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]

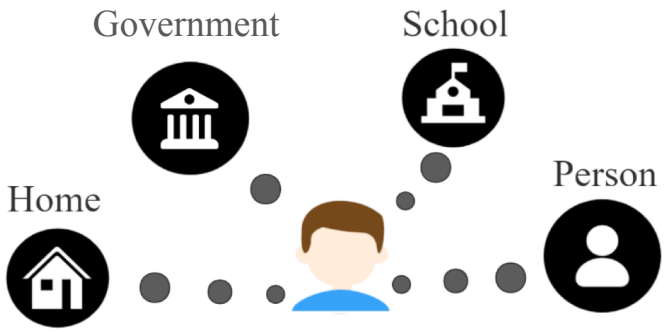
1University of Science and Technology of China,
{qiliuql,huangzhy,cheneh,jianhui}@ustc.edu.cn; {zyq1998,hxwjz,wf314159,tongustc}@mail.ustc.edu.cn;
2 the State University of New Jersey,
hxiong@rutgers.edu

Outline

- Background
- Related work
- Method
- Experiment
- Conclusion

Background

- **Educational context features**
  - Context features related to student learning progress:
    - **Person**: learning habit、self efficiency、test anxiety…
    - **Home**: learning environment、parents' education、wealthy、parents support…
    - **School**: teaching method、learning time、school climate…
    - **Government**: education policy…
    - …

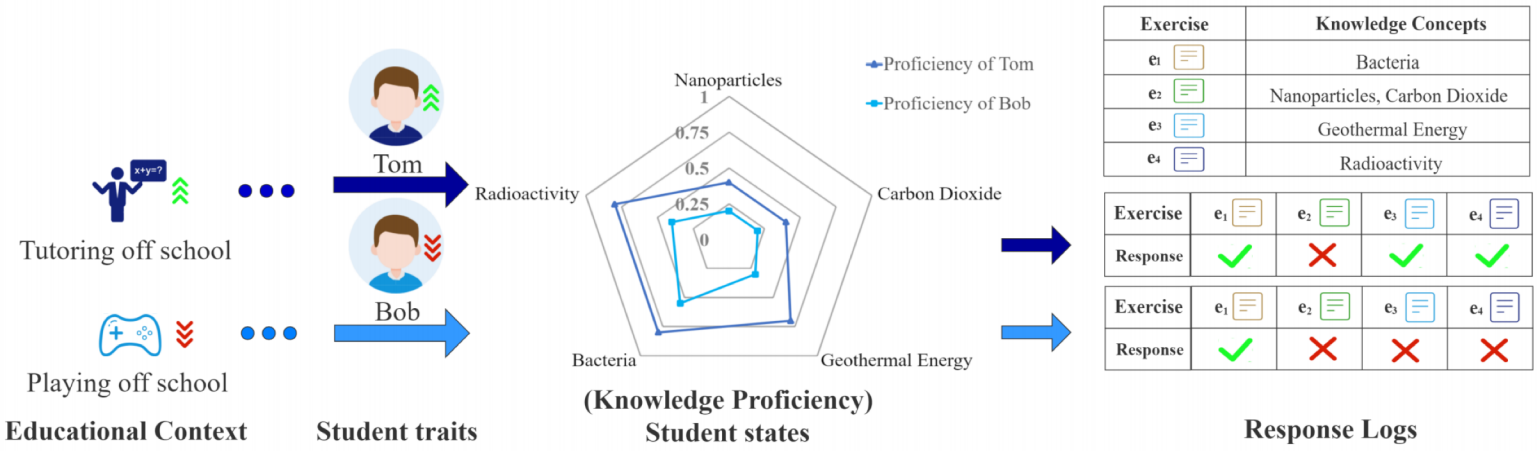


Educational Context

Background

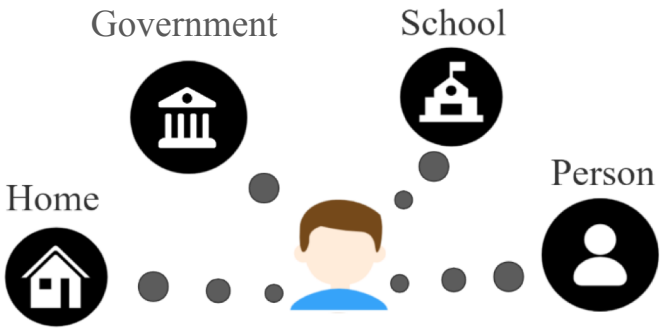
Cognitive diagnosis

- Diagnosing the cognitive states (e.g. knowledge proficiency) of student :
  - the problem of **how educational contexts affect student's knowledge proficiency** is still underexplored in cognitive diagnosis.
  - Meanwhile, influence of educational contexts are widely discussed in traditional education field.



