

教育大数据中认知跟踪模型研究进展

胡学钢 刘菲 卜晨阳

(大数据知识工程教育部重点实验室(合肥工业大学) 合肥 230601)

(合肥工业大学计算机与信息学院 合肥 230601)

(合肥工业大学大知识科学研究院 合肥 230601)

(jsjxhuxg@hfut.edu.cn)

Research Advances on Knowledge Tracing Models in Educational Big Data

Hu Xuegang, Liu Fei, and Bu Chenyang

(Key Laboratory of Knowledge Engineering with Big Data (Hefei University of Technology), Ministry of Education, Hefei 230601)

(School of Computer Science and Information Engineering, Hefei University of Technology, Hefei 230601)

(Research Institute of Big Knowledge, Hefei University of Technology, Hefei 230601)

Abstract With the in-depth advancement of informational education and the rapid development of online education, a large amount of fragmented educational data are generated during the learning process of students. How to mine and analyze these educational big data has become an urgent problem in the education and the knowledge engineering with big data fields. As for the dynamic education data, knowledge tracing models trace the cognitive status of students over time by analyzing the students' exercising data generated in the learning process, so as to predict the exercising performance of students in the future time. In this paper, knowledge tracing models in educational big data are reviewed, analyzed, and discussed. Firstly, knowledge tracing models are introduced in detail from the perspective of their principles, steps, and model variants, including two mainstream knowledge tracing models based on Bayesian methods and deep learning methods. Then, the application scenarios of knowledge tracing models are explained from five aspects: student performance prediction, cognitive state assessment, psychological factor analysis, exercise sequence, and programming practice. The strengths and weaknesses in Bayesian knowledge tracing models and Deep Knowledge Tracing models are discussed through the two classic models BKT and DKT. Finally, some future directions of knowledge tracing models are given.

Key words educational big data; knowledge tracing; student model; Bayesian knowledge tracing; deep learning

摘要 教育信息化的不断推进和在线教育的蓬勃发展产生了海量的教育数据,如何挖掘和分析教育大数据成为了教育领域和大数据知识工程领域亟待解决的问题。认知跟踪模型通过获取学生作答习题的得分表现,追踪学生随时间变化的认知状态,从而预测学生在未来时间的作答表现。对教育大数据中

收稿日期:2019-10-31;修回日期:2020-06-15

基金项目:国家重点研发计划项目(2016YFB1000901);国家自然科学基金项目(61806065);中央高校基本科研业务费专项资金项目(JZ2020HGQA0186)

This work was supported by the National Key Research and Development Program of China (2016YFB1000901), the National Natural Science Foundation of China (61806065), and the Fundamental Research Funds for the Central Universities (JZ2020HGQA0186).

通信作者:卜晨阳(chenyangbu@hfut.edu.cn)

认知跟踪模型进行了回顾、分析和展望.首先从模型的原理、步骤和方法等维度详细介绍了认知跟踪模型,包括基于贝叶斯方法和深度学习方法 2 类认知跟踪模型.同时,从学生作答表现预测、认知状态评估、心理因素分析、习题序列分析和编程练习 5 个方面阐述认知跟踪模型的应用情景.最后,以经典的贝叶斯认知跟踪模型和深度认知跟踪模型为例分析了 2 类模型的优缺点,并探讨和展望认知跟踪模型未来可能的研究方向.

关键词 教育大数据;认知跟踪;学生模型;贝叶斯认知跟踪;深度学习

中图法分类号 TP391

教育信息化的不断发展产生了海量的教育数据,如何对教育大数据进行挖掘和分析成为了教育领域^[1]和大数据知识工程领域^[2-5]亟待解决的问题.由于学生在学习过程中的行为表现、认知状态和心理因素是评估其学习成效的关键因素^[6],且这些因素随着时间的推移不断变化,因此构建面向动态学习数据的学生模型具有重要的研究意义.认知跟踪(knowledge tracing, KT)模型^[7]旨在面向学生不同时刻作答习题的得分表现数据,跟踪学生的认知状态并预测学生在未来时刻的作答表现,具有重要的研究价值和广泛的应用前景^[8-10].

如图 1 所示,认知跟踪模型属于面向动态数据的学生模型.在认知跟踪模型中,学生对习题的作答表现为可观测的变量,学生的认知状态为不可观测的隐变量^[11].1994 年 Corbett 等人^[7]基于包含观测

变量和隐变量的隐 Markov 模型(hidden Markov model, HMM)构建了贝叶斯认知跟踪(Bayesian knowledge tracing, BKT)模型.随后,学者们分别从变量表示方式、数据特征信息、建模方法等方面展开了对认知跟踪模型的研究.针对变量表示方式,部分认知跟踪模型(包括 BKT 模型)采用“0-1”二元表示方法,分别表示隐变量中学生的“未掌握-掌握”认知状态及观测变量中学生的“错误-正确”作答表现^[12-14];除二元表示方法外,部分认知跟踪模型采用多元表示方法增强模型的表示能力^[15-18].针对数据特征信息,部分认知跟踪模型通过结合数据中的习题^[10-11,19-20]、知识点^[21-23]、时间^[13-14,24]等信息提升模型拟合的精度.针对建模方法,现有认知跟踪模型中 2 类主流的建模方法为贝叶斯方法^[18,25-30]和深度学习方法^[10,31-38].

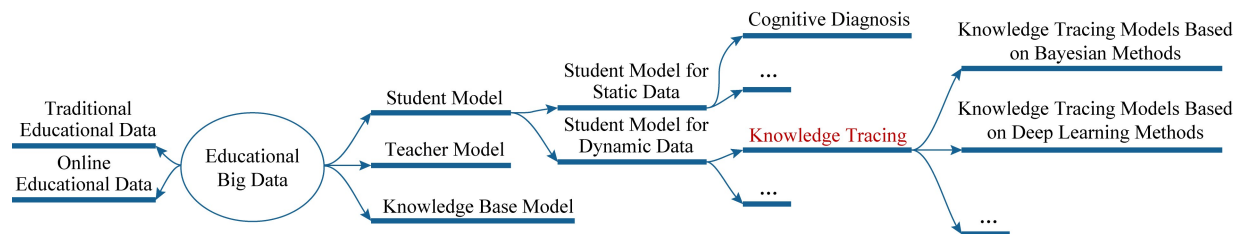


Fig. 1 Mind map of knowledge tracing models in educational big data

图 1 教育大数据中认知跟踪模型的思维导图

本文从建模方法的角度回顾了认知跟踪模型的发展,详细介绍了基于贝叶斯方法和深度学习方法 2 类认知跟踪模型.据我们所知,目前关于认知跟踪模型的综述型研究较少.文献[39]主要分析了贝叶斯方法的认知跟踪模型,未分析其他建模方法的认知跟踪模型.文献[40-43]是学生模型的综述,其中认知跟踪模型作为综述中的一个部分,并未详细展开介绍.文献[44-47]在性能、精度等方面比较了认知跟踪模型与其他模型.文献[39-47]未对认知跟踪模型进行系统、全面地梳理.本文通过检索计算机领域及教育领域的期刊和会议,从建模方法的角度详

细分析了教育大数据中认知跟踪模型的发展趋势,同时介绍了模型的应用情况,对不同类型的认知跟踪模型进行比较、分析和展望.本文的具体贡献有 3 个方面:

- 1) 对教育大数据中的认知跟踪模型进行全面梳理,分别从模型原理、算法步骤、模型的扩展和改进等方面详细介绍认知跟踪模型;
- 2) 分别对基于贝叶斯和深度学习方法的 2 类认知跟踪模型进行介绍,将 2 类模型按照数据特征信息、方法策略等方面归类整理;
- 3) 对比了现有认知跟踪模型的优缺点,并从模

型的参数寻优、可解释性与精度的权衡、部分假设与实际情况不符等角度分析了未来认知跟踪模型有待研究的方向。

1 相关工作

本节首先介绍教育数据挖掘的相关工作,随后分别对面向静态学习数据的认知诊断模型及面向动态学习数据的认知跟踪模型进行概述,最后分析本文与其他相关综述文章的区别。

1.1 教育数据挖掘

教育数据挖掘^[48]通过数据挖掘技术处理教育大数据,从而获取教育系统中有价值的信息。图2是一个教育系统示意图,模型中有3个主要角色,分别是知识库(knowledge base)、学生(students)和教师(teachers)。知识库是教育系统的基础;学生与教师

之间通过评估系统(assessment system)或习题(items)交互,具体为:

1) 知识库.在教育数据挖掘的相关研究中,通常将知识库视为知识点的集合,知识点之间可能存在树状或网状关系。知识库的自动构建是知识库模型^[49]的重要研究方向。

2) 学生.基于学生的学习过程可评估认知状态、分析心理因素、预测行为表现。学生模型的研究包括针对单个学生的认知状态提出的个性化学习方案和基于学生群体的相似性和差异性提出的协同学习方案等。

3) 教师.教师可通过追踪学生对习题的作答表现掌握学生的学习状态;通过评估系统的反馈分析学生的认知情况、交互心理等。此处的“教师”为广义上的“教师”,即按照课堂教育和在线教育2种教育方式,教师可被划分为教职人员和在线平台2类。

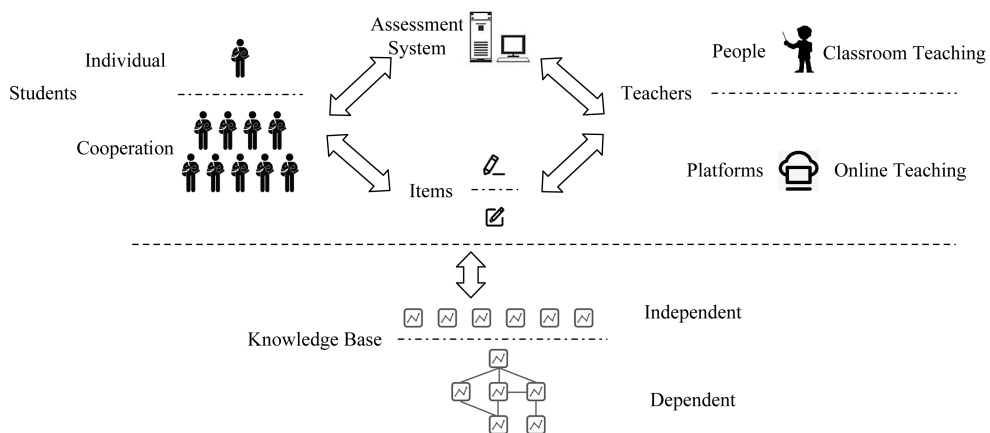


Fig. 2 Toy model of educational system

图2 教育系统示意图

根据研究对象的不同,教育数据挖掘的相关研究包括面向学生、面向教师以及面向知识库的相关模型和方法:1)面向学生的模型和方法:包括评估或跟踪学生认知状态的模型(如认知诊断模型^[50]、认知跟踪模型^[7,31]等)及学习资源的个性化推荐^[51-54]、协同学习的推荐算法^[55-57]等;2)面向教师的模型和方法:包括试卷的自动评估和评分等,如计算机自适应测验(CAT)可针对学生的测验情况精准定位测验内容、评估试卷的合理性等^[58-59];3)面向知识库的模型和方法^[49]。

1.2 认知诊断模型概述

由图1可知,学生模型包括面向静态学习数据和动态数据的模型。与面向动态学习数据的认知跟踪模型不同,经典的认知诊断模型旨在对静态学习数据

进行建模分析。认知诊断模型^[50](cognitive diagnosis model, CDM)基于教育学、统计学和计算机科学,通过对习题与知识点的关联关系及学生的答题情况,得到学生的知识点掌握情况。目前已有上百种认知诊断模型,包括规则空间模型^[60](rule space model, RSM)、属性层次模型^[61](attribute hierarchy model, AHM)、DINA^[62](deterministic input, noisy AND-gate)模型等。

DINA模型是CDM中应用十分广泛的模型。该模型假设每个习题都关联1个或多个知识点,通过分析学生的作答表现,并考虑“猜测”和“失误”等心理因素对学生进行诊断分析。其中,习题与知识点的关联关系通常使用Q矩阵表示。与认知跟踪模型不同的是,认知诊断模型中的学生作答数据不包含学生

作答习题的时间特性.然而,时间特性在分析学生的认知状态中具有重要的意义.因此,作为对动态学习数据建模的模型,认知跟踪模型在教育大数据领域值得深入研究.

1.3 认知跟踪模型概述

认知跟踪模型面向动态学习数据建模,基于学生在学习过程中对习题作答表现的数据,跟踪学生随时间的变化对知识点的掌握情况^[7].如图3所示,

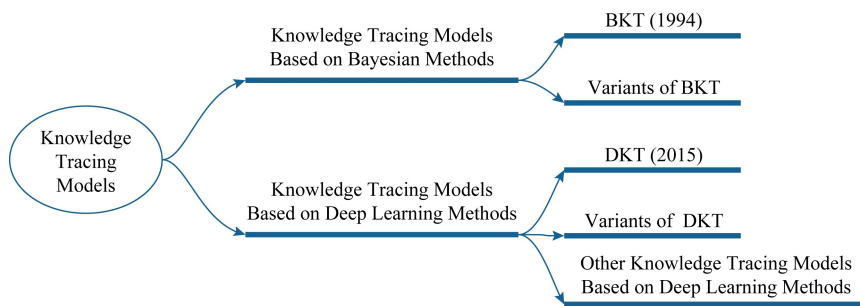


Fig. 3 Classification of knowledge tracing models

图3 认知跟踪模型的分类

1.3.1 问题背景

认知跟踪旨在根据学生在不同时刻的习题作答数据,分析学生认知状态的变化、并预测他们在未来

认知跟踪模型从建模方法上可分为基于贝叶斯方法的认知跟踪模型(详见第2节)和基于深度学习的认知跟踪模型(详见第3节).

前者出现最早的模型是 Corbett 等人^[7]提出的 BKT 模型,后者是以 2015 年 Piech 等人^[31]提出的 DKT(deep knowledge tracing)模型为代表.BKT 和 DKT 模型提出之后,学者们基于 BKT 和 DKT 提出了很多的扩展模型.

