Convolutional Knowledge Tracing: Modeling Individualization in Student Learning Proces

Shuanghong Shen, Qi Liu, Enhong Chen, Han Wu, Zhenya Huang, Weihao Zhao, Yu Su, Haiping Ma, Shijin Wang

SIGIR 2020

Outline

Background

- Benifiting from the technological progress and large-scale application of **online learning systems**, such as Coursera and MOOCs, the vast majority of students have received or are receiving **computer-assisted learning.**
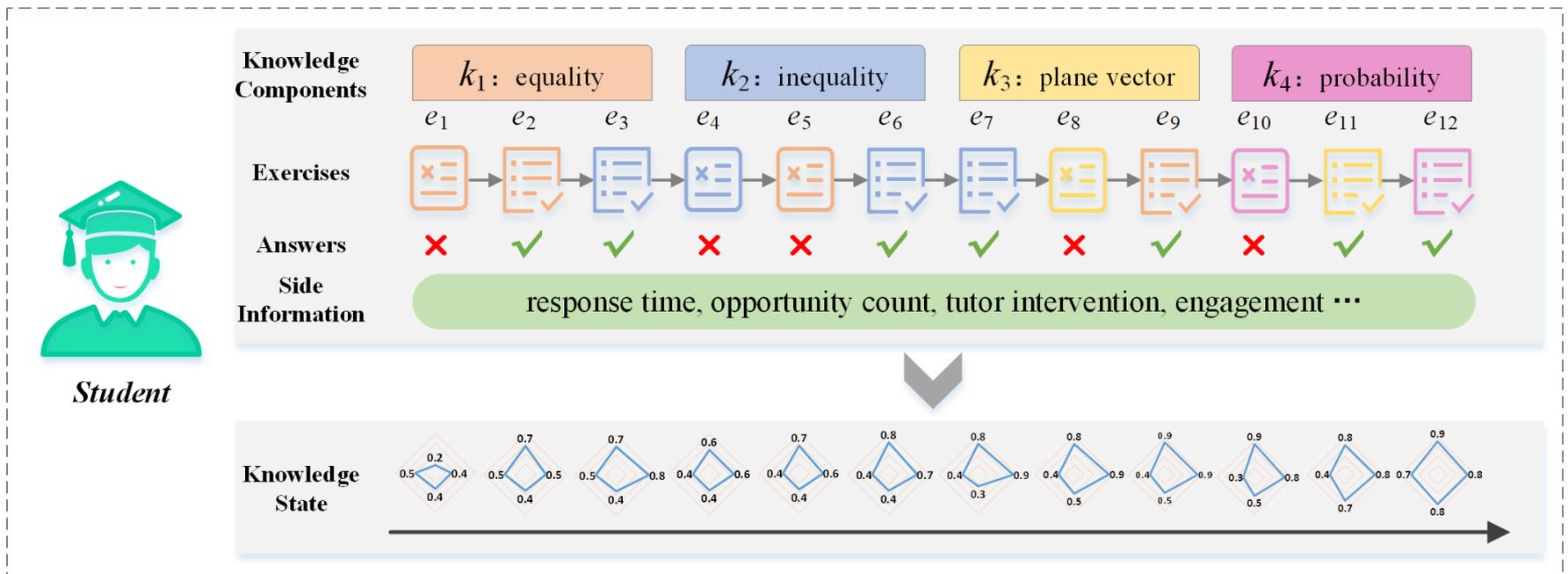
Background

- Knowledge tracing (KT): an emerging research area under this context that aims to assess students' changing knowledge state during learning process.

