Modeling Context-aware Features for Cognitive Diagnosis in Student Learning

Yuqiang Zhou[1], Qi Liu[1,*], Jinze Wu[1], Fei Wang[1], Zhenya Huang[1],
Wei Tong[1], Hui Xiong[2], Enhong Chen[1], Jianhui Ma[1]
[1]Anhui Province Key Laboratory of Big Data Analysis and Application, School of Computer Science and Technology, University of Science and Technology of China,
{qiliuql,huangzhy,cheneh,jianhui}@ustc.edu.cn; {zyq1998,hxwjz,wf314159,tongustc}@mail.ustc.edu.cn;
[2]the State University of New Jersey, hxiong@rutgers.edu

ABSTRACT

The contexts and cultures have a direct impact on student learning by affecting student's implicit cognitive states, such as the preference and the proficiency on specific knowledge. Motivated by the success of context-aware modeling in various fields, such as recommender systems, in this paper, we propose to study how to model context-aware features and adapt them for more precisely diagnosing student's knowledge proficiency. Specifically, by analyzing the characteristics of educational contexts, we design a two-stage framework ECD (*E*ducational context-aware *C*ognitive *D*iagnosis), where a hierarchical attentive network is first proposed to represent the context impact on students and then an adaptive optimization is used to achieve diagnosis enhancement by aggregating the cognitive states reflected from both educational contexts and students' historical learning records. Moreover, we give three implementations of general ECD framework following the typical cognitive diagnosis solutions. Finally, we conduct extensive experiments on nearly 52 million records of the students sampled by PISA (Programme for International Student Assessment) from 73 countries and regions. The experimental results not only prove that ECD is more effective in student performance prediction since it can well capture the impact from educational contexts to students' cognitive states, but also give some interesting discoveries regarding the difference among different educational contexts in different countries and regions.

CCS CONCEPTS

• **Computing methodologies → Artificial intelligence**; • **Applied computing → Education**.

KEYWORDS

Cognitive diagnosis; learning process; educational context

ACM Reference Format:

Yuqiang Zhou, Qi Liu, Jinze Wu, Fei Wang, Zhenya Huang, Wei Tong, Hui Xiong, Enhong Chen, Jianhui Ma. 2021. Modeling Context- aware Features for Cognitive Diagnosis in Student Learning. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '21), August 14-18, 2021, Virtual Event, Singapore.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3447548.3467264.

1 INTRODUCTION

It is a common notion that context features usually influence people's implicit states to make a difference in their explicit behaviors [12]. For instance, contexts are essential to infer customers' preference on specific items and redirect the consumption behaviors [33]. Similarly, in the domain of education, a better modeling in the educational contexts (e.g., those features from school, home or person [2, 16]) is also of significant importance to understanding students' learning process [24], analyzing their knowledge proficiency and further helping improve equity in education [28].

Indeed, governments may wonder if the education policy [14] for education reform works, teachers may be curious about the effect of different teaching methods in class [29], and parents may be interested in the influence of their involvement in children's study [25]. As shown in Figure 1, in students' learning process, the educational contexts influence students' traits which then reflect in the cognitive states, and finally result in the difference between their performance (i.e., response results). Since diagnosing the cognitive states of each student (more specifically, quantifying the proficiency level of the student on specific knowledge concepts, e.g., *Carbon Dioxide*, ranging from 0 to 1) is one of the most fundamental tasks in intelligent education [23, 36], a number of existing methods have tried to improve the accuracy of diagnostic results by fully exploiting the students' explicit response records (e.g., introducing the difficulty of exercise [6], the related knowledge concepts [7, 30, 35], the exercise texts [4, 36]), and the exercise relations [38] or considering more exceptions (e.g., slip and guess [5, 21]) in students' learning process. However, to the best of our knowledge, the problem of how educational contexts affect student's knowledge proficiency is still underexplored.

In this paper, to quantify the influence from educational contexts to students' knowledge proficiency and then diagnose the cognitive states of each student more precisely, we propose a focused study on introducing educational context features into the cognitive diagnosis process. By analyzing the characteristics of educational contexts, we find several domain and technical challenges along this line. Firstly, the educational contexts may involve contents from different aspects (e.g., parents' education, school resource, personal interest), which is hard to be analyzed uniformly. Those

*Qi Liu is the corresponding author.

KDD '21, August 14–18, 2021, Virtual Event, Singapore

ACM ISBN 978-1-4503-8332-5/21/08...$15.00
https://doi.org/10.1145/3447548.3467264

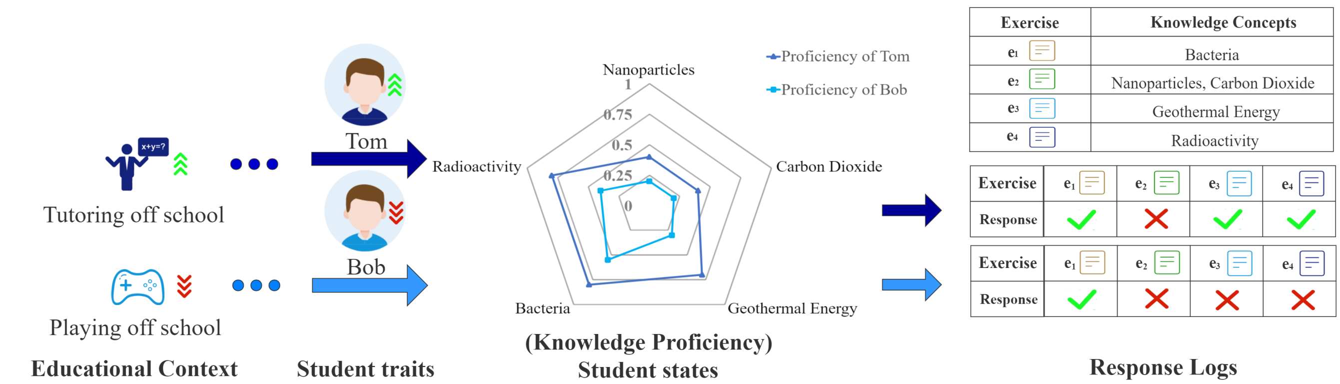


Figure 1: An illustration of students' learning process.

inspire us to model the contexts influence by aspects respectively. Secondly, as shown in Figure 1, the educational contexts should be modeled by influencing students' latent traits and are not directly concerned with specific knowledge concepts, while student states through cognitive diagnosis should then connect to the knowledge concepts. Thirdly, different students may get different influence even from the same educational context. For instance, while educational context "*get tutoring*" is generally a positive context for students' learning, truant students may get less positive feedback from "*tutoring*" and hardworking students can get more. As a result, the influence intensity of context should be personalized in terms of the different fitness between context and students' character. Finally, educational contexts may interact with each other while influencing students. For instance, the effect of context "*get tutoring*" will be affected by other educational contexts, e.g., parents' support, student's attitude and home ESCS (Economic, Social and Cultural Status) [18]. Therefore, how to construct and mine the inherent relevance among different contexts is worth exploring.

