

# 面向在线智慧学习的教育数据挖掘技术研究

刘洪<sup>1</sup> 陈恩红<sup>1</sup> 朱天宇<sup>1</sup> 黄振亚<sup>1</sup> 吴润泽<sup>1</sup> 苏喻<sup>2</sup> 胡国平<sup>2</sup>

**摘要** 随着教育信息化进程的深入,学生在线学习数据得到不断积累,为数据驱动的教育评估和智能辅助教学提供良好条件.然而,已有的面向在线智慧学习的教育数据挖掘模型很难从海量、稀疏、高噪的数据中准确分析试题特征和学生学业水平,也较少考虑学生及教师的个性化需求.文中针对上述问题开展若干面向在线智慧学习的教育数据挖掘技术研究工作,以教育学习所涉及的试题、学生、教师为对象,以个性化推荐等技术同教育领域知识相结合为手段,以提高学生学业水平为目标.具体介绍用于试题分析和检索的试题文本表征模型、基于认知诊断的个性化学习资源推荐方法、针对教师的教学建议和指导等方法,以及这些技术所依托的应用平台——科大讯飞在线教育系统“智学网”.最后简单讨论面向在线智慧学习的教育数据挖掘技术未来可能的研究方向.

**关键词** 在线智慧学习,教育评估,认知诊断,推荐系统

**引用格式** 刘洪,陈恩红,朱天宇,黄振亚,吴润泽,苏喻,胡国平.面向在线智慧学习的教育数据挖掘技术研究.模式识别与人工智能,2018,31(1):77-90.

**DOI** 10.16451/j.cnki.issn1003-6059.201801007

## Research on Educational Data Mining for Online Intelligent Learning

LIU Qi<sup>1</sup>, CHEN Enhong<sup>1</sup>, ZHU Tianyu<sup>1</sup>, HUANG Zhenya<sup>1</sup>, WU Runze<sup>1</sup>, SU Yu<sup>2</sup>, HU Guoping<sup>2</sup>

**ABSTRACT** With the rapid informationization of education, extensive data records from online education of students are accumulated, and it provides a good opportunity for both data-driven educational assessment and intelligent tutoring. However, existing models are hard to accurately analyze the characteristics of questions and the academic levels of students from the massive and sparse data with high noise. Meanwhile, it is difficult for these models to satisfy the personalized needs of students and teachers. In this paper, educational data mining studies on these problems are summarized. To improve the student academic level, these studies focus on modeling three objects in education (i. e., questions, students and teachers) and apply effective techniques, such as personalized recommendation methods, combined with the domain knowledge from education. Specifically, a question text embedding framework is presented for question analysis and question retrieval. Then, personalized recommendation methods on learning resources are illustrated based on the cognitive diagnosis of students. Moreover, the way of providing effective guidance and suggestions for teachers is showed. Some of these research achievements are applied to the online educational system “ZHIXUE” in iFlyTek. Finally, the possible research directions in the future are discussed.

**Key Words** Online Intelligent Learning, Educational Assessment, Cognitive Diagnosis, Recommender System

本文责任编辑 高新波

Recommended by Associate Editor GAO Xinbo

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 合肥 230027
2. 科大讯飞股份有限公司 合肥 230088

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
2. USTC iFLYTEK Co., Ltd., Hefei 230088

收稿日期:2017-12-08;录用日期:2017-12-22

Manuscript received December 8, 2017;

accepted December 22, 2017

国家自然科学基金项目(No. 61672483, U1605251)、中国科学院青年创新促进会会员专项项目(No. 2014299)资助

Supported by National Natural Science Foundation of China(No. 61672483, U1605251), Project of Youth Innovation Promotion Association of CAS(No. 2014299)

**Citation** LIU Q, CHEN E H, ZHU T Y, HUANG Z Y, WU R Z, SU Y, HU G P. Research on Educational Data Mining for Online Intelligent Learning. *Pattern Recognition and Artificial Intelligence*, 2018, 31(1): 77-90.

近些年,信息化进程已影响到当前社会生活的各方面,教育作为一个传统的领域,也正受到信息化的影响<sup>[1-2]</sup>。随着教育信息化的持续深入以及互联网的迅猛发展,在线教育已成为计算机融合传统教育领域而形成的一个新的重要研究和应用方向。

当前,已出现一批优秀的在线教育平台。例如大规模开放在线课堂(Massive Open Online Course, MOOC)就是具有代表性的一类在线学习平台<sup>[3]</sup>。MOOC平台借助发达的视频技术和网络技术,面向大众提供海量的网络课程,在一定程度上缓解教育资源匮乏、分配不均衡的问题。除了传统的课程材料(如课程视频、扩展阅读和问题考核),MOOC平台还提供诸如讨论论坛、学习社区等用于网络学习。除MOOC平台之外,智能辅导系统(Intelligent Tutoring Systems, ITS)<sup>[4]</sup>也是广受关注的在线教育平台形式之一。ITS旨在通过不断收集反馈信息为每位学生提供个性化定制教育,并且ITS提供的个性化教育通常是在没有人工干预的情况下进行。除了MOOC平台和ITS,还有移动端数组图书馆、在线学习社区等,这些在线学习平台的兴起为学生营造一个具有海量教育资源、不限时间、不限地点的在线学习环境。

在线教育模式的涌现受到各界关注,Coursera、edX、Udacity等发展较早,在这些成熟的国外在线教育平台的影响和激励下,催生许多国内在线教育平台,如MOOC中文网、微课网、国家教育资源公共服务平台等<sup>[5-6]</sup>。与此同时,在国内还兴起如猿题库(<http://www.yuantiku.com>)、智学网(<http://www.zhixue.com>)等一些在线试题库练习平台。区别于MOOC平台提供海量的网络课程,这类在线试题库练习平台主要面向K12中小学教育,帮助学生练习、巩固课堂中学习的知识。由于目前国内的K12教育仍以学校的离线教育为主,在线教育存在一定的覆盖规模较小、使用频率较低等问题,因此部分试题库练习平台在提供在线练习的同时,也收集学生的离线学习数据(包含学生作业、考试等)。此外,学习平台通过智能分析学生的答题数据,向学生、教师反映学生的个性化学习情况,并进行有针对性的试题训练,旨在帮助学生提高学业水平。

随着这些不同类型在线教育平台的快速建设,各个平台都收集大量的课程信息、学生信息及学生

学习记录。针对在线学习的数据分析和挖掘<sup>[7]</sup>逐渐成为教育数据挖掘(Educational Data Mining, EDM)的重要研究方向之一。EDM将数据挖掘的相关技术和方法应用于教育领域,对这些教育平台或线下教育产生的数据进行挖掘和分析,更好地了解学习情况,有效帮助学生进行学习等<sup>[8]</sup>。2010年,KDD Cup竞赛(<http://www.kdd.org/kdd-cup>)首次以使用学生在线答题记录预测学生学习成绩(认知能力)为任务,向全球的教育学、计算机科学的研究者发起挑战。2015年,KDD Cup竞赛再次使用教育数据,此次的任务是基于清华大学学堂在线MOOC平台数据,进行学生在线学习退课预测。由此可见,面向在线学习的教育数据挖掘已成为研究者普遍关注的重要课题。

目前,相关学者已研究多项针对MOOC平台和ITS的数据分析与挖掘方法<sup>[9-10]</sup>。Yudelson等<sup>[11]</sup>根据课程设置、参考书目和在线检测研究学生在MOOC平台中的行为模式挖掘与激励。Chakor<sup>[12]</sup>根据学生在课程周期中的行为特征进行学生退课预测。Bruff等<sup>[13]</sup>根据学生影响力(在线论坛、讨论组、测试成绩等)的知识传播建模与引导研究知识传播问题。朱天宇等<sup>[14]</sup>和Kongsakun等<sup>[15]</sup>研究在线学习平台中的个性化信息过滤问题,如课程、论坛专家文章推荐、学生分组、学友推荐等。这些已有的研究成果可以帮助学生更便捷地使用在线教育平台中的资源。

此外,学者们针对在线学习的试题库和学习系统书籍进行分析与挖掘。de la Torre<sup>[16]</sup>根据学生的学习情况记录进行认知诊断分析,了解学生的学习状态、知识点掌握情况和粗心程度等。Midgley<sup>[17]</sup>研究学生个性化教学与辅导,为每个学生提供个性化的试题推荐和学习策略。Lü等<sup>[18]</sup>根据在线学习平台试题库中试题依赖关系、学习顺序等构建知识图谱。Hudak等<sup>[19]</sup>研究试题内容自动标识等问题。这些工作可以更好地帮助研究人员了解学生的学习过程以及当前的学习状态,为每位学生智能定制更个性化的学习方案。

然而,在线智慧教育领域仍存在许多问题需要进一步解决。例如,由于从在线教育平台获得的数据具有海量、稀疏、高噪等特点,因此现有的智能辅助模型很难对试题特征和学生学业水平进行准确分

析<sup>[1]</sup>. 其次, 现有模型较少考虑学生、教师的个性化需求, 难以对每位学生进行个性化分析与推荐<sup>[20]</sup>. 为此, 本文针对在线学习涉及到的试题表示、学生学习和教师教辅等, 综述和讨论教育数据挖掘技术研究与应用的部分工作.