Background

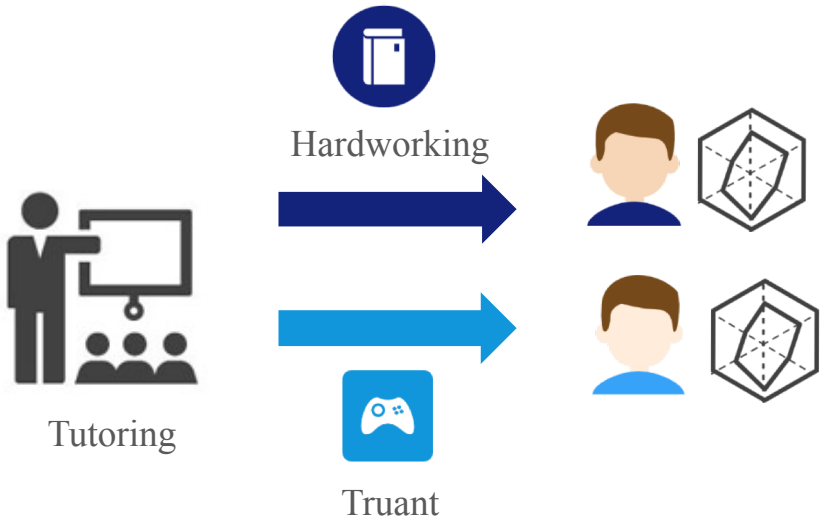
- **Challenge**
  - Educational contexts involve content from various aspects and are not concerned with specific knowledge concept
  - Influence from the same educational context can be personalized
  - Educational contexts may interact with others



Educational Context

Background

- **Challenge**
  - Educational contexts involve content from various aspects and are not concerned with specific knowledge concept
  - Influence from the same educational context can be personalized
  - Educational contexts may interact with others

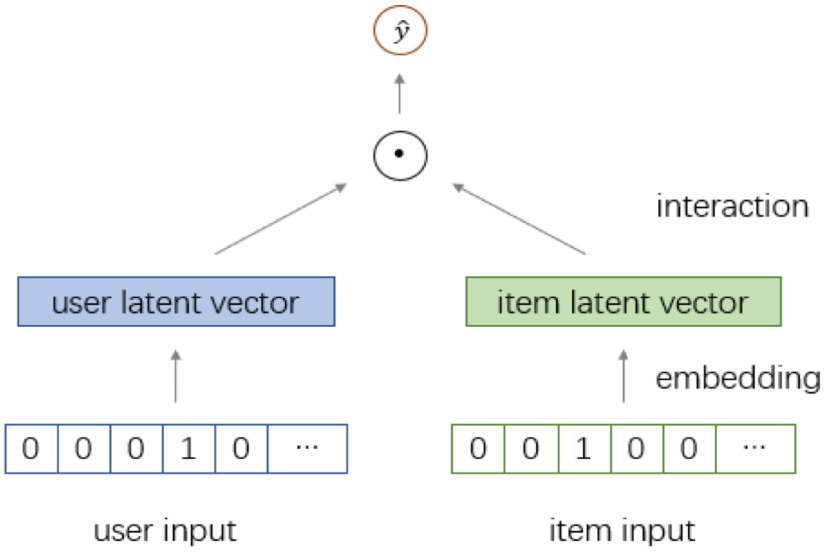


Outline

- Background
- **Related work**
  - Context-aware Modeling
  - Educational context analysis
  - Cognitive Diagnosis
- Method
- Experiment
- Conclusion

Related work

- **Context-aware modeling**
  - An important role in information retrieval (IR) related tasks, ranging from *web search* , *recommendation* to *online advertising*
  - eg: recommendation
    - Probabilistic graphical models
      - Latent Dirichlet Allocation (LDA) related ones
      - **Build the graph manually**
    - Neural network
      - Wide & Deep、Deep FM、NFM
      - DCN、xDeepFM
      - AFM、AutoInt
      - **Ingnore domain knowledge or weak interpretability**

