



中国计算机学会人工智能会议
(CCFAI 2023)

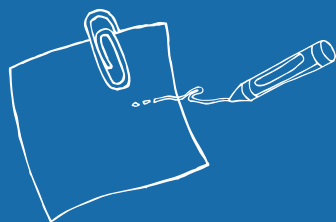


可信的端到端深度学生知识画像建模方法

王士进, 吴金泽, 张浩天, 沙晶, 黄振亚, 刘淇



认知智能全国重点实验室
STATE KEY LABORATORY OF COGNITIVE INTELLIGENCE



2023年7月29日





1 研究背景

2 问题定义

3 模型方法

4 实验分析

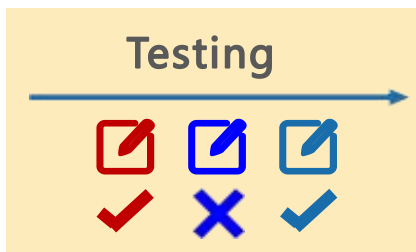
5 结论展望



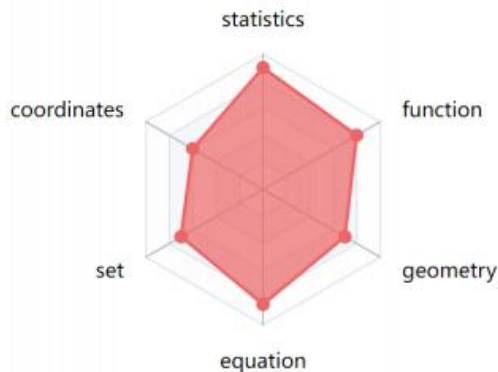
学生知识画像



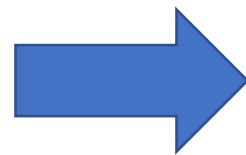
- 用户建模是一项基本任务，它旨在通过分析用户显式的行为数据，从而推断出用户难以观察的隐式画像特征，如能力，偏好，习惯，倾向等等。
- 学生知识画像是学生在不同知识概念上的掌握程度的全面精准的表现，基于这个表示智能教育系统可以为学生提供个性化的服务（习题推荐，学习路径规划等）。



学生作答数据



学生知识画像



- 用列举法表示下列集合：
 - 大于10而小于20的合数；
 - 方程组 $\begin{cases} x+y=1 \\ x^2-y^2=9 \end{cases}$ 的解集。
- 用描述法表示下列集合：
 - 直角坐标平面内X轴上的点的集合；
 - 抛物线 $y=x^2-2x+2$ 的点组成的集合；
- 使 $y=\frac{1}{x^2+x-6}$ 有意义的实数x的集合。
- 含两个元素的数集 $\{a^2-a\}$ 中，实数a满足的条件是。
- 若 $B=\{x^2+x-6=0\}$ ，则 $3 \in B$ ；若 $D=\{x \in \mathbb{Z} | -2 < x < 3\}$ ，则 $1.5 \in D$ 。
- 下列关系中表述正确的是（ ）
 - $0 \in \{x^2=0\}$
 - $0 \in \{(0,0)\}$
 - $0 \in \emptyset$
 - $0 \in \mathbb{N}$
- 对于关系：① $\sqrt{2} \in \{x | x \leq \sqrt{17}\}$ ，② $\sqrt{3} \in \mathbb{Q}$ ，③ $0 \in \mathbb{N}$ ，④ $0 \in \emptyset$ ，其中正确的个数是 A. 4 B. 3 C. 2 D. 1
- 下列表示同一集合的是（ ）
 - $M=\{2, 1\}$ ， $\{3, 2\}$ ； $N=\{1, 2\}$ ， $\{2, 3\}$
 - $M=\{1, 2\}$ ， $N=\{2, 1\}$
 - $M=\{y=y^2+1, x \in \mathbb{R}\}$ ， $N=\{y=y^2+1, x \in \mathbb{N}\}$
 - $M=\{x | x \geq 1, y=x^2-1, x \in \mathbb{R}\}$ ， $N=\{y=y^2-1, x \in \mathbb{N}\}$

