

认知诊断

本章将首先介绍认知诊断的基本概念、原理和模型,然后结合自适应框架,对自适应学习系统模型与实现结合实际案例进行详细阐述。最后,基于学习路径推荐算法,对个性化学习路径挖掘模型进行描述。我们还对个性化学习路径的典型案例分析,并探讨个性化学习路径推荐的未来与发展。

5.1 认知诊断简介

认知诊断(Cognitive Diagnosis)是指通过学生在测验上的试题作答记录对其知识状态(如知识掌握水平)进行评价,并在此基础上提出改善学生认知状态的方案与建议。

5.1.1 认知诊断概述

1. 教育测量理论发展

学生解决问题的行为表现都与认知前因有关,甚至有时会是

知识不足或策略选择的成熟心理的直接结果,因此,基于测试推理学生心理过程的研究有利于教师改变对学生的误解,并替换错误的策略。根据 Messick 的观点,从学生回答或解决试题的心理过程的角度来理解考试成绩是结构效度理论的一个核心特征。从这些测试中寻找的信息本质上是关于学生在思维和学习中的认知状态。也就是说,结构效度较高的测试可以反映与学习相关的某种形式的思维和高阶认知过程,从而确定教育者如何更好地改进与提供教学服务,以及如何最大限度地增加学生的学习机会。

教育测验理论可以预测测验者的能力,它认为每个人对于任何给定的属性都具有某种能力,这种属性可以被测量。至今,测验理论的产生及发展对解决许多教育及其他学科中的实际问题起到重要作用,主要有经典测量理论(Classical Test Theory,CTT)、概化理论(Generalizability Theory, GT)、项目反应理论(Item Response Theory, IRT)和认知诊断(Cognitive Diagnosis,CD),如图 5-1 所示。

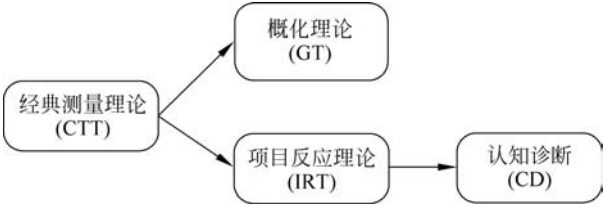


图 5-1 测验理论的产生及发展

17 世纪,基于测量误差服从正态分布思想,经典测试理论开始萌芽。1904 年查尔斯·斯皮尔曼解决了如何校正由于测量误差引起的衰减的相关系数,以及如何获得进行校正所需的可靠性指标。一部分人认为他的发现是经典测试理论的开端。经典测试理论是

一种具有数十种项目分析指标及评价标准体系的测试理论,它于20世纪中期关于测量误差服从正态分布思想而被提出并进行实践研究改进而日趋丰富、完善,最终形成。经典测试理论是目前应用最为广泛的一种测量理论,它建立了一系列项目分析的指标,包括平均分(Mean Score)、难度(Difficulty)、区分度(Discrimination)、信度(Reliability)和效度(Validity)等。

1972年,基于CTT标准化技术和项目分析技术等方法,Cronbach等提出了概化理论。它应用实验设计及其分析、方差分析模型等统计工具在一定范围内对误差进行控制,提高了测验的信度。它与CTT同属于随机抽样理论。概化理论可以更加有效且有预控性地去改善和提高测量精度与准确性。

在心理计量学中,项目反应理论(IRT)也称为题目反应理论或潜在特质理论,它用来分析考试成绩或者问卷调查数据的一系列心理统计学模型。项目反应理论基于3个假设,分别是:潜在特质空间的单维性假设、局部独立性假设和项目特征曲线假设。项目反应理论建立了项目性能,被试潜在特质水平与项目应答正确概率关系。项目反应理论认为被测的潜在特质即被测潜在的能力的掌握水平是连续的,被测在测验项目上的反应与被测的潜在特质存在着特殊的关系。通过项目反应理论建立的项目参数具有恒久性的特点,因此,不同测量量表的分数可以统一。不仅对被测的分析,而且对测验项目的分析,项目反应理论均克服了经典测试理论的部分局限,比经典测试理论更加细致与具体。项目反应理论是对经典测量理论的一种改进,因为它是建立在非线性的概率模型之上的。

基于项目反应理论(IRT)或经典测验理论(CTT),传统的教育

评估通常以学生在单个熟练度连续体中的位置来确定其分数,因此,分数也体现了学生个体在特定领域上以某种标准进行比较或排名,从而用作汇总评估程序的一部分。这样的分数在传统的教育评估中被广泛应用,例如:区别及格与不及格的学生,选择课程候选人,检验学生入读大学资格,确定奖学金获得者等。通过确定学生单个熟练度在连续体中的位置,把教育评估与教育决策者目标相关联。

1993年,Mislevy认为,测验理论大致可分为标准测验理论(Standard Test Theory)阶段和新一代测验理论(Test Theory for a New Generation of Tests)阶段。经典测试理论、概化理论和项目反应理论都属于标准测验理论。标准测验理论在本质上是将被测的心理特质视为一个心理学意义并不明晰的“统计结构”,其目的在于从宏观上给被测一个整体的评估,因此,标准测验理论在强调被测宏观层次能力水平测量及评估的同时,忽略了被测微观的内部心理加工过程的测量及评估,从而缺乏对被试内部心理加工机制的研究,这种研究视野称为“能力水平研究范式”。

随着心理测量学和认知心理学的进一步发展和现代教育技术水平的进步,教育者与学习者都希望得到更具体的、细微水平的测量与诊断评估。1993年,Frederiksen、Mislevy和Bejar在所著的*Test Theory for a New Generation of Tests*一书中正式提出了新一代测验理论的概念,标志着新一代测验理论的诞生。新一代测验理论强调测验应同时在宏观和微观、能力水平和认知水平的评估并举两种水平的研究视野下进行。因此,具有心理学理论支撑的新一代测量理论的结果,个体宏观能力水平和微观内部加工过程评估同时兼顾。新一代测验理论除了对能力的定位以外,深入

研究并拓展认知水平,即认知能力结构和状态的评估。

2. 认知诊断评价

认知诊断是以基于模型的测量和形成性评估相结合为特征的心理和教育测量,认知诊断的测试结果能对特定领域作可解释性的预测。标准测验理论对表现模式符合心理测量模型的测试结果能进行合理的解释分析,但它缺乏对个体差异定性方面的测量。被试者在处理能力的模式、解决问题的策略、相关背景知识、动机和身体方式(如障碍和残疾)方面可能有所不同。新一代测验理论更注重测量这些定性差异,认知诊断测量评估可用于指导分数的可解释性分析或为测试分数的外部相关因素定义调节变量。认知诊断是完全基于模型的,提供了关于被测者的优缺点反馈,因此,被测收到的不是总结性的分数,而是一份个人简介,详细说明他们掌握了哪些概念和技能,哪些没有掌握。

认知诊断评估(CDA)旨在测量学生的特定知识结构和处理技能,以便提供关于他们认知优势和劣势的信息。Xiangdong Yang 和 Susan E. Embretson 认为在心理或教育背景下的认知诊断测试主要集中于至少测试认知特征的以下 3 方面。

(1) 给定特定认知领域中必不可少的技能档案或知识列表,这些技能和知识集代表了该领域最重要的技能和概念,并且是发展其他高阶能力的基本组成部分。

(2) 结构化程序和/或知识网络。一个领域的知识结构不仅由该领域所拥有的基本技能或知识片段的数量来表示,还由人们对这些技能和知识进行组织的结构化来表示。

(3) 认知过程,成分或能力。认知研究的信息处理范式提供了

发掘认知内部过程的方法,因此可以为特定类型的认知任务开发特定的认知模型。教育者可以通过观察被测者执行此类任务时的潜在认知过程来解释所观察到的表现。

认知心理学家与认知心理测量学家把认知心理学对人类认知加工过程的内在机制的研究成果与研究范式相结合,创造性地开发了具有认知诊断功能的计理评估模型——认知诊断模型(Cognitive Diagnosis Models, CDMs)。CDMs 是一个种测量模型,它包含某个领域相关的离散的潜在属性变量,这些属性变量是指对二进制技能或知识的掌握,它代表测试项目所涵盖领域的各个基本特征。当属性向量的二进制分量指的是条件的存在或不存在时,该方法可能同样适用,而条件可能与医学、心理或精神病学构造有关。CDMs 可以通过观察被测者某个属性变量的回答估算学生的能力概况。当测试的目的是估计学生的个人资料或属性掌握模式时,可以使用认知诊断模型(CDM),而不是提供一般的能力估计。与 CTT 和 IRT 不同,CDM 是潜在类模型的特例。具体而言,CDM 根据属性精通模式对正确回答项目的概率进行建模。

5.1.2 认知诊断评价理论的基础概念

1. 属性与属性层级

在教育测量中,认知诊断是对个体认知加工过程中所涉及的知识结构体系的特定能力单元的认知诊断评价。被测者完成某个特定领域任务的知识结构或加工技能在认知诊断评估中称为属性。Tatsuoka 认为属性指生产规则、程序操作、项目类型等认知子任务。在认知诊断模型的建模中,用“认知属性”来描述测量任务

中被测的认知过程或认知技能并作为诊断评估单元,它可以是认知过程中所涉及的知识或技能,是对被测心理内部加工过程的诊断基础。如果确定了用于构建认知诊断测试的内容域,则逻辑任务分析可以指定认知属性。

Leighton(2004)认为认识属性在认知过程中,不是独立存在的个体,相反,认知属性从属于某个相关联的网络,并且认知属性之间在心理、逻辑或层级上存在着一定的层级关系。Leighton 将认知理论和心理测验实践联系起来,以促进教育和心理测验的发展和分析,提出了认知项目反应理论模型,称为属性层次方法(Attribute Hierarchy Method, AHM)。在属性层次方法中,首先使用认知心理学中的方法(例如项目评论和协议分析)来识别和研究认知属性,认知属性简称为属性,每个对象的属性配置文件由代表潜在有限属性集的二进制潜在变量向量表示,再用属性层级关系(Attribute Hierarchy)图对认知模型进行表征建模,属性层次关系结构表征某个属性的掌握是对另一属性的掌握的先决条件。属性层次方法假设认知诊断测试性能取决于属性的分级能力,被测只有具备这些属性才能正确回答测试项目。

Leighton(2004)等把属性层级关系划分成线性型、收敛型、分支型和无结构型 4 种基本类型,如图 5-2 所示。

(1) 在线性型层级关系模型中,属性 A1 是属性 A2、A3 和 A4 的先决条件,属性 A2 是属性 A3 和 A4 的先决条件,属性 A3 是属性 A4 的先决条件。从该线性型层次可得被测者只有掌握属性 A1 所代表的认知或技能,才有可能正确完成属性 A2、A3 和 A4 所测内容。

(2) 在收敛型层级关系模型中,属性 A3 和 A4 同时是属性 A5

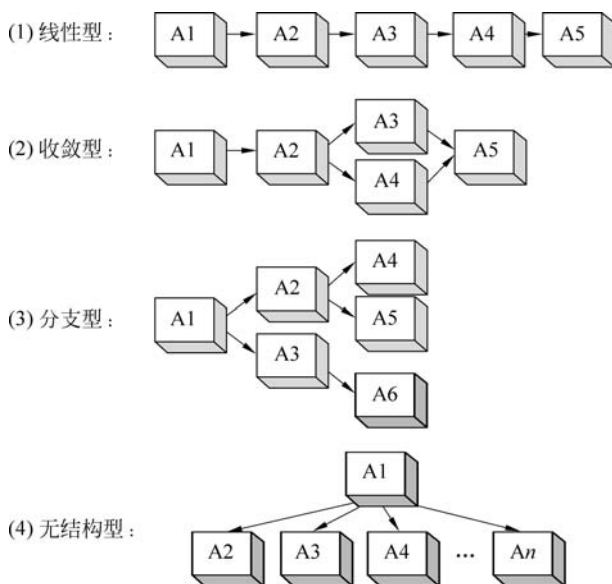


图 5-2 4 种属性层级关系基本类型

的先决条件,被测者不仅必须掌握属性 A3,还必须掌握 A4 才有可能正确完成属性 A5 所测内容。

(3) 分支型层级关系模型描述解决某一领域问题心理内部认知属性加工层级与顺序过程。

(4) 在无结构型层级关系模型中,属性 A1 是属性 A2, A3, ..., An 的先决条件,属性 A2, A3, ..., An 之前是并列关系且无先后顺序之分。

基于以上 4 种基本属性层级关系,可以复合构成表达更复杂并且相互关联的属性层级关系,从而可以描述各种相关联的认知过程或技能。AHM 的发展是为了解决与认知模型发展和统计模式识别相关的两个具体问题。

2. Q 矩阵理论

认知诊断可以看成模式识别和统计决策理论的问题。认知诊断的第一部分的工作是从不可观察的知识状态或认知过程中提取特征变量。正因为认知过程是不可观察的,并且作为分类类别的知识状态不能直接从观察中获得,使模式识别变得额外困难。Embreston^[1]提出用于构建认知诊断测试的内容域,可以用 n 项目 K 个指定属性来表征,并用 K 个元素向量表示认知属性。Tatsuoka 和她的同事提出通过可观察的项目反应模式—— Q 矩阵理论 (Q -Matrix Theory) 把认知过程中提取特征变量——认知属性,用关联矩阵(Q 矩阵)对正确解决问题所需的基本认知属性用数学表示。通过这样做,可以利用布尔代数和图论的基本定理来解释认知属性之间的关系,以及它们与知识结构相关联的属性(概率的和逻辑的)。Tatsuoka 将回答测试项目所需的基本认知任务组织成一个 Q 矩阵,其中的行代表属性,属性可以是子任务、认知过程和技能等。列代表项目,每列中的条目表示每个项目的解决方案中涉及哪些属性。这里需要注意的是如果两位专家使用不同的方法来解决给定的一组问题,则他们可能会获得完全不同的 Q 矩阵。下面以基于图 5-3 所示的层级关系为例,对 Q 矩阵理论相关概念进行介绍。

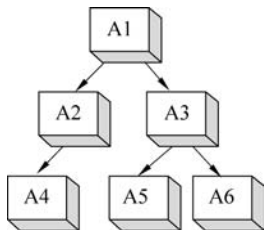


图 5-3 6 个属性的层级关系模型

1) 邻接矩阵(Adjacency Matrix)

属性层级关系除了可用图表示之外,也可用邻接矩阵表示,通常记为 **A** 矩阵。邻接矩阵由 K 行 K 列(K 指属性的个数)的 0 和 1 矩阵组成,如果认知属性间存在直接关系,则在邻接矩阵相应的元素中用“1”表示,否则用“0”表示。图 5-3 所示的属性层级关系的邻接矩阵表示如图 5-4(a)所示,第一行表示属性 1 是属性 2 和属性 3 的直接前提,第二行表示属性 2 是属性 4 的直接前提,第三行表示属性 3 是属性 5 和属性 6 的直接前提。

$$\begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

(a) 邻接矩阵

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

(b) 可达矩阵

图 5-4 用邻接矩阵与可达矩阵表示属性关系

2) 可达矩阵(Reachability Matrix)

可达矩阵通常记为 **R** 矩阵,它可以同时表示属性之间直接和间接的层级关系,图 5-3 所示的属性层级关系蕴含的属性之间的层级关系如图 5-4(b)可达矩阵所示,可达矩阵第一行表明属性 1 是所有属性的前提,可细分为直接前提或间接前提。第二行表明属性 2 是属性 2 和属性 4 的直接或间接前提,第三行表明属性 3 是属性 3、属性 5 和属性 6 的前提,第四行表明属性 4 是自身的前提,其他同理。