本文研究基本框架如图 1 所示: 1) 面向试题, 介绍基于试题表征的难度预测相关工作; 2) 面向学生, 介绍学生认知诊断分析和学生心理建模方法等; 3) 面向教学, 介绍面向自适应学习的个性化推荐、协同分组等的探索工作. 下面详细介绍各方面的具体进展, 以及部分技术的应用场景和依托的应用平台——科大讯飞在线教育系统“智学网”.

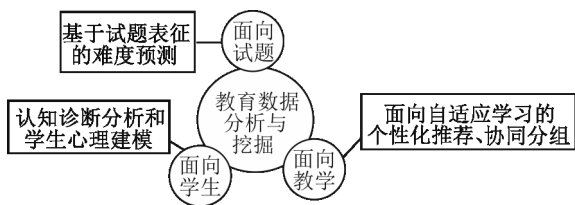


图 1 本文研究基本框架

Fig. 1 The proposed basic research framework

# 1 面向在线智慧学习的教育数据挖掘

本节从在线学习相关的 3 方面入手, 分别介绍教育数据挖掘在面向试题、学生、教学辅助等方面的研究工作实例, 并介绍部分技术的应用场景和所依托的科大讯飞在线教育系统“智学网”.

## 1.1 试题分析

试题作为教育中的一类重要的学习资源, 在传统教育和在线教育中都发挥重要作用, 针对试题分析的相关研究吸引许多研究人员的注意. 目前, 许多在线学习网站都提供在线题库供学生使用、学习<sup>[2]</sup>, 智学网、猿题库等学习平台更是搭建以试题练习为主的在线学习环境. 因此, 在线学习系统需要提供针对海量试题的收集、整理、存储, 并提供试题下载或在线答题等服务.

### 1.1.1 试题分析研究现状

图 2 为智学网中的试题样例, 包含试题类型、题面、知识点等一系列的试题信息. 如何有效地对试题题面、知识点及学生在试题上的答题记录等信息进行表征, 并使用试题的数据表征进行试题难度预测、相似试题判定或试题知识点识别, 是在线教育中面向试题研究的重要内容. 本文着重介绍基于试题表

征方法的难度预测研究.

题号: 1014XXXX143    题型: 解答题 难度: 一般    日期: 2016-05-13 标题/来源: XXX 年 XXX 试卷 知识点: 不等式和绝对值不等式 [题文] 已知关于 $x$ 的不等式 $ ax - 1  +  ax - a  \geq 1 (a > 0)$ 1) 当 $a = 1$ 时, 求此不等式的解集; 2) 若此不等式的解集为 $R$ , 求实数 $a$ 的取值范围.
--

图 2 智学网试题样例<sup>[20]</sup>

Fig. 2 Question sample of ZHIXUE in iFlyTek<sup>[20]</sup>

试题难度是试题的一项重要特征, 测试试题难度的预估对教学的各方面都具有一定的现实意义<sup>[21]</sup>. 首先, 对试题难度的准确估计和标注可以帮助系统构造更合理的试题库. 其次, 做好试题难度估计和标注工作可以辅助教师自动组卷, 在减少人工工作量的同时, 提高组卷的科学性和合理性. 最后, 准确的试题难度估计可以帮助在线学习系统为每个学生筛选合适的练习题.

教育心理学中的经典测量理论 (Classical Test Theory, CTT) 认为, 试题的难度可由一道试题被错误回答的比率表示, 即对一道试题, 答错的学生越多, 那么该试题就越难<sup>[22]</sup>. 然而, 在某些实际情况下, 需要在学生对试题进行作答之前, 进行试题难度的预测和评估. 例如, 对于标准化测试 (如 TOEFL, GRE 等), 每个学生可以多次参加考试, 并选择最高分数作为最终分数用于学校申请等工作<sup>[23]</sup>. 因此, 多次考试的试题难度应当一致, 以保障考试成绩的公平性; 对于一些大型的考试 (如中国的普通高等学校招生全国统一考试, 简称高考), 对试题的难度预测可以保证测试的质量.

传统的试题难度预测方法可以基于专家经验对试题难度进行人工评估或使用教育心理学中的简单回归模型<sup>[24]</sup>. 然而, 前者存在专家主观性强、费时费力的问题, 后者存在难以使用海量样本、对试题文本特征等利用不足的问题. Huang 等<sup>[25]</sup> 提出基于试题文本的难度预测框架 (Test-Aware Attention-Based Convolutional Neural Network, TACNN), 用于自动完成试题的难度预测任务.

### 1.1.2 基于试题表征方法的难度预测

图 3 为 TACNN 框架示意图. 具体地, 以英语阅读理解试题为例, TACNN 通过使用试题文本信息及学生在试题上的历史答题记录, 预测新的考试中出



现的英语阅读理解试题的难度(答错率)。

从图3可以看出,TACNN 主要由4部分组成,分别是输入层、句子卷积神经网络层、注意力层、预测层.输入层包含英语阅读理解选择题  $Q_i$  的各部分信息:阅读文本张量  $TD_i$ 、问题矩阵  $TQ_i$ 、选项矩阵  $TO_i$ . TACNN 的第2层,基于卷积神经网络(Con-volutional Neural Network, CNN)对试题中的

每个句子进行表征,将每句话表示为一个固定长度的向量.第3层注意力层结合学生的阅读习惯处理试题,包括本文每个句子对试题难度的贡献  $DA_i$  及每个选项的难度贡献  $DO_i$ . 在最终的预测层中,TACNN 根据文本难度贡献  $DA_i$ 、选项难度贡献  $DO_i$  及问题的形式化表征,设计考试范围依赖的偏序训练目标,消除不同考试的误差,预测试题难度  $\tilde{P}_i$ .

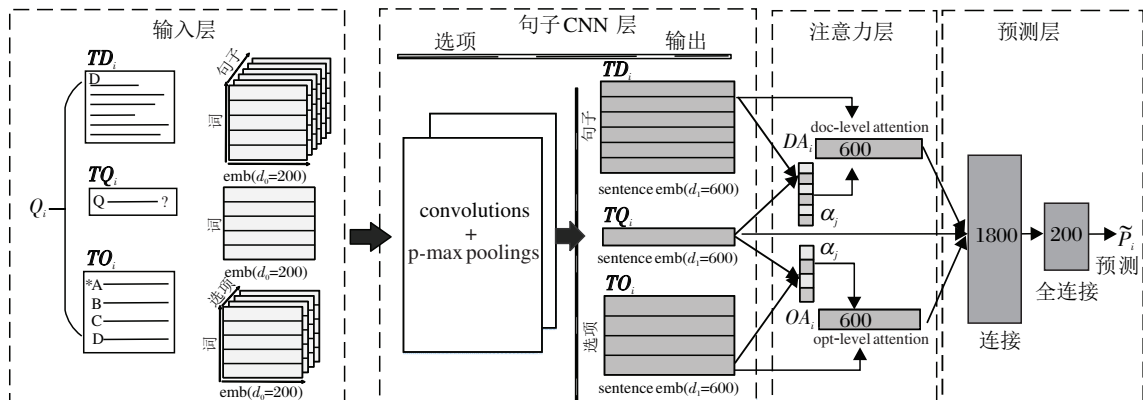


图3 TACNN 框架<sup>[25]</sup>

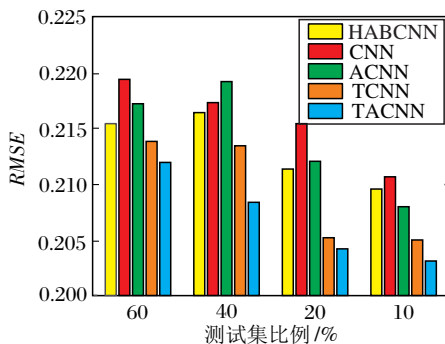
Fig. 3 TACNN Framework<sup>[25]</sup>

TACNN 设计统一的表征方法对英语阅读理解试题的各个组成部分(阅读文本、问题、选项)进行语义理解和表示,并针对每道题量化各部分文本的特征表示,以区分阅读理解试题阅读文本材料部分中各句子对不同问题的重要性.最重要的是,TACNN 设计可以消除不同考试范围误差的框架,以试题表征的结果均衡因参加考试的学生不同而导致的试题难度无法直接比较的问题,从而对每道试题进行精确的难度预测。

