



智能系统学报

CAAI TRANSACTIONS ON INTELLIGENT SYSTEMS

面向智能教育的自适应学习关键技术与应用

陈恩红, 刘淇, 王士进, 黄振亚, 苏喻, 丁鹏, 马建辉, 竺博

引用本文:

陈恩红, 刘淇, 王士进, 等. 面向智能教育的自适应学习关键技术与应用[J]. 智能系统学报, 2021, 16(5): 886–898.

CHEN Enhong, LIU Qi, WANG Shijin, et al. Key techniques and application of intelligent education oriented adaptive learning[J]. *CAAI Transactions on Intelligent Systems*, 2021, 16(5): 886–898.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202105036>

您可能感兴趣的其他文章

基于知识图谱和用户长短期偏好的个性化景点推荐

Personalized attraction recommendation based on the knowledge graph and users' long-term and short-term preferences
智能系统学报. 2020, 15(5): 990–997 <https://dx.doi.org/10.11992/tis.201904064>

知识图谱的推荐系统综述

Review of recommendation systems based on knowledge graph
智能系统学报. 2019, 14(2): 207–216 <https://dx.doi.org/10.11992/tis.201805001>

面对智能导诊的个性化推荐算法

A personalized recommendation algorithm for intelligent guidance
智能系统学报. 2018, 13(3): 352–358 <https://dx.doi.org/10.11992/tis.201711036>

智能交互的物体识别增量学习技术综述


Incremental learning and object recognition system based on intelligent HCI: a survey
智能系统学报. 2017, 12(2): 140–149 <https://dx.doi.org/10.11992/tis.201701006>

基于事件驱动的多智能体强化学习研究

Reinforcement learning for event-triggered multi-agent systems
智能系统学报. 2017, 12(1): 82–87 <https://dx.doi.org/10.11992/tis.201604008>

基于用户移动轨迹的个性化健康建议推荐方法

Personalized recommendation algorithm of health advice based on the user's mobile trajectory
智能系统学报. 2016, 11(2): 264–271 <https://dx.doi.org/10.11992/tis.201511026>

 微信公众平台



关注微信公众号, 获取更多资讯信息

吴文俊人工智能科技进步奖一等奖

成果名称：面向智能教育的自适应学习关键技术与应用

获 奖 人：陈恩红、王士进、刘淇、丁鹏、黄振亚、苏喻、马建辉、竺博、付瑞吉、张丹、沙晶、汪洋、朱林波、何春江

完成单位：中国科学技术大学、科大讯飞股份有限公司



陈恩红

中国科学技术大学教授，现任中国科学技术大学大数据学院执行院长兼计算机科学与技术学院副院长、中科大智慧城市研究院（芜湖）院长、CAAI会士、CCF会士。科技部重点领域创新团队“大数据分析及应用团队”负责人、大数据分析与应用安徽省重点实验室主任、中国计算机学会大数据专家委员会副主任、安徽省计算机学会理事长。担任IEEE TKDE、ACM TIST、计算机研究与发展等多个国内外学术期刊编委，多次担任KDD、AAAI、ICDM等重要国际学术会议的程序委员会委员，担任ICKG2020、CCDM 2015、CNCC 2014等的程序委员会主席以及BDTC2019大会主席等。作为项目负责人，承担了国家自然科学基金重大科研仪器研制项目、国家杰出青年基金项目、面上项目、联合重点基金项目，以及863计划、科技部重点研发计划课题等项目，以及与科大讯飞、华为、今日头条、阿里巴巴、美团等的合作项目。主要研究领域包括：数据挖掘、社会网络、推荐系统、教育大数据分析等。在国内外重要学术期刊TKDE、TOIS、TIST等和国际学术会议KDD、AAAI、WWW、SIGIR、NIPS等发表学术论文150余篇，获KDD2008最佳应用论文奖、ICDM2011最佳研究论文奖、SDM2015最佳论文提名奖、KDD2018最佳学生论文奖等。获得教育部自然科学一、二等奖各一项，多次获得中科院优秀导师奖、中科院朱李月华优秀教师奖。

团队简介

项目组成员来自中国科学技术大学和科大讯飞股份有限公司，依托双方单位共建的语音及语言信息处理国家工程实验室、大数据分析与应用安徽省重点实验室以及科大讯飞AI研究院、教育事业群。项目组织积极响应国家对教育强国的号召，在科技部、教育部、国家自然科学基金委等的项目资助下，经过十余年协同攻关，最终完成了面向智能教育的自适应学习关键技术与应用项目研究工作。团队成员共14人，含中国科学技术大学4人（陈恩红、刘淇、黄振亚、马建辉），科大讯飞AI研究院6人（王士进、竺博、付瑞吉、沙晶、汪洋、何春江）和科大讯飞教育事业群4人（丁鹏、苏喻、张丹、朱林波）。项目执行期间，中国科大团队主要负责项目的核心技术研发，科大讯飞AI研究院团队主要负责项目技术转化，科大讯飞教育事业群团队主要负责项目系统的开发和应用推广。项目团队成员间紧密配合，突破了多个技术难点，构建了教学资源表示新框架，提出学习者认知诊断新方法，设计了自适应学习推荐新技术，研发了面向基础教育的智能教育系统“智学网”并已在我国全国范围内广泛应用，产生了非常显著的经济效益和社会效益。

DOI: 10.11992/tis.202105036

网络出版地址: <https://kns.cnki.net/kcms/detail/23.1538.TP.20210713.1348.005.html>

面向智能教育的自适应学习关键技术与应用

陈恩红¹, 刘淇¹, 王士进², 黄振亚¹, 苏喻^{1,2}, 丁鹏², 马建辉¹, 竺博²

(1. 中国科学技术大学 计算机科学与技术学院 大数据分析与应用安徽省重点实验室, 安徽 合肥 230027; 2. 科大讯飞股份有限公司, 安徽 合肥 230088)

摘要: 本文是关于我们获得 2020 年度吴文俊人工智能科学技术奖主要工作的一个介绍。该成果针对自适应学习中面临的教学资源表示困难、学习状态诊断困难以及学习策略设计困难等关键技术难题, 首先构建数据驱动的教学资源无监督表示新框架, 提高了教学资源质量评估和内容检索的精度和效率。其次提出基于深度学习的学习者认知诊断新方法, 突破了以量表为基础的教育测量理论研究范式。然后设计基于知识匹配的个性化推荐技术以及多目标匹配的自适应推荐技术, 满足了智能教育场景的复杂约束与学习者的多样目标需求。最后, 本文成果研发了面向基础教育的智能教育系统——智学网, 已在全国推广使用, 对我国智能教育发展具有积极意义。

关键词: 自适应学习; 智能教育; 教学资源表示; 质量评估; 内容检索; 认知诊断; 知识追踪; 个性化推荐; 自适应推荐; 智能教育系统

中图分类号: TP18 **文献标志码:** A **文章编号:** 1673-4785(2021)05-0886-13

中文引用格式: 陈恩红, 刘淇, 王士进, 等. 面向智能教育的自适应学习关键技术与应用 [J]. 智能系统学报, 2021, 16(5): 886-898.

英文引用格式: CHEN Enhong, LIU Qi, WANG Shijin, et al. Key techniques and application of intelligent education oriented adaptive learning[J]. CAAI transactions on intelligent systems, 2021, 16(5): 886-898.

Key techniques and application of intelligent education oriented adaptive learning

CHEN Enhong¹, LIU Qi¹, WANG Shijin², HUANG Zhenya¹, SU Yu^{1,2},
DING Peng², MA Jianhui¹, ZHU Bo²

(1. Anhui Province Key Laboratory of Big Data Analysis and Application, School of Computer Science and Technology, University of Science and Technology of China, Hefei 230027, China; 2. iFLYTEK Co., Ltd, Hefei 230088, China)

Abstract: This paper is an introduction to our main work in winning of the 10th Wu Wenjun Artificial Intelligence Science and Technology Award. In response to key technical problems in adaptive learning, such as the difficulties in the representation of teaching resources, the diagnosis of knowledge state and the design of learning strategies, we first constructed a new framework for data-driven unsupervised representation of teaching resources, which significantly improves the accuracy and efficiency of quality evaluation and content retrieval of teaching resources. Then proposed a new method for learner cognitive diagnosis based on deep learning, which breaks through the scale-based research paradigm of educational measurement theories. And further designed a personalized recommendation technology based on knowledge matching and an adaptive recommendation technology based on multi-objective matching, meeting the complex constraints of intelligent education scenarios and the diverse learning targets of learners. Finally, an intelligent education system—“Zhixue.com” was developed for basic education, which has been popularized and used throughout the country. This system plays a positive role in the development of China’s intelligent education.

Keywords: adaptive learning; intelligent education; teaching resources representation; quality evaluation; content retrieval; cognitive diagnosis; knowledge tracing; personalized recommendation; adaptive recommendation; intelligent education system

收稿日期: 2021-05-25. 网络出版日期: 2021-07-13.