To address the challenges mentioned above, we design a novel *E*ducational context-aware *C*ognitive *D*iagnosis (ECD) framework. Generally, ECD has two stages, namely, the educational context modeling stage and the diagnosis enhancement stage. In the educational context modeling stage, the diverse contexts are first grouped into several fields by their content, then a hierarchical attentive network is proposed to represent the personalized influence from each context, and the students' external traits reflected by educational contexts will be generated. In the diagnosis enhancement stage, the students' external traits are adaptively integrated with their inner traits. In this way, the student states can be refined, and our general ECD framework is well defined to be implemented by combining with existing cognitive diagnosis solutions. For instance, we can have ECD-IRT by combining with IRT [6], ECD-MIRT with MIRT [30] and ECD-NeuralCD with NeuralCD [36], respectively. Finally, we conduct extensive experiments on nearly 52 million records of the students sampled by PISA (Programme for International Student Assessment) from 73 countries and regions over the world. The experimental results not only prove that ECD is more effective in student performance prediction since it can well capture the impact from educational contexts to students' cognitive states, but also show the superior interpretability of the ECD framwork. For instance, we give some interesting discoveries regarding the different influence from different educational contexts in different countries and regions.

2 RELATED WORK

The related researches can be grouped into following categories.

Context-aware Modeling. Context-aware user modeling plays an important role in information retrieval (IR) [31] related tasks, ranging from web search [13], recommendation [22, 39] to online advertising [41]. Following, we take context-aware recommendation as an example for better illustration, and there are generally two main streams. Researchers in the first stream processed the contexts with probabilistic graphical models, like the Latent Dirichlet Allocation (LDA) related ones [19, 42]. The work in the second stream usually adopt neural networks to get deep representations of context [10], where the interaction learning of various contexts and the extendablility of the methods are the major focus. For instance, Wide & Deep [3] jointly considers both memorization and generalization abilities with its Wide and Deep architecture. It adopts Deep Neural Network (DNN) as Deep part for generalization ability and generalizes linear model as Wide part for memorization ability. Besides, DeepFM [9] and NFM [11] combine neural Factorization Machine (FM) with DNN where neural FM introduces the interaction of different features. Then, DCN [37] and xDeepFM [17] design specific network module to capture high order interactions. And AFM [40] and AutoInt [34] conduct attention on features' interactions. In summary, context-aware modeling has been successfully applied to enhance the quality of service in user preference prediction related fields. However, from this data-driven perspective, the problem of how contexts affect people's knowledge proficiency is still underexplored.

Educational Context Analysis. Educational context is mainly discussed in the education domain, where the researchers often propose a set of assumptions to get a qualitative analysis of the influence from environment on students' learning process. For instance, the social cognitive theory [24] has already introduced the idea of interaction between environment and study. Recently, a number of works use professor-defined profiles to collect target educational contexts and analyze the results with traditional statistical methods. Since educational contexts have complex content from various of topics, most of these works focus on several specific contexts, including features from various aspects, like country (e.g., education policy [14]), school (e.g., school teaching [29], school climate [16]), home (e.g., parents involvement [25]), person (e.g., epistemological belief [2]) and so on. However, due to the inexplicit and complex influence mechanism of educational context on student performance, the traditional analysis methods usually

Table 1: Educational context examples.

Aspect	Context examples
Home	Highest education degree of parents
	Parents involvement in children's study
	Home Economic, Social and Cultural Status (ESCS)
School	Method of school teaching and learning
	Teacher' attitude to teaching and students
	Information and Communication Technology (ICT)
Person	Duration in early childhood education
	Whether students have a grade repetition experience
	Science activities experience out of school

can not build a quantitative relationship between educational contexts and specific student's learning states.

Cognitive Diagnosis. Cognitive diagnosis is a fundamental task in intelligent education. Most of existing methods can be classified into two types based on the data/information exploited, namely raw-response type and exercise information enhanced type. The raw-response methods only use students' numeric response records to infer their states, such as Item Response Theory (IRT) [6] and Probabilistic Matrix Factorization method (PMF) [35]. Specifically, IRT adopts a logistic-like function to describe integrated knowledge state of students, in which probability of right response P is denoted as the interaction of student's proficiency level θ, the exercise difficulty α and discrimination β. The assumption lies in that the higher level of student's proficiency, the higher probability that she will correctly response the exercise, and a simple version of IRT is represented as:

$$P = sigmoid(a(\theta - b)), \tag{1}$$

where the parameters θ, α and β are learned by fitting the students' historical response data. One step further, the exercise information enhanced methods [20] could introduce extra information from exercises. To be specific, MIRT [30] extends the dimension of scalar parameters in IRT (e.g., θ and β) to the number of knowledge concepts. DINA [5] adopts the exception assumptions, i.e., slip and guess. FuzzyCDF [21] further takes the partially correct responses on subjective problems into consideration with fuzzy set theory. More still, Deep Learning Enhanced Item Response Theory (DIRT) [4] consides extracting the exercises' text materials with elaborately designed network to restrict the exericise-related parameters. Besides, Neural cognitive diagnosis (NeuralCD) [36] is an extendable framework, and it models complex interaction between students and exercise by neural network. Generally, these existing methods just exploit students' historical response logs or exercise-related information for diagnosing the cognitive states of each student, and can not handle the educational contexts.

3 PRELIMINARIES

In this section, we give a brief introduction of the educational context, the datasets and a formal definition of the educational context-aware cognitive diagnosis problem.

Educational Context Description *Educational context* refers to the various features related to the students' learning process, which may come from different aspects. Table 1 lists several examples of educational contexts, which may contain different feature answers. For instance, answers to "Highest education degree of parents" may include "1 (*General senior*)", "2 (*Vocational senior*)" , "3 (*Junior*)" , "4 (*Primary*)" and "5 (*Not complete Primary*)", respectively.

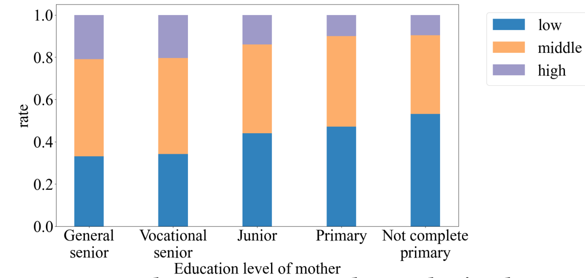
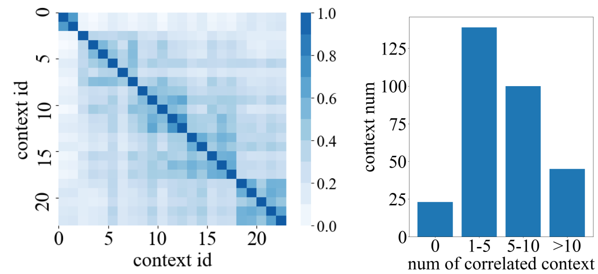


Figure 2: Student average score by mother's education.



(a) Pearson heatmap (b) Correlation distribution

Figure 3: The correlations between context features.

In fact, educational context features are usually difficult to collect. Fortunately, we now can address this issue with the help of OECD's Programme for International Student Assessment (PISA). PISA is one of the most famous worldwide testing programme, which is honored as Olympic Games in testing project and attracts nearly one hundred regions or countries to take part in. Specifically, PISA measures 15-year-olds' ability to use their reading, mathematics and science knowledge and skills to meet real-life challenges. Besides, it releases questionnaires to collect the students' educational contexts (like those contexts in Table 1) and make analysis [27]. Generally, the PISA puts out every three years with growing scale and releases the data and technical reports successively in the following three years. In this paper, we take the public dataset of PISA 2015[1] as an example, which focus on science assessment.

