



研究与开发

基于知识追踪机的多特征融合习题推荐模型

诸葛斌, 汪盈, 肖梦凡, 颜蕾, 王冰雁, 董黎刚, 蒋献
(浙江工商大学信息与电子工程学院, 浙江 杭州 310020)

摘要: 个性化习题推荐是智慧教育个性化服务领域的重要课题, 然而传统的习题推荐算法对于学生特征的研究不够彻底, 对于学生知识掌握与答题行为之间的关联信息挖掘未能充分, 导致推荐精准度不佳。为解决上述问题, 结合知识追踪机(KTM)和基于用户的协同过滤算法, 提出一种基于KTM多特征融合的习题推荐模型SKT-MFER。该模型首先构造了一个融入学生学习行为和学习能力的知识追踪模型KTM-LC, 精准挖掘学生的知识掌握水平; 接着设置两次筛选, 先利用知识点掌握矩阵初步筛选出相似的学生, 再根据认知状态相似度和习题难度相似度组合而成的综合相似度进行二次筛选, 双重过滤以保障习题推荐的准确度。通过广泛的实验证明, 所提方法相比于一些现有的基线模型有更好的效果。

关键词: 智慧教育; 习题推荐; 知识追踪; 协同过滤; 因子分解机

中图分类号: TP301

文献标志码: A

doi: 10.11959/j.issn.1000-0801.2024210

A multi-feature fusion exercise recommendation model based on knowledge tracing machines

ZHUGE Bin, WANG Ying, XIAO Mengfan, YAN Lei, WANG Bingyan, DONG Ligang, JIANG Xian
College of Information and Electronic Engineering, Zhejiang Gongshang University, Hangzhou 310020, China

Abstract: The subject of personalized exercise recommendation holds significant relevance within the domain of personalized services in smart education. Nevertheless, traditional algorithms have often lacked a deep understanding of student characteristics and failed to adequately explore the relationship between knowledge mastery and question-answering behaviors, leading to low recommendation accuracy. To address these issues, combining the knowledge tracing machine and the user-based collaborative filtering algorithm, as a KTM-based multi-feature fusion exercise

收稿日期: 2024-06-18; 修回日期: 2024-08-16

通信作者: 蒋献, jiangxianxd123@126.com

基金项目: 浙江省领雁研发项目(No.2023C03202); 2022年度浙江省高等教育学会高等教育研究重点立项课题(No.KT2022017); 2023年浙江省教育厅一般科研项目(No.Y202353291); 浙江省普通本科高校“十四五”教学改革项目(No.jg20220247)

Foundation Items: Zhejiang Lingyan R&D Project (No.2023C03202), Key Project of Higher Education Research of Zhejiang Higher Education Association in 2022 (No.KT2022017), 2023 General Research Project of Zhejiang Provincial Department of Education (No.Y202353291), The “14th Five-Year Plan” Teaching Reform Project of General Undergraduate Universities in Zhejiang Province (No.jg20220247)



recommendation model, SKT-MFER was proposed. Firstly, as a knowledge tracking model, KTM-LC, incorporating student learning behaviors and learning abilities, was constructed to accurately assess the student's knowledge mastery level. Subsequently, two filters were implemented to ensure the exercise recommendation's accuracy: the first was an initial screening utilizing the knowledge point mastery matrix to eliminate students who were similar to the target student, and the second was a filtering process considering the combined similarity of cognitive state similarity and exercise difficulty similarity. Through extensive experiments, it proves that the proposed method yields better results than some existing baseline models.

Key words: smart education, exercise recommendation, knowledge tracking, collaborative filtering, factorization machine

0 引言

当今世界,教育发展面临很多挑战,如教学质量不高、资源分配不均、服务单一等问题。随着科技进步,智慧教育逐渐成为可能。智慧教育以人工智能、大数据分析和云计算等信息技术为支撑,将信息技术和教育紧密结合,更注重学生个性、学习兴趣、学习方式等多元化因素,更加强调整个性化的学习,让教育更加智慧,提高教育效果,改善教育服务和管理。然而,信息技术的飞速发展也带来了“信息过载”问题。学习者面对海量习题,难以高效选择适合自己的内容,个性化学习需求日趋迫切。个性化推荐是解决这个问题的有效方法之一^[1],在线教育应该运用个性化推荐技术,根据学习者的特点,精准推荐适合的习题^[2]。

通过人工智能等前沿技术,对海量学习信息加以研究和管理,可以更加准确地了解学生的学习情况,从而为教育者提供更有效的学习指导。这些学习数据分析还可以用来建立模型,以便自动追踪学生在学习过程中的每一个阶段,实现个性化学习的目标。这种建模过程被称为知识追踪(knowledge tracing, KT),其目标是预测学生在下一次答题时是否正确。知识追踪机(KTM)^[3]是一种具有代表性的模型,使用因子分解机来估计学生的知识情况。基于因子分解机估计学生知识情况,并扩展自项目反应理论模型,延伸出因

子分析模型^[4],解决了数据稀疏性问题,同时通过对学生做题情况的分析,提供了更具解释力的预测。

协同过滤(collaborative filtering, CF)是一种在推荐系统中广泛使用的算法,其主要思想是通过分析大量用户的历史数据,找出相似用户之间的行为模式,从而预测目标用户对物品的兴趣。而UserCF^[4]是基于用户的协同过滤算法,寻找与目标用户行为或偏好相似的其他用户,并提出相应推荐。虽然基于UserCF的习题推荐方法结构简单、可解释性强,但这类算法忽略了不同学生知识掌握水平的差异,只能粗粒度对学生进行分类,无法满足智慧教育的个性化习题推荐要求。

为了进一步完善上述建模模型和推荐算法,本文提出了基于KTM多特征融合的习题推荐模型研究,研究内容和主要贡献如下。

(1) 本文提出了一种融合学生学习行为和学习能力特征的学生知识点掌握程度预测模型,利用融合了遗忘因素的长短期记忆(long short-term memory, LSTM)网络获得学生的知识点情况,然后联合习题难度构建学生的学习能力,最后通过改进KTM模型得到学生的知识掌握程度。

(2) 本文设计了两次过滤习题的方式:采用协同过滤算法计算出相似度较高的学生,引入削减因子提高未做题权重,从而初筛出候选集;利用认知诊断模型得到学生知识点认知状态,计

