Boosting Neural Cognitive Diagnosis with Student's Affective State Modeling

Shanshan Wang[1,2], Zhen Zeng[2], Xun Yang[3*], Ke Xu[1], Xingyi Zhang[1*]

[1]Information Materials and Intelligent Sensing Laboratory of Anhui Province, Anhui University, HeFei, China
[2]Institutes of Physical Science and Information Technology, Anhui University, HeFei, China
[3]University of Science and Technology of China, HeFei, China
{wang.shanshan, xuke}@ahu.edu.cn, q21201141@stu.ahu.edu.cn, xyang21@ustc.edu.cn, xyzhanghust@gmail.com

Abstract

Cognitive Diagnosis Modeling aims to infer students' proficiency level on knowledge concepts from their response logs. Existing methods typically model students' response processes as the interaction between students and exercises or concepts based on hand-crafted or deeply-learned interaction functions. Despite their promising achievements, they fail to consider the relationship between students' cognitive states and affective states in learning, *e.g.,* the feelings of frustration, boredom, or confusion with the learning content, which is insufficient for comprehensive cognitive diagnosis in intelligent education. To fill the research gap, we propose a novel **A**ffect-aware **C**ognitive **D**iagnosis (ACD) model which can effectively diagnose the knowledge proficiency levels of students by taking into consideration the affective factors. Specifically, we first design a student affect perception module under the assumption that the affective state is jointly influenced by the student's affect trait and the difficulty of the exercise. Then, our inferred affective distribution is further used to estimate the student's subjective factors, *i.e., guessing* and *slipping*, respectively. Finally, we integrate the estimated *guessing* and *slipping* parameters with the basic neural cognitive diagnosis framework based on the DINA model, which facilitates the modeling of complex exercising interactions in a more accurate and interpretable fashion. Besides, we also extend our affect perception module in an unsupervised learning setting based on contrastive learning, thus significantly improving the compatibility of our ACD. To the best of our knowledge, we are **the first** to unify the cognition modeling and affect modeling into the same framework for student cognitive diagnosis. Extensive experiments on real-world datasets clearly demonstrate the effectiveness of our ACD. Our code is available at https://github.com/zeng-zhen/ACD.

Introduction

Cognitive Diagnosis Modeling (CDM) serves as a fundamental task in educational data mining (Anderson et al. 2014; Nguyen 2015), aiming at revealing students' proficiency levels on specific knowledge concepts based on their response logs (Lord 1952). The diagnosis results can effectively support the downstream intelligent education tasks, such as knowledge tracing (Piech et al. 2015; Nakagawa, Iwasawa, and Matsuo 2019; Shen et al. 2021), computerized adaptive testing (Zhuang et al. 2022b; Wu et al. 2020; Zhuang et al. 2022a), and recommendation systems.

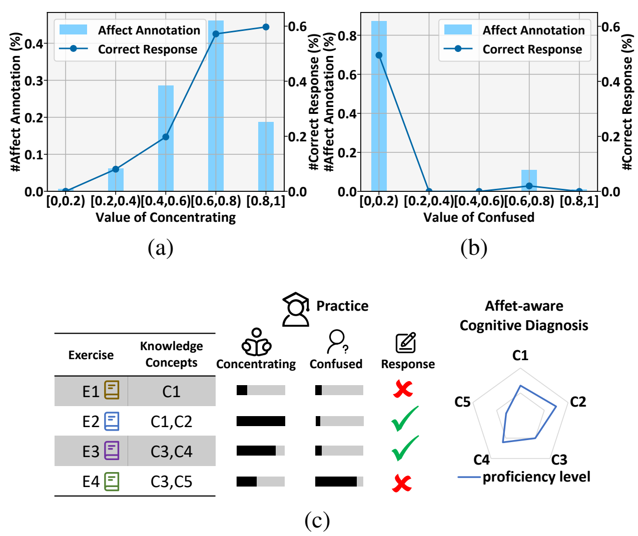


Figure 1: The statistical correlation between student-exercise cognitive response and affective state is illustrated in (a) and (b) based on the ASSIST17 dataset. A toy example of the affect-aware cognitive diagnosis is shown in (c).

The quality of CDM largely depends on the design of the interaction function that models the complex interactions between students and exercises or concepts. Early methods primarily relied on manually designed interaction models (Lord 1952; Reckase 2009). Recent methods have achieved significant performance by incorporating deep neural networks (Wang et al. 2020) and graph neural networks (Gao et al. 2021; Wang et al. 2023b) into the interaction functions. Currently, the mainstream paradigm for deep interaction functions mostly adopts the classic Item Response Theory (IRT (Lord 1952)) to model the probability of a student correctly answering an exercise based on the student learning ability and exercise difficulty. Despite their remarkable achievements, existing efforts in CDM fail to consider the affective states of students in learning. In this paper, we argue that the affective states of students is an indispens-

*Corresponding Authors.

able subjective factor for cognitive behavior analysis.

Some early studies (Pedro et al. 2013; San Pedro et al. 2014) in the field of educational data mining have investigated the relationships between affect and learning behavior in tutoring systems. They found that students who were bored or confused while answering the questions tended to do poorly on the test, and students with high levels of concentration obviously tended to make fewer mistakes. Besides, as shown in Figure 1, we illustrate the probability of student practicing exercises correctly in different *affect* dimensions based on the ASSIST17 dataset. Our observations are consistent with the aforementioned research findings, which clearly supports our argument in this paper that the affective factor should be carefully modeled into the interaction function for CDM. How to effectively perceive the students' affects and further leverage the affective cue to boost the students' cognitive diagnosis is a crucial research question, but has not been carefully studied so far in CDM.