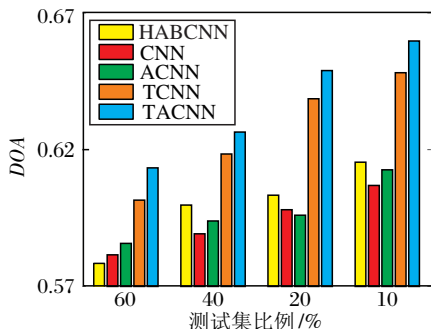
将 TACNN 框架应用于科大讯飞提供的学生英语阅读答题数据,使用近 300 万的学生答题记录 and 对应题面,对试题难度进行预测.为了衡量 TACNN 的性能,实验中首先使用广泛采用的均方根误差(Root Mean Square Error, RMSE)对比准确度.然后,使用排序准确度(Degree of Agreement, DOA)从排序的角度衡量试题对之间的难度排序正确百分比.其次,使用皮尔逊相关系数(Pearson Correlation Coefficient, PCC)衡量试题难度和学生答题表现之间的关系.最后,使用  $t$  检验通过率(Passing Ratio, PR)检测难度评估的置信水平。

这 4 个不同的评价指标中, RMSE 越小,模型表现越好;其余指标越大,模型表现越好.图 4 为 TACNN 难度预测结果,对比方法包括:分层含注意力策略的 CNN (Hierarchical Attention-Based CNN,

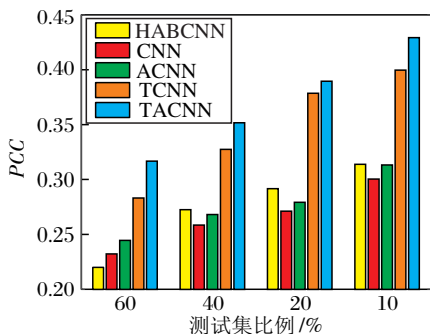
HABCNN)、CNN、含注意力策略的 CNN (Attention-Based CNN, ACNN)、文本感知的 CNN (Text-Aware CNN, TCNN).从图中可以看出,相比其它方法,TACNN 可更准确地对英语阅读类型试题进行难度预测。



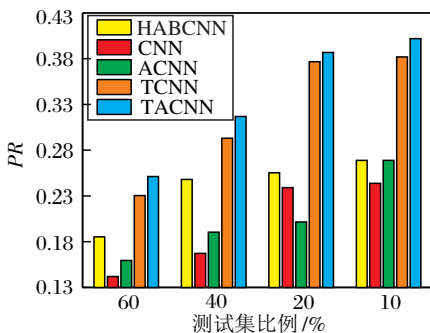
(a) RMSE



(b) DOA



(c) PCC



(d) PR

图 4 TACNN 难度预测结果<sup>[25]</sup>

Fig. 4 Difficulty prediction result of TACNN<sup>[25]</sup>

## 1.2 学生认知诊断

学生是在线教育系统服务的主要对象,准确分析学生的认知水平,对帮助了解学生状态、构建适当的教学计划都具有极大的帮助<sup>[9]</sup>. 为此,教育心理学的研究者提出认知诊断评价 (Cognitive Diagnosis Assessment, CDA),进行学生认知水平的全面分析.

### 1.2.1 认知诊断研究现状

CDA 是一个基于认知心理学、统计学和计算机科学的评价体系,通过对学生的答题数据进行建模分析,引入试题与知识结构的关联关系,对学生的认知状态进行诊断,定量考察学生的个体差异和认知水平<sup>[8]</sup>. 认知诊断理论起源于 20 世纪 50 年代,经过多年的发展,教育心理学家提出多个不同的认识诊断模型 (Cognitive Diagnosis Model, CDM). 根据“试题-知识点结构”的关系与“学生-知识点”的掌握类型<sup>[26]</sup>,常见的 CDM 可以分为单维连续型模型和多维离散型模型. 这里,“维度”指学生能力(试题所关联知识点、技能)的数目,“离散”、“连续”对应模型

诊断得到的学生能力值是离散型(0 或 1)还是连续型.

项目反映理论 (Item Response Theory, IRT) 是最常用的单维连续认知诊断模型之一<sup>[27]</sup>,它假设学生对试题的作答结果服从独立同分布,并将每位学生的认知状态表示为一个单维连续的能力值,再结合试题特征(区分度、难度等)建模学生的答题情况. IRT 模型的评估函数有多种形式,较常见的 IRT 模型的函数形式之一如下:

$$P(X_{ij} = 1 | \theta_j, a_i, b_i, c_i) =$$

$$c_i + \frac{1 - c_i}{1 + \exp(-1.7a_i(\theta_j - b_i))},$$

具体地,上式通过估计学生  $j$  答对试题  $i$  的概率诊断分析学生认知能力  $\theta_j$ ,其中,  $a_i$  为试题的区分度,  $b_i$  为试题的难度,  $c_i$  为试题的猜测度. 该模型为三参数逻辑斯蒂回归 IRT 模型 (Three-Parameter Logistic IRT, 3PL-IRT). 当不考虑试题猜测度  $c_i$  时,模型退化为双参数逻辑斯蒂回归 IRT 模型 (Two-Parameter Logistic IRT, 2PL-IRT). 若只考虑试题的区分度  $a_i$ ,该模型为单参数逻辑斯蒂回归 IRT 模型 (One-Parameter Logistic IRT, 1PL-IRT). 根据应用场景及诊断要求的不同,这 3 种类型的 IRT 模型在教育数据挖掘领域均具有较广泛应用.

在所有的多维离散型认知诊断模型中, DINA (Deterministic Input, Noisy AND-gate)<sup>[16]</sup> 模型是应用较广泛的模型之一. DINA 模型认为每个试题由显式的知识点表述,并将此关联关系使用  $Q$  矩阵表示<sup>[26]</sup>.  $Q$  矩阵通常由教育领域专家进行标注,从图 2 的智学网试题样例中可以看出,该题所关联的知识点为“不等式和绝对值不等式”. 表 1 为一个  $Q$  矩阵示例,从表中可看出,试题  $i_1$  关联知识点  $S_1$  和  $S_4$ . DINA 模型将学生能力描述为在多维知识点上的离散掌握向量,表 2 为 2 位学生在知识点上的掌握程度示例,从表中可看出学生  $j_1$  掌握知识点  $S_1$  和  $S_3$ .

表 1 试题-知识点关联  $Q$  矩阵示例

Table 1 Sample of question-knowledge point corresponding  $Q$ -matrix

试题	知识点				
	$S_1$	$S_2$	$S_3$	$S_4$	$S_5$
$i_1$	1	0	0	1	0
$i_2$	0	1	0	1	0

表 2 学生-知识点掌握矩阵示例

Table 2 Sample of student-knowledge point corresponding matrix

学生	知识点				
	$S_1$	$S_2$	$S_3$	$S_4$	$S_5$
$j_1$	1	0	1	0	0
$j_2$	1	1	0	1	0

DINA 模型同时考虑学生的认知能力及“猜测”和“失误”等心理因素进行认知诊断分析,对于一个包含  $K$  个知识点的测试,学生  $j$  答对试题  $i$  的概率  $X_{ij}$  可表示为

$$P(X_{ij} = 1 | \theta_j, g_i, s_i) = \begin{cases} 1 - s_i, & \eta_{ij} = 1 \\ g_i, & \eta_{ij} = 0 \end{cases}$$

其中,  $g_i$  和  $s_i$  分别为试题  $i$  的猜测和失误因素,  $\eta_{ij} =$

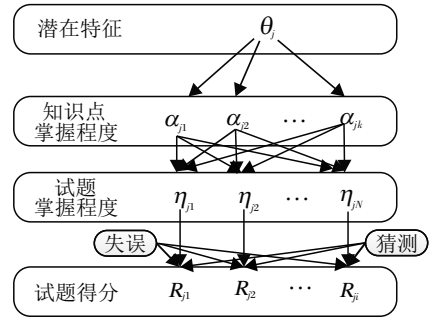
$\prod_{k=1}^K \theta_{jk}^{q_{ik}}$  是学生  $j$  对试题  $i$  的掌握程度. DINA 模型认为,学生  $j$  掌握试题  $i$  关联的全部知识点  $k$  时,可正确作答试题  $i$  ( $\eta_{ij} = 1$ ). 由专家标注得到的  $Q$  矩阵具有极好的解释性,因此, DINA 模型通过引入试题-知识点关联  $Q$  矩阵,不仅将诊断得到的学生单维认知能力扩展至多维认知状态,还对每个认知维度赋予实际含义,诊断得到的学生能力也具有很强的可解释性.