时刻的作答表现.假设学生在每个时间步作答一道习题,该模型的形式化定义中涉及到的符号如表1所示:

Table 1 Notations and Descriptions in Knowledge Tracing

表1 认知跟踪模型中涉及的符号及描述

Notation	Range of Parameters	Definition
S_i	$i \in \{1, 2, \dots, I\}$	The i -th student, where the number of students is I .
t	$t \in \{1, 2, \dots, T\}$	The t -th time step, where the number of time steps is T .
KC_q	$q \in \{1, 2, \dots, Q\}$	The q -th knowledge component, where the number of knowledge components is Q .
E_i^t	$i \in \{1, 2, \dots, I\}, t \in \{1, 2, \dots, T\}$	The item of student S_i at time step t .
y_i^t	$i \in \{1, 2, \dots, I\}, t \in \{1, 2, \dots, T\}$	The observation item score of student S_i at time step t .
y_i^{T+1}	$i \in \{1, 2, \dots, I\}$	The prediction item score of student S_i at time step $T+1$.
$k_{q,i}^t$	$t \in \{1, 2, \dots, T\}, i \in \{1, 2, \dots, I\},$ $q \in \{1, 2, \dots, Q\}$	The cognitive state for knowledge components q of student S_i at time step t .
K_i^t	$K_i^t \subseteq \{KC_q\}, q \in \{1, 2, \dots, Q\}$	The set of knowledge components related to the item E_i^t .
$k_{q,t}$	$q \in \{1, 2, \dots, Q\}, t \in \{1, 2, \dots, T\}$	Latent variables of knowledge component q at time step t .
$y_{q,t}$	$q \in \{1, 2, \dots, Q\}, t \in \{1, 2, \dots, T\}$	Observation variables of knowledge components q at time step t .
$P(L_0), P(L_t)$		Probability of learned state at time step 0 and t , respectively.
$P(T), P(G), P(S)$		Transition, guessing, and slipping probabilities, respectively.
i		An item.
j		A student.
$d_{q,i}$	$q \in \{1, 2, \dots, Q\}, i \in \{1, 2, \dots, I\}$	Difficulty of item i related to knowledge component q .
$\theta_{q,j}$	$q \in \{1, 2, \dots, Q\}, j \in \{1, 2, \dots, J\}$	Ability of student j related to knowledge component q .
$a_{q,t}$	$q \in \{1, 2, \dots, Q\}, t \in \{1, 2, \dots, T\}$	Affect component of knowledge component q at time step t .
$b_{q,t}$	$q \in \{1, 2, \dots, Q\}, t \in \{1, 2, \dots, T\}$	Engagement component of knowledge components q at time step t .
α, δ		Parameters to personalize the guessing and slipping probabilities.
$\sigma_a^2, \sigma_\delta^2$		Variances drawn from an inverse-Gamma-distributed conjugate prior.
y_t, k_t, o_t	$t \in \{1, 2, \dots, T\}$	Input, hidden, and output vectors at time step t , respectively.

已知 I 个学生的集合 $\{S_i\}, i \in \{1, 2, \dots, I\}$; Q 个知识点的集合 $\{KC_q\}, q \in \{1, 2, \dots, Q\}$; I 个学生在时间步 t 作答习题 E_i^t 的得分为集合 $\{y_i^t\}, i \in \{1, 2, \dots, I\}, t \in \{1, 2, \dots, T\}$. 学生 S_i 在时间步 t 作答的习题 E_i^t 涉及到知识点集合 $K_i^t \subseteq \{KC_q\}, q \in \{1, 2, \dots, Q\}$, 其中 K_i^t 可能未知. 认知跟踪问题中 2 个常见的求解目标分别是:

1) 求解学生在不同时刻的认知状态 $\{k_{q,i}^t\}, i \in \{1, 2, \dots, I\}, t \in \{1, 2, \dots, T\}, q \in \{1, 2, \dots, Q\}$, 其目标函数为

$$\arg \min_{\bar{y}^{T+1}} \left(\sum_{i=1}^I g(\bar{y}_i^{T+1}, y_i^{T+1}) \right), \quad (1)$$

其中, $\bar{y}^{T+1} = \{\bar{y}_i^{T+1}\}, i \in \{1, 2, \dots, I\}$; $g(\bar{y}_i^{T+1}, y_i^{T+1})$ 为预测值 \bar{y}_i^{T+1} 与真实值 y_i^{T+1} 的误差.

2) 预测学生在时刻 $(T+1)$ 的作答表现 $\bar{y}_i^{T+1}, i \in \{1, 2, \dots, I\}$, 其目标函数为

$$\arg \max_k \left(\sum_{i=1}^I \sum_{q=1}^Q f(k_{q,i}^t; y_i^t, K_i^t) \right), \quad (2)$$

其中, $k = \{k_{q,i}^t\}, t \in \{1, 2, \dots, T\}, q \in \{1, 2, \dots, Q\}, i \in \{1, 2, \dots, I\}$; K_i^t 表示学生 S_i 在时间步 t 作答的习题 E_i^t 涉及到知识点集合, $f(k_{q,i}^t; y_i^t, K_i^t)$ 表示学生 S_i 在认知状态 $k_{q,i}^t$ 条件下作答习题 E_i^t 的得分 y_i^t 的概率.

认知跟踪框架示意图如图 4 所示: 1) 模型的输入为 I 个学生在 T 个时间步作答不同知识点习题的观测得分. 任意习题涉及到 1 个或多个知识点, 习题涉及到知识点的不同情况如图 4 中圆角框所示. 即该习题可能只涉及到单个知识点, 也可能涉及到多个知识点; 用户标注的多个知识点可能相互独立、也可能是相关的; 习题涉及到的知识点可能是已知的, 也可能是未知的. 2) 针对观测得分数据, 通过贝叶斯网络、深度学习等模型进行拟合、参数寻优, 可得到每个学生的认知跟踪子模型. 3) 认知跟踪模型可应用于对学生在未来时刻的作答情况进行预测, 以及在已知习题与知识点的关联关系的情况下, 分析学生在每个时间步下对每个知识点的掌握状态.

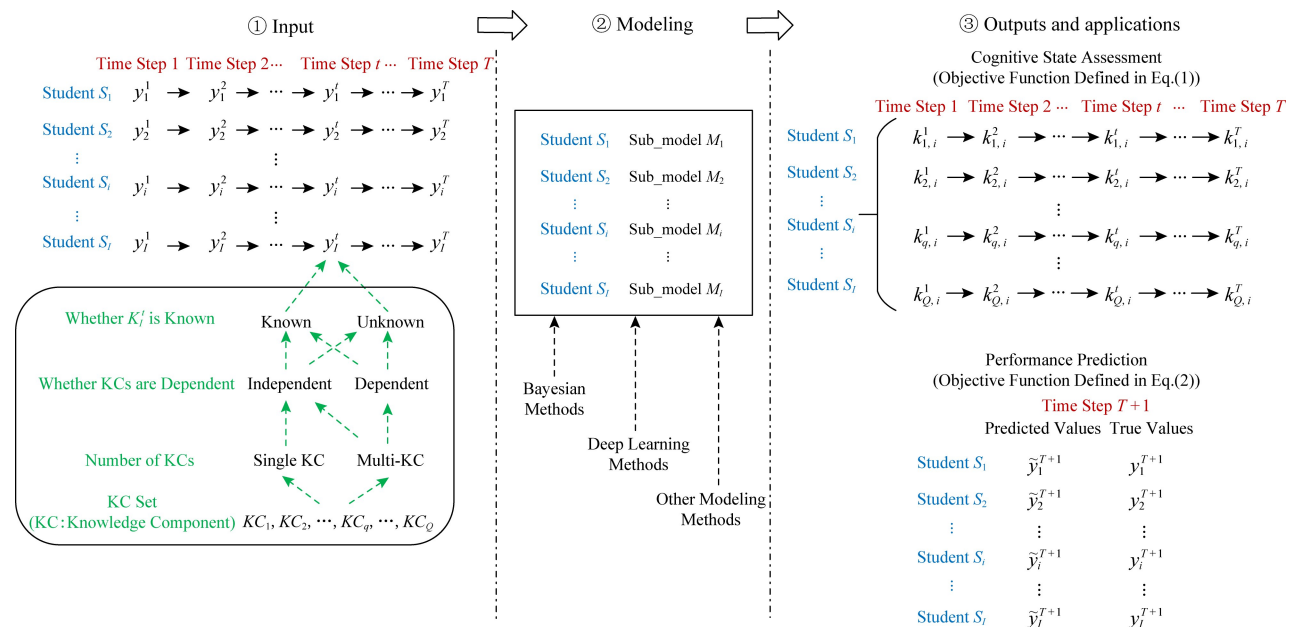


Fig. 4 Framework of knowledge tracing (assuming that a student answers only one item at a time step)

图 4 认知跟踪框架示意图 (假设学生在每个时间步作答 1 道习题)

1.3.2 模型的评价指标

在认知跟踪模型中常见的评价指标为: 对数似然 (log-likelihood, LL)、均方根误差 (root mean squared error, RMSE)、平均绝对误差 (mean absolute error, MAE)、ROC 曲线下方的面积大小 (area under the curve, AUC)^[63-64]. 不同的应用场景下可选择不同的评价指标, 如经典的参数估计方法 EM 算法

(expectation maximization)^[65] 通常使用 LL 拟合参数. LL^[66], RMSE^[67], MAE^[68] 的计算分别为

$$\text{metrics}_{\text{LL}} = \sum_{i=1}^n [c_i \lg(p_i) + (1 - c_i) \lg(1 - p_i)], \quad (3)$$

$$\text{metrics}_{\text{RMSE}} = \sqrt{\frac{1}{n} \sum_{i=1}^n (c_i - p_i)^2}, \quad (4)$$

$$metrics_{MAE} = \frac{1}{n} \sum_{i=1}^n |c_i - p_i|, \quad (5)$$

其中, n 表示预测数据的个数, c_i 表示第 i 个预测数据, p_i 表示第 i 个真实数据. 对应至认知跟踪模型, n 表示学生作答的习题个数, c_i 和 p_i 分别表示学生作答第 i 道习题得分的预测值和真实值.

在认知跟踪模型的评价指标研究中, 文献[63]对比了 LL, RMSE, AUC 这 3 个指标在认知跟踪的参数估计中的效果, 发现 RMSE 比其余 2 个评价指标对参数估计具有更强的指引效果, 使模型预测达

到最佳性能. 文献[64]通过对比分析得到 RMSE 和 LL 更加适合参数估计的场景, 而 AUC 在评价模型最终结果方面效果更佳, 因此大多数研究以 AUC 作为模型的最终评价指标.

1.3.3 公开数据集

现有认知跟踪模型的实验拟合部分的数据多数采用学生与在线平台交互的习题作答数据. 为方便研究者对认知跟踪模型进行更好的研究和发展, 本文对目前网络上公开的数据集进行了不完全归纳统计, 如表 2 所示:

Table 2 List of Public Datasets

表 2 公开数据集列表

Dataset	Website	Author
ASSISTmentsData	https://sites.google.com/site/assistmentsdata/home	Heffernan Worcester Polytechnic Institute
Datashop	https://pslcdatashop.web.cmu.edu/	Carnegie Mellon University
Anonymizeddata	https://code.org/research	Code.org
Synthetic	https://github.com/chrispiech/DeepKnowledgeTracing/tree/master/data/synthetic	Piech et al.

ASSISTmentsData 是 Heffernan 教授团队公开发布的数据集, 该数据集是由教辅系统获得的小学数学习题的学生答题记录, 是目前公开的认知跟踪领域最大的数据集^[31]. ASSISTmentsData 提供了 3 组数据: 1) 2009—2010 ASSISTment Data 数据集包括 Skill builder 数据(401 757 条学生作答记录)和 Non skill builder 数据(603 129 条学生作答记录) 2 个子数据集. 其中, Skill builder 的含义为: 若某学生作答数据满足某一条件(如连续正确作答 3 次与该知识点相关的习题), 则视为该学生已掌握该知识点, 从而不会再给该学生推荐该知识点的相关习题. 2) 2012—2013 School Data with Affect 数据集含 6 123 271 条学生作答记录. 其中, Affect^[69] 的含义为: 数据中包括与学生心理因素相关的 4 个特征 *Average_confidence* (FRUSTRATED), *Average_confidence* (CONFUSED), *Average_confidence* (CONCENTRATING), *Average_confidence* (BORED). 3) 2015 ASSISTments Skill Builder Data 数据集含 708 632 条学生作答记录. 该数据集为 2015 年收集的 Skill builder 数据.

Datashop 是最大的学习交互数据存储库, 供教育数据挖掘者、课程开发人员、教育技术研究者、心理学家等各界人士上传和使用. Datashop 中包括数学、化学、物理、英语、汉语等多门课程的学习交互数据, 时间跨度为 1969—2018 年且持续更新中, 每个

数据集的数据特征也不尽相同. 因此, Datashop 对于包括认知跟踪领域在内的教育大数据研究工作具有十分重要的支撑作用.

Anonymizeddata 公开了 2 次计算机编程挑战 (Hoc4 和 Hoc18) 的用户交互数据集, 时间跨度为 2013-12—2014-03. 其中, Hoc4 和 Hoc18 分别包含 509 405 个学生、1 138 506 次提交记录; 263 569 个学生、1 263 360 次提交记录.

Synthetic 是深度认知跟踪模型 DKT 的作者 Piech 等人^[31] 构建的模拟数据, 包含 4 000 个学生、50 个习题、20 万条学生答题记录.

本节介绍的 4 类认知跟踪领域中常用的公开数据集, 不同数据集的规模、数据特征有所不同, 研究者可以根据研究内容选择不同的数据集.

1.4 本文与现有综述的区别

本节将介绍现有认知跟踪相关综述, 并分析与本文的区别.

认知跟踪模型属于学生模型, 因此本节首先列出学生模型的相关综述. 此类综述以学生模型为对象, 认知跟踪模型作为子类模型, 未对其进行系统地介绍和分析. 文献[40-43]是关于学生模型的综述. 文献[40]围绕智能辅导系统 (intelligent tutoring system, ITS) 的学生模型展开. 文献[41]从理论技术的角度梳理学生模型, 包括叠加模型 (overlay model)、机器学习 (machine learning)、认知理论

(cognitive theories)、模糊学生模型(fuzzy student modeling)、贝叶斯网络(Bayesian network)等.文献[42]详细阐述了在特定的情境下如何选择学生模型的问题,对开发者和研究者提出了可能的研究方向.文献[43]从知识状态、认知行为、情感因素和综合类型4个方面对学生模型进行了梳理.这些文献涉及到认知跟踪模型的内容较少,而本文着重从原理、方法步骤、比较分析、发展历程、应用等不同的维度详细介绍认知跟踪模型.

据我们所知,现有认知跟踪模型的综述只有文献[39].文献[39]从知识点、学生和数 据 3 个方面分析了认知跟踪模型在教育领域上的应用.但该文献侧重于分析贝叶斯方法的认知跟踪模型,没有讨论其他建模方法的认知跟踪模型.与该文献相比,本文对教育大数据中认知跟踪模型的回顾和分析更加系统和全面,梳理并比较了不同建模方法的认知跟踪模型.