To fill the research gap, we develop a novel and flexible Affect-aware Cognitive Diagnosis (ACD) approach which can effectively diagnose students' knowledge proficiency levels on specific knowledge concepts in a more accurate and interpretable fashion. To be specific, our proposed ACD framework mainly consists of three parts. (1) We first design a student affect perception module under the assumption that students' affective states should be influenced by not only the students' personalized affect traits but also the difficulty of exercises. It is inspired by the Flow theory in (Csikszentmihalyi 1992) stating that specific affective states emerge depending on the degree of challenge and skill that is present for an activity. (2) Then, we utilize the predicted affective distribution to infer the two important subjective factors of students, *i.e., guessing* and *slipping*, in the Deterministic Inputs, Noisy And gate (DINA) (De La Torre 2009) model, which is significantly different from original design in DINA where *guessing* and *slipping* are simply treated as two exercise-specific parameters. (3) Finally, we can easily integrate our learned guessing and slipping parameters with the estimated response score from basic neural cognitive diagnosis frameworks (Gao et al. 2021; Wang et al. 2023b) based on the DINA model. Moreover, our affect perception module can not only be optimized in supervised learning manner on datasets with auxiliary affect annotations, but also be extended in an unsupervised learning setting based on contrastive learning, thereby being easily integrated with existing CDM frameworks.

To the best of our knowledge, we are the first to unify the cognition modeling and affect modeling into the same framework for student cognitive diagnosis. Overall, our proposed ACD firstly demonstrates the great potential of studying the relationships between affect and cognition in CDM. Despite the simplicity of our implemented ACD approach, we can significantly improve the strong CDM baselines, *e.g.,* NCD (Wang et al. 2020) and RCD (Gao et al. 2021), in different settings. The main merit of this work is that we present a simple yet effective solution to exploit the complementation of DINA model and existing IRT based CDM methods, benefiting from the students' affective modeling. Our key contributions are summarized as follows:

- We propose a novel and effective affect-aware CDM approach, which clearly validates that the affective state is an indispensable subjective factor for CDM.
- We develop a plug-and-play affect perception module which can be optimized either in fully-supervised or unsupervised learning setting, showing high compatibility.
- Extensive experiments and analysis on several benchmark datasets with different CDM baselines clearly demonstrate the rationale and effectiveness of our ACD.

Preliminaries

We first briefly formulate the affect-aware CDM task. Considering that our ACD method depends on the DINA paradigm, we further describe the DINA model in detail.

Problem Definition: Let M, N, K, and Z denote the number of students, exercises, knowledge concepts, and affect labels, respectively. $S = \{s_1, \cdots, s_M\}$, $E = \{e_1, \cdots, e_N\}$, and $C = \{c_1, \cdots, c_K\}$, denote the sets of students, exercises, and knowledge concepts. Let $\mathbf{Q} = \{Q_{ij}\}_{N\times K} \in \{0,1\}_{N\times K}$ be Q-matrix that records the relationship between exercises and knowledge concepts, where $Q_{ij} = 1$ if exercise e_i relates to the concept c_j and $Q_{ij} = 0$ otherwise. The student practice records R are denoted as the set of (s, e, r), where $r \in \{0, 1\}$ represents the binary score of student s on exercise e. The corresponding affect vector of student s is denoted as $\boldsymbol{a} = \{a_1, \cdots, a_z, \cdots, a_Z\}$, where $a_z \in (0, 1)$ denotes the value of z-th affect label, *e.g.,* concentrating or confused while student s doing exercise e.

Given the students' practice logs R, the Q-matrix, and the annotation of affective state vector $\boldsymbol{a}$, the goal of our affect-aware cognitive diagnosis is to learn an effective affect-aware cognitive diagnosis interaction function $F_\theta(\cdot)$ that can jointly estimate students' mastery levels on each knowledge concept through student performance prediction and also infer students' affective states in a multi-task learning manner.

DINA Model: The DINA (De La Torre 2009) model is one of the most typical CDM theories. It introduces two exercise factors: ***slipping*** $\hat{s}$ and ***guessing*** $\hat{g}$. $\hat{s}$ denotes the probability that a student has sufficient ability but makes an incorrect response due to a slipping error for the exercise, and $\hat{g}$ denotes the probability that the student does not know how to answer the exercise but guesses correctly. In DINA, $\hat{s}$ and $\hat{g}$ are only defined as exercise-specific factors, ignoring the individual differences between students. Let η_{ij} be the mastery of student s_i on exercise e_j, which is calculated by $\eta_{ij} = \prod_k \theta_{ik}^{\beta_{jk}}$, where both θ_{ik} and β_{jk} are binary variables. θ_{ik} denotes whether the student s_j has mastered the knowledge concept c_k. β_{jk} means whether the exercise e_j contains c_k. Based on the two factors, the probability of student s_i practising exercise e_j correctly is defined as:

$$\hat{y}_{i,j} = \hat{g}_j^{1-\eta_{i,j}} (1 - \hat{s}_j)^{\eta_{i,j}}. \tag{1}$$

Methodology

We provide a comprehensive overview of our Affect-aware Cognitive Diagnosis model (ACD) based on the NCD framework (Wang et al. 2020). Next, we briefly introduce how to extend our ACD in the unsupervised setting without affect

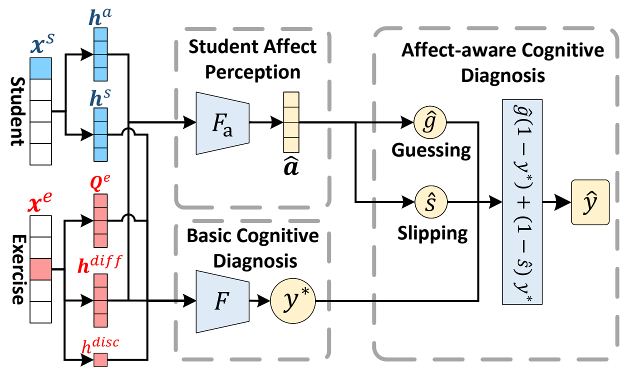


Figure 2: The pipeline of Affect-aware Cognitive Diagnosis.

labels. Note that our method can also be easily applied to other cognitive diagnosis frameworks.