基金项目: 国家自然科学基金项目 (61922073, U20A20229, 61727809).

通信作者: 刘淇. E-mail: qiliuql@ustc.edu.cn.

党的十九大报告中, 习近平总书记将“科教兴国”确定为决胜全面建成小康社会需要坚定实施的七大战略之一, 并指出“建设教育强国是中华民

族伟大复兴的基础工程”。发展能够融合大数据和人工智能技术的智能教育则是建设教育强国的必由之路^[1]。作为智能教育的核心,自适应学习旨在根据学习者个体的认知水平、学习能力等,动态调整相适应的教学资源与学习方式,提升学习者的学习效率,使其获得最佳发展。早在春秋末期,我国古代著名教育家孔子就开创了“因材施教,有教无类”的全新教育理念。这一理念也是自适应学习思想的开端,即使在两千余年后的今天依然指导着教育现代化发展的方向。遵循这一思想,联合国教科文组织在《教育2030行动纲领》中将教育质量的提升作为未来十几年内全世界教育改革发展的关键之一。与此同时,我国各级政府与教育主管部门也制定了多项教育政策来促进智能教育和自适应学习的发展。2019年,中共中央、国务院印发《中国教育现代化2035》,指出“利用现代技术加快推动人才培养模式改革,实现规模化教育与个性化培养的有机结合”。同时印发的《加快推进教育现代化实施方案(2018-2022年)》则具体指出“促进信息技术与教育教学深度融合,加快推进智慧教育创新发展”。

在传统的课堂教学模式下,由于时间和空间的限制,教师无法为每一个学生都制定自适应的学习方案。近年来,伴随着教育现代化程度的不断加深,互联网技术涌入到了教育教学的各个环节当中,新兴的在线学习模式方兴未艾^[2],国内外涌现出一系列优秀的在线学习系统,例如:Coursera、智学网、学堂在线等。特别地,在2020年初,受新型冠状病毒引发的肺炎疫情影响,约2.5亿学生利用网络、电视等在线工具学习知识,在线教育平台保障了正常的教育学习秩序。目前,我国在线教育用户规模已达4.23亿,规模空前^[3]。在线教育模式已经成为了教育现代化的重要组成部分,它突破了传统课堂教学中的时空约束限制,实现了优质教学资源共享,使得融合大数据和人工智能等先进技术来促进智能教育的发展成为了可能。

智能教育的核心之一是自适应学习,它以数据为基础,利用相关理论与方法,旨在实现更加精准、规模化的因材施教的教育目标。尽管在线教育模式为真正实现自适应学习带来了新的机遇,但其市场高度自由和开放的特点还无法满足智能教育的要求^[4]。一方面,尽管在线教育平台提供了开放共享、颇具规模的教学资源,但这些资源大多缺乏有效标记,使得自适应学习面临教学资源表示困难的技术难题^[5-7]。另一方面,由于学习者的学习状态隐蔽易变,如何精准地诊断每

一个学习者的认知状态是又一技术难题。此外,学习目标的复杂多样性也对自适应学习的学习策略设计造成了很大的困难。因此,研究自适应学习关键技术具有重要意义。

1 本文研究介绍

总结起来,自适应学习的关键是要解决教学过程中教学资源的理解难、学习者知识的诊断难、学习策略的设计难等三大难题。针对上述关键技术难题,本文成果构建了教学资源表示新框架,提出了学习者认知诊断新方法,设计了自适应推荐新技术,研发了面向基础教育的智能教育系统——智学网,已在全国推广使用,对我国智能教育发展具有积极意义。本文成果的研究应用方案如图1所示。具体来说,针对智能教育涉及的3类主要对象,即教学资源、学习者和学习策略,本文成果分别提出系列模型算法以解决现有分析技术面临的困难。首先,针对海量教学资源多源异构、语义丰富、标记缺失等特点导致的教学资源理解难题,本文成果构建了一种数据驱动的教学资源无监督表示新框架,提出“深度表征-精准评估-高效检索”的教学资源表示和应用新模式,揭示教学资源与知识空间的深层关系,解决专家评估操作低效、主观有偏的问题,提高了教学资源质量评估和内容检索的精度和效率。其次,针对学习者学习活动自由、状态隐蔽易变等特点带来的学习状态诊断难题,本文成果提出一系列基于深度学习的学习者认知诊断新方法,通过直接挖掘学习者的作答表现,阐明了“学习者-知识/兴趣-资源”的高阶交互和潜在关联,突破以量表为基础的教育测量理论研究范式,增强学习者认知建模的可解释性。最后,针对自适应学习推荐任务中场景多样、目标复杂等特点导致的学习策略设计困难,基于上述建立的教学资源表示框架和提出的学习者认知诊断方法,本文成果设计了基于知识匹配的个性化推荐新技术以及多目标匹配的自适应推荐新技术,分别解决了在静态学习场景和交互式学习场景中,如何将教学资源质量与学习者认知状态进行智能匹配的难题。进一步,本文成果集成上述关键技术难题,攻关所研发的创新方法,研发了面向基础教育的智能教育系统——“智学网”,已在全国推广使用。“智学网”通过伴随式地采集日常“教与学”过程性的动态数据,建立了面向中小学的大规模高质量教学资源库,构建了以学习者为中心的学业评价体系,实现了优质资源精准推荐,提供了智能环境与服务。

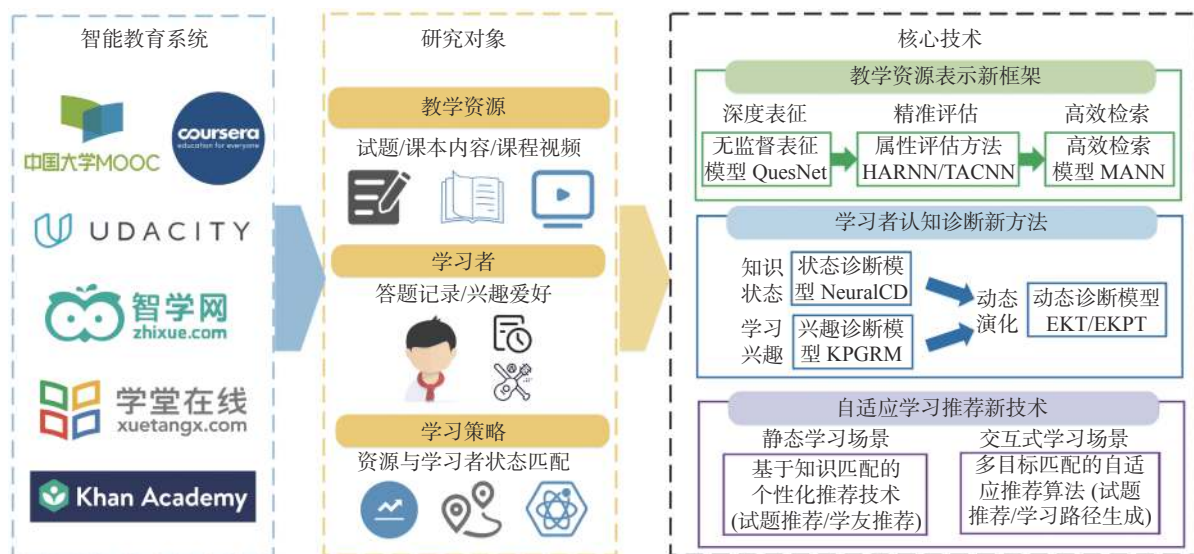


图1 研究总体介绍

Fig. 1 General introduction of the research

2 教学资源表示新框架

教学资源的理解是发展智能教育系统的基础。针对海量教学资源多源异构、语义丰富、标记缺失等特点导致的教学资源理解难题,本文成果构建了一种数据驱动的教学资源无监督表示新框架,提出“深度表征-精准评估-高效检索”的教学资源表示和应用新模式,揭示了教学资源与知识空间的深层关系,解决专家评估操作低效、主观有偏的问题,显著提高了教学资源质量评估和内容检索的精度和效率。

2.1 基于语义理解的教学资源无监督表征框架

教学资源一般由文本、图像和语音等多种不同类型的数据组成,具有明显的多源异构特征。此外,教学资源通常基于一定的教学目的编写,如用于评估或提高学习者对于特定知识或技能的掌握水平,因此往往蕴含着丰富的语义信息,并且具有更深层次的逻辑知识特征,以数学试题为例,计算题通常与代数知识相关,几何证明题则更关注几何知识,而传统自然语言理解方法难以精准建模这些逻辑特征^[8-10]。此外,教学资源的难度、知识点、相似资源等信息的标注对专家的经验具有较高要求,导致了教学资源的标签数据有限,标签缺失现象严重,具有明显的数据稀疏性特征。如何对多源异构、语义丰富且标签稀疏的教学资源进行充分理解与刻画,建立统一的数学表征,在教学资源的组织与自动标注、教学资源与学习者学习需求的自适应匹配等实际应用中具有重要价值,是教学资源表示技术研究中亟需解决的关键问题。