3) 关联矩阵(Incidence Matrix)

关联矩阵即 **Q** 矩阵,描述了测试项目与认知属性之间的关系,

它一般由 M 行 K 列的 0 和 1 矩阵组成,其中 M 指测验项目个数, K 指所有测试项目所测试属性的个数。假设给定测试域中的认知过程或技能被建模为 K 个潜在属性,项目 m 建模为向量 $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_k$ 的组合。 K 维二进制向量 $\mathbf{q}_m = (q_{m1}, q_{m2}, \dots, q_{mk})$ 表示项目 m 是否测试了属性 $k, m=1, 2, \dots, M$, 其中 $q_{mk}=1$, 代表项目 m 是否测试属性 k , 若 $q_{mk}=0$, 则代表项目 m 未测试属性 k 。用包含 5 个项目 6 个属性的测试为例,其 \mathbf{Q} 矩阵如图 5-5 所示。图 5-5 表示项目 1 测试了属性 1 和属性 2, 其他属性没有测试; 项目 2 测试了属性 2 和属性 3, 其他属性没有测试, 其他项目解释亦然。需要注意的是, 某个测试项目可以测试所有属性, 但不存在不测试任何属性的项目。

1	1	0	0	0	0
0	1	1	0	0	0
0	0	0	1	0	0
0	1	0	0	1	1
0	0	1	0	1	0
0	1	0	1	0	0

图 5-5 \mathbf{Q} 矩阵示例

4) 简化 \mathbf{Q} 矩阵(Reduced \mathbf{Q} -Matrix)

简化 \mathbf{Q} 矩阵(Reduced \mathbf{Q} -Matrix)也称为缩减矩阵, 记为 \mathbf{Q}_r 。 \mathbf{Q} 矩阵的第一行表示某个测试项目所测试关联的属性, 假设某个认知诊断评估所涉及的属性有 k 个, 每个属性可取值 1 或 0, 理论上这 k 个属性可以组合成 $2^k - 1$ (因为不可能存在不测试任何一个属性的项目) 种不同的测试项目。如图 5-3 所示, 6 个测试属性则可能有 $2^6 - 1 = 63$ 种不同的测试项目, 即 \mathbf{Q} 矩阵可能是 63 行 6 列, \mathbf{Q} 矩阵的行数将随着 k 值的增大, 呈指数级增长, 然而, 认知诊断评估中属性之间通常存在某种属性层级结构, 例如, 属性 1 代表

整数加法,属性 2 代表整数的乘法,这意味着属性 1 必须是属性 2 的前提,因此不可能存在某个测试项目只测试了属性 2,但不测试属性 1。基于某种属性层级结构,构造测试项目可能的属性组合往往比 $(2^k - 1)$ 小。如图 5-3 所示,测试项目如果测试属性 3,则必然测试属性 1 和属性 2,不存在项目只测试了属性 6,但不测试属性 1 或属性 4,以此类推。

Q_r 矩阵是通过减少矩阵,即通过减少 RSM 和 AHM 中的 Q -matrix 生成的。假设 Q_r 矩阵被认为是代表所有属性组合的那些项。如果属性数为 k ,则 Q 矩阵中的列数为 $(2^k - 1)$ 。根据 R 矩阵表示的属性层级的约束,把 Q 矩阵中不符合属性层级结构的项目删除,将 Q 矩阵简化为 Q_r 矩阵,这种生成 Q_r 矩阵的方法称为 Tatsuoka 法。2010 年,杨淑群在逐步向前回归的思想,提出关于 Q_r 矩阵的渐增式扩张算法。图 5-6 是基于图 5-4 与图 5-3 的简化 Q_r 矩阵示例。简化 Q 矩阵可以用于指导测试项目设计或者用于检测已编制测试项目的合理性。

1	0	0	0	0	0
1	1	0	0	0	0
1	0	1	0	0	0
1	1	0	1	0	0
1	0	1	0	1	0
1	0	1	0	0	1
1	1	1	0	0	0
1	1	1	0	1	0
1	1	1	0	0	1
1	1	1	1	0	0
1	1	1	1	1	0
1	1	1	1	0	1
1	0	1	0	1	1
1	1	1	0	1	1
1	1	1	1	1	1

图 5-6 简化 Q_r 矩阵示例

5) 理想属性掌握模式(Ideal Attribute Mastery Pattern)

理想属性掌握模式指的是在某个认知诊断评估测试中,被试者在理论上可能掌握的测试属性组合。假设认知诊断评估测试包含了 k 个属性,理论上,考生可能掌握的属性组合为 2^k 种,这里包含了被测没有掌握的任何一个属性的情况。由于测试项目的属性之间受属性层级关系影响,因此可能掌握的属性组合表的列中的某些组合是不可能存在的。理想属性掌握模式的计算方式是在简化矩阵 Q_r 的计算结果中添加一行全0组合,它指的是没能掌握任何一个属性的情况,对于如图5-3所示的属性层级关系,理想属性掌握模式如图5-7所示。认知诊断测试中观察到的被试表现是被测内在的属性掌握模式的反应,是对被测者心理学的洞察。把被试者观察到的属性掌握模式适当地分配到测试领域的知识状态时,就可以做出关于认知技能的推论。被测的属性掌握组合也称为被测的知识状态,它提供了认知和评估之间的联系。

1	0	0	0	0	0
1	1	0	0	0	0
1	0	1	0	0	0
1	1	0	1	0	0
1	0	1	0	1	0
1	0	1	0	0	1
1	1	1	0	0	0
1	1	1	0	1	0
1	1	1	0	0	1
1	1	1	1	0	0
1	1	1	1	1	0
1	1	1	1	0	1
1	0	1	0	1	1
1	1	1	0	1	1
1	1	1	1	1	1
0	0	0	0	0	0

图 5-7 理想属性掌握模式示例

6) 理想反应模式(Ideal Response Pattern)

理想反应模式指的是被测在认知诊断测试中不存在猜测或失误的情况下,在测试项目上作答时的反应模式。也就是说,当被测掌握了某个测试项目所测试的所有属性时,被测就一定能够答对测试项目,反之,如果被测没能掌握测试项目中的任何一个或一个以上属性,被测就不可能答对该测试项目。理想反应模式矩阵的行对应于满足属性层级结构约束的所有可能的属性掌握模式的集合,各列表示某个可能的属性掌握组合的被测对所有 Q_r 中测试项目的理想反应。丁树良(2017)提出:在测试项目所涉及的属性之间不可以补偿的前提下,用非补偿的定义和朴素集合论计算理想反应模式。基于如图 5-3 所示的属性层级关系,以图 5-6 为测试项目,以图 5-7 为假设的被测属性掌握模式,理想反应模式如图 5-8 所示。

1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0
1	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0
1	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0
1	1	1	0	1	0	1	1	0	0	0	0	0	0	0	0
1	1	1	0	0	1	1	0	1	0	0	0	0	0	0	0
1	1	1	1	0	0	1	0	0	1	0	0	0	0	0	0
1	1	1	1	1	0	1	1	0	1	1	0	0	0	0	0
1	1	1	1	0	1	1	0	1	1	0	1	0	0	0	0
1	0	1	0	1	1	0	0	0	0	0	0	1	0	0	0
1	1	1	0	1	1	1	1	1	0	0	0	1	1	0	0
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

图 5-8 理想反应模式示例

5.1.3 认知诊断代表性模型

认知诊断模型是潜在变量模型,用于推断潜在的技能、知识或个性,这些是对教育、心理和社会科学测试和测量的响应基础。认知诊断模型评估从被测的测试结果中得到所测领域相关属性的诊断反馈。具体而言,CDM 分析并确定被测是否对一组细粒度属性(潜在类别)精通。与总分数相比,教师、临床医生和其他测试用户可以使用每个被测的特定信息来更有效地调整和改进他们的干预措施。国内外的学者根据不同的研究需求开发出了许多认知诊断模型。2017 年,高旭亮等把认知诊断模型划分为以下几种类型,如图 5-9 所示。本节将对一些具有代表性的典型认知诊断模型进行介绍。

1. 规则空间方法

规则空间方法(Rule Space Method)是早期的先驱 CDM 模型。1971 年,Tatsuoka 提出并研究了规则空间方法。她首次使用了“ Q 矩阵”这个术语,引入了 Q 矩阵理论和规则空间模型来诊断被测者的知识水平。如 5.1.2 节所述, Q 矩阵用于将项目与测试开发过程中要测量的属性进行匹配,它表达了测试项目的分布和属性。

规则空间是一个二维笛卡儿坐标系,规则空间根据被测者的状态信息与观察到的响应模型,计算被测的潜在特质水平参数(θ)和偏离警戒指标变量(ζ)并映射构成二维空间。



图 5-9 认知诊断模型分类

第一维是被测潜在特质水平参数(θ),它对应于技能或能力变量,与被试能力水平的差异有关。被测潜在特质水平参数(θ)是由项目反映理论模型定义的变量,可以通过人才评估方法进行评估。项目反应理论常用的数学模型有单参数逻辑模型、双参数逻辑模型和三参数逻辑模型,测试者可以选择其中任意一个合适的模型来估计被测潜在特质水平参数(θ)的值,例如由双参数逻辑模型定义的潜在特质水平参数为 θ_j ,被测正确作答项目概率如式(5-1)所示。

$$P_{ij}(x_i = 1 | \theta_j) = \frac{1}{1 + e^{-D_{a_i}(\theta_j - b_i)}} \quad (5-1)$$

第二维 ζ 对应于偏离警戒指标变量,该变量用于测量项目响应模式的异常性。这个维度表示潜在特质水平为 θ 的被测,实际测试项目作答模式偏离理想项目反应模式的程度。 ζ 指数作为个人调节指数之一,已经被开发用于检测由异常行为(如欺骗或猜测)引起的异常反应模式。

假设被测者 i 实际的项目反应模式向量为 $\mathbf{X} = (x_1, x_2, \dots, x_n)$,其中 n 为测试项目的个数,则被测 i 对应的测试项目作答正确概率为 $P(\theta_i) = [P_1(\theta_i), P_2(\theta_i), \dots, P_n(\theta_i)]$,被测的作答反应函数向量如式(5-2)所示。

$$T(\theta_i) = \frac{\sum_{j=1}^n P_j(\theta_i)}{n} \quad (5-2)$$

由定义:

$$f(\mathbf{X}) = -[\mathbf{X}, P(\theta_i) - T(\theta_i)] + [P(\theta_i), P(\theta_i) - T(\theta_i)] \quad (5-3)$$

得

$$\zeta = \zeta(\theta_i, \mathbf{X}) = \frac{f(\mathbf{X})}{\sqrt{\text{Var}[f(\mathbf{X})]}} \quad (5-4)$$

以上是建立规则空间的方法,接着将观察的被测作答反应模式进行归类,并判断其对应的属性掌握模式。

假设观察被测的反应模式服从多变量正态分布,理想项目反应模式记为 R ,对应的能力水平 IRT 估计为 θ_R ,警戒指标变量为 ζ_R 并以理想反应模式 R_l 为中心,则对于双参数逻辑模型,测验信息函数如式(5-5)所示。

$$I(\theta) = \sum_{j=1}^n a_j^2 P_j(\theta_{R_l}) Q_j(\theta_{R_l}) [P_j(\theta_{R_l}) - T(\theta_{R_l})]^2 \quad (5-5)$$

相对地观察被测的项目反应模式记为 X ,对应的能力水平 IRT 估计为 θ_X ,警戒指标变量为 ζ_X ,得马氏距离平方如式(5-6)所示。

$$D_{xl}^2 = \frac{(\theta_X - \theta_{R_l})^2}{I(\theta_{R_l})} + \frac{(\zeta_X - \zeta_{R_l})^2}{\text{Var}(\zeta_{R_l})} \quad (5-6)$$

根据被测观察反应模式与理想反应模式之间的马氏距离,取最小的理想反应模式对应的属性掌握模式作为被测的最佳属性掌握模式。马氏距离的平方遵循两个自由度的卡方分布。由于距离度量本身不提供错误分类概率,因此最小误差的贝叶斯决策规则也被应用。当将 RSM 应用于定义明确的域时,RSM 已被证明特别有效,然而在从 Q 预测的反应模式中寻找反应模式行为的资源方面存在局限性。

2. 属性层级方法

属性层级方法(Attribute Hierarchy Method)是 Leighton 等于 2004 年提出的一种认知项目反应理论模型,这种方法是 Tatsuoka 的规则空间方法的一种变体。属性层级方法的设计目的在于将认知理论和心理测验实践联系起来,以促进教育和心理测验的发展和分析。

认知研究表明,认知技能不是孤立运行的,而是属于一个相互关联的过程网络(例如,库恩,2001;沃斯尼亚杜布鲁尔,1992),因此,AHM引入属性依赖的假设。基于这样的假设,即测试性能取决于一组称为属性并且按相互之间层级关系排序的能力。换句话说,属性层次结构的识别对于AHM是最重要的输入变量,因为它用于预测学生表现的类别并推断考生的认知能力、Leighton、Gierl和Hunka(1999)对属性进行了描述:属性是对在特定领域中执行任务所需的过程或声明性知识的描述。属性不是策略,但是属性为策略提供了构建块,特定层级关系的属性集可以成为解决某个具体问题的策略。属性是动态实体,具体地说,某个属性组合在时间点A是解决某个具体问题的属性组合成分,随着被测能力的提高与发展,这个属性组合在时间点B可能不再充当解决问题策略的有用描述。可以使用不同的方法(例如专家意见、任务分析、学生的书面回答等)来确定测试的属性。

在AHM的应用中,在测试之前,必须先确定测试域中正确回答测试项目所需的属性及属性层级关系。测试项目的开发必须以属性的层级结构为指导并需验证其合理性。通过使用属性层级结构来开发测试项目,项目开发人员可以最大程度地控制每个项目所测量的特定属性。当从层级结构开发测试项目时,可以为构造的项目标识唯一的邻接矩阵。相反,如果在开发测试项目时,没有以特定的属性层级结构为指导,则要想从项目中抽象层次结构(简化的 Q 矩阵),甚至为这些项目确定唯一的邻接矩阵是非常困难的。AHM在生成理想响应模式的过程中利用了Tatsuoka矩阵(邻接矩阵、可达性矩阵、关联矩阵和简化的 Q 矩阵)。最后从

Q-Matrix 生成被测者反应模式的预测及对被测观察到的反应模式进行分类。AHM 中观察到的响应模式的分类方法有以下两种。

1) 初步分类法(Preliminary Classification Method)

在这种方法中,将观察到的被测反应模式与理想响应模式进行比较,在其中识别出形式为 $0 \rightarrow 1$ 和 $1 \rightarrow 0$ 的误差。对应理想反应模式,计算得出被测的属性掌握模式。具体地说,令 V_j 为 n 个项目第 j 个理想响应模式,而 X 为观察到的被测属性掌握模式。那么, $d_j = V_j - X$ 产生由 $\{-1, 0, +1\}$ 元素组成的向量,它代表着相对应项目属性上观察到的属性掌握模式与理想反应模式之间的误差向量。其中 $d_j = 0$ 表示无错误; $d_j = -1$ 表示被测理想反应模式本应该答错的测试项目,却被观察到是答对的,即 $0 \rightarrow 1$ 误差,其概率等于 $P_{jk}(\theta)$, k 指 k 个 $0 \rightarrow 1$ 误差,相反 $d_j = 1$ 表示理想反应模式中为 1 的项目,被测却答错的,即 $1 \rightarrow 0$ 误差,其概率等于 $1 - P_{jm}(\theta)$, m 指 m 个 $1 \rightarrow 0$ 误差。那么 m 个 $1 \rightarrow 0$ 误差和 k 个 $0 \rightarrow 1$ 误差的联合概率如式(5-7)所示。

$$P_{j\text{Expected}}(\theta) = \prod_{k=1}^K P_{jk}(\theta) \prod_{m=1}^M [1 - P_{jm}(\theta)] \quad (5-7)$$