算出认知相似度和习题难度相似度,联合两种相似度得出综合相似度,提高习题再次筛选的推荐准确度,更好地实现了学生个性化习题推荐。

(3)在4个真实的开源知识追踪数据集上对本文提出的推荐模型效果进行评估,并与其他主流模型进行对比,在相应的评价指标上均有所提高。

1 相关工作

本节对知识追踪及习题个性化推荐的相关研究分别进行了简要回顾。

1.1 知识追踪

知识追踪的目标是观察、表示和量化学生的知识状态^[5]。由于学生在学习过程中会反复进行学习和练习,因此知识追踪本质上是一个时间序列问题,许多离散时间序列模型都可以用于建模。早期的知识追踪模型,如贝叶斯知识追踪(Bayesian knowledge tracing, BKT),基于一阶马尔可夫模型,由Corbett和Anderson首次提出并应用于智能导学系统^[6]。斯坦福大学首次引入深度学习方法,提出了深度知识追踪(deep knowledge tracing, DKT)^[7],利用递归神经网络对学生的学习过程建模,取得了突破性成果。

为了更好地把握学生做题情况的历史关联性,NPA模型^[8]引入双向长短期记忆(Bi-LSTM)网络并结合一个注意力层对学生的学习历史进行分析;EKT模型^[9]使用Bi-LSTM生成练习数据向量,综合考虑题目难度和区分度,解决了冷启动问题;Zhang等^[10]提出了一种动态键值记忆网络(DKVMN)模型,能更准确地输出学生知识掌握情况;DKVMN-CA模型^[11]在DKVMN模型的基础上,将题目概念的等级结构引入神经网络的设计中,对题目概念的多级结构进行建模;Yang等^[12]基于遗忘曲线理论,采用3DConvNets强化学生近期知识状态和短期效果,结合LSTM捕获长期特征提出卷积知识追踪(CKT);Nakagawa

等^[13]提出了基于图的知识追踪GKT模型,利用图神经网络处理学生做题历史数据。随着技术不断进步,如注意力知识追踪(AKT)^[14]、基于对抗训练的ATKT^[15]、基于问题难度匹配的DIMKT^[16]、以问题为中心的可解释QIKT^[17]、基于辅助任务增强的AT-DKT^[18],为知识追踪进一步改进提供了多样化思路。

1.2 习题推荐算法

当前主要的推荐技术包括基于内容的推荐(content-based, CB)、协同过滤(collaborative filtering, CF)推荐和混合推荐。基于内容的推荐通过分析用户过去的偏好推荐相似项目,例如,Huang^[19]设计了基于内容的MOOC课程推荐模型;Camposet^[20]利用非负矩阵分解计算学生相似性并进行课程推荐。协同过滤推荐算法可以分为用户协同过滤(UserCF)、项目协同过滤(ItemCF)和模型协同过滤(ModelCF),例如,柏茂林^[21]通过习题正解概率与用户能力特征进行推荐;付兴华等^[22]构建学习兴趣和教学资源模型进行组合推荐;吴云峰^[23]提出基于多分类器的迁移Bagging习题推荐算法。混合推荐结合多种技术和过滤算法以提高效果,如利用模糊树匹配^[24]、LDA主题模型和文本卷积神经网络建模^[25]、深度学习技术^[26]等,进一步提升了推荐上精准度。

2 方法

推荐系统框架采用一种先筛选后过滤的模式,首先利用知识追踪模型获得学生知识掌握水平,找到与其目标学生相似的学生,初步筛选以形成试题初筛候选集;在此基础上,利用学生认知矩阵进行认知相似度的计算,接着根据习题难度指标进行习题难度的相似度计算,然后把这两种相似度进行综合计算,而候选试题集经过综合相似度计算后得出最后待推荐的试题列表,真正实现个性化的习题推荐,习题推荐框架如图1所示。

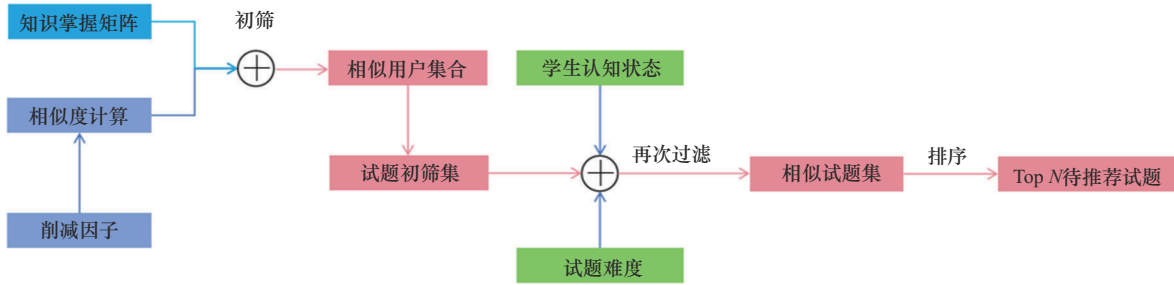


图1 习题推荐框架

2.1 知识追踪方法

2.1.1 KTM

基于因子分解机 (FM) 的方法在多个数据集中经过验证, 不仅能够处理基于序列的数据, 也能很好地处理稀疏数据, 同时还能很方便地引入多种边信息, 如知识概念、练习、做题次数、知识掌握情况等。KTM通过根据学生历史记录, 形成了三元组, 设学生为 s , 问题为 q , 回答情况为 r , 所以在某一时刻学生回答情况的三元组表示为: (s_t, q_t, r_t) 。

(1) 学生得分矩阵 M

学生得分矩阵 M 示例见表1, 其中 M_{ij} 表示第 i 个学生在题目 j 上的回答情况, 若 $M_{ij}=1$, 则表示学生 i 回答正确题目 j , 反之 $M_{ij}=0$ 表示学生回答错误题目 j 。

表1 学生得分矩阵 M 示例

M	试题1	试题2	试题3	...	试题 J
学生1	1	0	1	...	1
学生2	0	0	1	...	0
...
学生 I	0	1	1	...	1

(2) 知识点考察矩阵 Q

知识点考察矩阵 Q 示例见表2, 其中 q_{jk} 表示题目 j 是否涉及知识点 k , $q_{jk}=1$ 表示题目 j 中涉及知识点 k , 反之 $q_{jk}=0$ 则不涉及。但这是静态的知识点考察矩阵, 随着时间变化, 学生的知识状态在不断地更新, 所以用动态的知识掌握矩阵 Q_t 来表示学生在 t 时刻的知识情况, 该动态矩阵随着学生学习情况而不断变化。

