

文章编号: 1672-5913(2024)04-0091-05

中图分类号: G642

面向个性化学习的认知诊断实践探索

袁 恩, 陈卫卫, 张所娟, 唐艳琴, 吴永芬, 张文字, 刘 凌

(陆军工程大学 指挥控制工程学院, 江苏 南京 210007)

摘 要: 针对目前开展混合式教学的现状, 提出将知识图谱引入到教学中, 利用知识图谱构建课程的知识结构, 对课程的在线平台资源进行标注, 以数据结构与程序设计课程为例, 阐述如何针对学生在在线平台上练习的结果, 选择合适的认知诊断技术进行分析, 获得学生各个知识点的掌握情况, 为个性化学习的开展提供有力支撑。

关键词: 认知诊断; 知识图谱; 个性化学习; 混合式教学

DOI:10.16512/j.cnki.jsjy.2024.04.021

1 背 景

因材施教是现代化教育的重要理念。2019年, 中共中央、国务院印发的《中国教育现代化2035》将“因材施教”作为推进教育现代化的八大基本理念之一, 并且指出“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”^[1]。2020年, 中共中央、国务院印发的《深化新时代教育评价改革总体方案》指出要“坚持面向人人、因材施教、知行合一”^[2]。可见, 在现代化教育中, 教师要以学生为中心, 坚持因材施教, 根据学生的个性特点开展个性化的教学活动, 使每一名学生能够个性化学习。

另一方面, 近年来由于信息技术的发展以及疫情的影响, 混合式教学成为当前的热点^[3]。2020年教育部公布的国家级一流本科课程中, 线上线下混合式课程占比约17%^[4]。混合式教学是将传统的线下教学与基于信息化的在线学习结合的混合教学方式。随着信息技术的发展, 出现了各种各样的在线教学平台, 例如, 中国大学MOOC平台^[5], 教学辅助平台PTA^[6]和

EduCoder^[7]等, 这些在线教学平台极大地推动了混合式教学的发展, 同时也积累了大量学生学习数据, 为个性化学习的开展创造了有利条件。

然而, 各个在线平台存在海量的学习资源, 需要使用信息化技术对学生的进行学习情况进行智能化分析, 为学生提供精准的教学资源推荐^[8]。为做到这一点, 一方面需要对课程的知识结构进行合理构建, 并利用课程构建的知识结构对课程资源进行分析标注, 另一方面, 需要对学生的学习情况进行智能分析以支持学生个性化学习。

2 知识图谱和认知诊断

2.1 知识图谱

知识图谱是谷歌公司为了提高搜索效率提出的, 它可以从非结构化的学习资源中抽取实体、关系等要素, 利用可视化描述知识资源及其载体, 并且利用知识图谱挖掘、分析、构建和显示知识之间的联系^[9]。知识图谱主要由节点和边组成, 节点对应现实世界中的“实体”, 边则代表实体之间的关系。知识图谱在教育领域的应用可以从多维视角进行认知: 从知识建模视角, 知

基金项目: 教育部产学研合作协同育人项目(202101151012, 202101244033); 教育部第二批虚拟教研室建设试点(教高厅函[2022]13号之216号); 全国教育科学国防军事教育学科“十四五”规划课题“个性化学习与智能教育深度融合的课程改造研究”(JYKY-D2022013); 国家自然科学基金项目“面向军事教育的情境感知认知诊断与学习推荐方法研究”(62207031); 陆军院校虚拟教研室探索建设项目(参训函[2022]215号); 陆军工程大学十四五首批精品课程建设项目“数据结构与程序设计”; 2021 教学成果立项培育项目。

第一作者简介: 袁恩, 男, 讲师, 研究方向为计算机应用、数据分析, yuanenem@foxmail.com。

识图谱是对学科知识本体进行建模的方法；从知识导航视角，知识图谱能够在大数据、人工智能等技术的支持下，生成面向学习目标的个性化学习路径^[10]。可见，知识图谱在教育中的运用可以从不同的角度切入，根据需要解决的问题或者需要达成的目标出发，合理地运用知识图谱。例如，文献[11]利用知识图谱挖掘课程之间的关系，辅助推荐个性化学习资源，以解决学习资源过载导致的学习者知识迷航。文献[12]则同样将知识图谱用于个性化学习资源推荐，通过基于知识图谱的相似资源匹配推荐学习资源。从上述研究来看，知识图谱和个性化学习之间有着密切的关系。然而，上述的工作主要是利用知识图谱挖掘课程之间或者学习资源与课程之间的关系，对于一门课程来说，更重要的是课程的知识点、知识点之间的关系以及课程资源与知识点之间的关系。

