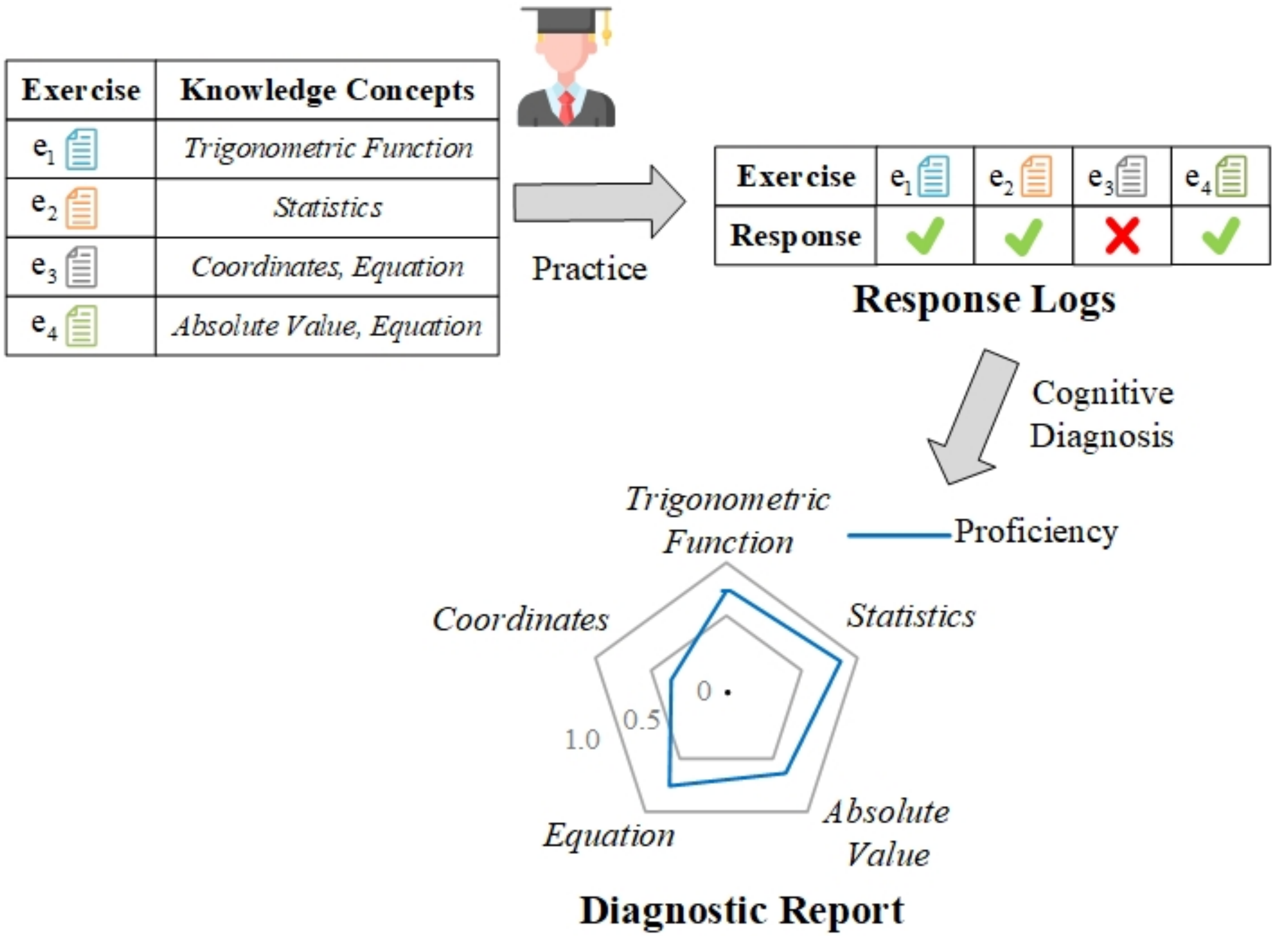
Neural Cognitive Diagnosis for Intelligent Education Systems

Fei Wang, Qi Liu, Enhong Chen, Zhenya Huang, Yuying Chen, Zai Huang, Shijin Wang

AAAI 2020

Background

- **Cognitive Diagnosis**: A fundamental task in many scenarios, especially **intelligent education**



Traditional Works

- **IRT, MIRT**: scalar or latent vectors for students and exercises; logistic-like interaction function

$$P(R_{uv}=1|\theta_u, a_v, b_v) = \frac{1}{1+\exp(-1.7 a_v(\theta_u - b_v))}$$

Skill proficiency | Discrimination | Difficulty

  - difficulty, discrimination, ability
- **DINA**: binary vectors for students and exercises; conjunctive assumption in interaction function

$$P(R_{ij}=1|\alpha_i) = (1-s_j)^{\eta_{ij}} g_j^{1-\eta_{ij}}$$

Skill proficiency vector | Slip | Guess

  - Q-matrix
- **MF**: latent vectors for students and exercises; inner productive interaction function

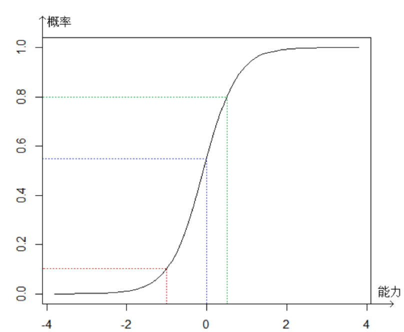
$$P(R_{uv}=1|\theta_u, b_v) = \theta_u \cdot b_v$$

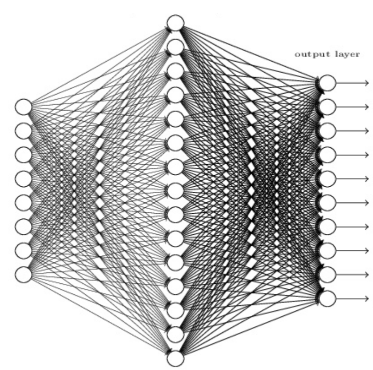
- Problems in the interaction functions:
  - manually designed → labor intensive
  - mostly linear → limited approximation ability
  - simplistic assumptions → restricted scope of applications

It is urgent to find an automatic way to learn the complex interactions for cognitive diagnosis.