表2 知识点考察矩阵 Q 示例

Q	知识点1	知识点2	知识点3	...	知识点 K
试题1	1	0	0	...	1
试题2	0	1	1	...	0
...
试题 J	0	0	1	...	1

用户 ID、题目 ID、成功的尝试、失败的尝试, 以及其他学习记录的其他信息作为 KTM 的输入, 且设所有特征为 N , 并以独热编码的方式将所有特征连接起来编码到一个具有稀疏特征的向量中, 记为 x , 且 $x_i > 0$, 且特征 i 的取值为 $1 \leq i \leq N$ 。将所有特征依次输入嵌入层、FM 层, 然后输出预测概率, KTM 模型结构如图2所示。

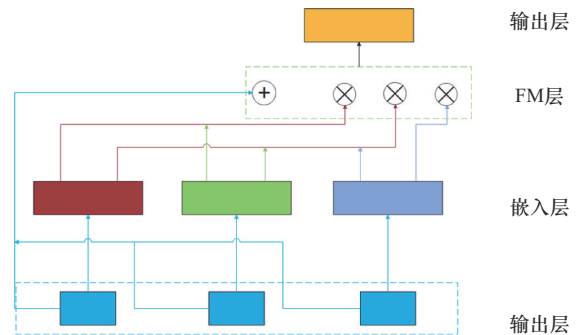


图2 KTM 模型结构

最后得出计算出学生回答正确问题的概率表示为:

$$p(y_t=1) = \sigma(\mu + \sum_{i=1}^N w_i x_{i,t} + \sum_{1 \leq i < j \leq N} \langle v_i, v_j \rangle x_{i,t} x_{j,t}) \quad (1)$$

其中, μ 全局偏差, $x_{i,t}$ 和 $x_{j,t}$ 是在 t 时刻收集的 N 个

特征向量中的第*i*个和第*j*个抽象特征。每个特征*i*都由偏差 $w_i \in R$ 和嵌入 $v_i \in R^d$ 构成，其中*d*为某个维度。

2.1.2 基于遗忘规律的学习行为模型

学生知识学习并非静态不变的，需要动态跟踪学生学习进程，考虑学生学习过程中的遗忘现象。在教育心理学领域，有一些研究关注人类的遗忘行为，其中最经典的就是艾宾浩斯遗忘曲线理论，其描述了学生遗忘知识的规律^[27]。遗忘曲线显示了遗忘进程的非均衡性，遗忘速度最初很快，之后则逐渐减缓，在相当时间段后，遗忘速度几乎不再变化，即遗忘的进程是“先快后慢”。学生遗忘行为受上一次学习与当前测试之间的时间间隔和学习材料的重复次数的影响。因此，本文主要考虑上述两个特征，借助全连接网络构建记忆擦除向量，并融合动态知识点情况，更新当前学生知识情况矩阵。

该方法受LSTM网络中忘记门的影响。遗忘处理过程如图3所示，为了应对学生遗忘的问题，可以将影响遗忘的因素 F_t 与学生知识点掌握的时序矩阵 $Q'_t(j)$ 进行处理。基于LSTM中忘记门和输入门的网络结构，该模型使用遗忘因素来更新知识点矩阵，以便在去除随时间逐渐遗忘的知识点后，得到新的知识点变化矩阵，并利用sigmoid激活函数的全连接层实现遗忘过程，具体计算式为：

$$f_t = \sigma[E_e F_t(k) + b_e] \quad (2)$$

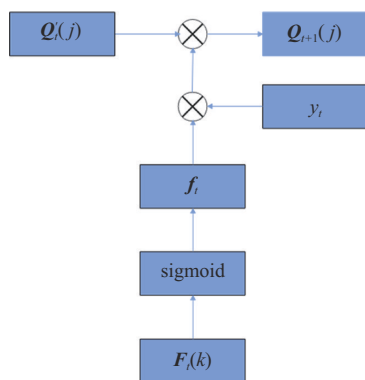


图3 遗忘处理过程

其中 σ 表示sigmoid激活函数， E_e 为权重矩阵，形状是 $d_k \times d_k$ ，全连接层的偏置向量 b_e 为 d_k 维。

对知识变化矩阵进行更新。具体计算式为：

$$y_t(j) = \text{soft max}[v_t Q(j)] \quad (3)$$

$$Q_{t+1}(j) = Q'_t(j)[1 - y_t(j)f_t] \quad (4)$$

其中， y_t 表示问题和知识点之间的相关权重矩阵， v_t 是嵌入向量。

通过加入遗忘因素，得到了考虑了时间因素的知识情况矩阵。该矩阵作为KTM的输入，可以进一步增强时序信息。在更新知识状态时，会考虑遗忘因素和学生回答的时间关系，从而提高预测准确率。

2.1.3 基于题目难度的学习能力模型

题目与题目之间是有难度区分的，因为每个题包含的知识点不同，有些题只有一个知识点，但有些题又有多个知识点涉及，且不同题目类型也有不同难度。但大部分知识追踪模型并没有考虑难度，所以本文通过对难度特征的引用来增强知识追踪模型。难度系数用*D*表示，表达式如下。

$$D(q_j) = \begin{cases} \frac{\sum_i |\{q_{ij}=0\}|}{|N_j|}, & |N_j| \geq 5 \\ 0.5, & \text{其他} \end{cases} \quad (5)$$

其中， N_j 是学生回答问题集合， q_{ij} 是学生*i*回答问题*j*的结果，所以难度系数就是回答错误的人数与总人数之比。

计算出问题系数后建立一个全连接神经网络，输入是问题编号，输出是问题难度，然后初始化 $d_t \in R^K$ 为问题难度的嵌入表示，当进入模型训练后就会转化为嵌入矩阵 $D \in R^{Q \times K}$ ，其中*Q*为总的问题数，*K*是嵌入向量的维度。

问题难度的评定主要基于学生犯错情况，因此可以通过转变思路来推断学生的学习能力，即通过正确率与错误率的差异来判断其对问题的掌握程度。



$$\text{Correct}(x_{ij}) = \frac{\sum_i^{|N_j|} \{q_{ij}=1\}}{|N_{ij}|} \quad (6)$$

$$\text{Incorrect}(x_{ij}) = \frac{\sum_i^{|N_j|} \{q_{ij}=0\}}{|N_{ij}|} \quad (7)$$

$$A(s_i) = \begin{cases} \text{Correct}(x_{ij}) - \text{Incorrect}(x_{ij}), & |N_{ij}| \geq 5 \\ D(q_j), & \text{其他} \end{cases} \quad (8)$$