Affect-aware Cognitive Diagnosis (ACD)

As shown in Figure 2, ACD can be mainly divided into three modules: student affect perception, basic cognitive diagnosis, and affect-aware cognitive diagnosis. The student affect perception module aims to predict the affects of students on specific exercises. Furthermore, through the affects, we can predict the probabilities of their guessing and slipping states. The basic cognitive diagnosis module can be an existing cognitive diagnosis model (here we present NCD for example) which utilizes the student's ability and exercise difficulty to predict the accuracy of response. Finally, the affect-aware cognitive diagnosis module combines the affect-aware guessing and slipping parameters along with the basic cognitive diagnosis results to make the final prediction of student responses based on the diagnostic formula of the DINA model.

Student Factors: In ACD, we design two factors $\boldsymbol{h}^s$ and $\boldsymbol{h}^a$ on each student. $\boldsymbol{h}^s$ denotes the mastery level of student on knowledge concepts. $\boldsymbol{h}^a$ is the latent affect trait of student. Let $\boldsymbol{x}^s$ be the one-hot vector of student, the factors $\boldsymbol{h}^s$ and $\boldsymbol{h}^a$ can be obtained by multiplying $\boldsymbol{x}^s$ and trainable matrices $\mathbf{A}$ and $\mathbf{S}$ respectively:

$$\boldsymbol{h}^s = \text{sigmoid}(\boldsymbol{x}^s \times \mathbf{S}), \tag{2}$$

$$\boldsymbol{h}^a = \text{sigmoid}(\boldsymbol{x}^s \times \mathbf{A}), \tag{3}$$

where $\boldsymbol{h}^s \in (0,1)^{1\times K}$, $\boldsymbol{h}^a \in (0,1)^{1\times d}$, $\boldsymbol{x}^s \in \{0,1\}^{1\times M}$, $\mathbf{S} \in \mathbb{R}^{M\times K}$ and $\mathbf{A} \in \mathbb{R}^{M\times d}$, and d denotes the dimension of the latent affect trait.

Exercise Factors: The exercise factors include the knowledge concepts correlation vector $\boldsymbol{Q}^e$, which represents the knowledge concepts involved in each exercise, the exercise difficulty $\boldsymbol{h}^{\text{diff}}$, and the exercise discrimination h^{disc}. Let $\boldsymbol{x}^e$ be the one-hot vector of exercise, and $\boldsymbol{Q}^e$ can be computed by $\boldsymbol{Q}^e_j = \boldsymbol{x}^e \times \mathbf{Q}$, where $\boldsymbol{Q}^e \in \{0,1\}^{1\times K}$, $\boldsymbol{x}^e \in \{0,1\}^{1\times N}$. $\boldsymbol{h}^{\text{diff}}$ represents the difficulty of the exercise on each knowledge concept. And h^{disc} reflects the exercise's capacity to discriminate between students exhibiting high and low mastery levels of knowledge concepts. We calculate them by:

$$\boldsymbol{h}^{\text{diff}} = \text{sigmoid}(\boldsymbol{x}^e \times \mathbf{E}), \tag{4}$$

$$h^{\text{disc}} = \text{sigmoid}(\boldsymbol{x}^e \times \mathbf{D}), \tag{5}$$

where $\boldsymbol{h}^{\text{diff}} \in (0,1)^{1\times K}$, $h^{\text{disc}} \in (0,1)$. $\mathbf{E} \in \mathbb{R}^{N\times K}$ and $\mathbf{D} \in \mathbb{R}^{N\times 1}$ are trainable matrices.

Student Affect Perception: This module is new-designed in our method. Considering that when the student is engaged in exercise-solving, the affect is influenced by both the affect latent trait and the difficulty of the exercise. Different students may exhibit different affects when facing the same exercise. Even the same student would exhibit different affects when facing different difficult exercises. We utilize $\boldsymbol{h}^a$ and $\boldsymbol{h}^{\text{diff}}$ to predict the affect of student facing exercise by a fully connected layer:

$$\hat{\boldsymbol{a}} = \text{sigmoid}(\mathbf{W}_a \times [\boldsymbol{h}^a, \boldsymbol{h}^{\text{diff}}]^T + \boldsymbol{b}_a), \tag{6}$$

where $\hat{\boldsymbol{a}} \in (0,1)^Z$ contains the predicted value of each affect. $\mathbf{W}_a \in \mathbb{R}^{Z\times(d+K)}$ represents the weight matrix and $\boldsymbol{b}_a$ is bias. $[\cdot]$ denotes the concatenation operation.

Since the affects of students is labeled as continuous values ranging from 0 to 1, we choose the mean square error (MSE) loss function for the affect perception module:

$$\mathcal{L}_{\text{a}} = \sum_i ||\hat{\boldsymbol{a}}^i - \boldsymbol{a}^i_{\text{gt}}||^2, \tag{7}$$

where $\boldsymbol{a}^i_{\text{gt}}$ denotes of the annotated affect vector of i-th student-exercise interaction. The affect loss $\mathcal{L}_{\text{a}}$ is finally averaged over the mini-batch. Note that when the affect annotation is missing in the training data, the student affect perception module in ACD can be optimized by unsupervised learning, which will be detailed later.

Basic Cognitive Diagnosis: We use the NCD (Wang et al. 2020) framework as our basic cognitive diagnostic module to implement ACD for its simplicity and effectiveness. The student response score y^* in NCD is predicted by an interaction function composed of multiple fully connected layers. Its input is formulated as follows:

$$\boldsymbol{f}_0 = \boldsymbol{Q}^e \circ (\boldsymbol{h}^s - \boldsymbol{h}^{\text{diff}}) \times h^{\text{disc}}, \tag{8}$$

where $\circ$ is the element-wise product. The output layer in NCD is formulated as follows:

$$y^* = \text{sigmoid}(\mathbf{W}_n \times \boldsymbol{f}_{n-1} + b_n). \tag{9}$$

The weight matrices of all layers are restricted to positive values to satisfy the monotonicity assumption.