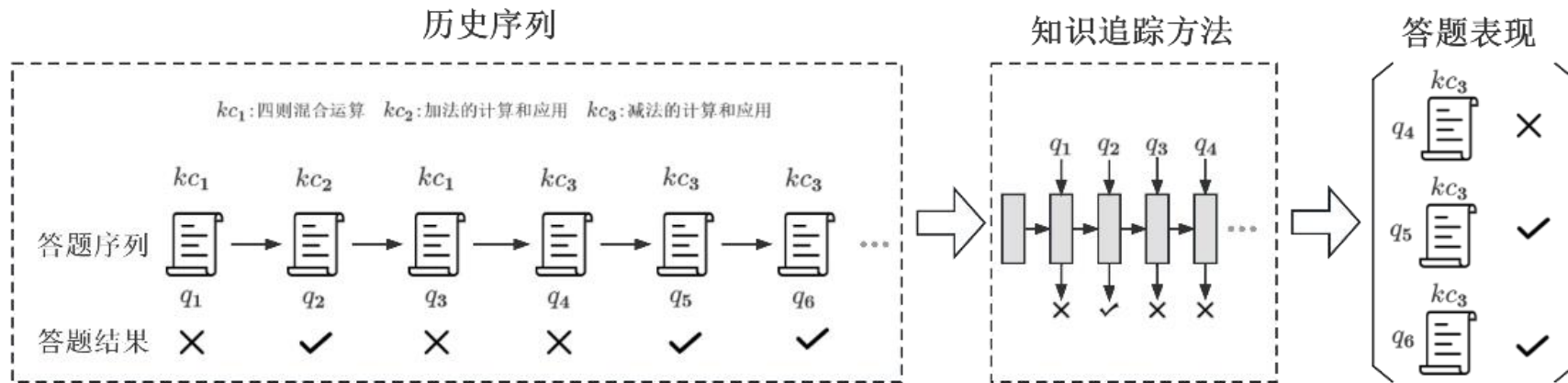
个性化服务



知识追踪



- 现有的学生知识画像建模任务及应用中，往往使用知识追踪（knowledge tracing, KT）类方法刻画学生。
- 知识追踪方法通常首先定义一个学生知识状态向量以表示学生在各个知识概念上的掌握程度，并基于学生每个时刻的表现预测目标，优化该隐式的知识状态向量，使其可以动态地表征学生知识状态变化趋势。

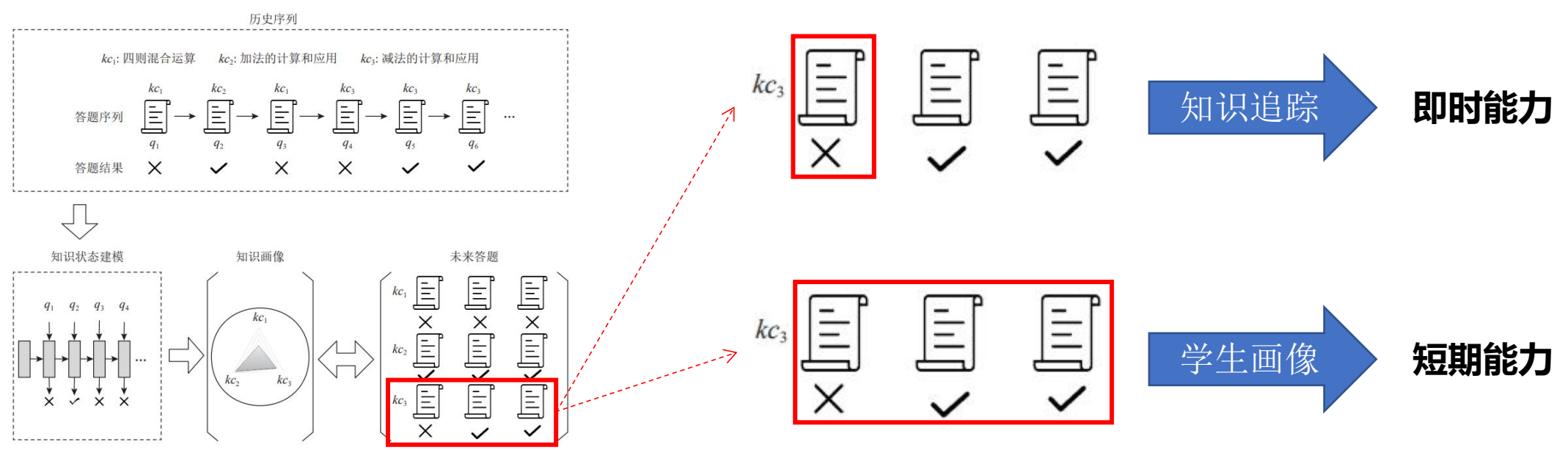




知识追踪



➤ **时序不一致**：知识追踪方法通常就学生答题表现预测这一优化目标进行优化。模型关注的是预测学生**下一时刻**对某个题目的作答情况的即时能力，而学生知识画像期望是一种足以描述学生在**未来一段时间**内表现的学生知识状态。