### 1.2.2 模糊认知诊断分析框架

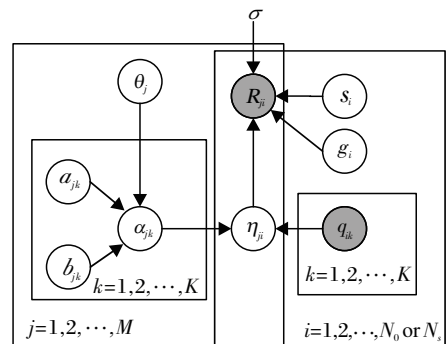
教育心理学家提出的传统认知诊断模型通常仅基于小样本,一般只适用于客观题(客观题是让学生从事先拟定的答案中辨认正确答案的题目,答案非对即错),难以考虑学生在主观题(学生自由陈述答案,评分是根据作答情况给出相应分值)上的答题情况进行诊断分析<sup>[8]</sup>. 针对这些传统认知诊断模型的局限性,随着大数据分析挖掘的不断发展,出现许多和数据挖掘方法相结合的认知诊断模型. Wu 等<sup>[29]</sup> 提出面向学生个性化学习的模糊认知诊断分析框架(Fuzzy Cognitive Diagnosis Framework, FuzzyCDF).

图 5 为 FuzzyCDF 的框架图,此框架同时考虑学生在客观题和主观题上的作答过程,形式化建模成一个 4 层的生成模型. 第 1 层是学生自身的潜在特性(能力),影响学生技能掌握程度的先验, FuzzyCDF 模型认为,学生的技能掌握程度受潜在特征的影响,假设每个学生  $j$  有一个潜在特征  $\theta_j$ . 第 2 层为学生当前的认知状态(针对技能的掌握程度),结合学生  $j$  潜在特性  $\theta_j$  和 IRT 理论中的试题参数共同决定,概率化学生  $j$  在技能  $k$  上的掌握程度  $\alpha_{jk}$ . 第 3 层根据学

生  $j$  的技能掌握程度,获得学生试题掌握程度  $\eta_{ij}$ ,在 FuzzyCDF 的第 3 层中,基于模糊系统分别建模学生在主观题和客观题上不同的答题模式. 具体地, FuzzyCDF 分别考虑主观题和客观题 2 种测试题型关联的知识点掌握模式(主观题上的技能掌握模式满足教育学中“补偿型”知识点假设,在客观题上的掌握模式满足“链接型”假设),并结合专家标注的  $Q$  矩阵和模糊系统理论生成学生对试题的掌握程度  $\alpha_{jk}$ (使用模糊并建模“补偿型”知识点关联的客观题,使用模糊交建模“连接型”知识点关联的主观题). FuzzyCDF 的第 4 层为学生答题结果的生成,进一步考虑传统认知心理学中学生答题时存在的“失误”和“猜测因素”,即学生  $j$  在掌握试题的情况下可能因为失误( $s_i$ )导致对试题  $i$  的答题错误,或是在没有掌握试题的情况下通过猜测( $g_i$ )答对试题  $i$ .

图 5 四层 FuzzyCDF 框架图<sup>[29]</sup>Fig. 5 4-step FuzzyCDF framework<sup>[29]</sup>

经过此四步的生成模型, FuzzyCDF 最终建模得到学生在试题上的最终得分结果. 图 6 为 FuzzyCDF 的图模型(该模型只需要两类观察变量:观察到学生的做题得分  $R_{ij}$ , 试题关联的知识点  $q_{ik}$ ),与图 5 中的步骤及本段中的描述相对应.

图 6 FuzzyCDF 图模型<sup>[29]</sup>Fig. 6 Graphical model of FuzzyCDF<sup>[29]</sup>

DINA 模型只能给出学生的离散化的认知状态

诊断结果(即输出的 1,0 值分别表示学生在某个知识点上掌握或未掌握),而 FuzzyCDF 可以进一步将诊断得到的学生认知状态细化为具有更多信息量的连续值(0 ~ 1 之间的数值,越大表示学生掌握程度越高).因此,通过 FuzzyCDF 诊断得到的结果具有更强的可扩展性、抗噪声性和可解释性.例如,图 7 在同个极坐标下分别展示 FuzzyCDF 和 DINA 模型诊断得到的学生知识点掌握程度结果对比.由图可见,对于知识点“去分母”而言,FuzzyCDF 诊断学生在该知识点上的掌握程度为 0.77,这种概率化的诊断结果更精确,且对在线学习提供的诸多面向学生的应用具有较大帮助.

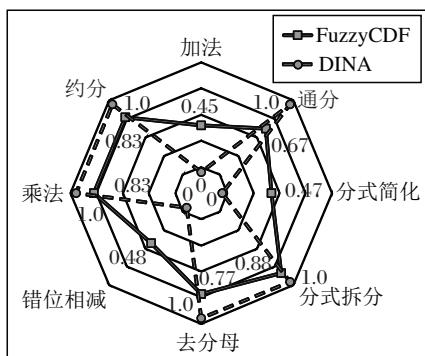


图 7 FuzzyCDF 与 DINA 的诊断结果对比<sup>[29]</sup>

Fig. 7 Comparison of diagnosis result of FuzzyCDF and DINA<sup>[29]</sup>

### 1.2.3 学生游戏心理估计与诊断

在线学习作为一种基于互联网的新型智能教学方式,在快速发展的同时也面临着许多传统的线下教育未曾遇到的问题<sup>[30-31]</sup>.在面向学生的认知诊断分析任务中,就面临这一情况:学生对试题的作答受自身能力和游戏心理的共同影响.学生自身能力即为知识点掌握程度等潜在特征,而游戏心理是指通过固定的答题方法或是猜测(敷衍)回答问题.那么,能否量化学生在作答时的游戏心理?Wu 等<sup>[32]</sup>提出“知识-猜测”叠加反应模型(Knowledge Plus Gaming Response Model, KPGRM),对在线学习中的学生游戏心理进行诊断和分析.

图 8 为学生在线答题的示例,从图中的答题数据部分可以看出,学生 3 通过不断猜测最终选择正确答案,学生 2 第一次作答就选择正确答案.对于这两类不同的答题情况,KPGRM 通过分析学生在多次答题(Multiple-Attempt Responses, MAR)下的显式游戏心理因子和单次答题(One-Attempt Response, OAR)下的显式和隐式猜测因子,建模并诊断学生的游戏心理及认知情况.

现有的研究认为,如果在线选择题的答题形式为学生只有答对当前试题才能进行下一题作答,那么学生的游戏行为主要有持续作答、系统作答和快速作答这 3 类<sup>[33-34]</sup>.KPGRM 基于这 3 种学生游戏行为,提出 4 个用于诊断显式的学生游戏心理的特征:1)在一道题上的答题尝试次数越多,对应的游戏心理因子越高(持续作答,Len);2)答题速度越快,游戏心理越强(快速作答,Spd);3)答题过渡越高(Cov),4)选择的选项对所有选项的覆盖度越大,都标志更高的游戏心理因子(系统作答,Trs).在对学生答题时的 4 个特征进行观测后,使用线性组合的方法对这 4 个特征进行组合,即可分析出学生的显式游戏心理因子.

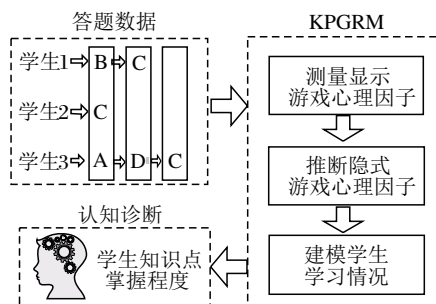


图 8 KPGRM 框架

Fig. 8 KPGRM framework

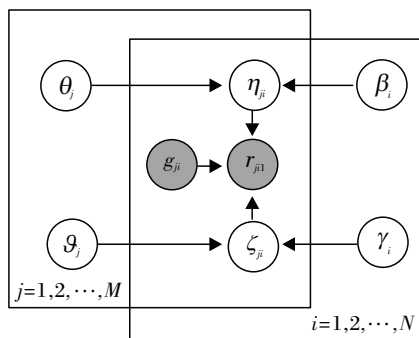


图 9 KPGRM 图模型<sup>[32]</sup>

Fig. 9 Graphical model of KPGRM<sup>[32]</sup>

相比学生对一道试题多次作答的情况,如果学生在第一次进行猜测时就选到正确答案,则没有那么丰富的信息用于分析学生的隐式游戏心理.因此,KPGRM 使用协同过滤(Collaborative Filtering, CF)对这种隐式的学生游戏心理进行分析.具体地,KPGRM 将每位学生、试题都映射到  $D$  维的空间中,使用一个向量描绘学生的潜在心理特征及对应试题的潜在特性.在应用潜在因子模型时,KPGRM 将学生的游戏心理因子看作学生与试题间的交互,并使



用已观测到的学生显式游戏心理因子作为基础,对相似学生的隐式游戏心理因子进行分析.

KPGRM 使用信号检测模型从含有噪声的观测(首次作答的结果  $r_{ji}$ ) 中检测学生  $j$  在回答试题  $i$  时的游戏心理因子  $g_{ji}$ . 分别考虑如下 2 种情况.