Learn the interaction function with neural network from data

$$P(R_{uv}=1|\theta_u,a_v,b_v,c_v)=f(\theta_u,a_v,b_v,c_v)=\frac{1}{1+e^{-1.7a_v(\theta_u-b_v)}}$$

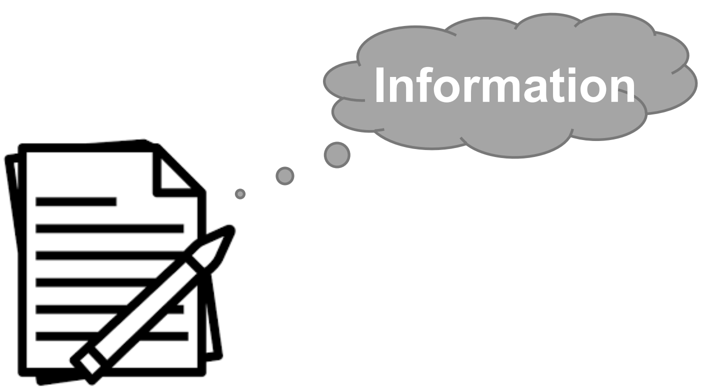




Challenges

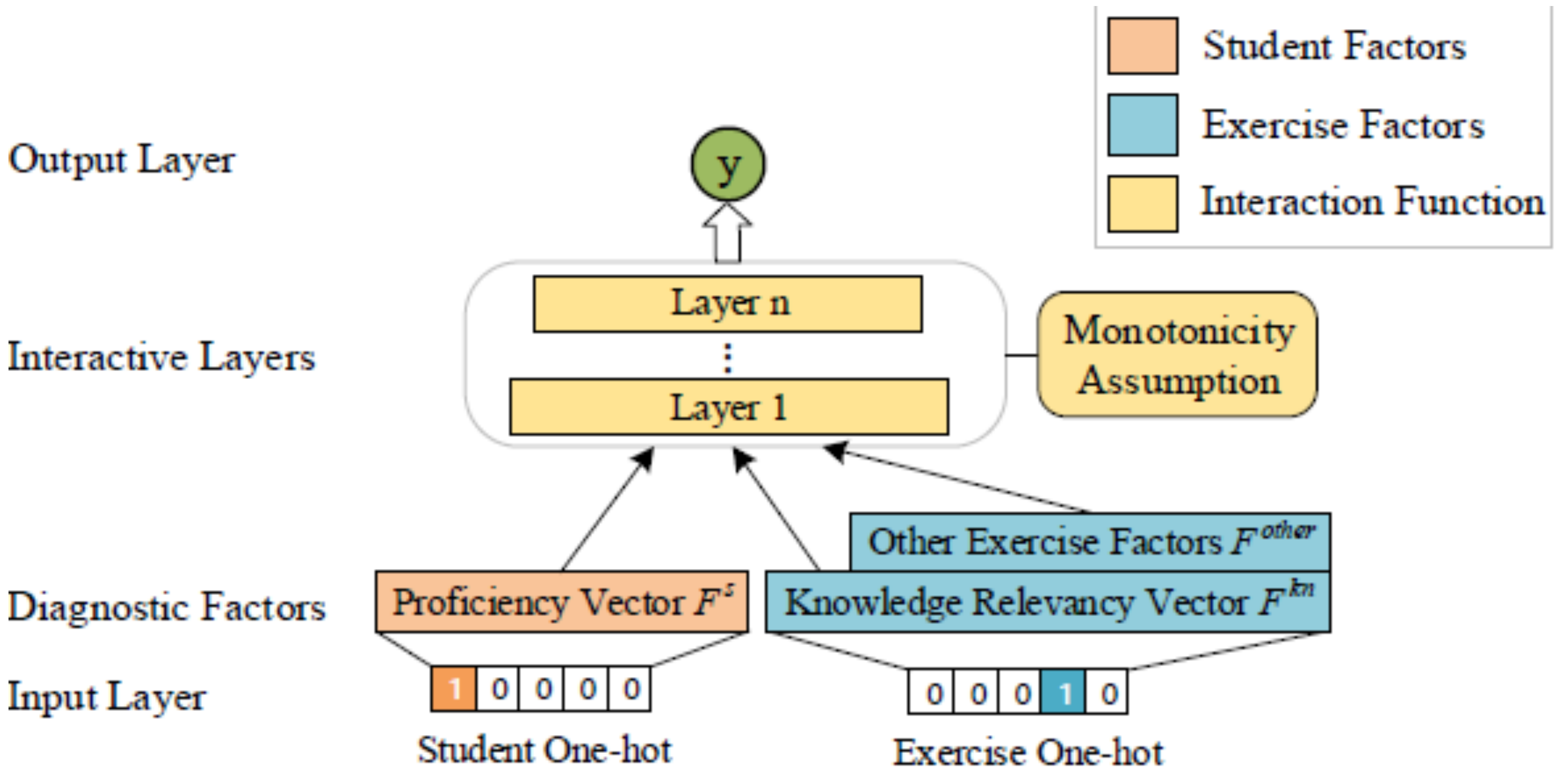
- Black-box nature of neural network
  - difficult to get explainable diagnosis results