其中, $\text{Correct}(x_{ij})$ 和 $\text{Incorrect}(x_{ij})$ 分别代表学生 i 正确回答问题 q_j 的概率, N_{ij} 表示学生 i 回答问题 q_j 的次数, 且学生 i 回答问题 q_j 的次数少于 5 次, 则用问题 q_j 的难度代替学生 i 的能力。

与难度嵌入的方法类似, 建立一个全连接网络, 输入是学生学号, 输出就是学生学习能力, 然后初始化 $se_i \in \mathbf{R}^K$ 为学生学习能力 000 嵌入, 且在模型训练中转化为嵌入矩阵 $\mathbf{A} \in \mathbf{R}^{S \times K}$, 其中 K 为嵌入向量的维度, S 为学生总数。在模型输入时这两个特征进行联合输入。

2.1.4 KTM-LC 模型

尽管 KTM 在处理稀疏数据和提供可解释性预测方面表现出色, 但由于未充分考虑时间因素和学生学习过程的动态变化, 其在预测学生知识掌握变化上的准确性有所局限。为此, 本文提出

了 KTM-LC (knowledge tracing machines with learning behavior and capacity) 模型, KTM-LC 模型框架如图 4 所示。

KTM-LC 模型一方面考虑学生在学习过程中的遗忘现象, 通过引入基于遗忘规律的学习行为模型, 利用 LSTM 网络的遗忘门机制动态追踪学生的知识状态, 剔除逐渐遗忘的知识点, 更新知识点矩阵; 另一方面, 模型结合题目难度, 构建学生的学习能力特征, 通过分析学生在不同难度题目上的表现, 精确评估其知识掌握情况。

KTM 的输入包括原来的学习记录, 学生正确尝试的特征向量和学生失败尝试的特征向量, 学生编号 (uid) 和题目编号 (qid) 外, 还有问题难度和学习能力联合的特征向量。这些输入最后会编码到一个具有稀疏特征的向量 \mathbf{x}_t 中, 然后与加入了遗忘因素的 LSTM 进行结合, 具体计算式如下:

$$\mathbf{x}'_t = Q_{t+1} \oplus \mathbf{x}_t \quad (9)$$

得到新的向量后就输入 FM 中进行特征交互, 从而得到学生对于知识点的掌握情况。改进后的概率计算式如下。

$$\hat{y}(\mathbf{x}') = w_0 + \sum_{i=1}^{N+K} w_i x'_i + \sum_{i=1}^N \sum_{j=i+1}^{N+K} \langle v_i, v_j \rangle x'_i x'_j \quad (10)$$

其中, M 为特征数量, K 为 LSTM 的隐藏状态

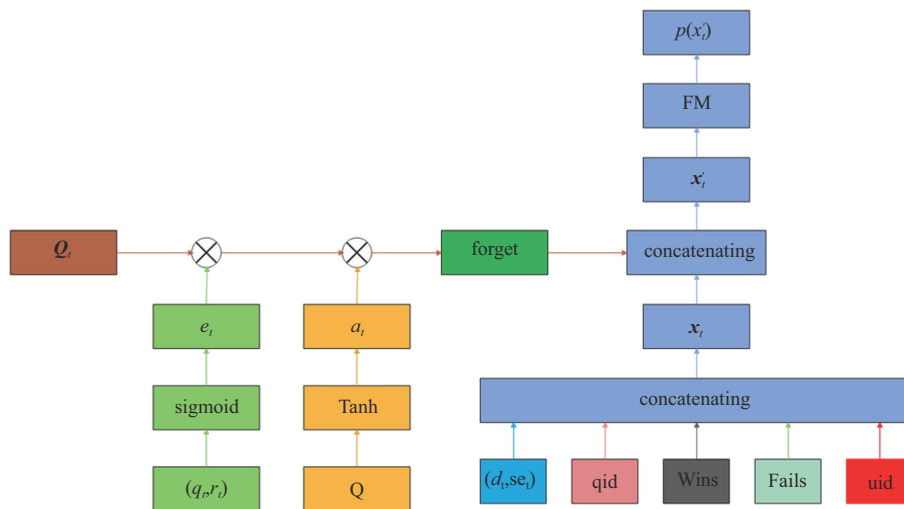


图4 KTM-LC模型框架

大小。

为了提高模型性能,本文使用标准交叉熵损失函数来优化模型的各个参数,该函数利用模型输出的预测值 $\hat{y}(\mathbf{x}')$ 和实际回答情况 r_t 之间的差异,具体计算式如下。

$$L = -\sum(r_t \log(\hat{y}(\mathbf{x}')) + (1 - r_t)(1 - \log(\hat{y}(\mathbf{x}')))) \quad (11)$$

2.2 SKT-MFER 习题推荐算法

本文通过知识追踪模型获取学生的知识掌握情况,并根据这些数据找到知识状态相似的学生,再基于这些相似学生进行个性化的习题推荐。为了提高推荐准确性,模型综合考虑了学生的认知状态特征和习题的难度特征,最终形成了基于学生知识追踪的多特征习题推荐模型(multi-feature exercise recommendation model based on student knowledge tracing, SKT-MFER),不仅能准确反映学生当前的认知水平,还能提供难度适宜的习题,以有效提升学习效果。

2.2.1 基于改进相似度计算方法的初步筛选

初筛阶段的相似度计算是根据皮尔逊相关系数进行改进的。对于待推荐学生来说,增加目标学生未做过题的关注度对目标学生来说更有意义,可以进行新内容的学习,完善自己的知识点掌握情况,所以本算法在皮尔逊相关系数的基础上加入了削减因子,其目的就是削减目标学生经常做的题目的权重,提高待推荐学生没涉及的问题的权重。削减因子计算式如下。

$$R(q) = e^{-\frac{n_q}{n}} \quad (12)$$

其中, q 为习题, n 为题目总数, n_q 为做过题目数, n 越大,则 $R(q)$ 值越小,即已做过题目对于相似度的计算影响越小,相应地也就提高了未做题权重,这样在待推荐题库中,增加了目标学生未做过题的概率。

综合上面两个计算式,把削减因子融入皮尔逊相关系数中,改进后的相似度计算式为:

$$\text{sim}(s, s') = \frac{\sum_{k \in K_{s, s'}} (a_{s, k} - \hat{a}_s)(a_{s', k} - \hat{a}_{s'}) e^{-\frac{n_q}{n}}}{\sqrt{\sum_{k \in K_{s, s'}} (a_{s, k} - \hat{a}_s)^2} \sqrt{\sum_{k \in K_{s, s'}} (a_{s', k} - \hat{a}_{s'})^2}} \quad (13)$$