另外,一些非综述类文献中介绍了认知跟踪模型与其他模型的对比分析情况^[44-47].文献[44]将认知跟踪模型与常规的随机对照试验(randomized controlled trial, RCT)、学习分解(learning decomposition)进行比较,分别从模型的目的、输入、假设条件3个方面进行评估.文献[45]将认知跟踪模型与表现因素分析(performance factor analysis, PFA)方法从预测精度和参数的合理性2方面进行比较;同时研究不同的参数拟合方法——最大期望(expectation maximization, EM)方法和BF(brute force)方法的认知跟踪模型KT+EM与KT+BF的性能,结果发现KT+EM的性能优于KT+BF.文献[46]从学生作答表现的预测方面比较了认知跟踪模型与PFA方法.文献[47]将认知跟踪模型与智能辅导系统中经常使用的连续正确响应进行了实验对比.文献[44-47]通过理论证明或实验分析的方式将认知跟踪模型与其他相关的学生模型进行比较,侧重点是分析了认知跟踪模型与其他特定模型在预测性能、参数拟合等方面的表现.而本文则系统地介绍了认知跟踪模型的发展历程和模型的应用,比较了不同建模方法模型的优缺点,并对未来的研究方向进行了展望.

2 基于贝叶斯方法的认知跟踪模型

本节介绍基于贝叶斯方法的认知跟踪模型,首

先解释了经典贝叶斯认知跟踪模型BKT的原理和步骤,对BKT模型的相关文献进行梳理;然后介绍了其他基于BKT扩展的贝叶斯认知跟踪模型、开发工具和公开源码.

2.1 BKT模型

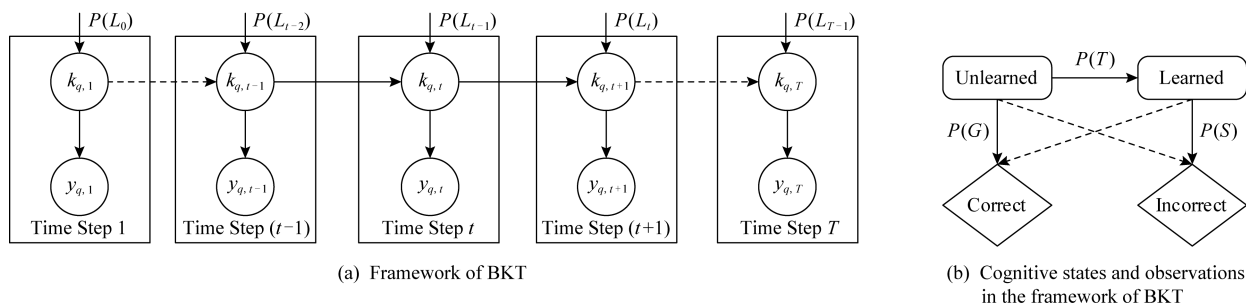
1994年Corbett等人^[7]提出BKT模型,在学生使用智能辅导系统过程中(此处指学生使用ACT Programming Tutor短程序编写系统进行Lisp语言编程练习),基于HMM^[70]构建学生认知状态变化的模型,监测和估计学生掌握知识点的概率.

HMM描述离散时间序列上的概率分布 $Y = (y_1, y_2, \dots, y_t, \dots, y_T)$,它取决于隐状态序列 $X = (x_1, x_2, \dots, x_t, \dots, x_T)$.其中, Y 称为观测序列(observation), X 称为隐状态(hidden state), Y 与 X 均是离散序列. t 表示时间步.即时间步 t 的观测变量 y_t 只与 x_t 有关,称 x_t 至 y_t 的概率为发射概率(emission probability),记为 $P(y_t | x_t)$.从时间步 $1 \sim T$,隐状态序列从 x_1 逐步转移成 x_T ,任意一个时间步的隐状态 x_t 至下一个时间步的隐状态 x_{t+1} 称为转移概率(transition probability),记为 $P(x_{t+1} | x_t)$.

如图5(a)所示,BKT模型将学生作答表现序列定义为HMM模型中的观测序列(对应图5(a)中的阴影圆圈);学生对知识点的认知状态定义为HMM模型中的隐状态(对应图5(a)中的非阴影圆圈). $P(L_t)$ 表示时间步 t 时学生掌握知识点 q 的概率, $P(L_0)$ 表示初始时间步时学生掌握知识点 q 的概率.如图5(b)所示,作答表现存在“正确”(correct)和“错误”(incorrect)2种取值(对应图5(b)中的2个菱形);认知状态的取值存在“已掌握状态”(learned state)和“未掌握状态”(unlearned state)2种取值(对应图5(b)中的2个切角矩形). $P(G)$ 是猜测概率,意为在学生的认知状态处于“未掌握状态”时,学生的作答表现为“正确”的概率; $P(S)$ 为失误概率,意为在学生的认知状态处于“掌握状态”时,学生的作答表现为“不正确”的概率; $P(T)$ 为转移概率,指学生的认知状态从“未掌握状态”变成“掌握状态”的概率.需要注意的是,BKT基于3点假设进行建模:

1) 每一次练习仅与一个知识相关;

2) 忽略学生的认知遗忘因素,即学生对知识的认知状态可以由“未掌握状态”转移至“掌握状态”,但反之的概率为0;

Fig. 5 Model of BKT^[17]图5 BKT模型^[17]

3) 学生的转移概率 $P(T)$ 不变^[15], 与学生作答表现、作答正确次数无关。

BKT 的具体步骤如下^[7]:

Step1. 将一个学生针对 1 个知识点的作答数据建模为如图 5 所示的 BKT 模型。

Step2. 初始化 $P(L_0)$, $P(T)$, $P(G)$, $P(S)$ 这 4 个参数的取值, 时间步 $t=1$ 。

Step3. 根据时间步 t 的观测值 $y_{q,t}$ 计算时间步 $t-1$ 的知识点掌握条件概率 $P(L_{t-1} | y_{q,t})$, 若 $y_{q,t} = \text{Correct}$, 则

$$P(L_{t-1} | y_{q,t} = \text{Correct}) = \frac{P(L_{t-1})(1 - P(S))}{P(L_{t-1})(1 - P(S)) + (1 - P(L_{t-1}))P(G)}. \quad (6)$$

若 $y_{q,t} = \text{Incorrect}$, 则

$$P(L_{t-1} | y_{q,t} = \text{Incorrect}) = \frac{P(L_{t-1})P(S)}{P(L_{t-1})P(S) + (1 - P(L_{t-1}))(1 - P(G))}. \quad (7)$$

Step4. 使用时间步 $t-1$ 的条件概率 $P(L_{t-1} | y_{q,t})$ 更新时间步 t 学生掌握该知识点的概率 $P(L_t)$ 。

$$P(L_t) = P(L_{t-1} | y_{q,t}) + (1 - P(L_{t-1} | y_{q,t}))P(T). \quad (8)$$

Step5. 更新 4 个参数的值, $t=t+1$ 。

Step6. 返回 Step3, 直至时间步 $t=T$, 得到 4 个参数的解。

Step7. 返回 Step2, 直至算法终止条件, 得到 BKT 模型。

在得到 4 个参数的最优解后, 即可评估学生的认知状态, 当学生作答正确时, 学生的认知状态为

$$P(L_t | y_{q,t} = \text{Correct}) = P(L_t)(1 - P(S)) + (1 - P(L_t)) \times P(G). \quad (9)$$

当学生作答错误时, 学生的认知状态为

$$P(L_t | y_{q,t} = \text{Incorrect}) = P(L_t)P(S) + (1 - P(L_t)) \times (1 - P(G)). \quad (10)$$

除了评估学生的认知状态, 还可以预测学生在下一时间步的作答表现, 即计算学生在下一时间步学生正确作答习题的概率:

$$P(y_{q,t+1} = \text{Correct}) = P(L_{q,t+1})(1 - P(S)) + (1 - P(L_{q,t+1}))P(G). \quad (11)$$

在此, 对式(6)~(11)进行详细解释:

由 HMM 性质知, 当前时间步的状态仅与当前时间步的观测结果及上一时间步的状态有关, 因此时间步 t 时学生对知识点 q 的掌握概率 $P(L_t)$ 的更新如式(8), 其中, $y_{q,t}$ 表示当前时间步 t 时的观测结果, $y_{q,t} \in \{\text{Correct}, \text{Incorrect}\}$; $P(L_{t-1} | y_{q,t})$ 表示上一时间步的掌握概率的条件概率, $1 - P(L_{t-1} | y_{q,t})$ 表示在当前时间步 t 的观测结果下上一时间步未掌握的概率; $P(T)$ 表示由未掌握状态转移至掌握状态的概率, 式(8)中的 $P(L_{t-1} | y_{q,t})$, $y_{q,t} \in \{\text{Correct}, \text{Incorrect}\}$ 的计算如式(6)(7)。

式(6)计算在时间步 t 的观测值为作答正确的情况下, 时间步 $t-1$ 的知识点掌握条件概率 $P(L_{t-1} | y_{q,t} = \text{Correct})$ 。分母为时间步 t 观测值为作答正确的所有情况的概率之和: 1) 时间步 $t-1$ 学生掌握知识点且学生对该习题无失误; 2) 时间步 $t-1$ 学生没有掌握知识点且学生对该习题猜测正确。分子为上述 2 种情况中“时间步 $t-1$ 学生掌握知识点”的概率, 因此求得 $P(L_{t-1} | y_{q,t} = \text{Correct})$ 。

式(7)计算在时间步 t 的观测值为作答错误的情况下, 时间步 $t-1$ 的知识点掌握条件概率 $P(L_{t-1} | y_{q,t} = \text{Incorrect})$ 。同式(6), 分母为时间步 t 观测值为作答错误的所有情况的概率之和: 1) 时间步 $t-1$ 学生掌握知识点且学生对该习题有失误; 2) 时间步 $t-1$ 学生没有掌握知识点且学生对该习题无猜测。分子为上述 2 种情况中“时间步 $t-1$ 学生掌握知识点”的概率, 因此求得 $P(L_{t-1} | y_{q,t} = \text{Incorrect})$ 。

式(8)计算时间步 t 学生掌握该知识点的概率

$P(L_t)$ 为下列 2 种情况的概率之和:1)基于时间步 t 的观测值情况下,时间步 $t-1$ 学生掌握知识点;2)基于时间步 t 的观测值情况下,时间步 $t-1$ 学生没有掌握知识点的概率与转移概率的乘积.因此求得 $P(L_t)$.

式(9)在时间步 t 学生作答正确的情况下,评估学生认知状态,即在时间步 t 学生作答正确的情况下,学生掌握该知识点的概率 $P(L_t | y_{q,t} = \text{Correct})$ 为以下 2 种情况的概率之和:1)学生在时间步 t 掌握该知识点且对该习题无失误;2)学生在时间步 t 未掌握该知识点且对该习题猜测正确.因此求得 $P(L_t | y_{q,t} = \text{Correct})$.

式(10)在时间步 t 学生作答错误的情况下,评估学生认知状态,即在时间步 t 学生作答错误的情况下,学生掌握该知识点的概率 $P(L_t | y_{q,t} = \text{Incorrect})$ 为以下 2 种情况的概率之和:1)学生在时间步 t 掌握该知识点且对该习题有失误;2)学生在时间步 t 未掌握该知识点且对该习题无猜测.因此求得 $P(L_t | y_{q,t} = \text{Incorrect})$.

式(11)用以预测时间步 $t+1$ 学生正确作答习题的概率 $P(y_{q,t+1} = \text{Correct})$ 为以下 2 种情况的概率之和:1)在时间步 $t+1$ 学生掌握该知识点且无失误;2)在时间步 $t+1$ 学生未掌握该知识点且猜测正确.因此求得 $P(y_{q,t+1} = \text{Correct})$.

2.2 BKT 模型的分析

BKT 模型在智能教辅系统等场景中发挥着重要的作用.学者们对 BKT 模型的参数、性能、数据等方面进行了具体分析.本节将从 BKT 模型的参数分析、BKT 模型的识别性问题和模型退化问题、BKT 模型的数据分析 3 个方面进行梳理.

2.2.1 BKT 模型的参数分析

在使用 BKT 模型进行认知状态评估时,可能会出现 2 类错误的评估:1)一个学生实际未掌握某知识,其认知状态却被评估为“掌握状态”,这种现象被称为“过度积极”(false positives);2)一个学生实际掌握某知识,其认知状态却被评估为“未掌握状态”,这种现象被称为“过度消极”(false negatives)^[71].为缓解上述错误评估的发生,可设置一个认知阈值,若学生掌握某知识的概率大于该认知阈值,则将学生对该知识的认知状态视为掌握;否则视为未掌握.比如,设定认知阈值为 0.95,若一个学生对一个知识点掌握的概率为 0.97,则判断该学生对该知识点的认知状态为掌握.

在课堂教学或 MOOC 在线学习场景下^[72-73],可能会存在学生对一些知识具有很强理解能力的情况.文献[74]探讨了当学生的初始参数 $P(L_0)$ 很高的极端情况下(如 $P(L_0) = 0.9$)认知跟踪的参数估计性能,实验结果显示这种情况的模型参数估计性能较低.

2.2.2 BKT 模型的识别性问题和模型退化问题

文献[75]指出 BKT 模型是不可识别的,即对于 BKT 的 4 个参数,不同的参数取值的组合可能导致相同的作答预测结果,此问题被称为识别性问题(identifiability problem).通过探索认知跟踪模型中参数对模型预测性能的敏感性问题,确认是否可以缩小参数的搜索空间.文献[76]发现模型对参数是不敏感的,并通过将参数聚类的方法缩小了参数的搜索空间,更快地找到最优解.文献[77]对 BKT 模型中学生的认知概率公式进行推演,证明了存在很多种关于猜测概率 $P(G)$ 和初始参数 $P(L_0)$ 的取值可能,解释了识别性问题出现的原因.此外,文献[78-79]均探讨了 BKT 模型的识别性问题.若识别性问题确实存在,必须对 BKT 模型重新建立新的模型评价标准.

文献[80]表示在符合实际情况的条件下,BKT 模型是可识别的,即认为 BKT 模型不存在识别性问题.该文献提出语义模型退化(semantic model degeneracy)问题,即在 BKT 模型中,求解的模型参数与模型的假设条件不一致.该问题出现在 BKT 模型的拟合过程中.由 2.1 节知,BKT 模型使用 2 状态的 HMM 拟合数据,得到符合该 HMM 的参数最优解;但是对猜测概率、失误概率及转移概率的最优解解释却与最初的 BKT 模型假设不完全一致.文献[80]认为分析模型退化的来源将是未来研究的一个重要方向.无论是识别性问题,还是模型退化问题,其根源可追溯到得到模型参数最优解的过程和方法,从而进行学生认知状态的估计和作答表现的预测^[81].

2.2.3 BKT 模型的数据分析

在 BKT 模型中,参数估计会影响 BKT 的评估性能,但由于数据规模的原因,参数最优解可能仍无法准确地拟合真实数据,这种情况被称为“样本误差”(sampling error)^[82].文献[82]通过实验得出结论:对于规模为 n 的学习数据, $1/\sqrt{n}$ 是最佳的训练样本规模.该文献使用的是模拟数据进行实验,未来可使用真实数据继续进行研究.