- Leverage rich exercise text information
  - difficult for traditional non-neural functions
  - worthy of exploring with the strong ability of neural network



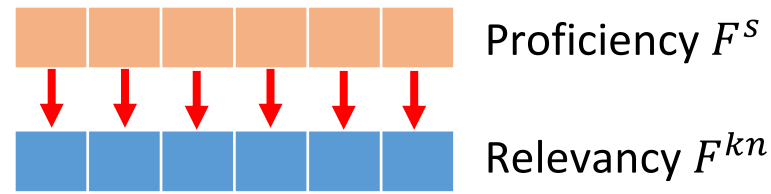
NeuralCD Framework

- **Student Factors**: knowledge proficiency vector F^s
- **Exercise Factors**: knowledge relevancy vector F^{kn}
  - other exercise factors F^{other} (optional): e.g., difficulty, discrimination
- **Interaction Function**: interactive multi layers
- **Output**: the probability that the student would correctly answer the exercise



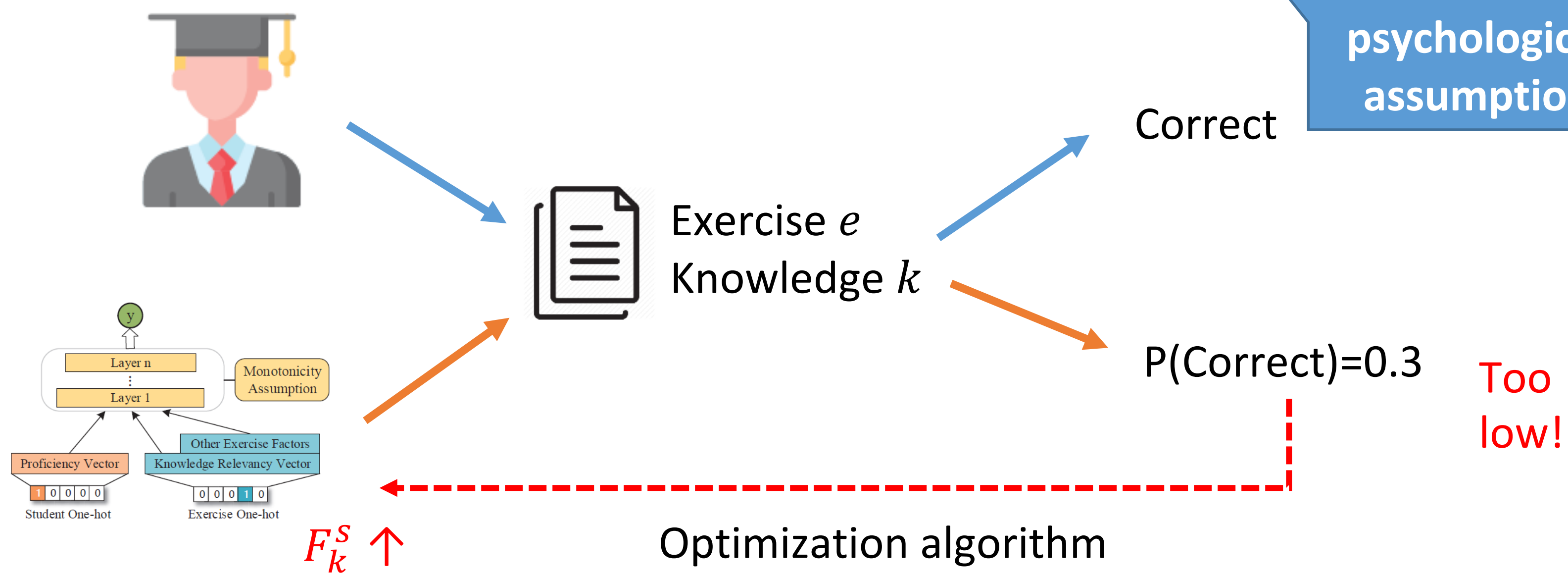
NeuralCD Framework

- **Explainable**
  - $\boldsymbol{F^s \circ F^{kn}}$: attach each entry of F^s to a specific knowledge concept