利用知识图谱可以对课程的知识进行建模，以知识点为实体，构建课程的知识图谱。对课程的知识图谱的构建，是对课程知识利用知识图谱进行梳理的过程。一方面，对课程所涉及的知识进行梳理，分析知识点之间的关系，构建课程知识点的知识图谱。这一过程是对课程知识点的深入分析和挖掘的过程，它是个性化学习的基础。另一方面，利用课程的知识图谱可以对课程资源进行标注，分析挖掘课程资源与课程知识点之间的关系。

2.2 认知诊断

对于课程的个性化学习不仅需要课程的知识图谱，而且需要对学习者的当前学习状态进行深入认知。认知诊断可以根据学生前期的学习数据，分析学生的学习状态，即学生对课程各个知识点的掌握情况。通过认知诊断对学生的认知进行建模，可以十分精确地获取学生在学习中的弱项，为后续学习的针对性提供基础。例如，可以根据学生当前的学习状态，智能地推荐个性化的试题^[13]。

典型的认知诊断方法包括项目反应理论IRT、受限潜在类别模型DINA和模糊认知诊断模型FuzzyCDF等^[14]。IRT是一种连续型认知诊断模型，它使用一个连续变量来刻画学生的学习能力，即学生学习的潜在特质。在一维的IRT

中，通过学生在一组试题上的得分来估计学生的潜在特质。研究者对IRT进行了拓展，引入了题目难度、区分度和猜测参数等进而发展出双参数IRT、三参数IRT和多参数IRT等。但是，IRT只能得到学生的整体学习水平，对于试题包含多个知识点的情况，没有办法获取学生在各个知识点上的掌握情况。

DINA和FuzzyCDF都是离散型认知诊断模型。在DINA中，将学生的能力用技能向量来表示，即1表示具备该技能，0表示不具备该技能。对于每一道题，都有对应的所需技能，构成Q矩阵，如图1所示。在DINA中，学生答对某一道题就必须具备该题所需的全部技能。因此，根据学生的技能向量和Q矩阵可得到学生针对选定试题的潜在响应向量，即学生在具备某些技能条件下针对选定试题的得分。考虑到学生可能在具备答题所需技能的条件下答错的概率和不具备答题所需技能条件下猜对答案的概率，分别引入失误概率和猜测概率。相对于IRT，DINA能够进一步挖掘题目与所需技能之间的关系，可用于评估学生技能掌握的情况。

	技能 1	技能 2	技能 3	技能 4
题目 1	1	0	1	0
题目 2	0	1	1	0
题目 3	1	1	1	0
题目 4	0	1	1	1
题目 5	1	1	1	1

图1 Q矩阵示例

DINA对学生技能的表示只有两种情况，即掌握和不掌握，没有中间状态，并且只有掌握了所有题目要求的技能，才能答对题目。这对于客观题来说比较合理，但是对于主观题则不太合理。FuzzyCDF将模糊集理论和模糊逻辑引入认知诊断，将学生的技能模糊化，映射到[0,1]的区间。同时，重新定义了学生对题目掌握程度的计算方式。对于客观题，根据学生的技能向量和Q矩阵，采用模糊交计算学生对题目的掌握情况，对于主观题，采用模糊并来计算学生对题目的掌握情况。

3 基于知识图谱的认知诊断实践

3.1 整体思路

混合式教学是当前教学改革的热点。它将传统课堂教学的优势和网络学习的优势结合起来,在教学过程中既发挥教师的主导作用,又体现学生的主动性和创造性。混合式教学利用了线上资源开展教学,积累了大量的学生学习数据,为学生个性化学习地开展提供了数据依据。在数据结构与程序设计课程中开展线上线下混合式教学的具体过程如图2所示。课程在中国大学慕课构建SPOC,供学生课前学习,利用雨课堂发布课前学习清单以及课前测,检验学员学习情况。课中,通过问题引导,与学生开展深入的交互式学习,并通过雨课堂完成后测。课后,利用PTA平台发布课后编程练习,并定期利用PTA平台开展阶段性测试。