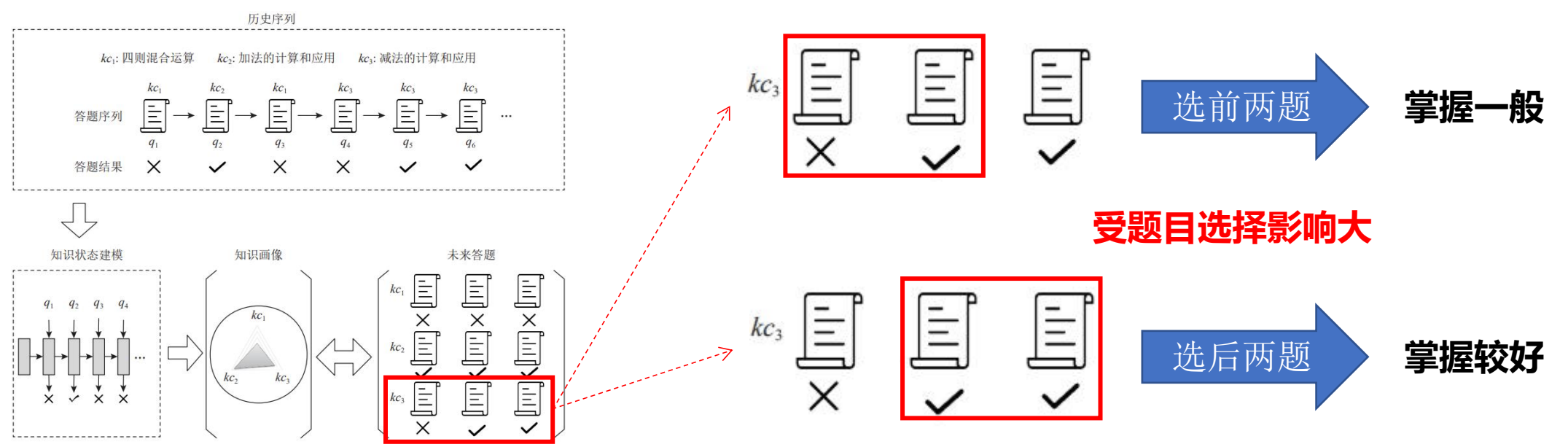




知识追踪



➤ **粒度不一致**：现有的知识追踪方法旨在对下一时刻学生在**某个试题**上的期望作答结果进行准确的预测。在此过程中，知识追踪方法将学生知识状态建模为一种抽象的表示，并假设更准确的试题作答预测与知识状态表示的更好估计是一致的。





1 研究背景

2 问题定义

3 模型方法

4 实验分析

5 结论展望



问题定义



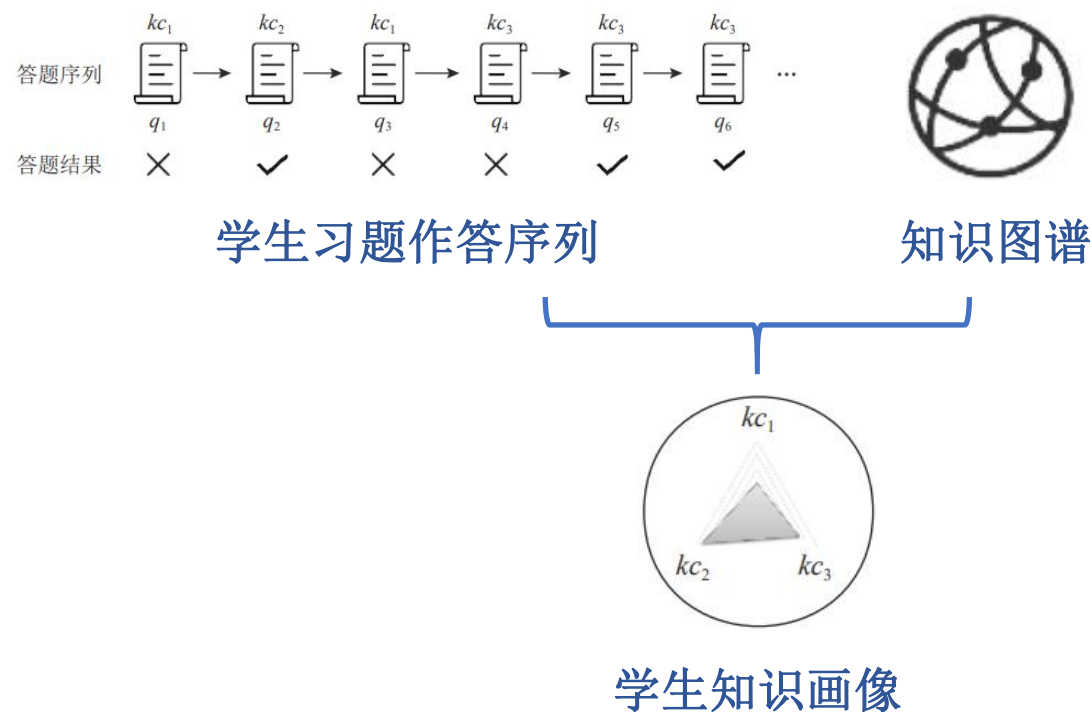
- 本文针对现有方法与学生知识画像预测目标的不一致的情况，为学生知识画像任务构建了**时序、预测粒度一致**的端到端的学生知识掌握度预测目标。

