

LFKT: 学习与遗忘融合的深度知识追踪模型*

李晓光¹, 魏思齐¹, 张昕¹, 杜岳峰¹, 于戈²

¹(辽宁大学 信息学院, 辽宁 沈阳 110036)

²(东北大学 计算机科学与工程学院, 辽宁 沈阳 110163)

通讯作者: 张昕, E-mail: zhangxin@lnu.edu.cn



摘要: 知识追踪任务旨在根据学生历史学习行为实时追踪学生知识水平变化, 并且预测学生在未来学习表现。在学生学习过程中, 学习行为与遗忘行为相互交织, 学生的遗忘行为对知识追踪影响很大。为了准确建模知识追踪中学习与遗忘行为, 提出一种兼顾学习与遗忘行为的深度知识追踪模型 LFKT (learning and forgetting behavior modeling for knowledge tracing)。LFKT 模型综合考虑了 4 个影响知识遗忘因素, 包括学生重复学习知识点的间隔时间、重复学习知识点的次数、顺序学习间隔时间以及学生对于知识点的掌握程度。结合遗忘因素, LFKT 采用深度神经网络, 利用学生答题结果作为知识追踪过程中知识掌握程度的间接反馈, 建模融合学习与遗忘行为的知识追踪模型。通过在真实在线教育数据集上的实验, 与当前知识追踪模型相比, LFKT 可以更好地追踪学生知识掌握状态, 并具有较好的预测性能。

关键词: 智慧教育; 知识追踪; 深度神经网络; 学习行为; 遗忘行为

中图法分类号: TP181

中文引用格式: 李晓光, 魏思齐, 张昕, 杜岳峰, 于戈. LFKT: 学习与遗忘融合的深度知识追踪模型. 软件学报, 2021, 32(3): 818-830. <http://www.jos.org.cn/1000-9825/6185.htm>

英文引用格式: Li XG, Wei SQ, Zhang X, Du YF, Yu G. LFKT: Deep knowledge tracing model with learning and forgetting behavior merging. Ruan Jian Xue Bao/Journal of Software, 2021, 32(3): 818-830 (in Chinese). <http://www.jos.org.cn/1000-9825/6185.htm>

LFKT: Deep Knowledge Tracing Model with Learning and Forgetting Behavior Merging

LI Xiao-Guang¹, WEI Si-Qi¹, ZHANG Xin¹, DU Yue-Feng¹, YU Ge²

¹(College of Information, Liaoning University, Shenyang 110036, China)

²(School of Computer Science and Engineering, Northeastern University, Shenyang 110163, China)

Abstract: The knowledge tracing task is designed to track changes of students' knowledge in real time based on their historical learning behaviors and to predict their future performance in learning. In the learning process, learning behaviors are intertwined with forgetting behaviors, and students' forgetting behaviors have a great impact on knowledge tracing. In order to accurately model the learning and forgetting behaviors in knowledge tracing, a deep knowledge tracing model LFKT (learning and forgetting behavior modeling for knowledge tracing) that combines learning and forgetting behaviors is proposed in this study. To model such two behaviors, the LFKT model takes into account four factors that affect knowledge forgetting, including the interval between students' repeated learning of knowledge points, the number of repeated learning of knowledge points, the interval between sequential learning, and the understanding degree of knowledge points. The model uses a deep neural network to predict knowledge status with indirect feedbacks on students' understanding of knowledge according to students' answers. With the experiments on the real datasets of online education, LFKT shows better performance of knowledge tracing and prediction in comparison with the traditional approaches.

* 基金项目: 国家自然科学基金(U1811261)

Foundation item: National Natural Science Foundation of China (U1811261)

本文由“支撑人工智能的数据管理与分析技术”专刊特约编辑陈雷教授、王宏志教授、童咏昕教授、高宏教授推荐。

收稿时间: 2020-08-23; 修改时间: 2020-09-03; 采用时间: 2020-11-06; jos 在线出版时间: 2021-01-21

Key words: intelligent education; knowledge tracing; deep neural network; learning behavior; forgetting behavior

在线教育系统提倡因材施教,即根据学生的知识水平为其推荐合适的学习资源^[1].学生的知识水平受其学习阶段和认知能力的影响,在学习过程中不断变化.实时追踪学生知识水平,对于个性化在线教育至关重要^[2,3].知识追踪(knowledge tracing,简称KT)任务主要包括:(1)通过学生的学习历史实时追踪其知识水平变化;(2)根据学生的知识水平预测其在未来学习中的表现.

传统的知识追踪技术主要有基于隐马尔可夫模型(HMM)^[4]的贝叶斯知识追踪(BKT)^[5].近年来,深度学习在自然语言处理^[6]及模式识别^[7]等任务上的表现优于传统模型.在知识追踪研究领域,亦提出了大量基于深度学习的知识追踪模型.基于循环神经网络的深度知识追踪(DKT)模型^[8]采用循环神经网络的隐藏向量表示学生的知识状态,并且据此预测学生的表现.动态键-值对记忆网络(DKVMN)^[9]借鉴了记忆网络(memory network)^[10]的思想,用值矩阵表示学生对于各个知识点的掌握程度,并以此预测学生表现,提升预测准确度.但是以上知识追踪方法都忽略了学生在学习过程中的遗忘行为对知识水平的影响.在教育心理学领域已经有很多学者认识到人类的遗忘行为,并且探索影响人类遗忘的因素.教育心理学理论艾宾浩斯遗忘曲线理论^[11,12]提出:学生会遗忘所学知识,遗忘带来的影响是知识掌握程度的下降.学生重复学习知识点的次数与距离上次学习知识点的时间间隔会影响学生的遗忘程度.教育心理学理论记忆痕迹衰退说^[13]认为:遗忘是记忆痕迹的衰退引起的,消退随时间的推移自动发生,原始的痕迹越深,则遗忘的程度越低.目前,在知识追踪领域,只有少数的研究考虑了学生的遗忘行为.BKT的扩展^[14,15]考虑了艾宾浩斯遗忘曲线提到的影响遗忘程度的两个因素:重复学习知识点的次数和间隔时间.DKT的扩展^[16]考虑了学生顺序学习的间隔时间,该研究认为:学科中的各个知识点是相通的,持续的学习会降低学生对于知识点的遗忘程度.然而,以上的研究仅仅考虑了部分影响学生遗忘的信息,忽略了学生原本知识水平对于遗忘程度的影响,学生对于其掌握的知识点和没有掌握的知识点的遗忘程度不同.与此同时,由于BKT算法与DKT算法的局限性,以往涉及遗忘的研究或者仅考虑了学生对于单个知识点的遗忘程度,或者仅考虑了学生对于整个知识空间的遗忘程度,没有建模学生对知识空间中的各个知识点的遗忘情况.