为解决上所述问题,本文成果提出一种多源

异构教学资源的统一表征框架——QuesNet。QuesNet设计了教学试题(一类典型的教学资源)中异构信息的统一表征方法,提出结合低层次语义和高层次逻辑知识的表征方法,利用无监督的预训练技术实现对海量无标签教学试题的深度理解和表征^[11]。

图2是QuesNet框架示意^[11]。该框架主要由3个部分构成,分别是嵌入层、内容表示层以及语义学习层。具体地,在嵌入层,QuesNet通过不同的嵌入表征模块将试题(如几何题)不同类型的输入信息(文本、图像以及元信息)映射为等长的数学向量表示,通过映射表将词映射为嵌入向量,通过卷积和池化层将图像编码为数学向量,通过全连接神经网络将元信息(如章节、知识点等)编码为特征向量。嵌入层将不同类型的试题内容映射到统一的数学空间,使模型能处理不同类型的输入。在内容表示层,QuesNet使用双向长短期记忆网络(long short-term memory networks, LSTM)建模题目内容局部性特征,捕获题目的深层次上下文语义信息与内容关联,生成试题内容的表征序列。最后,在语义学习层,在建模低层次语言特征的基础上,QuesNet通过多头自注意力机制聚合试题的表征序列,生成统一的试题向量,挖掘复杂的全局关联,捕获资源的高层次逻辑与知识特征。

QuesNet通过无监督任务在海量无标签试题数据上预训练模型。在嵌入层,QuesNet利用Word2Vec工具从数据中生成试题词嵌入向量的映射表,通过自编码器的重建损失函数预训练图像(如几何图形)与元信息如知识点的编码模

型。在内容层,通过掩码语言模型进行预训练,随机掩盖或替换部分输入内容,以预测被掩盖的部分作为预训练任务,使内容层模型能够从大量训练数据中学习建模资源的语言特征。在语义学习层,以与领域特定的逻辑和知识相关的任务进行模型预训练,在选择题中,以预测选项是否为问题的正确答案作为预训练任务,使模型能够捕

获资源的高层次逻辑与知识特征。本文成果将 QuesNet 学习试题表征向量应用于如知识点预测、难度预测等多个基于教学资源的应用任务中。特别地,以知识点预测任务为例,如表 1 所示, QuesNet 模型的准确率 (ACC) 指标较现有预训练方法相对提升近 10%, F_1 指标较现有方法也有 5% 以上的提升。

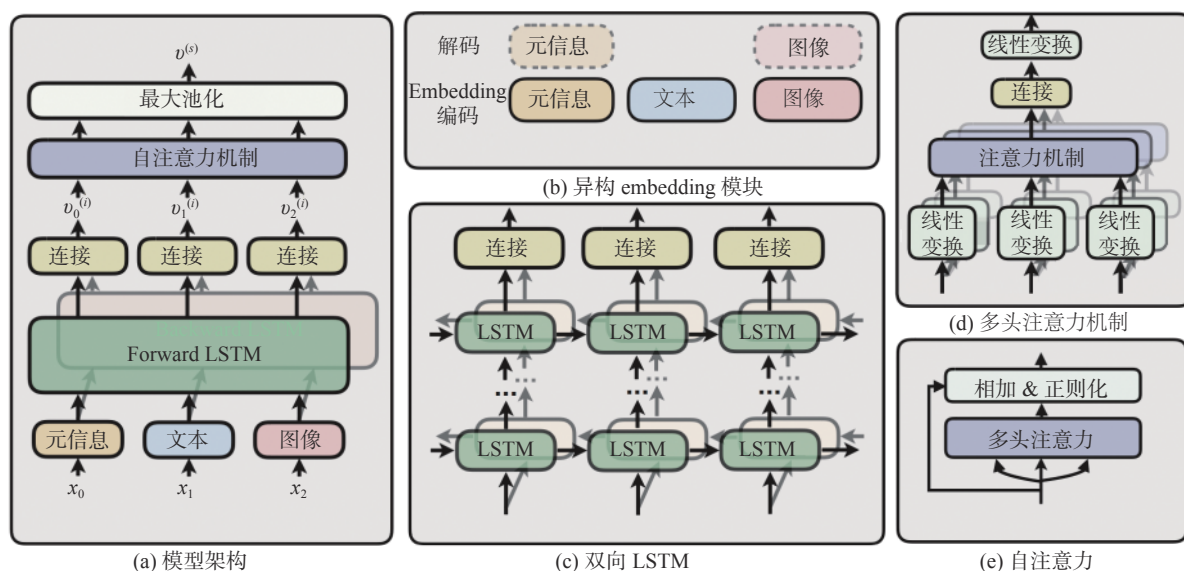


图 2 QuesNet 框架示意

Fig. 2 Overview of QuesNet framework

表 1 知识点映射任务实验结果

Table 1 Experimental result of knowledge mapping

模型	准确率	精确率	召回率	F-1分数
Original	0.574 4	0.414 7	0.787 2	0.543 2
ELMo	0.694 2	0.796 0	0.768 5	0.782 0
BERT	0.622 4	0.732 6	0.671 1	0.700 5
H-BERT	0.626 1	0.760 8	0.691 1	0.724 3
QuesNet	0.774 9	0.865 9	0.807 5	0.835 7

2.2 数据驱动的教学资源质量精准评估方法

评估教学资源的质量,如试题、课程的难度与知识点,不仅为教学资源的表征方法提供一种有效的验证方案,也为教学资源的检索与匹配等相关应用提供了支撑。难度与知识点是试题等教学资源的重要属性,在资源检索、测试评估与个性化推荐等应用中具有重要的意义。例如,在线教育平台中经常会根据学生在练习上的作答情况,检索具有特定难度与知识点的试题以巩固或加强训练。在能力评估测试等允许学生重复参加且选取最高成绩的测试中,每一次测试的试题需要具有基本一致的难度,以保证成绩的稳定性与公平性^[12]。在个性化推荐应用中,难度与知识点

属性具有重要的价值,通常情况下,难度过低或过高的练习都无法使学习者提高知识掌握程度,而知识点属性不合适的练习无法使学习者进行针对性的提高。因此,难度和知识点等教学资源质量的精准评估是一个具有重要价值的研究课题。

传统方法通过专家对教学资源(如试题)的质量进行标注。然而,对大量的教学资源进行标注需要花费大量的时间与人力,并且对标注人员的专业经验具有较高要求,难以通过传统的众包等方法进行大范围推广。此外,教学资源的质量标准难以通过具体标准准确给出,评估结果依赖专家经验,由不同专家评估的结果不统一,具有明显的强主观性特征。为了解决以上瓶颈问题,本文成果提出一种数据驱动的教学资源质量自动化评估方法,构建了资源属性预测模型,在知识结构多层依赖关系刻画的基础上,准确预测教学资源的属性特征,形成了高效且有数据依据的资源质量评估新模式。

首先,本文成果针对教学资源的知识点属性评估任务进行研究。教学资源具有多个知识点标签,且多个知识点具有层级组织结构,如“代数”、“函数”、“一次函数”。教学资源的多个知识点标签分布于层级结构的不同层上,不同层间的标签具

有依赖关系。针对这一特征,本文成果提出 HARNN (hierarchical attention-based recurrent neural network) 框架结合教学资源文本与知识点的层级结构,提出基于层级注意力的循环层以自顶向下的方式建模不同层间的依赖关系,并设计混合方法准确预测各层的标签,在层级结构上逐层预测与资源最相关的知识点标签。在试题多标签知识点预测任务上, HARNN 框架的正确率较现有方法

具有明显提高^[13]。

在教学试题的难度属性评估任务中,本文成果针对英语阅读理解试题提出如图3所示的 TACNN (test-aware attention-based convolutional neural network) 框架进行难度评估^[14]。TACNN 框架能够结合试题的阅读材料、问题与选项等不同部分文本的语义进行试题难度预测,并创新性地提出了测试感知的训练方法以解决不同测试的难度偏差。