- 输入

- 学生习题作答记录 $R = \{(q, kc, a)\}$
- 知识图谱 RK

- 输出

- 学生知识画像 r





1 研究背景

2 问题定义

3 模型方法

4 实验分析

5 结论展望



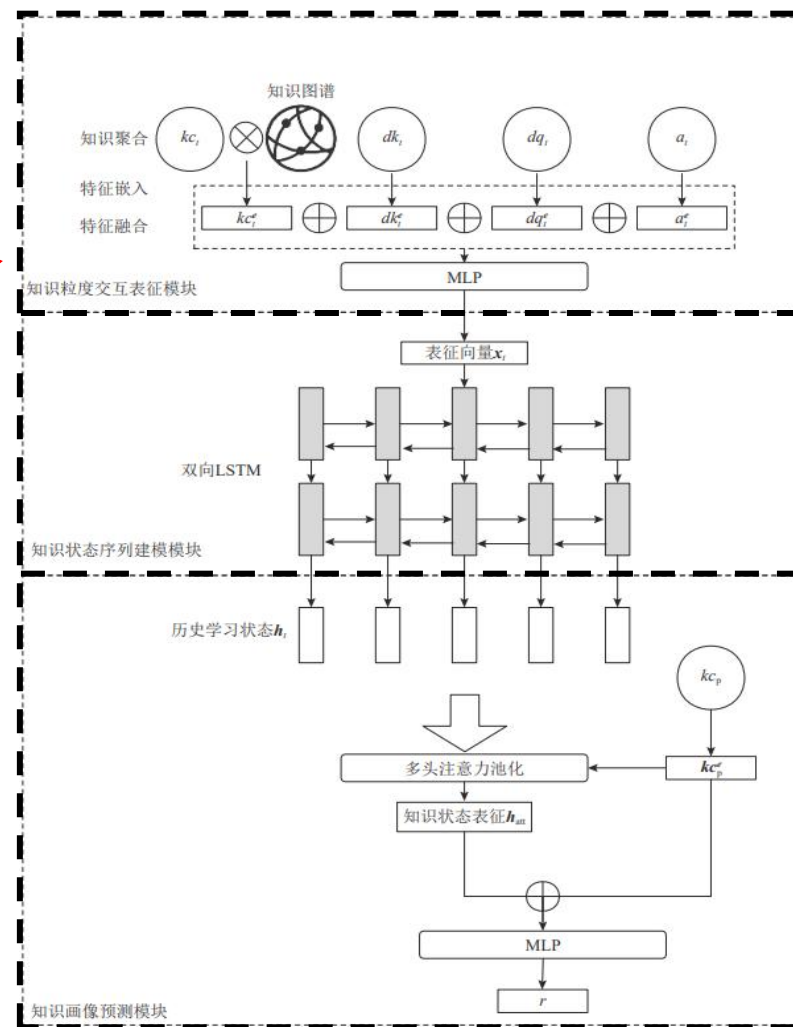
模型方法

➤ 本文提出了用于学生知识概念掌握度预测的端到端模型
Deep Knowledge Portrait (DKP), 模型框架如右图所示:

➤ **知识粒度交互表征模块**

➤ **知识状态序列建模模块**

➤ **知识画像预测模块**





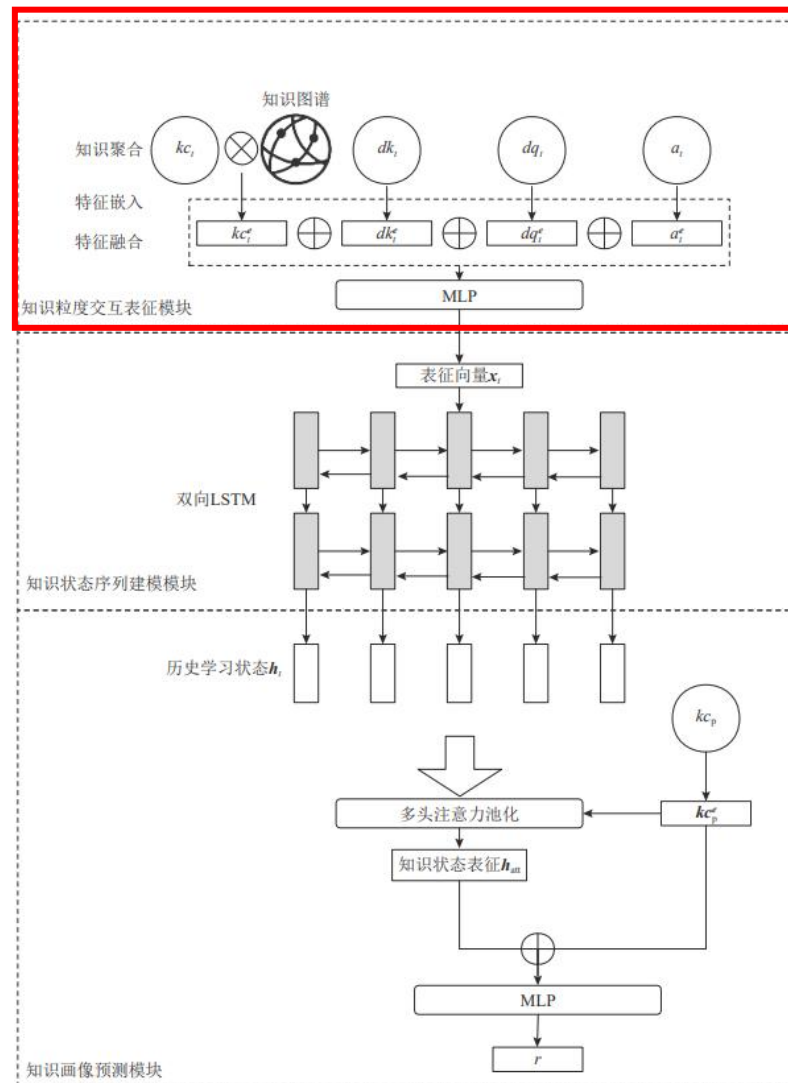
➤ 知识粒度交互表征模块

有效地区分不同知识概念上的交互

➤ 知识聚合 $\tilde{k}c_i^e = kc_i^e + RK_i \cdot EK.$

➤ 特征嵌入
$$dq = \frac{\sum_{i=0}^{|S_i|} a_i}{|S_i|} \bmod \lambda, dk = \frac{\sum_{i=0}^{|S_i|} a_i}{|S_i|} \bmod \lambda$$

➤ 特征融合
$$x_i = W_1^T [\tilde{k}c_i^e \oplus dq_i^e \oplus dk_i^e \oplus a_i^e] + b_1$$