  - ✓ teaching students in accordance with their aptitude

  - ✓ maximuming the learning efficiency of students

- Traditional methods

  - ✓ BKT: a special case of Hidden Markov Model (HMM)

  - ✓ DKT: introduces deep learning into KT, utilizing RNNs or LSTMs to model students' knowledge state

  - ✓ DKVMN: introduces external memory module to store the knowledge and update the corresponding knowledge state

Problems and challenges

- Individualization problem

  ✓ Students have different prior knowledge

  ✓ The learning rates differ from student to student
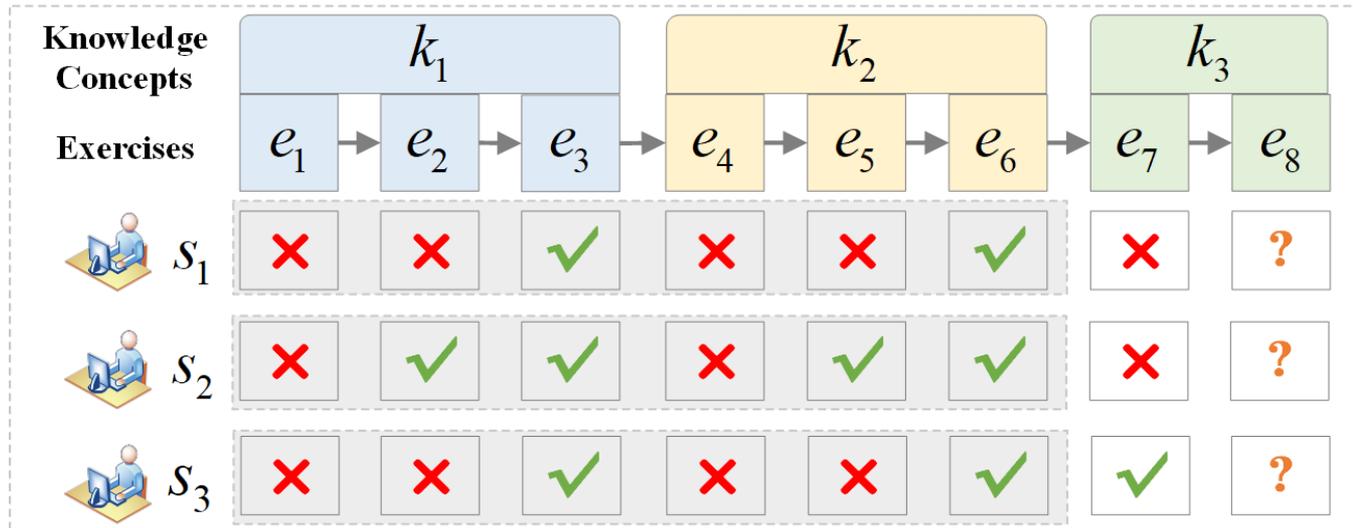
- Without related information given in advance, it is very challenging to measure the individualization of students. We proposed CKT to model the individualization of students with only their learning interactions

Model architecture

- Problem Definition

    Given the learning sequence $X_N = (x_1, x_2, ..., x_t, ..., x_N)$ with *N* learning interactions of a student, KT aims to assess the student's knowledge state after each learning interaction. In the learning sequence, x_t is an ordering pair $\{e_t, a_t\}$, which stands for a learning interaction. Here e_t represents the exercise being answered at learning interaction *t* and $a_t \in (0, 1)$ indicates whether the exercise e_t has been answered correctly (1 stands for correct and 0 represents wrong).

Model architecture

- Embedding

$$x_t = \begin{cases} [e_t \oplus a_t], & \text{if} \quad a_t = 1, \\ [a_t \oplus e_t], & \text{if} \quad a_t = 0, \end{cases}$$

- Individualized Prior Knowledge

  ✓ Historical Relevant Performance: measuring student historical performance relevant to the exercise to be answered

  $$\begin{cases} r_t(i) = Masking(e_i \cdot e_t), & i \in (t, N), \\ w_t(i) = Softmax(r_t(i)), & i \in (1, N), \end{cases} \qquad HRP_t(t) = \sum_{i=1}^{t-1} w_t(i) x_i.$$