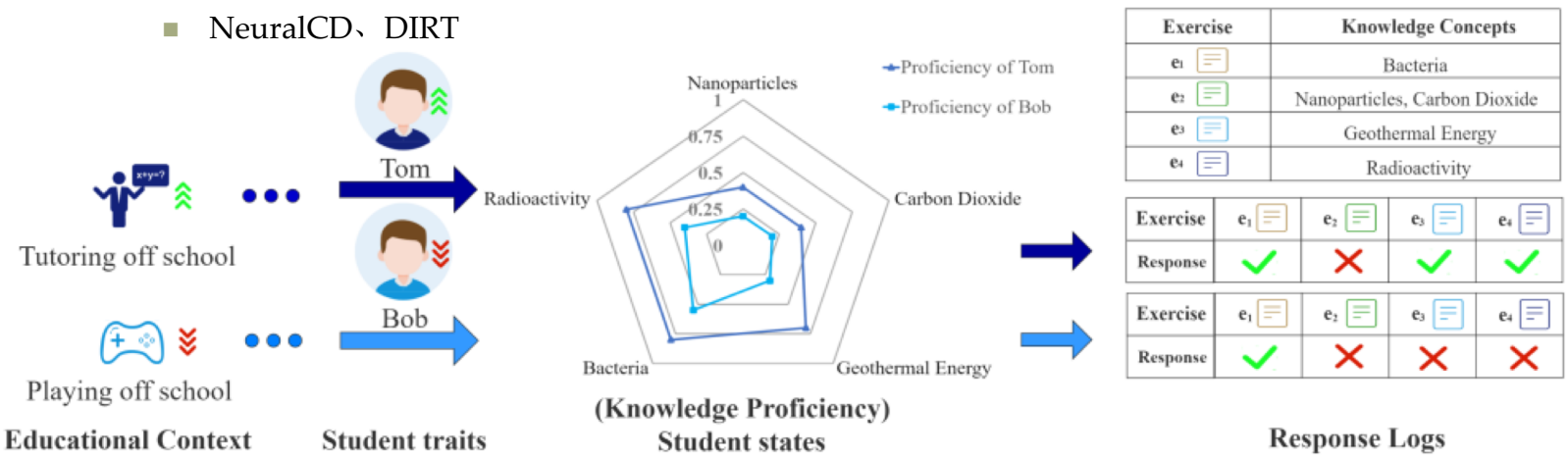


Related work

- **Educational context analysis**
  - Mainly discussed in traditional education domain from the empirical research perspective
    - Propose assumptions->Collect data->Analyze with traditional method(e.g. cluster, linear regression) -> Conclusion.
    - Get educational contexts with **well-designed questionnaire**
  - The relationship between educational contexts and student performance:
    - Measure of performance (e.g. grade) is not flexible.
    - It's hard to analyze complex influence of different contexts

Related work

- **Cognitive diagnosis**
  - Diagnosing the cognitive states(e.g. knowledge proficiency) of student
  - Focus on exercise related information:
    - **Raw response records**
      - IRT、PMF
    - **Q-matrix**(knowledge concept of exercise)
      - MIRT、 DINA、FuzzyCDF
    - **Exercise material**
      - NeuralCD、DIRT