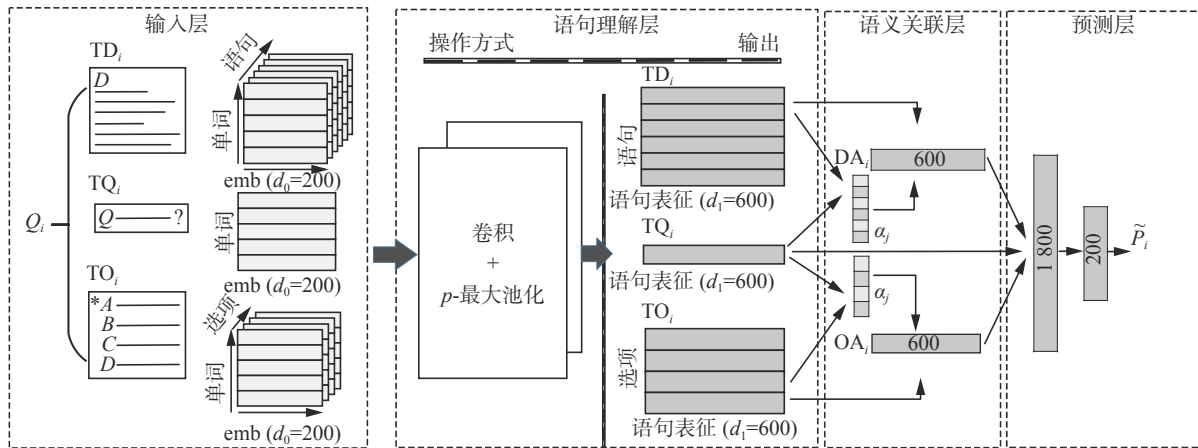


图3 TACNN 框架示意

Fig. 3 Overview of TACNN framework

具体地, TACNN 框架主要由输入层、语句理解层、语义依赖层与预测层4部分构成。输入层将试题的阅读材料作为语句序列,问题以及每个选项作为单独的语句,将每个语句中的单词映射为固定长度的嵌入向量表示。语句理解层通过卷积层和最大池化层以从局部到全局的方式捕获每个语句中的关键信息,从词嵌入向量中学习语句表征,将每个语句表示为固定长度的向量。语义依赖层针对阅读材料中不同语句对问题贡献不同的特点,通过加权求和的方式聚合语句表示向量,生成阅读材料和选项的注意力向量表征,以余弦相似度作为阅读材料和选项中每个句子对问题的重要性的度量。预测层拼接阅读材料和选项的注意力向量以及问题的向量表征,通过全连接

神经网络学习试题的难度表征并预测问题的难度值。

由于试题难度标签无法直接获得, TACNN 依据试题的学生作答平均得分作为试题的难度属性进行模型训练。由于参与不同测试的学生不同,不同测试的试题难度无法直接比较,但同一测试的试题难度可以直接比较,针对这一问题, TACNN 提出测试依赖的损失函数训练模型,而训练后的模型仅需要试题的文本信息对试题难度属性进行评估。在英语阅读理解试题难度预测任务上, TACNN 较现有方法在皮尔逊相关系数指标上平均提升约10%,如表2^[14]所示,与专家评估结果对比, TACNN 预测难度的稳定性(平均方差 std 小)更高。

表2 TACNN 与专家结果对比

Table 2 Comparison between TACNN and experts

测试	TACNN	专家平均	专家1	专家2	专家3	专家4	专家5	专家6	专家7
T_1	0.41	0.21	0.18	0.13	0.38	-0.08	-0.04	0.01	0.14
T_2	0.63	0.68	0.45	0.32	0.52	-0.01	-0.44	0.53	0.37
T_3	0.78	0.70	0.52	0.63	0.28	0.44	-0.29	0.45	0.52
T_4	0.63	0.40	-0.09	0.07	0.31	0.48	-0.40	0.58	-0.08
T_5	0.53	0.56	0.39	0.32	0.29	0.29	0.43	0.51	0.47
T_6	0.47	0.22	0.21	0.01	0.27	-0.23	0.10	0.24	0.17

续表 2

测试	TACNN	专家平均	专家1	专家2	专家3	专家4	专家5	专家6	专家7
T_7	0.81	0.73	0.58	0.29	0.72	0.72	0.70	0.59	0.69
T_8	0.77	0.45	0.35	0.45	0.24	0.14	0.19	0.45	0.64
T_9	0.81	0.55	0.25	0.54	0.35	0.53	0.13	0.32	0.36
T_{10}	0.76	0.57	0.49	-0.13	0.72	0.25	0.22	0.32	0.60
T_{11}	0.90	0.77	0.44	0.57	0.59	0.41	0.36	0.08	0.83
T_{12}	0.60	0.62	0.59	0.73	0.60	0.54	0.48	0.62	0.54
平均值	0.68	0.54	0.36	0.33	0.44	0.29	0.12	0.39	0.44
标准差	0.14	0.18	0.19	0.26	0.17	0.27	0.34	0.19	0.25

2.3 大规模教学资源的高效检索技术

现有在线教育平台一般具有大量的用户以及海量的教学资源,不同用户的学习需求存在巨大的差异,满足其需求的教学资源也各不相同,因此为了支撑大规模教学资源的智能应用,需要解决“如何精准检索满足学习需求的资源”以及“如何从海量资源中高效获得检索结果”两个关键问题。

传统检索技术仅基于教学资源的文本内容进行检索^[15-18],无法利用教学资源中的图像、知识等非文本内容,难以理解学习者的检索意图,难以从海量教学资源中精准检索满足学习者需求的资源,导致检索的精度差、效率低等不足。

针对以上问题,本文成果分别提出基于教学资源语义表征的相似资源检索方法和高效检索技术,不仅能够利用教学资源异构信息的内在含义,还能充分理解学习者的检索意图,进而提高教学资源表示的效率和资源检索的精度。在相似试题检索任务中,本文成果提出 MANN(multimodal attention-based neural network) 框架通过从异构数据中学习统一的语义表征进而在大规模教学资源中精准检索与目标试题相似的教学试题^[19]。MANN 通过卷积神经网络编码图像数据,通过嵌入模型编码知识点数据,从而利用试题的文本、图像、知识点的异构信息,通过基于注意力的长短期记忆网络从异构信息中学习统一的语义表征,通过注意力机制捕获试题的文本与图像、知识间的内在联系,并通过相似注意力评估目标试题与候选试题的相似部分,提高方法的可解释性。在相似试题检索任务中, MANN 的检索精度较现有方法相对提升约 30%,并且检索结果具有较好的可解释性。

3 学习者认知诊断新方法

学习者作为学习的主体,分析其学习状态具有重要的意义。然而,由于学习者学习活动自由、状

态隐蔽易变等特点,对其学习状态的诊断是非常困难的。为此,本文成果提出了基于深度学习的学习者认知诊断新方法,通过直接挖掘学习者的作答表现,阐明了“学习者-知识/兴趣-资源”的高阶交互和潜在关联,突破了以量表为基础的教育测量理论研究范式,增强了学习者认知建模的可解释性,为个性化学习推荐提供理论和技术支撑。

3.1 基于深度学习的认知诊断通用框架

学习者的知识状态是影响其表现的根本因素,学习者答对已掌握的知识相关的试题的可能性很高,而答对未掌握的知识相关的试题的可能性则很低,因此准确诊断学习者的知识状态是认知诊断的核心目标,同时也是自适应学习研究的关键课题。然而,学习者、知识、教学资源三者间存在复杂的交互关系。具体来说,学习者对不同的知识有着不同的掌握状态,这直接影响学习者与相关的教学资源的交互结果。学习者的知识状态也会在其与相关教学资源的交互中不断更新。此外,不同知识间、教学资源与知识间都存在多种多样的复杂关联。学习者、知识、教学资源三者间的复杂交互关系与学习者的认知诊断结果密切相关。而现有的教育测量理论研究大多基于 IRT^[20]或 DINA^[21]等传统认知诊断模型,需要人工设计学习者的认知特征并定义学习者与教学资源的交互关系,通常使用简单的线性关系对其进行建模,模型表达能力弱,难以捕获“学习者-知识-资源”间的复杂交互关系。因此如何建模这种交互关系一直是现有认知诊断技术的瓶颈。针对这一问题,本文成果首次构建了基于深度学习的通用认知诊断框架 NeuralCD(neural cognitive diagnosis),通过伴随式地挖掘学习者在各类试题上的作答表现,利用深度学习技术自动发现学习者与试题间的复杂交互关系,建立“学习者-知识-资源”之间的高阶交互和潜在关联。通过利用深度神经网络强大的特征学习能力,NeuralCD 模型突破了以人

工量表为基础的教育测量理论研究范式表达能力弱的缺陷,提高了对学习者的建模能力。

如图4^[22]所示,NeuralCD框架考虑认知诊断过程中的3个因素:学习者因素、试题因素与两者的交互。对每条作答记录,NeuralCD通过独热编码对应的学习者与试题作为输入并获得两者的诊断因素。学习者因素建模学生的特征,通过连续向量建模学习者在各知识上的掌握程度;试题因素建模试题的特征,并分为与知识相关的特征以及与知识无关的特征两种,前者是知识试题与知识概念间的关联,并与学习者因素对应,通过非负向量建模试题与每个知识概念的相关程度,而后者并非认知诊断所必需但可以与其结合,如知识的难度与试题难度、区分度等。在此基础上,交互层通过神经网络学习两个诊断因素的交互函数并输出学习者正确作答试题的概率。进一步,为了保证诊断结果的可解释性,NeuralCD引入单调性假设在训练中约束交互函数,即学习者正确作答试题的概率随其知识掌握度向量任一维的提高单调上升。经过模型训练,以学生的知识掌握度向量作为诊断结果。