  ✓ Concept-wised Percent Correct: accounting for the overall knowledge mastery of student on all knowledge concepts

  $$CPC_t(m) = \frac{\sum_{i=0}^{t-1} a_i^m == 1}{count(e^m)},$$

- Individualized Learning rate

  design hierarchical convolutional layers to extract the learning rate features by processing several continuous learning interactions simultaneously within a sliding window

- Objective Function

  $$L = -\sum_{t=1}^{N} (a_t \log p_t + (1 - a_t) \log(1 - p_t)).$$

Model architecture
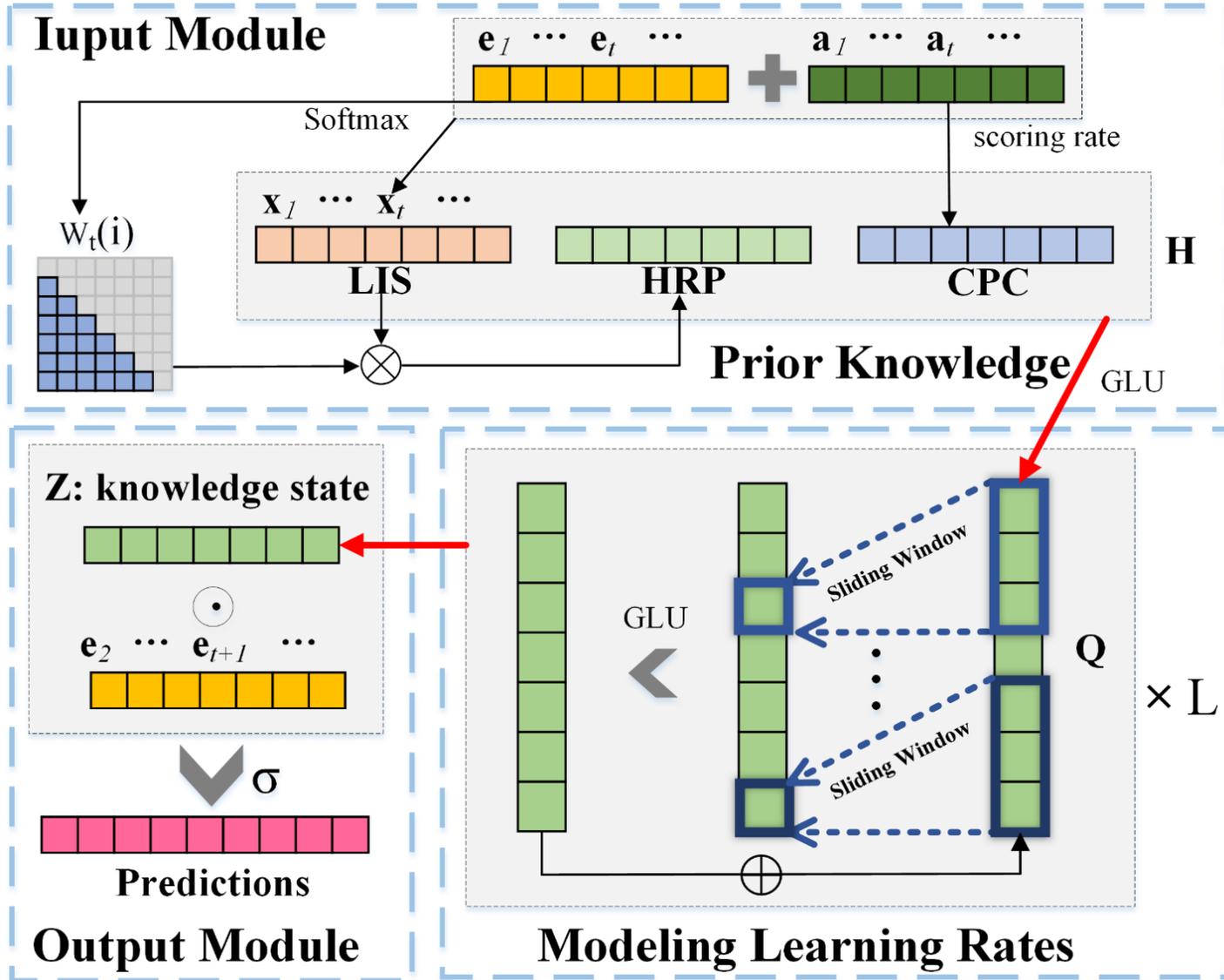
Experiments

- Datasets

Table 1: Statistics of all datasets.

Datasets	Statistics			
	Students	Concepts	Records	Avg.length
ASSIST2009	4,151	110	325,637	78
ASSIST2015	19,840	100	683,801	36
ASSISTchall	1,709	102	942,816	552
Statics2011	333	1,223	189,297	568
Synthetic-5	4,000	50	200,000	50

- Comparison methods
  - CKT-ONE with only one convolutional layer.