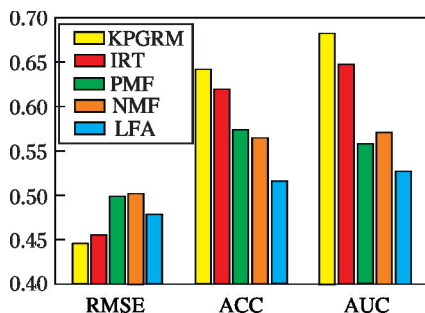
1) 学生  $j$  在没有任何游戏心理因子的情况下正确回答试题  $i$ , 建模学生学习状态  $\eta_{ji}$ .

2) 学生  $j$  完全依靠游戏心理因子(如猜测等)正确回答试题  $i$ , 建模学生学习状态  $\zeta_{ji}$ .

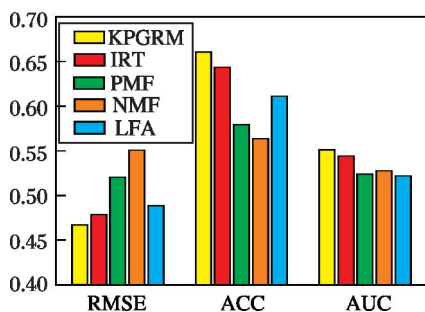
假设学生在每道试题上的作答是独立的情况,使用伯努利分布建模学生所有的第一次作答情况(答对或答错),表示如下:

$$P(r_{ji} = 1 | \theta_j, \beta_i, \vartheta_j, \gamma_i, g_{ji}) = \eta_{ji}^{1-g_{ji}} \zeta_{ji}^{g_{ji}}.$$

KPGRM 模型通过学生多次回答问题时的特征观测显式游戏心理因子,并使用协同过滤的方法推断隐式游戏心理因子. 对比实验方法包括:IRT、概率矩阵分解(Probabilistic Matrix Factorization, PMF)、非负矩阵分解(Non-negative Matrix Factorization, NMF) 和因子分析(Learning Factor Analysis, LFA), 各方法在学生得分预测表现在 Alone 数据集和 Earth 数据集上的对比结果如图 10 所示. 从图中可以看出, KPGRM 可以对学生得分进行更准确的预测.



(a) Alone



(b) Earth

图 10 KPGRM 的得分预测结果对比<sup>[32]</sup>

Fig. 10 Predicting result comparison of KPGRM<sup>[32]</sup>

此外,文献[32]中的实验还补充对游戏心理因子的分析.具体地,随机选取 Alone 数据集中 20 道 MAR 模式下的学生答题记录,并邀请 11 名志愿者进行学生猜测情况标记 F(1、0.5 和 0 分别表示“游戏心理”、“不确定”和“非游戏心理”),并计算平均受试者工作特征曲线下的面积(Area Under Receiver Operating Characteristic Curve, AUC).表 3 展示模型诊断得到的游戏心理因子和志愿者标注的对比情况,达到可接受的属性一致分析结果(Fleiss's  $\kappa = 0.73$ ).从表中可以看出, KPGRM 的测量结果和志愿者标注结果相符度较高.

表 3 KPGRM 结果和志愿者标记结果对比<sup>[32]</sup>

指标	Len	Spd	Trs	Cov	F
AUC	0.598	0.723	0.923	0.945	<b>0.966</b>

综上所述,在结合学生游戏心理因子的情况下, KPGRM 使用信号检测模型诊断学生学习状态,实现学生在线答题时游戏心理的分析,同时还能更精确地诊断学生的认知状态.

### 1.3 教学辅助应用

对于智慧教育的研究,是为了更好地辅导学生学习.因此,除了面向试题(试题表征)和面向学生(认知诊断),初步探索在线学习中的智能教学辅助.

#### 1.3.1 教学辅助研究现状

在面向在线学习的教育数据挖掘中,认知诊断方法可以帮助分析学生的学习状况<sup>[8]</sup>.学生通常需要不断地进行试题练习,用于学习、巩固学习的内容,因此,在线学习面向教学的核心任务之一就是如何基于学生不同的认知状态(知识点掌握程度)向每个学生推荐适合的习题(试题).

在传统的线下教学过程中,即使是对于知识点掌握程度不同的学生,通常也由教师以班级为单位进行作业、习题的布置.从图 7 的学生知识点掌握雷达图中可以看出,该学生在“分式化简”和“分式拆分”这两个知识点上的掌握程度区别较大,因此,每个具有不同认知状态的学生,对练习的试题需求也应当不同.伴随在线学习系统中试题库的不断扩充,传统方法不但难以对每个学生不同的学习状态进行有针对性的训练,人工筛选试题也会花费大量的时间.所以,如何自动为每个学生生成、推荐个性化试题,是在线教育领域中面临的一个重要问题.

近些年,Thai-Nghe 等<sup>[35]</sup>基于教育领域中的认



知诊断思想,研究基于学生得分预测的试题推荐方法,将学生的答题结果映射为“试题-知识点关联”及“学生-知识点掌握”两个低秩矩阵(表 1、表 2 即为某种映射的示例).该方法通过预测学生在未知试题上的得分,选择合适的题目进行推荐.因为学生的知识点掌握状态具有一定的隐蔽性,诊断结果可能不够准确,因此,再使用这类基于认知诊断的试题推荐方法进行试题推荐时,推荐准确度也会较差.此外,朱天宇等<sup>[14]</sup>使用传统的数据挖掘方法开展学生试题推荐工作,这类方法基于协同过滤的思想对学生的共性进行建模,较好地解决教师在筛选试题时耗时耗力的问题.然而,这类数据挖掘方法的效果易受学生做题数据稀疏性的影响,最终的题目推荐结果存在可解释性不足的问题,即难以向学生、教师说明推荐某一道试题的原因(如学生因为对某一知识点掌握程度不足而进行推荐<sup>[16]</sup>).因此,在实际应用中也具有一定的局限性.

针对学生知识点掌握情况的认知诊断分析不仅可用于在线教育的试题推荐,还可将分析得到的学生认知诊断情况应用于离线的课堂教育中,对学生群体的分析也是教育数据挖掘中重要的一部分<sup>[36]</sup>.

班级作为学生离线学习的一个重要单位,对班级里的学生进行建模,发现班内学生之间的相似性和差异性对离线教学具有较大帮助.学生协同学习分组<sup>[37]</sup>是离线学习中的一种重要学习形式,它通过将班级内的学生划分为 5~8 人学习小组的形式,组织每个学习小组内的学生进行学习,同一学习小组内的学生通过互相交流、讨论,达到提升学业水平的目的.传统的协同学习分组有人工选择和自动生成等方式,学生自主构建学习小组的人工选择方法容易聚集相似的学生,难以达到分组学习、互相提高的目的,而教师主导分组的方法分组效率较低,难以保证分组效果.此外,基于模拟特征的协同学习分组方法缺乏对分组效果的验证,而基于学生得分特征的分组方法粒度过粗,无法根据学生的实际知识点掌握水平进行合理的分组.

### 1.3.2 基于认知诊断的个性化试题推荐

针对认知诊断在个性化推荐教学辅助方面的应用,朱天宇等<sup>[14]</sup>通过结合数据挖掘的推荐方法,提出可增强试题推荐结果可靠性和可解释性的试题推荐方法(Probabilistic Matrix Factorization-Cognitive Diagnosis, PMF-CD),如图 11 所示.

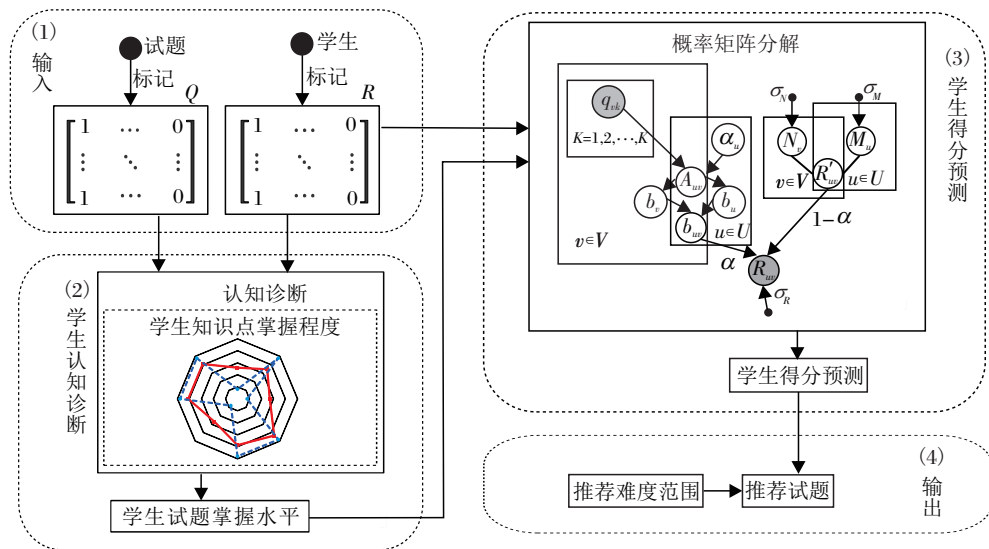


图 11 PMF-CD 个性化试题推荐<sup>[14]</sup>

Fig. 11 Personalized questions recommended PMF-CD framework<sup>[14]</sup>

PMF-CD 试题推荐框架主要包括数据输入、学生认知诊断、学生得分预测及试题推荐 4 个步骤.首先根据已有的学生答题记录,利用教育数据挖掘的认知诊断分析方法获得每个学生的认知状态.然后,