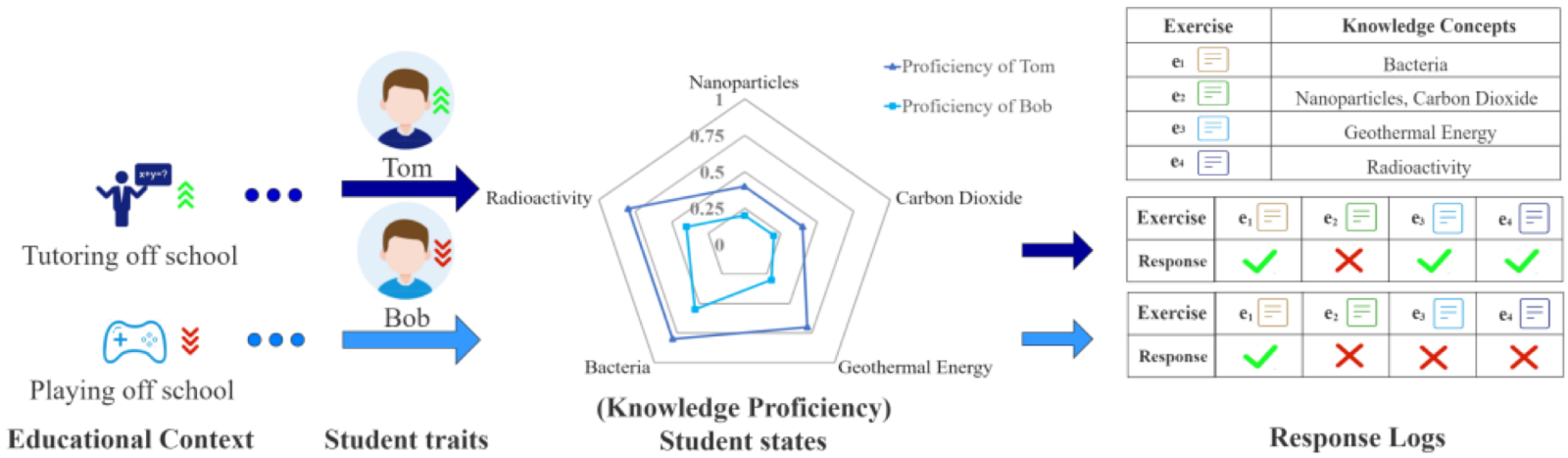
Outline

- Background
- Related work
- **Method**
- Experiment
- Conclusion

Method

Problem Definition

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



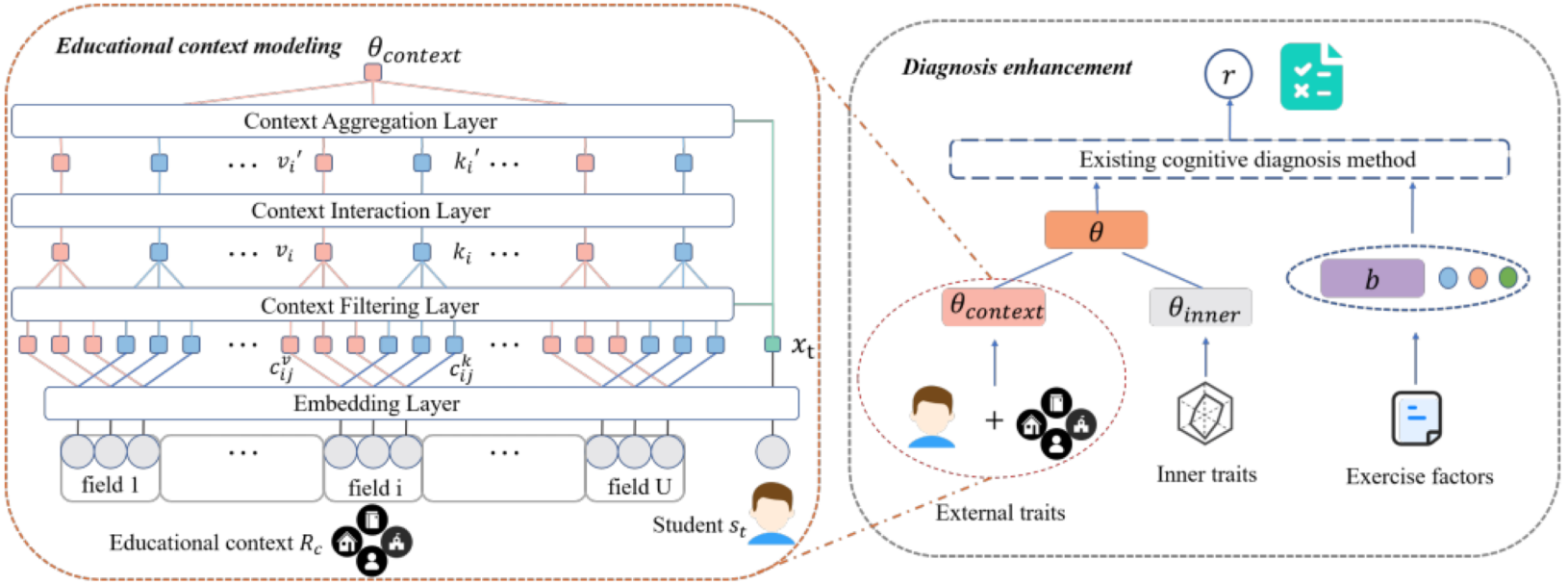
Method

Educational context-aware Cognitive Diagnosis(ECD)

- An two-stage architecture:
  - Educational context modeling
  - Diagnosis enhancement

$$r = F(\theta, \phi_e),$$

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



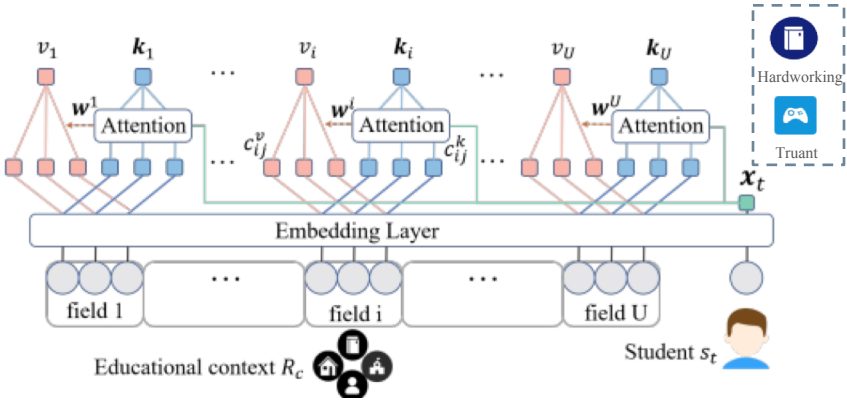
Method

Educational context modeling

- Embedding layer
  - student s_t -> latent character vector x_t
  - Context feature c-> influence key vector c^k and influence value vector c^v
- Context filtering layer