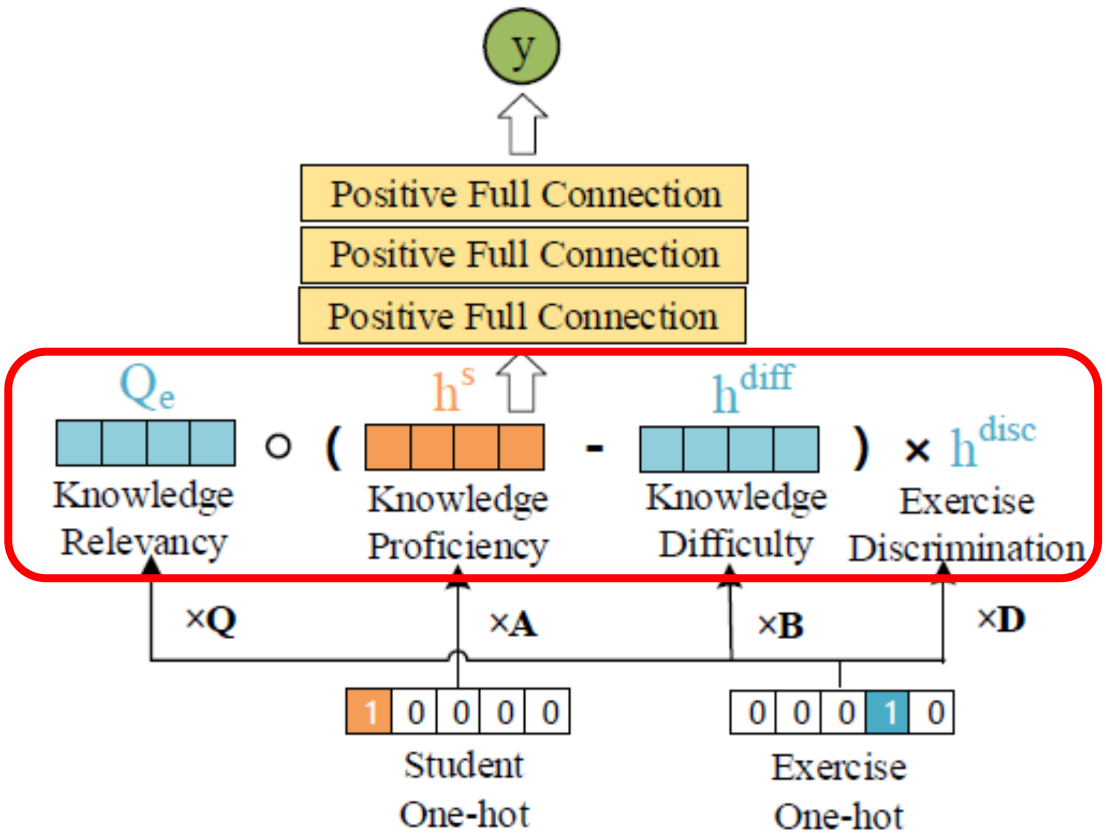
  - **Monotonicity Assumption**: *The probability of correct response to the exercise is monotonically increasing at any dimension of the student's knowledge proficiency.* (widely applicable)

Educational psychological assumption



NeuralCDM

- **Feasible and effective –** basic implementation with Q-matrix



input layer:

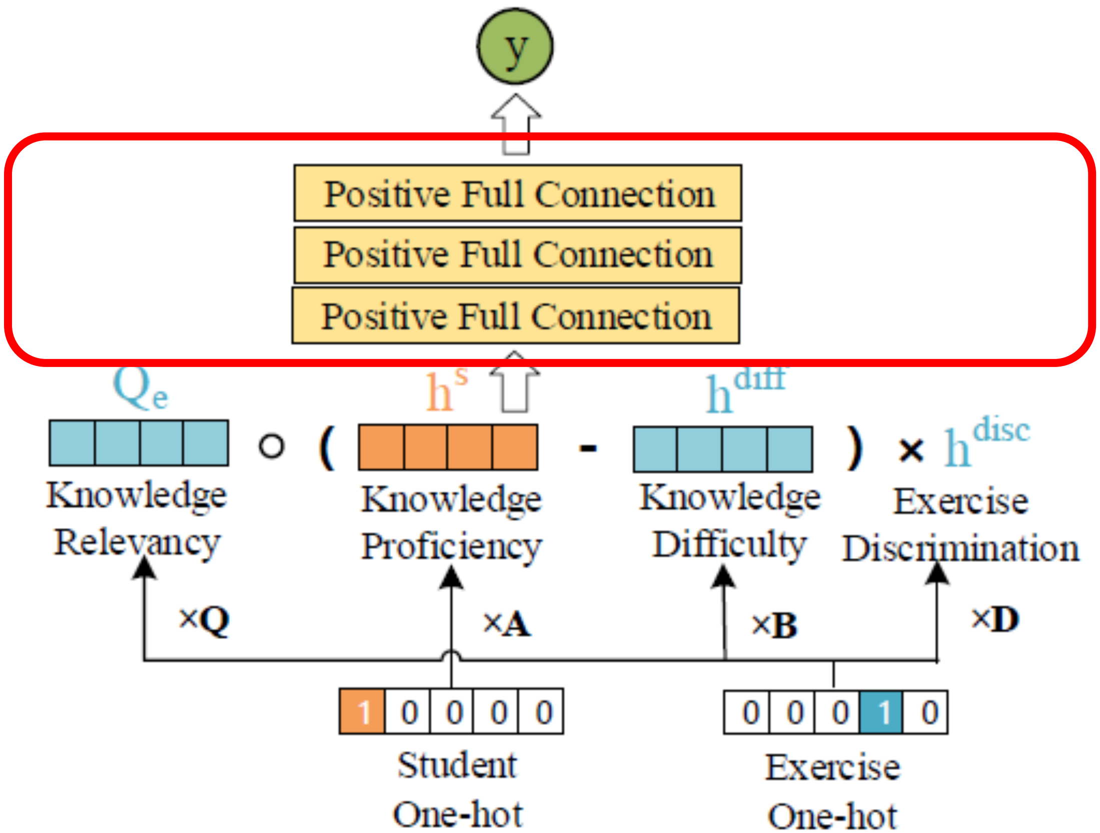
$$\boldsymbol{Q}_e \circ (\boldsymbol{h}^s - \boldsymbol{h}^{diff}) \times h^{disc}$$