其中, (s, s') 分别为目标学生和相似学生, $K_{s, s'}$ 为目标学生的相似学生重合的知识点掌握程度的数据, $a_{s, k}$ 和 $a_{s', k}$ 分别为学生 s 和学生 s' 对知识点 k 的掌握情况, \hat{a}_s 和 $\hat{a}_{s'}$ 分别为学生 s 和学生 s' 的对知识点平均掌握情况。

通过相似度计算式计算后,就能找到目标学生的最相似的学生,即可根据相似学生知识掌握情况矩阵里的习题组合成第一阶段的习题集 $G(k)$ 。

2.2.2 综合学生认知状态和试题难度相似度的试题再过滤

在初步筛选之后,为了进一步优化推荐算法,不仅需要考虑相似学生认知状态特征,还应结合习题的相似度和难度相似度来进行最终推荐。

(1) 学生认知状态相似度

将认知状态矩阵与题目认知属性关系矩阵进行相似度计算,得到认知状态相似度:

$$w_k = \text{sim}(\vec{K}_i, \vec{L}_i) = \frac{\vec{K}_i \cdot \vec{L}_i}{|\vec{K}_i| |\vec{L}_i|} \quad (14)$$

其中, \vec{K}_i 为习题库中的认知属性向量的夹角余弦, \vec{L}_i 为认知状态向量的夹角余弦。

通过学生认知状态矩阵和知识点认知属性矩阵进行结合,能够筛选出认知属性欠缺的知识点对应的试题。

(2) 试题难度相似度

难度是衡量试题的一个标准,每个人认知标准不一样,所需要设置的试题难度标准也不同,所以通过对初筛后的试题集与学生本身做题记录中的难度进行相似度比较,如果筛选出的难度



高, 则代表学生需要更难的题进行推荐, 反之则为其推荐简单的题, 需要筛选出适合学生难度的题目。

$\text{sim}(d, d') =$

$$\frac{\sum_{q \in Q_{g, g'}} (d_{g, q} - \dot{d}_g)(d_{g', q} - \dot{d}_{g'})}{\sqrt{\sum_{q \in Q_{g, g'}} (d_{g, q} - \dot{d}_g)^2} \sqrt{\sum_{q \in Q_{g, g'}} (d_{g', q} - \dot{d}_{g'})^2}} \quad (15)$$

其中, d 和 d' 分别为学生本身做题记录中结合学生能力的难度和初筛后的试题的难度, g 和 g' 分别为学习记录中的题库和初筛后形成的试题集, $d_{g, q}$ 和 $d_{g', q}$ 分别为在学习记录中的题库对应的题目难度和初筛后的题库中对应的题目难度, \dot{d}_g 和 $\dot{d}_{g'}$ 分别代表两个题库中的平均试题难度。

经过式 (5) 可以得到与学生比较匹配的难度。考虑有些学生可能没有学习记录, 无法通过初筛形成试题集, 那么可以根据题库里的题目难度进行推荐, 难度分为 3 个等级, 分别为简单、中等和难。最后通过对关于试题的两个相似度的联合考虑, 计算出一个综合性的相似度, 这样能够通过对比试题的多个相似度结合进行更加精准的试题推荐。

(3) 综合相似度计算

为了联合学生认知状态相似度和试题难度相似度, 从而确定出一个最后的相似度计算结果, 这里就需要加入两个权重参数 $\lambda_1, \lambda_2 \in [0, 1]$, 且 $\lambda_1 + \lambda_2 = 1$, 根据试题中的实际情况, 实时调整两个相似度的权重, 其中 λ_1 属于学生认知状态相似度的权重, λ_2 属于试题难度相似度权重, 其取值能够影响最后试题的相似度, 所以其取值很重要, 其计算式如下。

$$\lambda_1 = \frac{w_k^2}{\text{sim}^2(d, d') + w_k^2} \quad (16)$$

$$\lambda_2 = \frac{\text{sim}^2(d, d')}{\text{sim}^2(d, d') + w_k^2} \quad (17)$$

权重计算完毕后就进行综合相似度的计算,

计算式如下。

$$\text{sim}(q, q') = \lambda_1 \times w_k + \lambda_2 \times \text{sim}(d, d') \quad (18)$$

其中, q 和 q' 分别是初筛后的试题集和学生做题记录中的试题集。

最后根据综合相似度的计算结果生成相应推荐列表, 并对其按照顺序进行推荐, 即习题的 TopN 推荐。

3 实验

为了评价本文所提习题推荐模型的有效性和可行性, 本节首先介绍实验所用数据集, 然后介绍用于比较的基线模型, 接着介绍评价指标以及实现细节, 最后对实验结果进行分析比较得出结论。

3.1 实验数据集介绍

本文所用数据集为公开教育数据集: Assistments2009, Assistments2017, KDD Cup2010, 以及本团队开发的教育系统豆豆云数据集, 各数据集简介见表 3。

(1) Assistments2009: 这个数据集来自于 Assistments 在线教育平台公开的对数学练习的记录, 其中包含 4 217 名学生 ID, 124 个问题, 346 860 条学习记录。

(2) Assistments2017: 该数据集也是来自于 Assistments 在线教育平台公开的数据集, 其中包含 1 709 名学生 ID, 4 117 个问题, 942 816 条学习记录。

(3) KDD Cup2010: 该数据集是 2010 年 KDD 杯比赛开发数据集, 其中包含 1 338 名学生 ID, 91 913 个问题, 2 270 384 条学习记录。

(4) 豆豆云: 该数据集采集自豆豆云小程序做题记录, 豆豆云小程序是本团队自主所开发的智慧在线教育系统, 目前已有近三万名用户, 数据库中有两千多万条历史学习数据, 其中包含 30 个题库, 其中马克思原理科目共有 7 562 名学

生ID, 4 372 193条学习记录, 本文采用马克思原理科目的学生数据。

表3 各数据集简介

数据集名称	学生数	题目数	做题记录数
ASSISTment2009	4 217	124	346 860
ASSISTment2017	1 709	4 117	942 816
KDD 2010	1 338	91 913	2 270 384
豆豆云	7 562	1 333	4 372 193

3.2 实验环境与参数设置

本实验所采用的是 Windows10 操作系统, 处理器为 Intel i7-6500U, 内存为 8 GB, 使用的语言为 Python, 版本为 Python3.7。

为了评估模型性能, 本实验采用了五重交叉验证方法。每次交叉验证中, 数据集被随机分成若干子集, 训练集、验证集和测试集分别占 70%、10% 和 20%。本模型和参与比较的模型均采用与一致的超参数设置, 批处理大小 batch-size 为 32, TopN 的 N 取值为 15, 学习速率为 0.001, 训练轮数 epoch 设为 200, 使用 Adam 优化算法进行训练。