为了实现个性化学习目标,首先梳理数据结构与程序设计课程知识点,以课程知识点为实体,构建数据结构与程序设计课程的知识图谱。然后,抽取PTA平台习题利用课程知识图谱对PTA平台习题进行标注,使得PTA资源与课程知识点之间建立联系。最后,基于学生在PTA平台的练习数据,利用认知诊断技术对学生知识点掌握情况进行分析,为个性化资源的推荐提供依据。具体过程如图3所示。

3.2 认知诊断方法对比实验

DINA和FuzzyCDF都具备对学员知识点掌握情况的评估能力,我们首先利用课程教学过程中PTA平台上的学生7次测试数据分析两种方

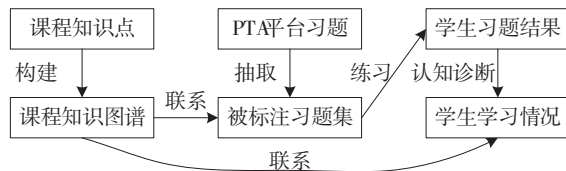


图3 学生学习情况认知诊断基本流程

法的有效性。数据集的基本信息见表1。PTA中题目的类型主要包括判断题、选择题、填空题、程序填空、函数题和编程题,将判断题、选择题、填空题归为客观题,程序填空、函数题和编程题归为主观题。从数据集可以看到,学生通过测试完成的题目数量不多,符合一般专业课程教学过程的实际情况。

表1 数据集基本信息

数据集	学生数量	知识点数	客观题数	主观题数
PTA 测验	141	10	158	23

可使用Python实现上述方法,并通过两种方法预测学生知识能力水平来对两种方法进行评估。FuzzyCDF模型的采样次数为5000,同时基于后2500次采样结果来估计模型参数。DINA中当前后两次参数估计小于0.0001时,停止迭代。实验结果如图4所示。

根据图4的结果,FuzzyCDF方法相比于DINA方法,预测的RMSE和MAE更小。并且,随着测试集比例增加、训练集比例减少,FuzzyCDF方法的RMSE和MAE变化不大,但是DINA方法的RMSE和MAE增加比较明显。

可见,FuzzyCDF方法更加能够适应数据量较小的情况。由此可见,FuzzyCDF方法更加适合用于我们的课程。因此,选择在课程中采用FuzzyCDF进行认知诊断分析。

3.3 认知诊断实践

在数据结构与程序设计课程中应用FuzzyCDF方法进行认知诊断的具体流程如图5所示。从PTA测试数据出发,根据前期构建的课程知识图谱以及PTA测试题,形成Q矩阵,同时根据PTA测试数据生成