模拟数据在教育数据挖掘及认知跟踪领域的使用十分广泛,如使用模拟数据验证 BKT 模型的收敛性问题^[83].文献[84]研究了是否可以区分一个数据集是模拟数据还是真实数据,分析了模拟数据与真实数据之间的相似性,但并未设计出可以自动识别模拟数据或真实数据的算法.

2.3 BKT 的扩展

自 1994 年 BKT 模型提出后,学者们从使用不同方法、以不同目的、适应不同场景等角度对 BKT 模型进行扩展.本文将 BKT 的扩展模型分为 3 类: 1)结合教育数据特征的扩展模型;2)混合方法扩展模型;3)其他模型.扩展模型的对比如表 3 所示:

Table 3 Comparison of the Variants of BKT Model

表 3 BKT 扩展模型的对比

Model	Year	Features	Method Combined	Sub-Skills*	Skill Hierarchy*	Learning State	Reference
BKT	1994						Ref [7]
BKT-ST	2014	Item Similarity					Ref [20]
Multi-Grained-BKT	2016				✓		Ref [22]
Historical-BKT	2016	Temporal Difference					Ref [22]
Template-Videos	2015	Help					Ref [85]
EEG-KT	2014	Student Mental States					Ref [86]
Affective BKT	2015	Student Mental States					Ref [87]
Spectral BKT	2015					✓	Ref [16]
MS-BKT	2019					✓	Ref [15]
Intervention-BKT	2016	Instructional Interventions					Ref [88]
FAST	2014	General		✓			Ref [89]
TD-BKT	2018	Temporal Difference					Ref [13]
DBN	2014		DBN		✓		Ref [90]
TLS-BKT	2018		Three-Way Decisions			✓	Ref [17]
KT & IRT Combined	2014	Item Difficulty & Student Ability	IRT				Ref [91]
KAT	2014	Gaming Behavior	HMM-IRT				Ref [92]
Hybrid LF-KT	2014	Item difficulty & Student Ability	LFM				Ref [93]

Note: ✓ means that the reference does includes the corresponding aspect.

2.3.1 结合教育数据特征的扩展模型

学生在学习过程中会产生大量的碎片化数据,这些教育数据含有丰富的语义信息^[89,94].传统的 BKT 模型只挖掘了语义信息中随时间变化的纵向信息,而忽略了数据中其余特征信息.因此,学者们结合教育数据中的各类特征信息提出 BKT 的扩展模型,包括结合个性化特征的扩展模型和结合通用特征的扩展模型.

1) 结合个性化特征的扩展模型

BKT 模型对单一知识点建模,使得模型对知识库模型及知识与习题关联模型的依赖性较强.如果知识库模型构建的粒度太粗或太细,会直接影响模

型认知状态评估的性能;如果学生在处理一个知识点关联的连续几道习题十分相似,可能会增加该学生对后面出现相似习题的猜测率^[20].因此,文献[20]提出 BKT-ST 模型,在 BKT 模型的基础上考虑连续习题之间的相似性问题,发现学生在连续作答多道相似习题后其猜测概率提高、失误概率降低、作答表现变好.

文献[22]在 MOOC 平台的 Coursera^①场景下,针对 BKT 模型忽略知识点之间丰富结构和关联性的问题,基于课程章节对应的讲座视频构建具有层次性和时间特性的 Multi-Grained-BKT 模型和 Historical-BKT 模型.文献[85]探讨了在线教育中

① <https://www.coursera.org/>

短视频对模型预测精度的作用.如作答某一习题涉及到工具的使用,该习题关联了记录使用该工具方法的短视频,那么在学生作答该习题时观看该短视频对学生的习题作答产生影响.Template-Videos 利用与习题高度相关的短视频信息提高了 BKT 模型的认知评估性能.脑电图(electroencephalography, EEG)设备可以有效探测学生的心理状态,如学生的注意力^[86].文献[86]将 EEG 方法引入对学生认知状态的评估中,提出的 EEG-KT 模型提高了 BKT 的预测性能.文献[87]将学生的“有效状态”(affective state)如困惑、无聊、投入或愉悦等因素融入 BKT 模型中,提出 Affective BKT 模型并将其应用至可实时评估学生认知状态的智能导学机器人中.

BKT 是一种 2 状态变量(变量仅为“0-1”取值)的一阶 HMM,模型假设处于一个无噪音环境中.文献[16]提出的 Spectral BKT 模型基于特征补偿和模型补偿范式,使用 3-gram 代替原始的“0-1”观测值(观测值仅为“0-1”取值,表示习题作答表现为“错误”或“正确”),给出 8 种光谱观测值,并在传统 BKT 模型 2 个认知状态(隐状态仅为“0-1”取值,表示认知状态为“未掌握状态”或“掌握状态”)之间增加了 2 个状态,提高了模型的预测性能.MS-BKT^[15]使用将传统的 2 个认知状态扩展为 21 个,使得模型可以更好地捕获可观测序列的信息.大多数 BKT 的扩展模型均基于 HMM,具有可解释性强的共性优点,同时也具有共性缺点:学生的学习速率保持不变,不符合客观事实^[15].在保留 BKT 参数可解释性优点的同时,MS-BKT^[15]使用权重参数代替静态学习率,经历多次练习的知识点意味着该学生对该知识点的学习率较低,被赋予较小的权重.实验表明与静态学习率相比,使用动态学习率进行认知状态估计的精度更高.

2) 结合通用特征的扩展模型

2.3.1 节介绍的扩展模型结合个性化特征可在某一方向或目的上达到较好效果,但可能因数据信息补全而具有一定的局限性.接下来介绍结合通用特征的扩展模型,缓解了模型的局限性问题.

文献[88]提出 Intervention-BKT 模型,认为各种类型的教学干预行为影响学生的认知状态,实验结果表明,提出的模型性能优于传统的 BKT 模型,可应用于自适应的教学策略推荐场景中.文献[89]改进估计参数的 EM 算法,从学生作答的观测向量中提取特征向量,并定义算法中 M 步骤的参数为特

征向量的函数,由此提出 FAST(feature aware student knowledge tracing)模型.文献[13]提出了一种基于知识点检测的数据切片算法,用于从学生观测数据中提取时差信息,并将时差信息整合至认知跟踪模型中,提出 TD-BKT 模型考虑了学生观测数据的时间特性,提高了认知状态评估的精度.

2.3.2 混合方法扩展模型

本节将介绍结合其他方法的 BKT 扩展模型.传统的 BKT 模型未考虑不同知识点之间的关联性,文献[90]使用动态贝叶斯网络(dynamic Bayesian network, DBN)模型构造知识点的层级关系,实验结果发现即使是简单的知识点层级关系,也可以将 BKT 的预测精度提高 10%.Zhang 等人^[17]受到三支决策的启发提出 TLS-BKT 模型,对 BKT 算法中 2 状态学习模型进行改进,将原有的“未掌握状态/掌握状态”改进为“未掌握状态/学习状态/掌握状态”,增强了状态模型的灵活性和普遍性.文献[95]应用权重中国餐厅过程(weighted Chinese restaurant process, WCRP)提出基于专家标注技能的认知技能自动发现方法.

为提升模型性能,部分文献将 BKT 模型与 IRT (item response theory), LFM(latent factors model)等方法结合起来.认知跟踪模型常被应用于学生认知状态的评估,未对不同学生和不同习题的差异性建模;而 IRT 可构建不同学生能力、习题难度的模型,但无法跟踪学生的认知状态^[91].文献[91]将 BKT 与 IRT 结合(本文将该模型称为 KT & IRT combined 模型),保留了 BKT 学生建模的优点,通过 IRT 方法提升模型认知状态评估的精度;文献[92]将 BKT 与可以评估学生心理因素的 HMM-IRT^[96]结合,提出 KAT (knowledge and affect tracing)模型可预测学生的作答表现并分析学生的游戏行为;文献[93]提出结合 LFM 与 KT 的 Hybrid LF-KT 模型,在层次贝叶斯模型中将学习的时间动力学理论(temporal dynamics of learning)与学生和习题的差异性相结合以预测学生的作答表现.

BKT, KT & IRT combined, KAT, Hybrid LF-KT 的模型示意图分别如图 6(a)~(d)所示.图 6 中非阴影圆圈表示模型中的隐变量、阴影圆圈方框表示模型中的观测变量,图 6(b)~(d)中的实现方框表示扩展模型比传统 BKT 模型增加的部分,具体含义对应图示部分.

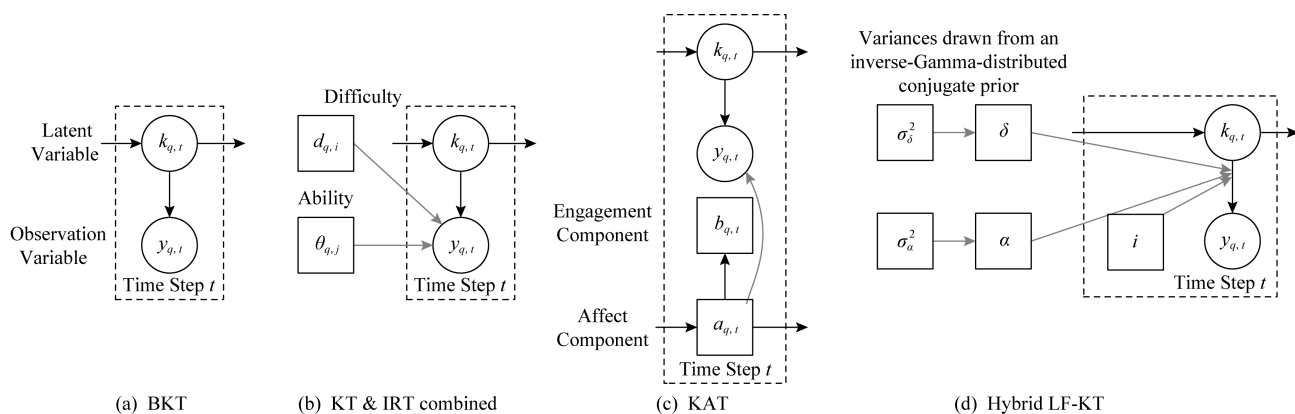


Fig. 6 Models of BKT^[7], KT & IRT combined^[90], KAT^[91], and Hybrid LF-KT^[92]

图6 BKT^[7],KT & IRT combined^[90],KAT^[91],Hybrid LF-KT^[92]模型

2.4 开发工具与公开源码

为降低开发和评估学生模型的成本,Chang 等人^[97]对贝叶斯网络包 BNT^①进行扩展,得到 BNT-SM 工具包.BNT-SM 通过隐藏大部分构造和训练 DBNs 的编码细节,使研究者可以更加专注于学生建模问题.研究这使用 BNT-SM 工具时,将学生模

型以 XML 文件输入,该工具会输出相应训练和评估模型的 BNT MatLab 代码,输出结果的大小至少是输入大小的 5 倍.BNT-SM 为后期学生模型研究者提供了便利有效地实验环境,作者公开该工具^②以供研究者使用.此外,表 4 给出部分基于贝叶斯方法的认知跟踪模型公开源码网址.

Table 4 Open Source Codes of KT Models Based on Bayesian Method

表 4 基于贝叶斯方法的认知跟踪模型公开源码列表

Model	Language	Website
BKT ^[7]	Python	https://github.com/yemao616
	Python	https://github.com/doneria-anjali/student-performance-predictor
	Python	https://github.com/karthickarya09/student-bkt
	C++	https://github.com/IEDMS/standard-bkt
	C++	https://github.com/myudelson/hmm-scalable
WCRP ^[95]	MatLab	https://github.com/CAHLR/xBKT
	Python	https://github.com/Ritesh17/WCRP
FAST ^[89]	C++	https://github.com/robert-lindsey/WCRP
	Java	https://github.com/ml-smores/fast

3 基于深度学习方法的认知跟踪模型

深度学习因其强大的表征能力被应用于各类研究中,基于深度学习方法的认知跟踪模型提高了传统认知跟踪模型的精度.本节介绍经典深度认知跟踪模型 DKT,并梳理 DKT 的扩展模型及其余基于深度学习的认知跟踪模型.

3.1 DKT 模型

人类的知识及大脑具有天然复杂性,因此使用复杂的网络结构构建认知跟踪模型更为合适.然而现有研究大多数基于具有严格约束函数的一阶 HMM. Piech 等人^[31]于 2015 年提出 DKT 模型,将神经网络(recurrent neural network, RNN)应用于认知跟踪问题,在不需要专家标注习题与知识点关系的情况下,显著提升了传统 BKT 模型作答表现预测的性能.

① <http://bnt.sourceforge.net>

② <http://www.cs.cmu.edu/~listen/BNT-SM>

RNN 是一种递归的动态模型,其当前信息基于历史信息及当前输入,与 HMM 相比,RNN 具有高维的、连续的潜在变量表示,因此,RNN 在处理时间序列问题时都得到了很好的成效.认知跟踪模型根

据每一个时间步 $(0, 1, \dots, t)$ 的学生作答表现序列 (y_0, y_1, \dots, y_t) 预测未来时间步的学生作答表现 y_{t+1} ,本质上是一个时间序列问题.基本的 RNN 结构示意图如图 7 所示:

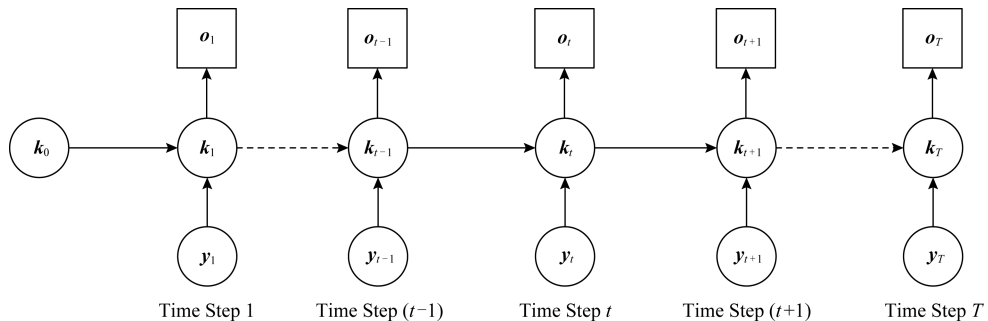


Fig. 7 Mode of DKT^[31]

图 7 DKT 模型^[31]

$(y_1, y_2, \dots, y_T), (k_0, k_1, \dots, k_T), (o_1, o_2, \dots, o_T)$ 分别对应 RNN 中的输入层、隐藏层和输出层.输入层 (y_0, y_1, \dots, y_T) 表示时间步 $1 \sim T$ 的学生的习题作答表现.关于输入层的具体含义见 one-hot 表示法及压缩感知表示法.隐藏层 k_0, k_1, \dots, k_T 中, k_0 为隐藏单元初值, (k_0, k_1, \dots, k_T) 分别为时间步 $1 \sim T$ 的隐藏单元.输出层 (o_1, o_2, \dots, o_T) 表示时间步 $1 \sim T$ 时,正确作答每一道习题的概率.阴影圆圈表示网络的输入,非阴影圆圈表示隐藏单元,方框表示网络的输出.隐藏单元可视作网络的记忆单元,存储着历史隐藏单元的所有信息.由图 7 可见,每一个时间步中输入层的信息向隐藏层单向传递,每一个时间步中隐藏单元的信息向输出层单向传递,上一个时间步隐藏单元的信息向当前时间步的隐藏单元单向传递.由此可得,当前时间步的隐藏单元由历史隐藏单元及当前时间步的输入共同决定,当前时间步的输出单元仅由当前时间步的隐藏单元决定.