3.3 评价参数

TopN 推荐的优势在于能够简化习题选择过程, 帮助学生在大量习题中快速找到最符合其需求的选项。其效果通常通过精确率 (Precision) 和召回率 (Recall) 等指标来评估。为评价该推荐算法是否行之有效, 本文采用精准率、召回率和 F1 值 (F1) 3 类指标对本文所提出的推荐算法进行综合评价, 3 个指标的具体计算式如下。

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (19)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (20)$$

$$\text{F1} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (21)$$

其中, TP 是指推荐结果中正确推荐习题的数目,

FP 是指推荐结果中错误推荐习题的数目, FN 是指应该推荐给学生但没有进入推荐列表的习题数。

3.4 实验结果与分析

3.4.1 KTM-LC 对比实验与分析

KTM-LC 算法与基准算法的 AUC 值比较见表 4, KTM-LC 在 4 个数据集上的性能均优于其他模型, 因为 KTM-LC 在输入上考虑了多个特征指标, 所以对于学习行为和学习能力有了更多依据, 且实验数据说明了该模型对于知识追踪的效率有了提升。

表4 KTM-LC 算法与基准算法的 AUC 值比较

数据集	BKT	DKT	KTM	AKT	KTM-LC
ASST2009	0.672	0.773	0.802	0.767	0.825
ASST2017	0.657	0.741	0.757	0.764	0.775
KDD2010	0.664	0.771	0.788	0.791	0.819
豆豆云	0.652	0.753	0.782	0.787	0.798

从实验结果分析, DKT 的 AUC 在数据集 ASST2009 上的值为 77.3%, 而 KTM 的值为 80.2%, 高出 2.9%, 可以得出知识追踪机性能是高于深度学习模型性能。本模型高出 KTM 模型 2.3%, 改进模型在性能方面有所提升, 证明加入处理时序信息的 LSTM 和多特征的交互融合能够提高 KTM 的性能。

ASST2009 数据集中各模型 AUC 对比如图 5 所示, 是 4 种模型在 ASST2009 数据集上的训练和验证迭代结果。通过图 5 可以看出, 前 3 个模型都有不同程度的过拟合化的现象, 而本文所构建的新模型能够很好地预防过拟合化的发生, 所以本文模型有一定的提升效果。

3.4.2 SKT-MFER 对比实验与分析

在对比实验中, 本文选择了 4 种不同的推荐模型: UB-CF、IB-CF、DKT-CF^[28] 和 DKVMN-CF。基于用户的协同过滤 (UB-CF) 通过分析用户之间的相似性进行内容推荐, 虽然实现简

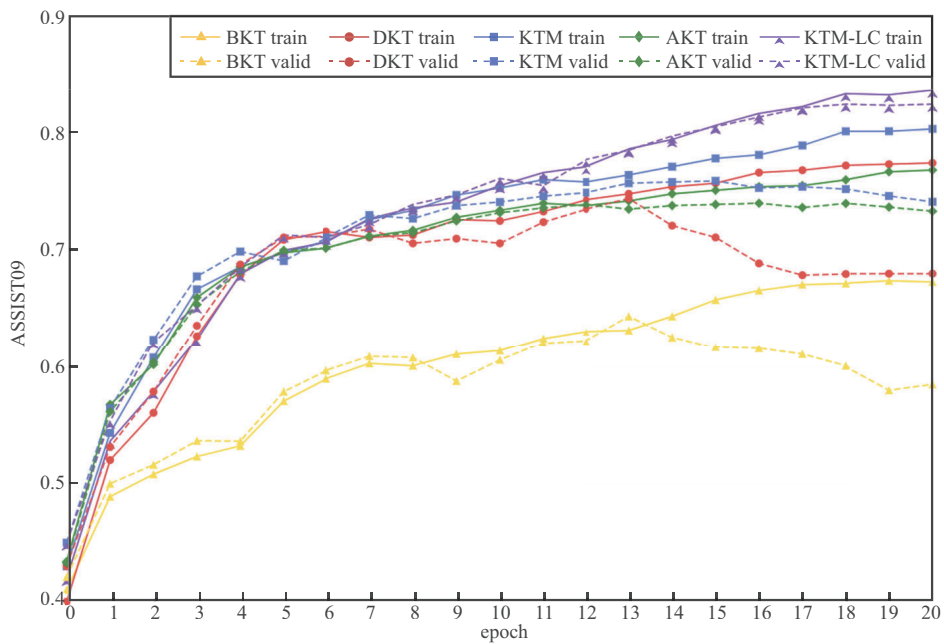


图5 ASSIST09数据集中各模型AUC对比

单且能够提供个性化推荐，但存在冷启动和数据稀疏问题；基于物品的协同过滤（IB-CF）根据物品之间的相似性推荐内容，其推荐稳定性较高，计算量相对较小，但对个性化需求的满足不如UB-CF；深度知识追踪与协同过滤结合（DKT-CF）利用深度学习模型捕捉用户知识状态的时间序列变化，提供高精度的动态推荐，适合教育领域的知识追踪应用，但对计算资源要求较高，且需要大量数据进行训练。DKVMN-CF（概念感知动态键值记忆网络协同过滤）在DKVMN-CA的基础上结合协同过滤技术，能有效实现习题推荐功能，但是高度依赖于输入数据的质量。

为了评估本文模型和以上对比模型的性能，实验将数据集按60%、70%和80%的比例划分为训练集，剩余部分作为测试集。随后，在4个数据集上，针对这5个模型计算了精准率、召回率和F1值3个指标，以全面衡量不同模型在推荐效果和精度方面的表现。

5种推荐模型的指标情况见表5。从表5中可

以看出，随着训练集的划分比重不断提高，4种推荐模型的准确率、召回率和F1值在不断提升的；其中ASSIST2009的推荐效果最好，是由于该数据集试题数量少，分布比较集中，便于进行相似度分析，而其他3个数据集题量比较大，分布比较分散，不利于模型计算；本推荐模型（SKT-MFER）在各个数据集上的指标评分最优，其PR指标在每个数据集中都显著高于其他模型，尤其在ASSIST2009数据集上，其3个指标分别达到93.1%、38.5%、54.5%。实验表明SKT-MFER在性能上有所提升。