$$w^i = \text{Softmax}(sim(x_t, c_i^k)),$$

$$v_i = \sum_{j=1}^{m_i} w_j^i * c_{ij}^v, \; k_i = \sum_{j=1}^{m_i} w_j^i * c_{ij}^k,$$



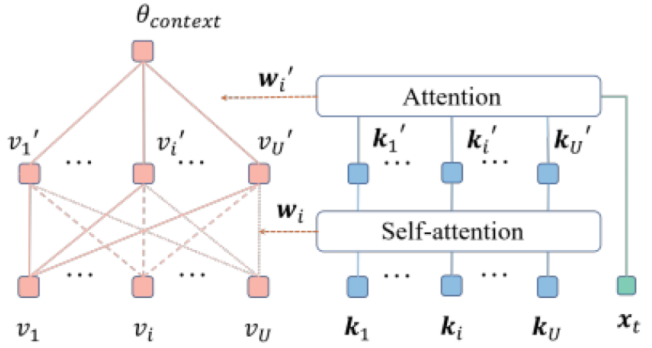
Method

- **Educational context modeling**
  - Context interaction layer

$$\boldsymbol{w}_i = \text{Softmax}(sim(\boldsymbol{k}_i, \boldsymbol{k})), i \in \{1, ..., U\},$$
$$\text{where } \boldsymbol{k} = (\boldsymbol{k}_1; \boldsymbol{k}_2; \ldots; \boldsymbol{k}_U),$$
$$\boldsymbol{k}_i' = \sum_{j=1}^{U} w_{i,j} * \boldsymbol{k}_j,\ v_i' = \sum_{j=1}^{U} w_{i,j} * v_j.$$

  - Context aggregation layer

$$\boldsymbol{w}' = \text{Softmax}(sim(\boldsymbol{x}_t, \boldsymbol{k}')),$$
$$\text{where } \boldsymbol{k}' = (\boldsymbol{k}_1'; \boldsymbol{k}_2'; \ldots; \boldsymbol{k}_U'),$$
$$\theta_{context} = \sum_{i=1}^{U} w_i' * v_i'.$$



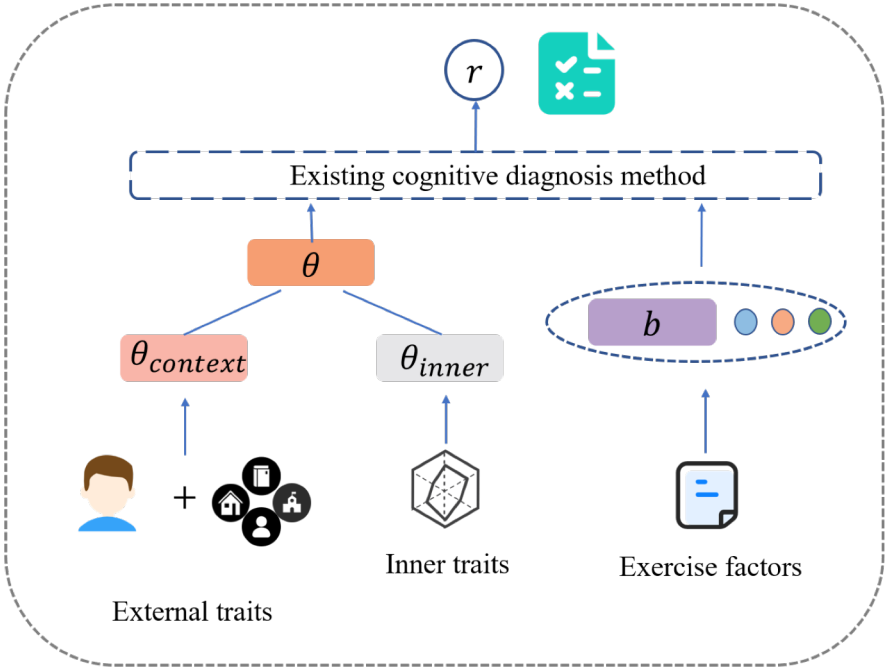
Method

- **Diagnosis enhancement**
  - Student state representation

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

  - Enhance existing method

$$r = \text{CDMethod}(\theta, \phi_e), \phi_e = \{b, \ldots\},$$



Outline

- Background
- Related work
- Method
- **Experiment**
- Conclusion

Experiment

- **Dataset**
  - Programme for International Student Assessment(PISA)
    - honored as Olympic Games in testing project
    - 15-year-olds'
  - PISA 2015:
    - Student questionnaire for educational context
    - Response records of exercise

