

◎热点与综述◎

个性化学习中学科知识图谱构建与应用综述

赵宇博, 张丽萍, 闫 盛, 侯 敏, 高 茂

内蒙古师范大学 计算机科学技术学院, 呼和浩特 010022

摘 要: 学科知识图谱是依赖大数据、人工智能等技术构建的支持教学活动的重要工具, 作为一种学科知识语义网络, 能够助力个性化学习体系的发展并促进数字教育资源新基建。对知识图谱的概念、分类等内容进行概述; 总结了学科知识图谱的概念、特点、优势、内涵及其对个性化学习的支持等内容; 重点梳理了学科知识图谱的构建流程: 学科本体构建、学科知识抽取、学科知识融合以及学科知识加工, 并介绍了学科知识图谱在个性化学习情境中的应用及其面临的挑战; 展望了学科知识图谱以及个性化学习的未来趋势, 为教育资源的组织方式及个性化学习的创新发展提供借鉴和启示。

关键词: 个性化学习; 学科知识图谱; 知识图谱构建技术

文献标志码: A **中图分类号:** TP18 **doi:** 10.3778/j.issn.1002-8331.2209-0345

Construction and Application of Discipline Knowledge Graph in Personalized Learning

ZHAO Yubo, ZHANG Liping, YAN Sheng, HOU Min, GAO Mao

College of Computer Science and Technology, Inner Mongolia Normal University, Hohhot 010022, China

Abstract: The discipline knowledge graph is an important tool to support teaching activities based on big data, artificial intelligence and other technologies. As a kind of discipline knowledge semantic network, it contributes to the development of personalized learning systems and the promotion of new infrastructure for digital education resources. Firstly, this paper outlines the concept and classification of knowledge graph. Secondly, this paper summarizes the concept, characteristics, advantages, connotation and the support for personalized learning of discipline knowledge graph. Next, this paper focuses on the sorting of construction process of discipline knowledge graph: discipline ontology construction, discipline knowledge extraction, discipline knowledge fusion and discipline knowledge processing, and it also introduces the application of discipline knowledge graph in personalized learning situations and the challenges. Finally, this paper prospects the future tendency of discipline knowledge graph and personalized learning. It provides the reference and inspiration for the organization of educational resources and the innovative development of personalized learning.

Key words: personalized learning; discipline knowledge graph; knowledge graph construction technology

在大数据、人工智能等技术迅速发展的时代背景下, 互联网变革成为大势所趋, Web 技术在此次技术变革中占据了主导地位, 正朝着语义网络^[1]的方向演变。语义网络对现有 Web 增加了语义支持, 对知识的表述更

加直观, 可以为用户提供更加便捷的服务, 提高人机协同工作的效率。2012 年 5 月, 谷歌公司提出并公布了知识图谱^[2]这一概念, 知识图谱的诞生最初是为了完善语义网络, 构建新一代智能化搜索引擎。随着人工智能技

基金项目: 国家自然科学基金(61462071); 内蒙古自然科学基金(2021LHMS06012); 教育部产学合作育人项目(2020 年); 内蒙古自治区高等学校科学研究项目(NJZY21560); 内蒙古哲学社会科学研究专项(ZSZX21097, ZSZX21102); 内蒙古自治区研究生科学基金创新项目(S20210278Z); 内蒙古师范大学研究生创新基金(CXJJS21157, CXJJS22137); 内蒙古师范大学基本科研业务费专项资金(2022JBXC018)。

作者简介: 赵宇博(1999—), 男, 硕士研究生, CCF 学生会员, 研究方向为计算机教育应用, E-mail: 1980977735@qq.com; 张丽萍(1974—), 女, 硕士, 教授, CCF 会员, 研究方向为软件工程、智能教育; 闫盛(1984—), 男, 硕士, 讲师, CCF 会员, 研究方向为计算机应用; 侯敏(1973—), 女, 硕士, 副教授, CCF 会员, 研究方向为软件分析、智能教育; 高茂(1997—), 男, 硕士研究生, 研究方向为计算机教育应用。

收稿日期: 2022-09-22 **修回日期:** 2023-01-12 **文章编号:** 1002-8331(2023)10-0001-21

术的愈加成熟和社会对高新技术的需求呈爆炸式增长,知识图谱作为热点技术逐渐被应用到各个领域,为人类知识的建模提供了新的思路,在各个行业都展现出了巨大的潜力。

由于人工智能等技术在教育领域的出色表现,教育模式的突破迎来了新的契机,个性化学习的相关研究取得了显著的成果^[3]。国务院在《新一代人工智能发展规划》中明确指出:“要重点突破知识图谱构建与学习、知识演化与推理等技术,要实现智能教育,建立以学习者为中心的教育环境^[4]”。个性化学习的教育模式是以学习者为中心,根据学习者认知状态、知识维度以及个体特征等提供最贴切的学习资源^[5]。学科知识图谱可以为个性化学习模式的形成与发展提供丰富的理论基础和坚实的技术支撑^[6]。它根据学科知识体系以及知识结构进行构建,目的是细粒度地划分知识点并建立知识点之间丰富的关联关系,从而为学习者提供丰富的个性化服务,如认知状态诊断、学科知识智能问答以及学习资源推荐等。基于此,本文在介绍知识图谱概述的基础上,阐述了学科知识图谱的概念、特点、优势、内涵及其对个性化学习的支持,总结了学科知识图谱的构建方法及其在个性化学习中的应用,指出了学科知识图谱应用在个性化学习情境中存在的问题与挑战,期望借助学科知识图谱促进人工智能技术与教育体系的深度融合,从而推动个性化学习的创新发展。

1 知识图谱概述

知识图谱作为人工智能领域中重要的研究内容,在知识组织方面具有极大优势。它主要用于从类型不同的庞杂数据中抽取实体、实体间的关系以及实体的属性,然后将三者通过图结构的数据模型或拓扑进行整合,从而将抽象零散的知识进行清晰的可视化展示。

知识图谱的形式化定义为 $G=\{E, R, F\}$, 其中 E 表示实体集合, R 表示关系集合, F 表示事实集合。 F 中的每一个事实元素都通过三元组 (h, r, t) 进行表示,其中 h 和 t 分别表示头部实体和尾部实体, r 表示头部实体 h 和尾部实体 t 之间的关系^[7]。知识图谱从逻辑上被划分为数据层和模式层。在数据层中,一系列三元组数据被存储在图数据库中,这些三元组数据组成了复杂的实体关系网络。模式层是知识图谱的核心,被用来规范数据层中三元组的具体形式。通常采用本体库对模式层进行管理,本体库是知识图谱的模具,详细描述了实体的类型以及实体间关系的类型。

知识图谱的构建方式分为自顶向下构建和自底向上构建^[8]。自顶向下的构建方式需要依托于高质量结构化数据,通过本体学习细化概念和关系,通过知识抽取等相关技术实现实体匹配和模式层完善,从而完成知识图谱的构建;自底向上的构建方式是从公开的非结构化

