testing in Asia, 2025, 15: 68. DOI: 10.1186/s40468-025-00409-1.

[10] Kabudi T, Pappas I, Olsen D H. AI-Enabled Adaptive Learning Systems: A Systematic Mapping of the Literature[J]. Computers and Education: Artificial Intelligence, 2021, 2: 100017. DOI: 10.1016/j.caeai.2021.100017.

[11] Tahiru F. AI in Education: A Systematic Literature Review[J]. Journal of Cases on Information Technology, 2021, 23(1): 1-20. DOI: 10.4018/JCIT.2021010101.

[12] Carlon M K J, Cross J S. Knowledge Tracing for Adaptive Learning in a Metacognitive Tutor[J]. Open Education Studies, 2022, 4(1): 206-224. DOI: 10.1515/edu-2022-0013.

[13] Ghiasabadi Farahani M, Akbari Torkestani J, Rahmani M. Adaptive Personalized

Recommender System Using Learning Automata and Items Clustering[J]. Information Systems, 2022, 106: 101978. DOI: 10.1016/j.is.2021.101978.

[14] Celik I, Dindar M, Muukkonen H, Järvelä S. The Promises and Challenges of Artificial Intelligence for Teachers: A Systematic Review of Research[J]. TechTrends, 2022, 66(4): 616-630. DOI: 10.1007/s11528-022-00715-y.

[15] da Silva F L, Slodkowski B K, da Silva K K A, Cazella S C. A Systematic Literature Review on Educational Recommender Systems for Teaching and Learning: Research Trends, Limitations and Opportunities[J]. Education and Information Technologies, 2023, 28(3): 3289-3328. DOI: 10.1007/s10639-022-11341-9.

[16] Gligorea I, Cioca M, Oancea R, Gorski A T, Gorski H, Tudorache P. Adaptive Learning Using Artificial Intelligence in e-Learning: A Literature Review[J]. Education Sciences, 2023, 13(12): 1216. DOI: 10.3390/educsci13121216.

[17] Crompton H, Burke D. Artificial Intelligence in Higher Education: The State of the Field[J]. International Journal of Educational Technology in Higher Education, 2023, 20: 22. DOI: 10.1186/s41239-023-00392-8.

[18] Shum L C, Rosunally Y, Scarle S, Munir K. Personalised Learning through Context-Based Adaptation in the Serious Games with Gating Mechanism[J]. Education and Information Technologies, 2023, 28(10): 13077-13108. DOI: 10.1007/s10639-023-11695-8.

[19] Farhood H, Joudah I, Beheshti A, Muller S. Advancing Student Outcome Predictions through Generative Adversarial Networks[J]. Computers and Education: Artificial Intelligence, 2024, 7: 100293. DOI: 10.1016/j.caeai.2024.100293.

[20] Megahed M, Mohammed A. A Comprehensive Review of Generative Adversarial Networks: Fundamentals, Applications, and Challenges[J]. WIREs Computational Statistics, 2024, 16(1): e1629. DOI: 10.1002/wics.1629.

[21] Rejeb A, Rejeb K, Appolloni A, Treiblmaier H, Iranmanesh M. Exploring the Impact of ChatGPT on Education: A Web Mining and Machine Learning Approach[J]. The International Journal of Management Education, 2024, 22(1): 100932. DOI: 10.1016/j.ijme.2024.100932.

[22] Siafis V, Rangoussi M, Psaromiligkos Y. Recommender Systems for Teachers: A Systematic Literature Review of Recent Research[J]. Education Sciences, 2024, 14(7): 723. DOI: 10.3390/educsci14070723.

[23] Amin S, Uddin M I, Alarood A A, Mashwani W K, Alzahrani A O, Alzahrani H A. An Adaptable and Personalized Framework for Top-N Course Recommendations in Online Learning[J]. Scientific Reports, 2024, 14: 10382. DOI: 10.1038/s41598-024-56497-1.

[24] Lagos-Castillo A, Chiappe A, Ramirez-Montoya M S, Becerra Rodríguez D F. Mapping the Intelligent Classroom: Examining the Emergence of Personalized Learning Solutions in the Digital Age[J]. Contemporary Educational Technology, 2025, 17(1): ep543. DOI: 10.30935/cedtech/15617.