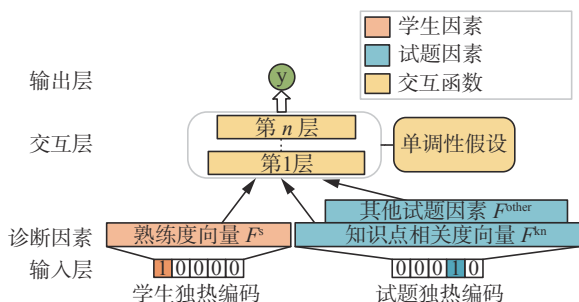


图4 NeuralCD框架示意

Fig. 4 Overview of NeuralCD framework

NeuralCD是一个泛化的通用认知诊断框架,在该框架下,许多传统认知诊断模型可以视为NeuralCD框架的特例,如MF^[23]、IRT、MIRT^[24]及其变种。例如设定学习者诊断因素为一维、试题诊断因素为一维以及交互函数为sigmoid函数,NeuralCD可以退化为传统IRT模型。因此NeuralCD框架统一了多种传统教育测量学技术。更进一步,本文成果基于NeuralCD框架实现了一个具体的认知诊断模型NeuralCDM。如表3^[22]所示,在学习者认知诊断任务上,误差较现有方法相对降低约5%。

3.2 多角度的学习者学习兴趣模型

学习者的学习兴趣是影响其学习表现的潜在因素,因此,如何识别学习者的学习兴趣是认知诊断的重要任务。然而,学习兴趣是个人隐性特征,难以直接测量,且影响因素众多。为此,本文成果提出了“学习者-兴趣-资源”的多角度建模方

法,从学习者的“同质化特点”、“长短期行为”和“心理活动”等3个影响学习兴趣的重要因素着手,数值化建模不同因素对于学习者兴趣的影响程度,增强学习兴趣建模的可解释性。为研究“游戏行为”(gaming behavior)等行为活动对学习者的学习兴趣的影响,本文提出了KPGRM(knowledge plus gaming response model)模型,对在线学习中的学生“游戏行为”进行诊断和分析^[25-26]。针对学生答题过程中不同的答题行为,KPGRM通过分析学生在多次答题(multiple-attempt responses, MAR)下的显式“游戏行为”因子和单次答题(one-attempt response, OAR)下的显式猜测因子和隐式猜测因子,建模并诊断学生的“游戏行为”以及相对应的认知情况。

表3 认知诊断任务实验结果

Table 3 Experimental result of cognitive diagnosis

模型	准确率	均方根误差	AUC
DINA	0.650±0.001	0.467±0.001	0.676±0.002
IRT	0.674±0.002	0.464±0.002	0.685±0.001
MIRT	0.701±0.002	0.461±0.001	0.719±0.001
PMF	0.661±0.002	0.476±0.001	0.732±0.001
NeuralCDM	0.719±0.008	0.439±0.002	0.749±0.001

在现有大多数研究假设的一种在线选择题答题形式下,学生只有答对当前试题才能进行下一题的作答。在该形式下,学生的可能存在的“游戏行为”主要可以分为3类:持续作答、系统作答和快速作答,KPGRM基于这3种学生“游戏行为”提出了4个用于诊断显式的学生“游戏行为”的特征:1)持续作答(keep answering):在一道题上的答题尝试次数越多,则对应的“游戏行为”因子越高。2)快速作答(quickly):答题速度越快,则“游戏行为”越强。3)系统作答(systematically):答题转换越多,系统性越差,则“游戏行为”越多。4)选项覆盖度(coverage):选择的选项对所有选项的覆盖度越大,都标志着更高的“游戏行为”因子。基于学生的答题特征观测,KPGRM对这4个特征进行线性组合,对学生的显式“游戏行为”因子进行分析。具体来说,KPGRM模型通过学生多次回答问题时的特征观测其显式“游戏行为”因子,并使用协同过滤的方法推断隐式“游戏行为”因子。在结合学生“游戏行为”因子的情况下,使用信号检测模型来进行学生学习状态的诊断。KPGRM在分析学生在线答题时“游戏行为”的同时,能更加精确地诊断学生的认知状态,具体内容可参考文献[26]。

3.3 构建认知状态的动态演化模型

学习者的知识状态与学习兴趣在长期的学习

过程中动态变化,这给学习者认知诊断演化建模提出了更高的要求^[27]。为此,本文成果分别从教学资源属性特征和人脑记忆特性两个角度,量化了学习过程中学习者的认知状态变化过程,基于马尔科夫等性质提出了融合教学资源表征和记忆曲线理论的时序建模方案,揭示了学习者知识状态和学习兴趣的动态演化规律。

本文成果提出一种称为 exercise-aware knowledge tracing(EKT)^[28]的动态知识跟踪框架,该框架通过题目语义来建模教学资源特征,进而分析其对学习者认知状态的影响。对比传统的认知状态动态演化模型,EKT还考虑了两类额外信息对于学习过程的影响,分别是题目中包含的知识信息和每个题目的个性化信息。具体来说,模型输入学生历史上做过的题目序列,输出学习者的认知状态预测。对于题目中包含的知识信息,模型设计了一种动态增强记忆网络,以存储题目之间

的知识共性,并量化题目知识对学生知识概念学习的影响。对于每个题目的个性化信息,模型设计了基于双向循环神经网络的题目特征提取器,可以自动学习题目的个性化特征。然后,模型分别对每一个时刻的共性和个性两种特征信息进行融合,利用多维 LSTM 网络追踪学习者在多个知识概念下的状态变化^[10]。最后,本文成果提出两个 EKT 的实例化模型,模型框架如图 5^[28]所示。其中图 5(a)所示的 EKTM(EKT with Markov property)假设学生的未来表现仅与当前时刻的知识状态相关,即对学生的认知动态演化建立一阶马尔科夫假设。图 5(b)所示的 EKTA(EKT with Attention mechanism)则假设学生的未来表现与其历史上的每一时刻的知识状态均可能相关,按照题目的相似性聚合历史的做题表现。本文成果在大规模数据集上对 EKT 框架进行了实验验证,结果表明 EKT 同时具有较高的预测精度和良好的可解释性。

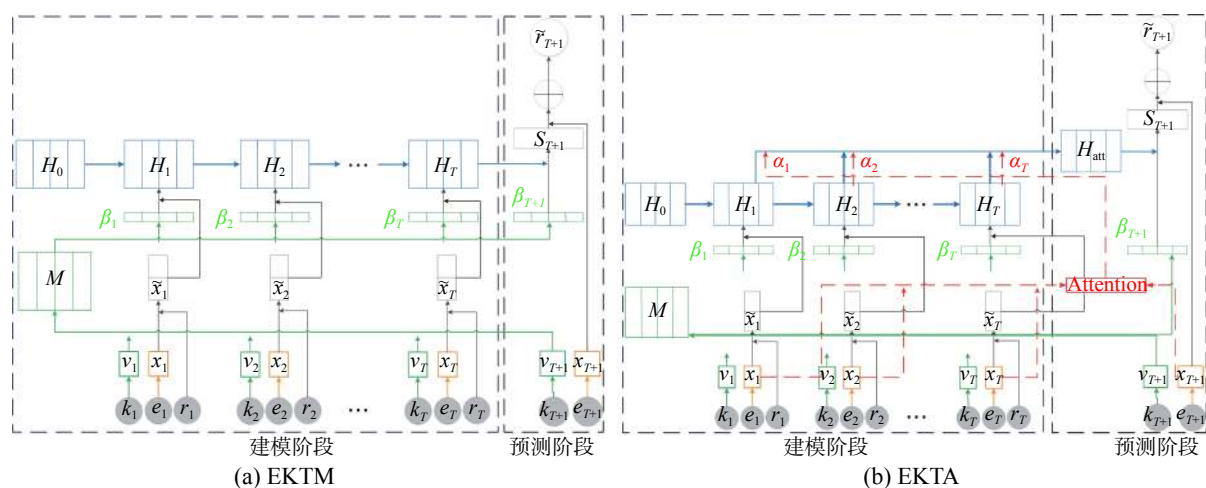


图5 基于 EKT 框架的两个模型框架

Fig. 5 Architectures of two implementations based on EKT framework

为了建模人脑记忆特性对学习者的认知状态的影响,本文成果提出一种融合学习因素的知识跟踪模型 exercise-correlated knowledge proficiency tracing (EKPT)^[29]。首先,模型通过一个预定义的关联矩阵将题目映射为一个向量,向量的每一维表示该题是否与某一特定知识概念相关联。为建立知识空间中题目之间的关联关系,EKPT 首先为题目寻找具有相近知识概念的邻居题目集合,然后聚合其邻居的知识信息,以使得具有相似知识概念的题目在空间中距离相近。为建立学生的动态演化模型,EKPT 基于学生的表现,将学生投影到相同的知识空间中的一个认知状态向量,该向量的每一维表示此学生在何种程度上掌握某一知识概念。最后,结合“学习曲线理论”和“遗忘曲线理论”,模型对学习者在学习过程中的记忆与遗忘