数据、半结构化数据中抽取出实体、关系以及属性信息,加入到知识图谱的数据层,通过本体对数据层的知识要素进行概念抽象形成知识图谱的模式层,最终完成知识图谱的构建。

知识图谱按照知识覆盖范围和知识挖掘深度的不同,可划分为通用知识图谱和领域知识图谱^[9]。通用知识图谱面向通用领域,覆盖的知识范围较为宽泛,但是对于知识挖掘的深度尚浅,以常识性知识为主。因为通用知识图谱涉及到大量数据,在构建初期,需要根据数据内容提炼特征,所以一般采用自底向上的构建方式。领域知识图谱面向的是固定领域,其特点是知识挖掘深度较深,对数据质量的要求较高,因此一般采用自顶向下的方式构建,从而更清晰地组织行业内数据,保证数据的权威性。

当前,国外具有代表性的通用知识图谱主要有 DBpedia^[10]、YAGO^[11]、Freebase^[12]、Probase^[13]等。国内代表性的通用知识图谱主要有搜狗知立方^[14]、百度知心^[15]、Zhishi.me^[16]、CN-DBpedia^[17]等。上述通用知识图谱的具体描述如表1所示。

表1 通用知识图谱描述

Table 1 Universal knowledge graph description

通用知识图谱	规模	构建特点
DBpedia ^[10]	2 800 多万个实体	信息来自维基百科
YAGO ^[11]	1 000 万个实体	属性信息描述详细
Freebase ^[12]	19 亿条三元组	自顶向下构建
Probase ^[13]	1 000 万个概念	包含概念最多
搜狗知立方 ^[14]	整合海量碎片信息	可理解用户搜索意图
百度知心 ^[15]	包含多领域知识	侧重于深度搜索
Zhishi.me ^[16]	2 000 多万个实体	具有用户查询终端
CN-DBpedia ^[17]	1 600 多万个实体	数据质量高

相对于通用知识图谱来说,领域知识图谱的知识覆盖范围十分有限,但是其更注重知识挖掘的深度,能够覆盖细粒度的知识。例如在电商领域,阿里巴巴构造的电商知识图谱^[18]是在已有的商品数据的基础上,与外部行业数据进行深度融合,从而实现知识图谱的扩展。医疗是高度知识密集型的行业,医疗信息的整理与归纳在医疗决策及医院管理中发挥着极其关键的作用,是医疗智能化革新的基石。国内构建的中医药知识图谱^[19]通过融合医学的基本知识、临床医疗数据、中医药领域文献等多种数据来源,使得知识结构更加深入。卫宁健康通过对知识图谱的构建,实现了从用户的角度适时地、合规地、可靠地传达医学知识,并成功应用于临床实践中^[20]。在金融领域,刘政昊等人^[21]通过收集股票基本数据、概念数据、股东持股数据以及价格数据,构造了知识关联视角下的金融证券知识图谱,从而建立了各类股票知识之间的关系,综合分析了股票的关联与未来可能的走势。在能源安全方面,国内研究者通过构建能源安全知识图谱,分析了国内外能源安全的研究现状、研究热

点以及前沿趋势,为能源安全问题、能源安全概念以及评价体系指标的构建提供了参考与借鉴^[22]。除上述领域外,人类生活中用到的许多人工智能产品的研发也采用了知识图谱这一技术,例如小米公司的小爱同学、百度公司的小度助手以及阿里巴巴的智能客服等。在未来,基于知识图谱的人工智能产品将会从智能搜索、智能决策以及智能推荐等更多方面为人类提供全方位、多层次的服务。

人类通过知识图谱可以对复杂的知识进行建模并获取有价值的结构化信息。随着知识图谱相关技术的发展与成熟,机器将具备更加完善的认知能力,知识图谱已经成为“感知智能”向“认知智能”过渡的重要战略手段和技术支撑。

2 学科知识图谱概述

教育学科的分类是十分庞杂的,表示了人类的知识在不同领域的具体应用^[23]。通过构建学科知识图谱,可以将零散的学科知识根据知识结构和内在特征形成统一的知识表示,并为学习者提供丰富的个性化服务,有助于学习者对知识进行更深层次的理解和运用。

学科知识图谱从数字教育资源新基建场域及多学科教学境脉的双重背景中衍生而来,是教育领域中支持教学活动的重要工具。学科知识图谱的构建是通过对学科知识进行实体和关系等信息的抽取,从而形成认知支架,并对不同标签的数字教育资源进行组织和管理^[24]。相对于其他领域知识图谱而言,学科知识图谱不仅能够对知识进行有效组织,还具有以下特点:首先,学科知识图谱的节点具有多样性。学科知识图谱的节点由知识点和相关学习资源组成,知识点可以为公式、概念或者命题等,学习资源可以是课程视频、教学课件以及习题等;其次,学科知识图谱节点之间的关系具有特殊性。学科知识图谱中节点之间具有包含、顺序、相关以及先修后继等多种关系,这些关系和节点共同组成了学科知识的脉络;最后,学科知识图谱对数据质量的要求极为严格。学科知识图谱最终是一个面向学习者的知识库,为保证教育过程的严谨和公平,其包含的内容必须准确无误,需要由学科专家和知识工程师进行严格的质量审核^[25]。

学科知识图谱作为人工智能技术与教育教学深度融合的产物,可以将学科知识资源进行有序组织和优化,进而映射学习者的思维模式和高阶认知活动。与传统的学科知识组织方式相比较(如教材、视频、学习网站等),通过学科知识图谱对学科知识进行组织不仅能够实现知识资源在语义层面的聚合,还能使学习者更好地从宏观上把握知识脉络。两种方式的对比如表2所示。

表2 学科知识组织方式的比较

Table 2 Comparison of discipline knowledge organization methods

学科知识组织方式	优势	劣势
学科知识图谱	(1)对学科资源的组织能力强 (2)可在个性化学习中发挥作用	需要强大的技术支持
传统的学科知识组织方式	(1)组织方式简单 (2)符合学习者传统的学习习惯	在个性化学习中难以发挥作用

学科知识图谱遵循结构化信息展示的应用范式,应用数学、学科教学、人工智能、教育资源组织以及教学策略组织等多种理论与技术相结合,共同构成了学科知识图谱的逻辑结构,其内部知识实体语义关系的建模具有高度结构化、集成化以及动态性强的特点。但是,现有大部分研究尚未关注融合多模态资源的学科知识图谱构建,致使知识资源形式过于单一,数字教育资源智能组织框架中“信息茧房”问题逐渐显现^[26]。因此,构建学科知识图谱的数据来源应该是多种的,在获取到相关文本资源的基础上,还可以通过学习网站、实践训练平台、优质公开课、问答社区等渠道获取多模态的学科知识资源(文本、图片、音频以及视频等)。融合了多模态资源的学科知识图谱对教师的学科教学素养提升和学习者的个性化深度意义建构具有重要作用,可应用于学科资源的智能组织、学习活动中个性化问题解决、学科教学设计与评价等,其逻辑架构与内涵如图1所示。

学科知识图谱的构建应该遵循以实现学科知识的认知智能为导向,以解决教学活动中的痛点问题为目标的原则。目前,许多研究者从这一原则出发,结合教学活动中的痛点问题对学科知识图谱进行了深入研究与探索。例如,在数学学科下,有研究者针对国内现有学科知识图谱数据来源单一的问题,提出了一种多源异构