输入层为学生的作答表现序列,将学生的作答表现(“正确”、“错误”)通过向量表示.在此介绍 2 种向量表示方法:one-hot 表示法^[98]和压缩感知表示法^[99].假设所有的习题关联的知识点总共有 Q 个.

1) one-hot 表示法.在任意一个时间步 t , 定义一个长度为 $2Q$ 的空向量 y_t 以存储该时间步的学生作答表现.设该时间步时学生作答的习题与知识点 q 关联,若学生正确作答该题,则将向量 y_t 的第 $Q+q$ 位设为 1,其余位均设为 0;若学生错误作答该题,则将向量的第 q 位设为 1,其余位均设为 0.以此类推,即可得输入层的所有向量表示.当知识点的数

目很大时,使用 one-hot 表示法表示的向量稀疏,因此需对向量进行压缩.

2) 压缩感知表示法.通过将学生作答表现分配给长度为 $\lg 2Q$ 的随机高斯输入向量来精确编码.

输出层为学生正确作答每道习题的概率向量序列.设习题数目为 J ,则输出层的向量长度为 J .对于任意一个时间步 $o_t = (o_t^1, o_t^2, \dots, o_t^j, \dots, o_t^J)$, o_t^j 表示学生正确作答习题 j 的概率.

DKT 模型通过采用小批次随机梯度下降法^[100](stochastic gradient descent on minibatches, SGDM)作为优化方法.对于一个学生,计算该学生从时间步 0 至时间步 t 的交叉熵信息之和,优化目标为

$$\min L = \sum_t l(o_t \cdot \delta(q_{t+1}), a_{t+1}). \quad (12)$$

其中, $l(\cdot)$ 表示交叉熵损失函数, (q_{t+1}, a_{t+1}) 表示在第 $t+1$ 时间步学生作答的习题标签 q_{t+1} ,即学生是否正确作答该习题的标签 a_{t+1} , $\delta(q_{t+1})$ 表示第 $t+1$ 时间步学生作答的习题 q_{t+1} 的 one-hot 编码向量.

由此,通过 DKT 模型,可以根据 $(0, 1, \dots, t)$ 时间步的学生作答表现向量 (y_0, y_1, \dots, y_t) 得到第 t 个时间步的习题正确作答概率向量 o_t ,根据第 $t+1$ 个时间步作答的习题得到正确作答该习题的概率 $P(y_{t+1} | x_1, x_2, \dots, x_t)$.

3.2 DKT 扩展模型

DKT 模型将 RNN 应用到认知跟踪模型中,并取得了较好的预测精度,受到了学者们的广泛关注.本节将介绍利用数据特征信息及结合其他方法 2 类 DKT 扩展模型.表 5 是本节介绍的所有 DKT 扩展模型的对比列表.

Table 5 Comparison of the Variants of DKT Model

表 5 DKT 扩展模型的对比

Model	Year	Features	Method Combined	Skill Hierarchy*	Reference
DKT	2015				Ref [31]
DKT-t	2019	Time-Gap			Ref [14]
PDKT-C	2018	Skill Relationship		√	Ref [101]
DKTS	2019	Question Relationship			Ref [38]
E2E-DKT	2018	Student Performance Logs			Ref [102]
DKT-DSC	2018	Student Level	K-Means		Ref [103]
Classifier-based DKT	2018	Heterogeneous Features	Tree-Based Classifiers		Ref [104-105]
Adaptive DKT	2017	Question & Student Level	Auto Encoder		Ref [34]
DKT+	2018		Regularization		Ref [37]

Note: √ means that the reference does includes the corresponding aspect.

3.2.1 结合教育数据特征扩展模型

文献[34]通过引入习题层面和学生层面的更多特征信息扩展 DKT 模型(本文称之为 Adaptive DKT),增加自编码网络层将输入转换成低维特征向量,减少了模型训练所需的资源和时间.文献[14]通过分析 Fuutoot 平台(一种个性化的自适应学习平台)获取的学生数据,学生尝试作答 2 个习题的时间间隔可能高至 1 h 或 1 d,甚至 1 周.Lalwani 等人[14]将时间间隔信息作为 DKT 模型的一个特征编码,提出 DKT-t 模型,实验发现 DKT-t 模型可提高 DKT 模型的预测性能,同时还可根据学生不同的作答序列及认知状态跟踪学生的遗忘曲线.

学生数据的稀疏问题一直影响着认知跟踪模型的预测精度和模型复杂度.Chen 等人^[101]认为从知识库中充分探讨知识点之间的相互依赖关系有助于解决上述数据稀疏问题.文献[101]提出 PDKT-C 模型,通过引入知识点之间的先决关系,将其作为模型中的一个约束,以提高模型的精度.文献[38]认为习题之间也存在相关关系(side relations).比如 2 个习题关联的知识点之间具有相似关系或其他隐含关系,那么这 2 个习题必然存在相关关系.这些相关关系可以构造成为一个习题子图,该子图可为认知跟踪提供更加丰富的特征信息.因此,Wang 等人^[38]提出 DKTS 模型,将习题之间关系的特征信息整合至 DKT 模型中,提升了模型的预测性能.

文献[102]提出无需知识点标签的 end-to-end 模型 E2E-DKT,通过学生作答习题记录的日志中自动学习习题和知识点的向量嵌入.实验证明了该模型学习的向量嵌入对 DKT 的性能具有促进作用.

3.2.2 混合模型

文献[103]将学习能力相似的学生聚类为一个

群体,基于 K-Means 聚类方法提出了 DKT-DSC 模型,提高了 DKT 模型个性化的认知跟踪能力.文献[104-105]引入学生数据的异构特征,分别结合决策树、随机森林、GBDT 这 3 种树形分类器预测学生的作答表现(本文称该方法为 Classifier-based DKT).

Yeung 等人^[37]发现 DKT 模型在预测学生的认知水平时,可能学生对一个知识点的作答表现较好,但预测结果反而下降;同时,作者认为学生的认知状态会随着时间的推移逐渐改变,而不是在掌握状态和未掌握状态之间交替变换.为解决上述 2 个问题,文献[37]提出 DKT+ 模型,定义了“重构错误”(reconstruction error)和“波动准则”(waviness measures)作为正则化损失函数来扩充 DKT 模型的原始损失函数.

3.3 基于深度学习方法的其他模型

DKT 模型将 RNN 应用于认知跟踪领域,开启了深度学习方法在该领域的应用.表 6 为本节介绍的基于深度学习方法的其他认知跟踪模型的对比列表.表 7 给出部分基于深度学习方法的认知跟踪模型公开源码网址.

在复杂交互学习中,学生解决某个问题需要掌握多个知识,甚至多个知识的组合会产生更多的额外知识.针对这种情况,Huang 等人^[21,106]考虑了知识的应用上下文,提出了一个数据驱动(data-driven)框架 CKM-HSC,追踪学生更深层次的认知状态.

文献[107]提出了一种可解释的概率知识能力跟踪模型(knowledge proficiency tracing, KPT),设计了一个概率矩阵分解框架将学生和习题先验知识结合起来,跟踪学生的知识熟练程度,并通过实验证明了 KPT 的有效性和可解释性.Su 等人^[19]考虑将习题信息与学生的做题记录相结合,设计了双向

Table 6 Comparison of Other Deep-Learning-Based KT Models**表 6 其余深度学习 KT 模型的对比**

Model	Year	Features	Method	Skill hierarchy*	Reference
CKM-HSC	2016	Skill Relationship	Data-Driven Techniques	✓	Ref [21, 106]
KPT	2017	Q-matrix Prior	Matrix Factorization		Ref [107]
EERNN	2018	Question Feature	RNN		Ref [19]
EKT	2019	Question Feature	RNN	✓	Ref [10]
DKVMN	2017		MANN		Ref [108]

Note: ✓ means that the reference does includes the corresponding aspect.

Table 7 Open Source Codes of KT Models Based on Deep Learning Methods**表 7 基于深度学习的认知跟踪模型开源码列表**

Type	Model	Language	Website
DKT and Its Variants	DKT ^[31]	Lua	https://github.com/chrispiech/DeepKnowledgeTracing
		Python	https://github.com/jiangxinyang227/dkt
	DKT+[37]	Python	https://github.com/mmkhajah/dkt
	DKT-DSC ^[103]	python	https://github.com/Songlielie/DKT-DSC
Other Deep-Learning-Based Methods	DKVMN ^[108]	Python	https://github.com/jennyzhang0215/DKVMN
		Python	https://github.com/yjhong89/DKVMN
	KTM ^[106]	Python	https://github.com/jilljenn/ktm

LSTM 学习习题信息的编码,提出 exercise-enhanced recurrent neural network(EERNN)框架;同时提出 2 种预测策略:基于 Markov 特性的 EERNNM (EERNNM with Markov property)和基于注意力机制的 EERNNA(EERNNA with attention).但该框架对学生的认知状态采用隐向量表示,无法显示跟踪学生的认知状态,因此 Huang 等人^[10]进一步提出 EKT(exercise-aware knowledge tracing)框架,实现了对学生关于每道题各知识点的认知状态跟踪。

为解决 DKT 无法显示跟踪学生对每一知识概念的掌握水平问题,文献^[108]将改进的 MANN (memory-augmented neural networks)应用于 KT 问题;通过引入 key-value 存储单元,提出 DKVMN (dynamic key-value memory networks)框架以发现每个输入习题的潜在知识概念、跟踪学生对知识点的认知状态。DKVMN 比 DKT 具有更大容量的外部存储空间,因此需要更少的参数。

4 认知跟踪模型的应用

认知跟踪模型对学生的习题作答表现进行建模,广泛应用于各个在线教育平台,如 MOOC 平台的 Coursera^[22], edX^[109], MATHEISIS Algebra Tutor^[110]

等。本节列出认知跟踪模型分别在作答表现预测、认知状态评估、心理因素分析、习题序列研究和特定的编程练习中的应用。

4.1 作答表现预测

通过跟踪学生的习题作答表现序列,预测未来时刻学生的习题作答表现,是认知跟踪模型的重要应用之一,在此举例说明。

文献^[111]探讨了集成方法(ensemble methods)是否可以在有导师学习中提高后试测验(post-test)的预测精度,通过实验发现集成方法未促进该预测精度的提升,并给出了可能的解释。文献^[22]基于 BKT 模型,结合知识点中丰富的关联关系信息提出了 Multi-Grained-BKT 和 Historical-BKT 模型,通过实验证明了提出方法在 Coursera 的 MOOC 课程数据中的预测性能较好。文献^[10]在 EERNN 框架^[19]的基础上,提出 2 种预测策略:基于 Markov 性质的 EERNNM 和基于注意力机制的 EERNNA,并进一步提出 EKT 框架实现了对学生认知状态的评估和学生作答表现的预测。

4.2 认知状态评估

本文 4.1 节介绍了认知跟踪模型在学生作答表现预测上的应用,事实上部分文献通过首先评估学生的认知状态,进而实现作答表现的预测,如文献

[10,19,22];部分文献提出的模型可以预测学生的作答表现,却无法评估学生的认知状态,如文献[31].

举例说明认知跟踪模型可评估认知状态的情况.在2.1节中介绍了使用BKT模型进行认知状态评估任务时,为缓解“过度积极”和“过度消极”问题,通常设置一个认知阈值作为学生是否掌握知识的判定条件.实际上,认知阈值可作为一个可调参数来控制2类问题发生的相对频率^[71].文献[71]将练习机会序列分成了“学习”(learning)、“过渡”(lag)和“过度练习”(over-practice)3个阶段,对这3个阶段进行量化分析,并阐明了BKT模型中的认知阈值与这3个阶段之间的关系.文献[112]使用BKT模型估计学生掌握知识成分的可能性,然后在整个学习过程中对学生的学习曲线进行建模.文献[107]设计了一个概率矩阵分解框架,结合学生练习的先验知识,跟踪学生的认知掌握程度.

4.3 心理因素分析

学生的习题作答表现不仅取决于学生对知识的掌握情况和学生应用知识的能力,还有很多其他因素影响作答表现,如遗忘、猜测、失误、游戏行为、情绪状态等心理因素.

Xu等人^[86]使用脑电图设备以探测学生学习时的心理状态,针对不同的心理状态将学生的学习、遗忘、猜测和失误因素设置为不同的参数,提高了BKT模型的预测性能.

在智能教辅系统ITS的学生模型应用中,大多数研究假设学生对知识是无遗忘的,却未有文献证明此假设是符合客观事实的.文献[113]考虑了遗忘(forgetting)和再学习(relearning)因素,对学生长期学习建模十分重要.Lalwani等人^[14]通过Fuutoot平台发现学生作答习题的间隔可能为1h或1d,甚至1周之久,因此作者使用时间间隔信息扩展DKT模型,提出了可以追踪学生遗忘曲线的DKT-t模型.

游戏行为(gaming the system)指学生不断地利用系统的反馈及帮助来获得习题的正确答案^[114].比如,某学生在作答多项选择题时,并没有试图找到正确答案,而是选择了每道题的所有选项,以此提高其点击正确答案的频次.文献[115]提出可以预测游戏行为的认知跟踪模型KTB(knowledge tracing with behavior),尝试在利用BKT模型预测学生作答表现的同时,预测学生的游戏行为.

Spaulding等人^[87]构建了一个智能教辅机器人,通过增加学生情绪状态的估计以扩展BKT模型.学生的情绪状态包括学生是否感到困惑、无聊、

精力集中或心情愉悦等.同时还探讨了相较于智能教辅系统,学生在与智能教辅机器人交互时是否可以展现出更多情绪状态的问题.

4.4 习题序列研究

认知跟踪模型中,学生的习题作答数据是以时间为顺序的.因此学生作答习题的次序可能会影响学生的学习效率,从而影响其作答表现.

文献[31]说明了DKT模型可以通过式(13)发现习题之间的先决关系.其中, $y(j|i)$ 表示在已知上一时间步学生正确作答了习题*i*的情况下,该学生在这一步正确作答习题*j*的概率, K 表示学生作答的习题数目:

$$J_{ij} = \frac{y(j|i)}{\sum_{k=1}^K y(j|k)}. \quad (13)$$

文献[116]应用具有天然的层次性RBT(revised bloom's taxonomy)^[117]作为习题次序策略,研究学生获得知识的次序对学生作答表现的影响,使学生先进行低层次的训练再进行高层次的训练,以提升模型精度.