针对于此,本文提出一种兼顾学习和遗忘的深度知识追踪模型LFKT(learning and forgetting behavior modeling for knowledge tracing).LFKT拟合了学生的学习与遗忘行为,实时更新、输出知识点掌握程度,并以此预测其未来表现.本文主要创新和贡献有:

- (1) 结合教育心理学,知识追踪模型LFKT考虑了4个影响知识遗忘的因素:学生重复学习知识点的间隔时间、重复学习知识点的次数、顺序学习间隔时间以及学生对于知识点的掌握程度.LFKT根据以上信息建模遗忘行为,拟合学生因遗忘行为导致的知识掌握程度变化;
- (2) 基于深度神经网络技术,LFKT设计了一个基于RNN和记忆神经网络的知识追踪神经网络.该网络包括注意力层、遗忘层、学习层、预测层和知识水平输出层.其中:遗忘层采用全连接网络计算记忆擦除向量和记忆更新向量,用以拟合遗忘行为;学习层采用LSTM网络,并利用学生在学习结束时的答题结果作为知识掌握程度的间接反馈.根据记忆擦除向量和记忆更新向量,获得经过遗忘后的知识掌握程度的中间嵌入,并以此作为学习层的输入,进而获得遗忘与学习相融合的知识掌握程度嵌入;
- (3) 通过在两个真实在线教育数据集上的实验验证结果表明:LFKT可以有效地建模学生的学习行为与遗忘行为,实时追踪学生的知识水平,并且LFKT模型的预测性能优于现有模型.

本文第1节介绍知识追踪相关工作.第2节给出相关概念和符号定义,并提出问题.第3节详细介绍LFKT模型结构以及训练方法.第4节给出LFKT对比实验结果和分析.最后总结全文.

1 相关工作

知识追踪大致可以分为基于隐马尔可夫模型的知识追踪与基于深度学习的知识追踪.

贝叶斯知识追踪是典型的基于隐马尔可夫模型实现知识追踪目标的模型之一.BKT将学生对某个知识点的潜在知识状态建模为一组二进制变量,每一个变量代表学生对于特定知识点“掌握”或“没掌握”,根据隐马尔可夫模型更新学生的知识状态,进而更新其在未来学习中的表现.在BKT的基础上,通过考虑其他因素,提出了

很多扩展方案,例如:Pardos 等人的研究^[17]引入了习题难度对预测学生表现的影响;Baker 等人的研究^[18]引入了学生猜测和失误对预测学生表现的影响;Khajah 等人的研究^[19]将人类的认知因素扩展到 BKT 中,从而提高预测精度.基于隐马尔可夫的知识追踪模型将学生对各个知识点的掌握程度分别建模,忽略各个知识点之间的关系.BKT 模型简单地将学生对于各个知识点的掌握程度分为“掌握”与“未掌握”,忽略了中间情况.

深度知识追踪是深度学习模型循环神经网络在知识追踪任务的首次尝试,将学生的学习历史作为输入,使用 RNN 隐藏状态向量来表示学生的知识水平,并且基于此预测学生未来学习表现.DKT 模型无法建模学生对于各个知识点的掌握程度,仅仅可以建模学生的整个知识水平.Chen 等人的研究^[20]考虑了知识点之间的先验关系,提升了 DKT 的预测性能.Su 等人的研究^[21]将习题文本信息和学生的学习历史作为循环神经网络的输入,利用 RNN 的隐向量建模学生的知识水平,取得了较好的预测效果,但是依旧无法建模学生对于各个知识点的掌握程度.DKVMN 模型借鉴了记忆网络的思想,用值矩阵建模学生对于各个知识点的掌握程度,考察习题与各个知识点之间的关系,追踪学生对于各个知识点的掌握程度.Abdelrahman 等人的研究^[22]利用注意力机制,着重考察学生作答相似习题时的答题历史,改进了 DKVMN.Sun 等人的研究^[23]将学生答题时的行为特征扩展到 DKVMN,从而取得更好的预测效果.Liu 等人的研究^[24]利用习题的文本信息与学生的学习历史进行知识追踪.Nakagawa 等人的研究^[25]用知识图建模各个知识点之间的关系,取得了良好的预测效果.Minn 等人的研究^[26]根据学生的知识水平对学生进行分类,再对每一类学生进行知识追踪.Cheng 等人的研究^[27]通过深度学习扩展了传统方法项目反映理论.Shen 等人的研究^[28]考虑了每个学生的个性化差异,取得了很好的效果.Ai 等人的研究^[29]考虑了知识点之间的包含关系,重新设计了记忆矩阵,改进了 DKVMN,取得了良好的效果.关于学生遗忘行为对知识掌握状态的影响,Qiu 等人^[15]考虑了学生距离上一次重复学习知识点的间隔时间,将新一天的标记加入到 BKT 中,建模先前学习之后一天的遗忘行为,但是无法对较短时间的遗忘行为进行建模.Khajah 等人的研究^[14]应用学生重复学习知识点的次数估计遗忘的概率,改进了 BKT,提升了预测准确度.Nagatani 等人的研究^[16]考虑了重复学习知识点的次数、距离上一次学习知识点的间隔时间和距离上一次学习的间隔时间,改进了 DKT 模型,但忽略了学生原本对于知识点掌握程度对学生遗忘程度影响.

总的来说,当前的研究可以一定程度上追踪学生对于各个知识点的掌握程度,但或者忽略了学生遗忘行为,默认学生在学习间隔前后知识水平一致;或者对于遗忘行为的建模不够全面.本文所提出的 LFKT 模型综合考察了影响学生遗忘程度的因素,将学生的练习结果作为学生知识掌握程度的间接反馈,建模学生的学习与遗忘行为,实时追踪学生对于各个知识点的掌握程度,预测学生未来学习表现.

2 问题定义

这里,我们定义 S 为学生集合, K 为知识点集合, E 为习题集合. 每个学生单独进行学习,互不影响. 学生 s 的答题历史 $H_s = [(e_1, r_1), (e_2, r_2), \dots, (e_t, r_t)]$, 其中: e_t 为学生 t 时刻所做的习题, r_t 为答题结果, $r_t=1$ 表示答题正确, $r_t=0$ 表示答题错误. $k_t \subseteq K$ 为习题 e_t 所涉及的知识点集合. 矩阵 $M^k (d_k \times |K|)$ 为整个知识空间中 $|K|$ 个知识点嵌入表示, 其中, d_k 维列向量为一个知识点的嵌入表示. 矩阵 $M_{t-1}^v (d_v \times |K|)$ 为学生在 $t-1$ 时刻的知识空间中知识点嵌入. 用 $|K|$ 维向量 $value_{t-1}$ 表示 $t-1$ 时刻学习结束时学生对于各个知识点的掌握程度, 其中, 向量每一维的值介于 $(0, 1)$ 之间: 值为 0, 表示学生对于该知识点没有掌握; 值为 1, 表示学生对于该知识点已经完全掌握. 学生在学习过程中会遗忘未复习的知识点, 同时, 学生在两次学习的间隔时间也会遗忘所学知识, 因此, $t-1$ 时刻学习结束时的知识水平与 t 时刻学习开始时的知识水平不尽相同. 本文用矩阵 $M_t^{fv} (d_v \times |K|)$ 建模学生在 t 时刻开始学习时的知识掌握状态. 矩阵 M_t^{fv} 与 M_{t-1}^v 形状相同, M_t^{fv} 是由 M_{t-1}^v 经过遗忘处理后得到. t 时刻学习结束, 系统获得学生答题结果, LFKT 模型根据答题结果将开始学习时知识掌握嵌入矩阵 M_t^{fv} 更新为学习结束时知识掌握嵌入矩阵 M_t^v , 以此建模学习行为. 预测学生在下一个候选习题 e_{t+1} 上的表现时, 由于 t 时刻学习结束到 $t+1$ 时刻开始答题之间存在时间间隔, 学生在时间间隔内的遗忘行为会影响知识掌握状态, 因此需要根据影响知识遗忘的因素将 M_t^v 更新为 M_{t+1}^{fv} , 从而预测学生答题表现.