Affect-aware Cognitive Diagnosis: Intuitively, the performance of the student is often correlated with his affect during the exercise-solving process. For example, even if a student possesses the capability to answer an exercise correctly, a low level of concentration during interaction could still lead to an incorrect response. To fit the influence of affect on the interaction process, we model the probabilities of

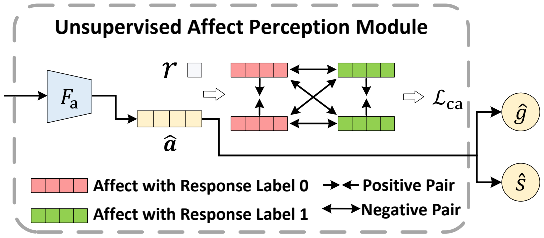


Figure 3: Unsupervised affect perception module.

student's guessing and slipping based on the estimated affect distribution, which are as follows:

$$\hat{g} = \text{sigmoid}(\mathbf{W}_g \times \hat{\boldsymbol{a}} + b_g), \quad (10)$$
$$\hat{s} = \text{sigmoid}(\mathbf{W}_s \times \hat{\boldsymbol{a}} + b_s), \quad (11)$$

where $\mathbf{W}_g, \mathbf{W}_s \in \mathbb{R}^{1\times Z}$ denote the weight matrices of guessing and slipping, respectively, and b_g, b_s are bias terms. $\hat{g} \in (0, 1)$ represents the probability that the student answers the exercise correctly due to guessing, and $\hat{s} \in (0, 1)$ represents the probability that the student makes a mistake on exercise due to slipping. By combining $\hat{g}$, $\hat{s}$, and y^*, we adopt the well-known DINA diagnostic formula to infer the affect-aware student response score $\hat{y}$ as follows:

$$\hat{y} = \hat{g}(1 - y^*) + (1 - \hat{s})y^*, \quad (12)$$

where the DINA formula in Eq. (12) is a reformulated version in FuzzyCDM (Liu et al. 2018) which is more suitable for predicting student's performance using neural networks.

Inspired by DINA, the first part in Eq. (12) represents the probability that student s doesn't know how to solve exercise e but guessed correctly, and the second part represents the probability that, based on the student's capability, he should have answered the exercise correctly but made a mistake. Given that students' response labels r are binary (0 for incorrect and 1 for correct), the binary cross-entropy loss is leveraged as the loss function for cognitive diagnosis:

$$\mathcal{L}_{\text{CDM}} = -\sum_i (r_i \log \hat{y}_i + (1 - r_i) \log(1 - \hat{y}_i)). \quad (13)$$

Training: In our approach, we jointly train the two losses in Eq. (7) and Eq. (13) in our ACD. The student affect perception module aims to predict the student's affective distribution based on the student's latent affect trait $\boldsymbol{h}^a$ and exercise difficulty $\boldsymbol{h}^{\text{diff}}$. The predicted affect provides assistance to the basic cognitive diagnosis. Specifically, the affect-aware cognitive diagnosis module utilizes affect-related parameters to predict students' responses and diagnose their levels of knowledge concept mastery. Synchronously, the accuracy of cognitive diagnosis results could also influence the student affect latent traits. We optimize the two losses simultaneously by employing a joint loss function:

$$\mathcal{L} = \mathcal{L}_{\text{CDM}} + \lambda \mathcal{L}_{\text{a}}, \quad (14)$$

where λ denotes the trade-off parameter for the affect loss.

Dataset	ASSIST17	ASSIST12	Junyi
Students	1709	27633	10000
Exercises	3162	53086	835
Knowledge concepts	102	265	835
Response records	311569	2013626	220799
AVG#score	0.4365	0.6971	0.6516

Table 1: Statistics of experimental datasets.

Unsupervised Contrastive ACD Model (CACD)

In fact, not all the datasets have the affect labels. When there are no available affect labels, it is not-trivial to optimize the student affect perception module. Considering the statistical correlation between the students' affects and their cognitive response results illustrated in Figure 1 (a) and (b), we utilize the contrastive learning strategy to design an unsupervised affect perception module, as illustrated in Figure 3, to replace the student affect perception module shown in Figure 2. We refer to this new designed model suitable for datasets without affect labels as Contrastive ACD (CACD). Here we assume that positive student-exercise interactions with the response results more likely correspond to the same affective state of students, while negative student-exercise interactions more likely correspond to different affective states. Under this assumption, we can collect sufficient positive pairs and negative pairs for contrastive learning to optimize the affect vector $\hat{\boldsymbol{a}}$ in Eq. (6). To replace $\mathcal{L}_{\text{a}}$, the contrastive affect perception loss $\mathcal{L}_{\text{ca}}$ is formulated as follows:

$$\mathcal{L}_{\text{ca}} = -\sum_i \frac{1}{|\mathcal{P}_i|} \sum_{j \in \mathcal{P}_i} \log \frac{\exp(\text{sim}(\hat{\boldsymbol{a}}^i, \hat{\boldsymbol{a}}^j)/\tau)}{\sum\limits_{k \in \mathcal{N}_i} \exp(\text{sim}(\hat{\boldsymbol{a}}^i, \hat{\boldsymbol{a}}^k)/\tau)}, \quad (15)$$

where $\mathcal{P}_i$ denotes the set of positive pairs constructed for i-th student-exercise interaction, and $\mathcal{N}_i$ denotes the set of negative pairs constructed for i-th interaction. $\text{sim}(\cdot, \cdot)$ denotes the cosine similarity between the two estimated affect vectors, and τ is the temperature of contrastive loss.