$$\boldsymbol{F}^{kn} \circ \boldsymbol{F}^s$$

Directly from Q-matrix

NeuralCDM

- **Feasible and effective –** basic implementation with Q-matrix



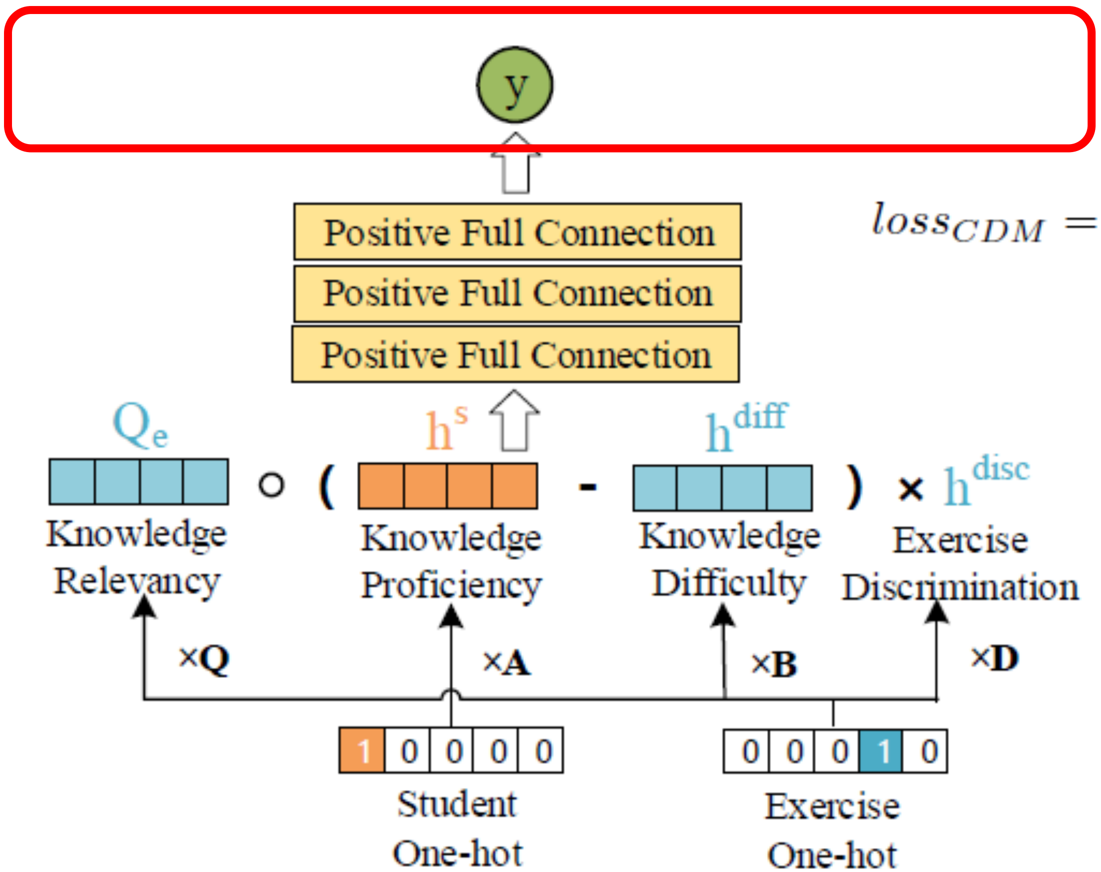
interaction layer:

- full connection
- positive weights

Monotonicity Assumption

NeuralCDM

- **Feasible and effective –** basic implementation with Q-matrix

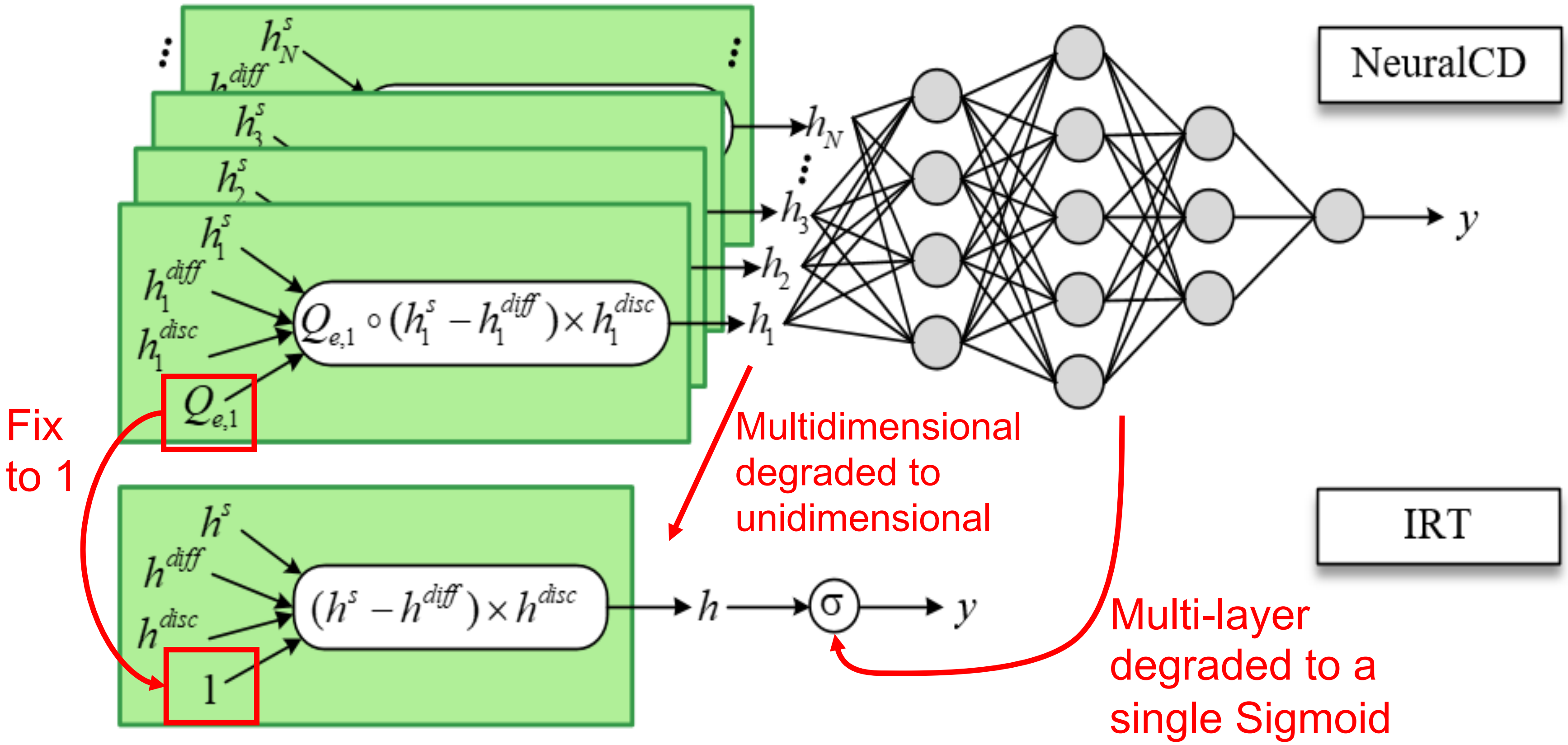


output layer:

$$loss_{CDM} = -\sum_i (r_i \log y_i + (1 - r_i) \log(1 - y_i))$$

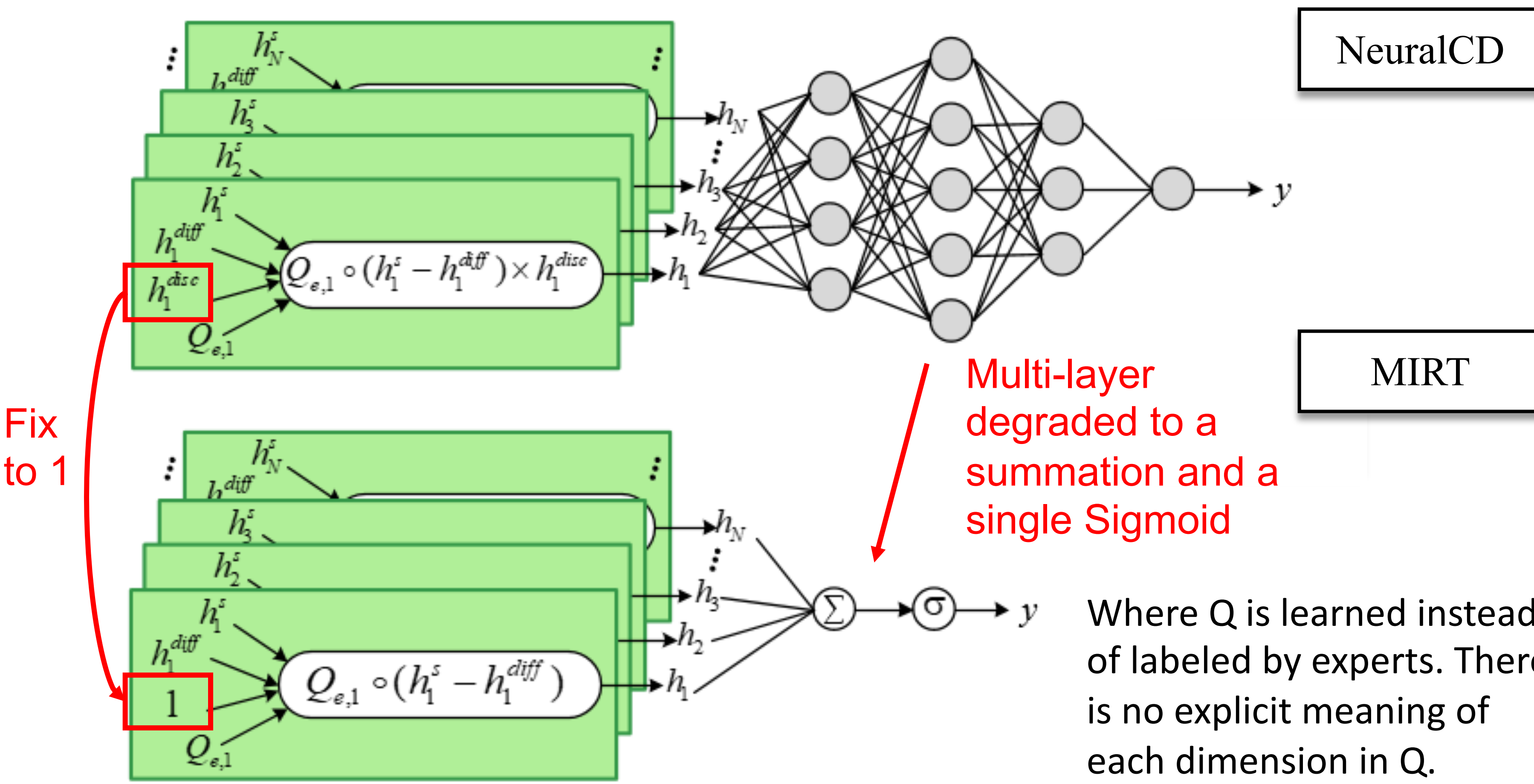
Generality of NeuralCD

- NeuralCD framework is general and can cover some traditional models
  - e.g., IRT, MIRT, MF



Generality of NeuralCD

- NeuralCD framework is general and can cover some traditional models
  - e.g., IRT, MIRT, MF