式(5-7)表明,给定潜在特质水平为 θ 的被测 $0 \rightarrow 1$ 的误差概率与 $1 \rightarrow 0$ 误差概率之积,计算得出潜在特质水平为 θ 的被测观察到属性掌握模式接近理想反应模式的预测接近程度,当相应的 $P_{j\text{Expected}}(\theta)$ 越大被测划分为具有第 j 组属性的可能性就越大。

2) 初步分类的验证法(Verification of Preliminary Classification)

初步分类验证法通过识别逻辑上属于观察到的被测反应模式子集所有理想反应模式来获得。例如,如果被测观察到的响应模式为(11011),则逻辑上理想反应模式(10011)、(01010)、(11001)等都属于该观察到的响应模式,其子集是可以被识别的,于是被测

就被标签为已掌握该属性模式。如果被测的理想反应模式在逻辑上不属于观察到的反应模式,例如理想反应模式(11101)是不属于观察模式(11011)的子集,就必须计算每个这类理想反应模式 $1 \rightarrow 0$ 的误差概率之积,并以此作为对被测进行预测分类的依据,其中 $P_{j\text{Expected}}(\theta)$ 的计算方法与初步分类方法一样。初步分类的验证法只计算 $1 \rightarrow 0$ 误差概率,因此比初步分类对观察到的反应模式的分类更保守。

研究表明,AHM 存在着局限性。不管是初步分类法还是初步分类验证法都是对大量应试者观察到的反应模式进行分类,分类是基于被测所观察到的反应模式的,因此,一旦对观察到的异常响应模式进行分类时,可能会产生非常低的似然估计。

3. DINA 模型

DINA 模型(Deterministic Inputs, Noisy And gate model)是一个易于处理和解释的认知诊断模型。DINA 模型设 X_{ij} 为被测 i 对测试项目 j 的反应,并设 $\alpha_i = \{\alpha_{ik}\}$ 为被测的二元答题技能向量,其中如果第 k 个元素是 1,则表示答对,是 0,则表示答错。元素 k 所代表的内容可以是技能、知识或认知过程。

跟大多数 CDM 一样,DINA 模型的实现需要构造一个 Q 矩阵,它是一个由 0 和 1 组成的 $J \times K$ 矩阵。 Q 矩阵可被看作一个认知设计矩阵,它明确地标识了每个测试项目的认知规范。

假设某个认知诊断测试的属性有减去基本分数、减少和简化、将整数与分数分离、从整数中借用并将整数转换为分数,那么某个测试项目 $6\left(\frac{3}{7} - 1\frac{4}{7}\right)$ 的 Q 矩阵向量则为(10110)。

在 DINA 模型中,被测 i 在项目 j 上的理想作答表示为 η_{ij} ,当被测掌握项目 i 考查所有属性,则 $\eta_{ij} = 1$,否则为 0,用式(5-8)

表示。

$$\eta_{ij} = \prod_{k=1}^K \alpha_{ik}^{q_{jk}} \quad (5-8)$$

设被测在作答过程中的失误概率为 $s_j = P(X_{ij}=0 | \eta_{ij}=1)$ 和猜测概率为 $g_j = P(X_{ij}=1 | \eta_{ij}=0)$ 。则被测 i 回答项目 j 的技能向量 α_i 正确的概率如式(5-9)所示。

$$P_j(\alpha_i) = P(X_{ij}=1 | \alpha_i) = g_j^{1-\eta_{ij}} (1-s_j)^{\eta_{ij}} \quad (5-9)$$

从式(5-9)可以看出答对一个项目,被测不止需要具备该测试项目必需的所有技能且不能存在误差,还要对至少缺少一项答对测试项目所需技能的项目能正确猜对。这里注意,如果不存在猜测和回答误差,则对一个项目正确反应的模型概率是 0 或 1。也就是说,被测响应模式完全由 α_i 和该项目的 Q 矩阵的共同作用决定,然而,正如 de la Torre 所指出的,试测答对项目并不局限于通过随机反应得到的正确反应,包括使用 Q 矩阵中没有阐述的替代策略。例如,如果一个项目可以用多种不同的技能来解决,则这些技能可能不是 Q 矩阵中规定的项目的回答规范,被测有可能是猜测,但实际上是在用不同的策略系统地解决问题。

DINA 模型是一个简洁且可解释的模型,不用考虑属性的数量是多少,每个项目只需两个参数(g_j 和 s_j)就可以提供良好的模型拟合。相对于项目反映模型(Item Response Models, IRM),DINA 模型具有处理多维二元的潜在技能,而传统的 IRM 具有处理一维连续的潜在特征。因为 DINA 模型中技能模式的数量是有限的,所以每个模式都可以被视为一个潜在的组或类,因此,DINA 模型和传统的 CDMs 归类于多个分类模型或受限潜在类模型。

4. 认知诊断模型的应用发展

近十多年来深度学习在许多领域带来了突破性的发展,例如用于图像识别、语音和自然语言处理、多模态认知能力识别和教育数据挖掘等,认知诊断模型在智慧教育中的应用成为研究的热点之一。许多学者对认知诊断模型进行了深入研究与扩展,例如自动编码器算法被用于开发深度学习认知诊断模型(DLCD),它不仅可以在完整的 Q 矩阵和不完整的 Q 矩阵上很好地工作,而且还显示出最有利的泛化能力;基于循环神经网络的知识追踪,可以通过贝叶斯网络、深度学习等模型进行拟合,从而得到学生在未来时刻的作答预测;有学者考虑了知识的应用上下文关系,提出基于数据驱动框架追踪学生更深层次的认知状态等。

目前,认知诊断模型的研究还处于发展阶段,需要进行更进一步的研究,例如:在实践中 CDM 参数估计软件缺乏可识别性的评估功能;认知诊断与机器学习相融合,用于计算机自适应学习,缺乏可解释性,不便于理解,不利于教育者提供有依据的干预策略等。人工智能与教育大数据飞速发展,给传统的认知诊断评估带来了新的挑战,推动着认知诊断模型的更新与内容的扩展,并且不断学习新的理论模型、研究方法,从而使智慧教育领域的研究不断发展。

5.1.4 认知诊断建模拟合评价

为了使来自 CDM 的推论有效,至关重要的是要确定模型对数据的适应性。认知诊断模型是主要用于评估应试者对技能或属性的精通和不精通的心理计量学模型。在认知诊断评估中,CDM 通

常与 Q 矩阵一起使用,以提供有关应试者的诊断信息。这些信息可以为学生的学习提供帮助,并有助于设计更好的教学。通常,认知诊断建模仅用于指代心理计量过程部分,但是,从整体角度来看,应将 Q 矩阵作为认知诊断建模过程的一部分。当需要考虑推论的有效性时, Q 矩阵作为建模过程中不可或缺的一部分,其作用将变得更加关键。之所以如此,是因为 CDM 和 Q -matrix 都可能在诊断建模环境下导致模型数据失配。

最近,已经开发了不同类型的 CDM,它们可以潜在地应用于广泛的环境。其中一些 CDM 是高度受限的模型,如 DINA 模型和 DINO(Deterministic Inputs, Noisy or gate, DINO)模型。有些是加法性质的,如加法 CDM(Additive CDM, A-CDM)。此外还有线性逻辑模型(the Linear Logistic Model, LLM)、对数线性 CDM 和广义 DINA(Generalized DINA, G-DINA)。

为了使来自不同 CDM 的推论有效,重要的是要确定模型对数据的拟合度(绝对拟合评估)。随着各种 CDM 的可用性,为特定应用选择最合适的模型(相对拟合评估)也很重要。CDM 在发展的同时,已经开发或使用了各种拟合统计或方法。其中一些统计数据包括:基于项目对的观察和预测相关性和对数优势比之间的残差及单个项目的观察和预测正确比例之间的残差;项目区分指数;基于观察和预测项目对反应的 χ^2 和 G 统计,以及观察到的和预测的项目条件成功概率之间的平均绝对差异和相关的近似均方根误差。常规拟合统计量有 Akaike 信息准则、贝叶斯信息准则、偏差信息准则,以及在经验上采用贝叶斯因子进行相对适合度评估等。

Kunina-Habenicht 等研究发现,在对数线性 CDM 中,当约 30% 的条目被随机置换或属性数量被过度指定或指定不足时, AIC

和 BIC 有助于选择被正确指定的 Q 矩阵与 Q 矩阵误差。此外,他们还发现,当忽略所有交互作用时,AIC 可以帮助针对参数误差的模型选择正确的模型,但是,平均绝对差和近似均方根误差在测试级别上对绝对拟合评估的有用性受到限制。此外,这些统计数据或方法中的大多数主要用于相对拟合评估。由于通常无法获得这些程序的 z 得分或 p 值,因此难以进行固定显著性水平的绝对拟合评估。

当认知诊断建模中出现模型-数据不匹配时,不匹配的来源可能来自属性的性质、属性结构、 Q matrix 或 CDM。其中, Q 矩阵误差和 CDM 设定误差是建模过程的组成部分,与其他不匹配的来源(例如属性性质的错误指定)相比,这两种误差都很容易调查,并且都会严重影响参数估计质量和分类精度,甚至会相互作用而导致估计过程的恶化。

1. Q 矩阵误差

开发认知诊断评估的一个关键步骤是通过指定 Q 矩阵来整合实质性知识。设 q_{jk} 表示 $J \times K$ 的 Q 矩阵的 J 行和 K 列的元素,其中 J 和 K 分别表示项目数和属性数。如果要求掌握属性 k 才能正确回答 j 项,则条目 q_{jk} 被指定为 1,否则为 0。在大多数 CDM 应用中,通过实质性知识建立质量矩阵的过程往往是主观的,并在研究人员中引起了严重的有效性问题的。当出现 Q 矩阵误差时,可能意味着为某个项目指定的属性指定不足(一些 1 被错误地指定为 0)、指定过度(一些 0 被错误地指定为 1),或者同时指定不足和指定过度(一些 1 被错误地指定为 0,一些 0 被错误地指定为 1)。在认知诊断建模中, Q 矩阵起着约束待估计项目参数数量的重要

作用。这种限制可能与不同类型的 CDM 相互作用。例如,尽管 DINA 模型的每个项目总是有两个参数,但 A-CDM 的每个项目都有 $(K_j^* + 1)$ 个参数,其中 K_j^* 是项目 j 所需的属性数量,因此,不同类型的 Q 矩阵误差与不同类型的 CDM 交织在一起,是混淆不匹配的来源之一。

2. CDM 参数误差

CDM 参数误差是指建模过程中心理测量成分的参数设置不正确。在选择 CDM 参数化时,研究人员将它们对回答测试项目所涉及的假设认知过程概念化、形式化。给定特定的 Q 矩阵和属性结构,可以发现本质上可能会有许多不同的参数设置(例如 DINA、DINO、A-CDM),因此,不能忽略 CDM 参数误差的可能性及这样做的后果。理论上,饱和的 CDM 始终比任何精简的 CDM 更适合数据,因为它们的参数设置更为复杂,但是,饱和模型并不一定是首选。原因之一是饱和的 CDM 需要较大的样本量才能精确估算。另一个原因是,假定 CDM 的适应性基本上不逊色,它们简化后易于解释。通过使用可以补偿模型参数复杂性的拟合统计量(例如 AIC、BIC),可以在 CDM 参数误差的范围内解决涉及饱和 CDM 和还原 CDM 的模型选择。

相对拟合评估是指在一组候选模型中选择最佳拟合模型的过程。在相对拟合评估下,对于匹配统计信息来讲,识别参数误差是一种用于选择一个模型作为最佳拟合模型的办法(前提是它在候选模型中),然而,在现实中,很难判断真正的模型是否是候选的模型之一,因为存在着参数误差的问题。绝对拟合评估是指确定现有模型是否充分拟合数据的过程。在绝对拟合评估下对参数误差

敏感的拟合统计应该以高概率拒绝参数误差的模型,然而,在实践中,很可能不止一个模型能够充分拟合数据。

仿真研究表明,对于相对拟合评估,BIC(在某种程度上是AIC)可用于检测 CDM 参数误差, Q 矩阵误差或两者同时存在。饱和模型可以在检测 CDM 参数误差或 Q 矩阵误差中发挥重要作用。对于 CDM 参数误差,它可以用于区分可能是真实的还是参数误差的 CDM。对于 Q 矩阵误差,它可以用作在 Q 矩阵之间进行比较的真实模型。对于绝对拟合评估,项目对的观察和预测相关性与 Fisher 变换之间的残差及成对项目响应的观察和预测对数胜算比率之间的残差具有相似的性能,并且在大多数情况下对不同的错误指定敏感,尽管两种统计数据都倾向于保守。对于这两个统计量,饱和模型在大多数情况下可以作为真实模型,然而,除非涉及高度受限的约束诊断矩阵,否则两者对过度指定的 Q 矩阵都不敏感。

5.1.5 小结

本节主要介绍了认知诊断的发展及基本概念,以及 3 个经典的认知诊断模型,同时对认知诊断模型在智慧教育中的应用做出了相关的介绍与展望。通过本节的介绍,读者对认知诊断及认知诊断模型的应用有了初步的了解。

5.1.6 思考与练习

- (1) 什么是认知诊断?
- (2) 认知诊断模型分为几种类型? 分别有什么?
- (3) DINA 模型的优点与缺点分别是什么?