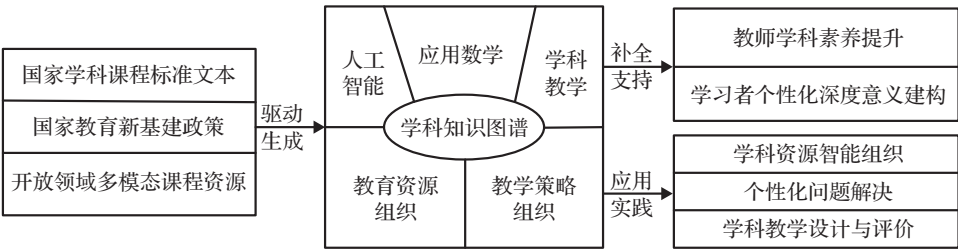


图1 学科知识图谱逻辑架构与内涵

Fig.1 Logical framework and connotation of discipline knowledge graph

数据融合的方法来实现数学学科知识图谱的构建,并基于该方法实现了知识问答和自动推荐系统,为破解数学教学个性化不足提供了思路^[27]。在物理学学科下,有研究者构建了基于物理学知识图谱的学习者画像模型并应用于教学实践中,有效帮助了学习者进行知识体系的构建,促进了规模化教学与个性化学习的深度融合^[28]。在医学学科下,杨云飞等人^[29]在人工智能技术飞速发展和医学知识等数据资源大规模增长的背景下,提出了面向医学知识图谱的多视图、交互式的可视化构建方法,从而使学习者更加有效地对知识的结构进行分析与理解,提升了医学知识在学习层面与应用层面的转化效率。除上述学科外,还有研究者构建了面向K-12教育的课程知识图谱,并将其成功应用于教学实践中。实验表明,该知识图谱能够帮助学习者对知识进行更深层次的理解,同时也为破解当前教育模式个性化不足问题提供了新思路^[30]。

随着人工智能技术和教育模式的发展与改革,学科知识图谱被更加重视,亟需在社会服务等活动中发挥更好的作用。作为教育新基建视域下的知识资源组织工具,学科知识图谱不仅可以有效组织学科资源,实现学科知识的结构化表示,还可以在在一定程度上打破现有的学科壁垒,促进学科之间的交叉融合和个性化学习的发展,进而实现优质学习资源的共享,有助于形成良好的教育生态。

3 学科知识图谱对个性化学习的支持分析

随着现代教育技术的充分运用,教育的个体价值与社会功能被进一步强化,知识的本质、教师的教学方式以及教育的模式正在被重新塑造^[31]。教育模式的改变使得传统的以教师为中心的学习模式被颠覆,如何激励学习者在学习过程中发挥主观能动性成为了教育发展的主要任务。个性化学习是一种针对学习者个体差异而开展的满足其个性化需求的学习方式^[32]。如今的个性化学习以满足学习者的个性化需求为服务方向^[33],旨在激励学习者发挥自身潜能,培养智慧型学习者^[34]。

个性化学习的内涵可以总结为以下三方面:第一,个性化学习要求教育者针对个体特征不尽相同的学习

者采取相应的教育技术和方法。教育者需要在承认并接受学习者在各方面存在差异的前提下,同时做到“因材施教”和“有教无类”,帮助学习者充分挖掘和发展个性,促使学习者在各方面自由、平衡地发展。第二,个性化学习过程是学习者的个性体现和学习潜力被挖掘的过程。个性化学习模式的设计可以采用“学习者驱动,自下而上”的方式。传统的自下而上的学习模式更多地是通过教师驱动,在整体环境中缺乏对学习者的强调,会在一定程度上限制其有效性。采用“学习者驱动,自下而上”的学习模式可以清晰地体现并记录学习者的学习活动和学习习惯^[35],充分挖掘学习者的学习潜力。第三,个性化学习揭示了目前教育模式存在的弊端。目前的教育模式的本质是教育者通过相关教育手段使学习者成为教育者所期望的人,这与学习的本质相悖,学习的本质是使学习者通过学习成为自己所期望的人。个性化学习的发展使得教育的壁垒得以突破,教育思想与教育模式的改变更加契合个性化学习的过程性和反思性,更加尊重人的独立性与主动性^[36]。

学科知识图谱作为融合全面的学科知识点与教学资源的知识库,不仅能够在教育数据分析、学科资料的整理归纳、学习者画像建模、学科知识问答系统实现以及学习资源的个性化推荐等方面发挥重要的基础性作用,还能够促进学习者由简单的知识学习向思维发展进行转变^[37],是促进个性化学习发展的重要手段。学科知识图谱对个性化学习的支持主要体现在以下几方面,如图2所示。

(1) 学科知识图谱为个性化学习资源库建设提供支撑

个性化学习资源库是教育变革的重要依托,是教育智慧沉淀、分享的载体^[38],还是个性化学习内涵第一要素——“因材施教,有教无类”的强有力支撑。近年来,随着教育资源的爆炸式增长,在命名实体识别、关系抽取等自然语言处理技术的支持下,越来越多的多模态教育资源被融合在学科知识图谱中^[39],学科知识图谱的自动化构建程度和构建质量也正在逐步提升。依托高质量的学科知识图谱进行个性化学习资源库的建设,可以设计并实现针对不同学习者的学习资源提供方案,进而促进优质教育资源的共享,形成良好的个性化学习生态。

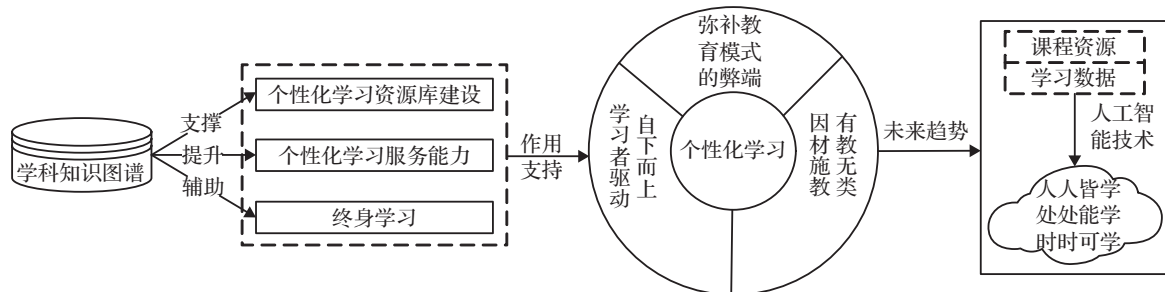


图2 学科知识图谱对个性化学习的支持分析

Fig.2 Analysis of support of discipline knowledge graph for personalized learning

(2)学科知识图谱可提升个性化学习服务能力

个性化学习的实现是新时代教育理念追求的目标,但是由于学习者群体规模庞大、学习者个人特征不尽相同等因素导致了“学习者驱动,自下而上”的学习模式难以实现,个性化学习对学习者的服务能力欠缺的问题也逐渐显现^[40]。随着知识图谱等人工智能技术在教育领域的应用,充分挖掘学习者的个人特征并对学习者画像进行建模已经成为现实。目前,针对学习者认知状态诊断这一问题,研究者在充分结合知识图谱与机器学习等技术的基础上,构建了教育知识图谱概念模型^[41],通过该模型的应用,有效实现了将学习者置于核心地位的教学模式,从而在一定程度上解决个性化学习服务能力不足的问题,使得学习者在汲取知识的同时获得成功体验感。在未来,学科知识图谱赋能个性化学习的相关研究不仅要在技术层面进行深入挖掘,还应对其下一步的发展进行多视角、多层次的深入探索。