特性进行量化,进而动态追踪其知识掌握情况。本文成果将 EKPT 应用于实际教育学任务中,实验结果验证了所提出模型的效果。

4 自适应学习推荐新技术

智能教育的核心目标是提供学习者自主学习环境。其中,如何设计和自适应学习的策略是一个重要的任务。然而,由于智能教育场景多样、目标复杂等特点,学习策略的设计和自动调整非常困难,本文成果首次设计了基于知识匹配的个性化推荐新技术以及多目标匹配的自适应推荐新技术,分别解决了在静态学习场景和交互式学习场景中,如何将教学资源质量与学习者认知状态进行智能匹配的难题,为智能教育应用和系统研发奠定了基础。

4.1 基于知识匹配的个性化推荐技术

自适应学习推荐所具备的教育背景与传统推荐场景(如电子商务)有所不同。传统的电商推荐关注用户在兴趣爱好上与商品的匹配,而教育中的自适应学习推荐关注学习者与教育资源在知识上的匹配。因此,实现自适应学习推荐技术的一个关键在于如何在知识空间中教学资源与学习者进行精准匹配。为解决传统推荐方法难以显式建模知识状态的难题,同时满足学习者不同的学习需求,本文成果面向教学资源推荐和学习伙伴推荐两个任务,分别设计了融合认知状态的矩阵分解技术和优化认知收益的异质聚类技术,在知识空间中,实现了“学习者-教学资源”和“学习者-学习者”之间的智能匹配,保证了教学资源推荐和学习伙伴推荐的可靠性^[30]。

为满足教学资源的个性化推荐需求,本文成果结合教育心理学中的认知诊断方法和数据挖掘的推荐方法,为了增强试题推荐结果可靠性以及

可解释性,提出了一种新的试题推荐方法。如图6所示^[30],该试题推荐框架主要包括数据输入、认知诊断、得分预测以及试题推荐4个步骤。根据已有的学生答题记录,该框架利用认知诊断方法获得每个学生的认知状态。随后,将学生的认知状态作为先验信息应用到得分预测的概率矩阵分解方法中,在考虑学生共性的同时加入了每个学生的个性化认知状态。最后,根据试题难度范围和用户认知状态的不同,可以筛选出合适的试题并形成个性化的推荐试题集,每个学生可以得到一份相应的个性化推荐试题。因此,根据学生的不同学习需求,该个性化试题推荐框架可以有针对性的向学生推荐不同难度的试题。如图7所示^[30],针对需要提升薄弱知识点的掌握程度、提高知识能力的学生,可以向其推荐难度较大的试题,帮助成绩提升;而对于每个知识点的基础巩固,则可以向学生推荐难度较低(简单)的试题。

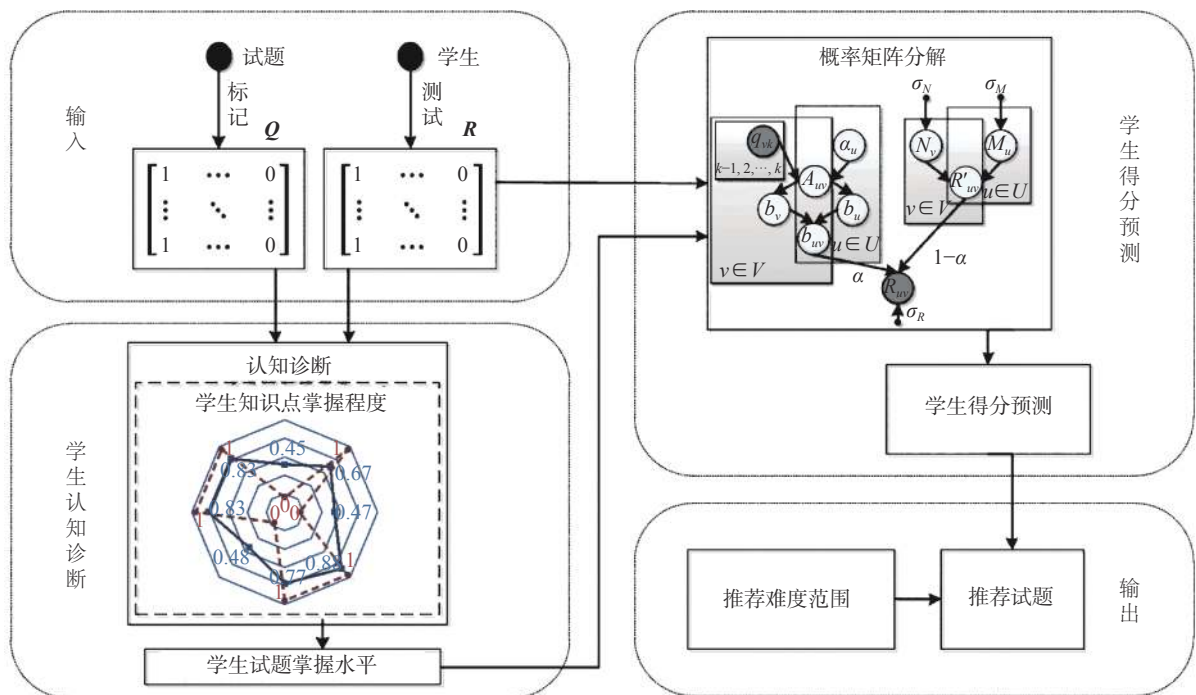


图6 结合认知能力结果的学生试题推荐框架

Fig. 6 Question recommendation with cognitive level

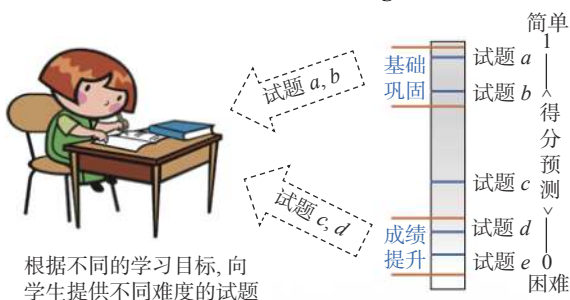


图7 有针对性的试题推荐

Fig. 7 Personalized question recommendation

为满足学习伙伴的个性化推荐需求,本文成果研究了一种基于学生认知诊断分析的学生协同分组方法^[22],并基于该方法为学生寻找合适的学习伙伴。具体地,本文成果提出了基于学生差异的分组算法 UKB(uniform k -means based)和基于收益的学生分组算法 BGB(balanced gain based)。UKB 以学生知识点掌握程度为特征对班级内的学生进行聚类,旨在将同类学生分配到不同的学习小组;BGB 通过最大化组内所有学生的平均

收益,对学生分组。本文成果将分组方法应用于学习伙伴推荐任务中,在班级中为每名学生学习最合适的学习小组伙伴,实验结果验证了所提出模型的效果^[12]。

4.2 多目标匹配的自适应推荐技术

在智能教育系统中,学习者通常与系统进行交互式的多轮学习,使得单一目标的推荐策略难以满足学习者动态多样的学习需求。因此,探索不同的推荐目标,使学习者始终保持沉浸感并能够获取更多知识,是自适应推荐所面临的主要难题之一。为此,本文成果提出了融合多个学习目标的自适应推荐技术,利用强化学习协同优化多目标长期收益,实现了推荐策略的自动调整;设计了满足知识结构约束的自适应推荐技术,生成合理的学习路径。

在实际应用场景中,一个理想的在线学习推荐系统需要对其推荐机制进行调整以适应实际情况。传统推荐系统的推荐机制一般较为单一,不足以满足实际需求,因此,本文成果提出如下3种在线题目推荐目标。1) 复习与探索: 题目推荐的根本目的是帮助学生掌握知识,不仅包括没有接触过的新知识,还包括未掌握需要复习的旧知识,因此需要根据学生实际情况,动态地进行复习与探索。2) 难度平滑性: 实践表明,一个良好的知识学习过程是循序渐进的,量化地说,即学习者接受的题目序列在难度上不应该变化过于剧烈,在推荐的过程中,算法需要根据之前的推荐序列自适应调整当前推荐的题目难度,使得被推荐题目的难度更加平滑。3) 参与度: 保持学习积极性和热情是决定学习者学习效率的重要因素,尤其是需要学习者持续参与的在线教育系统,因此,算法需要在适当的时候推荐一些有趣的或具有挑战性的题目以激发学习者的兴趣。因此,本文成果提出一种基于多学习目标优化的个性化推荐算法 DRE (deep reinforcement exercise recommendation), 其框架如图 8^[31] 所示。