基于以上描述,本文将问题定义为:给定每个学生的学习历史记录,实现以下两个目标.

- 跟踪学生的知识状态变化;
- 预测学生在下一个候选习题 e_{t+1} 上的表现.

3 方法描述

3.1 影响知识遗忘的因素

学生如果没有及时复习所学知识,就会产生遗忘,其对知识的掌握程度便会衰减.教育心理学理论艾宾浩斯曲线理论^[11,12]表明,学生对于所学知识点保留率受以下两方面影响:学生重复学习次数与时间间隔.时间间隔可以分为重复学习知识点的时间间隔和顺序学习的时间间隔.除此之外,教育心理学理论记忆痕迹衰退说^[13]认为,学生对于知识点的掌握状态也影响着学生的遗忘程度.因此,本文考虑了以下 4 个影响知识遗忘的因素.

- RT(repeated time interval):距离上次学习相同知识点的时间间隔;
- ST(sequence time interval):距离上次学习的时间间隔;
- LT(repeated learn times):重复学习知识点的次数;
- KL(past knowledge level):学生原本对于该知识点的掌握程度.

RT,ST 和 LT 这 3 个标量组合在一起得到向量 $C_i(i)=[RT_i(i),ST_i(i),LT_i(i)]$,表示影响学生对于知识点 i 遗忘程度的前 3 个因素.每个知识点所对应的向量 $C_i(i)$ 组成矩阵 $C_i(d_c \times |K|)$.本文用向量 $M_{t-1}^V(i)$ 度量学生原本对于知识点 i 的掌握程度,并且将其作为影响学生对于知识点遗忘程度的第 4 个因素(KL).将矩阵 C_i 与 M_{t-1}^V 组合在一起,得到矩阵 $F_t=[M_{t-1}^V, C_i]$,表示 4 个影响知识遗忘的因素.

3.2 LFKT模型

本文提出了一种融合学习与遗忘的深度知识追踪模型 LFKT,如图 1 所示.

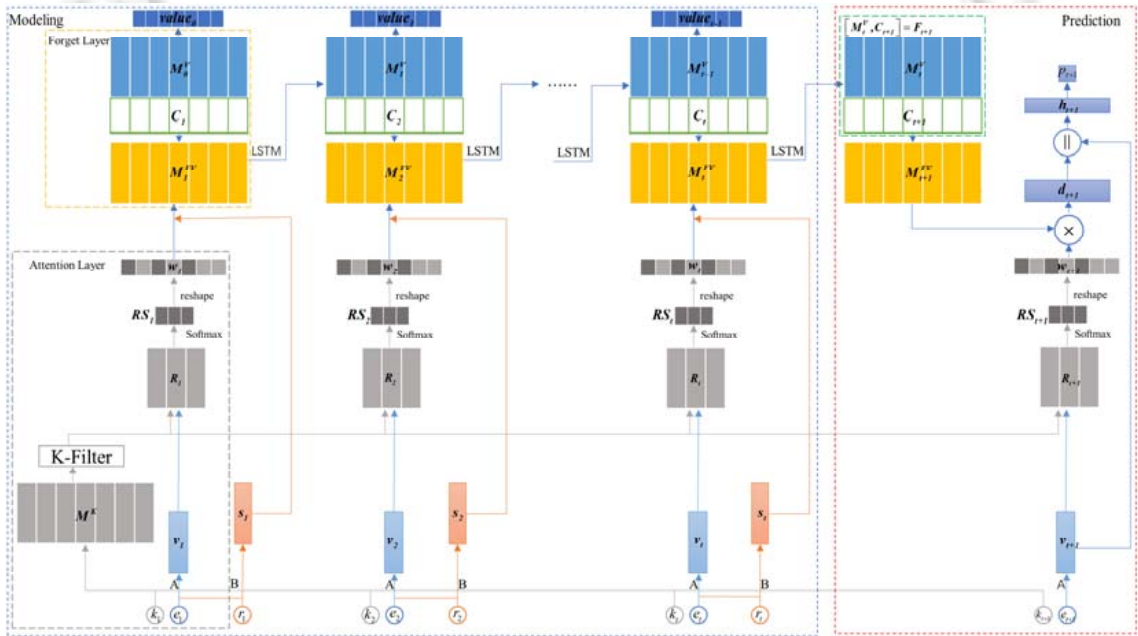


Fig.1 LFKT model

图 1 LFKT 模型

LFKT 模型分为注意力层(attention layer)、遗忘层(forget layer)、预测层(prediction layer)、学习层(learn layer)

以及知识水平输出层(knowledge level output layer):注意力层以习题 e_t 以及习题所涵盖的知识点集合 k_t 作为输入,计算习题与各个知识点的知识相关权重;遗忘层根据第 3.1 节中提出的影响知识遗忘的 4 个因素将学生上一时刻学习结束时知识掌握嵌入矩阵 M_{t-1}^V 更新为当前时刻学习开始时知识掌握嵌入矩阵 M_t^{FV} , 有效地对学生遗忘行为进行建模;预测层根据 $t+1$ 时刻开始答题时的 M_{t+1}^{FV} 预测学生的答题表现;学习层通过 LSTM 网络将本次学习开始时的 M_t^{FV} 更新为本次学习结束时的 M_t^V , 对学生学习行为进行建模;知识水平输出层以学生本次学习结束时的 M_t^V 作为输入,输出学生知识水平向量 $value_t$, 实时输出学生对于各个知识点的掌握程度.

3.2.1 注意力层

注意力层的作用是计算习题与知识点之间的知识相关权重.注意力层的输入是学生当前练习题目 e_t 以及题目所涉及的知识点集合 k_t .为了将 e_t 映射到连续的向量空间,用 e_t 乘以习题嵌入矩阵 $A(d_k \times |E|)$, 生成一个 d_k 维度的习题嵌入向量 v_t .其中,矩阵 A 中每一个 d_k 维列向量为一道习题的嵌入表示.专家标注了每道习题所涵盖的知识点,存放于集合 k_t 中.本文通过知识点过滤器(K-Filter)滤掉不相关知识点,保留习题涵盖知识点.矩阵 R_t 存储习题涵盖知识点的嵌入向量,其中,矩阵 R_t 中每一个 d_k 维列向量为习题涵盖的一个知识点的嵌入向量.计算习题嵌入向量 v_t 与涵盖知识点嵌入向量 $R_t(i)$ 之间的内积,再计算内积的 *Softmax* 值并存入向量 RS_t 中.向量 RS_t 表示习题 v_t 与习题涵盖知识点之间的知识相关权重,如公式(1)所示:

$$RS_t(i) = \text{Softmax}(v_t^T R_t(i)) \tag{1}$$

$|K|$ 维向量 w_t 为习题与全部知识点之间的知识相关权重向量.由于习题涵盖知识点被知识点过滤器滤出,模型需要将习题涵盖知识点的知识相关权重放入到 w_t 对应的位置上.首先,初始化 $|K|$ 维全零向量 w_t , 即 $w_t \leftarrow [0, \dots, 0]$; 之后,将 RS_t 每一维的权重放入到 w_t 的相应位置上,即 $w_t[k_t[i]] \leftarrow RS_t[i]$, 得到习题与各个知识点的知识相关权重.