(3)学科知识图谱可辅助学习者实现终身学习

个性化学习是终身的过程,终身学习是个体生存的的必要条件和适应时代发展的基本生存方式,也是弥补当前教育弊端根本途径。个性的塑造是个体发展的终极目标,而个性的发展则是个体活力和创造力的表现。持续地学习更新迭代的知识是实现终身学习的主要途径,21 世纪的一个显著特征是每三到五年,人类的知识就会增加一倍。学科知识图谱具备完善的知识更新机制,可以对外界的知识进行动态的感知与更新,不断吸收优质的学科知识资源。此外,通过在学科知识图谱中嵌入个性化引导引擎的方式,可以有效提升学习者的思维能力并帮助学习者完善对知识的感知和对自身的评价与反思^[42]。学习者不仅可以借助学科知识图谱的动态更新机制对新的知识进行学习,还可以在不断的进行学习中进行自我评价和反思,从而完善自身的知识体系,实现终身学习。

虽然目前的个性化学习体系还不够完善,但是实现真正意义上的个性化学习已经成为可能。杨宗凯教授^[43]曾指出:未来的个性化学习将以满足每一个学习者的个性化发展需求为目标,在大数据等技术的支持下,持续地对学习者相关学习数据进行收集和分析。刘三女牙教授^[44]曾指出:未来的个性化学习模式将更加智能化地

适应学习者在多个时空场景的不断迁移和变换,“人人皆学、处处能学、时时可学”的学习新生态也将成为现实。个性化学习模式的实施与发展是教育改革的必经之路,也是推进教育新基建、完善教育体系的重要手段。

4 学科知识图谱构建

学科知识图谱的构建是一项学科知识工程,目的是对不同教育学科分类下的学科知识进行高度的凝练和可视化展示,从而形成信息、知识、数据和智慧四者之间转化的闭环。学科知识图谱的构建不仅为个性化学习的实现提供了理论基础和技术支撑,还为学科知识在社会层面的具体应用提供了思路和指导。

考虑到学科知识图谱对学科知识数据质量的高要求,主要采用自顶向下的方式构建。构建流程如图3所示。首先,在获取教育领域和开放领域的学科知识资源的基础上构建学科知识图谱的本体结构,用来对学科知识进行建模和规范;其次,进行学科知识抽取,以获得学科知识实体以及实体间的关系;接下来,通过学科知识融合实现实体和实体间关系的融合过滤;最后,通过学科知识加工实现学科知识图谱的构建。

4.1 学科本体构建

本体是同一领域内不同主体之间进行交流的语义基础^[45],位于知识图谱的模式层,可以清晰地描述概念的层次体系^[46]。构建本体的目标就是要将一个或多个特定领域的概念和术语规范化,为其在该领域的应用提供便利。

用于本体构建的工具主要有Protégé、Web Onto以及Onto Edit等。Protégé具有图形化的操作界面,能够为用户提供详细的帮助文档,可实现模块化设计^[47]。Web Onto提供多重继承机制,支持用户浏览、构建和编辑本体,便于信息检索^[48]。Onto Edit支持用图形化的方式进行本体构建,并支持多种插件,具有良好的扩展功能^[49]。

学科知识本体是对学科概念进行建模的规范,用于描述实体的类型和实体间关系的类型。其构建方式可分为手动构建、自动构建和半自动构建三种。三种构建方式的特点以及优势如表3所示。

手动构建的方式需要人工进行编辑,并结合学科专家的指导来定义实体类型和实体间关系的类型,这种方

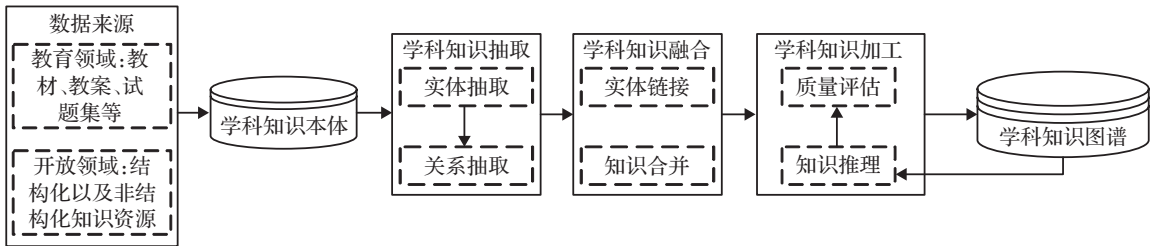


图3 学科知识图谱构建流程

Fig.3 Construction process of discipline knowledge graph

表3 本体构建方式对比

Table 3 Comparison of ontology construction methods

本体构建方式	构建特点	优势
手动构建 ^[50]	人工定义规则	准确率高
自动构建 ^[51]	利用机器学习等技术实现规则定义	降低构建成本,减轻人工负担
半自动构建 ^[52-53]	以计算机为主,人工干预为辅	保证准确率的同时减轻人工负担

式能有效提高构建结果的准确率。目前,有研究者采用手动构建的方式实现了学科知识图谱的本体构建,并建立了课程标准、学科知识以及学习资源等内容之间的联系^[50],如图4所示。

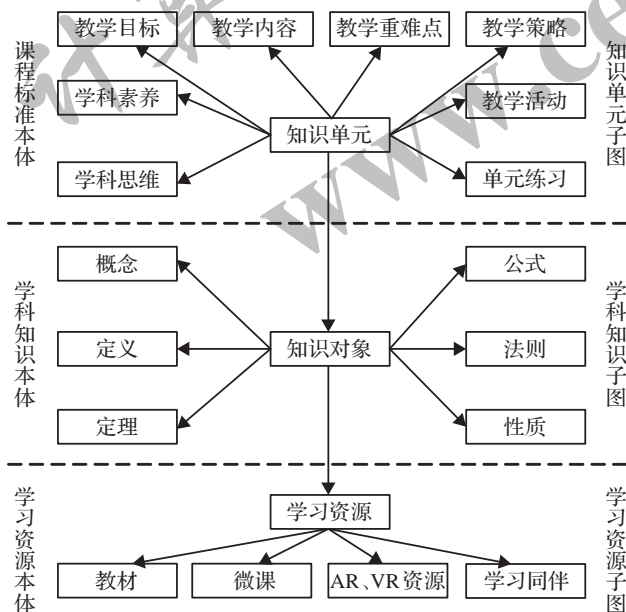


图4 手动构建的学科知识图谱本体

Fig.4 Manually constructed ontology of discipline knowledge graph

自动构建的方式主要是从已有的本体库出发,利用机器学习等技术进行逐步扩展,这种构建方式能有效降低构建成本,减轻人工负担。目前,有研究者在获取了大量的网络课程资源的基础上,通过网络爬虫、机器学习以及深度学习等技术对课程关键词、概念以及关系等信息进行抽取,实现了课程本体的自动构建,并利用本体建模工具建立了课程本体的OWL(ontology web language)描述,完成了课程本体的表示与存储^[51]。