Remarks Our approach can be regarded as a plug-and-play module which is added to existing cognitive diagnostic models to improve the performance.

Experiments

To demonstrate the generalization and effectiveness of affect perception in ACD, we first compare the model incorporating affect perception with its baselines. Then we will analyze and interpret the models.

Datasets: In order to verify the generalization, we evaluate our model on three real datasets, in which two have affect labels, while one does not. A brief overview of the datasets is described as follows:

- **ASSIST17**[1] is collected for ASSISTments Longitudinal Data Mining Competition in 2017.
- **ASSIST12**[2] is the data for the school year 2012-2013 with affect. The affect data was extracted by researchers

[1]https://sites.google.com/view/assistmentsdatamining/dataset
[2]https://sites.google.com/site/assistmentsdata/2012-13-school-data-with-affect

Datasets		ASSIST17			ASSIST12		
Models		ACC	RMSE	AUC	ACC	RMSE	AUC
DINA	baseline	64.84±0.09	46.66±0.02	69.64±0.06	71.45±0.07	43.80±0.04	69.49±0.11
	ACD	**71.15±0.26**	**43.62±0.18**	**77.72±0.29**	**73.76±0.08**	**42.12±0.05**	**74.21±0.06**
IRT	baseline	65.96±0.10	46.53±0.03	72.37±0.07	73.11±0.04	42.67±0.01	72.67±0.03
	ACD	**71.64±0.23**	**43.25±0.10**	**78.55±0.22**	**74.26±0.07**	**41.71±0.02**	**75.44±0.07**
MIRT	baseline	68.17±0.02	46.48±0.03	74.13±0.00	73.81±0.00	44.44±0.00	72.58±0.00
	ACD	**72.43±0.13**	**42.78±0.07**	**79.46±0.13**	**74.38±0.07**	**41.63±0.06**	**75.64±0.11**
NCD	baseline	69.21±0.87	45.13±0.60	75.34±1.01	74.23±0.05	41.95±0.01	74.78±0.04
	ACD	**72.42±0.10**	**42.84±0.07**	**79.49±0.10**	**74.98±0.05**	**41.37±0.06**	**76.02±0.16**
RCD	baseline	71.55±0.15	43.39±0.10	78.10±0.03	74.49±0.04	41.75±0.03	75.22±0.05
	ACD	**72.31±0.10**	**42.84±0.05**	**79.27±0.08**	**74.61±0.02**	**41.43±0.00**	**76.19±0.01**
SCD	baseline	71.59±0.10	43.35±0.05	78.19±0.01	74.64±0.03	41.48±0.02	75.94±0.04
	ACD	**72.69±0.05**	**42.79±0.02**	**79.40±0.07**	**74.70±0.01**	**41.37±0.00**	**76.10±0.02**

Table 2: Experimental results on student performance prediction in percentage.

from student logs (Wang, Heffernan, and Heffernan 2015; Botelho, Baker, and Heffernan 2017). This dataset has been widely used in research related to affect.

- **junyi**[3] is collected from the Chinese e-learning website Junyi Academy (Chang, Hsu, and Chen 2015). This dataset is widely used in CDM.

Both ASSIST17 and ASSIST12 include four types of affect: *bored*, *concentrating*, *confused*, and *frustrated*. Following the setting in (Gao et al. 2021), we only retain the initial response log of students, and exclude students with less than 15 response records. The processed data statistics are presented in Table 1. 80% data of each student is leveraged for training and the remaining 20% for testing.

Baselines: We use the following baselines for experiments:

- DINA (De La Torre 2009) modeled the student and exercise factors as binary vectors, incorporating guessing and slipping parameters to estimate the performance.
- IRT (Lord 1952) modeled students and exercises as unidimensional traits and utilized the logistic model to represent their interactions.
- MIRT (Reckase 2009) extended the traits of students and exercises in IRT to multi-dimension.
- NCD (Wang et al. 2020) leveraged the neural network to model the interactions between students and exercises.
- RCD (Gao et al. 2021) modeled the relationships between knowledge concepts and introduced GCN.
- SCD (Wang et al. 2023b) utilized self-supervised graph learning to address the long-tailed problem in CDM.

Experimental Settings: The *Prediction Accuracy* (ACC), *Root Mean Square Error* (RMSE), and *Area Under an ROC Curve* (AUC) are selected as evaluation metrics to evaluate the results in our method. For fairness, we use the same hyperparameter settings for all models. We set the coefficient λ for the affect prediction loss in Eq. 14 to 1 to avoid over adjusting. and the dimension of the affect latent traits for students, denoted as d, to 128 for good performance. The

[3]https://pslcdatashop.web.cmu.edu/DatasetInfo?datasetId=1198

Dataset		junyi		
Models		ACC	RMSE	AUC
DINA	baseline	74.22	41.76	78.72
	CACD	**76.96**	**40.05**	**82.18**
IRT	baseline	67.60	42.68	77.50
	CACD	**77.17**	**39.76**	**82.64**
MIRT	baseline	75.13	41.17	79.89
	CACD	**77.27**	**39.73**	**82.64**
NCD	baseline	74.43	41.72	79.09
	CACD	**77.35**	**39.70**	**82.73**
RCD	baseline	77.16	39.63	82.62
	CACD	**77.42**	**39.56**	**82.94**
SCD	baseline	77.30	39.61	82.77
	CACD	**77.45**	**39.59**	**82.90**

Table 3: The results on the dataset without affect labels.

network parameters are initialized using Xavier initialization following the (Wang et al. 2020). All the weights are sampled from $\mathcal{N}\left(0, \frac{2}{n_{\text{in}}+n_{\text{out}}}\right)$, where n_{in} refers to the input dimensionality of a layer, and n_{out} refers to the output dimensionality. We implement all the baselines and the ACD version by PyTorch. The experiments are conducted on the Intel Core i9-10900X CPU and a GeForce RTX 3090 GPU.