3.2.2 遗忘层

遗忘层根据第 3.1 节中介绍的影响知识遗忘的因素 F_t 对上一次学习结束时学生的知识掌握嵌入矩阵 M_{t-1}^V 进行遗忘处理,得到学生本次学习开始时的知识掌握嵌入矩阵 M_t^{FV} .受 LSTM 中遗忘门与输入门的启发,根据影响知识遗忘的因素 F_t 更新 M_{t-1}^V 时,首先要擦除 M_{t-1}^V 中原有的信息,再写入信息.对于遗忘行为建模来说,擦除过程控制学生知识点掌握程度的衰退,写入过程控制学生知识点掌握程度的更新.遗忘层如图 2 所示.

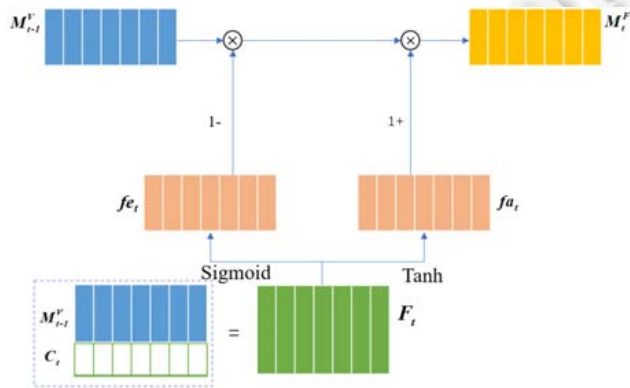


Fig.2 Forget layer

图 2 遗忘层

利用一个带有 *Sigmoid* 激活函数的全连接层将影响学生对知识点 i 遗忘程度的因素 $F_t(i)$ 转换为知识点 i 对应的记忆擦除向量 $fe_t(i)$:

$$fe_t(i) = \text{Sigmoid}(FE^T F_t(i) + b_{fe}) \tag{2}$$

全连接层的权重矩阵 FE 的形状是 $(d_v + d_c) \times d_v$, 全连接层的偏置向量 b_{fe} 为 d_v 维.记忆擦除向量 $fe_t(i)$ 是一个维

度为 d_v 的列向量,向量中的所有值都是在(0,1)范围之内.利用一个带有 Tanh 激活函数的全连接层将影响学生对知识点 i 遗忘程度的因素 $F_i(i)$ 转换为知识点 i 对应的记忆更新向量 $fa_i(i)$:

$$fa_i(i) = \text{Tanh}(\mathbf{FA}^T \mathbf{F}_i(i) + \mathbf{b}_{fa}) \quad (3)$$

全连接层的权重矩阵 \mathbf{FA} 的形状是 $(d_v + d_c) \times d_v$,全连接层的偏置向量 \mathbf{b}_{fa} 为 d_v 维,记忆更新向量 $fa_i(i)$ 是一个维度为 d_v 的列向量.根据记忆擦除向量与记忆更新向量对 \mathbf{M}_{t-1}^V 更新得到 \mathbf{M}_t^{FV} :

$$\mathbf{M}_t^{FV}(i) = \mathbf{M}_{t-1}^V(i)(1 - fe_i(i))(1 + fa_i(i)) \quad (4)$$

遗忘层输出本次学习开始时知识掌握嵌入矩阵 \mathbf{M}_t^{FV} .遗忘层可以为分别建模学生对不同知识点的遗忘过程,根据学生对于不同知识点学习历史的不同,计算学生对于不同知识点的遗忘程度.

3.2.3 学习层

学习层根据学生答题结果追踪学习过程中的知识掌握变化,将学生开始学习时知识掌握嵌入矩阵 \mathbf{M}_t^{FV} 更新为学生学习结束时知识掌握嵌入矩阵 \mathbf{M}_t^V ,建模学习行为.元组 (e_t, r_t) 表示学生在时间 t 的答题结果,为了将元组 (e_t, r_t) 映射到连续的向量空间,元组 (e_t, r_t) 与大小为 $d_v \times 2|E|$ 的答题结果嵌入矩阵 \mathbf{B} 相乘,得到 d_v 维答题结果嵌入向量 s_t ,学习层将答题结果嵌入向量 s_t 和习题对应的知识相关权重向量 w_t 作为输入,利用学生在线教育系统中的练习结果作为学生学习效果的间接反馈,通过 LSTM 网络更新学生学习过程中的知识掌握状态,建模学习行为:

$$\mathbf{M}_t^V(i) = \text{LSTM}(s_t, w_t(i), \mathbf{M}_t^{FV}(i); \theta) \quad (5)$$

3.2.4 预测层

预测层的目的是预测学生在下一个候选习题 e_{t+1} 上的表现.由于学生在两次答题间隔内的遗忘行为会影响其知识掌握状态,预测层根据当前时刻开始答题时的知识掌握嵌入矩阵 \mathbf{M}_{t+1}^{FV} 预测学生正确回答习题 e_{t+1} 的概率.将知识相关权重 w_{t+1} 与当前时刻开始答题时学生知识掌握嵌入 \mathbf{M}_{t+1}^{FV} 加权求和,得到的向量 d_{t+1} 为学生对于习题涵盖知识点的加权掌握程度嵌入向量:

$$d_{t+1} = \sum_{i=1}^K w_{t+1}(i) \mathbf{M}_{t+1}^{FV}(i) \quad (6)$$

学生成功解答习题,不仅与学生对于习题涵盖知识的综合掌握程度有关,还与习题本身有关,所以将向量 d_{t+1} 和 v_{t+1} 连接得到的组合向量 $[d_{t+1}, v_{t+1}]$ 输入到带有 Tanh 激活函数的全连接层,输出向量 h_{t+1} 包含了学生对习题涵盖知识的综合掌握程度和习题本身的特点.其中,矩阵 \mathbf{W}_1 和向量 \mathbf{b}_1 分别表示全连接层的权重与偏置:

$$h_{t+1} = \text{Tanh}(\mathbf{W}_1^T [d_{t+1}, v_{t+1}] + \mathbf{b}_1) \quad (7)$$

最后,将向量 h_{t+1} 输入到一个带有 Sigmoid 激活函数的全连接层中,得到表示学生答对习题的概率 p_{t+1} .其中,矩阵 \mathbf{W}_2 和向量 \mathbf{b}_2 分别表示全连接层的权重与偏置:

$$p_{t+1} = \text{Sigmoid}(\mathbf{W}_2^T h_{t+1} + \mathbf{b}_2) \quad (8)$$

3.2.5 知识水平输出层

知识水平输出层以学生学习结束时知识掌握嵌入矩阵 \mathbf{M}_t^V 为输入,输出表示学生本次学习结束时知识掌握程度的 $|K|$ 维向量 $value_t$,向量的每一维介于(0,1)之间,表示学生对于该知识点的掌握程度.知识水平输出层如图 3 所示.