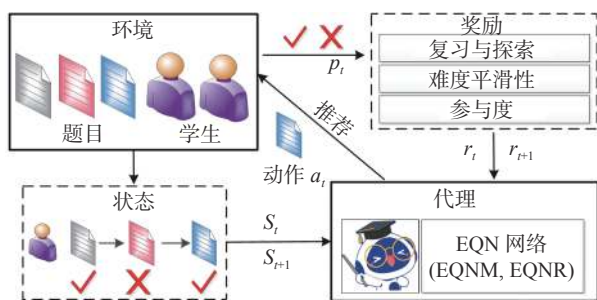


图 8 DRE 框架

Fig. 8 The proposed DRE framework

这是一个基于深度强化学习的通用框架,它

将在线学习推荐过程建模成一个马尔科夫决策过程 (Markov decision process, MDP)。由于题目资源规模过大,传统的基于有限 MDP 的强化学习算法在在线教育系统中难以实现。因此, DRE 转而对“状态-动作”价值函数进行参数化估计。具体来说, DRE 设计了两种基于不同策略的神经网络结构 (exercise Q-network, EQN)。第 1 种结构基于 Markov 性质,称为 EQNM (EQN with Markov property), 即假设推荐策略仅取决于学习者在当前时刻的状态。第 2 种结构基于序列特性,称为 EQNR (EQN with recurrent manner), 即假设推荐策略取决于学习者之前所有的历史行为序列。两种神经网络对应的强化学习训练算法都是模型无关的,不需要借助或学习 Markov 过程的状态转移信息。进一步地, DRE 针对前述 3 种推荐目标,即复习和探索、难度平滑性、参与度,分别定义了量化的奖励函数,以同时对 3 个目标进行优化。因此, DRE 在能够通过调节 3 个推荐目标的平衡来自适应地调整其推荐机制,以优化推荐效果。最后,本文对 DRE 分别进行了离线测试和在线模拟测试,实验结果验证了 DRE 的自适应题目推荐效果的有效性和鲁棒性^[10]。

为生成合理的符合知识结构的推荐路径,本文成果提出了一种基于知识结构的 learners 学习路径推荐方法 CSEAL (cognitive structure enhanced framework for Adaptive Learning)^[32]。CSEAL 首先根据目标学习路径中包含的学生需要学习的所有知识单元及学生当前学习的第一个知识单元,确定候选知识单元集合;然后, CSEAL 利用知识追踪算法获取学生的当前知识状态,并根据学生不同的知识状态,确定候选知识单元集合中每一个知识单元作为目标知识单元时为最优解的概率,最后选取其中最大概率的知识单元作为目标知识单元,即学生下一个需要学习的知识单元。在 CSEAL 中,通过将学习路径推荐问题转化为逐步的马尔可夫决策问题,并应用演员-评论家算法,动态更新推荐策略,从而顺序地向不同学生推荐能实现高效学习的知识单元,详细方法可参考文献^[32]。

5 智能教育系统——“智学网”

融合上述 3 项自适应学习核心技术,本文成果进一步研发了一款面向基础教育的智能教育系统,名为“智学网”。该系统通过伴随式地采集日常“教与学”过程性的动态数据,建立面向中小学的大规模高质量教学资源库,构建以学习者为中

心的学业评价体系,实现优质资源精准推荐,为老师“针对性教”和学生“个性化学”提供了智能环

境与服务。“智学网”系统核心模块主要由数据层、技术层和应用层构成,核心模块结构如图9所示。

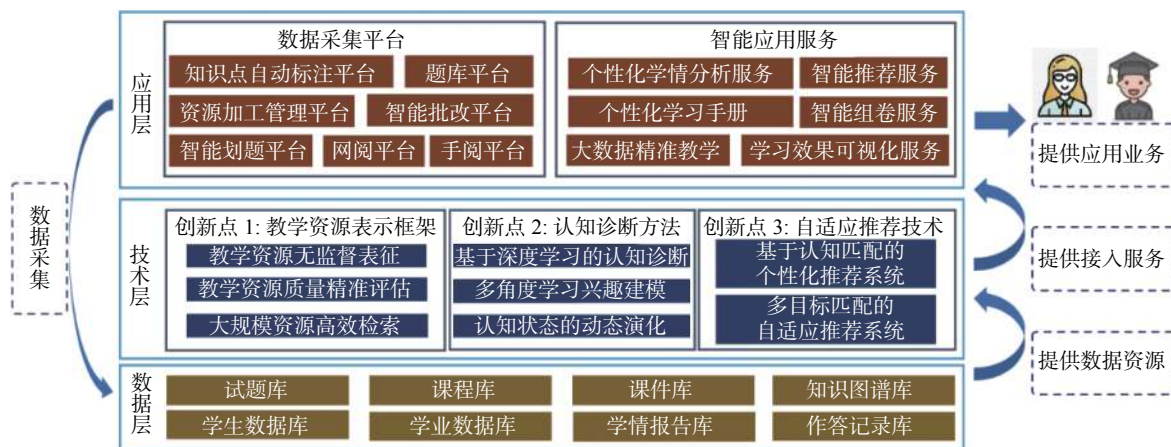


图9 “智学网”系统核心模块结构

Fig. 9 Structure diagram of key modules of “Zhixue”

数据层主要负责对上层应用产生的教学资源(试题、课程、课件等)和学习者学习行为(学业数据、作答记录等)数据进行存储管理。现阶段,“智学网”已构建了面向中小学的大规模高质量教学资源库。技术层基于本文成果核心技术创新成果,即教学资源表示新框架、学习者认知诊断新方法和自适应学习推荐新技术,技术层为应用层提供了核心接口服务。应用层包含数据采集平台和智能应用服务,数据采集平台通过如知识点标注平台、题库平台等多个数据平台采集教学资源与学习者数据,并回灌数据层进行存储和管理;智能应用服务为学习者提供了多个服务平台,如个性化学情分析、个性化学习手册,构建了以学

习者为中心的评价体系。通过数据层、服务层、应用层的联合协作,“智学网”实现了优质教学资源精准推荐,同时,也伴随式记录了更高质量的学业数据,促进良性循环。

“智学网”的智能教育技术服务体系框架如图10所示,包含智能基础设施、上层对象和服务场景等3层架构。截止至2020年7月,“智学网”系统已深入教学核心,其产品服务覆盖智能课堂、智能教研、智能考试、智能评价等各类教与学场景。系统已在安徽、江西等全国各省市16000余所学校推广并实现常态化应用,惠及师生约2500万。全国范围每月组织联考数千场,服务各类班级、年级测试数万场,提供评价报告800万份。

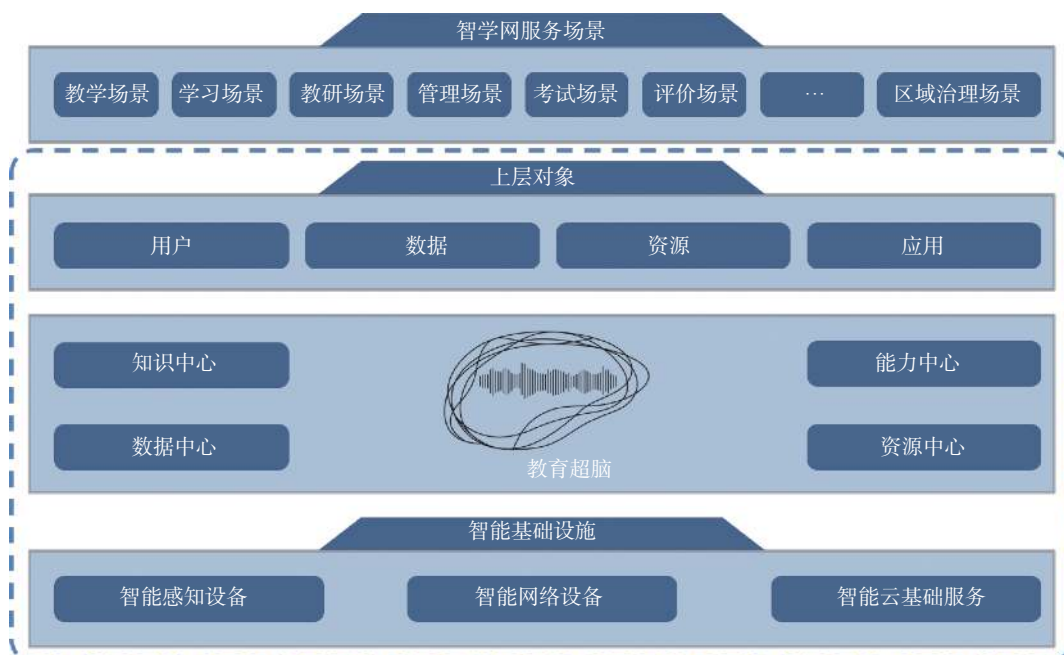


图10 “智学网”智能教育技术服务体系框架

Fig. 10 Framework of intelligent technology and service system of “Zhixue”

6 结束语

教育作为民族振兴、社会进步的基石,是国之大计、党之大计,对实现中华民族伟大复兴具有决定性意义。本文成果经过10余年的潜心研发,突破了面向智能教育的自适应学习技术难点,构建了教学资源表示新框架,提出了学习者认知诊断新方法,设计了自适应推荐新技术,研发了面向基础教育的智能教育系统——“智学网”。其关键技术和系统平台已取得重大的技术突破和大规模行业应用。未来,我们将依托所研发的“智学网”系统,逐步开展教学资源知识库的构建、学习者多种学习行为数据的收集以及在线与离线数据的对标融合,旨在设计智能程度更高、解释性更强的自适应学习技术,力求在关键技术上再一次突破,为我国智能教育发展应用和教育现代化事业做出进一步的贡献。