  - CKT-HRP measures prior knowledge only from HRP.
  - CKT-CPC measures prior knowledge only from CPC.
  - CKT-ILR only models individualized learning rate.
  - CKT-IPK only models individualized prior knowledge.
  - DKT leverages recurrent neural network to assess student knowledge state [12]. We utilized LSTM in our implemention.
  - DKVMN takes advantage of memory network to get interpretable student knowledge state.

Experiments

- Student performance prediction
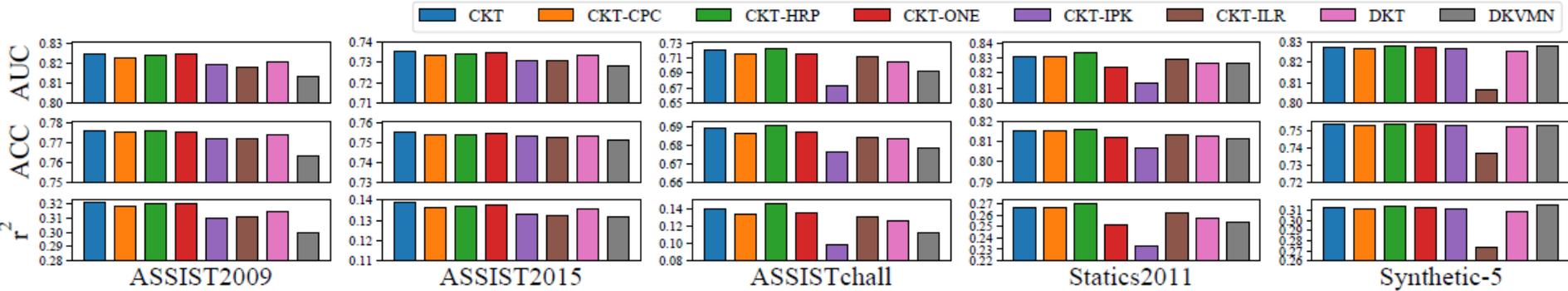
Figure 3: Results of comparison methods on student performance prediction.

- Visualization of knowledge tracing results

16: Probability of a Single Event; 4: Table; 29: Counting Methods; 3: Probability of Two Distinct Events
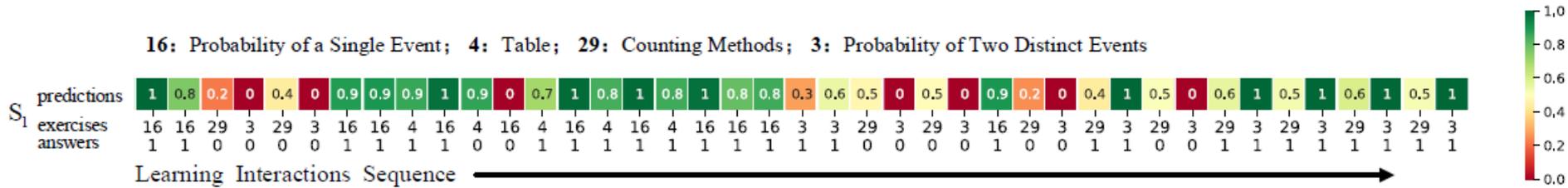
Figure 4: Visualization cases of individualized knowledge tracing result of student.