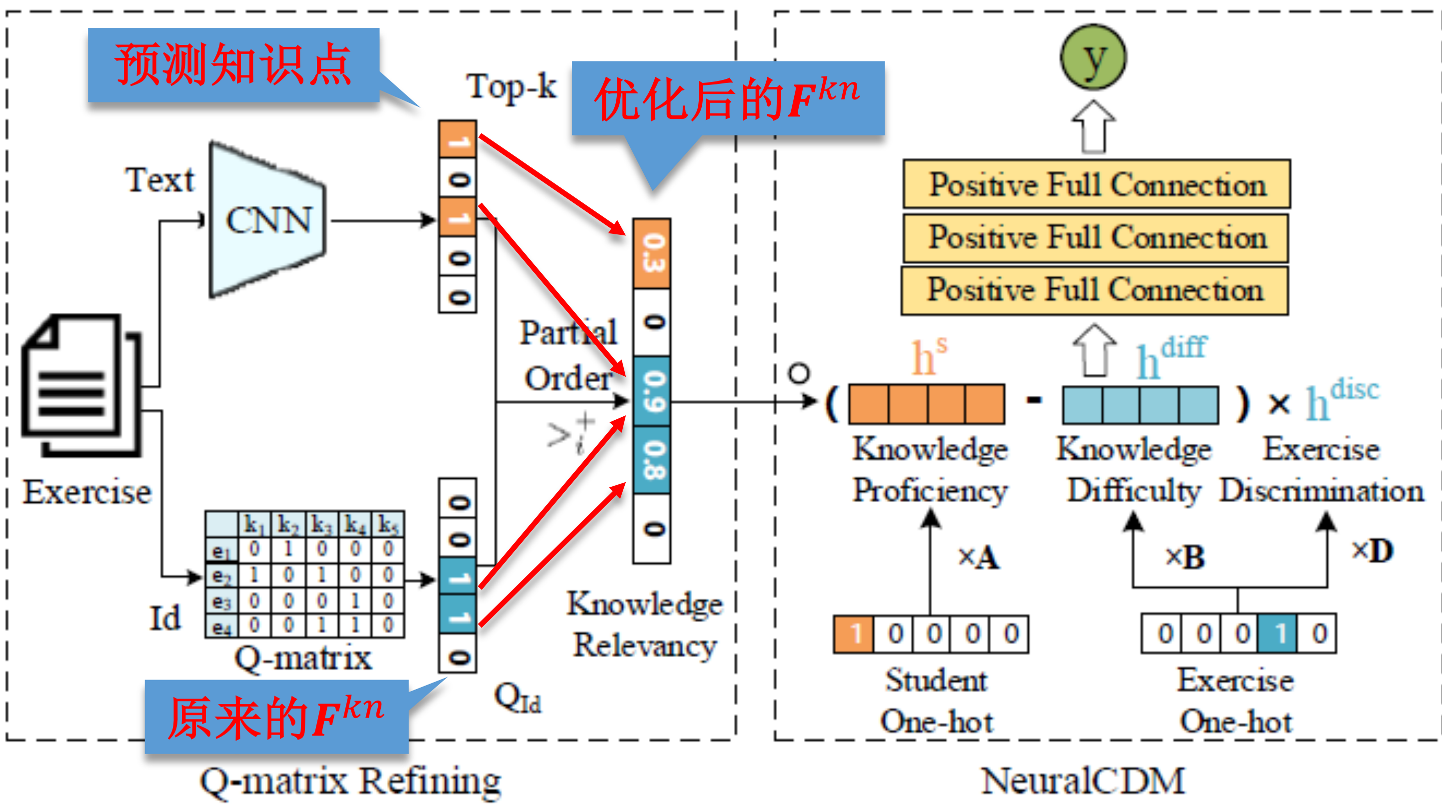


Where Q is learned instead of labeled by experts. There is no explicit meaning of each dimension in Q.

NeuralCDM+

- **Extendible –** refine Q-matrix with exercise texts
  - pre-train a CNN to predict knowledge concepts of the input exercise
  - combine with Q-matrix through a partial order probabilistic scheme:

knowledge relevancy: Q-matrix >= predicted > other = 0



Experiments

Datasets

Dataset	Math	ASSIST
#Students	10,268	4,163
#Exercises	917,495	17,746
#Knowledge concepts	1,488	123
#Response logs	864,722	324,572
#Knowledge concepts per exercise	1.53	1.19
$\mathrm{AVG}_{\#log}$	2.28	8.05
$\mathrm{STD}_{\#log>1}$	0.305	0.316

- **Math**: Zhixue[1], mathematical exercises (with texts) and logs
- **ASSIST**: Assistment[2], mathematical exercises (without texts) and logs

Student performance prediction

	Math			ASSIST		
Model	Accuracy	RMSE	AUC	Accuracy	RMSE	AUC
DINA	0.593±.001	0.487±.001	0.686±.001	0.650±.001	0.467±.001	0.676±.002
IRT	0.782±.002	0.387±.001	0.795±.001	0.674±.002	0.464±.002	0.685±.001
MIRT	0.793±.001	0.378±.002	0.813±.002	0.701±.002	0.461±.001	0.719±.001
PMF	0.763±.001	0.407±.001	0.792±.002	0.661±.002	0.476±.001	0.732±.001
NeuralCDM	0.792±.002	0.378±.001	0.820±.001	**0.719±.008**	**0.439±.002**	**0.749±.001**
NeuralCDM+	**0.804±.001**	**0.371±.002**	**0.835±.002**	-	-	-

Best

[1]Private dataset, provided by iFLYTEK Co., Ltd.