在预测层中,每一个时间节点 t ,公式(7)、公式(8)根据两种输入预测学生在特定习题 e_t 上的表现:学生对于该习题涵盖知识的综合掌握嵌入向量 d_t 和习题嵌入向量 v_t .因此,如果只是想估计在没有任何习题输入情况下学生对于第 i 个知识点的掌握情况,可以省略习题嵌入 v_t ,同时,让学生知识掌握嵌入矩阵 \mathbf{M}_t^V 的第 i 列 $\mathbf{M}_t^V(i)$ 作为等式的输入.图 3 展示了知识水平输出层的详细过程.具体来说:学习层输出矩阵 \mathbf{M}_t^V 后,为了估计对于第 i 个知识点的掌握程度,构造了权重向量 $\beta_i = (0, \dots, 1, \dots, 0)$,其中,第 i 维的值等于 1,并用公式(9)提取第 i 个知识点的知识掌握嵌入向量 $\mathbf{M}_t^V(i)$,之后,使用公式(10)、公式(11)估计知识掌握水平:

$$\mathbf{M}_i^V(i) = \beta_i \mathbf{M}_i^V \quad (9)$$

$$y_i(i) = \text{Tanh}(\mathbf{W}_1^T [\mathbf{M}_i^V(i), \mathbf{0}] + \mathbf{b}_1) \quad (10)$$

$$\text{value}_i(i) = \text{Sigmoid}(\mathbf{W}_2^T y_i(i) + \mathbf{b}_2) \quad (11)$$

向量 $\mathbf{0}=(0,0,\dots,0)$ 与习题嵌入 \mathbf{v}_i 维度相同,用于补齐向量维度.参数 $\mathbf{W}_1, \mathbf{W}_2, \mathbf{b}_1, \mathbf{b}_2$ 与公式(7)和公式(8)中的完全相同.依次计算知识空间中每一个知识的掌握程度,得到学生知识掌握程度向量 value_i .

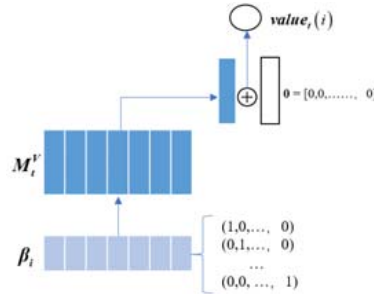


Fig.3 Knowledge level output layer

图3 知识水平输出层

3.2.6 模型优化目标

模型需要训练的参数主要是习题嵌入矩阵 \mathbf{A} 、答题结果嵌入矩阵 \mathbf{B} 、神经网络权重与偏置以及知识点矩阵 \mathbf{M}^K .本文通过最小化模型对于学生答题结果的预测值 p_i 和学生答题的真实结果 r_i 之间的交叉熵损失函数来优化各个参数.损失函数如公式(12)所示,本文使用 Adam 方法优化参数:

$$L = -\sum_i (r_i \log p_i + (1 - r_i) \log(1 - p_i)) \quad (12)$$

4 实验

本文在两个真实在线教育数据集上进行实验,通过对比 LFKT 模型与其他知识追踪模型的预测性能以及知识追踪表现,证明 LFKT 模型的有效性.

4.1 数据集

首先介绍实验所使用的两个真实在线教育数据集.

- ASSISTments2012:该数据集来自于 ASSISTments 在线教育平台.本文删除了学习记录数过少(<3)的学生信息,经过预处理后,数据集包含 45 675 个学生、266 个知识点以及总计 5 818 868 条学习记录.数据集中, user_id 字段表示学生编号, skill_id 字段表示知识点编号, problem_id 字段表示题目编号, correct 字段表示学生真实答题结果, start_time 字段表示学生本次开始学习的时间, end_time 字段表示学生本次学习结束的时间;
- slepemapy.cz:该数据集来自于地理在线教育系统.本文同样删除学习记录数过少(<3)的学生信息,经过预处理后,数据集包含 87 952 个学生、1 458 个知识点以及总计 10 087 305 条学习记录.数据集中, user 字段表示学生编号, place_answered 字段表示学生真实答题结果, response_time 字段表示学生学习时间, place_asked 字段表示问题编号.其中,由于该数据集中每个问题仅仅考察一个知识点,因此 place_asked 也是知识点编号.

4.2 评测指标

本文使用平均 AUC(area under the curve)、平均 ACC(accuracy)和平均 RMSE(root mean squared error)作为评估预测性能的指标.AUC 被定义为 ROC 曲线与下坐标轴围成的面积,50%的 AUC 值表示随机猜测获得的预

测性能,高 AUC 值说明模型具有较高的预测性能.ACC 为准确率,即正确预测结果占全部结果的百分比,高 ACC 值说明模型具有较高的预测性能.RMSE 被定义为预测值与真实值的均方根误差,RMSE 的值越低,说明模型的预测性能越好.对于每一个模型,本文都进行 30 次测试,取平均 AUC 值、平均 ACC 值和平均 RMSE 值.

4.3 对比方法与参数设置

为了评估 LFKT 模型的性能,本文选择贝叶斯知识追踪^[5]、深度学习追踪^[8]和动态键-值对记忆网络^[9]这 3 个经典 KT 模型作为对比方法.对比方法的参数设置如下.

- BKT^[5]:实验设置 4 个主要参数: L_0 表示学生未开始作答时对题目考察知识点的掌握程度, T 为学生作答后对知识点掌握程度从不会到会的概率, G 为学生没掌握该知识点但是答对题目的概率, S 表示学生掌握知识点但是答错题目的概率.BKT 利用以上参数构造一个隐马尔可夫模型,并使用 EM 算法求解参数;
- DKT^[8]:实验按照 Piech 等人^[8]的方法设置超参数.循环神经网络 GRU 隐藏层的大小为 200,批处理大小设置为 30,使用 Adam 作为优化器,学习率为 0.001;
- DKVMN^[9]:本文使用 Zhang 等人^[9]的方法设置超参数.对于 ASSISTments2012 数据集,记忆矩阵列数为 20,隐藏向量的大小为 30.对于 slepemapy.cz 数据集,记忆矩阵列数为 320,隐藏向量大小为 128.批处理大小为 30,使用 Adam 作为优化器,学习率为 0.001.

LFKT 模型:对于 ASSISTments2012 数据集,知识嵌入矩阵 M_i^K 的列数为 266.对于 slepemapy.cz 数据集,知识嵌入矩阵 M_i^K 的列数为 1 458.批处理大小为 30,采用 Adam 优化器,学习率为 0.001.对于知识点嵌入向量的维度 d_k 和学生知识掌握程度嵌入向量的维度 d_v 的超参数设置,本文通过比较模型在测试数据集上的平均 AUC 值进行选取.为了减少参数数量,本文设 $d=d_k=d_v$,测试结果见表 1.ASSISTments2012 数据集中,当 $d=d_k=d_v=32$ 时,平均 AUC 值为 0.7513,高于其他超参数设置情况.在 slepemapy.cz 数据集中,当 $d=d_k=d_v=128$ 时,平均 AUC 值为 0.8032,高于其他超参数设置.通过对比可知:当 $d=d_k=d_v$ 设置过低时,模型的学习能力较低;当 $d=d_k=d_v$ 设置过高时候,模型参数过多,容易造成过拟合.因此,对于 ASSISTments2012 数据集,设 $d_k=d_v=32$;对于 slepemapy.cz 数据集,设 $d_k=d_v=128$.