在已知学生知识点掌握程度的情况下,通过将已观测到的学生答题情况和试题对应的知识点作为先验,计算学生在每道试题上的实际掌握水平(即排除猜测和失误后的学生真实水平).在获得学生的

试题掌握水平后,再计算学生和试题的先验得分情况,并将其应用于概率矩阵分解 (Probabilistic Matrix Factorization, PMF). PMF 可以获得学生和试题的隐向量,并对学生的答题情况进行预测. 最后,模型根据所需试题的难度范围,筛选得分预测在适合范围内的试题并形成待推荐试题集. 经过 PMF-CD 模型的处理,每个学生会得到一份结合自身的知识点掌握情况和相似学生答题情况共同生成的个性化试题.

在提出的个性化试题推荐框架中,可以根据学生的不同学习需求,有针对性地向学生推荐不同难度的试题. 如图 12 所示. 针对需要提升薄弱知识点的掌握程度,提高知识点掌握程度的学生,可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,可以向学生推荐难度较低(简单)的试题.

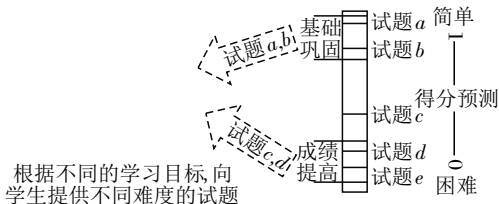
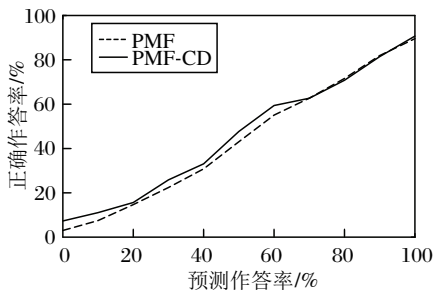


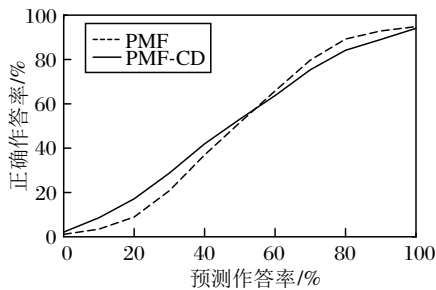
图 12 有针对性的试题推荐<sup>[14]</sup>

Fig. 12 Question Recommendation<sup>[14]</sup>

实验将 PMF-CD 框架应用于两组某市的高中学生实际答题记录的数据集 (DATASET1、DATASET2), 并进行试题推荐. 从图 13 中的结果可以看出, PMF-CD 在向学生进行不同难度的试题推荐时, 相比未引入认知诊断的推荐方法 (PMF), 均可以得到更好的推荐效果 (PMF-CD 的预测结果更贴近学生真实的答题率, 即推荐的简单题不会过于简单, 推荐的难题不会过于难).



(a) DATASET1



(b) DATASET2

图 13 不同推荐难度对推荐效果的影响<sup>[14]</sup>

Fig. 13 Influence of difficulty on results<sup>[14]</sup>

### 1.3.3 学生协同学习分组推荐

如前所述,除将对学生的认知诊断结果应用于个性化试题推荐之外,还可以根据班级内学生的知识点掌握水平,进行协同学习分组的推荐. 针对 1.3.1 节中提到的学生协同学习分组方法存在的相关问题, Liu 等<sup>[38]</sup> 研究基于学生认知诊断分析的学生协同学习智能分组方法. 如图 14 所示,该框架分为输入、学生建模、学生分组及输出 4 个部分. 输入部分为班级中待分组的学生及收集的学生答题情况. 首先根据学生的试题作答情况对学生进行认知诊断,分析软性确定性输入噪音与门模型 (Soft-Deterministic Inputs Noisy and "Gate", S-DINA), 获得学生在每个独立的知识点上的概率化掌握程度. 然后根据每个学生不同的知识点掌握情况,进行学生分组,并分别提出基于学生差异的分组算法 (Uniform  $k$ -means Based, UKB) 和基于收益的学生分组算法 (Balanced Gain Based, BGB). 如图 15 所示, 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,旨在将同类的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,对学生进行分组.

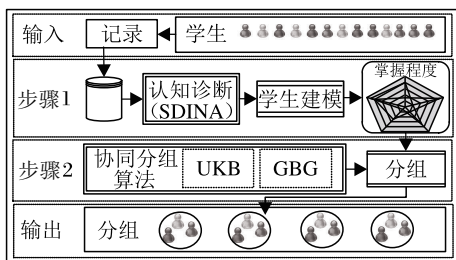


图 14 学生协同学习分组框架<sup>[38]</sup>

Fig. 14 Collaborative learning team formation framework<sup>[38]</sup>

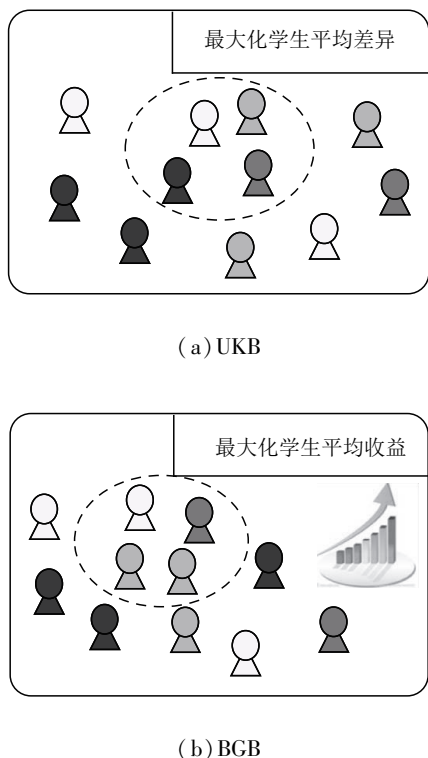


图 15 UKB、BGB 分组方法示意图<sup>[38]</sup>

Fig. 15 Sketch map of team formation result of UKB, BGB<sup>[38]</sup>

将学生协同学习分组方法应用于某市高中的班级学生协同学习分组推荐,图 16 为某分组中学生知识点掌握水平的柱状图表示,从图中可以看出,该分组推荐方法可以平衡小组内学生 ( $SD-Stu_1 \sim SD-Stu_5$ ) 的知识点 ( $S_1 \sim S_5$ ) 掌握水平,让学生可以互相学习、共同进步.此外,实验中还引入授课教师对协同学习分组结果的真实评估.具体地,该调查分别使用根据学生得分、DINA 模型诊断得到的二元技能掌握程度及本文中 S-DINA 模型诊断得到的概率化技能掌握程度 3 种特征,输出学生分组.

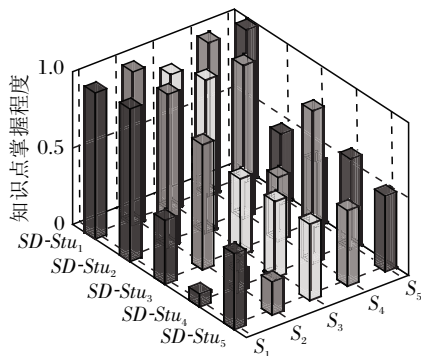


图 16 协同学习分组示例<sup>[38]</sup>

Fig. 16 Team formation example<sup>[38]</sup>

表 4 为授课教师对这些分组的盲选结果,从命中率来看,本文所提的概率化技能掌握程度特征可以取得与授课教师意见更相符的协同学习分组结果.

表 4 授课教师调查结果<sup>[38]</sup>

Table 4 Teacher evaluation

特征	命中率/%
学生得分	6.6
二元技能掌握程度	26.7
概率化技能掌握程度	66.7

## 2 面向在线智慧教育的数据挖掘技术应用

对于在线学习各方面的研究工作,最终目的是将研究成果转化为智慧教育应用,帮助在线教育系统更好地辅导学生学习.