5.2 自适应学习

移动互联网、智能教育的发展,极大地延伸了传统的学习空间和教育实践,技术作为人的存在方式,正在促使教学模式和学习方式发生深刻变革。在人工智能(AI)、大数据、移动互联网和云计算等新兴技术的支持下,数据驱动下的自适应学习日益成为教育界关注的焦点,并逐渐发展成为以大数据为基础的教育技术新范式。我国《教育信息化“十三五”规划》、《教育信息化 2.0 行动计划》和《新一代人工智能发展规划》等文件均明确了个性化学习的重要性,指出“关注学习者不同特点和个性差异”“为每个学习者提供适合的教育”“探索在信息化条件下实现差异化教学、个性化学习、精细化管理、智能化服务的典型途径”等内容。由此可见,关注个性化学习、尊重个体差异、促进学习者的个性化发展,是我国未来教育发展的重要内容。目前有关自适应学习(Adaptive Learning)的研究,本质均指向个性化,认为将其作为一种实现个性化学习的具体方法,可以为学习者提供个性化服务。在数字化学习环境中,自适应学习通过特定的科学手段挖掘学习者的个性特征差异,全程记录和分析学习行为数据,对学习行为、学习过程和学习结果进行价值判断,以此开展个性化的干预和指导,从而促进知识意义的主动构建和有效学习。

5.2.1 自适应学习概述与框架

近年来,随着如 Knewton、InterBook、ELM-ART、AcroBatiq、Smart Sparrow、英语流利说和猿题库等系统的大规模普及与应用,自适应学习的发展尤为迅速。对国内外相关研究发现,当前自适应学习的研究主要集中在系统模型构建、框架设计和平台开发、技术应用研究、系统模块组件开发和构建及对研究现状综述分析等领域,知识图谱及其表征学习、知识追踪、路径推荐等为其研究热点,神经网络、深度学习、贝叶斯网络技术也被广泛应用,然而,根据文献梳理,当前与系统模块相关的组件的关键技术多采用自然语言处理、机器学习、深度学习等方法实现,可解释性效果不佳,在某种程度上容易形成“黑箱”问题,因此,针对教育技术领域面临的 key 问题和现有方法的缺陷,本节在对自适应学习框架进行解读的基础上,开展相关组件关键技术的研究和存在问题的梳理分析,并对近年来解释性较好的技术研究做出介绍。这些研究和应用经验的借鉴和综合运用,将成为推动学习者个性化发展行之有效的策略。

当前自适应学习框架主要由领域知识模型、学习者特征模型和教学模型三部分组成,其又称为经典的“三角模型”。其中,领域知识模型包含学习领域的逻辑体系、基本概念和原理、规则定义及探究方式等。学习者特征模型动态描述学习者的认知状况、学科知识、学习历史、学习风格和偏好及情感状态等方面的个性特征。教学模型决定适合学习者的学习活动和教学策略。此三角模型正是教学过程的“三要素”(教师、学生和教学内容)在计算机辅助教学(Computer Aided Instruction,CAI)中智能化、程序化的实现,相关

工作原理(如图 5-10 所示)为:教学模型根据领域知识及其推理,对学习 者特征模型反映出的知识水平、认知能力、学习风格和偏好等加以诊断和分析,做出适应性决策,动态安排高切合度的学习内容、学习资源及其呈现方式,有针对性地向学习者提供个性化推荐服务。同时,对学习过程进行实时监测和管理,动态获取学习者的表现数据,由此不断训练、更新学习者的特征模型。

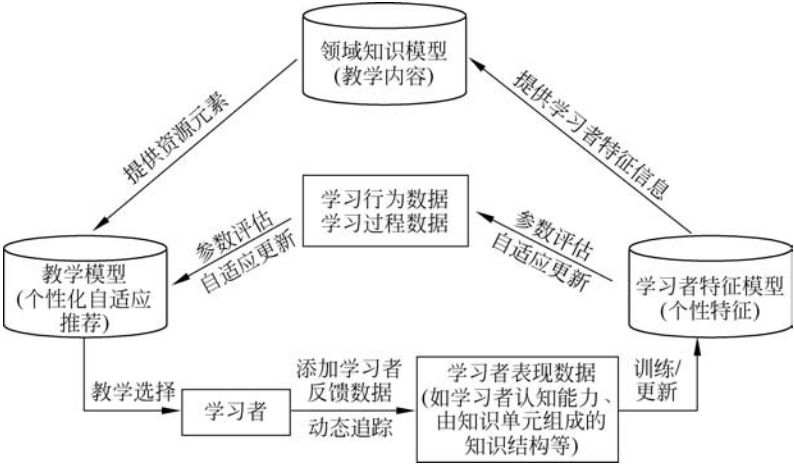


图 5-10 自适应学习基本框架

1. 领域知识模型

根据关联主义学习理论,知识是一个由相互联系的事实、概念、命题和规则等要素构成的网络,而学习是为了增加对这个网络的认知和理解,从而促进有基础、有意义的创造,更好地开展个体构建。领域知识模型对应用领域的组成元素及其结构进行描述,表示内部各组成元素及其之间的相互关系,一般由语义万维网、领域本体、层次结构和知识图谱等技术构建,其核心作用在于支持知

识的获取、组织和推理等。在大数据视角下,CHAPLOT 和 KOEDINGER 提出使用教育知识图谱来引导一个过程中多个单元间的先决条件关系,其他学者也通过引入神经网络、机器学习、深度学习等技术对教育知识图谱构建过程中的关键环节(如实体识别、关系提取、知识表示等)做了深入研究。在知识表征学习领域,当前绝大多数工作聚焦于通用知识图谱方面,BORDES 等将“关系”解释为低维向量空间上头部和尾部实体间的转换操作,Ji 等提出了一种使用两个向量来表示实体和关系的 TransD 细粒度模型,Xu 等提出了一种用于学习实体结构和文本信息联合表示的深层体系结构,Kazemi 等提出了一种基于张量分解方法的双线性模型来解决头尾实体关联性问题等,然而,以上技术大多仅停留在表层学习概念间的链接关系,缺乏对实体重要性及不同类型实体间的关系研究,同时对海量多元异构媒体资源的跨图谱表征学习也存在稳健性不足的问题,与实际的应用需求仍有较大差距。

2. 学习者特征模型

学习者特征模型是自适应学习的核心和基础,反映个体自身及其行为所受强化关系上的个体差异,预示不同的学习行为表现。每个学习者的个性特征各不相同且动态变化,因此需借助 AI 技术对学习者的行为序列进行动态检测和建模,预测其对知识的掌握程度及学习趋势,相关方法主要有知识追踪、覆盖模型和贝叶斯网络等。如 Corbett 和 Anderson 提出了一种贝叶斯知识追踪(Bayesian Knowledge Tracing, BKT)模型,该模型从结构上来讲是一个隐马尔可夫模型(Hidden Markov Model, HMM),其将学习者的知识状态表示为一个二元组{掌握该知识点,未掌握该知识

点},并根据当前知识状态来预测隐变量的概率分布。Piech 等提出一个深度知识追踪(Deep Knowledge Tracing,DKT)模型,通过循环神经网络(Recurrent Neural Network,RNN)对学习者的知识状态进行建模后,利用长短期记忆网络(Long Short-Term Memory,LSTM)模型追踪学习者随时间变化的知识熟练程度,相关效果被证明优于 BKT 等传统模型。Yeung 等在原始 DKT 模型的损失函数中引入与重建了与波动相对应的正则项和正则约束,增强了跨时间补偿预测性能的一致性。根据以上研究和实验发现,基于深度学习的知识追踪模型已被证明在无须人工特征的情况下优于传统的知识追踪模型,但也存在未考虑到学习者知识状态将随知识难度和遗忘规律影响等问题,并且相关参数和表示一直被指出不可解释。

3. 教学模型

教学模型根据知识间蕴含的前驱和后继关系,综合考虑学习者当前的知识状态、认知能力、学习风格及偏好等特征,有针对性地推送个性化的学习路径和学习资源,其构建方法主要有基于内容的推荐、基于协同过滤的推荐及混合推荐等。在早期的学习推荐系统中,基于内容的推荐技术被广泛使用,但这些方法存在采用人工标记耗时且仅适用于小规模数据等方面的问题。后来,有关学者应用协同过滤技术或混合过滤的方式提供相关的自适应信息,如应用本体技术、语义、上下文关系等改善了这一情况,但也发现混合过滤并不能轻易通过改变结构而改善结果。陈敏等以“学习元”平台为例提出泛在学习的内容个性化推荐模型,Tang 等通过应用 RNN,实现下一步学习推荐可高达 60% 的预测准确度。以

上提到的绝大多数推荐方法,只是在现有学习数据支持下就学习者的知识背景和学习条件而开展的推荐,其可解释性效果不佳。

5.2.2 数据驱动下自适应学习系统的支撑模型与实现机制

在大数据时代,越来越多的学习行为能够被追踪和记录,教育从“用经验说话”向“用数据驱动决策、管理与创新”的方向发展。数据驱动下的自适应学习通过大数据分析学习过程和学习行为,能够精准识别学习者的个性特征、动态监控学习过程、实时预测学习趋势和有效评价学习结果,给予学习者个性化的干预和自适应的指导,因此,如何对海量教育数据的概念特征进行自动提取并建立关系,如何追踪学习者在学习过程中不断变化的个性特征,如何精确定义学习者每一步要学习的知识单元等,均是当前教育实践中亟须解决的科学难题。鉴于此,针对经典的“三角模型”建立一个可解释的自适应学习技术框架,需着重对教育知识图谱的构建与表征学习、知识追踪和个性化学习路径推荐等核心技术难点进行研究,如图 5-11 所示,以解决教育数据中的概念边界检测、教育知识图谱表征学习的实体间语义信息传播、深度知识追踪的数据稀疏化和不可解释、个性化学习路径推荐融合课程序列等问题。

1. 教育知识图谱的精准构建与表征学习方法

知识图谱(Knowledge Graph, KG)作为诊断学习者对知识和技能的真实掌握状况的底层依托,知识元抽取在本质上来看属于序列标记问题,因此,可将教育数据的概念提取并看作一个词语序列标记问题。考虑到教育数据的顺序性及其内部词语的依赖性,

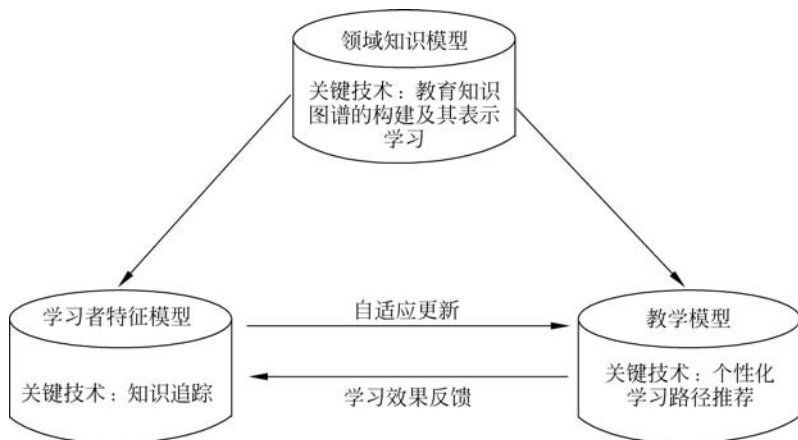


图 5-11 自适应学习框架的相关组件及对应关键技术

Huang 等在融合双向长短期记忆网络 (Bidirectional LSTM, BiLSTM) 和条件随机场算法 (Conditional Random Fields, CRF) 的基础上提出了 BiLSTM-CRF (Bidirectional LSTM-CRF) 模型, 该模型的精准度高且对词向量的依赖性较低, 李振等对基于深度学习的知识元抽取并做出深入分析后指出, BiLSTM-CRF 模型是当前序列标注问题解决领域中较为成熟的应用, 该模型结合了 BiLSTM 和 CRF 的特点, 具有序列建模能力强、特征抽取自动化的优点。Ma 等将卷积神经网络 (Convolutional Neural Networks, CNN) 融入 BiLSTM-CRF 模型中并提出一种“端到端”的 BiLSTM-CNN-CRF 模型, 该模型通过采用 CNN 进行文本字向量特征学习, 识别效果得到显著提升。此外, Li 等专注命名实体边界检测, 提出了 AT-BDRY (Adversarial Transfer for Named Entity Boundary Detection) 模型, 通过无监督的传输学习方法来减少源域和目标域之间数据分布中的差异, 无须任何手工特征或任何先前

语言知识便能从文本中检测出实体边界。为了融合实体的文本和结构信息, Kipf 等提出一种图卷积网络(Graph Convolutional Networks, GCN)的半监督学习模型, 该模型通过谱图卷积的局部一阶近似确定卷积网络结构的选择, 学习隐藏层表示, 编码局部图形结构和节点特征, 并可直接用于图结构数据处理。此外, Schlichtkrull 等引入关系图卷积神经网络(Relational Graph Convolutional Networks, R-GCN)构建知识图谱, 对链接预测和实体分类的2个实验均验证了 R-GCN 作为实体分类的独立模型的有效性。领域知识建模是构建自适应学习系统的关键, 然而, 当前的研究主要依赖可编码、可量化的显性知识进行建模, 忽略了隐性知识的应用, 因此, 如何利用教育知识图谱对学习者在学习过程中的隐性知识进行表示和建模, 如何有效实现隐性知识和显性知识间的转化, 将成为未来教育知识图谱研究和发展的重点和难点。

2. 基于深度学习的知识追踪

知识追踪基于学习者的行为序列进行建模, 其能够预测学习者对知识的掌握程度, 是自适应学习系统构建的核心和基础。近年来, 基于 RNN 的知识追踪方法因具备捕获人类学习的复杂表示能力, 其效果优于其他所有传统方法而被广泛应用, 但同时也不可避免地存在输入序列重构、预测结果波动和处理稀疏数据时无法泛化等问题。这些问题虽可以采用在损失函数中引入正则项并对输出结果进行正则约束等方法进行优化, 但在效果提升方面却不显著且缺乏足够的可解释性。为此, Nakagawa 等提出了一种基于图神经网络(Graph Neural Networks, GNN)的知识追踪方法, 该方法将知识结构转化为图形, 从而间接将知识追踪任务重构成

GNN 中时间序列节点级分类问题。从数据结构的角度来看,知识结构可以以图形 $G=(V,E,A)$ 的形式进行组织,将有关数据图形结构性质的先验知识整合到模型中,提高知识追踪的性能和可解释性。相关实验表明,该方法可以潜在改善对学习者的预测效果,在无须增加其他信息的情况下更具解释性。近年来,通过深度学习处理图结构数据的 GNN 方法研究备受关注,各种泛化框架和重要操作陆续问世,并在相关研究领域取得了成功结果。

另一方面,知识追踪也可以看作对学习者的参与学习活动时知识概念(Knowledge Concepts, KCs)的掌握程度进行建模的一项任务。Pandey 等认为:在知识概念的学习中,学习者在各项学习活动中所掌握的技能彼此关联,并且取决于与该学习活动相关联的过去学习表现;为有效解决数据稀疏时无法泛化等问题,他们提出了一种基于自我注意(Self-Attention)的知识追踪模型,该模型能够在不使用任何 RNN 的情况下模拟学习者的互动历史,并通过学习者在历史互动中的学习行为表现进行推理和预测,相关实验表明该模型比基于 RNN 的方法快一个数量级。此外, González-Brenes 等研究表明,通过知识追踪和其他建模方法的组合应用,可有效提升模型的预测精度,如 CAI 等采用知识追踪和回归分析模型相结合的方法研究学习者的整体学习趋势,预测未来学习趋势和表现, Khajah 等结合知识追踪和项目反应理论(Item Response Theory, IRT)模型来预测学习者的知识掌握,获得了显著成效。

总体来讲,训练知识追踪的目标是利用学习者的历史学习数据去预测其未来学习表现, DKT 模型在优化学习效率、发现不同知识点间的内在联系、动态反映学习者的连续知识水平变化等方面表现出强大优势,但同时也存在着模型无法重构、学习者对知识

点掌握程度不连续等问题,未来需进一步对各种相关的DKT+模型进行探索和研究。

3. 个性化学习路径推荐

目前有关自适应学习的研究,本质上均指向个性化,认为可将其作为一种实现个性化学习的具体方法,为学习者提供个性化服务。当前,海量的学习资源、碎片化的学习时间、复杂的学习情境及师生分离的学习空间等,加剧了学习者的“信息过载”和“知识迷航”问题。研究发现,在自适应学习系统中,如果缺乏准确的导航性学习路径的支持,学习者就可能会难以准确、快速达成既定的学习活动要求和确定的学习活动目标。结合学习者的智能认知学习状态和智能认知学习目标,为其智能规划和推荐合适的学习路径,是提升自适应学习系统个性化服务质量的关键。学习路径生成的科学本质是根据学习者的学习目标和认知状态来正确对待被组织学习的各种知识元素并进行确定顺序的科学过程。学习路径生成问题可以描述为:在已有学科知识元及其拓扑关系、学习者的学习目标及先验知识结构的前提下,对学习者的待学习的各门学科专业知识点的元数据集进行了分类排序,生成一个涵盖学习目标的各门学科专业知识元序列。

与常规的推荐系统类似,个性化自适应学习路径推荐除了需解决常见的数据过大和冷启动问题外,还需考虑系统中其他模块的输出问题。个性化学习路径推荐是一个详细的推荐过程,如果仅使用学习者数据,难度是非常大的。为了提供更加准确的多样和可解释的推荐,Wang等在基于KG与用户-项目图(User-Item Graph)的混合结构中提出了一种知识图注意力网络