模型方法



➤ 知识状态序列建模模块

挖掘知识粒度上的学生知识状态变化

知识追踪

关注**连续时刻**学生的表现

预测时保障**未来信息不可见**

连续单向单步推断

单向LSTM

知识画像

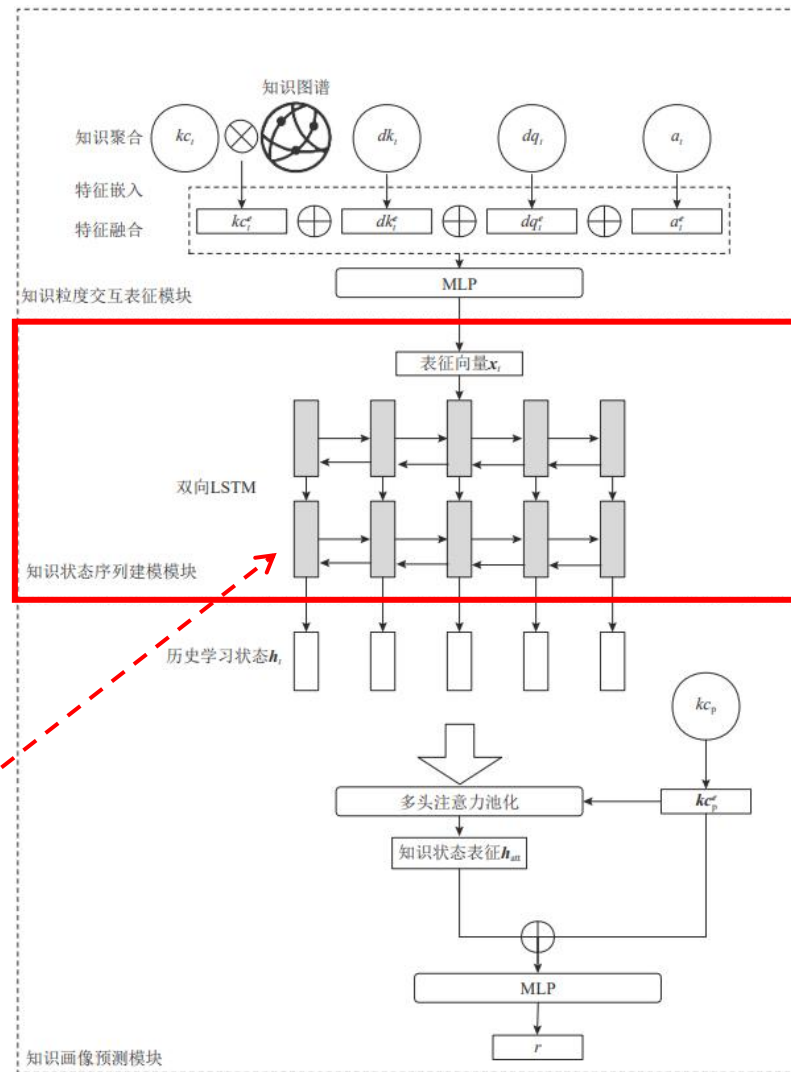
阶段性诊断，预测**时刻不连续**

画像时所有**序列均为可见的**

阶段静态序列推断

双向双层LSTM

$$h_t = [\vec{h}_t, \overleftarrow{h}_t]$$





模型方法



➤ 知识画像预测模块

给予不同的历史状态不同的注意力

➤ 相似度计算

$$\alpha_i = W_a[W_n h_i, W_n k c_p^e],$$

➤ 注意力状态

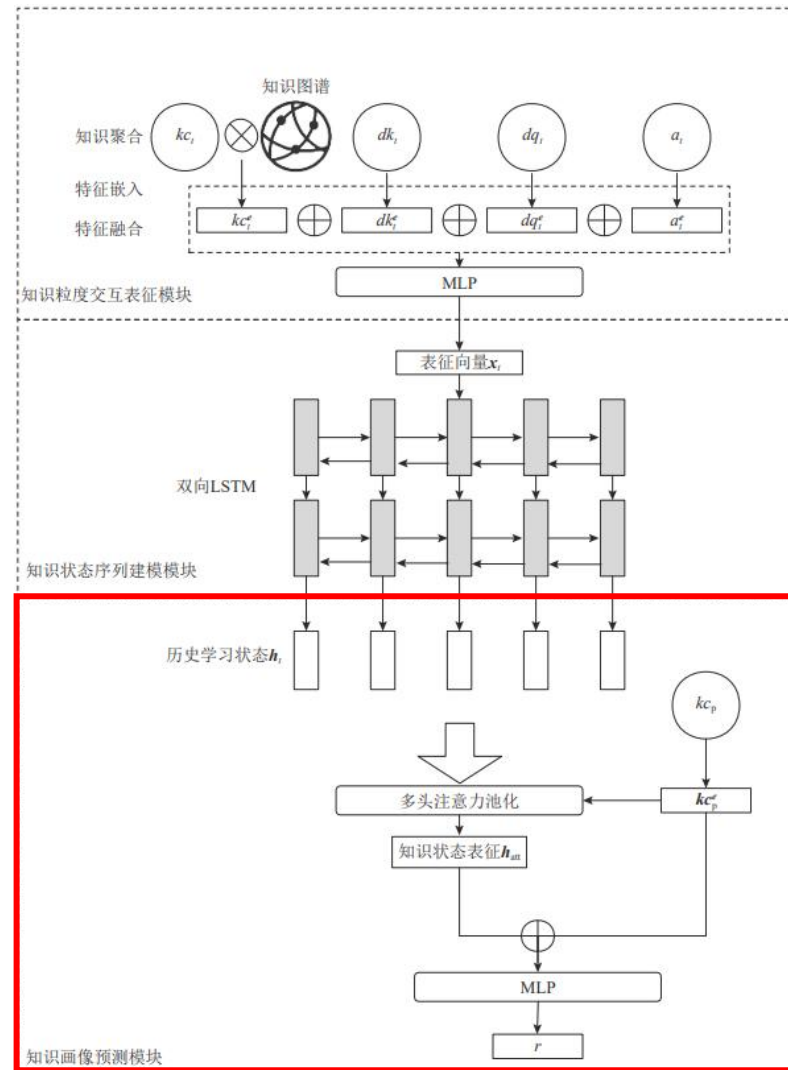
$$h_{\text{attn}} = \sum_{i=0}^t h_i \cdot \frac{\exp(\text{LeakyReLU}(\alpha_i))}{\exp\left(\sum_{j=0}^t \text{LeakyReLU}(\alpha_j)\right)}$$

➤ 画像值计算

$$r = W_2^T [k c_p^e \oplus h_{\text{attn}}] + b_2$$

➤ 损失值计算

$$\text{loss} = -\sum (y \log(r) + (1 - y) \log(1 - r)).$$





1 研究背景

2 问题定义

3 模型方法

4 实验分析

5 结论展望



➤ 数据集

➤ Assist

收集自在线教辅平台ASSISTments数据.

➤ JunYi

收集自基于可汗学院发布的开源代码和 2012 年建立的在线学习平台 Junyi Academy 中的答题日志.

➤ MATH

收集自某知名企业 2019—2022 年初中学生的在线答题记录.

数据集	Assist	JunYi	MATH
学生数	4217	247606	16661
知识点数	124	39	1427
题目数	26687	720	49853
交互数	401756	25925992	4319270



实验分析



➤ 画像场景构建

➤ MATH

使用该数据中自然存在的学生测评时刻作为画像点，使用画像点前最长 150 条交互记录作为输入，对于短于该长度的序列，使用零向量将其填充到固定长度；画像点后最长 50 条交互记录作为预测序列，其中出现的知识概念作为待画像知识概念

➤ Assist&JunYi

每隔 50 个交互记录定义一个画像点，划分同上

➤ 真实画像值

定义为预测序列中每个知识概念**前 3 次交互**的整体作答情况

数据集	画像计算题数	类别	未来表现
MATH	2	正例	0.823 4
		负例	0.672 5
	3	正例	0.804 9
		负例	0.609 9
	5	正例	0.799 6
		负例	0.658 6