ST005 **What is the <highest level of schooling> completed by your mother?**

ST005Q01TA

If you are not sure which box to choose, please ask the <test administrator> for help.

(Please select one response.)

<ISCED level 3A>	☐ 1
<ISCED level 3B, 3C>	☐ 2
<ISCED level 2>	☐ 3
<ISCED level 1>	☐ 4
She did not complete <ISCED level 1>	☐ 5

PISA Programme for International Student Assessment OECD

Experiment

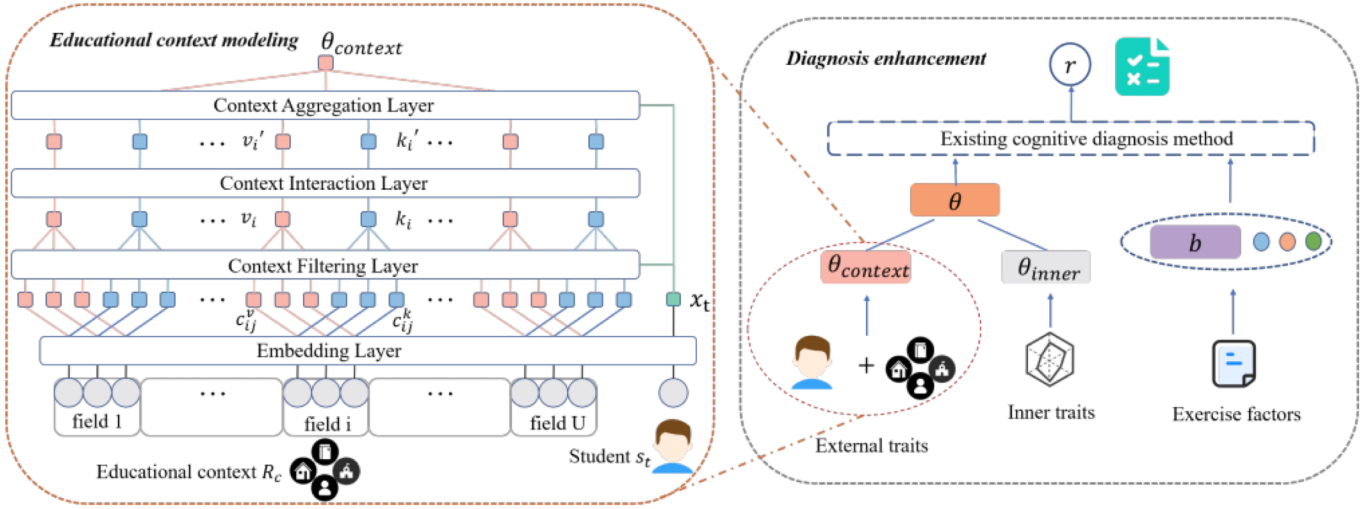
- **Preprocess**
  - Extract three datasets from PISA 2015 by area, namely *Asia*, *Europe*, *America*.
  - Filter out the students whose records are less than 20
  - Statics:

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

Experiment

- **Baseline:**
  - Existing cognitive diagnosis methods:
    - IRT, MIRT, Neural CD
  - Different network in educational context modeling stage:
    - DeepFM, NFM