(Knowledge Graph Attention Network, KGAT)方法,在 GNN 框架下以“端到端”的方式实现知识图的高阶关系建模。该模型采用递归方式传播来自节点邻居的嵌入并引入 Attention 机制,用于区分邻居嵌入的重要性,相关实验证明了其在理解高级关系重要性方面的可解释性。

学习路由不同的节点组成,每个节点都代表着一个知识点,每门课程均包含着或多或少的知识点。Nabizadeh 等提出了一条融合课程知识关系的学习路径推荐模型,该模型从课程序列图中选择所要学习的知识点后,组合知识关系和深度优先算法对所选知识点的课程序列进行深度搜索,估算学习时间及得分,再根据学习目标等要素推荐一系列满足学习者有限时间限制的学习路径。相关实验证明,该模型在最大程度上提高学习者分数的同时可满足时间限制,但仍存在学习者和学习对象冷启动问题。

冷启动问题在推荐系统中最为普遍,如果没有学习者的学习行为数据,则后续的神经网络等模型便无法正常运行。为有效地解决此问题,Pliakos 等提出了一种将 IRT 模型和机器学习相结合的混合方法,该方法将 IRT 与基于学习者辅助信息的分类树、回归树相集成,并对学习者能力评估和项目反应进行预测。实验结果表明:IRT 与随机森林相结合可提供误差最低和响应最高的预测准确性,有效减轻学习环境中冷启动问题的影响。Zhou 等提出了一种基于 LSTM 神经网络融合知识关系的全路径学习推荐模型:先基于个性特征相似度对学习者集合进行聚类并训练 LSTM 模型,预测学习路径及其表现,再从路径预测结果中选择个性化的学习路径并进行推荐,从而有效地解决了没有学习行为数据的学习者路径推荐问题。学习路径中每个节点都具有关于学习者个体特

征、学习内容和学习资源等的多维数据,与文本数据的特征相似,学习路径数据同样具有序列结构,实验证明该模型对各种数据集都有较好的权衡。知识关系来源于知识追踪,由于知识追踪模型具备可解释性,因此该推荐模型也具备了一定的可解释性。

5.2.3 自适应学习典型案例分析

1. Knewton

随着我国大数据技术日趋成熟,在数据驱动下的应用和创新也越来越多地涌现出活力。在以教育质量和教育方式为重点的教育改革中,智慧教育的探索步伐从未停止,其中,如何实现学生个性化学习与教育需求成为重点研究方向之一。Knewton 作为一家大型美国自动化适应职业学习服务基础设施服务提供商,其所致力开发的自适应学习平台为学习者提供个性化教育内容成为可能。该平台针对学习者学习目标的模糊性和动态性等特点,从空间强化、学习记忆曲线和学生档案等主件出发,实现对学习者的认知状态诊断和学习目标动态规划。该平台借用知识间的转移概率对推荐引擎进行优化,使用层级聚类对学生进行分组,所提供的服务主要基于对学习者数据的分析,基本要素主要包括教学过程中数据的采集和处理、基于教学数据的认知推理及个性化学习路径推荐,其主要技术核心包括以下几方面。

(1) 基于分布式框架的大规模数据处理能力,对海量学习者的学习过程、行为、档案等数据的收集及整合处理。

(2) 以学习者的个性特点、学习风格、认知能力水平、知识结构和所掌握的知识对象的难易程度等作为推理的特征,定义了关联

匹配原则,以学习者需要的学习服务(如所掌握的知识内容等)为输出特征,通过大规模教育领域知识图谱及逻辑推理引擎的构建,实现认知推理。

(3) 通过对学习者的各种个体特点、学习行为等多维信息进行整合,调用各种领域性的知识对象的困难程度和样式进行风格化的描述,与其他学习者的行为特点进行相应匹配计算,对学习目标、学习内容、学习路径进行分析,在获取知识间关系及知识难度系数的基础上,采用机器学习等技术途径,构建更加优质的推荐引擎。

Knewton 自适应学习平台能够实现高效的学习行为分析和优质的学习资源综合推荐,离不开 4 类核心学习数据模型的综合支持,包括知识图谱、学生事件、目标管理及推荐与分析 API,如图 5-12 所示。知识图谱是 Knewton 提供个性化的知识学习和应用信息技术服务的技术基础,该平台基于学科领域专家知识建立的知识图谱,并与学习者的过程要素进行有机关联。

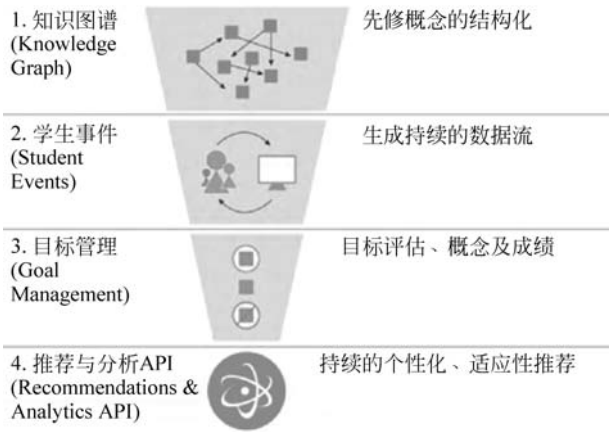


图 5-12 Knewton 自适应学习平台核心数据模型

学生事件是学习者在学习过程的行为记录,是学习者的学习途径分析及学习效果评估的重要数据资源。目标管理是学习平台针对每个学习者的智能化学习史、状态和评价进行量身定制的目标,随着每个学习者的特点变化而自动化地对其进行智能化的规划和调整。推荐与分析 API 是指学习者对于一个学习系统在使用时所遇到的内容进行推荐和分析的结果报表,以 API 的形式为用户提供了服务,在学习者的学习过程中即时提供匹配的合适的内容。

空间强化是指新的概念或技巧在被学习吸收时,以前学过的概念或技巧也被强化。由于新的知识往往是采用循序渐进、以其他学习者所熟悉的形式或基于其他学习者所熟悉的知识内容(如先验性知识等)进行组织和编排,与传统的强化不同,Knewton 采用空间强化以注重其他知识在一定时间上的扩展延伸性分布和甄别筛选抽取,使新旧知识之间融会贯通,保证了学习者在掌握和习得新知识的同时,旧的知识也能在实践中得到极大程度上的温习。与此同时,受到艾宾浩斯遗忘曲线的影响,Knewton 运用了指数的增长与衰退曲线,将其作为一种描述被学习者自我学习与遗忘能力的曲线。相关模型的基本假设是:如果一个学习者已经接触过一些与相关的学习课程的内容,则其对该课程主题的知识或技能的记忆将更加深入一层;若一个学习者从未真正地接触到相关的主题内容,则很可能在自己学习完一段时间后被遗忘。

鉴于学习者能力的发展变化及多种智能之间是相互联系和连接的,Knewton 对其在传统的项目反应理论基础进行了拓展,认为多种智能学习者自身的知识和能力参数将会随其在一定学习阶段的时间内发生动态变化,并且其与智能相关的表征非仅限于某一固定参数,由此从问题层级上的表现方式对学习者的知识概念

参数进行了建模,通过运用聚焦于概念层面的知识图谱来分析学习者的知识概念参数并做出表征和评价。

2. ALEKS

ALEKS 是专门针对 K-12 及其他高等教育中的各种数学学习情况进行的自适应性评价。1994 年,加州大学欧文分校研究室基于人脑神经测试和数学算法等多个领域的一些相关知识研发并推出了 ALEKS(Assessment and Learning in Knowledge Spaces,知识空间评价与学习),该平台已经可以广泛地用来对每位学习者在课堂上所学的数学知识进行智能化评测。2013 年,ALEKS 被麦格爾-希尔(Mcgraw-Hill)集团收购。该平台采用的是一款基于互联网络的人工智能课程评价和学习系统,在线智能化的课程体系是其最广泛而又核心的业务。围绕着智能化课程这一核心产品,ALEKS 逐步开发了 Quicktables、ALEKS360 和 Response to Intervention 等多种教学辅助产品,形成了多用户端的产品矩阵。ALEKS 作为一个线上学习体系,覆盖了从 K-12 到高等教育的绝大多数数学项目,现在它还可以为在家进行自主学习(Home-Schooled)的学生提供在线学习和辅导。

ALEKS 基于知识空间理论(Knowledge Space Theory),通过评测学习者对数学知识的掌握程度和适应性提问,在系列测试问题中获取学习者的认知水平、知识结构信息后,快速准确定位学习者已掌握的知识和未掌握的知识,并选择最适合该学习者学习的知识点来推进学习。当学习者首次使用 ALEKS 时,将进入该系统教程学习界面,该教程将对 ALEKS 的每个操作流程进行详细介绍。接着,学习者将进入测试评估界面,需要在 45 分钟左右的时

间内完成 20~30 道数学题的测试,通过测试评估,系统将获取每个学习者的认知水平和知识结构等,并通过学习者的大体情况精准定位其相关知识的掌握阶段,从而为学习者推荐个性化的学习路径。测评工作完成后,学习者将能够观察和看到系统自动制作生成的 ALEKS 知识饼图(Aleks Pie),该图以不同的颜色进行标注,让学习者迅速地了解当前自身的实际学习状况与各方面对知识的掌握程度。

在学习者的测试过程中,为了避免学习者直接跳过测试内容,造成测试结果的不具参考性,教学者也需确保学习者对所学课程内容有个清晰了解。在测试过程中,系统将根据学习者在上一道题中所回答的正确情况,智能地判断下一个问题的难易和复杂程度,同时系统也将根据学习者上一次的测验成绩,智能决定下一次测验的难易程度。在学习者的测验过程中,ALEKS 将不断重新评估学习者的学习情况,巩固强化所掌握的内容。每次的测验结果都将以 ALEKS 知识饼图的形式呈现给学生,以便学生对自身知识的掌握程度、认知水平和历史成绩有个清晰认识,测验结果也可随时向教师和家长发送。

在这个课程的实际应用场景中,为了有效避免大部分学生在测试的环节中遇到许多项目的选择题而胡乱进行选择的情况,ALEKS 模拟了传统的线下测试形式,采用主观题要求学生在一个作答框内输入正确的选择题答案。在课堂的初始,当学习者登录 ALEKS 后,将首先进入使用说明视频界面,该视频教程将对 ALEKS 每个操作流程进行详细介绍。随后,学习者将顺利地通过第一次自适应评估和测试,然后继续进行一次为期 45 分钟左右的新技术和课程知识的掌握。在学习了所有的新知后,学习者将需

要在系统上按时地完成 20~30 道数学题的实验,以便于检验当前知识点掌握的程度。相关测评的题目均应是根据学习者初次测评的结果、所学课程的内容及对上一个问题的反馈情况进行智能化地生成并调整。ALEKS 能够准确地定位到每个学习者正处于的认知状态,所以为每个学习者所推送的检测内容都是各不相同的。

也就是说, ALEKS 一开始对每个学习者进行教学设置都应该是相同的。当每个学习者都开始针对系统自动产生的问题进行作答时, ALEKS 的机器学习源代码程序便会自动地进行分析各位学习者的作答过程,从而可以准确地判断其中的各种相关知识和概念在理解上所掌握的程度。若学习者对某种类型的背景知识问题出现错误率比较高,系统则将会自动地向学习者推荐相关背景知识材料,引导其进行视频学习或给予一些内容提示。如果学习者轻松完成了问题测试,系统则会让他们进行更多的相关领域的知识测试。当学习者熟练掌握该领域的知识后,系统将对学习者进行深度测试。

ALEKS 线上课程中的测试题目主要来源于教材各章节后的习题、教研团队人工出题和算法自动生成题目。ALEKS 的教研团队将挑选一批具有多年教学经验、高专业学历、多领域知识丰富的技术人员进行测试性习题的人工开发,例如每个细微的知识点往往需要 400 道试题,那么我们的教研团队至少将为此做 650 道试题。同时, ALEKS 还将通过引入深度学习、机器学习等新型人工智能算法,不断地学习有关针对不同的学习阶段、不同领域知识的各种类型题目,为学习者提供海量的题库内容。相关算法自动生成的引入防止了“回音壁”效应,有效地规避了学习者在同一时间内对习题重复练习的情况。

此外,ALEKS 还推出 QuickTables,通过分布式适应训练考核机制(Distributed Practice Mechanism)为每个班级的学生提供了自定义适应课程考核能力评估和其他相关的自编程适应训练内容,让学习者的学习进度一目了然。该辅导课程通过定期的多次测试结果评估和再次模拟测试,对每位优秀学习者的基础知识及基本掌握数据进行了多次虚拟测试监测和多次实时测试反馈,并根据每位优秀学习者的再次测试评估结果为其进行量身设计,以便为其提供一种更为个性化的基础学习辅导内容,有效地提升了每位优秀学习者对基础知识的基本掌握和学习效率,同时也有效地做到了“因材施教”。

5.2.4 小结

本节从数据驱动的视角出发,通过国内外研究综述分析,对自适应学习系统框架和相关组件进行阐述和解读,重点从领域知识模型、学习者特征模型和教学模型三方面对其实现机制进行了探析,在此基础上对个性化学习路径推荐研究进行探讨。教育是一个复杂的系统工程,自适应学习和个性化学习路径推荐的进一步丰富和发展需回归教育本质。从教育的角度出发,将学习路径的个性化推荐研究与教育全过程相结合,在教育教学理论的指导下开展推荐模型的构建和系统的研发,真正为学习者提供更具个性化的学习服务,将成为未来自适应学习研究的重中之重。

5.2.5 思考与练习

(1) 什么是自适应序列? 请简要分析自适应学习中可能出现的自适应序列类型。

(2) 自适应学习是怎样实现的？请简述自适应学习的内部系统结构。

(3) 自适应学习需要什么样的数据？请简述自适应学习的教学数据分析方法。

5.3 个性化学习路径推荐

随着自适应学习的飞速发展,个性化学习路径的推荐越来越重要。与向所有学习者提供相同资料的传统学习不同,自适应学习旨在为每个学习者量身定制个性化的学习项目和路径。自适应学习系统通过学习者当前的知识水平和学习先决条件关系等为其推荐一条个性化的学习路径 $A \rightarrow B \rightarrow C \rightarrow D \rightarrow E \rightarrow F \rightarrow \dots$ 。个性化学习路径有助于学习者有效理解新的学习内容和学习项目。关于个性化学习路径推荐的研究发展现状 5.2 节已进行了论述,在此不再赘述。教育研究表明,认知结构对适应性学习具有很大影响。认知结构描述了知识点的定性和发展,包括学习者的知识水平和学习内容的知识结构(如先决关系等)。知识水平反映的是对不断演化且无法直接观察到的学习者对学习内容的掌握程度,同时知识结构捕捉的是各个学习内容之间的拓扑关系。目前关于个性化学习路径推荐的研究,多是利用人工智能系统中用户模型的构建和推荐算法等方式进行。相关流程为依据学习者在线学习过程中的个性化需求的信息内容进行检索,通过对学习对象信息资源的收集、整理、归纳和分类,向潜在学习者推荐所需的学习路径和学

习资源。近年来推荐算法主要根据学习者的在线答题记录和学习者之间的相似度,找到最相似的学习者,而依据最相似学习者的相关数据对目标学习者进行得分预测,然后进行学习路径和学习资源的推荐。在学习路径规划中,通常可以使用基于内容的自适应推荐、协同过滤推荐和混合式推荐等算法来帮助实现个性化自适应推荐,本节内容首先遵循推荐系统中的一般分类方法,对常用的学习路径推荐算法做出介绍。