### 2.1 在线自适应学习

Brusilovsky 等<sup>[39]</sup>提出自适应学习的概念后,得到国内外研究人员的重视.自适应学习通常指为学生提供学习环境、实例或平台,通过分析学生的学习过程,发现并总结理论,用于学生自主解决问题的学习形式.自适应学习可以看作个性化教育的一种实际应用,目前,已有在线学习网站向学生提供自适应学习的解决方案以满足每个学生不断变化的学习需求.“智学网”作为场景覆盖较广、智能化较高的学生答题数据采集、分析、应用平台之一,全程跟踪学生的学习过程,为每位学生打造个性化学习路径,促进学生的个性化学习,提高学业水平.

当前的自适应学习系统要求及时诊断学生的学习状态,根据每位学生的不同状态生成个性化学习内容,保证学习的高效性和合理性<sup>[39]</sup>.

1.3.1 节介绍的个性化试题推荐就是自适应学习系统的一个重要功能.由于国外的自适应学习系统发展较早,因此,相对于国内近些年才起步、发展的自适应学习平台,发展略为成熟. Knewton (<https://www.knewton.com>) 为美国一个个性化教育初创公司开发的较典型的个性化学习产品,该平台的核心就是向学生提供一系列的自动资源内容评估(评估系统内试题、视屏等学习资源的质量)、自适应诊断技术(学生认知能力诊断)以及推荐功能(针对每个学生的个性化学习资源推荐).目前,该平台已经覆盖全球 120 个国家,《快公司》(Fast



Company) 杂志 (<https://www.fastcompany.com>) 曾将其评为全球大数据领域最具创新力的十大公司之一。

## 2.2 离线教学辅助

对试题、学生的分析和诊断结果,不仅可应用于在线的自适应学习平台,还可用于辅助教师的离线教学<sup>[38]</sup>。1.3节提到的学生协同学习分组推荐,就是将学生分析诊断结果应用于离线(课堂)教学的案例之一。除了面向学生分析的离线教学应用之外,对试题的分析也对离线教学辅助具有一定的帮助,如根据试题的难度、知识点考察范围等进行自动组卷,保证试卷整体的科学性与合理性<sup>[39]</sup>。

## 2.3 智学网

目前,国内的自适应学习在线教育平台大多以系统资源构建和学生能力分析为主要目标,较难有效利用学习资源向不同能力的学生提供个性化服务,也难渗透到学生的日常学习教学场景中。为了探索国内 K12 教育与自适应学习系统结合的可能性,为学生提供科学、个性化的在线学习环境,科大讯飞公司智学网(zhixue.com)系统集成一套用于学生自适应学习的在线模块,并率先与全国各地的中学进行合作试点,取得一定的初步成效。

在提供在线自适应学习平台的同时,智学网还伴随式收集离线(在校)的教与学数据,全面分析学生的学习状态,并及时向授课教师反馈,帮助教师精准教研,帮助学生推荐个性化学习资源,促使有效提升。也就是说,在对学生进行认知诊断分析、试题推荐的同时,也能提供面向教师的班级学生学习状态反馈、智能组卷等服务。

文中介绍的部分教育数据挖掘技术及相关应用成果(如认知诊断分析、试题推荐、自动组卷等)已应用于智学网的学习平台。与此同时,通过学生对所推荐资源的学习反馈,形成“学习-诊断-推荐-反馈-学习”的闭环结构。目前,智学网初步搭建的大数据在线教育平台已覆盖全国 29 个省 230 多个城市,提供的服务让近 13 000 所学校的千万师生受益。

## 3 结束语

随着信息技术的发展和网络的普及,一方面,各种在线教育平台层出不穷,积累海量的教育资源以及教学数据,产生许多的智慧学习应用需求(如自适应学习等)。另一方面,在线教育的持续发展为面向智慧教育的数据挖掘技术研究提供良好的契机和条件,越来越丰富、高质量的教育数据让教育心理

学、认知心理学等和计算机科学更紧密结合,从而设计更有效的模型、方法和应用系统。在此背景下,本文综述针对在线教育中面向试题(教学资源)、面向学生和面向教学 3 个方面进行若干教育数据挖掘技术的研究探索工作。

可以预测,随着在线学习的不断发展和完善,该领域内的数据挖掘基础研究与应用工作仍将在诸多方面不断产生技术革新<sup>[20]</sup>。

1)学生的多维认知因素建模。众所周知,学生的学习过程受到诸多因素的共同影响,如“猜测”、“粗心”等教育心理学中已经发现的因素。通过数据挖掘技术对学生学习中出现的众多不确定因素进行准确的建模,突破传统认知诊断技术基于小样本或单因素分析的局限性,可以更好地帮助在线教育系统、授课教师了解学生的学习状态,帮助提高学生的学业水平。

2)跨学科的知识迁移学习分析。当前数据挖掘中针对学生认知诊断的工作通常围绕某一门独立学科的学生答题数据展开。然而,学生所学的不同学科(如数学、物理等)之间并不完全独立,相互之间可能存在知识转移现象。因此,利用迁移学习等数据挖掘方法寻找学生在学习中共有的模式、知识结构,对于全面了解学生的学习状态具有重要意义。

3)在线学习与离线教育的融合。纵然面向在线学习的教育数据挖掘研究与应用具有良好的前景和预期,但是为学生的日常学习、教师的课堂教学进行辅助服务也应当是在线学习平台的根本目标之一。因此,针对中小学的在线教育平台当前应以辅助传统的离线(课堂)教育为主,不能完全取而代之。所以,面向中小学的在线学习研究与应用的关注点除本文提到的面向试题、面向学生和面向教学辅助等内容,如何实现在线教育和离线教育的有效融合,也是一个重要的研究方向。

除此之外,相关研究人员应当关注在线学习和离线辅助教育效果的衡量标准选择问题,如考虑到仅利用学生的成绩进行学业水平的衡量过于片面,学者们引入认知诊断的分析方法。但是,由于学生的知识点掌握水平属性仍存在一定的隐蔽性<sup>[41]</sup>等原因,如何综合多方面的因素,综合衡量学生的学习状态仍是需要关注、研究的问题。

## 参 考 文 献

[1] 傅钢善. 教育技术发展轨迹探讨. 电化教育研究, 2005(9): 22-26.