半自动构建的方式主要是将手动构建和自动构建两种方式进行有效结合,在保证准确率的前提下减少人工负担。目前,有研究者针对法律问题的自然语言答案难以确定的问题,利用互联网上的结构化文本数据和本体论的思想,实现了法学本体的半自动构建,为法律规则的高效组织和法律问题的精准回答提供了思路^[52]。此外,高劲松等人^[53]在深入分析学科知识结构和本体特征后,提出了一种改进的学科知识本体半自动构建方法。首先,在领域专家和知识工程师的指导下确定学科

知识范畴并获取学科知识;其次,依据学科类别、学科术语、同义词及主题索引对学科知识进行组织;再次,进行本体构建,实现本体形式化表示;最后,对本体进行检测,以保证本体的正确性、一致性和可扩展性。

构建高质量的学科知识本体是构建完备、准确的学科知识图谱的先决条件。但是,在实际应用中,当现实的知识体系变化时,之前构造的本体就必须进行相应的演化,从而保持与知识体系的一致性。而且,不同本体之间对实体类型、关系类型的定义有很大出入,导致本体之间映射、扩充与合并等操作是十分复杂的,构建高质量的跨学科知识本体具有一定的难度。针对这一问题,亟需在现有构建方式的框架、步骤以及适用范围的基础上,提出一种泛化性强、健壮性强以及可扩展性强的本体构建方法,在进行映射、扩充以及合并等操作时对实体类型以及关系类型进行统一,以保证操作的规范性,同时提升跨学科知识本体的质量以及自动化构建水平。

4.2 学科知识抽取

学科知识抽取是在获取到的学科资源数据中抽取出现概念实体和实体间关系,形成“实体-关系-实体”的三元组表示形式的过程。

知识抽取是构建学科知识图谱的重要步骤,具体分为实体抽取和关系抽取两部分^[54]。面向结构化数据的学科知识抽取主要采用直接映射的方式,这种方式可以将关系型数据库中的数据直观地映射在资源描述框架(resource description framework, RDF)中,为定义复杂的转换奠定了基础。早期的面向非结构化数据的学科知识抽取主要依赖于人工制定规则和机器学习的方式,但是这两种方式在面对大规模数据时表现乏力,导致知识抽取效率不高。随着以卷积神经网络(convolutional neural networks, CNN)^[55]和循环神经网络(recurrent neural networks, RNN)^[56]为代表的深度学习^[57]技术的出现,面向非结构化数据的学科知识自动抽取成为现实。

4.2.1 实体抽取

实体抽取又称命名实体识别,其目标是将文本中的实体元素提取出来,包括人名、组织名、地理位置和时间等^[58]。学科知识的实体抽取就是从获取到的学科数据中将知识主题相关的概念、公式以及原理等提取出来。

目前,实体抽取的方法主要可以划分为基于规则的实体抽取、有监督的实体抽取和无监督的实体抽取。其中,有监督的实体抽取又可以划分为基于特征的实体抽取和基于深度学习的实体抽取^[59]。实体抽取方法分类如图5所示。

基于规则的方法首先需要由领域专家和语言学者构建一系列的实体抽取规则,并在此基础上对文本进行字符串匹配,以实现实体抽取。基于规则的实体抽取方法具有较高的准确率,但是这种方式存在耗时耗力、覆盖面不足和自动化程度低等问题,在面对海量的学科知

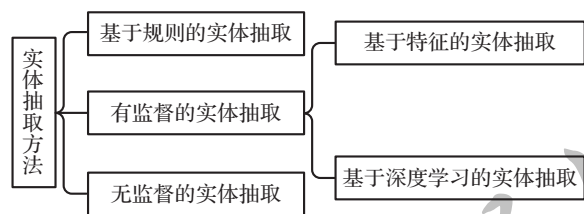


图5 实体抽取方法分类

Fig.5 Classification of entity extraction methods

识数据时并不适用。因此,在学科知识实体抽取方面,无监督的实体抽取、基于特征的实体抽取以及基于深度学习的实体抽取相关研究成为主流。

基于无监督学习的实体抽取的一个典型方法是聚类^[60],该方法的主要思想是根据文本相似度从聚类组中提取实体。无监督的实体抽取方法优势是能够处理无标签数据。但是,因为无标签数据所体现出来的样本特征较少,所以最终的分类结果也会存在准确率不高的问题。目前,有研究者将无监督学习思想融入到了化学学科的实体抽取研究中。首先,利用布朗聚类 and 词嵌入的方式生成单词表示特征;其次,结合领域知识增强样本数据的语义特征;最后,通过机器学习模型抽取化学学科的命名实体。实验表明,通过领域知识与无监督特征学习相结合的方式能够有效解决缺乏标注数据的问题^[61]。

基于特征的实体抽取是将实体抽取作为序列标注任务,该方法应用起来比较方便,具有较高的可移植性,但是也存在过度依赖特征工程的缺点。基于特征的实体抽取通常需要通过语料标注、特征定义和训练模型三步实现。斯坦福大学的NER(named entity recognition)^[62]是一个被广泛应用于特征定义的工具,具有较高的准确率。经常采用的模型主要有条件随机场(conditional random field,CRF)^[63]等。CRF是一种用于数据分割和序列标记的概率模型,被广泛应用于实体抽取、句法分析、图像识别等领域。目前,有研究者在构建学科词汇领域词典的基础上,利用CRF模型实现了针对学科知识资源的词级别特征提取和序列标注,成功提取了学科知识命名实体,并取得了80%以上的F1值^[41]。F1值的计算方法如式(3)所示。其中P为精确率,其含义是在所有被预测为正样本中实际为正样本的概率;R为召回率,其含义是在实际为正样本中被预测为正样本的概率。在式(1)以及式(2)中,TP代表预测为正例的正样本,FP代表预测为正例的负样本,FN代表预测为负例的正样本。F1值同时兼顾了精确率和召回率,可以看作是模型精确率和召回率的一种调和平均。

$$P = \frac{TP}{TP + FP} \times 100\% \tag{1}$$

$$R = \frac{TP}{TP + FN} \times 100\% \tag{2}$$

$$F1 = \frac{2 \times R \times P}{P + R} \times 100\% \tag{3}$$

基于深度学习的实体抽取方法无需人工制定复杂的特征,可以从输入的语料中自动提取特征,但是该方法往往缺乏可解释性。目前,长短时记忆神经网络(long short-term memory neural network, LSTM)及其双向形式应用极为广泛。LSTM的双向形式被命名为BiLSTM^[64],在实体抽取方面,BiLSTM-CRF模型^[65]表现优异。该模型自底向上分别是Embedding层、BiLSTM层和CRF层。Embedding层中词的向量表示是BiLSTM层的输入,BiLSTM层利用正向LSTM和反向LSTM分别计算每个词考虑左侧和右侧词时对应的向量,并将两个向量相结合,形成词向量输出,CRF层以BiLSTM层输出的词向量作为输入,对语句中命名实体进行序列标注。目前,有研究者在BiLSTM-CRF的基础上融合了基于变换器的双向编码器表示技术(bidirectional encoder representations from transformers, BERT)对模型进行优化,构建了BERT+BiLSTM-CRF模型,模型结构如图6所示。BERT是一种预训练语言模型,能够很好地学习到上下文信息并有效解决一词多义等问题^[66]。在利用BERT+BiLSTM-CRF模型进行实体抽取时,首先需要对BERT模型进行训练,以获得文本数据的向量表示,然后通过BiLSTM和CRF进行特征提取和序列标注,最终抽取相应实体。在学科知识实体抽取方面,张毅等人^[67]在BERT+BiLSTM-CRF模型的基础上融合了IDCNN,有效改善了传统命名实体识别模型无法表征一词多义以及忽略部分局部特征的问题,并在初等数学试题数据集上进行实验,取得了82.29%的P值,94.95%的R值以及93.91%的F1值,相比于BiLSTM-CRF模型和BERT+BiLSTM-CRF模型在F1方面分别提升了4.29%和1.23%。上述学科知识实体抽取方法的总结与对比分析如表4所示。