得分一致性最高





实验分析



➤ 对比试验方法

知识追踪类 画像方法

隐层状态预测

学生知识状态作为学生知识画像，其中知识状态的每个位置的标量表示了对对应知识概念的掌握程度。

DKT DKVMN SAKT

平均题目预测

对某个知识概念下所有题目的预测作答结果的平均值作为学生在某个知识概念上的掌握程度。

AKT LPKT



实验分析



➤ 学生知识画像预测

模型方法	数据集								
	Assist			JunYi			MATH		
	ACC	AUC	RMSE	ACC	AUC	RMSE	ACC	AUC	RMSE
DKT	0.628 2±0.002	0.701 2±0.009	0.662 7±0.008	0.623 2±0.007	0.685 8±0.009	0.613 7±0.005	0.569 1±0.007	0.701 1±0.005	0.656 3±0.005
DKVMN	0.639 5±0.005	0.663 0±0.006	0.578 5±0.003	0.630 9±0.004	0.660 4±0.006	0.523 3±0.007	0.615 6±0.006	0.649 3±0.004	0.552 0±0.003
SAKT	0.614 6±0.009	0.708 9±0.007	0.620 7±0.005	0.633 9±0.0002	0.699 7±0.0002	0.605 0±0.0003	0.593 9±0.005	0.698 9±0.002	0.640 3±0.005
AKT	0.637 0±0.008	0.664 4±0.006	0.600 2±0.002	0.626 7±0.006	0.682 6±0.003	0.559 4±0.005	0.662 7±0.001	0.718 5±0.003	0.532 5±0.001
LPKT	0.668 1±0.003	0.680 3±0.003	0.554 1±0.004	0.636 1±0.001	0.703 1±0.005	0.549 3±0.003	0.645 5±0.002	0.689 3±0.001	0.525 8±0.000 2
DKP	0.695 5±0.000 2	0.720 6±0.000 2	0.439 8±0.000 1	0.673 4±0.000 7	0.704 9±0.007	0.451 2±0.000 3	0.685 9±0.003	0.747 8±0.002	0.448 9±0.002



方法	DKT	DKVMN	SAKT	AKT	LPKT	DKP
推理速度 (ms/概念)	1.37	5.67	3.58	48.61	103.97	3.10



在保障一定的推理速度的基础上，实现了可信的学生知识画像预测。



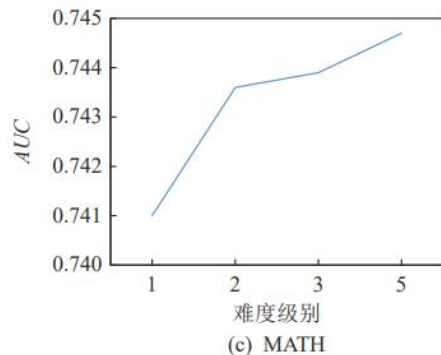
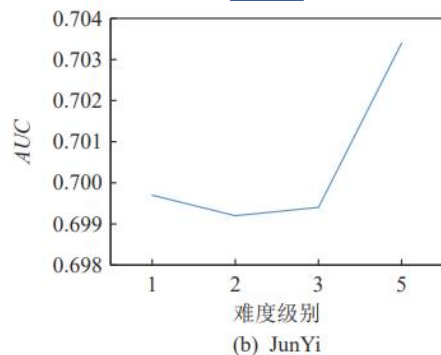
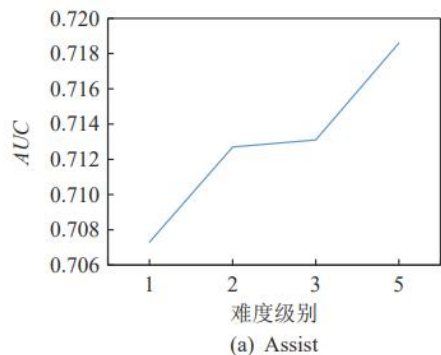
实验分析