5.3.1 学习路径推荐算法

1. 基于内容的推荐

基于内容的推荐算法,主要利用商品的内容或者一些固有的属性进行推荐,具体做法为通过收集项目、用户的各种相关资料信息及用户对该项目的运营管理操作行为等数据来设计和构建一个推荐算法的模型,其原理是根据目标用户的历史行为,获取用户的兴趣偏好,并通过对历史行为分析完成目标用户兴趣偏好建模,从而为目标用户推荐与其兴趣偏好相符合的项目,属于 Schafer 划分的项目到项目关联。该算法主要采用机器学习方法从关于项目内容的一个特征描述事件中获取一些关于用户感兴趣的偏好数据资料,即直接通过对项目内容的信息进行推断,无须再依据使用者对该项目进行评价。其中核心步骤分别为:基于项目的用户信息及其他用户的操作行为构建用户特征描述表示、基于项目的信息构建特征描述表示、基于用户及其他项目的特征描述表示。

该算法首先通过显式反馈、隐式反馈或者显隐式相结合等多种方式从网络中获取行为项目用户交互的行为项目数据,然后在

网络中通过行为项目的数据特性学习出待推荐目标用户的兴趣偏好,在此基础上进行计算,以此得出待推荐的目标用户和其他行为项目的内容特性相似程度,最终根据特征相似度对所有待预测项目进行综合排序,从而为目标用户推荐其潜在兴趣项目。其中,显式反馈是指使用者明确地表示自己对项目的偏爱和行为,如点赞、收藏、评分、评级等;反之,隐式反馈则表示不能直接表现用户兴趣倾向的行为,如浏览、搜索、观看、单击、购买等。常用的基于具体用户的相关信息分析模型数据构建的分析方法主要包括决策树、神经网络和基于用户向量的数据表示分析方法等,相关信息模型的分析构建则通常需要一个基于用户的具体历史活动行为模型为数据依托点进行分析数据为支撑,因而可能随用户的兴趣偏好转移而改变。

基于内容的推荐无须大量评分记录的数据,依靠了目标用户的历史行为资料、兴趣偏好的内容及项目特点信息等,因而很好地避免了评分记录资料稀疏的问题。同时对于一个新的项目而言,该模型推荐算法仅仅需要对相关项目的特征进行提取,便可直接向目标用户进行推荐,通过学习模拟推荐项目的相关内容和特征,便能够对其被推荐的原因做出解释,有效地彻底解决了一个新项目冷启动的问题。当前如分类学习等相关推荐技术已然发展成熟,然而,由于特征性的内容需要具有良好的结构性,并且在特征推荐的过程中目标用户的需求必须是能够以特征性的内容为基础和形式进行表达,无法显示获取其他用户的判断数据,因而该推荐算法也常遭遇特征提取困难等问题。

2. 协同过滤推荐

协同性的过滤,简单来讲就是一种利用个人和使用者间拥有

的一种共同体验,并且具有类似的兴趣和偏好的方式,来发现目标用户对项目的潜在偏好并做出推荐,简单来讲,就是结合相似用户间的兴趣商品所做出的推荐。在教学实践中,协同过滤推荐可以推荐所有形式的内容,如文字、图片、视频、音频、动画等。该算法一般采用最相关邻近的技术,根据每个目标用户的历史兴趣偏爱行为资料进行计算后得到与用户之间的距离后,通过借鉴目标用户的最相关邻近用户对项目进行加权评估,以此来准确地预测每个目标用户对相关项目的偏爱程度,从而依据该喜好程度完成项目推荐。因而,该类推荐计划算法主要通过将未经评分的项目进行评级和预测来设计和实现,所推荐的项目不一定是目标对象特别感兴趣的。

协同过滤推荐主要分为基于启发式协同推荐过滤算法(Memory-based Algorithms)和基于组合模型的协同推荐过滤算法(Model-based Algorithms)两种类型,不同的协同过滤算法间也存在较大差异。启发式推荐算法先通过每个历史用户评价和项目效用评估值之间的差异比例去计算每个使用者之间或每个项目之间的相似度,然后依照每个使用者的项目历史使用评价和项目使用者之间相似度比例计算每个使用者的推荐效率。启发式的数据推荐分析算法更加容易被用户实现,并且其中所推荐的数据结果同样具有很高的可解释性。这种过滤算法又可分为基于用户的协同过滤(User-based Collaborative Filtering)和基于物品的协同过滤(Item-based Collaborative Filtering)两大类。基于模型的推荐算法利用基于模型矩阵计算函数、可分解计算方法的模型技术及其应用模型来自动构建一个关于预测用户消费倾向的推荐模型,预测出一个用户对于某个项目的潜在使用倾向。该种分析算法能

够有效地分析并缓解这些类型数据的信息稀疏化缺失问题,常用的分析方法主要有基于数据规律的分析模型、决策树、贝叶斯方法及潜在直接影响数据因素的分析模型等。

综上所述,协同过滤主要的目标之一就是寻找和筛选用户最近的邻居,从而依照最近邻居的偏爱和喜好来做出潜在对该项目的评级和预测。这种计算方法主要可以分为以下 3 个步骤。

1) 用户评分

收集可以代表用户信息的数据,一般采用评分或给予评价的方式进行,主要有显性评分和隐性评分。显性评分指用户直接对项目的数值进行评分,而隐性评分则表现为用户对项目的单击、浏览、购买等行为。

2) 最近邻搜索

搜索与用户兴趣偏好相同的另一组用户,然后计算两组用户的相似度。一般采用 Pearson Correlation Coefficient、Cosine-based Similarity、Adjusted Cosine Similarity 等方法来测算用户间的距离。

3) 产生推荐结果

产生了最近一个邻居的聚类集合后,根据这个聚类集合对于一个目标客户感兴趣的某个未知项目做出了评分和预测,再把这个评分最高的一个项目引入其他目标用户。较常用的推荐结果有 Top-N 推荐和关系推荐。Top-N 推荐主要是针对每个个体目标的用户所产生,对每个目标的用户所推荐不一样的结果,而关系推荐则主要针对最近一次推荐相邻的目标用户进行的一种类似关系推荐规则(Association Rules)的分析挖掘。

协同推荐算法可以通过过滤一些机器难以自动进行的内容分

析的信息,帮助用户找到潜在但自身尚未被发现的兴趣偏好,能够有效地使用其他类似用户的反馈信息,减少了用户的反馈数量,加快了个性化学习的步伐。尽管如此,该算法仍有许多问题需要解决,最为典型的问题是用户对一个项目的评级数据相对比较少,常常遭遇到评级数据稀疏的问题。此外,对于一个新用户或者一个新项目而言,系统刚刚开始时所提供的推荐数据质量比较差及所提供的推荐数据集的质量主要取决于所提供的历史数据集,往往会出现因为没有提供历史评级的数据而导致无法实际进行推荐的情况,存在冷启动、算法健壮性等问题。

3. 混合推荐

鉴于基于内容的推荐、协同过滤推荐等单一的推荐算法均可能存在各自不足,如数据稀疏、冷启动、算法健壮性等问题,因此我们可以考虑通过将不同的算法分别进行组合推荐,避免或有效地弥补各自的推荐技术的薄弱点,从而使更好的推荐在实际运行中产生。混合式推荐体系也被认为是推荐体系的另外一个重要研究热点,协同过滤技术和其他技术相结合,不再单独使用一种推荐算法对商品进行推荐。常见的组合策略主要包括3种,分别为后融合、中融合和前融合。

(1) 后融合的本质为决策层面的混合,通过运用各个评分预测结果线性化地组合、设计评价投票的机制或根据评价结果的可信性进行选取组合等手段,在两种或两种以上的推荐算法下对最终产生的评价结果及最终产生情况的数据进行最终的预测和计算生成。

(2) 中融合的本质为模型结构层面上的推荐混合,通过在一种

模型推荐混合算法的模型基础上,将另一种模型推荐混合算法进行模型融合,例如在协同过滤算法的基础上加入基于内容推荐的特征元素,该方法可有效缓解数据稀疏问题。

(3) 前融合在本质上是特征层面的融合,通过将多种推荐算法进行融合并应用到统一的模型中,然后将从各种特征数据中提取出来的特征值作为统一模型的信息进行输入,由统一模型得到推荐的结果。例如,将所有关于用户属性、客户行为等的数据作为信息的输入,通过训练一个统一的分类器来生成推荐结果,此种类别中最著名的技术就是在一系列基于用户和服务的属性中采取降维技术,本质上是对特征层次的融合。

与此同时,在组合方式上,有研究人员提出了以下 7 种组合思路。

(1) 加权型:将多种加权推荐综合算法的分析数据和综合计算结果,进行多次加权分析混合计算后而产生的综合推荐。

(2) 变换型:根据提出的问题背景和实际情况进行实时变化推荐,并采用不同的推荐技术,例如系统首先尝试使用基于内容的推荐与协同过滤推荐相结合,系统首先尝试使用协同过滤推荐算法,若无法形成一个具有高可信度的推荐,则将会进行变换式尝试。

(3) 合并型:同时运用多种类型的合并推荐分析算法,产生基于多种类型的合并推荐算法结果,反馈给目标用户参考使用。

(4) 特征组合:将不同新型推荐分析算法和各种数据源中不同的特征信息分别进行合理组合在一起,由另外一种新型推荐分析算法对其进行组合使用。如将基于内容的信息作为增加的特征向量,然后在增加的数据特征向量集上使用协同过滤推荐技术。

(5) 层叠: 简而言之, 就是通过综合地使用后一种推荐技术算法去优化前一种推荐的方法, 如先通过这种技术算法综合运用来直接生成一个较为简单粗放的推荐结果, 在此基础上再通过综合使用第二种推荐算法技术对其结果作进一步推荐。

(6) 特征扩充: 将前一种技术推荐的应用技术的基本特征输出和各种特征拓展信息分别输入, 以此作为后一种技术推荐应用技术的各种基本特征信息输入, 第二种技术推荐应用技术的各种基本特征输入, 以此拓展信息输入, 包括了第一种技术推荐应用技术所需的可能直接产生的各种技术的基本特征拓展信息。

(7) 元层次型: 使用前一种推荐技术后所形成的模型, 作为后一种推荐技术的输入, 与其他特征扩充不同的地方是, 元层次型中整个模型均以特征为输入, 而其他的特征扩充只能使用该模型后所形成的某些特征作为后一种推荐技术的输出。

4. 基于关联规则的推荐

基于关联规则的推荐以关联规则的理论为依据, 关联规则的分析能够找到不同商品之间在市场上进行销售时的相互关联性。以购物篮为例, 根据用户的历史购物记录, 为用户推荐潜在采购商品, 因此, 该类商品的推荐算法的主要特点之一是, 可根据用户在历史上购买的物品情况进行大规模的数据采集, 将已经被采购的商品作为一个规则标签, 将潜在需要进行推荐的物品和对象作为一个规则体, 在大量的数据收集中来搜索和寻找两种物品之间的一些隐含的相互关系。该算法的关键在于, 找出一个不同用户可能需要同时采购的物品资源数据, 此数据可能由多次活动产生并且不断出现的物品项目集合和活动时间序列, 挖掘不同资源项之

间可能存在的价值关联项,然后基于其已有的价值关联准则向其他用户推荐一些更加感兴趣的物品和资源。也就是说,使用数据挖掘过程中的关联规则挖掘技术,找出两个或多个对象之间的关联性,并以此进行推荐。关于关联规则的推荐一般分为以下两个步骤。

(1) 通过分析用户与资源项之间的历史数据,然后通过 Apriori 算法、FP-Growth 算法等生成所有的频繁项集(Frequent Item Sets)。

(2) 在频繁项集中,通过计算支持度(Support)、置信度(Confidence)、提升度等来提取强关联规则。

但是,关联规则的发现和制订最为关键且非常耗时,关联规则一旦生成便无法自动更新,这成为该算法的瓶颈问题。同时,商品名称的同义性问题,也是关联规则的又一难点。

5. 蚁群算法

蚁群算法是一种群智能算法,最初由一位意大利学者 Colormi A. 和 Dorigo M. 等于 1991 年提出,其模拟蚂蚁自行搜索和自动寻找各类食物的计算过程,能够求出一条从蚂蚁原点觅食开始,经过若干个蚂蚁给定的食物需求点后返回原点的最简易计算路径。它是由一群基本没有智能或者只具备轻微智能的个体(Agent)通过利用彼此之间的智能相互配合和互相协作而表现出智能行为,从而给我们求解一个复杂的数学问题带来了一个全新的数学途径和一种可能。蚂蚁在其自己行走的整个过程中不断地释放出信息素,以便准确标注其行走路径,随着行走时间和距离的推移,此时会有若干蚂蚁找到食物,这样至少存在若干条从蚂蚁洞穴到食物的路

径。在单位时间内,短路径上的蚂蚁总体数量要比长路径上的蚂蚁数量多,因此短路径上积累的信息素浓度也会因此得到响应程度的增高,这为后面的蚂蚁提供了明确的方向指引,选择该路径的蚂蚁数量也会越来越多。最终,蚂蚁群会在正向反馈因素的影响和作用下归集到最短路径上,与该路径相对应的解即待优化问题的最优解。可以使用随机蚁群分布中的一只蚂蚁的具体行走实时路径空间来随机表示待进行优化问题的一个实际可行性数据来理解,整个随机蚁群所有的蚂蚁行走路径便形成了一个待优化问题的实际理解数据空间。

5.3.2 个性化学习路径挖掘结构模型

与常规的推荐系统类似,个性化自适应学习路径推荐除了需解决常见的数据过大和冷启动问题,还要考虑系统中其他模块的输出问题。教育知识图谱是知识领域模型和学习者特征模型的结合,因此,在个性化自适应学习系统的设计中,需融合教育知识图谱和知识追踪两大模块的知识细粒度,构建一个既能解决推荐系统的常规问题,又能通过具备可解释性的神经网络模型体现个性化差异的学习路径推荐框架,如图 5-13 所示。该框架将学习者在学习过程中所呈现的学习效果作为评估指标,若学习效果未达到有关期待值,则系统会基于最新学习数据重新训练和推荐路径。若学习者在计数器控制下于一定预测次数内仍未达到相关期待值,系统则会从优化路径集里直接选择与学习者个性化特征最切合的学习路径作为最终推荐结果,整个运行机制本身是自适应的。