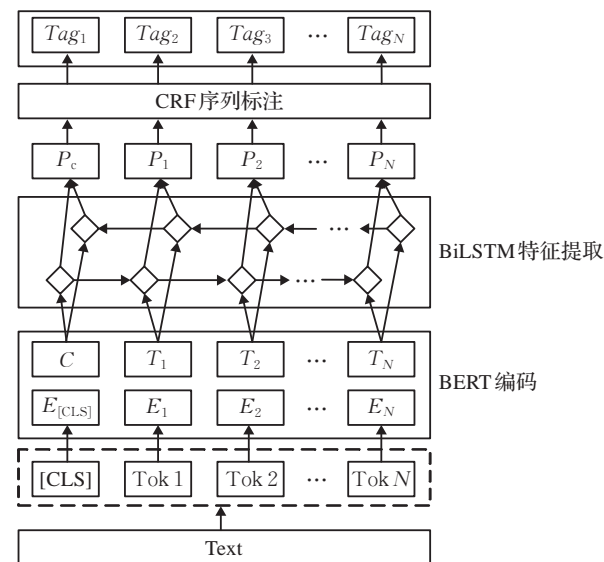


图6 BERT+BiLSTM-CRF模型结构

Fig.6 BERT+BiLSTM-CRF model structure

表4 学科知识实体抽取方法总结

Table 4 Summary of discipline knowledge entity extraction methods

实体抽取方法	主要思想	优点	缺点	相关工作
基于规则的实体抽取	由领域专家和语言学者构建一系列的实体抽取规则,并对文本进行字符串匹配	准确率高	耗时耗力,覆盖面不足,自动化程度低,不适用于海量的学科知识数据	—
无监督的实体抽取	根据文本相似度从聚类组中提取实体	能够处理无标签数据	分类结果不够准确	将无监督学习思想融入到特征工程中,有效解决缺乏标注数据的问题 ^[61]
基于特征的实体抽取	将实体抽取作为序列标注任务	可移植性较高	过度依赖于特征工程	利用 CRF 模型实现词级别的特征提取和序列标注,提升了知识图谱自动化构建程度 ^[41]
基于深度学习的实体抽取	利用深度学习模型从输入的语料中自动提取特征	无需复杂的特征工程,能够学习高维潜在语义信息	缺乏可解释性	将 IDCNN 融入 BERT+BiLSTM-CRF 模型中,有效解决一词多义以及特征提取过程中局部特征被忽略的问题 ^[67]

近年来,在学科知识实体抽取方面基于有监督学习和无监督学习的方法都取得了大量的研究成果,但是在对较为复杂的学科语料的研究中还存在实体抽取准确率不高等问题,还有较为广阔的发展空间。传统的命名实体识别模型是将实体抽取当作序列标注任务,但是这种方式无法很好地解决嵌套命名实体识别的问题。嵌套命名实体是一种特殊的命名实体,即一个实体内部还存在一个或多个其他实体。例如,“字符串常量”这个实体中包含“常量”这个实体,用传统的序列标注模型进行识别时可能只会将“常量”这个实体识别出来,但是“字符串常量”这个实体会识别失败。针对这一问题,可以将序列标注任务转换为预测实体在文本中实体头部索引和实体尾部索引的任务,从而提升实体抽取任务的准确率。此外,现有的预训练语言模型如BERT,需要消耗大量的计算资源,时间开销也会随着文本长度以及模型层数的增加而变大。针对这一问题,可以将知识蒸馏思想融入到实体抽取任务中。知识蒸馏的本质是利用具有卓越性能的大型模型对简单模型进行指导,从而让简单模型学习到大型模型已经训练好的参数和权重,以实现减少资源和时间成本的开销,进而在不过分降低准确率的同时提升模型的训练速度。

4.2.2 关系抽取

为形成具有语义关系的学科知识网络,在对学科知识进行实体抽取后,还需要对实体之间的关系进行挖掘。学科知识关系抽取是指从学科知识数据中抽取出实体间关系的过程,是构建学科知识图谱的重要步骤。

目前,用于关系抽取的常用工具主要有NLTK、DeepDive、Stanford CoreNLP以及LTP-Cloud等。NLTK是一个基于Python语言开发的自然语言处理工具,可用于文本分类、文本分析以及推断实体间存在的关系等方面^[68]。DeepDive由斯坦福大学发布,能有效减少数据噪声并剔除不精确的数据,提高关系抽取的准确率^[69]。Stanford CoreNLP由众多的语法分析工具集成,如命名实体识别器、词性标注器等。Stanford CoreNLP可以利

用句子中的语法等信息实现实体间关系抽取^[70]。LTP-Cloud由科大讯飞和哈尔滨工业大学联合推出,可用于语义角色标注和语义依存分析等,对关系抽取任务具有一定的帮助^[71]。

学科知识关系抽取的方法主要分为三种:基于规则的方法、基于机器学习的方法以及基于深度学习的方法。其中,基于深度学习的方法又可分为有监督的关系抽取和远程监督的关系抽取。

基于规则的关系抽取重点是依靠学科专家和语言学者人工制定高质量的关系抽取规则,并结合同义词词林以及依存句法分析等技术从文本中抽取出实体间的关系,从而形成“实体-关系-实体”的三元组表示形式。基于规则的关系抽取方法的优点是准确率高,但是由于需要人工参与制定规则,存在费时费力、难以维护等问题。目前,有研究者通过利用同义词词林对关系触发词进行词频统计,确定了最大词频候选关系和次大词频候选关系,再结合特定的关系判别规则实现了人物关系的抽取,并在主流的人物关系抽取公开数据集上取得了较好的结果^[72]。此外,在生物医学学科知识图谱构建方面,赵哲焕等人^[73]针对传统基于规则的关系抽取方法召回率较低的问题,提出了一种改进的基于规则的关系抽取方法。该方法首先利用机器学习技术在生物医学文献中抽取蛋白质实体,组成相应的实体二元组;其次利用句法分析技术对关系抽取规则进行改进和丰富,制定相应的句法模板;最后通过词典匹配的方法抽取出实体二元组中两个实体间关系类型的关系词。该方法最终在公开蛋白质关系抽取数据集上取得了86.98%的 P 值(相比于只使用词典匹配和句法模板分别提升了4.0%和8.4%),79.40%的 R 值(相比于只使用词典匹配和句法模板分别提升了6.3%和49.7%)以及83.02%的 $F1$ 值(相比于只使用词典匹配和句法模板分别提升了5.3%和39.9%)。