Table 1 AUC results of different hyperparameters
表 1 不同超参数设置下 AUC 值对比实验结果

Model	ASSIST2012		slepemapy.cz	
	d	AUC	d	AUC
LFKT	8	0.713 2	32	0.785 6
	16	0.742 3	64	0.789 7
	32	0.751 3	128	0.803 2
	64	0.747 7	256	0.793 3

4.4 学生表现预测结果分析

知识追踪的基本任务之一是,预测学生当前对于下一道候选习题的答题表现.通过对比模型的预测结果与学生的真实答题结果,比较模型的预测性能.本文随机地将数据集中 70%的数据作为训练集,另外 20%的数据作为测试集,10%的数据作为验证集,用于超参数的调整与早期停止.

表 2 给出了 LFKT 与 3 个对比方法在两个数据集上的平均 AUC 值、平均 ACC 值和平均 RMSE 值的对比结果.LFKT 在两个数据集上的平均 AUC 值、平均 ACC 值和平均 RMSE 值均优于其他 3 种对比方法.这一结果表明,LFKT 模型在预测学生未来表现方面性能是优于现有模型的.从实验对比结果可以看出,BKT 在两个数据集上的预测性能最低.这说明 BKT 将学生对于某一知识点的潜在知识水平建模为二进制变量是具有局限性的. DKT 利用循环神经网络的潜在向量建模学生整体的知识水平,无法建模学生对于各个知识点的掌握程度,因此 DKT 在两个数据集上的预测性能低于 DKVMN 和 LFKT.DKVMN 与 LFKT 都可以建模学生对于各个知识点的掌握程度,但是 DKVMN 忽略了学生在学习期间的遗忘行为,默认学生对未复习知识点的掌握程度一直

不变,这存在一定的局限性,因此 LFKT 的预测性能强于 DKVMN.

Table 2 Prediction results of models

表 2 模型预测性能对比实验结果

Model	AUC		ACC		RMSE	
	ASSIST 2012	slepemapy.cz	ASSIST 2012	slepemapy.cz	ASSIST 2012	slepemapy.cz
BKT	0.673 2	0.701 7	0.631	0.673	0.440	0.429
DKT	0.723 5	0.787 1	0.679	0.731	0.413	0.407
DKVMN	0.732 3	0.792 1	0.686	0.743	0.404	0.401
LFKT	0.751 3	0.803 2	0.723	0.762	0.387	0.381
LFKT(without forget layer)	0.732 1	0.790 2	0.693	0.747	0.398	0.392

除此之外,本文测试了去掉遗忘层后 LFKT 模型的预测表现.LFKT 模型去掉遗忘层以后,忽略了学生的遗忘行为,默认学生在学习间隔内,知识水平不会发生任何变化,即 $M_t^V = M_{t+1}^{FV}$. 实验结果显示:去掉遗忘层后,LFKT 模型的预测性能较没去掉遗忘层时有所下降.这表明学生的遗忘行为对其知识水平有所影响,本文所提出的影响知识遗忘的因素是十分有效的.

4.5 知识追踪结果分析

知识追踪的另一个基本任务是实时输出学生对于各个知识点的掌握程度.本文进行以下实验验证 LFKT 模型在知识追踪任务上的有效性.

本文截取了数据集 ASSISTments2012 中一位学生在一段时间内的学习记录,并分别使用 LFKT 模型和 DKVMN 模型追踪学生对于 5 个知识点掌握程度的变化,如图 4 和图 5 所示.图中横轴表示学生的学习历史,元组 (k_t, r_t) 中 k_t 表示学生学习的知识点, r_t 表示学生作答情况.纵轴表示模型追踪的 5 个知识点.

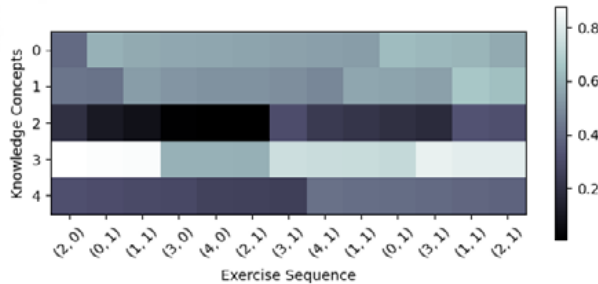


Fig.4 Knowledge level output result of LFKT in ASSISTments2012

图 4 ASSISTments2012 数据集 LFKT 知识水平输出结果

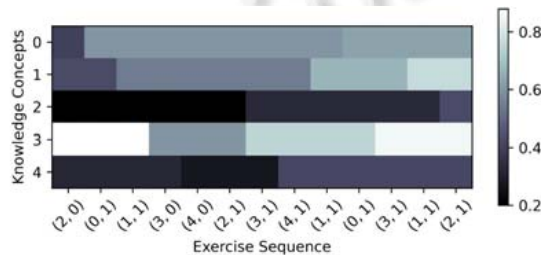


Fig.5 Knowledge level output result of DKVMN in ASSISTments 2012

图 5 ASSISTments 2012 数据集 DKVMN 知识水平输出结果

实验结果显示:在第 2 时刻,学生对知识点 0 答题正确后,DKVMN 和 LFKT 模型对学生知识点 0 掌握程度的追踪结果都有所提高(输出值增加);在第 4 时刻,学生对知识点 3 对应习题答题错误后,DKVMN 模型和 LFKT 模型对学生知识点 3 的掌握程度的追踪结果都有所下降.以上结果说明:DKVMN 模型和 LFKT 模型在获得学

生作答结果以后,都会根据学生的作答结果更新对应知识点的掌握程度,以此建模学生的学习过程.图 4 和图 5 中,学生第 2 时刻学生学习知识点 0 到第 10 时刻第 2 次学习知识点 0 期间,DKVMN 模型并没有更新学生对于知识点 0 的掌握程度,而 LFKT 模型显示学生对于知识点 0 的掌握程度一直在下降.由此可以看出:DKVMN 模型没有考虑学习者在学习期间的遗忘行为,而 LFKT 模型考虑了学习者在学习期间的遗忘行为.以上结果表明:LFKT 模型与 DKVMN 模型都可以建模学生的学习行为,DKVMN 模型无法建模学生的遗忘行为,而 LFKT 模型可以建模学生的遗忘行为,并实时追踪学生对于各个知识点的掌握程度.

本文在 slepemapy.cz 测试数据集中随机抽取 3 名学生在一段时间内的学习记录,并分别使用 LFKT 模型与 DKVMN 模型输出其对 5 个知识点的掌握程度.3 名学生的部分答题序列如下.

- $H_a=[(1,1),(2,1),(3,1),(4,0),(5,0),(6,0),(1,0),(2,1),(3,1),(4,0),(5,1),(6,1)]$;
- $H_b=[(6,1),(5,1),(4,1),(3,1),(2,1),(1,1),(2,1),(3,1),(4,1),(5,1),(6,1),(1,1)]$;
- $H_c=[(1,1),(2,1),(3,1),(4,0),(5,1),(6,1),(1,1),(2,1),(3,0),(4,1),(5,0),(6,0)]$.