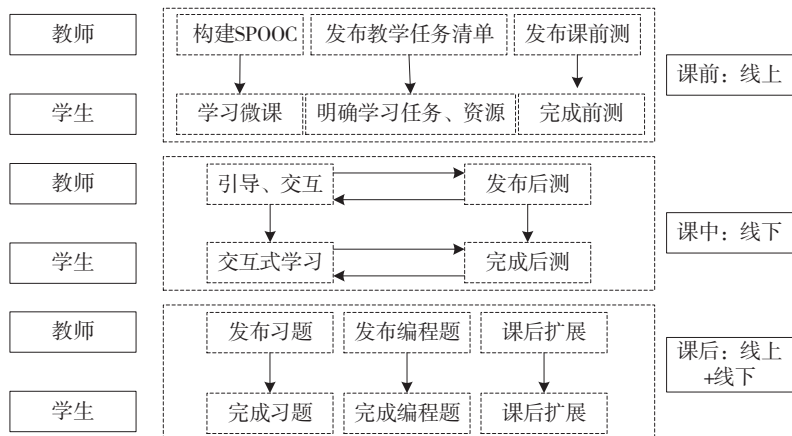


图2 数据结构与程序设计课程线上线下混合式教学过程

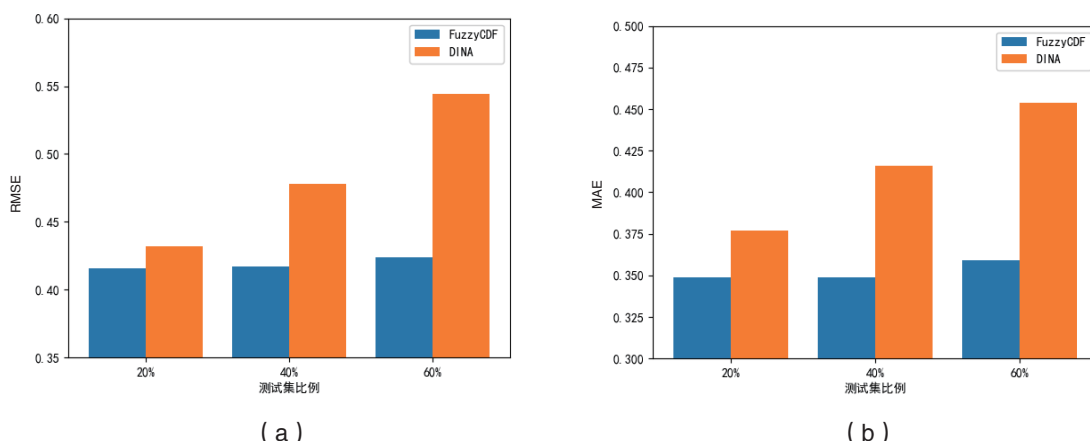


图4 DINA和FuzzyCDF预测情况评估

学生试题得分矩阵。利用 Q 矩阵以及学生试题得分矩阵应用 FuzzyCDF 模型, 估计模型参数, 最后利用 FuzzyCDF 模型评估学生的学习能力。

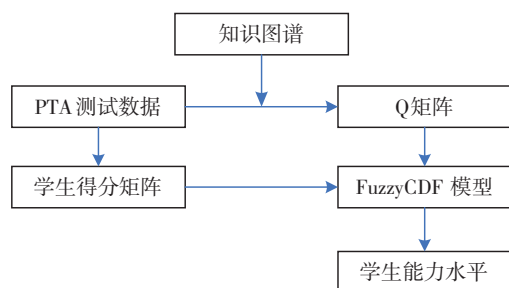


图5 基于FuzzyCDF认知诊断过程

根据知识图谱生成的数据结构与程序设计(上)的知识点见表2。在数据结构与程序设计课程实施过程中分为上、下两个部分, 分别在两个学期实施, 数据结构与程序设计(上)侧重程序设计。

利用 PTA 测试数据得到每个学生的知识点掌握水平, 如图6所示。在图6中, 最好的学生是指该学生各个知识点水平求和后值最大。学生的平均水平是指各个知识点上学生掌握水平的平均值。图6显示了两名学生各个知识点的掌握水平。可见, 采用 FuzzyCDF 可以直观地表达每个学生知识点的掌握水平。

利用 FuzzyCDF 可以获得知识点 k 对于不同学生 j 的难度 ajk 和区分度 bjk , 利用中位数可得到各个知识点对学生的难度和区分度。同一个知识点, 对于不同的学生具有不同的难度和区分

表2 数据结构与程序设计(上)课程知识点

知识点	子知识点
基本概念	基本概念, 标识符
基本数据类型	常量定义、整型、实型、字符型、转义字符、字符串、变量定义
表达式	运算符优先级和结合性、算术表达式、关系表达式、逻辑表达式、条件表达式、赋值表达式、逗号表达式、位运算表达式
输入输出	printf 输出、scanf 输入
分支	基本 if 语句、if 语句嵌套、switch
循环	While、for、do-while、多重循环、continue 和 break
数组	一维数组、字符数组、二维数组
构造类型	结构体、联合类型、枚举类型
文件	文本文件处理、二进制文件处理
函数	函数定义、函数调用、函数返回值、参数传递、变量作用域、变量存储属性、递归调用

度, 体现了不同学生个性化的学习特征。进一步分析, 可以得到学生在更加详细知识点上的掌握水平, 即对于每个学生可以获得表2中子知识点的掌握水平, 使得对学生知识点掌握水平的分析更加精确, 有利于更加精准地开展个性化学习。

4 结 语

因材施教是当前教育的基本理念之一。要实现因材施教就要充分认识到不同学习者个性化的差异, 注重个性化学习。混合式教学结合大量线上资源给个性化学习创造了有利条件, 然而从学