Experimental Results and Analysis

Performance Comparison: Table 2 shows the performance comparison on datasets with affect labels. The error bars after '±' represent the standard deviations of 5 evaluation runs for each model. From the table, it can be observed that all the methods incorporating our affect perception outperform their respective baselines. The improvements are particularly remarkable for the earlier methods. The superior performance of NCD (ACD) and MIRT (ACD) over the latest baselines suggests that introducing affect perception brings greater benefits than improving the diagnostic function. Compared with RCD and SCD, RCD (ACD) and SCD (ACD) show relatively smaller improvements. This is because the graph neural networks employed in SCD and RCD, although enhancing performance, could weaken the discriminative power of features, which conflicts with our intention of extracting personalized parameters from

Models	Student Performance			Affect	
	ACC	RMSE	AUC	RMSE	MAE
NCD	69.21	45.13	75.34	45.34	36.39
ACD-w/o-L	71.54	43.37	78.40	42.48	33.12
CACD	71.76	43.25	78.54	32.41	23.28
ACD	72.42	42.84	79.49	22.10	14.64
Oracle	**79.84**	**40.27**	**80.35**	**0**	**0**

Table 4: The results of affect prediction on ASSIST17.

student affect conception		ACD	RMSE	AUC
student	exercise			
✔		63.54	47.19	67.83
	✔	65.51	46.35	70.59
✔	✔	**72.42**	**42.84**	**79.27**

Table 5: The results of ablation study on ASSIST17.

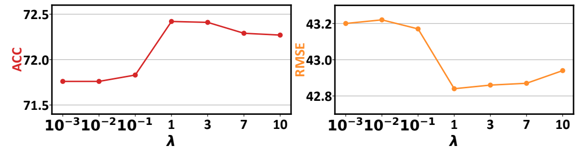


Figure 4: Performance on different drade-off parameter λ.

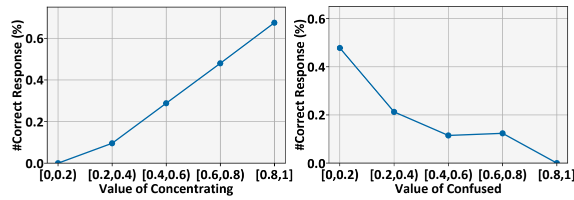


Figure 5: The impact of predicted affect on response results.

affect. Nevertheless, the introduction of affect perception still leads to improvement. Table 3 shows the results on the dataset without affect labels. The CACD indicates the use of contrastive loss in affect perception. Even in the absence of affect labels, our approach still achieved large improvements. This demonstrates the effectiveness of affect attributes. Without loss of generality, we opt to consider NCD as the basic cognitive diagnosis module of ACD in the context of our subsequent analysis.

Reliability of Affect Perception: To explain the improvements brought by the affect perception, we evaluate the accuracy of the predicted affect on ASSIST17. As the affect prediction can be regarded as the regression task, besides the conventional metrics of ACC, RMSE and AUC on student performance prediction, we use RMSE and *Mean Absolute Error* (MAE) as evaluation metrics on affect prediction. In comparison, in addition to the baseline model NCD, our model ACD and CACD, we designed two variants, ACD-w/o-L and "Oracle". ACD-w/o-L represents the variant where no affect prediction loss is constrained. "Oracle" means directly extracting personalized parameters from the affect labels. Noteworthily, due to using the test set affect labels as input in Oracle, it is actually an upper bounded model and it is designed only for comparison. In NCD, as there is no affect prediction involved, random affect was used instead. The results are shown in Table 4. It is obvious that a more improved accuracy in affect prediction corresponds to a more accurate cognitive diagnostic results. This suggests that the integration of student affect prediction and cognitive diagnosis is meaningful and effective, and our student affect prediction module can accurately predict affects. ACD-w/o-L exhibits a significant improvement compared with NCD, indicating that the personalized student parameters incorporated into our designed interaction function are effective. While CACD shows marginal improvement in student performance prediction compared with ACD-w/o-L, it demonstrates a considerable enhancement in affect prediction results. This suggests that CACD has stronger interpretability, which aligns with the requirements of CDM.

Ablation Study on Student Affect Perception: In ACD, we undertake the prediction of students' affects during exercise-solving based on two critical factors: the latent affect traits of students and the difficulty of exercises. To verify the essentiality of simultaneously considering both two factors, we conducted a comparative study by incorporating only one of these factors within the student affect perception module. The results are shown in Table 5. When only considering exercise difficulty, the model is similar to the DINA model and the model's performance is superior to that achieved by solely considering student affect traits. This indicates that different students might experience similar affect while attempting exercises of the same difficulty. When simultaneously incorporating both factors, *i.e.* ACD, a significant enhancement in performance is achieved. This substantiates our hypothesis that students' affects during exercise-answering is concurrently correlated with both student affect traits and exercise difficulty.

Hyper-parameter Analysis: λ is the trade-off parameter in Eq. (14). We set it from 0.001 to 10. As shown in Figure 4, as the value of λ gradually increases, the performance of ACD improves due to more accurate affect predictions and the optimal performance is achieved when λ is set to 1. When λ becomes excessively large, it would lead to descent in performance due to the neglect of crucial CDM.

Correlation Between Affect and Response Results: As shown in Figure 1, there is a certain correlation between students' affects and their response. In order to verify the interpretability of our approach, *i.e.*, whether the predicted affects reflect the aforementioned correlation, we computed the distribution between the predicted affects and response logs. We present the correct response probability on different affects *concentrating* and *confused* predicted by ACD on ASSIST17 in Figure 5. The Figure 5 demonstrates that higher levels of *concentrating* have a positive impact on response quality, whereas increased *confused* often lead to unsuccessful responses. The impact of predicted affects on response results remains consistent with the ground truth. Comparing Figure 1 and Figure 5, it is obvious that there are

Exercise	#1	#2	#3	#4	#5	#6
Concept	A	B	B	C	D	E
Student#1	✔	✔	✘	✘	✘	✔
Student#2	✘	✘	✔	✘	✔	✔

Table 6: Response logs of two students in ASSIST12.