With the guidance of technical report in "PISA 2015 Results" [28], we select the students' questionnaire context data and science cognitive data from different regions to compose three subset datasets (i.e., Asia, Europe and America) and their statistical information will be given in the experimental part (Section 5.1.1). To intuitively understand the characteristics of educational contexts, we analyze the relationship between educational contexts and student performance. In Figure 2, we visualize the distribution of students' average score with their answers to the context of "highest level of schooling completed by mother". We can find that mother's education level is positively correlated to students' average score.

However, a small number of students with answer of "*Not complete primary*" may still have high average score, which means they are less influenced by this negative educational context feature. Therefore, the influence of the same educational context can be personalized. We also briefly analyze the correlation distribution between context features. Specifically, except the contexts with digital answers (e.g., "*Learning time*"), there are also natural ordinal relation in many contexts of non-digital answers (e.g., "*Not*

[1] http://www.oecd.org/pisa/data/2015database/

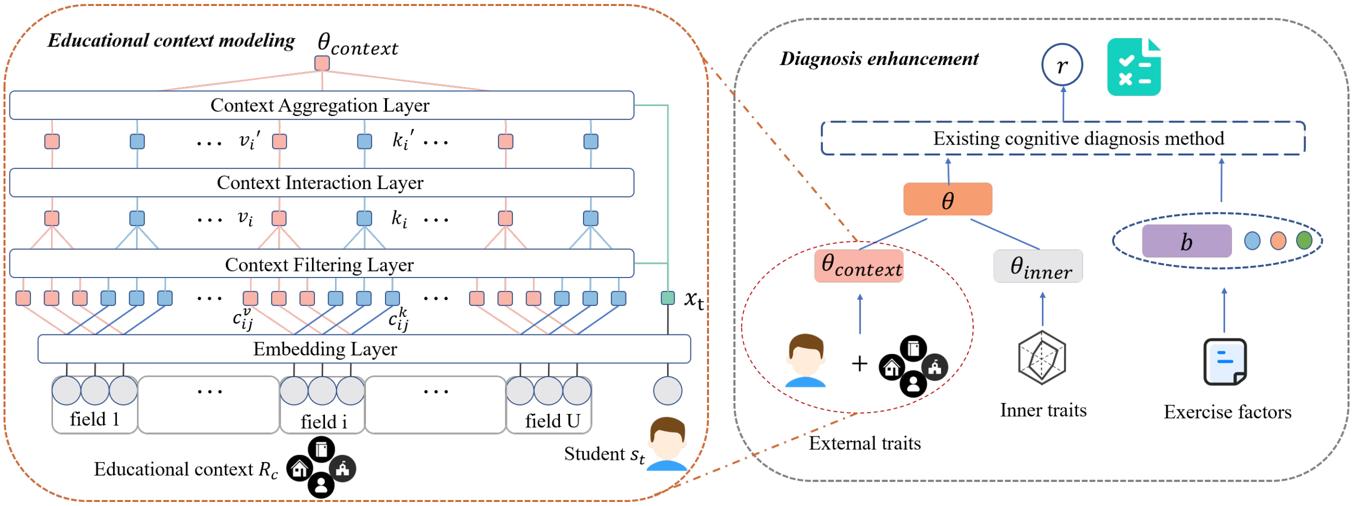


Figure 4: Overview of ECD Framework .

complete Primary - Primary - Junior - Senior" for "*Education level*", "*Strongly disagree - Disagree - Agree - Strongly agree*" for "*Good class atmosphere*"). We coded those context answers to numbers by their order and then compute the Pearson correlation coefficient between the coded answers of different contexts, and two contexts are treated as correlated if the coefficient is greater than 0.3. We visualize Pearson coefficient between context features in Figure 3(a) , which shows inherent relevance between contexts. Moreover, we summarize the correlation distribution of contexts in Figure 3(b), where most educational contexts are correlated with others.

Problem Definition. We formally introduce the educational context-aware cognitive diagnosis problem.

Suppose we have N students, T educational context questions, and M exercises in a learning system, which can be represented as $\mathcal{S} = \{s_1, s_2, ..., s_N\}$, $\mathcal{Q} = \{q_1, q_2, ..., q_T\}$, $\mathcal{E} = \{e_1, e_2, ..., e_M\}$ respectively. The logs R, consisting of educational context question response records R_q and exercise response records R_e, are denoted as set of triplet (s, q, r_q) and (s, e, r_e) respectively, where $s \in \mathcal{S}$, $e \in \mathcal{E}$, $q \in \mathcal{Q}$, r_q is the response (e.g., "*General senior*") that student s answered on educational context question q (e.g., "Highest education degree of parents") and r_e is the score (transferred to percentage) that student s got on exercise e. Then the problem can be formally defined as:

Definition 3.1. Educational context-aware cognitive diagnosis: Given students' logs $R = \{R_q, R_e\}$, our goal is to infer students' proficiency on knowledge concepts (i.e., student states in Figure 1) through student performance (i.e., exercise answering) prediction.

4 EDUCATIONAL CONTEXT-AWARE COGNITIVE DIAGNOSIS FRAMEWORK

In this section, we first give an overview of the two-stage Educational context-aware Cognitive Diagnosis (ECD) architecture, and then describe these two stages in detail.

4.1 Model Overview

Generally, the exercise answering process can be formulated as:

$$r = F(\theta, \phi_e), \tag{2}$$

where r refers to student response (score or correctness), θ denotes student's knowledge proficiency, ϕ_e denotes the exercise parameters e (e.g., difficulty b, discrimination a). F denotes the manually designed cognitive behavior function (e.g., item response function in IRT, Eq.(1)) that models the interaction between student and exercise parameters and output the response. Considering the educational contexts of a student, we further divide θ into two parts:

$$\theta = G(\theta_{context}, \theta_{inner}), \text{ where } \theta_{context} = H(C), \tag{3}$$

where $\theta_{context}$ is the student external traits affected by educational contexts, θ_{inner} is the student inner traits that traditional cognitive diagnosis concerns and G denotes the influence function of the two student traits. C is the educational context of the student, and H denotes the influence function of educational context on student external traits $\theta_{context}$, which we mainly focus on in this work.

As Figure 4 shows, the overall architecture of ECD framework consists of two stages: educational context modeling stage and diagnosis enhancement stage. In educational context modeling stage, we design attention networks to model the influence of educational contexts on $\theta_{context}$. Specifically, we first group the diverse educational contexts into several fields and simulate personalized influence respectively with the context filtering layer. Then, we model the inherent relevance among different fields in context interaction layer. After that, the context aggregation layer captures field level personality and generate the external student traits $\theta_{context}$. In the diagnosis enhancement stage, we combine $\theta_{context}$ and θ_{inner} (Eq.(3)) and then output the predicted score with the cognitive behavior function (Eq.(2)). After training with students' logs R, we will get θ of each student as the diagnostic results.

4.2 Educational Context Modeling

In the educational context modeling stage, we design a hierarchical attentive network to model the personalized and inherently relevant influence of educational contexts on students. The whole network consists of four layers: embedding layer, context filtering layer, interaction layer and aggregation layer.