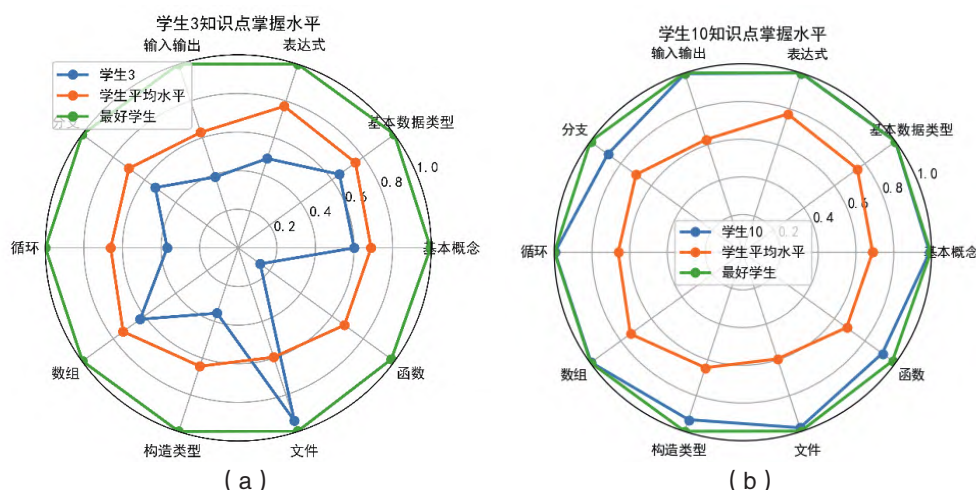


图6 学生的知识点掌握水平

生角度来看,对于一门新的课程完全是陌生的。这就需要教师首先对课程知识进行梳理,知识图谱就是一个很好的知识建模工具。课程知识图谱不但对学生的学学习起到引导作用,同时也对课程资源进行标注,但是个性化的学习同样需要了解学生的学习情况。利用认知诊断技术可以分析学生在各个知识点的掌握情况,获取学生的学习

情况。

在数据结构与程序设计课程中采用混合式教学模式,以知识点为实体构建课程知识图谱,并对在线教学平台PTA的习题进行标注。在这种教学过程中,根据学生在PTA平台的习题测试情况,利用FuzzyCDF对学生在各个知识点的掌握情况进行实时评估,为个性化学习提供了依据。

参考文献:

- [1] 国务院. 中国教育现代化2035[EB/OL]. [2023-10-09]. http://www.gov.cn/xinwen/2019-02/23/content_5367987.htm.
- [2] 国务院. 深化新时代教育评价改革总体方案[EB/OL]. [2023-04-15]. http://www.moe.gov.cn/jyb_xxgk/moe_1777/moe_1778/202010/t20201013_494381.html.
- [3] 何文滢, 钟志勇. 国内混合式学习研究评述与展望(2005—2021年): 基于VOSviewer的可视化分析[J]. 教育文化论坛, 2022, 14(5): 124-135.
- [4] 中华人民共和国教育部. 关于公布首批国家级一流本科课程认定结果的通知[EB/OL]. [2023-04-15]. http://www.moe.gov.cn/srcsite/A08/s7056/202011/t20201130_502502.html.
- [5] 刘领兵. C语言程序设计课程的MOOC/SPOC混合教学实践[J]. 计算机教育, 2022(4): 5-8.
- [6] 窦燕. 基于CSP和PTA的数据结构与算法课程设计策略[J]. 计算机教育, 2021(6): 5-9.
- [7] 李薇. 基于EduCoder平台的面向对象程序设计课程实践教学探索与实践[J]. 计算机教育, 2022(6): 179-183.
- [8] 王一岩, 郑永和. 智能时代个性化学习的现实困境、意蕴重构与模型构建[J]. 电化教育研究, 2023, 44(3): 28-35.
- [9] 吴彦文, 孙晨辉, 李斌. 知识图谱助力学科教学: 以大学模拟电子技术基础课程为例[J]. 软件导刊, 2020, 19(12): 195-198.
- [10] 李振, 周东岱, 王勇. “人工智能+”视域下的教育知识图谱: 内涵、技术框架与应用研究[J]. 远程教育杂志, 2019, 37(4): 42-53.
- [11] 李春英. 融合知识图谱的学习者个性化学习资源推荐[J]. 小型微型计算机系统, 2023(1): 1-10.
- [12] 张棚翔. 基于知识图谱和图嵌入的个性化学习资源推荐[J]. 计算机系统应用, 2023, 32(5): 180-187.
- [13] 朱天宇. 基于认知诊断的个性化试题推荐方法[J]. 计算机学报, 2017, 40(1): 176-191.
- [14] 吴润泽. 基于认知诊断的学生作答行为建模方法及应用研究[D]. 北京: 中国科学技术大学, 2019.

(编辑: 孙怡铭)