随着机器学习技术的出现及其在教育领域中的应用,基于规则的关系抽取方法存在的弊端在一定程度上

得以解决,关系抽取任务的召回率也有所提升。但是,基于机器学习的关系抽取仍需要人工提取特征,未能实现完全自动化的关系抽取。目前,有研究者从MOOC课程资源出发,针对学习者在学习后续课程知识之前应该先学习什么先验知识的问题,通过随机森林等机器学习方式挖掘了学科知识资源间的潜在关系,并充分分析了学科知识资源之间的关系对学习者的学习效率的影响,对个性化学习的发展具有重要意义^[74]。还有研究者针对传统的基于机器学习的关系抽取方法效率难以提高的问题,提出了融合特征增强的机器学习关系抽取。首先通过计算中学课程知识点与核心谓语、语义角色之间的距离以及知识点之间的相似度,获得了知识点的位置特征以及基于词向量的特征,得到了知识点之间更深层次的语义关系;其次利用支持向量机(support vector machine, SVM)与K-近邻算法的机器学习技术实现了知识点之间的关系预测,并构建了中学课程领域的多学科知识图谱。该方法最终取得了80.99%的P值,78.93%的R值以及79.55%的F1值,相比于不使用特征增强的方法分别提升了4.02%、4.00%以及3.62%^[75]。

相比于基于规则以及基于机器学习的关系抽取方法,有监督的关系抽取方法和远程监督的关系抽取方法可以在通过大量数据训练的基础上自动获得模型,不需要人工提取特征。其中,有监督的关系抽取方法主要以LSTM形式为主。Song等人^[76]针对现有关系抽取方法会忽略实体上下文特征的问题,利用双向LSTM提取实体对之间的上下文特征,结合词法、句法、语义和位置特征实现了学科资源的实体关系抽取,并在人工标注的数据集上进行实验,取得了82.36%的P值,82.11%的R值以及82.23%的F1值,相比于不添加特征的方法分别提升了16.23%、17.86%和17.05%。

虽然有监督关系抽取方法在针对领域知识进行关系抽取时能够更有效地提取特征,但是不能很好地处理

大量的无标签数据,而远程监督的关系抽取方法可以通过数据自动对齐远程知识库的方式来有效解决这一问题^[77]。目前,有研究者针对有监督的关系抽取受限于文本质量低、语料稀缺、标签数据难以获取以及模型复杂程度高等问题,结合远程监督的思想,利用基于关系表示的注意力机制提取了知识点之间的有向关系信息,并基于BERT模型实现了初等数学的知识点的关系抽取,有效降低了模型的复杂度,同时也为其他学科知识点之间的关系抽取奠定了基础。最终该方法在人工构建的初中数学知识点关系抽取数据集上取得了75.35%的P值,70.07%的R值以及72.61%的F1值,相比于PCNN模型分别提升了1.20%、4.24%、2.87%^[78]。此外,还有研究者在获取计算机学科教学大纲、教材、教学计划以及互联网中百科全书文本等资源的基础上,将远程监督的思想引入到学科知识关系抽取数据集构建任务中,发现利用远程监督方法标记的训练数据只包含少量的噪声数据,可以有效减轻人力负担^[79]。上述学科知识关系抽取方法的总结与分析如表5所示。

在学科知识图谱构建过程中,关系抽取任务是建立在实体抽取任务基础上进行的,也就是将实体抽取和关系抽取视作两个单独的任务。例如,首先对文本中实体进行检测和提取,然后预测实体之间的关系类型,从而获取学科知识的“实体-关系-实体”三元组表示形式。这种流水线的方式虽然已经取得了很大的成功,但是还存在误差传播的缺点,导致关系抽取任务容易受到实体抽取任务引入的误差的影响。针对这一问题,可以将实体抽取任务与关系抽取任务进行融合,在充分利用句法结构知识以及深度学习模型的基础上提出一种联合抽取机制,促进实体、关系等信息的融合,在一定程度上缓解实体抽取模型和关系抽取模型之间的误差传播问题,从而最大程度地减少外界因素的影响,实现高质量、高覆盖、高平衡的学科知识抽取。

表5 学科知识关系抽取方法总结

关系抽取方法	主要思想	优点	缺点	相关工作
基于规则的关系抽取	依靠人工制定的关系抽取规则,从文本中抽取实体间的关系	准确率高	费时费力,难以维护	利用句法分析技术对关系抽取规则进行完善,通过词典匹配技术抽取实体间关系 ^[73]
基于机器学习的关系抽取	将关系抽取任务当作分类问题,利用机器学习模型实现关系预测	可有效提升召回率	需要人工提取特征	通过机器学习方式挖掘学习资源之间的关系,有效解决了学习者在学习后续课程知识之前应该先学习什么先验知识的问题 ^[74]
有监督的关系抽取	对模型投入足量的已标记关系类别的训练语料,然后进行特定关系的抽取	可减少特征提取误差	处理无标签数据能力较差	通过特征增强的方式获取知识点之间更深层次的语义信息,有效提升了基于机器学习的关系抽取方法的效率 ^[75]
远程监督的关系抽取	通过数据自动对齐远程知识库的方式实现关系抽取	可处理无标签数据	容易出现错误标注	利用双向LSTM提取了实体对之间的上下文特征,有效解决了现有关系抽取方法会忽略实体上下文特征的问题 ^[76]
				提出了适用于知识点之间有向关系抽取的BERT模型,有效解决了有监督的关系抽取受限于文本质量低、语料稀缺、模型复杂程度高的问题 ^[78]

4.3 学科知识融合

知识融合是一种高水平的知识组织方式^[80]。知识融合可以在实体抽取和关系抽取的基础上实现知识的融合过滤。知识融合的过程包括两方面:实体链接与知识合并。

4.3.1 实体链接

实体链接是指将自然语言文本中的实体指称与知识图谱中的实体相关联的过程^[81]。学科知识的实体链接分为实体指称识别、候选实体生成和候选实体消歧三步,如图7所示。

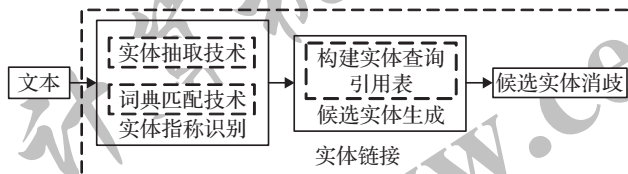


图7 学科知识实体链接流程

Fig.7 Entity linking process of discipline knowledge

目前,用于实体链接的工具主要有Dexter、TAGME以及AGDISTIS等。Dexter是一个用于实体链接的标准程序,不需要其他软件和高性能的硬件作为依托,用户可以很容易地掌握Dexter的使用方法^[82]。TAGME能够实现短文本的注释,并能够将短文本链接到Wikipedia中^[83]。AGDISTIS能够分析出文本中实体的具体地址,并将实体链接到DBpedia中^[84]。

实体指称识别是学科知识实体链接的第一步,这一步骤主要利用实体抽取技术和词典匹配技术来实现。词典匹配技术需要依赖事先构建的实体指称词典,从而与文本进行匹配实现指称识别^[85]。目前,有研究者从个性化学习的角度出发,从学生对问题的回答中抽取知识单元作为实体,从而构建实体词典,并基于所构建实体词典进行学习状态诊断的相关研究,有效促进了个性化学习的发展进程^[86]。此外,还有研究者针对手工构建实体词典费时费力的问题,利用主动学习技术实现了实体词典的半自动构建,并在维基百科上进行实验,对其中的29个实体类别进行识别, $F1$ 值达到了93.5%^[87]。