4.2.1 Embedding layer. This layer is used to assign trainable embeddings to each educational context entry u_j and each student s_t. We assign each context entry with an influence key vector $\boldsymbol{c}^k \in \mathbb{R}^{d_1}$ and an influence value scalar c^v, which indicate the latent feature of the context entry and its influence intensity respectively. For each student s_t, we assign a trainable student latent character vector $\boldsymbol{x}_t \in \mathbb{R}^{d_1}$ to capture s_t's fitness of different context entries. d_1 is the vector dimension manually set. It is worth noting that different context entries (e.g., "*Female*" and "*Male*") of the same

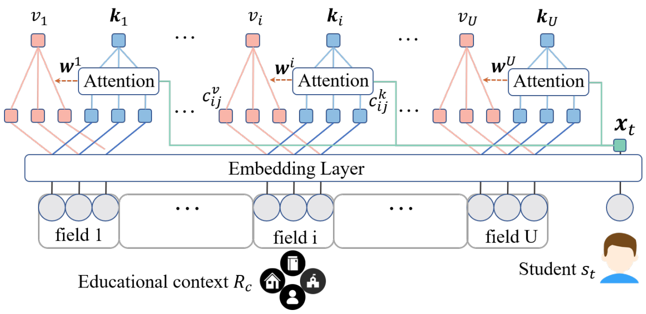


Figure 5: Context Embedding and Filtering Layers.

context feature (e.g., "*Student Gender*") will not appear in a single student. As a result, for m different context features, each student would have at most m context entries.

4.2.2 Context filtering layer. The educational contexts may involve features from different aspects, causing it unsuitable to model their influence uniformly. Therefore, we group the educational context entries into U context fields $C = \{C_1, C_2, \dots, C_U\}$ according to their content, where $C_i = \{c_{i1}, c_{i2}, \dots, c_{im_i}\}$. Generally, a context field C_i contains m_i context entries for m_i context features.

Then, with the consideration of personalizing the context influence for each student, we use attention mechanism to obtain the weights of different context entries. Figure 5 shows the attention based filtering method applied in each context field. We first calculate the cosine similarity between the student character $\boldsymbol{x}_t$ and the context influence key vectors $\boldsymbol{c}_i^k = (\boldsymbol{c}_{i1}^k; \boldsymbol{c}_{i2}^k; \dots; \boldsymbol{c}_{im_i}^k)$:

$$\boldsymbol{w}^i = \text{Softmax}(sim(\boldsymbol{x}_t, \boldsymbol{c}_i^k)), \tag{4}$$

where sim denotes the cosine similarity function, and the weights in $\boldsymbol{w}^i$ reflect the student's fitness on different context entries. Finally, the influence of each educational context field is generated with the weighted-summation of different features in the field:

$$v_i = \sum_{j=1}^{m_i} w_j^i * c_{ij}^v, \ \boldsymbol{k}_i = \sum_{j=1}^{m_i} w_j^i * \boldsymbol{c}_{ij}^k, \tag{5}$$

where w_j^i is the j-th weight in $\boldsymbol{w}^i$. The influence intensity v_i is the overall influence intensity from context field C_i on student s_t before considering the relevance among different fields. The influence type $\boldsymbol{k}_i$ describes the personalized latent feature of C_i.

4.2.3 Context interaction layer. To model the inherent relevance between different educational context fields, we introduce a self-attention module in the interaction layer. As shown in Figure 6, after the context filtering layer, we get the influence type $\boldsymbol{k}_i$ and intensity v_i of educational context field C_i, which are both personalized from Eq. (5). By applying the self-attention mechanism to the $\boldsymbol{k}_i$ and sharing the weights to v_i, we can simulate the personalized inherent relevance between educational context fields.

$$\begin{aligned} &\boldsymbol{w}_i = \text{Softmax}(sim(\boldsymbol{k}_i, \boldsymbol{k})), i \in \{1, ..., U\}, \\ &\text{where } \boldsymbol{k} = (\boldsymbol{k}_1; \boldsymbol{k}_2; \dots; \boldsymbol{k}_U), \end{aligned} \tag{6}$$

$$\boldsymbol{k}_i' = \sum_{j=1}^{U} w_{i,j} * \boldsymbol{k}_j, \ v_i' = \sum_{j=1}^{U} w_{i,j} * v_j. \tag{7}$$

4.2.4 Context aggregation layer. After the interaction layer, we have obtained the educational context field representations with consideration of both personalization in feature level and inherent relevance in field level. In context aggregation layer, we utilize another attention module to assemble the influence from different context fields and finally get the student external trait $\theta_{context}$. Again, the student's latent vector $\boldsymbol{x}_t$ is used to calculate the similarities between $\boldsymbol{k}_i'$ to ensure the personality in context fields:

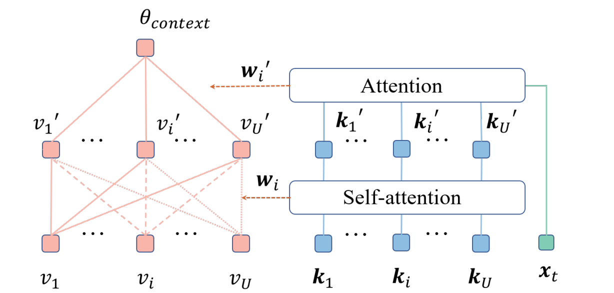


Figure 6: Context Interaction and Aggregation Layers.

$$\begin{aligned} &\boldsymbol{w}' = \text{Softmax}(sim(\boldsymbol{x}_t, \boldsymbol{k}')), \\ &\text{where } \boldsymbol{k}' = (\boldsymbol{k}_1'; \boldsymbol{k}_2'; \dots; \boldsymbol{k}_U'), \end{aligned} \tag{8}$$

$$\theta_{context} = \sum_{i=1}^{U} w_i' * v_i'. \tag{9}$$

The student's external trait $\theta_{context}$ is a global and imperceptible influence on students' cognitive states, while the students' inner trait θ_{inner} inferred from student exercise logs is the local reflection of students' ability or knowledge proficiency (depending on the chosen cognitive diagnosis model). Therefore, we need to take them both into consideration.

4.3 Diagnosis Enhancement

In this stage, we integrate the external student trait reflected by educational contexts with existing cognitive diagnosis models.

Specifically, we first represent the inner trait θ_{inner} with a student latent vector. Then, we choose an adaptive optimization of personalized weight-added method (G in Eq. (3)) to aggregate the student's external and inner traits as the students' final states.

$$\theta = d_t * \theta_{context} + (1 - d_t) * \theta_{inner}, \tag{10}$$

where d_t is the trainable weight personalized for each student s_t, and $\theta, \theta_{context}, \theta_{inner}$ are the student's final knowledge proficiency, student's external trait extracted from educational contexts and student's inner trait inferred from response records respectively. We adopt existing cognitive diagnosis models as F in Eq. (2). The prediction of students' responses on exercises is formulated as:

$$r = \text{CDMethod}(\theta, \phi_e), \phi_e = \{b, ...\}, \tag{11}$$

where CDMethod denotes existing cognitive diagnosis models, like IRT [6], MIRT [30] and NeuralCD [36], ϕ_e denotes the parameters related to the exercise (e.g., exercise difficulty b) in the adopted traditional cognitive diagnosis model.

4.4 Model Learning

With the summarization of the whole model in Figure 4, during model learning, we train the parameters in context modeling $\phi_c = [\{\boldsymbol{c}_{ij}^k, c_{ij}^v\}, \boldsymbol{x}_t]$, weight variable d_t and parameters in existing methods ϕ_e and θ_{inner}. We apply the multi-learning methods to handle it. Specifically, like the common methods in two-classify task, we use the cross entropy as the loss function.

$$loss = Entropy(r, label), \tag{12}$$

where r is the predictions of the model and $label$ is the ground truth of the students' responses on the exercises. In order to get all the parameters well-trained, we adjust the loss function as:

$$loss_c = Entropy(r_c, label), \tag{13}$$

$$loss_{exist} = Entropy(r_{exist}, label), \tag{14}$$

$$loss' = loss + \alpha * loss_c + \beta * loss_{exist}, \tag{15}$$

Table 2: Results on student performance prediction.