序列中的每一项 (k_t, r_t) 表示学生作答的结果,其中 k_t 表示学生当前作答习题所涵盖的知识点, r_t 表示学生的答题结果,结果为 0 表示答错,结果为 1 表示答对.

LFKT 模型与 DKVMN 模型对学生各个知识点掌握程度的追踪情况如图 6 所示,其中:实线表示 LFKT 模型对于学生知识点掌握程度的追踪情况,虚线表示 DKVMN 模型对于学生知识点掌握程度的追踪情况.以图 6(c) 为例,学生 a 在第 3 时刻正确回答涵盖知识点 3 的习题,LFKT 与 DKVMN 输出的知识点掌握程度值均在第 3 时刻后开始增加.该结果说明,LFKT 与 DKVMN 均可以根据学生的答题结果建模学生学习行为.图 6(a)中,学生 a 在第 8 时刻后一直没有复习知识点 1,在第 8 时刻以后,DKVMN 输出的对应知识点掌握程度值变化很小,但 LFKT 输出值变化较大,呈持续下降趋势.该结果说明:LFKT 模型可以建模学生的遗忘行为,追踪到学生由于遗忘而产生的知识掌握程度变化;但 DKVMN 忽略了学生的遗忘行为,无法追踪学生由遗忘产生的知识水平变化.

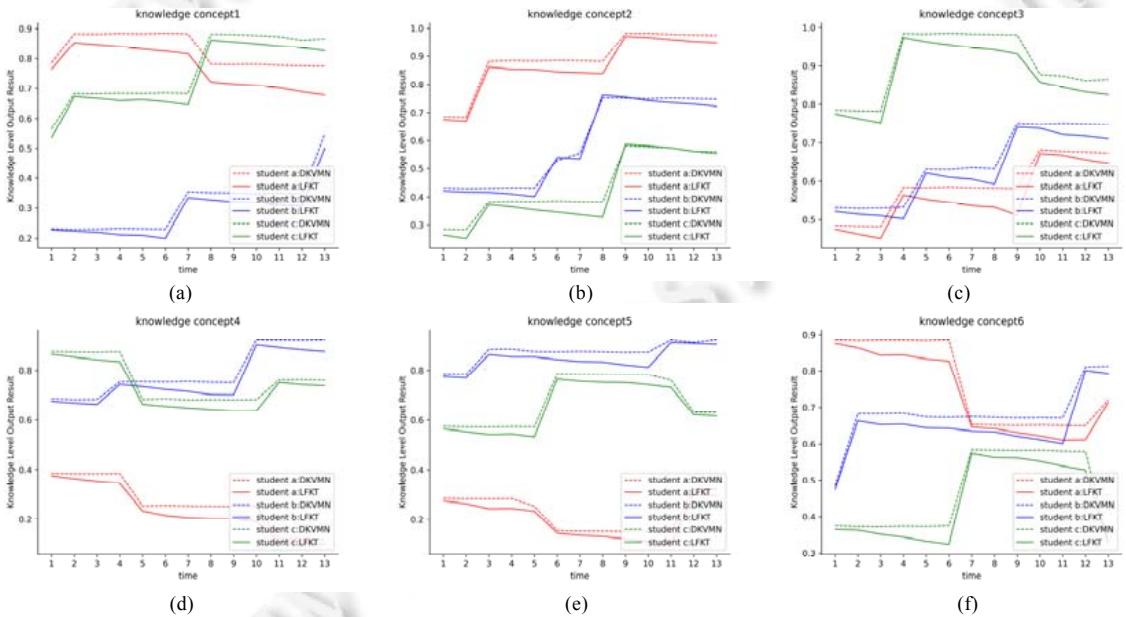


Fig.6 Knowledge level output result in slepemapy.cz

图 6 slepemapy.cz 数据集知识水平输出结果

本文汇总了 ASSISTments2012 测试集中所有学生答对一道习题后的答题序列,记录 LFKT 输出的知识点掌握程度,并记录学生下一次作答涉及该知识点习题时 LFKT 输出的知识点掌握程度,求解下一次答题时 LFKT 输出的知识点掌握程度距上一次答题结束时 LFKT 输出的知识点掌握程度的衰减比例,即 $value_{t+1}(i)/value_t(i)$.

图 7 显示了学生平均知识掌握程度衰减比例与重复学习知识点次数与时间间隔之间的关系.纵轴表示重复学习知识点次数,单位为“次”;横轴表示距离上次学习知识点的间隔时间,单位为“分钟”.数据集中的间隔时间从 0 分到几天不等,因此本文通过 $\log_2(RT)$ 的方法降低横轴的长度.由结果可知,LFKT 模型再现了学生复杂的遗忘行为,即:在相同间隔时间的情况下,重复学习次数越多,由于遗忘引起的知识点掌握程度的衰减比例越低;在相同重复学习次数的情况下,间隔时间越长,由于遗忘引起的知识点掌握程度衰减比例越大.由此可以看出:LFKT 可以根据影响知识遗忘的因素挖掘遗忘规律,建模遗忘行为,准确计算出由于遗忘引起的知识掌握程度变化.

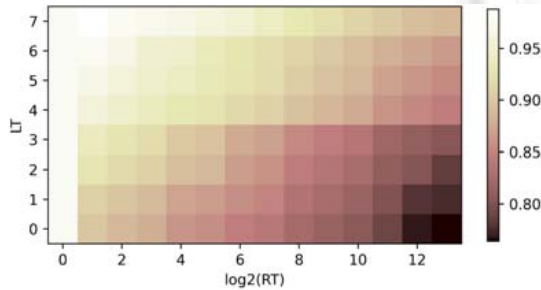


Fig.7 Quantized distribution map of knowledge level decline ratio of LFKT against repeated time interval and repeated learn times

图 7 LFKT 知识水平输出衰减比例与间隔时间和重复次数的量化分布图

5 总结与未来工作

本文重点考虑了学生遗忘行为对其知识掌握程度的影响,提出了融合学习与遗忘的深度知识追踪模型 LFKT.实验表明,LFKT 在预测学生答题表现方面优于传统知识追踪模型 BKT,DKT 和 DKVMN.LFKT 在追踪学生知识水平变化时,不但可以根据学生答题情况追踪学生学习过程中的知识水平变化,还可以体现由复杂知识遗忘因素导致的学生遗忘行为,并实时追踪由于学生遗忘造成的知识水平变化过程.

在未来的研究中,我们将针对以下方面进行探索.

- (1) 知识点之间的关系特征对知识追踪的影响.知识点之间存在先验、后验、包含等关系,需要明确知识点之间的关系特征,进而准确推测学生对于各个知识点的掌握程度;
- (2) 本文所采用数据集中习题涉及的知识点相对较少,比较容易根据学生的答题结果追踪学生对于习题考察知识点的掌握程度.对于覆盖知识点数量较高的习题,如综合性习题,会导致知识点掌握程度的不确定性.知识点高覆盖率下的知识追踪,也是我们下一步的研究工作.