Experiment

Student performance prediction

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**

Ablation

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

Experiment

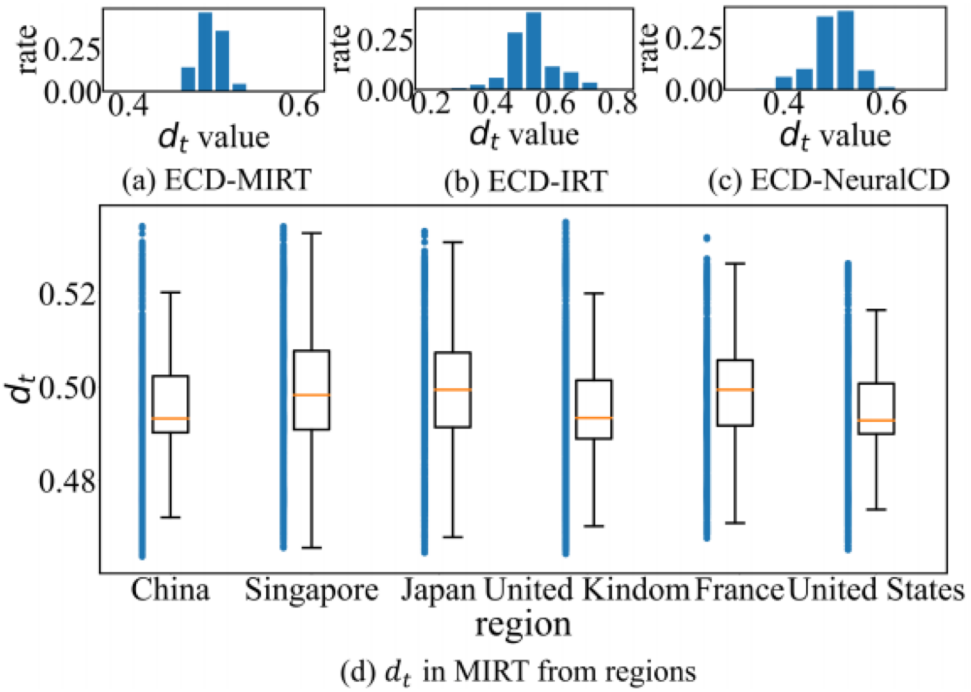
Parameter Analysis

- Distribution of d_t

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

  - d_t denotes the weights of influence from context features
  - Most students have a dt value in [0.4,0.6] for all ECD implements

- Both the context influence and the historic exercise records are not ignorable for a general diagnosis.



Experiment

- **Visualization**

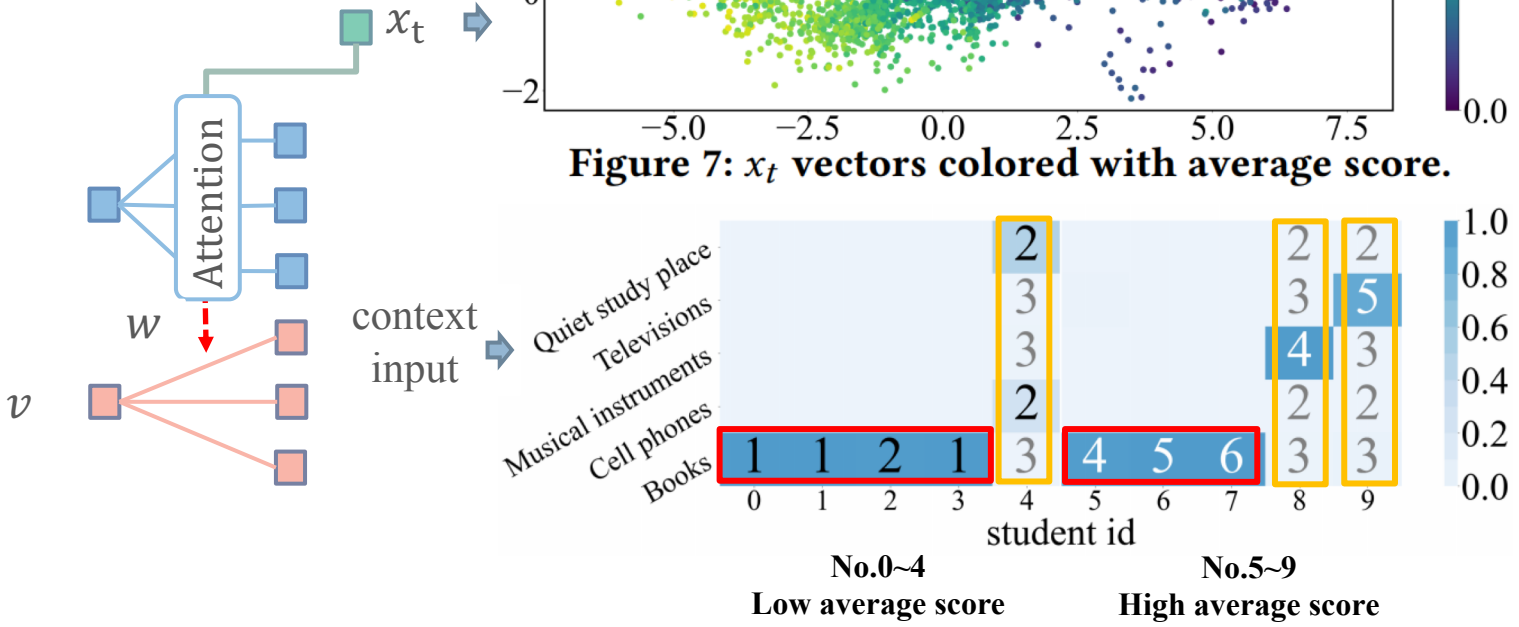


Figure 7: x_t vectors colored with average score.