Model	Asia AUC	Asia RMSE	Asia ACC	Europe AUC	Europe RMSE	Europe ACC	America AUC	America RMSE	America ACC
Random	0.499	0.578	0.499	0.500	0.577	0.501	0.502	0.577	0.501
NeuralCD	0.714	0.490	0.658	0.718	0.476	0.659	0.712	0.495	0.665
DeepFM-NeuralCD	0.728	0.488	0.660	0.745	0.455	0.688	0.743	0.472	0.661
NFM-NeuralCD	0.722	0.483	0.660	0.718	0.494	0.667	0.717	0.486	0.652
ECD-NeuralCD	**0.745**	**0.468**	**0.677**	**0.770**	**0.443**	**0.700**	**0.764**	**0.445**	**0.699**
IRT	0.734	0.460	0.675	0.741	0.456	0.687	0.736	0.455	0.678
DeepFM-IRT	0.736	0.459	0.673	0.753	0.450	0.689	0.768	0.443	0.701
NFM-IRT	0.724	0.464	0.670	0.752	0.452	0.679	0.771	0.441	0.703
ECD-IRT	**0.757**	**0.449**	**0.689**	**0.760**	**0.447**	**0.699**	**0.773**	**0.439**	**0.703**
MIRT	0.669	0.484	0.622	0.696	0.493	0.650	0.691	0.475	0.655
DeepFM-MIRT	0.744	0.460	0.676	0.741	0.454	0.684	0.738	0.459	0.678
NFM-MIRT	0.736	0.463	0.665	0.757	0.452	0.692	0.755	0.449	0.688
ECD-MIRT	**0.786**	**0.435**	**0.704**	**0.790**	**0.432**	**0.710**	**0.795**	**0.427**	**0.715**

Table 3: The statistics of datasets from PISA.

Datasets	Students	Educational contexts	Context records	Exercise	Exercise records
Asia	76,609	300	14,586,482	260	2,172,516
Europe	69,016	300	18,127,964	260	1,952,577
America	62,091	300	14,205,515	260	1,746,899

where r_c denotes the prediction with only the students' external trait from educational contexts (the weight variable in Eq. (10) d_t = 1), r_{exist} denotes the predictions with only existing methods (the weight variable in Eq. (10) $d_t = 0$). α, β are hyper parameters that trade off these three losses, and $loss'$ is the final loss function.

5 EXPERIMENTS

We conduct experiments to demonstrate the effectiveness of ECD framework with several baselines. Besides, based on ECD, we deeply analyze the influence of educational context features on students' learning states, make discussion on some typical observations. [2]

5.1 Experimental Setup

5.1.1 Data partition and preprocessing. The overall PISA 2015 dataset contains more than 0.5 million students from 73 different countries and regions. We further extract three datasets from PISA 2015 by area, namely Asia, Europe, America. In each dataset, we have 300 different questions related to different educational context features and get students' response records on totally 260 science cognitive exercises, respectively. Specially, we filter out the students whose records are less than 20 to ensure that there is sufficient data for training. Further, in our datasets, the questionnaire problems are manually grouped into fields ($U = 23$) according to their content and the guidance in reports of PISA 2015 [26]. Some basic statistics of these datasets are shown in Table 3. Finally, we radomly partition all the datasets into 80%/20% for training/testing.

5.1.2 Baseline Approaches. To verify the influence of educational context features on students' performances, we present three implementations based on ECD framework that combine typical diagnosis methods. In particular, we implement ECD-IRT, ECD-MIRT and ECD-NeuralCD following IRT, MIRT, NeuralCD, respectively.

- IRT [6]: IRT is a cognitive diagnosis method which models the cognitive processes from students' exercising records with a logistic-like function.
- MIRT [30]: MIRT is a vairant of basic IRT model, where it extends the latent trait value of each student in IRT to a multi-dimension knowledge proficiency vector.
- NeuralCD [36]: NeuralCD is a deep neural cognitive diagnosis framework which models the interaction from students' exercising records with a multilayer perceptron (MLP).

Specifically, to demonstrate the effectiveness of context modeling in our ECD, we compare it with two typical context modeling methods widely used in context-aware recommendation works.

- DeepFM [9]: DeepFM utilizes a neural factorization machine to learn the interaction of features in bit wise. Then it combines the output of the two parts as the final result.
- NFM [11]: NFM uses a neural factorization machine to model the features' interaction in vector wise and pass the output to a DNN module to get the final result.

Besides, since PISA datasets are new to cognitive diagnosis task, we also compare with the basic random algorithm which predicts the students' scores randomly from $Uniform(0, 1)$.

5.1.3 Metrics. The performance of cognitive diagnosis models is difficult to evaluate as we can't obtain the true knowledge proficiency of students. Following existing works, we evaluate these models indirectly through the results of students' performance prediction, from both regression and classification perspectives. As a regression task, we quantify the distance between the predicted and actual scores with Root Mean Square Error (RMSE). The smaller the values are, the better the results have. Treating the problem as a classification task, where a record with score 1(0) indicates a positive (negative) instance, we adopt Area Under an ROC Curve (AUC) and Prediction Accuracy (ACC) to measure the effectiveness [1, 32], and the larger values are, the better the results have.

5.1.4 ECD Setting. We specify the experimental setups in ECD, including ECD framework settings and diagnosis model settings. In ECD, for all implementations, that is ECD-IRT, ECD-MIRT and ECD-NeuralCD, we set the dimensions of student type vector and context influence type vector as 10 and assign the similarity measure in attention mechanism with cosine similarity. In addition, for diagnosis models, typical 2-PL model of IRT [6] and MIRT [30] are respectively adopted to ECD-IRT, ECD-MIRT, and origin settings of NeuralCD are applied in ECD-NeuralCD following [36].

5.1.5 Training Setting. To set up the training process, we initialize the parameters with Xavier initialization [8], which fill the weights with random values sampled from $N(0, std^2)$, where $std = \sqrt{\frac{2}{n_{in}+n_{out}}}$, n_{in} is the number of neurons feeding into the weights, and n_{out} is the number of neurons the results is fed to. Besides, we set the mini batches as 128 and select different weight variables α and β and lr (learning rate) for different models (i.e., $\alpha = 4, \beta = 0, lr = 0.005$ for IRT-based models, $\alpha = 0, \beta = 4, lr = 0.001$ for MIRT-based models

[2]Our code of ECD is available at https://github.com/bigdata-ustc/ECD.

Table 4: Results of ablation experiment.