References:

- [1] Adamopoulos P. What makes a great MOOC? An interdisciplinary analysis of student retention in online courses. In: Proc. of the ICIS. 2013. 1–21.
- [2] Zhou Q, Mou C, Yang D. Research progress on Int'l Conf. on educational data mining: A survey. Ruan Jian Xue Bao/Journal of Software, 2015,26(11):3026–3042 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/4887.htm> [dor: 10.13328/j.cnki.jos.004887]
- [3] Zhu TY, Huang ZY, Chen EH, Liu Q, Wu RZ, Wu L, Su Y, Chen ZG, Hu GP. Cognitive diagnosis based personalized question recommendation. Chinese Journal of Computers, 2017,40(1):176–191 (in Chinese with English abstract).
- [4] Rabiner L, Juang B. An introduction to hidden Markov models. IEEE ASSP Magazine, 1986,3(1):4–16.
- [5] Corbett AT, Anderson JR. Knowledge tracing: Modeling the acquisition of procedural knowledge. User Modeling and User-Adapted Interaction, 1995,4(4):253–278.

- [6] Graves A, Mohamed A, Hinton G. Speech recognition with deep recurrent neural networks. In: Proc. of the IEEE Int'l Conf. on Acoustics, Speech and Signal Processing. Vancouver: IEEE, 2013. 6645–6649.
- [7] LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-Based learning applied to document recognition. Proc. of the IEEE, 1998, 86(11):2278–2324.
- [8] Piech C, Bassen J, Huang J, Ganguli S, Sahami M, Guibas L, Sohl-Dickstein J. Deep knowledge tracing. In: Proc. of the Neural Information Processing Systems. 2015. 505–513.
- [9] Zhang JN, Shi XJ, King I, Yeung D. Dynamic key-value memory networks for knowledge tracing. In: Proc. of the Int'l Conf. on World Wide Web. 2017. 765–774.
- [10] Sukhbaatar S, Weston J, Fergus R. End-to-End memory networks. In: Proc. of the Neural Information Processing Systems. 2015. 2440–2448.
- [11] Ebbinghaus H. Memory: A contribution to experimental psychology. *Annals of Neurosciences*, 2013,20(4):155–156.
- [12] Murre JM, Dros J. Replication and analysis of ebbinghaus' forgetting curve. *PLoS One*, 2015,10(7):e0120644.
- [13] Bailey CD. Forgetting and the learning curve: A laboratory study. *Management Science*, 1989,35(3):340–352.
- [14] Khajah M, Lindsey RV, Mozer MC. How deep is knowledge tracing. In: Proc. of the Int'l Conf. on Educational Data Mining. 2016. 94–101.
- [15] Qiu YM, Qi YM, Lu HY, A Z. Does time matter? Modeling the effect of time with Bayesian knowledge tracing. In: Proc. of the Int'l Conf. on Educational Data Mining. 2011. 139–148.
- [16] Nagatani K, Zhang Q, Sato M, Chen YY, Chen F, Ohkuma T. Augmenting knowledge tracing by considering forgetting behavior. In: Proc. of the World Wide Web Conf. 2019. 3101–3107.
- [17] Pardos ZA, Heffernan NT. KT-IDEM: Introducing item difficulty to the knowledge tracing model. In: Proc. of the Int'l Conf. on User Modeling, Adaptation, and Personalization. Springer-Verlag, 2011. 243–254.
- [18] Baker D, Corbett AT, Aleven V. More accurate student modeling through contextual estimation of slip and Guess probabilities in Bayesian knowledge tracing. In: Proc. of the Int'l Conf. on Intelligent Tutoring Systems. Springer-Verlag, 2008. 406–415.
- [19] Khajah M, Wing RM, Lindsey RV, Mozer MC. Integrating latent-factor and knowledge-tracing models to predict individual differences in learning. In: Proc. of the Int'l Conf. on Educational Data Mining. 2014. 99–106.
- [20] Chen PH, Lu Y, Zheng VW, Pian Y. Prerequisite-Driven deep knowledge tracing. In: Proc. of the 2018 IEEE Int'l Conf. on Data Mining (ICDM). IEEE, 2018. 39–48.
- [21] Su Y, Liu QW, Liu Q, Huang ZY, Yin Y, Chen EH, Ding C, Wei S, Hu GP. Exercise-Enhanced sequential modeling for student performance prediction. In: Proc. of the National Conf. on Artificial Intelligence. 2018. 2435–2443.
- [22] Abdelrahman G, Wang Q. Knowledge tracing with sequential key-value memory networks. In: Proc. of the Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval. 2019. 175–184.
- [23] Sun X, Zhao X, Ma Y, Yuan XR, He FJ, Feng J. Multi-Behavior features based knowledge tracking using decision tree improved DKVMN. In: Proc. of the ACM Turing Celebration Conf. China, 2019. 1–6.
- [24] Liu Q, Huang ZY, Yin Y, Chen EH, Xiong H, Su Y, Hu GP. EKT: Exercise-aware knowledge tracing for student performance prediction. *IEEE Trans. on Knowledge and Data Engineering*, 2019:1–15.
- [25] Nakagawa H, Iwasawa Y, Matsuo Y. Graph-Based knowledge tracing: Modeling student proficiency using graph neural network. In: Proc. of the IEEE/WIC/ACM Int'l Conf. on Web Intelligence (WI). IEEE, 2019. 156–163.
- [26] Minn S, Desmarais MC, Zhu FD, Xiao J, Wang JZ. Dynamic student classification on memory networks for knowledge tracing. In: Proc. of the Pacific-Asia Conf. on Knowledge Discovery and Data Mining. Cham: Springer-Verlag, 2019. 163–174.
- [27] Cheng S, Liu Q, Chen EH, Huang Z, Huang ZY, Chen YY, Ma HP, Hu GP. DIRT: Deep learning enhanced item response theory for cognitive diagnosis. In: Proc. of the ACM Int'l Conf. on Information and Knowledge Management. 2019. 2397–2400.
- [28] Shen SH, Liu Q, Chen EH, Wu H, Huang ZY, Zhao WH, Su Y, Ma HP, Wang SJ. Convolutional knowledge tracing: Modeling individualization in student learning process. In: Proc. of the Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval. 2020. 1857–1860.
- [29] Ai FZ, Chen YS, Guo YC, Zhao YX, Wang ZZ, Fu GW. Concept-Aware deep knowledge tracing and exercise recommendation in an online learning system. In: Proc. of the Int'l Conf. on Educational Data Mining (EDM 2019). 2019. 240–245.

附中文参考文献:

- [2] 周庆,牟超,杨丹.教育数据挖掘研究进展综述.软件学报,2015,26(11):3026-3042. <http://www.jos.org.cn/1000-9825/4887.htm> [doi: 10.13328/j.cnki.jos.004887]
- [3] 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.计算机学报,2017,40(1):176-191.



李晓光(1973—),男,博士,教授,主要研究领域为机器学习,数据挖掘,智慧教育.



杜岳峰(1986—),男,博士,讲师,CCF 专业会员,主要研究领域为数据质量管理,推荐系统,机器学习.



魏思齐(1996—),男,硕士,主要研究领域为智慧教育,数据挖掘,深度学习.



于戈(1962—),男,博士,教授,博士生导师,CCF 会士,主要研究领域为数据库理论与技术,分布与并行式系统,云计算与大数据管理,区块链技术与应用.



张昕(1979—),男,博士,副教授,CCF 专业会员,主要研究领域为数据挖掘,复杂网络,机器学习.

www.jos.org.cn