	Student#1		Student#2	
	Ex-#2	Ex-#3	Ex-#2	Ex-#3
Frustrated	0.3	0.3	0.3	0.5
Confused	0	0	0	0.6
Concentrating	0.7	0.1	0.7	0.7
Bored	0.4	0.7	0.1	0.3
guessing	-	-	-	✔
slipping	-	✔	-	-

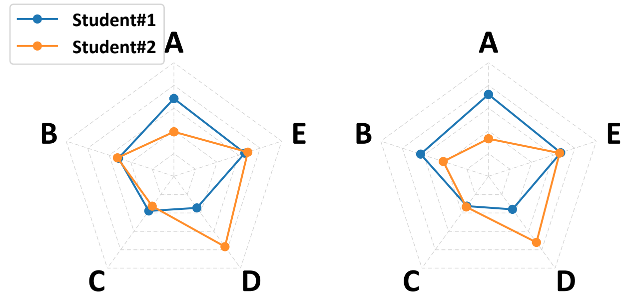
Table 7: The affect of students and ACD's prediction of guessing and slipping probability. Instances with predictive outcomes exceeding 0.5 are considered to have occurred and are denoted by ✔. Note that "Ex-" refers to "Exercise".

some disparities between the predicted distribution of affect and the ground truth, the reason may be that the responses are influenced by a combination of various affects and student abilities rather than single affect.

Case Study: In order to conduct a more comprehensive analysis of the role of affect, here we present a diagnosis example of NCD and ACD on ASSIST12. Table 6 displays a part of the response logs. Figure 6 illustrates the diagnostic results of NCD and ACD. Both models show similar diagnostic results for the two students across all concepts except for concept B. However, in the context of concept B, NCD did not provide interpretable diagnostic results. Student#1 answered exercise#2 correctly but failed on exercise#3, while student#2 exhibited the opposite pattern. The diagnostic results from NCD indicate a similar proficiency for concept B of both students, which fails to provide a satisfactory explanation for this phenomenon. In fact, all CDMs based on the IRT paradigm are confronted with this issue. Different form them, our ACD takes students' affects into account. Table 7 presents not only the labels of affect, but also the prediction of ACD on guessing and slipping probability based on students' affects. For student#1, when answering exercise#3, ACD concludes that although student#1 answered exercise#3 incorrectly, he actually possesses a solid understanding of concept B but made a slipping. In contrast, student#2, even though he answered exercise#3 correctly, ACD suggests that student#2 might have guessed the answer without a true grasp of concept B. As a result, student#1 possesses a higher proficiency on concept B compared to student#2.

Related Work

Cognitive Diagnosis Modeling: Cognitive diagnosis modeling (CDM) is a fundamental task in intelligent education. Existing work primarily derives from IRT (Lord 1952) and DINA (De La Torre 2009) paradigms. IRT (Lord 1952) aimed to project the latent features of students and items and predict the performance with a manually designed function.



(a) NCD (b) ACD

Figure 6: The diagnosis results of NCD and ACD.

It is the basic model in CDM. Then, MIRT (Reckase 2009) expanded the features of the IRT (Lord 1952) model into multidimensional vectors to enhance its expressive power. NeuralCD (Wang et al. 2020) replaced the manually designed function with neural networks to fit the interactions between students and items. RCD (Gao et al. 2021) pioneered the application of graph networks in cognitive diagnosis and modeled the relationships between knowledge concepts. Lately, SCD (Wang et al. 2023b) enhanced the model's performance on long-tailed problem. As another foundational paradigm, DINA (De La Torre 2009) modeled the student and exercise factors as binary vectors, incorporating guessing and slipping parameters to estimate the performance. However, existing works based on the DINA (De La Torre 2009) model (*e.g.*, FuzzyCDM (Liu et al. 2018)) considered guessing and slipping parameters as factors associated with the exercises, ignoring the personalized response patterns of individual students.

Student Affect Related Research: Affect as the expression of personalized states is involved in our method. Existing works (San Pedro et al. 2013; Ocumpaugh et al. 2014; Wang, Heffernan, and Heffernan 2015) primarily focused on detecting affect states during students' exercise-solving processes based on their interaction logs. (Botelho, Baker, and Heffernan 2017) refitted the detectors with deep learning. However, existing works mainly focused on utilizing the interaction logs between students and exercises to label the affect. There has been no work specifically addressing affect perceptive cognitive diagnosis.

Conclusion

This paper presents an affect-aware cognitive diagnosis (ACD) model. Specifically, we design a student affect perception module to predict the affects exhibited by students during exercise solving and extract personalized interaction parameters from the predicted affects to enhance the interaction process in CDM. We then introduce a contrastive loss based on response results to extend the model to datasets without affect labels. Extensive experiments validate the effectiveness and generality of our model. In the future, we will integrate more techniques, such as domain generalization (Wang et al. 2023a; Chang et al. 2023) or federated learning (Dai et al. 2023) to enhance our ACD model.

Acknowledgments

This work is supported by National Natural Science Foundation of China (Grant No. 62106003, 62206003, 62272435, U22A2094, U21A20512), National Key Research and Development Project (NO. 2018AAA0100105), and the University Synergy Innovation Program of Anhui Province (Grant GXXT-2022-047).

References

Anderson, A.; Huttenlocher, D.; Kleinberg, J.; and Leskovec, J. 2014. Engaging with massive online courses. In *Proc. Int. Conf. World Wide Web, WWW*, 687–698.