4.5 编程练习中的认知跟踪

学生在开放习题中的训练可以为学生作答数据提供更加丰富的信息.编程题^[8]即属于开放习题训练中的一种.

文献[7]对学生在ACT Programming Tutor短程序编写系统进行Lisp语言编程练习建模,提出了经典的BKT模型.在学生学习程序设计技术方面,为解决学生应用哪种编程结构以及有多少学生理解编程语言的概念等问题,文献[118]提出三相评估(three-phase measuring)方法,用于观察学生的编程错误、应用编程结构,同时应用贝叶斯学习模型确定编程知识概念.文献[9]基于抽象语法树(abstract syntax tree, AST)提出了一种认知跟踪模型,可以在给定的编程练习中评估学生的认知状态.基于DKT模型的思想,文献[8]将嵌入式程序输入到一个RNN中,并将其用以预测未来编程练习中学生的表现情况.

5 比较、分析与展望

认知跟踪模型在智慧教育^[119]、智能教辅系统^[120]、在线教育^[121]等方面具有广泛的应用需求,因此对认知跟踪模型的研究是计算机领域和教育领域的热点问题.通过对2类认知跟踪模型的梳理,将

当前研究归纳为 5 点:1)认知跟踪模型的参数误差和评价指标;2)现有认知跟踪模型的分析;3)新的认知跟踪模型的提出;4)认知跟踪模型开发工具的研究;5)认知跟踪模型应用的研究。

本文从建模方法上将认知跟踪模型分为基于贝叶斯方法和基于深度学习方法 2 类分别介绍。2 类方法均是根据习题作答表现跟踪学生的认知状态,但因建模方法的不同而各自具有优缺点,本节首先以认知跟踪模型中经典的 BKT 模型和 DKT 模型为例进行分析。

5.1 BKT 与 DKT 的比较

本文分别在 2.1 节和 3.1 节详细阐述了 BKT 模型和 DKT 模型,这 2 个模型分别是基于贝叶斯方法和基于深度学习方法经典的认知跟踪模型,二者各有优劣,本节将以二者为例对 2 类认知跟踪模型进行比较分析。

根据图 6 的 BKT 模型和图 7 的 DKT 模型,可发现 BKT 针对单个知识点建模,其观测变量与知识点之间的关系已知,而 DKT 将所有知识点构建为一个模型,其观测变量与知识点之间的关系未知。这是因为在面向某学生的作答数据时,BKT 模型首先将学生作答数据按照习题关联的知识点分组,每一个知识点关联的所有习题作答形成一个序列;相反地,DKT 模型将一个学生作答的所有知识点关联的习题构建为一个序列。

举例如下:目前有学生 i 的作答数据,将与知识点 A 关联的习题记作 A_x , x 为以习题出现的时间为序的编号,以此类推,学生 i 作答的习题序列为 $A_1A_2B_1A_3C_1B_2B_3$ 。在 BKT 模型中,对应学生 i 的作答表现序列被列为 3 组,分别是:1)对应原习题序列第 1 题、第 2 题和第 4 题的作答表现 $y_{A,1}, y_{A,2}, y_{A,3}$;2)对应原习题序列第 3 题、第 6 题和第 7 题的作答表现 $y_{B,1}, y_{B,2}, y_{B,3}$;3)对应原习题序列第 5 题的作答表现 $y_{C,1}$,将这 3 组作答表现序列分别建模成 HMM 模型进行计算。而在 DKT 模型中,学生 i 的作答表现序列仅被列为 1 组,即 $y_1, y_2, y_3, y_4, y_5, y_6, y_7$,分别对应原习题序列的第 1~7 题,将此作答表现序列建模成一个 RNN 模型进行计算。由举例可得:

1) 由于 BKT 模型需要将原习题序列按照关联的知识点进行分组,因此其建模过程依赖于习题与知识点之间的关联情况,需要进行人工标注;而 DKT 模型无需将原习题序列分组,节省了大量人工标注的代价。

2) 原习题序列中的第 4 题在按照知识点分组后出现在了第 1 个分组的第 3 题,即 $A_1A_2B_1A_3C_1B_2B_3$ 中 A_3 的作答表现在分组过后出现在了 $y_{A,1}, y_{A,2}, y_{A,3}$ 中的 $y_{A,3}$,习题出现的时间特征丢失,而 DKT 模型却保留了这一特性。文献[122]分析,DKT 模型正是因为有效利用了学生作答表现数据中的特征和规律,性能得到了大幅度地提升;同时,若总结数据中的特征和规律并应用至 BKT 模型中,也可以得到很好的效果。

3) 通过 BKT 的隐状态 $k_{q,t}$,可追踪学生在任意时间步 t 对知识点 q 的认知状态,这体现了 BKT 模型强大的可解释性;而 DKT 模型的隐藏层虽起到了至关重要的信息传递作用,但是却无法解释其中的含义,即 DKT 模型无法追踪学生的认知状态。

通过以上分析,BKT 模型和 DKT 模型各自的优点和不足之处总结为 2 点,同时也适用于一般基于贝叶斯方法和基于深度学习方法 2 类认知跟踪模型。

1) BKT 可对学生的认知状态进行跟踪,具有强大的可解释性;但是,BKT 也因模型本身的表示能力,丢失了学生数据中的一些重要特征和规律,其预测性能有待提升;

2) DKT 因其拥有较强的表示能力,可以有效利用学生数据中一些重要的特征和规律,具有较好的预测性能;但是,DKT 目前无法显式得到学生对每个具体的知识点的认知状态,可解释性有所欠缺。

5.2 分析与展望

基于已有研究工作及模型的优缺点,本节从认知跟踪模型的建模任务、数据特点、建模方法及优缺点、建模的难点与挑战 4 个方面对认知跟踪的方向进行分析与展望。

1) 建模任务。认知跟踪模型面向动态的教育数据建模,基于学生在学习过程中作答习题的数据,跟踪学生对知识点的掌握随时间的变化情况,预测学生在未来时刻作答习题的表现^[7]。如图 5 所示阐述了认知跟踪的问题模型包含输入、建模及输出。输入为学生在不同时刻作答不同知识点习题的观测得分;建模为将输入层的输入数据使用建模方法进行模型拟合、参数寻优,得到对每个学生的认知跟踪子模型;输出为通过建模层得到的认知跟踪模型,根据不同的应用需求得到不同的输出结果;其中,认知跟踪模型可应用于学生认知状态的跟踪、作答表现的预测等场景中。

2) 数据特点。认知跟踪模型的输入为学生在不同时刻作答不同知识点习题的观测得分数据。由于

认知跟踪模型是面向动态学习数据的,因此模型的数据应具有时间特性;由于认知跟踪模型针对每个学生训练建模,因此模型的数据应有足够数目学生的作答数据,以确保训练的效果;此外,在认知跟踪模型应用于学生认知状态的跟踪时,若要得到显性的学生对知识点的认知状态,数据应包含课程的知识点集合、学生作答的习题与知识点的关联关系。从实际情况考虑,知识点应是树状结构或网状结构的多粒度状态,而非知识点之间相互独立。

3) 建模方法及优缺点,目前主流的建模方法包括贝叶斯方法(如 BKT 模型)和深度学习方法(如 DKT 模型)。^①贝叶斯方法:贝叶斯方法将学习过程建模为 HMM。依据 HMM 的时间特性,假设学生在当前时刻的作答表现只与学生上一时刻的作答表现以及学生当前时刻的认知状态相关,从而可构建出基于 HMM 的贝叶斯认知跟踪模型。因此,贝叶斯方法的认知跟踪模型要求对每一学生的不同知识点作答数据分开独立建模。其优点是可对学生的认知状态进行跟踪,具有较好的可解释性。而由于模型本身的表示能力,该模型的缺点是丢失了学生数据中的一些重要特征和规律,使其预测性能有待提升;并且该模型要求标注每道习题涉及到的知识点集合。^②深度学习方法:在习题涉及到的知识点集合未标注的情况下,可使用深度学习方法对输入数据进行特征表示及建模,将学生对知识点的认知状态作为隐藏单元,对学生的作答数据进行拟合。其优点是具有强大的表示能力,能有效利用学生数据中一些重要的特征和规律,因此具有较好的预测性能。而该模型的缺点是目前的深度学习模型无法得到学生的认知状态,因此可解释性需要加强。

4) 建模的难点与挑战。认知跟踪问题的建模依然存在 3 方面难点与挑战。

① 贝叶斯认知跟踪的模型参数寻优问题。

贝叶斯认知跟踪模型的参数取值直接影响模型的性能。有学者甚至发现模型的参数虽然取值不同但可能导致相同的预测结果,即模型的识别性问题。现有的贝叶斯认知跟踪模型中,参数寻优方法可能存在 4 方面问题:Ⅰ.在保证模型精度的情况下降低参数寻优的时间复杂度和空间复杂度问题;Ⅱ.在数据有限的情况下防止参数的过拟合问题;Ⅲ.设计参数估计算法使其不陷入局部最优解问题;Ⅳ.结合其他类型的教育数据以更好地估计参数值问题。

② 认知跟踪模型的精度与可解释性的权衡问题。

结合 5.1 节的比较分析可知,基于贝叶斯方法

和深度学习方法的 2 类主流认知跟踪模型均有一定的局限性。因此,部分研究通过结合其他类型的数据特征或引入混合方法以提高贝叶斯认知跟踪模型的精度,但是现有模型的精度问题仍有待提升。而基于深度学习的认知跟踪模型虽然具有较高的精度,但在可解释性方面有所欠缺。因此综合贝叶斯认知跟踪模型及深度认知跟踪模型的优点,权衡精度和可解释性的认知跟踪模型仍需要进一步研究。

③ 模型部分假设与实际应用不符的问题。

现有认知跟踪模型的部分假设不符合实际情况。例如现有的大多数模型有 5 种假设:Ⅰ.假设学生在学习过程中无遗忘,而实际上学生在学习过程中会随着时间的推移而遗忘某些知识;Ⅱ.假设标注的知识点之间完全独立,而实际中不同课程之间、同一课程之内的知识点之间可能存在包含、继承、前驱、等价等不同的关系;Ⅲ.假设不同知识点处于同一层级级别,而知识点可能因包含关系存在不同的级别,一个完整的知识体系应该拥有多粒度的层级关系;Ⅳ.假设习题之间没有关联,而实际上在知识点之间存在包含前继等关系的情况下,习题之间也存在关联;Ⅴ.假设每道习题只与一个知识点相关联,而实际上很多综合性习题涉及到多个知识点。因此,认知跟踪模型的研究还需针对实际场景进行进一步的研究。

6 总 结

随着教育信息化的推进,教育领域迎来了大数据时代。从传统教育和在线教育中产生的海量碎片化数据中挖掘与学生相关的认知状态、行为表现、心理因素等信息具有重要的意义。认知跟踪模型根据学生动态的习题作答表现跟踪学生的认知状态、预测学生的未来作答表现,广泛应用于 MOOC 等在线教育平台中。本文从建模方法上将认知跟踪模型分成贝叶斯认知跟踪模型和深度认知跟踪模型,对其进行详细地梳理和分析,并介绍了认知跟踪方法的 5 类应用。同时,以经典的 BKT 和 DKT 模型为例分析了 2 类认知跟踪模型的优缺点,并对未来可能的研究方向进行了探讨和展望。

参 考 文 献

- [1] Bakhshinatogh B, Zaiane O R, ElAtia S, et al. Educational data mining applications and tasks: A survey of the last 10 years [J]. Education and Information Technologies, 2017, 23 (1): 537-553

- [2] Wu Minghui, Wu Xindong. On big wisdom [J]. Knowledge and Information Systems, 2019, 58(1): 1-8
- [3] Lu Ruqian, Jin Xiaolong, Zhang Songmao, et al. A study on big knowledge and its engineering issues [J]. IEEE Transactions on Knowledge Data Engineering (TKDE), 2019, 31(9): 1630-1644
- [4] Wu Xindong, Chen Huanhuan, Liu Jun, et al. Knowledge engineering with big data (BigKE): A 54-month, 45-million RMB, 15-institution national grand project [J]. IEEE Access, 2017, 5(1): 12696-12701
- [5] Wu Xindong, He Jin, Lu Ruqian, et al. From big data to big knowledge: HACE + BigKE [J]. Acta Automatica Sinica, 2016, 42(7): 965-982 (in Chinese)
(吴信东, 何进, 陆汝钤, 等. 从大数据到大知识: HACE + BigKE [J]. 自动化学报, 2016, 42(7): 965-982)
- [6] Yang Fan, Li F W B. Study on student performance estimation, student progress analysis, and student potential prediction based on data mining [J]. Computers & Education, 2018, 123(1): 97-108
- [7] Corbett A T, Anderson J R. Knowledge tracing: Modeling the acquisition of procedural knowledge [J]. User Modeling and User-Adapted Interaction, 1994, 4(4): 253-278
- [8] Wang L, Sy A, Liu L, et al. Deep knowledge tracing on programming exercises [C] //Proc of the 4th ACM Conf on Learning@Scale (L@S). New York: ACM, 2017: 201-204
- [9] Jiang Bo, Ye Yun, Zhang Haifeng. Knowledge tracing within single programming exercise using process data [C] //Proc of the 26th Int Conf on Computers in Education (ICCE). Taoyuan, Taiwan: APSCE, 2018: 89-94
- [10] Huang Zhenya, Yin Yu, Chen Enhong, et al. EKT: Exercise-aware knowledge tracing for student performance prediction [J/OL]. IEEE Transactions on Knowledge and Data Engineering (TKDE), 2019 [2020-10-30]. <https://ieeexplore.ieee.org/abstract/document/8744302/>
- [11] Meng Lingling, Zhang Mingxin, Zhang Wanxue, et al. CS-BKT: Introducing item relationship to the Bayesian knowledge tracing model [J/OL]. Interactive Learning Environments, 2019 [2020-10-30]. <https://www.tandfonline.com/doi/abs/10.1080/10494820.2019.1629600>
- [12] Minn S, Zhu Feida, Desmarais M C. Improving knowledge tracing model by integrating problem difficulty [C] //Proc of IEEE Int Conf on Data Mining Workshops (ICDMW). Piscataway, NJ: IEEE, 2018: 1505-1506
- [13] Zhu Junhu, Zang Yichao, Qiu Han, et al. Integrating temporal information into knowledge tracing: A temporal difference approach [J]. IEEE Access, 2018, 6(1): 27302-27312
- [14] Lalwani A, Agrawal S. What does time tell? Tracing the forgetting curve using deep knowledge tracing [G] //LNCS 11626: Proc of the 20th Int Conf on Artificial Intelligence in Education (AIED). Berlin: Springer, 2019: 158-162
- [15] Agarwal D, Baker R S. Dynamic knowledge tracing through data driven recency weights [C/OL] //Proc of the 27th Int Conf on Computers in Education. Taoyuan, Taiwan: APSCE, 2019 [2020-10-30]. <https://eric.ed.gov/?id=ED607821>
- [16] Falakmasir M, Yudelso M, Ritter S, et al. Spectral Bayesian knowledge tracing [C] //Proc of the 8th Int Conf on Educational Data Mining (EDM). Worcester, MA: IEDMS, 2015: 360-363
- [17] Zhang Kai, Yao Yiyu. A three learning states Bayesian knowledge tracing model [J]. Knowledge-Based Systems (KBS), 2018, 148(1): 189-201
- [18] Pardos Z A, Heffernan N T. KT-IDEM: Introducing item difficulty to the knowledge tracing model [G] //LNCS 6787: Proc of User Modeling, Adaptation, and Personalization (UMAP). Berlin: Springer, 2011: 243-254
- [19] Su Yu, Liu Qingwen, Liu Qi, et al. Exercise-enhanced sequential modeling for student performance prediction [C] //Proc of the 32nd AAAI Conf on Artificial Intelligence (AAAI). Menlo Park, CA: AAAI, 2018: 2435-2443
- [20] Hawkins W J, Heffernan N T. Using similarity to the previous problem to improve Bayesian knowledge tracing [C] //Proc of the 7th Int Conf on Educational Data Mining (EDM). Worcester, MA: IEDMS, 2014: 136-140
- [21] Huang Yun, Hollstein J D G, Brusilovsky P. Deeper knowledge tracing by modeling skill application context for better personalized learning [C] //Proc of the 24th Int Conf on User Modeling, Adaptation and Personalization (UMAP). New York: ACM, 2016: 325-328
- [22] Wang Zhuo, Zhu Jile, Li Xiang, et al. Structured knowledge tracing models for student assessment on coursera [C] //Proc of the 3rd ACM Conf on Learning@Scale (L@S). New York: ACM, 2016: 209-212
- [23] Zhou Xuan, Wu Wenjun, Han Yong. Modeling multiple subskills by extending knowledge tracing model using logistic regression [C] //Proc of IEEE Int Conf on Big Data. Piscataway, NJ: IEEE, 2018: 3994-4003
- [24] Qiu Yumeng, Qi Yingmei, Lu Hanyuan, et al. Does time matter? Modeling the effect of time in Bayesian knowledge tracing [C] //Proc of the 4th Int Conf on Educational Data Mining (EDM). Worcester, MA: IEDMS, 2011: 139-148
- [25] Baker R S J d, Corbett A T, Aleven V. More accurate student modeling through contextual estimation of slip and guess probabilities in Bayesian knowledge tracing [G] //LNCS 5091: Proc of Intelligent Tutoring System (ITS). Berlin: Springer, 2008: 406-415
- [26] Pardos Z A, Heffernan N T. Modeling individualization in a Bayesian networks implementation of knowledge tracing [G] //LNCS 6075: Proc of User Modeling, Adaptation, and Personalization (UMAP). Berlin: Springer, 2010: 255-266