Model	Asia AUC	Asia RMSE	Asia ACC	Europe AUC	Europe RMSE	Europe ACC	America AUC	America RMSE	America ACC
ECD-NeuralCD	**0.745**	**0.468**	**0.677**	**0.770**	**0.443**	**0.700**	**0.764**	**0.445**	**0.699**
- Filtering	0.743	0.469	0.669	0.764	0.445	0.699	0.762	0.445	0.699
- Interaction	0.736	0.471	0.665	0.752	0.451	0.687	0.746	0.463	0.684
- Aggregation	0.738	0.465	0.668	0.747	0.456	0.678	0.747	0.450	0.690
ECD-IRT	**0.757**	**0.449**	**0.689**	**0.760**	**0.447**	**0.699**	**0.773**	**0.439**	**0.703**
- Filtering	0.745	0.456	0.680	0.752	0.451	0.695	0.757	0.447	0.694
- Interaction	0.745	0.455	0.677	0.756	0.449	0.694	0.768	0.442	0.699
- Aggregation	0.739	0.456	0.680	0.755	0.450	0.688	0.754	0.448	0.687
ECD-MIRT	**0.786**	**0.435**	**0.704**	**0.790**	**0.432**	**0.710**	**0.795**	**0.427**	**0.715**
- Filtering	0.781	0.440	0.695	0.787	0.433	0.706	0.788	0.434	0.709
- Interaction	0.779	0.443	0.695	0.787	0.433	0.708	0.788	0.433	0.704
- Aggregation	0.773	0.443	0.698	0.777	0.438	0.700	0.763	0.442	0.692

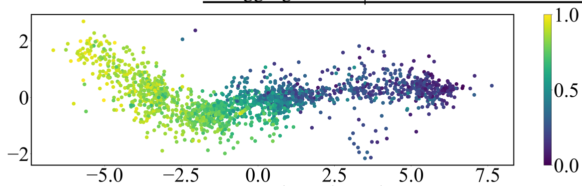


Figure 7: x_t vectors colored with average score.

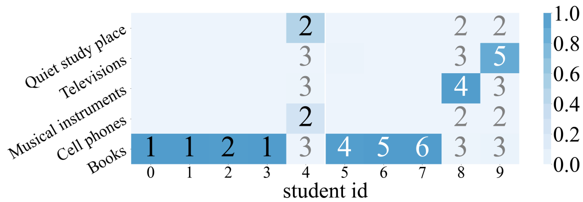


Figure 8: Attention weight and coded response of context features for different students.

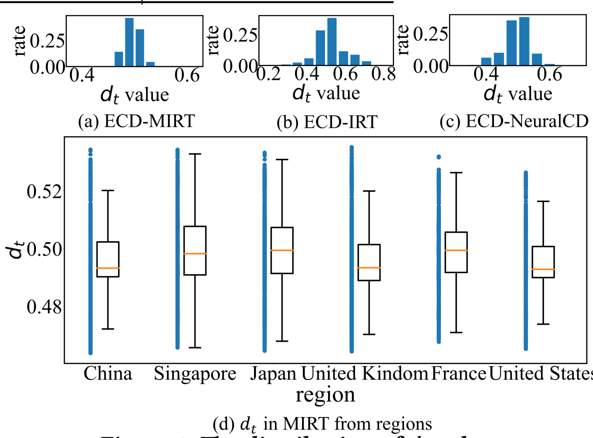


Figure 9: The distribution of d_t values.

and $\alpha = 1, \beta = 1, lr = 0.005$ for NeuralCD-based models). All models are implemented by PyTorch using Python, and all experiments are run on a Linux server with Tesla K80 GPU.

5.2 Experimental Results

5.2.1 Student Performance Prediction. The overall performances is shown in Table 2, there are several key observations: Firstly, in general, context modeling methods (e.g., DeepFM-MIRT, NFM-MIRT, ECD-MIRT) outperform the original cognitive diagnosis methods (e.g., MIRT). It indicates that the abundant educational context features deserve consideration for cognitive diagnosis. Secondly, our proposed ECD framework performs better than DeepFM or NFM method on all the cognitive diagnosis methods in three datasets. It notes that our method can model educational contexts more effectively. Thirdly, our ECD methods have a stable great improvement on all the cognitive diagnosis methods in three datasets, while the performances of DeepFM and NFM methods are unstable and not always positive for the student performance prediction task. As mentioned before, there are critial diverse content, personalized influence and inherent relevance challenges in educational context modeling. In our opinion, modeling contexts with consideration of all the characteristics contributes the superiority of ECD.

5.2.2 Ablation. To verify the effectiveness of all the three layers in ECD framework, i.e., context filtering, context interaction and context aggreration, we conduct the ablation experiments. Specifically, we successively replace each layer with an ordinary aggregation layer, which simply averages the inputs, while maintain the other two layers. Table 4 reports the results of each case, which we conclude as following. Firstly, no matter which layer is replaced, final performances decrease to some degree. It shows every layer contributes to the final performances, which indicates the effectiveness of these attentive modules modeling personized influence and inherent relevance. Secondly, final performances suffer the greatest damage when the context aggregation layer is replaced, which indicates personalization in context field level plays the most important role in our educational context modeling.

5.2.3 Parameter Analysis. In our work, we define student latent character vector, and students from different character will receive different influence from the same educational context. One intuition is that students with high performance may be affected by positive contexts more. For instance, a hardworking students may get more positive influence from the educational context "*get tutoring*" than a truant student. Further to observe the relationship between student character and student performance, we visualize the student character vectors after reducing their dimension by t-SNE [15] and color each vector by the corresponding student's average response score in Figure 7. It is noting that the distribution of students' character vector has close relationship with the students' average score, which proofs the personalization characteristic of educational contexts may reflect to students' general ability and the students will receive different influence from the same context.

Besides, as mentioned in Section 4.2, variable d_t denotes the weights of influence from context features. We visualize the distribution of d_t values for different models and regions in Figure 9. Specifically, we first summarize the distribution of d_t in all regions for each ECD model in Figure 9(a), 9(b), and 9(c). Further, we analyze the d_t values of students from six specific regions across Asia, Europe and America in Figure 9(d). We can find that distribution of d_t values from basic cognitive diagnosis methods or regions vary slightly. Moreover, for all models, most students have a d_t value in

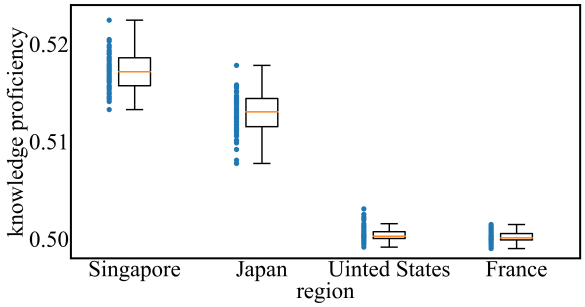


Figure 10: Distribution of average knowledge proficiency.

[0.4,0.6], which means both the context influence and the historic exercise records are not ignorable for a general diagnosis.

5.2.4 Attention Visualization. In ECD framework, we use an attention mechanism in context filtering layer to model the importance of different context features. To verify the effectiveness of the context filtering layer, we analyze the relationship between attention weights of educational context features and students' performances. Here, we visualize an example of ten different students' context features from educational context field "*home ESCS*" in Figure 8. Specially, the former 5 students (No. 0-4) have low average scores while the last 5 students (No. 5-9) have a high level performances. Then, we note the responses on the widely focused context "*Books*" and other focused responses, which is coded to number 1-6. Specifically, the context "*Books*" refers the problem "How many books are there in you home?", and the higher number refers to the more books. Apparently, there is a correlation between the students' answers and their performances that more books have a more positive influence on students' performances. As the Figure 8 shows, the low score level students have higher attention on the negative answers (e.g. 1, 2) and the positive answers (e.g. 4,5,6) weight more in high score level students. Besides, it is worth noting that there are not always many books for some students with high scores (No. 8-9). In this case, some other features, like "*Televisions*" and "*Musical instruments*", may make a positive influence for students just as the samples shown in Figure 8. This indicates that the attention mechanism in context filtering layer can reasonably model the personalized influence of different features on students.