Experiment

Cognitive States Visualization

Distribution of average knowledge proficiency

score	economy
556	Singapore
538	Japan
534	Estonia
532	Chinese Taipei
531	Finland
529	Macao (China)
528	Canada
525	Viet Nam
523	Hong Kong (China)
518	B-S-J-G (China)
516	Korea
513	New Zealand
513	Slovenia
510	Australia
509	United Kingdom
509	Germany
509	Netherlands
506	Switzerland
503	Ireland
502	Belgium
502	Denmark
501	Poland
501	Portugal
498	Norway
496	United States
495	Austria
495	France

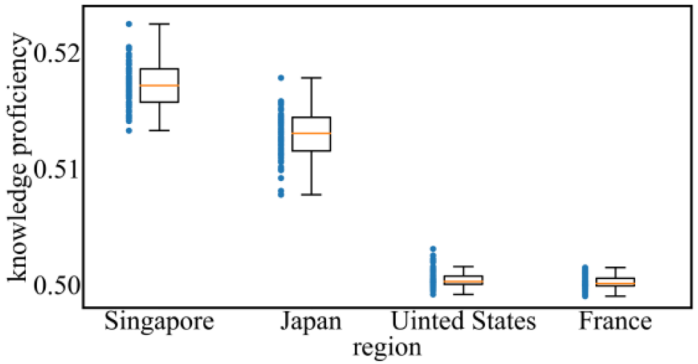


Figure 10: The distribution of average knowledge proficiency in regions.

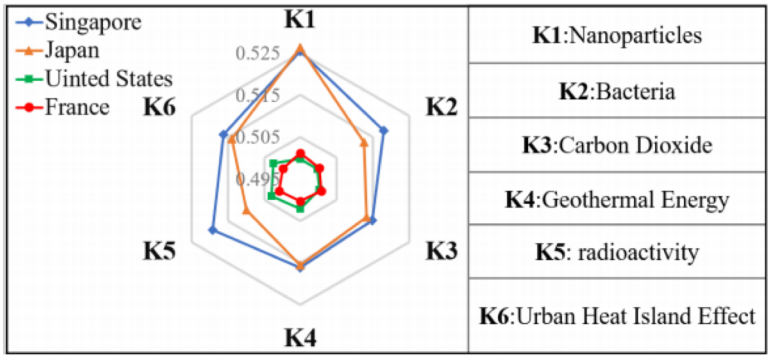


Figure 11: Visualization of average knowledge proficiency in different regions.

Experiment

Discussion

- Record the 3 most important context for each student by w_i'
- Summarize the important context by regions

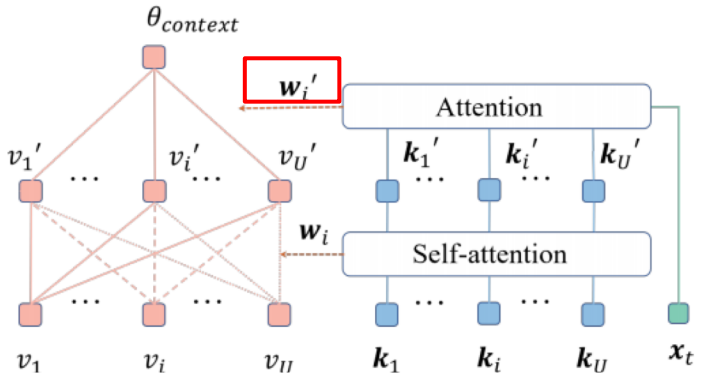


Table 5: Important educational contexts in different regions.

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"

Conclusion

Conclusion

- presented a novel framework ECD for students' cognitive diagnosis, which is also a quantitive perspective for educational context understanding
  - We designed a two-stage solution with a hierarchical attentive network modeling the influence of educational contexts and an adaptive optimization for student traits aggregation
  - We **implemented three specific models** with different existing methods under the framework, (i.e., ECD-IRT, ECD-MIRT, ECD-NeuralCD) and **conducted extensive experiments on real-world datasets** to demonstrate the effectiveness as well as interpretability of ECD framework
  - We analyzed and discussed the difference of influencial context features for students from different regions with our ECD framework

Thanks!