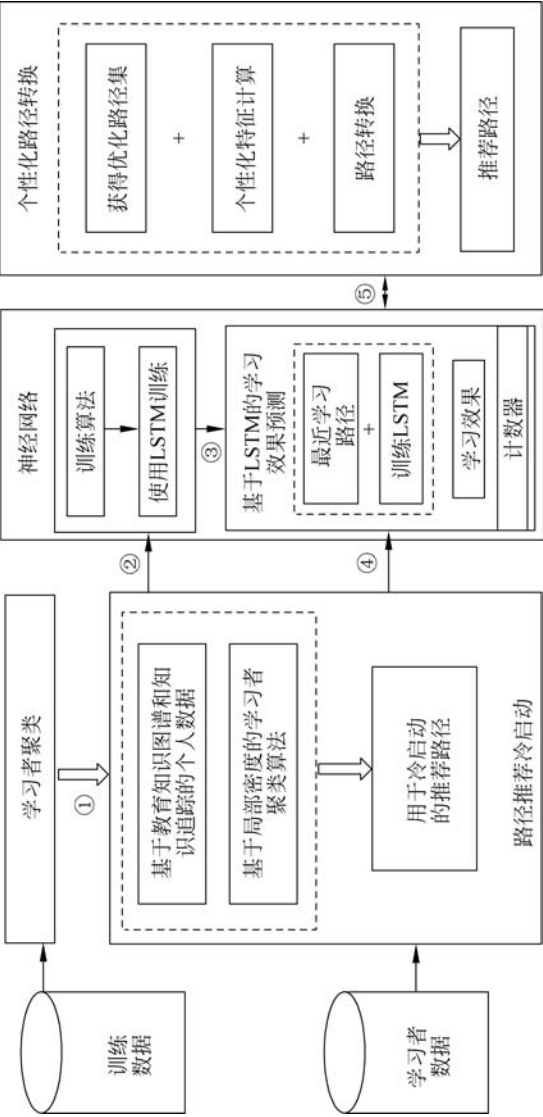


图 5-13 个性化自适应学习路径推荐框架

1. 基于知识图谱的课程推荐

个性化学习路径推荐是一个详细的学习路径推荐,如果仅仅用学生数据进行学习路径推荐,则难度是非常大的,因此在具体实践中,需先对学习者的课程推荐,针对不同的课程序列生成不同的学习路径,又因每个课程均包含众多知识点,因此可构建一个基于知识图谱的可解释推荐模型。在前期的领域知识模型中已完成了教育知识图谱的构建工作,因此可直接从该知识图谱中获知每门课程的信息。在此基础上结合学习者在特征模型中的知识追踪,动态监控学习者对各知识点及各门课程的认知程度,生成学习者个性特征数据。最后利用网络表示方法对课程信息和学习者信息进行嵌入,从而进行课程推荐。因前期的知识追踪是可解释的,所以此环节的课程推荐也具有可解释性。

2. 融合知识关系的学习路径推荐模型

学习路径由不同的节点组成,每个节点都代表着一个知识点,每个课程均包含着或多或少的知识点。通常情况下,领域的知识是教师和学生进行个性化学习所需要提供的的数据资料来源,其能够表征为课程、知识单元和知识点共3种内容的颗粒度。内容之间的关系分别为前提、包含和并列,各个知识单元和知识点都应包含难度、风格与所属任务等属性。知识领域是一个具有整体性的知识层面的数据集实体,知识领域的各个数据集以各种知识单元的组合形式来表现和描述,知识单元由各种知识点数据集共同构成。实体关系用来记录一个知识单位与其他的知识点单位之间、其他的知识点与其他的知识点单位之间的一种逻辑关系。首先,

学习者可以通过在课程的序列图中选择所需要学习的一个知识点,由这些知识的追踪来获得学习者的性格特征。其次,集合知识关系并通过使用深度优先算法对所有学习途径的数据进行深度搜索,估算出每位学习者的所有学习与得分。再次,结合学习者的目标等因素,为学生们推荐出一系列符合学习者有限的时间约束力所要求的学习路径,使之达到最佳的得分效果,从而找到最适宜于每个学习者的具有独特个性化的学习途径。由于课程推荐具备可解释性,同时在推荐过程中亦对每个知识点进行探讨研究,由此实现的学习路径推荐模型同样具备可解释性。

3. 基于局部密度的学习者快速聚类算法

冷启动问题在推荐模型系统中最为普遍,如果一旦没有了学习者的这些学习行为和数据,则后续的神经网络等模型便无法正常工作,因此,在开展个性化自适应学习路径推荐时,需要先对学习者进行聚类,确保每个类别中学习者的先前数据可用,由此可为不同类别生成不同的学习路径,从而相当于给某个类里不能获得学习行为数据的学习者提供一个推荐学习路径。假设有限局部空间内聚类中心的特征密度高于邻居,并且与密度较高点的距离相对较大。这个假设能够直观地计算出一个被嵌入空间的聚类次数,并且能够自动地发现一个异常的值,无须考虑它们的形态及被嵌入空间的维度。由于前面已构建出教育知识图谱,所以可从中获得学习者特征信息,结合知识追踪便可掌握学习者的个人数据,再基于以上假设,通过对每个学习者数据所处的空间密度进行排序后选择聚类中心,由此便可快速完成聚类。

4. 融合知识关系的学习路径推荐模型

学习路径由不同的节点组成,每个节点均是一种具有相关学习者的个体特点和相应的学习信息资源等序列化结构的多维度数据。众所周知,LSTM神经网络具有处理不同长度序列的能力,其在不完整路径预测结果的可靠性方面有良好的性能,因此可采用基于LSTM神经网络融合知识关系的模型进行学习路径推荐,如图5-14所示。知识点关系由知识图谱和知识追踪通过向量嵌入融合后导入 h_t , h_t 是神经元的输出,由遗忘门来控制。 rls_t 表示知识点之间的关系,LSTM无须学习即可直接使用给定的 rls_t ,当效果不好时亦可以舍去 rls_t 。由此,该模型在不同的数据集上均具有很好的准确性和权衡,即使用户可在大规模的数据集上降低训练的时间,又使用户可以通过在小规模的数据集上增强其准确性。由于知识跟踪模型具有很强的可解释性,并且知识之间的关系来自于知识跟踪,因此这个模型也就具备了一定的可解释性。

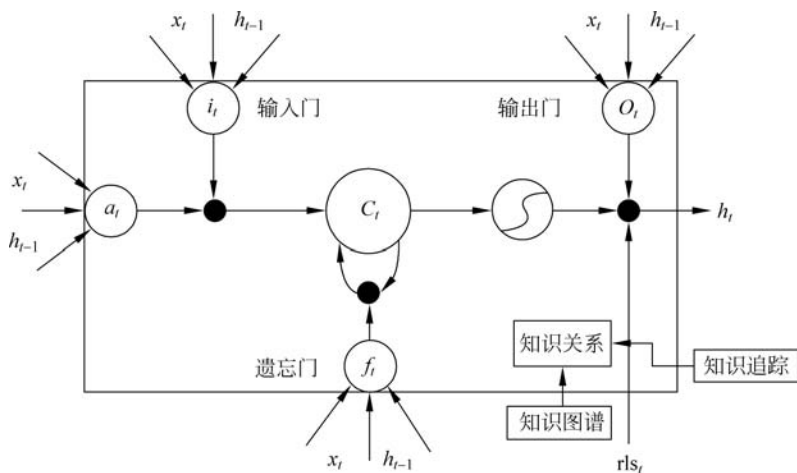


图 5-14 融合知识关系的学习路径推荐模型

5.3.3 未来个性化学习路径推荐研究探讨

传统个性化推荐的主要目标是通过对被推荐用户的相关个人信息和与其所推荐项目的有关资料或者信息进行综合分析,以便及时地获取使用者的消费习惯及其他兴趣点和偏好,从而向其他使用者推荐自己认为感兴趣的个性化推荐项目,主要个性化推荐技术手段包括基于个性化信息内容的自适应过滤式推荐、协同式信息过滤混合推荐、使用关联规则的过滤推荐及混合式过滤推荐。这些应用技术对于现代电子商务与社交网络等各个领域来讲具有很强的技术实用性,但在个性化学习路径推荐方面,相关研究与应用相对较少。学习路径生成与推荐的本质是根据学习者的学习目标和认知状态对所需要学习的知识元信息进行排列的过程,较好的系统性和连续性是个性化知识学习路径序列的基础和要求。相关工作原理为:适应性引擎根据领域知识及其推理,对学习者的特征模型反映出的知识水平、认知能力、学习风格和偏好等加以诊断和分析,做出适应性决策,动态安排高切合度的学习内容、学习资源及其呈现方式,有针对性地向学习者提供个性化推荐服务。同时,在开展学习路径推荐的同时也对学习过程进行实时监测和管理,动态获取学习者的表现数据,由此不断训练、更新学习者的特征模型,然而,现有的推荐方法通常将每次推荐的项目作为独立过程进行考虑,较少关注推荐结果的系统性和连续性,并且缺乏对用户当前知识水平和目标知识体系进行综合分析。如果在自适应学习路径推荐中直接应用上述推荐方法,将面临“推荐产生的知识序列缺乏连续性”和“知识需求分析片面化”等问题,同时特征抽取困难、数据稀疏和冷启动等问题也与之并存。

1. 过去基于本体的推荐系统研究

理想的学习路径生成需要建立在一种学习者已经掌握并具备的认识状态和知识拓扑网络结构的基础上,知识拓扑网络所承载的学习路径更加可靠,因此它们能够较为精准地匹配到每一位学习者对于各种个性化的学习需求。值得注意的是,当涉及自适应学习和个性化学习路径推荐时,没有单一的学习者概要或结构化内容模型,这使对相关研究中对本体的需求更为必要,尤其是当需要系统相互操作时,本体便显得尤为重要。Weng 和 Chang 等提出了一种利用本体和扩散激活的推荐系统,其中本体用于帮助创建用户概要文件,扩散激活被用于理解在学习者网络中其他有影响力的学习者。Pukkhem 等利用 Agent 本体,设计了基于多 Agent 的学习对象推荐系统。Rani 等在 Felder-Silverman 学习风格模型的基础上创建了学习风格本体,并将多智能本体概念用于实现推荐系统相关推荐的个性化。Saleena 和 Srivatsa 利用模糊领域本体和本体专家设计了一个自适应学习系统。Martinez-Cruz 等利用本体和模糊语言建模设计了一个推荐系统,其中使用本体显示用户的满意度。当满意度没有明确表述时,再使用语言模型计算信任因子。Montuschi 等描述了一种基于语义的自适应学习推荐系统,在该推荐系统中,本体被用来表示学习结果。Dascalu 等提出了一个推荐系统来帮助专业社区内的学习者。当学习者通过 LinkedIn 的个人资料登录 Protégé 时,其详细信息就会被提取出来。Labib 等设计了一个本体来提取不同学习风格维度之间的关系,因此使用本体将学习风格与学习者的特征联系起来。Ouf 等在通过提出 4 个本体来代表学习者、学习者对象、学习活动和教学

方法的基础上,开发了一个个性化电子学习资源的框架,并且通过在本体上使用规则来添加个性化维度。Tarus 等采用混合方法提供推荐,在学习者、学习资源上实现本体集成序列模式挖掘,从 Weblog 中识别学习者的历史序列模式。Yanes 等提出了一个基于本体的软件工程推荐系统,该推荐系统使用 COTS 组件,其中本体被用来表示该组件,而用户模型用来表示用户的兴趣。Obeid 等通过为高等教育机构、学生和雇员创建本体来生成个性化推荐,然后使用机器学习技术来聚集类似的学习者,最后将聚类后的学习者数据发送到混合推荐引擎。

许多通过互联网提供的符合标准的学习对象库和开放教育资源库为自使用学习和个性化学习路径推荐提供了强大的支持,然而,学习对象的搜索和选择可能涉及访问各种存储库的固有的复杂操作,每个存储库都可能涉及不同的软件工具,以及学习资源的不同组织形式和规范格式。这种复杂性可能会阻碍自适应学习的成功。跨存储库聚合器,即可在不同存储库中漫游以满足用户/教师查询的系统,可以帮助降低这种复杂性,尽管交付的问题可能仍然存在。Medio 等提出了一个混合推荐系统 MoodleRec 作为 Moodle 学习管理系统的插件。MoodleRec 可以对一组受支持的标准兼容的学习对象库进行排序,并根据一个简单的基于关键字的查询给出学习对象的排序列表,并且各种建议策略在两个层面上运作。首先,创建一个学习对象的排序列表,按照它们与查询的对应关系及它们的质量(由原始存储库表示)进行排序。其次,使用社会生成的功能向教师展示所列出的学习对象是如何在其他课程中被利用的。文中还进行了实际实验研究,并讨论了 MoodleRec 方法的有效性。

Moodle 是世界上最流行的学习管理系统之一,在 222 个国家拥有超过 8000 万用户。在自适应学习系统开发中可以配置一个 Moodle 插件,将其作为一个可以将学习活动添加到课程中的模块,该模块则作为插件嵌入在 Moodle 中的推荐系统运行。它允许教学者:

(1) 发送查询。

(2) 从不同的 LORs 收到一个 LOs 排名列表。

(3) 研究列表中有关课程管理系统的各种资料,例如它们在其他课程中的使用情况,以及同一课程的相似教师如何使用它们。

(4) 从教师正在编辑的课程列表选择一个或多个 LOs。

该模块根据以下标准赋予检索资料的特权:

(1) LO 元数据和查询的关键字之间的对应关系。

(2) 在原产地主中授予原产地主的等级(如有)。

(3) 在同一 LMS 的其他教师所创建的课程中实际使用该 LO。

教育资源建设的一大挑战是为学习者提供智能化、个性化的资源推荐。Wu 等提出了基于语义网和教育学的教育资源语义推荐框架。在该框架中,可构建领域本体来描述领域的知识结构,采用本体技术和资源描述框架对所有资源和用户组合进行描述,支持语义推理。Wu 等在语义资源组织的基础上,制定了一套基于教学法的推理规则。这些规则是由知识的类型、知识的内部结构和学习者的学习表现综合而成的。以《数据库理论与实践》课程为例进行了实例研究。在这种情况下,根据学习者不同的知识结构和不同的学习成绩,推荐不同的学习材料。提出了 3 种典型的学习模式来描述个性化的学习体验。这个框架可以作为教师和资源设

计者的指南。

2. 基于知识图谱的个性化学习路径推荐

然而,基于本体的推荐系统在自适应学习中仍存在一定的局限性,例如创建本体耗时,或者在无标准数据集时,对基于本体的推荐系统的评估则会变得尤为困难,再如使用本体需要知识工程方面的技能等。由于自适应学习系统完全依赖于唯一输入本体的覆盖程度和细节,因此导致其中基于本体的推荐系统只依赖于本体知识,这成为相关推荐的一大缺点。同时,在传统的推荐系统中往往存在着数据稀疏和冷启动等问题,基于这些限制,近年来将知识图谱作为一种边信息导向推荐系统,吸引了业界研究者的广泛兴趣,知识图谱中通过其自身包含的大量信息刚好可以解决以上问题。随着链接数据计划的日益普及,许多公共的知识图谱已被广泛使用,如 Freebase、DBpedia、Wikidata 和 YAGO 等,它们是新型的语义网络,记录了数百万个概念、实体及其关系。

知识图谱以实体及它们之间的关系对事实信息进行建模,从语义上表示世界各事物之间的联系。从技术上讲,知识图谱是一种图,不同类型的实体作为定点,实体之间的各种关系作为边。这里的实体是现实生活中真实的物体,或者人类头脑中的抽象概念。关系描述着两个实体之间的关系,两个实体及其之间的关系构成了知识图谱的一个三元组。知识图谱通常包含着数十亿个顶点、多类型和三元组。形式上知识图谱可以表示为 $KG=(E,R)$,其中 E 和 R 分别表示实体集和关系集。知识图谱中每个事实存储为一个三元组 (s,r,o) ,其中 $s \in E, o \in E$ 分别为主体和对象实体, $r \in R$ 则表示它们之间的关系。从理论上来讲,知识库中丰富的结构化

知识有助于人工智能各个领域的发展,包括但不限于决策、信息检索、问答、词义消歧、文本分类(或聚类)和自然语言理解等。

研究发现,知识图谱作为一种异构网络,将项目及其属性映射其中,可以了解项目及其之间的关系。亦可将用户及用户端信息集成于其中,从而更精准地捕捉用户和项目之间的关系及获悉用户的偏好,以及将知识图谱作为一种教育资源和学习路径推荐的新方法,可以解决自适应学习中基于本体推荐系统的不足。除能够给人提供一种更加精确的推荐之外,基于知识点图谱的推荐结果还可以给人们带来额外的多元化、可解释的意义。基于知识图谱的推荐方法是完全自动化的,无须预定义的语料库或领域知识本体,可以独立于自适应学习推荐系统中的领域应用。有关研究表明,理想的学习路径的生成需要建立在学习者已经具备认知状态及知识拓扑结构的基础上,而知识图谱所承载的网络学习路径能更加精准地匹配每一位学习者的各种个性化学习需要。使用知识图谱交互式推荐方法的优势主要可以体现为两大方面,一是该方法不涉及预先构建的信息地图,可以自动从现有的信息系统中获取地图信息;二是由于学习者和推荐系统之间存在一些(如规则生成等)交互行为,相关的交互式决策方法为推荐结果提供了可解释性。由于信息地图中的许多推荐参数本质上是个性化和主观的,因此对于每个推荐参数实际上并不存在一个近似元素。由此,此种交互式推荐方法更适用于现实教学过程中的应用程序。