5.2.5 Cognitive States Visualization. To indicate the effectiveness of our ECD framework, we compare the average diagnosis results of regions with the PISA report [26]. Specifically, in the PISA report, the science literacy of students have been evaluated as scores in region level. Typically, we take four representative regions from different continents, i.e., Singapore, Japan, United States and France as example, where the scores are 556, 538, 496, 495, respectively. In each region, we first compute the average knowledge proficiency of students: $\theta_g = \frac{\sum_{i=1}^{n} \theta_i}{n}$, ($g \in \{Singapore, Japan, USA$ and $France\}$), where g denotes the region, n denotes the numbers of students and θ_i refers the knowledge proficiency of i-th student. Here we take the final θ (Eq. (10)) from ECD-NeuralCD as an example. Then we visualize the distribution of θ_r in all knowledge concepts in Figure 10. We can find that the order of regions are consistent with the report of PISA. Moreover, we also visualize θ_g of 6 specific knowledge concepts in Figure 11 to intuitively reflect the difference. Similarly, the θ_{USA} and θ_{France} is lower than $\theta_{Singapore}$ and θ_{Japan} in all six concepts. Besides, compared with θ_{France}, θ_{USA} is lower in concepts of "*Nanoparticles*", "*Bacteria*" and "*Carbon Dioxide*" but

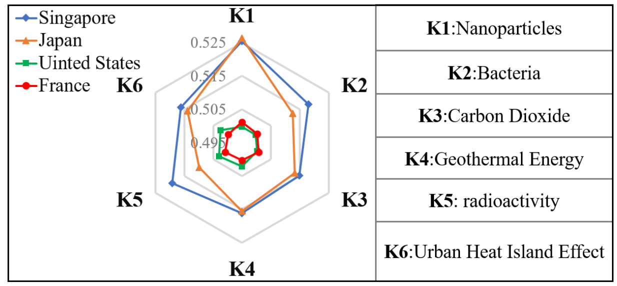


Figure 11: Visualization of average knowledge proficiency.

higher in "*Urban Heat Island Effect*", "*Radiotherapy*" and "*Geothermal Energy*". Genernally, ECD can further discriminate these specific difference in knowledge concepts between regions, while the assessment in region level keeps consistent with the PISA report.

5.3 Discussion

In addition to promote student diagnosis, we also wonder the importance of different educational contexts. As menthioned in Section 4.2, different educational contexts will be aggregated in the context aggregation layer with an attention module. For specific student, the attention weight of a certain context denotes its importance. Following this line, we record the 3 most important contexts for each student and summarize the results by region in Table 5, where contexts of different aspects are noted with different color (i.e., red for home, blue for school and black for person).

Here are some interesting and instructive observations. Firstly, context "*Parent education*" are focused in China and Korea. In our opinion, that can be concerned with the similar local tradition in education. For instance, national college entrance exam plays an important role in the students' education in China and Korea. It puts a heavy stress on students and even their parents. In other regions, "*Parent education*" does not attract a wide attention. However, that does not mean the education contexts of home aspect can always be ignored. On the contrary, all regions give much attention in educational context "*home ESCS* (Economic, Social and Cultural Status)", which suggests the considerable impact of family support in learning process. Similarly, in school aspect, contexts "*School learning*" and "*Teaching attitude*" shows difference between regions from Aisa and the others, while "*School ICT*" and "*ICT Usage*" play an important role in all regions (data of the two contexts is lack in USA). All those note that the educational resources are vital for the student learning. Finally, compared with other environment context features, contexts of person aspect weights less in most regions, which indicates that external features are generally more important. This inevitably alert us to worry about and attach importance to educational fairness.

6 CONCLUSION

In this paper, we presented a novel framework ECD for students' cognitive diagnosis, which is also a quantitive perspective for educational context understanding. Specifically, we first designed a two-stage solution with a hierarchical attentive network modeling the influence of educational contexts and an adaptive optimization for student traits aggregation. Then, we implemented three specific models with different existing methods under the framework, (i.e., ECD-IRT, ECD-MIRT, ECD-NeuralCD). Besides, we conducted extensive experiments on real-world datasets to demonstrate the effectiveness as well as interpretability of ECD framework. Finally,

Table 5: Important educational contexts in different regions.

Regions	Context
Uinted States	"Home ESCS", "School learning", "Teacher Attitude", "Self-efficacy"
Uinted Kindom	"Home ESCS", "School learning", "Teacher Attitude", "School ICT", "ICT Usage", "Self-efficacy"
France	"Home ESCS", "School learning", "Teacher Attitude", "School ICT", "ICT Usage"
Germany	"Home ESCS", "School learning", "Teacher Attitude", "School ICT", "ICT Usage"
Italy	"Home ESCS", "School learning", "Teacher Attitude", "School ICT", "ICT Usage"
Singapore	"Home ESCS", "School ICT", "ICT Usage", "Interest on science", "Self-efficacy"
Japan	"Home ESCS", "School ICT", "ICT Usage", "Self-efficacy"
Korea	"Parent education", "Home ESCS", "School ICT", "ICT Usage"
China	"Parent education", "Home ESCS", "School ICT", "ICT Usage"

we analyzed and discussed the difference of influencial context features for students from different regions with our ECD framework. We hope this work will lead to more studies in the future.

7 ACKNOWLEDGMENTS

This research was partially supported by grants from the National Natural Science Foundation of China (No.s 61922073, U20A20229 and 91746301), the Fundamental Research Funds for the Central Universities (Grant No. WK2150110021). Qi Liu acknowledges the support of the Youth Innovation Promotion Association of CAS (No. 2014299).

REFERENCES

[1] Ghodai Abdelrahman and Qing Wang. 2019. Knowledge Tracing with Sequential Key-Value Memory Networks. In *SIGIR* (Paris, France) *(SIGIR'19)*. Association for Computing Machinery, New York, NY, USA, 175–184.

[2] Michelle M Buehl and Patricia A Alexander. 2005. Motivation and performance differences in students' domain-specific epistemological belief profiles. *AERJ* 42, 4 (2005), 697–726.

[3] Heng-Tze Cheng, Levent Koc, Jeremiah Harmsen, Tal Shaked, Tushar Chandra, Hrishi Aradhye, Glen Anderson, Greg Corrado, Wei Chai, Mustafa Ispir, et al. 2016. Wide & deep learning for recommender systems. In *DLRS*. 7–10.

[4] Song Cheng, Qi Liu, Enhong Chen, Zai Huang, Zhenya Huang, Yiying Chen, Haiping Ma, and Guoping Hu. 2019. DIRT: Deep Learning Enhanced Item Response Theory for Cognitive Diagnosis. In *CIKM*. 2397–2400.

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

[6] Susan E Embretson and Steven P Reise. 2013. *Item response theory*. Psychology Press.

[7] Aritra Ghosh, Neil Heffernan, and Andrew S. Lan. 2020. Context-Aware Attentive Knowledge Tracing. In *ACM SIGKDD*. 2330–2339.

[8] Xavier Glorot and Yoshua Bengio. 2010. Understanding the difficulty of training deep feedforward neural networks. *J MACH LEARN RES* 9 (2010), 249–256.

[9] Huifeng Guo, Ruiming Tang, Yunming Ye, Zhenguo Li, and Xiuqiang He. 2017. DeepFM: A Factorization-Machine based Neural Network for CTR Prediction. In *IJCAI*.