➤ 消融实验

去除部分模块得到简化模型：完整方法效果最好，特别是注意力池化增强了知识粒度上的一致性。

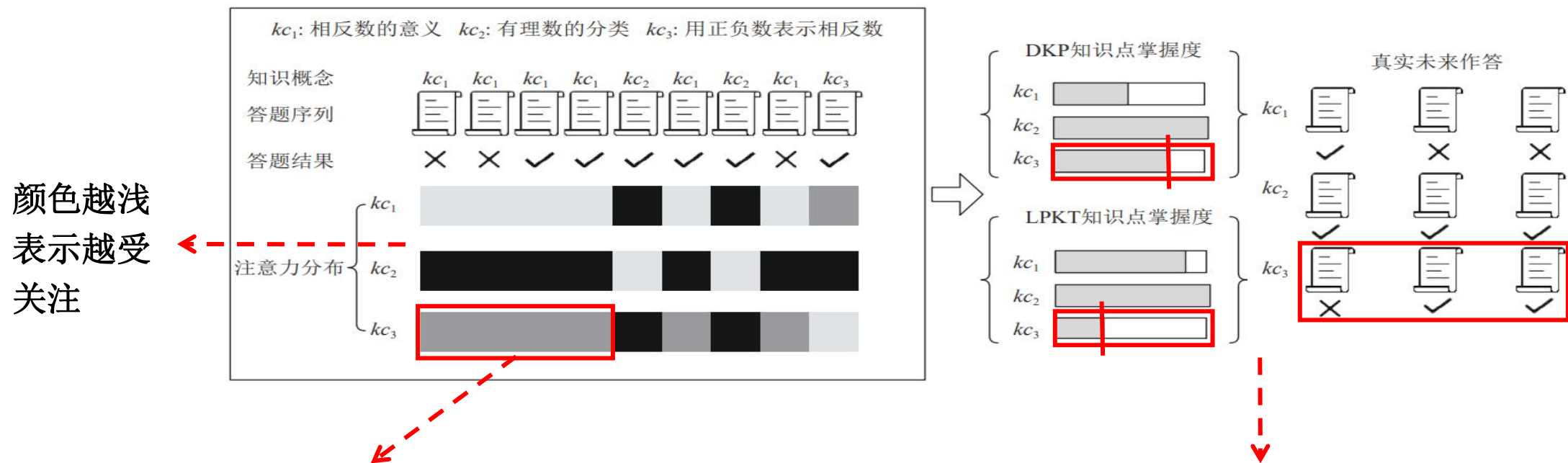
设置不同难度级别：较精细的难度划分通常能获得更优秀的结果，说明难度特征能够提供有效画像信息。



数据集	消融版本	ACC	AUC	RMSE
Assist	完整的 DKP 方法	0.695 5	0.720 6	0.439 8
	w/o 知识聚合	0.691 8	0.718 6	0.440 9
	w/o 难度表征	0.690 7	0.707 3	0.441 7
	w/o Bi-LSTM	0.690 8	0.710 8	0.441 2
	w/o 注意力	0.684 0	0.698 7	0.446 4
JunYi	完整的 DKP 方法	0.673 4	0.704 9	0.451 2
	w/o 知识聚合	0.671 4	0.703 4	0.456 4
	w/o 难度表征	0.669 9	0.699 7	0.457 4
	w/o Bi-LSTM	0.670 5	0.702 5	0.456 7
	w/o 注意力	0.665 9	0.695 7	0.459 2
MATH	完整的 DKP 方法	0.685 9	0.747 8	0.448 9
	w/o 知识聚合	0.682 8	0.744 7	0.449 8
	w/o 难度表征	0.691 9	0.741 0	0.451 5
	w/o Bi-LSTM	0.681 8	0.740 4	0.451 7
	w/o 注意力	0.674 9	0.731 1	0.455 3



➤ 学生知识画像过程可视化



对“用正负数表示相反数”进行预测时，相关的知识概念，即“相反数的意义”也获得了一定的关注

LPKT做出的估计与学生未来总体表现并不完全一致
DKP能准确的基于学生历史作答输出可信的知识画像



➤ 学生知识画像可信性分析

学生知识画像的分布与学生未来表现的分布是否一致

数据集	预测类别	TOP3	TOP5	TOP10
Assist	正例	0.844 4	0.802 3	0.824 6
	负例	0.644 3	0.642 1	0.674 7
JunYi	正例	0.878 8	0.860 5	0.866 2
	负例	0.698 4	0.624 4	0.648 3
MATH	正例	0.832 2	0.807 8	0.833 5
	负例	0.622 1	0.588 1	0.629 0

- DKP预测为正例学生的平均作答得分要明显高于负例学生的平均作答得分，即DKP对学生在正例产生了较高的估计，对应了真实记录中学生在该知识概念上短期内较好的作答情况。
- 进一步地，本文提出的方法已经在某智能教育平台上进行了开放场景的应用. 在初中数学主流用户版本上对46066个用户进行在线测试，最终学生知识画像准确率达到70%。同时，通过发放问卷调研用户对画像的满意度，回收20000份问卷得到75%的满意度。



1 研究背景

2 问题定义

3 模型方法

4 实验分析

5 结论展望



结论展望



- 学生知识画像是学生在不同知识概念上的掌握程度的全面精准表示。
- 现有大多数知识追踪类方法存在不一致的问题，导致画像结果不可信。
- 我们提出一种可信的端到端深度学生知识画像建模方法 **DKP**。
 - 知识粒度交互表征模块
 - 知识状态序列建模模块
 - 知识画像预测模块
- 大量实验证明 **DKP** 方法在学生知识画像任务上的优越性。
- 未来继续探究智能教育场景中学生知识画像预测任务的独特问题。
 - 结合知识图谱，解决答题记录更稀疏场景下的冷启动问题。
 - 尝试将题目文本等丰富的信息融入到表征建模当中。



Thank you



Q&A



sosweetzhang



hxwjz@mail.ustc.edu.cn



<http://cogskl.iflytek.com/>



bigdata.ustc.edu.cn