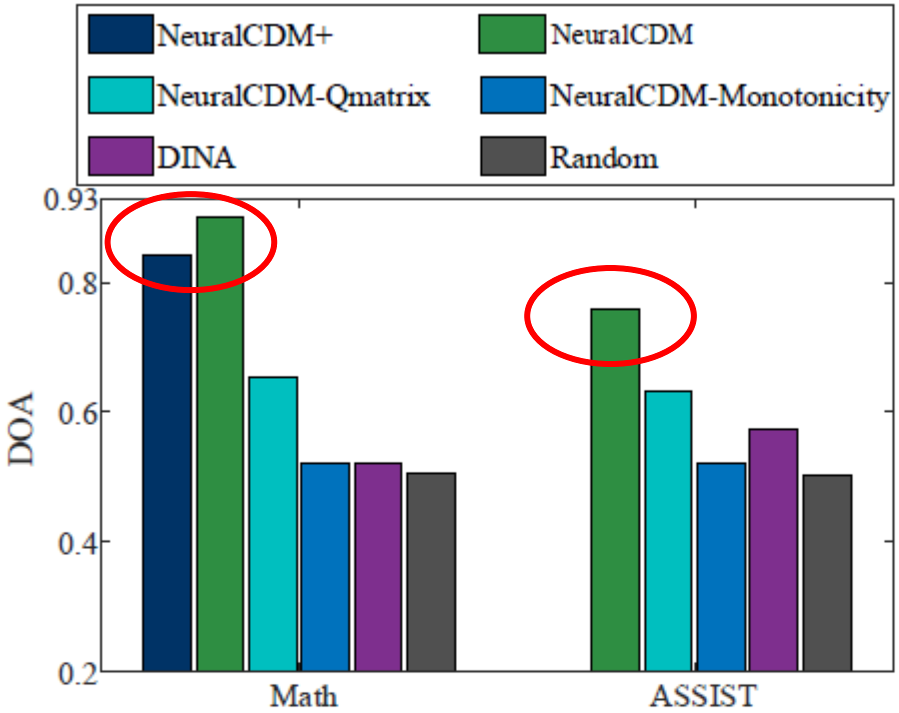
[2]https://sites.google.com/site/assistmentsdata/home/assistment-2009-2010-data/skill-builder-data-2009-2010

Experiments

Model interpretation

- If student a has a better mastery on knowledge concept k than student b, then a is more likely to answer exercises related to k correctly than b.

$$DOA(k) = \frac{1}{Z} \sum_{a=1}^{N} \sum_{b=1}^{N} \delta(F^s_{ak}, F^s_{bk}) \sum_{j=1}^{M} I_{jk} \frac{J(j,a,b) \wedge \delta(r_{aj}, r_{bj})}{J(j,a,b)}$$



Higher DOA: students who perform well on certain knowledge concept get higher diagnosed knowledge proficiency

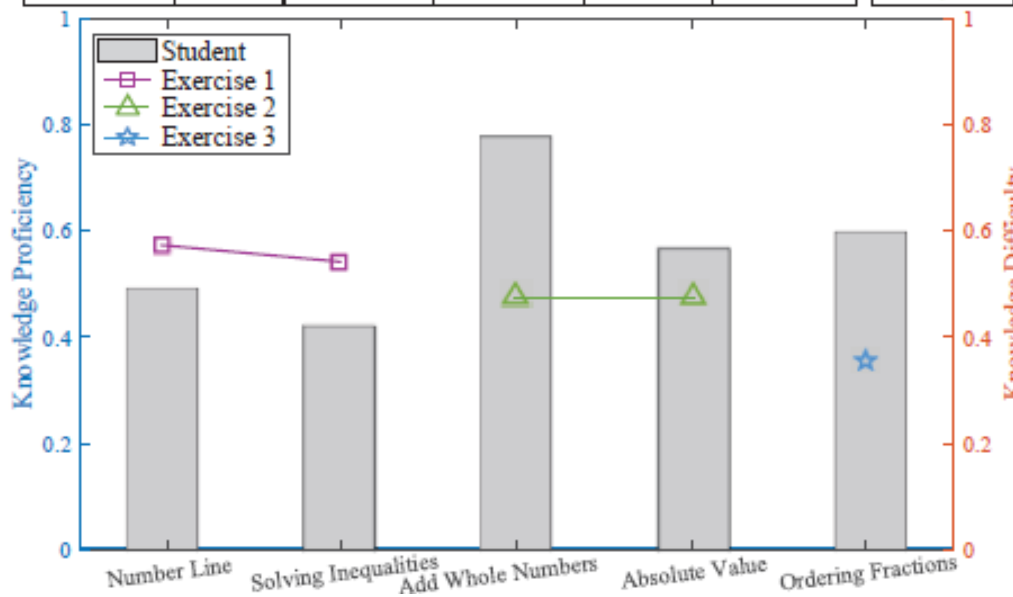
Experiments

- **Case study**
  - a student's performance on 3 exercise in ASSIST
  - and his/her diagnosed result

The student is more likely to response correctly when his/her proficiency satisfies the requirement of the exercise.

Q-matrix | Logs

	Number Line	Solving Inequalities	Add Whole Numbers	Absolute Value	Ordering Fractions	Student Response
Exercise 1	1	1	0	0	0	✗
Exercise 2	0	0	1	1	0	✓
Exercise 3	0	0	0	0	1	✓



Exercise Knowledge Difficulty (points)
Student Knowledge Proficiency (bars)

Conclusion

- We propose a neural cognitive diagnostic framework: NeuralCD
  - student and exercise factors, neural network interaction layers
  - monotonicity assumption

- Feasibility: NeuralCDM with Q-matrix
- Extendibility: NeuralCDM+ with refined Q-matrix that leverages exercise texts
- Generality: covers some traditional models

- Effective and explainable: experiments on two real-world datasets

Code for NeuralCDM is available at https://github.com/bigdata-ustc/NeuralCD

Thank you for listening

Q & A