- (FU G S. Exploration of Educational Technology Development. e-Education Research, 2005(9): 22-26. )
- [2] HONG C M, CHEN C M, CHANG M H, *et al.* Intelligent Web-Based Tutoring System with Personalized Learning Path Guidance. Computers & Education, 2008, 51(2): 787-814.
- [3] BRESLOW L, PRITCHARD D E, DEBOER J, *et al.* Studying Learning in the Worldwide Classroom Research into edX's First MOOC. Research & Practice in Assessment, 2013, 8: 13-25.
- [4] POLSON M C. Foundations of Intelligent Tutoring Systems. Hove, UK: Psychology Press, 2013.
- [5] 陈 池,王宇鹏,李 超,等. 面向在线教育领域的大数据研究及应用. 计算机研究与发展, 2014, 51(S1): 67-74.  
(CHEN C, WANG Y P, LI C, *et al.* The Research and Application of Big Data in the Field of Online Education. Journal of Computer Research and Development, 2014, 51(S1): 67-74. )
- [6] 管 佳,李奇涛. 中国在线教育发展现状,趋势及经验借鉴. 中国电化教育, 2014(8): 62-66.  
(GUAN J, LI Q T. China's Online Education Current Situation, Trend and Experience for Reference. Chinese Journal of Computers, 2014(8): 62-66. )
- [7] PEÑA-AYALA A. Educational Data Mining: A Survey and a Data Mining-Based Analysis of Recent Works. Expert Systems with Applications, 2014, 41(4): 1432-1462.
- [8] LEIGHTON J P, GIERL M J. Cognitive Diagnostic Assessment for Education: Theory and Applications. Cambridge, UK: Cambridge University Press, 2007.
- [9] ANDERSON T. The Theory and Practice of Online Learning. Athabasca, Canada: Athabasca University Press, 2008.
- [10] MCAULEY A, STEWART B, SIEMENS G, *et al.* The MOOC Model for Digital Practice[M/OL]. [2017-11-20]. [http://elearning.org/Articles/MOOC\\_Final.pdf](http://elearning.org/Articles/MOOC_Final.pdf).
- [11] YUDELSON M, HOSSEINI R, VIHAVAINEN A, *et al.* Investigating Automated Student Modeling in a Java MOOC // Proc of the 7th International Conference on Educational Data Mining. New York, USA: ACM, 2014: 261-264.
- [12] CHAKOR Y A. Abandonment of Learners MOOC Problematic Analysis and Proposed Solutions. International Journal of Computer Applications, 2016, 153(2): 35-37.
- [13] BRUFF D, FISHER D, MCEWEN K, *et al.* Wrapping a MOOC: Student Perceptions of an Experiment in Blended Learning[J/OL]. [2017-11-20]. <https://my.vanderbilt.edu/douglasfisher/files/2013/06/JOLTPaperFinal6-9-2013.pdf>.
- [14] 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, 41(1): 176-191.  
(ZHU T Y, HUANG Z Y, CHEN E H, *et al.* Cognitive Diagnosis Based Personalized Question Recommendation. Chinese Journal of Computers, 2017, 41(1): 176-191. )
- [15] KONGSAKUN K, FUNG C C. Neural Network Modeling for an Intelligent Recommendation System Supporting SRM for Universities in Thailand. WSEAS Transactions on Computers, 2012, 11(2): 34-44.
- [16] DE LA TORRE J. DINA Model and Parameter Estimation: A Didactic. Journal of Educational and Behavioral Statistics, 2009, 34(1): 115-130.
- [17] MIDGLEY C. Goals, Goal Structures, and Patterns of Adaptive Learning. New York, USA: Routledge, 2014.
- [18] LÜ Y J, ZHAO G, MIAO P, *et al.* Construction of Multidimensional Dynamic Knowledge Map Based on Knowledge Requirements and Knowledge Connection // WEN Z, LI T, eds. Knowledge Engineering and Management. Berlin, Germany: Springer, 2014: 83-94.
- [19] HUDAK G M, KIHN P. Labeling: Pedagogy and Politics. New York, USA: Routledge, 2014.
- [20] 刘 淇,陈恩红,黄振亚,等. 面向个性化学习的学生认知能力分析. 中国计算机学会通讯, 2017, 13(4): 28-34.  
(LIU Q, CHEN E H, HUANG Z Y, *et al.* Cognitive Analysis for Personalized Student Learning. Communications of the CCF, 2017, 13(4): 28-34. )
- [21] BOOPATHIRAJ C, CHELLAMANI K. Analysis of Test Items on Difficulty Level and Discrimination Index in the Test for Research in Education. International Journal of Social Science & Interdisciplinary Research, 2013, 2(2): 189-193.
- [22] ALAGUMALAI S, CURTIS D D. Classical Test Theory // MACLEAN R, WATANABE R, BAKER R, *et al.*, eds. Applied Rasch Measurement: A Book of Exemplars. Berlin, Germany: Springer, 2005: 1-14.
- [23] ZHANG L, YANG L. Repeater Analyses for TOEFL iBT[J/OL]. [2017-11-20]. <https://www.ets.org/Media/Research/pdf/RM-08-05.pdf>.
- [24] FUCHS L S, FUCHS D, HAMLETT C L, *et al.* Effects of Expert System Consultation within Curriculum-Based Measurement, Using a Reading Maze Task. Exceptional Children, 1992, 58(5): 436-450.
- [25] HUANG Z Y, LIU Q, CHEN E H, *et al.* Question Difficulty Prediction for READING Problems in Standard Tests // Proc of the 31st AAAI Conference on Artificial Intelligence. Palo Alto, USA: AAAI Press, 2017: 1352-1359.
- [26] DI BELLO L V, ROUSSOS L A, STOUT W. 31A Review of Cognitively Diagnostic Assessment and a Summary of Psychometric Models. Handbook of Statistics, 2006, 26(6): 979-1030.
- [27] EMBRETSON S E, REISE S P. Item Response Theory. Hove, UK: Psychology Press, 2013.
- [28] RUPP A A, TEMPLIN J. The Effects of Q-Matrix Misspecification on Parameter Estimates and Classification Accuracy in the DINA Model. Educational and Psychological Measurement, 2008, 68(1): 78-96.
- [29] WU R Z, LIU Q, LIU Y P, *et al.* Cognitive Modelling for Predicting Examinee Performance // Proc of the 24th International Joint Conference on Artificial Intelligence. New York, USA: ACM, 2015: 1017-1024.
- [30] BAKER R S, CORBETT A T, KOEDINGER K R. Detecting Student Misuse of Intelligent Tutoring Systems // Proc of the International Conference on Intelligent Tutoring Systems. Berlin, Germany: Springer, 2004: 531-540.
- [31] VAESSEN B, PRINS F, JEURING J. University Students' Achievement Goals and Help-Seeking Strategies in an Intelligent Tutoring System. Technical Report, UU-CS-2013-019. Utrecht, The Netherlands: Utrecht University, 2013.

- [32] WU R Z, XU G D, CHEN E H, *et al.* Knowledge or Gaming? Cognitive Modelling Based on Multiple-Attempt Response // Proc of the 26th International World Wide Web Conference. New York, USA: ACM, 2017: 321–329.
- [33] BAKER R S, CORBETT A T, KOEDINGER K R, *et al.* Adapting to When Students Game an Intelligent Tutoring System // Proc of the International Conference on Intelligent Tutoring Systems. Berlin, Germany: Springer, 2006: 392–401.
- [34] BAKER R S, CORBETT A T, KOEDINGER K R, *et al.* Off-Task Behavior in the Cognitive Tutor Classroom: When Students Game the System // Proc of the SIGCHI Conference on Human Factors in Computing Systems. New York, USA: ACM, 2004: 383–390.
- [35] THAI-NGHE N, DRUMOND L, HORVATH T, *et al.* Factorization Techniques for Predicting Student Performance [ J/OL ]. [2017-11-20]. [https://www.ismll.uni-hildesheim.de/pub/pdfs/Nguyen\\_et\\_al\\_ERSAT\\_2011.pdf](https://www.ismll.uni-hildesheim.de/pub/pdfs/Nguyen_et_al_ERSAT_2011.pdf).
- [36] AGRAWAL R, GOLSHAN B, TERZI E. Grouping Students in Educational Settings // Proc of the 20th International Conference on Knowledge Discovery and Data Mining. New York, USA: ACM, 2014: 1017–1026.
- [37] GALL M D, GALL J P. The Discussion Method. *The Psychology of Teaching Methods*, 1976, 75(S1): 166–216.
- [38] LIU Y P, LIU Q, WU R Z, *et al.* Collaborative Learning Team Formation: A Cognitive Modeling Perspective // Proc of the 21st International Conference on Database Systems for Advanced Applications. Berlin, Germany: Springer, 2016: 383–400.
- [39] BRUSILOVSKY P, PEYLO C. Adaptive and Intelligent Web Based Educational Systems. *International Journal of Artificial Intelligence in Education*, 2003, 13(2/3/4): 159–172.
- [40] 路平, 王敏娟, 万昆. 试题库自动组卷策略研究. *电脑开发与应用*, 2001, 14(2): 9–11.  
(LU P, WANG M J, WAN K. Tactics Study on Automatic Creating Test Paper from Item Pool. *Computer Development & Application*, 2001, 14(2): 9–11.)
- [41] 戴海奇, 张峰, 陈雪枫. *心理与教育测量*. 第3版. 广州: 暨南大学出版社, 2011.  
(DAI H Q, ZHANG F, CHEN X F. *Psychological and Educational Measurement*. 3rd Edition. Guangzhou, China: Jinan University Press, 2011)

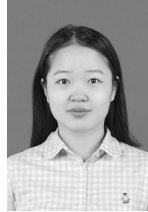
## 作者简介



刘淇, 博士, 副教授, 主要研究方向为数据挖掘、知识发现、机器学习. E-mail: qiliuq@ustc.edu.cn.  
(LIU Qi, Ph. D., associate professor. His research interests include data mining, knowledge discovery and machine learning.)



陈恩红(通讯作者), 博士, 教授, 主要研究方向为机器学习、数据挖掘、社会网络、个性化推荐系统. E-mail: cheneh@ustc.edu.cn.  
(CHEN Enhong (Corresponding author), Ph. D., professor. His research interests include machine learning, data mining, social network and personalized recommender system.)



朱天宇, 硕士研究生, 主要研究方向为数据挖掘、推荐系统. E-mail: zhtianyu@mail.ustc.edu.cn.  
(ZHU Tianyu, master student. Her research interests include data mining and recommender system.)



黄振亚, 博士研究生, 主要研究方向为教育数据挖掘、推荐系统. E-mail: huangzhy@mail.ustc.edu.cn.  
(HUANG Zhenya, Ph. D. candidate. His research interests include educational data mining and recommender system.)



吴润泽, 博士研究生, 主要研究方向为教育数据挖掘、认知诊断. E-mail: wrz179@mail.ustc.edu.cn.  
(WU Runze, Ph. D. candidate. His research interests include educational data mining and cognitive diagnosis.)



苏喻, 博士研究生, 主要研究方向为数据挖掘、图像识别. E-mail: firesysysy@163.com.  
(SU Yu, Ph. D. candidate. His research interests include data mining and image identification.)



胡国平, 博士, 工程师, 主要研究方向为智能语音、语言核心技术研究. E-mail: gphu@iflytek.com.  
(HU Guoping, Ph. D., engineer. His research interests include intelligent voice and language core technology.)