此外，需要验证TopN中的N为多少最为合适。实验中设置推荐试题数量为5、10、15、20、25、30，通过精准度指标找出最为合适的推荐列表。不同推荐题数下的数据集精准度如图6所示。从图6中可以看出选择N=15即前15道题进行推荐是准确率最高的，这是由于ASSIST2009题数少，每道题分布比较集中，所以ASSIST2009数据集效果最好；而其他两个数据集题量比较大，所以每道题下的学生做题记录比较分散，不利于准确率计算。

表5 5种推荐模型的指标情况

指标	模型	数据集											
		ASST2009			ASST2017			KDD2010			豆豆云		
		60%	70%	80%	60%	70%	80%	60%	70%	80%	60%	70%	80%
PR	UB-CF	45.5%	54.4%	64.8%	43.6%	50.7%	60.9%	35.7%	43.3%	55.7%	43.5%	50.3%	58.5%
	IB-CF	58.8%	63.5%	67.6%	55.8%	60.7%	63.7%	37.2%	46.7%	57.3%	40.2%	47.7%	64.7%
	DKT-CF	81.9%	85.8%	88.2%	77.7%	80.5%	84.7%	74.1%	75.6%	78.6%	75.6%	79.5%	81.9%
	DKVMN-CF	82.1%	86.9%	90.7%	79.5%	82.3%	85.6%	74.9%	77.5%	80.0%	77.1%	80.7%	83.9%
	SKT-MFER	88.9%	90.4%	93.1%	83.3%	85.4%	87.3%	76.5%	81.4%	82.2%	80.2%	83.1%	84.7%
RR	UB-CF	15.4%	20.7%	24.7%	12.7%	15.5%	16.2%	12.1%	12.4%	12.6%	12.6%	13.4%	14.5%
	IB-CF	17.6%	21.6%	27.4%	14.3%	16.5%	18.7%	12.5%	13.6%	14.7%	13.2%	16.5%	17.6%
	DKT-CF	31.1%	33.6%	35.1%	28.8%	29.4%	30.8%	24.7%	26.7%	28.4%	26.5%	29.7%	30.2%
	DKVMN-CF	32.5%	35.4%	37.0%	30.8%	31.2%	31.9%	26.9%	27.7%	29.4%	25.8%	29.1%	31.6%
	SKT-MFER	36.5%	37.8%	38.5%	31.5%	32.1%	33.6%	28.6%	29.1%	30.5%	30.6%	31.1%	32.5%
F1	UB-CF	23.0%	30.0%	35.8%	19.7%	23.7%	25.6%	18.1%	19.3%	20.6%	19.5%	21.2%	23.2%
	IB-CF	27.1%	32.2%	39.0%	22.8%	25.9%	28.9%	18.7%	21.1%	23.4%	19.9%	24.5%	27.7%
	DKT-CF	45.1%	48.3%	50.2%	42.0%	43.1%	45.2%	37.1%	39.5%	41.7%	39.2%	43.2%	44.1%
	DKVMN-CF	47.5%	50.9%	52.3%	43.2%	44.6%	46.5%	39.4%	41.1%	42.6%	41.5%	44.5%	45.3%
	SKT-MFER	51.8%	53.3%	54.5%	45.7%	46.7%	48.5%	41.6%	42.9%	44.5%	44.3%	45.3%	47.0%

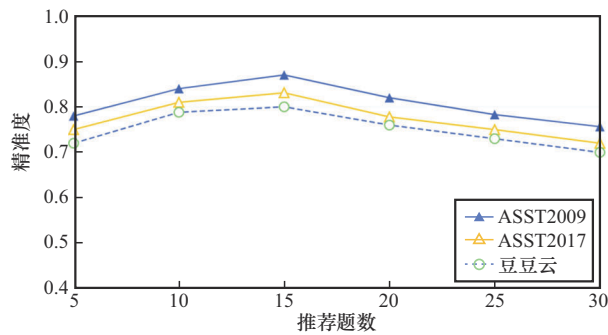


图6 不同推荐题数下的数据集精准率

4 结束语

如何进行个性化习题推荐一直是智慧教育研究领域中的一个重要方面，不断地完善个性化习题推荐模型对于提高学生学习效率具有重要意义。为此，本文首先对学生知识追踪模型方面优化，利用融入遗忘因素的LSTM来学习学生知识情况得出学生学习行为，并考虑基于习题难度的学习能力方面，提高知识追踪模型预测学生知识掌握水平的准确性；在知识追踪训练结果的基础上

与协同过滤进行融合来进行个性化习题推荐，对习题进行两次筛选，初筛通过改进相似度计算出依据相似学生的习题预选集，二次筛选通过相似题目进行最后推荐列表输出。最后在4个数据集上验证了基于KTM多特征融合的习题推荐模型的有效性，能够在一定程度上提高学生习题个性化推荐的准确性。

参考文献:

- [1] 叶俊民, 宋家琦, 张珂, 等. 一种知识图谱增强的在线评测系统习题推荐算法研究[J]. 小型微型计算机系统, 2023(11): 44. YE J M, SONG J Q, ZHANG K, et al. A problem recommendation algorithm for online assessment system based on knowledge map enhancement[J]. Journal of Chinese Computer Systems, 2023(11): 44.
- [2] 诸葛斌, 尹正虎, 斯文学, 等. 基于学生知识追踪的多指标习题推荐算法[J]. 电信科学, 2022, 38(9): 129-143. ZHUGE B, YIN Z H, SI W X, et al. A multi-index problem recommendation algorithm based on student knowledge tracing[J]. Telecommunications Science, 2022, 38(9): 129-143.
- [3] VIE J J, KASHIMA H. Knowledge tracing machines: factoriza-