- [27] Wang Yutao, Heffernan N T. Extending knowledge tracing to allow partial credit: Using continuous versus binary nodes [G] //LNCS 7926: Proc of Int Conf on Artificial Intelligence in Education (AIED). Berlin: Springer, 2013; 181-188
- [28] Yudelson M V, Koedinger K R, Gordon G J. Individualized Bayesian knowledge tracing models [G] //LNCS 7926: Proc of Int Conf on Artificial Intelligence in Education (AIED). Berlin: Springer, 2013; 171-180
- [29] Pu Yanjun, Wu Wenjun, Han Yong, et al. Parallelizing Bayesian knowledge tracing tool for large-scale online learning analytics [C] //Proc of IEEE Int Conf on Big Data (Big Data). Piscataway, NJ: IEEE, 2018; 3245-3254
- [30] Feng Junchen, Zhang Bo, Li Yuchen, et al. Bayesian diagnosis tracing: Application of procedural misconceptions in knowledge tracing [G] //LNCS 11626: Proc of the 20th Int Conf on Artificial Intelligence in Education (AIED). Berlin: Springer, 2019; 84-88
- [31] Piech C, Bassen J, Huang J, et al. Deep knowledge tracing [C] //Proc of the 28th Int Conf on Neural Information Processing Systems (NeurIPS). Cambridge, MA: MIT, 2015; 505-513
- [32] Tang S, Peterson J C, Pardos Z A. Deep neural networks and how they apply to sequential education data [C] //Proc of the 3rd ACM Conf on Learning@Scale (L@S). New York: ACM, 2016; 321-324
- [33] Zhang Jiani, King I. Topological order discovery via Deep Knowledge Tracing [G] //LNCS 9950: Proc of Int Conf on Neural Information Processing (ICONIP). Berlin: Springer, 2016; 112-119
- [34] Zhang Liang, Xiong Xiaolu, Zhao Siyuan, et al. Incorporating rich features into deep knowledge tracing [C] //Proc of the 4th ACM Conf on Learning@Scale (L@S). New York: ACM, 2017; 15-25
- [35] Lalwani A, Agrawal S. Validating revised bloom's taxonomy using deep knowledge tracing [G] //LNCS 10947: Proc of Int Conf on Artificial Intelligence in Education (AIED). Berlin: Springer, 2018; 225-238
- [36] Swamy V, Guo A, Lau S, et al. Deep knowledge tracing for free-form student code progression [G] //LNCS 10948: Proc of Int Conf on Artificial Intelligence in Education (AIED). Berlin: Springer, 2018; 348-352
- [37] Yeung C K, Yeung D Y. Addressing two problems in deep knowledge tracing via prediction-consistent regularization [J/OL]. //Proc of the 5th ACM Conf on Learning@Scale (L@S). New York: ACM, 2018 [2020-10-30]. <https://dl.acm.org/doi/abs/10.1145/3231644.3231647>
- [38] Wang Zhiwei, Feng Xiaoqin, Tang Jiliang, et al. Deep knowledge tracing with side information [G] //LNCS 11626: Proc of the 20th Int Conf on Artificial Intelligence in Education (AIED). Berlin: Springer, 2019; 303-308
- [39] Li Feiming, Ye Yanwei, Li Xiaofei, et al. Application of knowledge tracing model in education: A review from 2008 to 2017 [J]. Distance Education in China, 2018, 1(7): 86-91 (in Chinese)
(李菲茗, 叶艳伟, 李晓菲, 等. 知识追踪模型在教育领域的应用: 2008—2017年相关研究的综述[J]. 中国远程教育, 2018, 1(7): 86-91)
- [40] Vanlehn K. Student Modeling [M]. East Sussex, UK: Psychology Press, 1988; 55-62
- [41] Chrysafiadi K, Virvou M. Student modeling approaches: A literature review for the last decade [J]. Expert Systems with Applications, 2013, 40(11): 4715-4729
- [42] Pelanek R. Bayesian knowledge tracing, logistic models, and beyond: An overview of learner modeling techniques [J]. User Modeling and User-Adapted Interaction, 2017, 27(3/4/5): 313-350
- [43] Xu Pengfei, Zheng Qinhu, Chen Yaohua, et al. Research on learner modeling in education data mining [J]. Distance Education in China, 2018, 521(6): 7-13 (in Chinese)
(徐鹏飞, 郑勤华, 陈耀华, 等. 教育数据挖掘中的学习者建模研究[J]. 中国远程教育, 2018, 521(6): 7-13)
- [44] Jiang Zhuoxuan, Zhang Yan, Li Xiaoming. Learning behavior analysis and prediction based on MOOC data [J]. Journal of Computer Research and Development, 2015, 52(3): 614-628 (in Chinese)
(蒋卓轩, 张岩, 李晓明. 基于MOOC数据的学习行为分析与预测[J]. 计算机研究与发展, 2015, 52(3): 614-628)
- [45] Gong Yue, Beck J E, Heffernan N T. Comparing knowledge tracing and performance factor analysis by using multiple model fitting procedures [G] //LNCS 6094: Proc of Int Conf on Intelligent Tutoring Systems (ITS). Berlin: Springer, 2010; 35-44
- [46] Gong Yue, Beck J E, Heffernan N T. How to construct more accurate student models: Comparing and optimizing knowledge tracing and performance factor analysis [J]. International Journal of Artificial Intelligence in Education, 2011, 21(1/2): 27-46
- [47] Wang Yan. Defining mastery: Knowledge tracing versus N-consecutive correct responses [D]. Worcester, MA: Worcester Polytechnic Institute, 2016
- [48] Zhou Qing, Mou Chao, Yang Dan. Research progress on educational data mining: A survey [J]. Journal of Software, 2015, 26(11): 3026-3042 (in Chinese)
(周庆, 牟超, 杨丹. 教育数据挖掘研究进展综述[J]. 软件学报, 2015, 26(11): 3026-3042)
- [49] Chen Penghe, Yu Lu, Zheng V W, et al. KnowEdu: A system to construct knowledge graph for education [J]. IEEE Access, 2018, 6(1): 31553-31563
- [50] Torre J D L. DINA model and parameter estimation: A didactic [J]. Journal of Educational and Behavioral Statistics, 2009, 34(1): 115-130

- [51] Lian Defu, Ye Yuyang, Zhu Wenya, et al. Mutual reinforcement of academic performance prediction and library book recommendation [C] //Proc of the 16th Int Conf on Data Mining (ICDM). Piscataway, NJ: IEEE, 2016; 1023-1028
- [52] Zhou Yuwen, Huang Changqin, Hu Qintai, et al. Personalized learning full-path recommendation model based on LSTM neural networks [J]. Information Sciences, 2018, 444(1): 135-152
- [53] Jiang Changmeng, Feng Jun, Sun Xia, et al. Personalized exercises recommendation algorithm based on knowledge hierarchical graph, ReKHG [J]. Computer Engineering and Application, 2018, 54(10): 229-235 (in Chinese)
(蒋昌猛, 冯筠, 孙霞, 等. 基于知识点层次图的个性化习题推荐算法[J]. 计算机工程与应用, 2018, 54(10): 229-235)
- [54] Xie Zhenping, Jin Chen, Liu Yuan. Personalized knowledge recommendation model based on constructivist learning theory [J]. Journal of Computer Research and Development, 2018, 55(1): 125-138 (in Chinese)
(谢振平, 金晨, 刘渊. 基于建构主义学习理论的个性化知识推荐模型[J]. 计算机研究与发展, 2018, 55(1): 125-138)
- [55] Liu Yuping, Liu Qi, Wu Runze, et al. Collaborative learning team formation: A cognitive modeling perspective [G] // LNCS 9643; Int Conf on Database Systems for Advanced Applications (DASFAA). Berlin: Springer, 2016; 383-400
- [56] Zhu Haiping, Ni Yifu, Tian Feng, et al. A group-oriented recommendation algorithm based on similarities of personal learning generative networks [J]. IEEE Access, 2018, 6(1): 42729-42739
- [57] Zhu Tianyu, Huang Zhenya, Chen Enhong, et al. Cognitive diagnosis based personalized question recommendation [J]. Chinese Journal of Computers, 2017, 1(1): 176-191 (in Chinese)
(朱天宇, 黄振亚, 陈恩红, 等. 基于认知诊断的个性化试题推荐方法[J]. 计算机学报, 2017, 1(1): 176-191)
- [58] Liu Qi, Huang Zai, Huang Zhenya, et al. Finding similar exercises in online education systems [C] //Proc of the 24th Int Conf on Knowledge Discovery & Data Mining (ACM SIGKDD). New York: ACM, 2018; 1821-1830
- [59] Ma Peijun, Wang Tiantian, Su Xiaohong. Automatic grading of student programs based on program understanding [J]. Journal of Computer Research and Development, 2009, 46(7): 1136-1142 (in Chinese)
(马培军, 王甜甜, 苏小红. 基于程序理解的编程题自动评分方法[J]. 计算机研究与发展, 2009, 46(7): 1136-1142)
- [60] Gao Huijian, Xin Tao, Li Feng. Equivalence of anchorless tests with the same Q matrix based on RSM [J]. Journal of Psychological Science, 2011, 34(4): 957-964 (in Chinese)
(高慧健, 辛涛, 李峰. 基于 RSM 对 Q 矩阵相同的无锚题测验的等值[J]. 心理科学, 2011, 34(4): 957-964)
- [61] Cai Yan, Tu Dongbo, Ding Shuliang. Theorems and methods of a complete Q matrix with attribute hierarchies under restricted Q-Matrix design [J]. Frontiers in Psychology, 2018, 9(1): 1413-1420
- [62] Wang Chao, Liu Qi, Chen Enhong, et al. The rapid calculation method of DINA model for large scale cognitive diagnosis [J]. Acta Electronica Sinica, 2017, 46(5): 1047-1055 (in Chinese)
(王超, 刘淇, 陈恩红, 等. 面向大规模认知诊断的 DINA 模型快速计算方法研究[J]. 电子学报, 2017, 46(5): 1047-1055)
- [63] Dhanani A, Lee S Y, Phothilimthana P M, et al. A comparison of error metrics for learning model parameters in Bayesian knowledge tracing [C] //Proc of the 7th Int Conf on Educational Data Mining (EDM). Worcester, MA: IEDMS, 2014; 153-154
- [64] Pelánek R. A brief overview of metrics for evaluation of student models [C] //Proc of the 7th Int Conf on Educational Data Mining (EDM). Worcester, MA: IEDMS, 2014; 151-152
- [65] Sammaknejad N, Zhao Yujia, Huang Biao. A review of the expectation maximization algorithm in data-driven process identification [J]. Journal of Process Control, 2019, 73(1): 123-136
- [66] Cruyff M J L F, Böckenholt U, Van D H P G M, et al. A review of regression procedures for randomized response data, including univariate and multivariate logistic regression, the proportional odds model and item response model, and self-protective responses [G] //Handbook of Statistics. Amsterdam: Elsevier, 2016; 287-315
- [67] Borji A, Cheng Mingming, Jiang Huaizu, et al. Salient object detection: A benchmark [J]. IEEE Transactions on Image Processing (TIP), 2015, 24(12): 5706-5722
- [68] Willmott C J, Matsuura K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance [J]. Climate Research, 2005, 30(1): 79-82
- [69] Pardos Z A, Baker R S J D, San P M O C Z, et al. Affective states and state tests: Investigating how affect throughout the school year predicts end of year learning outcomes [C] // Proc of the 3rd Int Conf on Learning Analytics and Knowledge. New York: ACM, 2014; 117-124
- [70] Baum L E, Petrie T. Statistical inference for probabilistic functions of finite state Markov chains [J]. The Annals of Mathematical Statistics, 1966, 37(6): 1554-1563
- [71] Fancsali S, Nixon T, Ritter S. Optimal and worst-case performance of mastery learning assessment with Bayesian knowledge tracing [C] //Proc of the 6th Int Conf on Educational Data Mining. Worcester, MA: IEDMS, 2013; 35-42