[10] Wei Guo, Can Zhang, Huifeng Guo, Ruiming Tang, and Xiuqiang He. 2020. Multi-Branch Convolutional Network for Context-Aware Recommendation. In *SIGIR*. Association for Computing Machinery, New York, NY, USA, 1709–1712.

[11] Xiangnan He and Tat-Seng Chua. 2017. Neural factorization machines for sparse predictive analytics. In *SIGIR*. 355–364.

[12] Heider and Fritz. 1936. *Principles of topological psychology*. McGraw-Hill,.

[13] Silu Huang, Jialu Liu, Flip Korn, Xuezhi Wang, You Wu, Dale Markowitz, and Cong Yu. 2019. Contextual Fact Ranking and Its Applications in Table Synthesis and Compression. In *ACM SIGKDD*. 285–293.

[14] David Hursh. 2007. Assessing No Child Left Behind and the Rise of Neoliberal Education Policies. *American Educational Research Journal* 44, 3 (2007), 493–518.

[15] Van Der Maaten Laurens and Geoffrey Hinton. 2008. Visualizing Data using t-SNE. *JLMR* 9, 2605 (2008), 2579–2605.

[16] Jeffrey Leiter and Matthew C Johnsen. 1997. Child maltreatment and school performance declines: An event-history analysis. *AERJ* 34, 3 (1997), 563–589.

[17] Jianxun Lian, Xiaohuan Zhou, Fuzheng Zhang, Zhongxia Chen, Xing Xie, and Guangzhong Sun. 2018. xdeepfm: Combining explicit and implicit feature interactions for recommender systems. In *SIGKDD*. 1754–1763.

[18] Xiangyi Liao and Xiaoting Huang. 2018. Who is More Likely to Participate in Private Tutoring and Does it Work?: Evidence from PISA (2015). *ECNU Review of Education* (2018).

[19] Qi Liu, Enhong Chen, Hui Xiong, and Yong Ge. 2013. A Cocktail Approach for Travel Package Recommendation. *TKDE* 26, 2 (2013), 278–293.

[20] Qi Liu, Zhenya Huang, Yu Yin, Enhong Chen, Hui Xiong, Yu Su, and Guoping Hu. 2019. Ekt: Exercise-aware knowledge tracing for student performance prediction. *TKDE* 33, 1 (2019), 100–115.

[21] Qi Liu, Runze Wu, Enhong Chen, Guandong Xu, Yu Su, Zhigang Chen, and Guoping Hu. 2018. Fuzzy cognitive diagnosis for modelling examinee performance. *TIST* 9, 4 (2018), 1–26.

[22] Zheng Liu, Xing Xie, and Lei Chen. 2018. Context-Aware Academic Collaborator Recommendation. In *ACM SIGKDD*. 1870–1879.

[23] Zitao Liu, Guowei Xu, Tianqiao Liu, Weiping Fu, Yubi Qi, Wenbiao Ding, Yujia Song, Chaoyou Guo, Cong Kong, Songfan Yang, and Gale Yan Huang. 2020. Dolphin: A Spoken Language Proficiency Assessment System for Elementary Education. In *WWW*. 2641–2647.

[24] Aleksandra Luszczynska and Ralf Schwarzer. 2005. Social cognitive theory. *Predicting health behaviour* 2 (2005), 127–169.

[25] Ralph B McNeal. 2015. Parent involvement and student performance: The influence of school context. *Educational Research for Policy and Practice* 14, 2 (2015).

[26] OECD. 2016. *PISA 2015 Results (Volume I)*. 492 pages. https://doi.org/https://doi.org/10.1787/9789264266490-en

[27] OECD. 2017. *PISA 2015 Assessment and Analytical Framework*. 260 pages. https://doi.org/https://doi.org/10.1787/9789264281820-en

[28] OECD. 2018. *Equity in Education*. 192 pages. https://doi.org/https://doi.org/10.1787/9789264073234-en

[29] Erin R. Ottmar, Sara E. Rimm-Kaufman, Ross A. Larsen, and Robert Q. Berry. 2015. Mathematical Knowledge for Teaching, Standards-Based Mathematics Teaching Practices, and Student Achievement in the Context of the Responsive Classroom Approach. *AERJ* 52, 4 (2015), 787–821.

[30] Mark D Reckase. 2009. Multidimensional item response theory models. In *Multidimensional item response theory*. Springer, 79–112.

[31] Dominic Seyler, Praveen Chandar, and Matthew Davis. 2018. An Information Retrieval Framework for Contextual Suggestion Based on Heterogeneous Information Network Embeddings. In *SIGIR* (Ann Arbor, MI, USA). Association for Computing Machinery, New York, NY, USA, 953–956.

[32] Shuanghong Shen, Qi Liu, Enhong Chen, Han Wu, Zhenya Huang, Weihao Zhao, Yu Su, Haiping Ma, and Shijin Wang. 2020. Convolutional Knowledge Tracing: Modeling Individualization in Student Learning Process. In *SIGIR*. Association for Computing Machinery, New York, NY, USA, 1857–1860.

[33] Naman Shukla, Arinbjörn Kolbeinsson, Ken Otwell, Lavanya Marla, and Kartik Yellepeddi. 2019. Dynamic Pricing for Airline Ancillaries with Customer Context. In *ACM SIGKDD*. 2174–2182.

[34] Weiping Song, Chence Shi, Zhiping Xiao, Zhijian Duan, Yewen Xu, Ming Zhang, and Jian Tang. 2019. Autoint: Automatic feature interaction learning via self-attentive neural networks. In *CIKM*. 1161–1170.

[35] Nguyen Thai-Nghe, Tomáš Horváth, and Lars Schmidt-Thieme. 2010. Factorization models for forecasting student performance. In *EDM 2011*.

[36] Fei Wang, Qi Liu, Enhong Chen, Zhenya Huang, Yuying Chen, Yu Yin, Zai Huang, and Shijin Wang. 2020. Neural cognitive diagnosis for intelligent education systems. In *AAAI*, Vol. 34. 6153–6161.

[37] Ruoxi Wang, Bin Fu, Gang Fu, and Mingliang Wang. 2017. Deep & cross network for ad click predictions. In *ADKDD'17*. 1–7.

[38] Zhiwei Wang, Xiaoqin Feng, Jiliang Tang, Gale Yan Huang, and Zitao Liu. 2019. Deep knowledge tracing with side information. In *AIED*. Springer, 303–308.

[39] Le Wu, Lei Chen, Richang Hong, Yanjie Fu, Xing Xie, and Meng Wang. 2019. A hierarchical attention model for social contextual image recommendation. *IEEE TKDE* 32, 10 (2019), 1854–1867.

[40] Jun Xiao, Hao Ye, Xiangnan He, Hanwang Zhang, Fei Wu, and Tat-Seng Chua. 2017. Attentional Factorization Machines: Learning the Weight of Feature Interactions via Attention Networks. In *IJCAI-17*. 3119–3125.

[41] Yi Zhang, Arun C. Surendran, John C. Platt, and Mukund Narasimhan. 2008. Learning from Multi-Topic Web Documents for Contextual Advertisement. In *ACM SIGKDD* (Las Vegas, Nevada, USA) *(KDD '08)*. 1051–1059.

[42] Hengshu Zhu, Enhong Chen, Hui Xiong, Kuifei Yu, Huanhuan Cao, and Jilei Tian. 2014. Mining Mobile User Preferences for Personalized Context-Aware Recommendation. *TIST* 5, 4 (2014), 1–27.