参考文献:

- [1] 张志祯, 张玲玲, 李芒. 人工智能教育应用的应然分析: 教学自动化的必然与可能 [J]. 中国远程教育, 2019(1): 23–35, 92.
ZHANG Zhizhen, ZHANG Lingling, LI Mang. A normative analysis of AI in education: the necessity and possibility to automate instruction[J]. Distance education in China, 2019(1): 23–35, 92.
- [2] 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): 209–224.
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): 209–224.
- [3] 陈丽, 林世员, 郑勤华. “互联网+”时代中国远程教育的机遇和挑战 [J]. 现代远程教育研究, 2016(1): 3–10.
CHEN Li, LIN Shiyuan, ZHENG Qinhua. Opportunities and challenges of Chinese distance education in the “Internet+” Era[J]. Modern distance education research, 2016(1): 3–10.
- [4] 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, 2018(1): 77–90.
LIU Qi, CHEN Enhong, ZHU Tianyu, et al. Research on educational data mining for online intelligent learning[J]. Pattern recognition and artificial intelligence, 2018(1): 77–90.
- [5] 黄振亚. 面向个性化学习的数据挖掘方法与应用研究 [D]. 合肥: 中国科学技术大学, 2020.
HUANG Zhenya. Data mining techniques and applications for personalized learning[D]. Hefei: University of Science and Technology of China, 2020.
- [6] 朱天宇. 结合认知诊断的方法的学生建模与应用研究 [D]. 合肥: 中国科学技术大学, 2018.
ZHU Tianyu. Student modeling and application research based on cognitive diagnosis[D]. Hefei: University of Science and Technology of China, 2018.
- [7] 刘玉苹. 基于认知诊断的辅助教育技术研究 [D]. 合肥: 中国科学技术大学, 2016.
LIU Yuping. A study of educational assistive technologies based on cognitive diagnosis theory[D]. Hefei: University of Science and Technology of China, 2016.
- [8] 黄仔. 基于多模态学习的试题建模方法与应用研究 [D]. 合肥: 中国科学技术大学, 2019.
HUANG Zai. Research of exercise modeling and application based on multimodal learning[D]. Hefei: University of Science and Technology of China, 2019.
- [9] 吴凡. 面向 2030 的教育质量: 核心理念与保障模式——基于联合国教科文组织等政策报告的文本分析 [J]. 教育研究, 2018, 39(1): 132–141.
WU Fan. The core conceptualization and assurance models in education 2030 framework for action——text analysis on policies and reports of UNESCO[J]. Educational research, 2018, 39(1): 132–141.
- [10] PETERS M E, NEUMANN M, IYYER M, et al. Deep contextualized word representations[C]//Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers). New Orleans, USA, 2018: 2227–2237.
- [11] YIN Yu, LIU Qi, HUANG Zhenya, et al. Quesnet: a unified representation for heterogeneous test questions [C]//Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. Anchorage, USA, 2019: 1328–1336.
- [12] FUCHS L S, FUCHS D, HAMLETT C L, et al. Effects of expert system consultation within curriculum-based measurement, using a reading maze task[J]. Exceptional children, 1992, 58(5): 436–450.
- [13] HUANG Wei, CHEN Enhong, LIU Qi, et al. Hierarchical multi-label text classification: an attention-based recurrent network approach[C]//Proceedings of the 28th ACM International Conference on Information and Knowledge Management. Beijing, China, 2019: 1051–1060.
- [14] HUANG Zhenya, LIU Qi, CHEN Enhong, et al. Question difficulty prediction for READING problems in standard tests[C]//Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence. San Francisco, USA, 2017: 1352–1359.
- [15] HAGE H, AIMERU E. ICE: a system for identification of conflicts in exams[C]//IEEE International Conference on Computer Systems and Applications. Dubai, United Arab

- Emirates, 2006: 980–987.
- [16] TSINAKOS A, KAZANIDIS I. Identification of conflicting questions in the PARES system[J]. *International review of research in open and distributed learning*, 2012, 13(3): 297–313.
- [17] WILLIAMS A E, AGUILAR-ROCA N M, TSAI M, et al. Assessment of learning gains associated with independent exam analysis in introductory biology[J]. *CBE-life sciences education*, 2011, 10(4): 346–356.
- [18] YU Jing, LI Dongmei, HOU Jiajia, et al. Similarity measure of test questions based on ontology and VSM[J]. *The open automation and control systems journal*, 2014, 6(1): 262–267.
- [19] LIU Qi, HUANG Zai, HUANG Zhenya, et al. Finding similar exercises in online education systems[C]//Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. London, United Kingdom, 2018: 1821–1830.
- [20] MELLEBERGH G J. Item bias and item response theory[J]. *International journal of educational research*, 1989, 13(2): 127–143.
- [21] DE LA TORRE J. DINA model and parameter estimation: A didactic[J]. *Journal of educational and behavioral statistics*, 2009, 34(1): 115–130.
- [22] WANG Fei, LIU Qi, CHEN Enhong, et al. Neural cognitive diagnosis for intelligent education systems[J]. *Proceedings of the AAAI conference on artificial intelligence*, 2020, 34(4): 6153–6161.
- [23] KOREN Y, BELL R, VOLINSKY C. Matrix factorization techniques for recommender systems[J]. *Computer*, 2009, 42(8): 30–37.
- [24] JEWSBURY P A, VAN RIJN P W. IRT and MIRT models for item parameter estimation with multidimensional multistage tests[J]. *Journal of educational and behavioral statistics*, 2020, 45(4): 383–402.
- [25] LIU Qi, WU Runze, CHEN Enhong, et al. Fuzzy cognitive diagnosis for modelling examinee performance[J]. *ACM transactions on intelligent systems and technology (TIST)*, 2018, 9(4): 48.
- [26] WU Runze, XU Guandong, CHEN Enhong, et al. Knowledge or gaming?: cognitive modelling based on multiple-attempt response[C]//Proceedings of the 26th International Conference on World Wide Web Companion (WWW' 2017 Companion). Perth, Australia, 2017: 321–329.
- [27] PIECH C, BASSEN J, HUANG J, et al. Deep knowledge tracing[C]//Proceedings of the 28th International Conference on Neural Information Processing Systems. Montreal, Canada, 2015: 505–513.
- [28] LIU Qi, HUANG Zhenya, YIN Yu, et al. EKT: exercise-aware knowledge tracing for student performance prediction[J]. *IEEE transactions on knowledge and data engineering*, 2021, 33(1): 100–115.
- [29] HUANG Zhenya, LIU Qi, CHEN Yuying, et al. Learning or forgetting? A dynamic approach for tracking the knowledge proficiency of students[J]. *ACM transactions on information systems*, 2020, 38(2): 19.
- [30] LIU Yuping, LIU Qi, WU Runze, et al. Collaborative learning team formation: a cognitive modeling perspective[C]//Proceedings of the 21st International Conference on Database Systems for Advanced Applications. Dallas, Texas, USA, 2016: 383–400.
- [31] HUANG Zhenya, LIU Qi, ZHAI Chengxiang, et al. Exploring multi-objective exercise recommendations in online education systems[C]//Proceedings of the 28th ACM International Conference on Information and Knowledge Management. Beijing, China, 2019: 1261–1270.
- [32] LIU Qi, TONG Shiwei, LIU Chuanren, et al. Exploiting cognitive structure for adaptive learning[C]//Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. Anchorage, USA, 2019: 627–635.

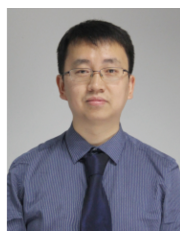
作者简介:



陈恩红, 教授, 博士生导师, IEEE 高级会员、CAAI 会士、CCF 会士、大数据专家委员会副主任, 主要研究方向为机器学习与数据挖掘、社会网络、个性化推荐。国家自然科学基金重大科研仪器研制项目、联合基金重点项目、国家 863 计划、科技部重点研发计划课题等多项。发表学术论文 150 余篇。



刘淇, 教授, 博士生导师, 主要研究方向为数据挖掘与知识发现、机器学习方法及其应用、教育大数据分析。入选中国科协“青年人才托举工程”、CCF 青年人才托举计划(2017 年)、微软亚洲研究院青年学者“铸星计划”、CCF-Intel 青年学者提升计划等。主持国家自然科学基金面上项目, 科技部重点研发计划课题等多项。发表学术论文 80 余篇。



王士进, 高级工程师, 主要研究方向为人工智能、模式识别、智能教育系统。主持和参与 863 计划重点项目、工信部电子信息产业发展基金项目等多项, 获授权专利和软件著作权 10 余项。发表学术论文 30 余篇。