- [72] Jiang Zhuoyan, Zhang Yan, Li Xiaoming. Learning behavior analysis and prediction based on MOOC data [J]. *Journal of Computer Research and Development*, 2015, 52(3): 614-628 (in Chinese)
(蒋卓轩, 张岩, 李晓明. 基于 MOOC 数据的学习行为分析与预测[J]. *计算机研究与发展*, 2015, 52(3): 614-628)
- [73] Li Haojun, Yang Lin, Zhang Pengwei. Method of online learning resource recommendation based on multi-objective optimization strategy [J]. *Pattern Recognition and Artificial Intelligence*, 2019, 1(4): 306-316 (in Chinese)
(李浩君, 杨琳, 张鹏威. 基于多目标优化策略的在线学习资源推荐方法[J]. *模式识别与人工智能*, 2019, 1(4): 306-316)
- [74] Nelimarkka M, Ghorri M. The effect of variations of prior on knowledge tracing [C] //Proc of the 7th Int Conf on Educational Data Mining (EDM). Worcester, MA: IEDMS, 2014: 146-150
- [75] Beck J E, Chang Kai Min. Identifiability: A fundamental problem of student modeling [G] //LNCS 4511: Proc of the 8th Int Conf on User Modeling. Berlin: Springer, 2007: 137-146
- [76] Ritter S, Harris T K, Nixon T, et al. Reducing the knowledge tracing space [C] //Proc of the 2nd Int Conf on Educational Data Mining (EDM). Worcester, MA: IEDMS, 2009: 151-160
- [77] Van D S B. Properties of the Bayesian knowledge tracing model [J]. *Journal of Educational Data Mining*, 2013, 5(2): 1-10
- [78] Lee J I, Brunskill E. The impact on individualizing student models on necessary practice opportunities [J]. *British Journal of Dermatology*, 2012, 169(4): 910-915
- [79] Huang Yun, González-Brenes J P, Kumar R, et al. A framework for multifaceted evaluation of student models [C] //Proc of the 8th Int Conf on Educational Data Mining (EDM). Worcester, MA: IEDMS, 2015: 203-210
- [80] Doroudi S, Brunskill E. The misidentified identifiability problem of Bayesian knowledge tracing [C] //Proc of 10th Int Conf on Educational Data Mining (EDM). Worcester, MA: IEDMS, 2017: 143-149
- [81] Gweon G H, Lee H S, Dorsey C, et al. Tracking student progress in a game-like learning environment with a monte carlo Bayesian knowledge tracing model [C] //Proc of the 5th Int Conf on Learning Analytics And Knowledge (LAK). New York: ACM, 2015: 166-170
- [82] Coetzee D. Choosing sample size for knowledge tracing models [C] //Proc of the 7th Int Conf on Educational Data Mining (EDM). Worcester, MA: IEDMS, 2014: 117-121
- [83] Pardos Z, Heffernan N T. Navigating the parameter space of Bayesian knowledge tracing models: Visualizations of the convergence of the expectation maximization algorithm [C] //Proc of the 3rd Int Conf on Educational Data Mining (EDM). Worcester, MA: IEDMS, 2010: 161-170
- [84] Rosenberg-Kima R B, Pardos Z. Is this data for real? [C] //Proc of the 7th Int Conf on Educational Data Mining (EDM). Worcester, MA: IEDMS, 2014: 141-145
- [85] Machardy Z, Pardos Z A. Toward the evaluation of educational videos using Bayesian knowledge tracing and big data [C] //Proc of the 2nd ACM Conf on Learning@Scale (L@S). New York: ACM, 2015: 347-350
- [86] Xu Yanbo, Chang Kaimin, Yuan Yueran, et al. EEG helps knowledge tracing! [C] //Proc of the 7th Int Conf on Educational Data Mining (EDM). Worcester, MA: IEDMS, 2014: 43-48
- [87] Spaulding S, Breazeal C. Affect and inference in Bayesian knowledge tracing with a robot tutor [C] //Proc of the 10th Annual ACM/IEEE Int Conf on Human-Robot Interaction Extended Abstracts. New York: ACM, 2015: 219-220
- [88] Chen Lin, Min Chi. Intervention-BKT: Incorporating instructional interventions into Bayesian knowledge tracing [C] //Proc of Int Conf on Intelligent Tutoring Systems (ITS). Berlin: Springer, 2016: 208-218
- [89] González-Brenes J, Huang Yun, Brusilovsky P. General features in knowledge tracing to model multiple subskills, temporal item response theory, and expert knowledge [C] //Proc of the 7th Int Conf on Educational Data Mining (EDM). Worcester, MA: IEDMS, 2014: 84-91
- [90] Kaser T, Klingler S, Schwing A G, et al. Beyond knowledge tracing: Modeling skill topologies with Bayesian networks [C] //Proc of Int Conf on Intelligent Tutoring Systems (ITS). Berlin: Springer, 2014: 188-198
- [91] Khajah M M, Huang Yun, González-Brenes J P, et al. Integrating knowledge tracing and item response theory: A tale of two frameworks [C] //Proc of the 7th Int Conf on Educational Data Mining (EDM). Worcester, MA: IEDMS, 2014: 7-15
- [92] Schultz S, Arroyo I. Tracing knowledge and engagement in parallel in an intelligent tutoring system [C] //Proc of the 7th Int Conf on Educational Data Mining (EDM). Worcester, MA: IEDMS, 2014: 312-315
- [93] Khajah M, Wing R, Lindsey R, et al. Integrating latent-factor and knowledge-tracing models to predict individual differences in learning [C] //Proc of the 7th Int Conf on Educational Data Mining (EDM). Worcester, MA: IEDMS, 2014: 99-106
- [94] Tong Wei, Wang Fei, Liu Qi, et al. Data driven prediction for the difficulty of mathematical items [J]. *Journal of Computer Research and Development*, 2019, 56(5): 1007-1019 (in Chinese)
(佟威, 汪飞, 刘淇, 等. 数据驱动的数学试题难度预测[J]. *计算机研究与发展*, 2019, 56(5): 1007-1019)
- [95] Lindsey R V, Khajah M, Mozer M C. Automatic discovery of cognitive skills to improve the prediction of student learning [C] //Proc of the 27th Int Conf on Neural Information Processing Systems (NeurIPS). Cambridge, MA: MIT, 2014: 1386-1394

- [96] Johns J, Woolf B P. A dynamic mixture model to detect student motivation and proficiency [C] //Proc of the 21st National Conf on Artificial Intelligence and the 18th Innovative Applications of Artificial Intelligence. Menlo Park, CA: AAAI, 2006: 163-168
- [97] Chang K M, Beck J, Mostow J, et al. A Bayes Net Toolkit for student modeling in intelligent tutoring systems [C] //Proc of Int Conf on Intelligent Tutoring Systems (ITS). Berlin: Springer, 2006: 104-113
- [98] Rodríguez P, Bautista M A, Gonzalez J, et al. Beyond one-hot encoding: Lower dimensional target embedding [J]. *Image and Vision Computing*, 2018, 75(1): 21-31
- [99] Cambareri V, Mangia M, Pareschi F, et al. A case study in low-complexity ecg signal encoding: How compressing is compressed sensing? [J]. *IEEE Signal Processing Letters*, 2015, 22(10): 1743-1747
- [100] Tang Junqi, Egiazarian K, Davies M. The limitation and practical acceleration of stochastic gradient algorithms in inverse problems [C] //Proc of Int Conf on Acoustics, Speech and Signal Processing (ICASSP). Piscataway, NJ: IEEE, 2019: 7680-7684
- [101] Chen Penghe, Lu Yu, Zheng V W, et al. Prerequisite-driven deep knowledge tracing [C] //Proc of IEEE Int Conf on Data Mining (ICDM). Piscataway, NJ: IEEE, 2018: 39-48
- [102] Nakagawa H, Iwasawa Y, Matsuo Y. End-to-end deep knowledge tracing by learning binary question-embedding [C] //Proc of IEEE Int Conf on Data Mining Workshops (ICDMW). Piscataway, NJ: IEEE, 2018: 334-342
- [103] Minn S, Yi Yu, Desmarais M C, et al. Deep knowledge tracing and dynamic student classification for knowledge tracing [C] //Proc of IEEE Int Conf on Data Mining (ICDM). Piscataway, NJ: IEEE, 2018: 1182-1187
- [104] Cheung L P, Yang Haiqin. Heterogeneous features integration in deep knowledge tracing [C] //Proc of Int Conf on Neural Information Processing (NeurIPS). Berlin: Springer, 2017: 653-662
- [105] Yang Haiqin, Cheung L P. Implicit heterogeneous features embedding in Deep Knowledge Tracing [J]. *Cognitive Computation*, 2018, 10(1): 3-14
- [106] Huang Yun, Guerra J, Brusilovsky P. A data-driven framework of modeling skill combinations for Deeper Knowledge Tracing [C] //Proc of the 9th Int Conf on Educational Data Mining (EDM). Worcester, MA: IEDMS, 2016: 593-594
- [107] Chen Yuying, Liu Qi, Huang Zhenya, et al. Tracking knowledge proficiency of students with educational priors [C] //Proc of ACM Conf on Information and Knowledge Management (CIKM). New York: ACM, 2017: 989-998
- [108] Zhang Jiani, Shi Xingjian, King I, et al. Dynamic key-value memory networks for knowledge tracing [C] //Proc of the 26th Int Conf on World Wide Web (WWW). New York: ACM, 2017: 765-774
- [109] Pardos Z A, Bergner Y, Seaton D T, et al. Adapting Bayesian knowledge tracing to a massive open online course in edx [C] //Proc of the 6th Int Conf on Educational Data Mining (EDM). Worcester, MA: IEDMS, 2013: 137-144
- [110] Sklavakis D, Refanidis I. The MATHESES Algebra Tutor: Web-based expert tutoring via deep model tracing [C] //Proc of Artificial Intelligence in Education: Building Learning Systems that Care: From Knowledge Representation to Affective Modelling. Ohmsha: IOS, 2009: 795-795
- [111] Baker R S J, Pardos Z A, Gowda S M, et al. Ensembling predictions of student knowledge within intelligent tutoring systems [C] //Proc of Int Conf on User Modeling, Adaptation, and Personalization (UMAP). Berlin: Springer, 2011: 13-24
- [112] Cai Yali, Niu Zhendong, Wang Yingwang, et al. Learning trend analysis and prediction based on knowledge tracing and regression analysis [G] //LNCS 9052: Int Conf on Database Systems for Advanced Applications (DASFAA). Berlin: Springer, 2015: 29-41
- [113] Wang Yutao, Heffernan N. T. Towards modeling forgetting and relearning in ITS: Preliminary analysis of ARRS data [C] //Proc of the 4th Int Conf on Educational Data Mining (EDM). Worcester, MA: IEDMS, 2011: 351-352
- [114] Baker R S, Corbett A T, Koedinger K R, et al. Off-task behavior in the cognitive tutor classroom: when students game the system [C] //Proc of the SIGCHI Conf on Human Factors in Computing Systems (CHI). New York: ACM, 2004: 383-390
- [115] Schultz S E, Arroyo I. Expanding knowledge tracing to prediction of gaming behaviors [C] //Proc of Int Workshop on Approaching Twenty Years of Knowledge Tracing. Worcester, MA: IEDMS, 2014: 157-158
- [116] Agrawal S, Lalwani A. Analysing problem sequencing strategies based on revised bloom's taxonomy using deep knowledge tracing [C] //Proc of Int Conf on Intelligent Tutoring Systems (ITS). Berlin: Springer, 2018: 407-410
- [117] Adams N E. Bloom's taxonomy of cognitive learning objectives [J]. *Journal of the Medical Library Association*, 2015, 103(3): 152-162
- [118] Kasurinen J, Nikula U. Estimating programming knowledge with Bayesian knowledge tracing [C] //Proc of the 14th Annual ACM SIGCSE Conf on Innovation and Technology in Computer Science Education. New York: ACM, 2009: 313-317
- [119] Zheng Qinghua, Dong Bo, Qian Buyue, et al. The state of the art and future tendency of smart education [J]. *Journal of Computer Research and Development*, 2019, 56(1): 213-228 (in Chinese)
(郑庆华, 董博, 钱步月, 等. 智慧教育研究现状与发展趋势 [J]. *计算机研究与发展*, 2019, 56(1): 213-228)

- [120] Li Manli, Huang Zhenzhong. Exploration of educational empirical research based on big data from MOOCs platforms [J]. Chinese Science Bulletin, 2015, 1(5): 570-580 (in Chinese)

(李曼丽, 黄振中. MOOCs 平台大数据的教育实证[J]. 科学通报, 2015, 1(5): 570-580)

- [121] Liu Qi, Chen Enhong, Zhu Tianyu, et al. Research on educational data mining for online intelligent learning [J] Pattern Recognition and Artificial Intelligence, 2017, 31(1): 77-90 (in Chinese)

(刘淇, 陈恩红, 朱天宇, 等. 面向在线智慧学习的教育数据挖掘技术研究[J]. 模式识别与人工智能, 2017, 31(1): 77-90)

- [122] Khajah M, Lindsey R V, Mozer M C. How deep is knowledge tracing? [C] //Proc of the 9th Int Conf on Educational Data Mining (EDM). Worcester, MA: IEDMS, 2016: 94-101



Hu Xuegang, born in 1961. PhD, professor. His main research interests include data mining and knowledge engineering.



Liu Fei, born in 1994. PhD candidate. Her main research interests include educational data mining.



Bu Chenyang, PhD, assistant professor, master supervisor. His main research interests include evolutionary algorithms and their applications in areas such as knowledge graphs, educational data mining, and power systems.

《计算机研究与发展》征订启事

《计算机研究与发展》(Journal of Computer Research and Development)是中国科学院计算技术研究所和中国计算机学会联合主办、科学出版社出版的学术性刊物,中国计算机学会会刊。主要刊登计算机科学技术领域高水平的学术论文、最新科研成果和重大应用成果。读者对象为从事计算机研究与开发的研究人员、工程技术人员、各大专院校计算机相关专业的师生以及高新企业研发人员等。

《计算机研究与发展》于1958年创刊,是我国第一个计算机刊物,现已成为我国计算机领域权威性的学术期刊之一。并历次被评为我国计算机类核心期刊,多次被评为“中国百种杰出学术期刊”。此外,还被《中国学术期刊文摘》、《中国科学引文索引》、“中国科学引文数据库”、“中国科技论文统计源数据库”、美国工程索引(EI)检索系统、日本《科学技术文献速报》、俄罗斯《文摘杂志》、英国《科学文摘》(SA)等国内外重要检索机构收录。

国内邮发代号:2-654;国外发行代号:M603

国内统一连续出版物号:CN11-1777/TP

国际标准连续出版物号:ISSN1000-1239

联系方式:

100190 北京中关村科学院南路6号《计算机研究与发展》编辑部

电话: +86(10)62620696(兼传真); +86(10)62600350

Email: crad@ict.ac.cn

http://crad.ict.ac.cn