通过知识图谱进行学习路径推荐的算法主要包括基于嵌入的方法、基于连接的方法和基于传播的方法,这3种方法的介绍如下:一是基于嵌入的方法,其主要利用知识图谱中的丰富语义内容进行学习路径的推荐;二是基于连接的方法,其主要通过知识图谱中

的实体连接方式进行推荐；三是基于传播的方法，其主要通过知识图谱中的连接路径进行推荐，然而，这种推荐方法仅对交互数据的特征进行建模，忽略了交互数据间的关系，从而无法实现从在线学习者的集体学习行为中提取基于属性内容的协同信息。随着推荐系统中对精读要求的日益增高，算法中对知识图谱的各个节点之间的复杂依赖关系的要求也越来越高，因此，具有复杂依赖关系的节点也会随着阶数的增加而相应增加。从另一方面来看，具有复杂依赖关系的节点的增加将给推荐模型带来巨大的计算压力，同时这些复杂的依赖关系也存在着对预测结果不均衡的问题。基于此，为解决以上问题，一般考虑在基于知识图谱推荐的基础上，采用基于邻居节点嵌入来更新相关节点嵌入及采取注意力机制学习传播期间每个邻居节点权重的方法，加以改善。

基于知识图谱的推荐系统受益于知识图谱中所蕴含的大量实体节点、边等信息，因此以上方法可以对相关推荐系统所产生的推荐结果的准确性产生有利影响，并且可增加相关推荐结果的可解释性。在实际教学应用中，利用基于知识图谱的推荐算法对学习者的学习路径进行个性化推荐（如测试试题推荐等）则显得尤为重要。类比传统的商业领域的推荐系统的相关研究，学者们将使用自适应学习系统中的测试试题推荐功能的学习者类比成商业领域推荐的用户，将测试试题类比为可供推荐的商品，学习者在学习过程中的测验得分则类比为用户兴趣偏好倾向的评分。学者们往往认为测试试题推荐考虑的是一组显式知识点之间的关联，也就是说所推荐的相关测试试题考察的知识点间是一组给定的且有关联的，学习者在测试评分过程中会得到所学知识点的学习掌握情况。由于知识点之间的关联存在着一定的隐秘性而无法直接观测，往

往将影响推荐效果,然而,知识图谱可以很好地挖掘知识点之间存在的隐性关联,在未来关于个性化学习路径的推荐研究中,可着眼于使用知识图谱技术对测试试题进行精准推荐。如前述的Knewton公司也开始尝试使用知识图谱技术对相应的学习内容和测试试题做个性化精准推荐,为后续知识图谱技术的全面应用做出探索和打下基础。

尽管当前学界关于基于知识图谱的推荐技术已有较多研究成果,但在自适应学习领域中的关于多任务学习、跨领域推荐和动态推荐等,也是基于知识图谱的推荐中亟须解决的技术问题。同时,在测试试题推荐的过程中,推荐系统不仅要保证推荐结果的精准性,还需要实现推荐过程的可解释性,并且由于在自适应学习中,大多数学习资源、测试试题的推荐均是在真实教学场景下的推荐,除需解决静态推荐问题外,还需考虑学习者与知识点之间的频繁交互问题。同时,除了对显性知识进行推荐,对知识点间隐藏的隐性内容进行推荐也是非常必要的。教育是一个极其复杂的事件组合,个性化学习路径推荐的进一步丰富和发展需回归教育本质。除了学习目标、任务、内容、资源和结果等实体信息,与过程相关的事件语义也是精准学习路径推荐所关注与表达的核心。未来的研究,可借助于事件知识图谱所具备的描述动态事件及其逻辑关系的功能,能够对抽取的事件、实体及其关系进行有效融合,并且包含大量的事件语义信息,更具丰富准确的语义表示能力等优势,从教育的角度出发,将事件知识图谱作为对在线学习者知识与技能真实掌握状况的底层依托,将学习路径的个性化推荐研究与教育全过程相结合,在教育教学理论的指导下开展推荐模型的构建和系统的研发。其中,如何对海量在线学习数据和事件实体进行采

集、抽取及其融合关联分析,如何对事件知识图谱中教学事件及其逻辑关系进行特征提取及可解释表示,如何追踪学习者在学习过程中不断变化的知识能力水平和掌握其学习能力及其学习状态,如何增加推荐过程的可解释性,如何利用细粒度知识图谱和精准学习者画像以便更精准地进行个性化推荐,以及如何对学生的教学过程、学习行为和学习结果进行认知推理和个性化学习路径推荐,将成为未来研究亟须解决的重难点问题。

5.3.4 个性化学习路径推荐典型案例分析

1. 基于多重智能算法的个性化学习路径推荐模型

学者申云凤于2019年3月提出基于多重智能算法的个性化学习路径推荐模型,该模型通过人工神经网络算法、蚁群算法等,搭建了相似的智能学习过程用户路径模型和框架,实现了个性化的学习路径推荐。相似学习用户模型可以通过对学习者的学习风格和学业水平相似性等因素进行综合计算和大规模数据分析等实现和获取,个性化学习路径推荐则可以对相似学习用户学习模块进行向量映射、网格数据聚类和高密度集聚类等,在计算学习路径并获得协同过滤推荐 TopN-1 的基础上再通过蚂蚁群算法实现。合并概率最佳和可以优化的协同推荐学习路径 TopN-2,有序地处理合并,最后推荐给学习用户,如图 5-15 所示。

学习行为预示着在线用户的学习样式、学习风格和其他学习层次及水平等各种个体差别的特点,它们都是实现个性化路径推荐的基础和关键,因此,相似学习用户模块的建构包含 3 个步骤:第一,对在线用户的学习活动和学习行为进行量化;第二,学习风

格分类预定阈值的判定与计算；第三,构建相似学习用户模型。个性化的学习路径推荐模块的主要任务之一就是通过计算学习路径,形成学习路径推荐列表。

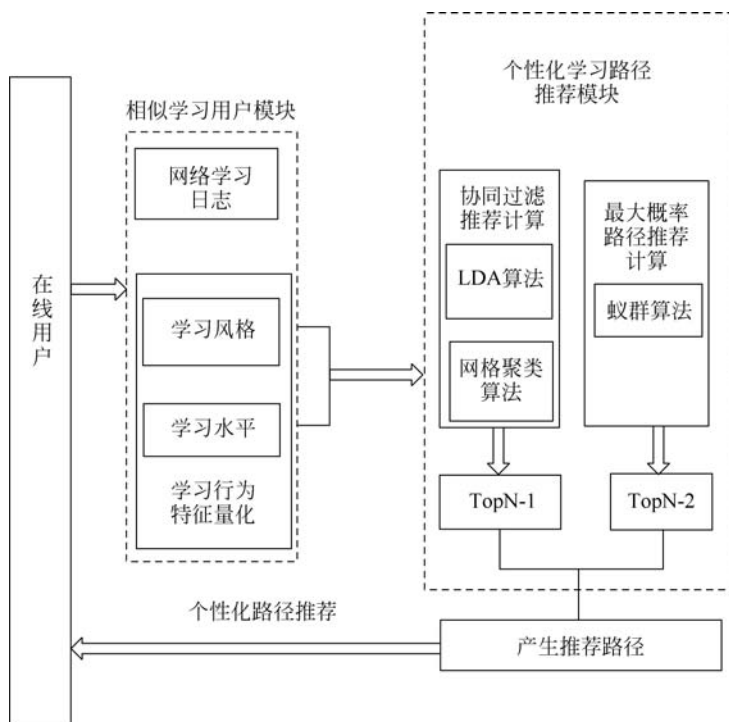


图 5-15 个性化学习路径推荐模型

学习路径计算在根据相似学习者模型所建立的协同过滤推荐的学习者路径的基础上,通过蚁群算法搜索并得到概率最大化的学习者路径,对其关联计算后进行个性化学习路径推荐。鉴于不同在线学习者的在线学习技术水平和认知风格等存在差异,因此先利用神经网络算法对在线学习者的在线学习行为特征进行

分析,建立具有相似在线学习特点的用户模型,在此基础上对相似用户在线学习行为数据进行网格聚类以便于获取其学习的路径,再进行协同过滤和推荐。同时,混合蚁群算法以有效减少个体差异,弥补协同过滤算法的不足。由于所有的学习路径兼具连续性和序列的知识项目集合,因此我们需要对路径的推荐方式进行项目之间的关联程度的计算。

2. 基于 AprioriAll 算法的个性化学习路径挖掘

学者姜强于 2018 年提出了基于 AprioriAll 算法的个性化学习路径挖掘模型,该模型通过对学习者的整个学习过程的数据的采集分析,全方位记录、追溯和准确掌握学习者的学习特征、需要、习惯和行为,最后通过 AprioriAll 算法挖掘学习者基于学习样式风格的同一簇群体学习的个性化的信息,生成准确的个性化的学习活动序列,如图 5-16 所示。相关流程为:一是学习风格判定,利用所罗门学习风格量表、Myers-Briggs 风格量表、Felder-Silverman 学习风格模型等问卷调查的显性主观判定、贝叶斯网络分析挖掘学习风格的行为模型(如查看学习材料的类型、学习时间、浏览次数及参与论坛讨论发帖量、读帖量等)的隐性方法推测学习风格,并通过这两种研究方法的相互结合,实现对于一种个性化的学习思维路径的精准推送。二是知识水平估测,一方面利用项目反应理论中的 Logistic 模型、等级反应模型和布鲁姆教学目标分类理论,综合检测学习者各个知识点所掌握的状况与目标的测试、练习难度的分布,实现了对学习者各个知识点在概念上的水平估计;另一方面,充分运用人工智能算法,如矩阵分解算法、隐马尔科夫模型等,依据学习者进行行为数据(其中包括案例学习的时间、数量

与观察单击次数、问题解答的持续时间与尝试次数等)的实时跟踪和判断,提高了学习者的概念层次,并通过对概念性知识的理解层次和困境层次两个维度来动态地呈现所需要进行学习的物体。三是学习路径挖掘及个性化推荐,利用 AprioriAll 关联规则算法,从每个群体的学习行为中发掘最优的学习路径,同时基于每个人学习的特征风格、知识技能水平等属性进行了个性化的挖掘和推送,解决了“学习迷航”“认知过载”等问题,提升了学习的内驱动力和学习要求。

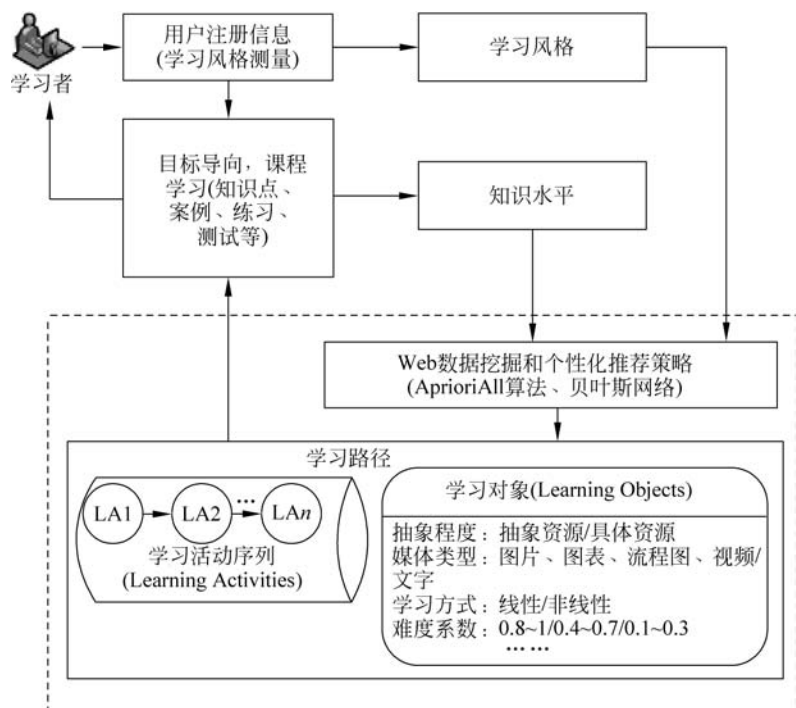


图 5-16 个性化学习路径挖掘结构模型

其中,利用 AprioriAll 算法挖掘学习行为数据,以便自动生成精准的个性化学习行为活动数据序列,整个数据挖掘过程主要包括排序阶段、大项目集阶段、转换阶段、序列阶段及选最大阶段等 5 个阶段,具体如下。

(1) 排序阶段。首先根据 Felder-Silverman 学习风格模型对学习者的分类,其次以学习者账号为主关键字、以学习行为时间为次关键字,将原始行为数据库转换成学习者序列数据库。

(2) 大项目集阶段。运用关联规则进行挖掘算法,找出了所有频繁的项集。

(3) 转换阶段。在对于学习序列模型进行寻找的过程中,不断地检测给定的大项目集是否被包含在学习序列中。

(4) 序列阶段。利用核心算法寻找频繁序列。

(5) 选最大阶段。通过修剪方法在大序列集中找出极大序列,减少冗余活动序列的出现。

此外,个性化学习路径推荐还需综合考虑每个领域各类学习课程内容的各种重点和难度综合系数,如以各类学习者的综合认知能力水平等级和所学领域专业知识难度级别相匹配进行计算作为依据,按照各个学习章节、知识点、难易综合程度、考试成绩比重等多种综合属性因素进行综合精确定位,再根据每个学生每次做题时的回答准确率(其中包括学生回答某个重要问题的时间及所需学习的时间)等,持续地对每个领域的学习者的专业知识水平情况进行综合评估,而后再做出适应性的智能化推荐。例如在测试环节,当诊断出学习者的专业知识能力水平和相关专业知识技术水平的等级级别属于高级时,一些相关测试难度较大的测试实体,如果通过测试,则标注学习者已完全掌握了相关的知识实体,同时

默认其也掌握了其他难度中、低层次的实体知识。当学习者的专业知识能力水平和相关专业知识技术水平等级诊断为中等级别时,系统自动推送中等难度试题。若通过推荐测试,则系统标记已经完全掌握了与此测试相关的专业知识,并适应性地推送高难度试题进行深度测试,否则推送较低知识层次的试题进行综合测试,以此类推。

5.3.5 小结

本节在对常用推荐算法进行概述的基础上,分别从基于知识图谱的课程推荐、融合知识关系的学习路径推荐模型、基于局部密度的学习者快速聚类算法和融合知识关系的学习路径推荐模型四大方面对个性化学习路径挖掘结构模型进行分析,并对未来个性化学习路径推荐研究进行探讨。最后介绍了基于多重智能算法的个性化学习路径推荐模型和基于 AprioriAll 算法的个性化学习路径挖掘两个典型案例,对个性化学习路径推荐进行进一步剖析。

5.3.6 思考与练习

- (1) 什么是学习风格?请简述学习风格的种类,并对不同的学习风格在个性化的学习路径中所推荐的影响进行分析。
- (2) 在个性化学习路径推荐逻辑中,如何兼顾“学习差异性”?