Experiments

- Exercise embeddings learning

81: Area Rectangle;   83: Area Parallelogram; 41: Finding Percents;
47: Percent Discount; 99: Linear Equations;   97: Choose an Equation

	81	83	41	47	99	97
81	1.9	0.86	0.21	-0.21	0.79	0.62
83	0.86	2.7	-0.63	-0.53	0.4	0.4
41	0.21	-0.63	2.3	0.79	-0.06	-0.63
47	-0.21	-0.53	0.79	1.8	0.18	-0.34
99	0.79	0.4	-0.06	0.18	2.8	1.2
97	0.62	0.4	-0.63	-0.34	1.2	2.6
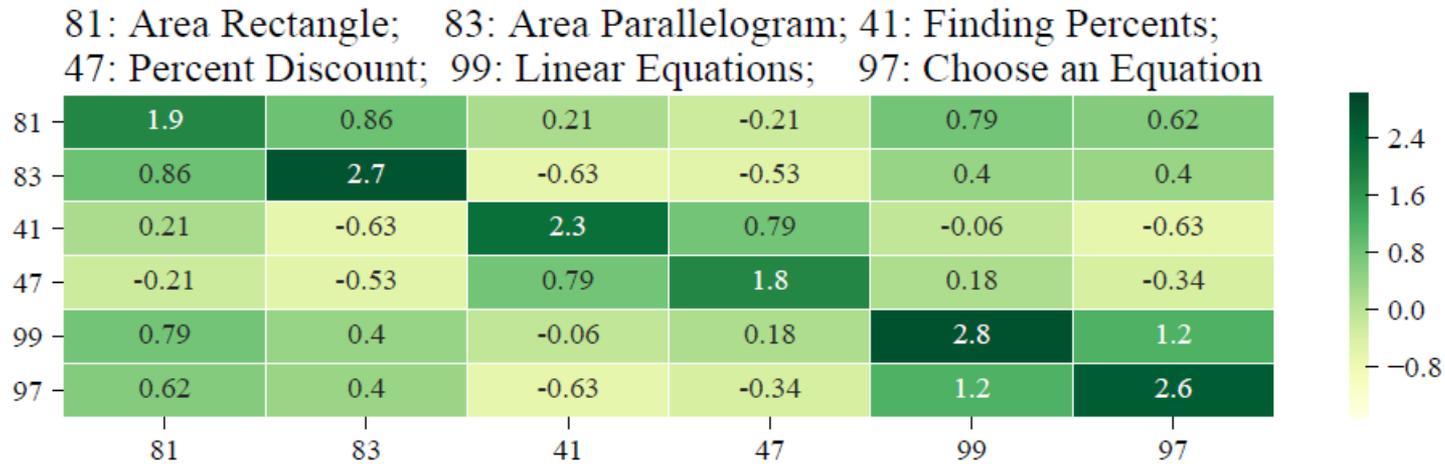
Figure 5: Visualization of exercise relevant coefficients

Conclusion

- We proposed a novel model called Convolutional Knowledge Tracing (CKT) to mode individualization of students in KT task.

- We measured individualized prior knowledge from students' historical learning interactions (i.e., HRP andCPC).

- We designed hierarchical convolutional layers to extract individualized learning rates based on continuous learning interactions.

- Extensive experiment results indicated that CKT could get better knowledge tracing results through modeling individualization in student learning process

Code for CKT is available at:
https://github.com/bigdata-ustc/Convolutional-Knowledge-Tracing

Thank you for listening

Q & A