- tion machines for knowledge tracing[C]//Proceedings of the National Conference on Artificial Intelligence. Association for the Advancement of Artificial Intelligence (AAAI).[S.I.:s.n.], 2019.
- [4] PAVLIK P I, CEN H, KOEDINGER K R. Performance factors analysis - a new alternative to knowledge tracing[C]//Proceedings of the Conference on Artificial Intelligence in Education: Building Learning Systems that Care: from Knowledge Representation to Affective Modelling. Amsterdam: IOS Press, 2009.
- [5] HERLOCKER J L, KONSTAN J A, BORCHERS A, et al. An algorithmic framework for performing collaborative filtering[J]. ACM, 1999.
- [6] CORBETT A T, ANDERSON J R. Knowledge tracing: modeling the acquisition of procedural knowledge[J]. *User Modeling and User-Adapted Interaction*, 4(4): 253-278.
- [7] PIECH C, SPENCER J, HUANG J, et al. Deep knowledge tracing[J]. *Computer Science*, 2015, 3(3): 19-23.
- [8] LEE Y, CHOI Y, CHO J, et al. Creating a neural pedagogical agent by jointly learning to review and assess[J]. 2019.
- [9] LIU Q, HUANG Z, YIN Y, et al. EKT: exercise-aware knowledge tracing for student performance prediction[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2019.
- [10] ZHANG J, SHI X, KING I, et al. Dynamic key-value memory networks for knowledge tracing[J]. *International World Wide Web Conferences Steering Committee*, 2017: 765-774.
- [11] 艾方哲. 基于知识追踪的智能导学算法设计[D]. 北京: 北京交通大学, 2019.
- AI F Z. Design of intelligent tutoring algorithm based on knowledge tracing[D]. Beijing: Beijing Jiaotong University, 2019.
- [12] YANG S, ZHU M, HOU J, et al. Deep knowledge tracing with convolutions[J]. 2020. arXiv preprint arXiv:2008.01169.
- [13] NAKAGAWA H, IWASAWA Y, MATSUO Y. Graph-based knowledge tracing: modeling student proficiency using graph neural network[C]//Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence. New York: ACM Press, 2019.
- [14] GHOSH A, HEFFERNAN N, LAN A S. Context-aware attentive knowledge tracing[C]//Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York: ACM Press, 2020: 2330-2339.
- [15] GUO X, HUANG Z, GAO J, et al. Enhancing knowledge tracing via adversarial training[C]//Proceedings of the 29th ACM International Conference on Multimedia. New York: ACM Press, 2021: 367-375.
- [16] SHEN S H, HUANG Z Y, LIU Q, et al. Assessing student's dynamic knowledge state by exploring the question difficulty effect[C]//Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval(SIGIR '22). New York: ACM Press, 2022: 427 - 437.
- [17] CHEN J H, LIU Z T, HUANG S Y, et al. Improving interpretability of deep sequential knowledge tracing models with question-centric cognitive representation[J] 2023, arxiv.org/abs/2302.06885.
- [18] LIU Z T, LIU Q Q, CHEN J H, et al. Enhancing deep knowledge tracing with auxiliary tasks[C]//Proceedings of the ACM Web Conference 2023 (WWW'23). New York: ACM Press, 2023: 4178-4187.
- [19] HUANG R, LU R. Research on content-based MOOC recommender model[C]//Proceedings of the 2018 5th International Conference on Systems and Informatics(ICSAI). Piscataway: IEEE Press, 2018: 676-681.
- [20] CAMPOS R, DOS S R P, OLIVEIRA J. A recommendation system based on knowledge gap identification in MOOCs ecosystems[C]//Proceedings of the XVI Brazilian Symposium on Information Systems. [S.I.:s.n.], 2020: 1-8.
- [21] 柏茂林. 基于协同过滤的数学学习个性化推荐系统的设计与实现[D]. 锦州: 渤海大学, 2018.
- BAI M L. Design and implementation of personalized recommendation system for mathematical exercises based on collaborative filtering[D]. Jinzhou: Bohai University, 2018.
- [22] 付兴华. 基于协同过滤算法的移动学习系统的设计与实现[D]. 北京: 北京工业大学, 2017.
- FU X H. Design and implementation of mobile learning system based on collaborative filtering algorithm[D]. Beijing: Beijing University of Technology, 2017.
- [23] 吴云峰, 冯筠, 孙霞, 等. 基于多分类器的迁移 Bagging 习题推荐[J]. *计算机应用*, 2013, 33(7): 1950-1954.
- WU Y F, FENG J, SUN X, et al. Recommendation of migrating Bagging problem based on multiple classifiers[J]. *Journal of Computer Applications*, 2013, 33(7): 1950-1954.
- [24] WU D, LU J, ZHANG G. A fuzzy tree matching-based personalized e-learning recommender system[J]. *IEEE Transactions on Fuzzy Systems*, 2015, 23(6): 2412-2426.
- [25] 田保军, 刘爽, 房建东. 融合主题信息和卷积神经网络的混合推荐算法[J]. *计算机应用*, 2020, 40(7): 1901-1907.
- TIAN B J, LIU S, FANG J D. Hybrid recommendation algorithm that combines subject information and convolutional neural network[J]. *Journal of Computer Applications*, 2020, 40(7): 1901-1907.
- [26] LIU J, YIN C, LI Y, et al. Deep learning and collaborative filtering-based methods for student' performance prediction and course recommendation[J]. *Wireless Communications and Mobile Computing*, 2021(9): 1-13.
- [27] 李晓光, 魏思齐, 张昕, 等. LFKT: 学习与遗忘融合的深度知识追踪模型[J]. *软件学报*, 2021, 32(3): 818-830.
- LI X G, WEI S Q, ZHANG X, et al. LFKT: deep knowledge

tracing model integrating learning and forgetting[J]. Journal of Software, 2021, 32(3): 818-830.

[28] 马骁睿, 徐圆, 朱群雄. 一种结合深度知识追踪的个性化习题推荐方法[J]. 小型微型计算机系统, 2020, 41(5): 990-995.

MA X R, XU Y, ZHU Q X. A personalized exercise recommendation method combined with depth knowledge tracing[J]. Journal of Chinese Computer Systems, 2020, 41(5): 990-995.

[作者简介]



葛斌 (1976-), 男, 博士, 浙江工商大学信息与电子工程学院教授, 主要研究方向为网络和通信技术、互联网技术和网络安全。



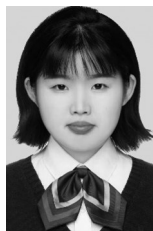
汪盈 (2000-), 女, 浙江工商大学信息与电子工程学院硕士生, 主要研究方向为智慧教育和个性化推荐。



肖梦凡 (2001-), 女, 浙江工商大学信息与电子工程学院硕士生, 主要研究方向为智慧教育和个性化推荐。



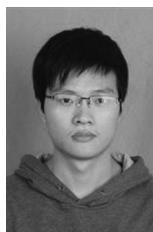
颜蕾 (1996-), 女, 浙江工商大学信息与电子工程学院硕士生, 主要研究方向为教育数据挖掘、深度学习和机器学习。



王冰雁 (2000-), 女, 浙江工商大学信息与电子工程学院硕士生, 主要研究方向为智慧教育和数字水印。



董黎刚 (1972-), 男, 博士, 浙江工商大学信息与电子工程学院教授, 主要研究方向为智能网络、在线教育。



蒋献 (1988-), 男, 浙江工商大学信息与电子工程学院讲师、实验员, 主要研究方向为在线教育。