Botelho, A. F.; Baker, R. S.; and Heffernan, N. T. 2017. Improving sensor-free affect detection using deep learning. In *Artificial Intelligence in Education*, 40–51. Springer.

Chang, H.-S.; Hsu, H.-J.; and Chen, K.-T. 2015. Modeling Exercise Relationships in E-Learning: A Unified Approach. In *EDM*, 532–535.

Chang, T.; Yang, X.; Zhang, T.; and Wang, M. 2023. Domain Generalized Stereo Matching via Hierarchical Visual Transformation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 9559–9568.

Csikszentmihalyi, M.; and Csikszentmihalyi, I. S. 1992. *Optimal experience: Psychological studies of flow in consciousness*. Cambridge university press.

Dai, R.; Yang, X.; Sun, Y.; Shen, L.; Tian, X.; Wang, M.; and Zhang, Y. 2023. FedGAMMA: Federated Learning With Global Sharpness-Aware Minimization. *IEEE Transactions on Neural Networks and Learning Systems*.

De La Torre, J. 2009. DINA model and parameter estimation: A didactic. *Journal of educational and behavioral statistics*, 34(1): 115–130.

Gao, W.; Liu, Q.; Huang, Z.; Yin, Y.; Bi, H.; Wang, M.-C.; Ma, J.; Wang, S.; and Su, Y. 2021. Rcd: Relation map driven cognitive diagnosis for intelligent education systems. In *SIGIR - Proc. Int. ACM SIGIR Conf. Res. Dev. Inf. Retr.*, 501–510.

Liu, Q.; Wu, R.; Chen, E.; Xu, G.; Su, Y.; Chen, Z.; and Hu, G. 2018. Fuzzy cognitive diagnosis for modelling examinee performance. *ACM Trans. Intell. Syst. Technol.*, 9(4): 1–26.

Lord, F. 1952. A theory of test scores. *Psychometric monographs*.

Nakagawa, H.; Iwasawa, Y.; and Matsuo, Y. 2019. Graph-based knowledge tracing: modeling student proficiency using graph neural network. In *IEEE/WIC/ACM International Conference on Web Intelligence*, 156–163.

Nguyen, T. 2015. The effectiveness of online learning: Beyond no significant difference and future horizons. *MERLOT Journal of online learning and teaching*, 11(2): 309–319.

Ocumpaugh, J.; Baker, R.; Gowda, S.; Heffernan, N.; and Heffernan, C. 2014. Population validity for educational data mining models: A case study in affect detection. *British Journal of Educational Technology*, 45(3): 487–501.

Pedro, M. O.; Baker, R.; Bowers, A.; and Heffernan, N. 2013. Predicting college enrollment from student interaction with an intelligent tutoring system in middle school. In *EDM*.

Piech, C.; Bassen, J.; Huang, J.; Ganguli, S.; Sahami, M.; Guibas, L. J.; and Sohl-Dickstein, J. 2015. Deep knowledge tracing. *Advances in neural information processing systems*, 28.

Reckase, M. D. 2009. Multidimensional item response theory models. In *Multidimensional item response theory*, 79–112. Springer.

San Pedro, M. O.; Ocumpaugh, J.; Baker, R. S.; and Heffernan, N. T. 2014. Predicting STEM and Non-STEM College Major Enrollment from Middle School Interaction with Mathematics Educational Software. In *EDM*, 276–279.

San Pedro, M. O. Z.; Baker, R. S. d.; Gowda, S. M.; and Heffernan, N. T. 2013. Towards an understanding of affect and knowledge from student interaction with an intelligent tutoring system. In *Artificial Intelligence in Education*, 41–50. Springer.

Shen, S.; Liu, Q.; Chen, E.; Huang, Z.; Huang, W.; Yin, Y.; Su, Y.; and Wang, S. 2021. Learning process-consistent knowledge tracing. In *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, 1452–1460.

Wang, F.; Liu, Q.; Chen, E.; Huang, Z.; Chen, Y.; Yin, Y.; Huang, Z.; and Wang, S. 2020. Neural cognitive diagnosis for intelligent education systems. In *Proc. AAAI Conf. Artif. Intell., AAAI*, volume 34, 6153–6161.

Wang, S.; Chen, Y.; He, Z.; Yang, X.; Wang, M.; You, Q.; and Zhang, X. 2023a. Disentangled Representation Learning with Causality for Unsupervised Domain Adaptation. In *Proceedings of the 31st ACM International Conference on Multimedia*, 2918–2926.

Wang, S.; Zeng, Z.; Yang, X.; and Zhang, X. 2023b. Self-supervised Graph Learning for Long-tailed Cognitive Diagnosis. In *Proc. AAAI Conf. Artif. Intell., AAAI*, volume 37(1), 110–118.

Wang, Y.; Heffernan, N. T.; and Heffernan, C. 2015. Towards better affect detectors: effect of missing skills, class features and common wrong answers. In *Proceedings of the fifth international conference on learning analytics and knowledge*, 31–35.

Wu, Z.; Li, M.; Tang, Y.; and Liang, Q. 2020. Exercise recommendation based on knowledge concept prediction. *Knowledge-Based Systems*, 210: 106481.

Zhuang, Y.; Liu, Q.; Huang, Z.; Li, Z.; Jin, B.; Bi, H.; Chen, E.; and Wang, S. 2022a. A Robust Computerized Adaptive Testing Approach in Educational Question Retrieval. In *SIGIR - Proc. Int. ACM SIGIR Conf. Res. Dev. Inf. Retr.*, 416–426.

Zhuang, Y.; Liu, Q.; Huang, Z.; Li, Z.; Shen, S.; and Ma, H. 2022b. Fully adaptive framework: Neural computerized adaptive testing for online education. In *Proc. AAAI Conf. Artif. Intell., AAAI*, volume 36(4), 4734–4742.