候选实体生成是确定识别出的实体指称可能指向的实体集合,以提高实体链接的准确率。学科知识候选实体生成主要通过构建查询实体引用表来实现。构建查询实体引用表可以建立实体指称与候选实体的对应关系^[85],实现候选实体集合的获取。实体引用表示例如表6所示。

在完成实体指称识别和候选实体生成后,需要为实体指称确定指向的实体,这一步骤被称为候选实体消歧。目前,有研究者针对规模化、分散化的在线教育资源给用户带来了信息过载、资源选择困难等问题,构建了MOOC资源知识图谱,根据文本相似度对不同平台的教师、院校进行候选实体消歧,从而有效避免了数据

表6 实体引用表示例

Table 6 Example of entity reference

实体指称	候选实体集合
二叉树	(1)完美二叉树
	(2)完满二叉树
	(3)完全二叉树
图	(1)有向图
	(2)无向图
	(3)完全图

的冗余,为学习者了解学习资源概况、查找与个人需求相匹配的课程提供了实质性的帮助^[88]。还有研究者基于层次过滤的思想,对通过爬虫获取的课程资源进行实体消歧,并构建了基于多源异构数据融合的初中数学知识图谱,为学习者提供了及时且智能的学习支持服务,为破解初中数学在线教学个性化不足问题提供了思路^[27]。

在教育大数据的环境下,学科知识的实体链接相关工作主要针对非结构化数据,但是对于目前学科知识图谱构建中一对多、多对多的实体关系挖掘仍面临噪声处理能力不足和先验数据不充分等挑战^[89]。其中最为突出的是现有研究还尚未关注跨模态实体语义计算的实体链接,致使学习者在学习过程中因知识资源模态单一,出现思维定式以及学习方式固化的问题。因此,提出一种完善的多模态教育资源智能组织策略,进一步拓展与深化学科知识实体链接的相关研究是至关重要的。此外,因为中文语料库的语义表达稀疏,存在大量难以区分的相似实体,这就导致一般模型过于依赖文本以外的特征信息,难以对文本特征进行全面且深入的学习,进而致使基于学科知识图谱的上层应用的性能存在限制。针对这一问题,可以将实体链接任务的重点聚焦于候选实体消歧,将候选实体消歧作为分类任务,利用多通道网络模型和注意力机制等技术进行多维度、深层次的文本特征提取,进而降低模型对外部特征的依赖,提升实体链接任务的准确率。

4.3.2 知识合并

除了结构化数据和非结构化数据外,构建学科知识图谱的数据来源可能还包括少部分的半结构化数据,半结构化数据多以XML、CSV、JSON等格式存储。对于这部分数据,同样可以采用RDF形式进行整理,但是前提是需要将其转化为结构化数据,目前可通过相关工具实现转换。例如,可通过XSPARQL^[90]实现XML与RDF的转换,通过Datalift^[91]可实现XML、CSV和RDF的转换。学科知识合并流程如图8所示。

目前,有研究者在获取多源异构学科数据的基础上,融合了基于文献计量学的科学知识图谱和基于知识库的知识图谱,实现了甲骨文知识图谱的构建,有效解决了不同结构的数据难以融合的问题,并为其他学科古籍类知识图谱的构建提供了借鉴^[92]。还有研究者通过

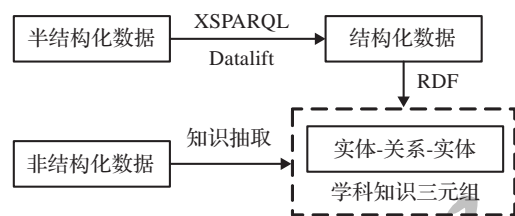


图8 学科知识合并流程

Fig.8 Process of merging discipline knowledge

对多源异构时空数据的分析与挖掘,研究了地理知识图谱构建的关键技术,并为实现地理信息的知识化,提高虚拟地理环境系统的智能化服务水平做出了贡献^[93]。

知识合并技术的优化对数据的质量具有很强的依赖,因此在获取学科知识数据时要加大对数据的筛查力度,将获取高质量的数据作为出发点。在高质量的学科知识数据的基础上进行学科知识图谱构建可以有效减少噪声,提升构建环节中神经网络模型对数据样本支持的鲁棒性。

4.4 学科知识加工

通过知识抽取与知识融合后,可以获得一系列事实表达。但是,事实并非知识,还需要通过知识加工才能获得清晰的知识体系。知识加工包括知识推理以及质量评估。

4.4.1 知识推理

知识推理在知识图谱发展与演变的过程中至关重要,主要用于对知识图谱进行知识补全。学科知识推理主要包括实体和关系的推理,目的是扩展和丰富学科知识网络^[94]。

学科知识推理较为成熟的方式可以分为四种,分别为基于规则的推理、基于图结构的推理、基于本体的推理以及基于分布式表示的推理。

基于规则的知识推理是通过人工制定一系列推理规则,并将这些规则应用到知识图谱中,实现知识的补全和纠错。基于规则的方法可以有效挖掘知识图谱中隐藏的语义信息,提高推理结果的准确率。但是,随着知识图谱规模的增大,对复杂规则的需求也越来越大,规则制定所需的人工和时间成本也在逐渐提高。目前,有研究者针对一般的答疑系统缺乏对知识的智能表示的问题,利用构建的推理规则,实现知识点的低阶推理,完成了知识图谱的补全。并利用补全后的知识图谱构建了知识教学答疑系统,在华东师范大学公共计算机课程上进行实验。结果表明,该系统能够更好地对问题进行分类和表示,对学生提问问题的分类准确率达到了80%左右^[95]。

基于图结构的知识推理首先是将整个知识图谱视为一个有向图,利用图中节点表示实体以及实体属性,利用边表示实体间的关系或者实体和其属性间的关系,然后利用相关算法实现知识推理。其中,最经典的算法

是路径排序(path ranking, PR)算法^[96],该算法的主要思想是将知识图谱看作由一组三元组组成的多关系图,对于一个给定的关系,利用枚举的方法将此关系所连接的路径一一列举,然后利用这些路径作为特征训练一个二分类器,进而实现实体间关系推理和预测。因为需要将所有路径作为特征进行训练,而且对于每一个实体对都需要判断它们是否满足每一种存在的关系,所以会消耗大量的计算资源,可扩展性较差。目前,有研究者针对学习路径难以挖掘的问题,将教育知识图谱中的问题图式视为图,问题或任务视为节点,关联关系视为边,从初始问题解决状态开始,利用PR算法在图上随机游走,到目标问题解决状态结束,进而实现教育知识图谱中学习路径的挖掘,为智慧学习环境的构造提供了支持^[97]。

基于本体的知识推理是将本体中蕴含的语义信息和逻辑提取出来,对实体以及实体间的关系进行推理,从而实现知识图谱的补全。基于本体的知识推理可以充分利用本体中的先验知识,并在一定程度上简化推理过程。但是,覆盖率广的本体难以获得,这就导致基于本体的知识推理过度依赖现有数据,泛化能力较差,可计算性较低。目前,有研究者针对学科知识推理类型难以界定的问题,构建了数学课程内容本体以及数学课程习题本体,实现了数学课程中概念关系、习题类型以及习题求解方法等多种类型知识的推理,为数学课程知识的有效组织以及知识服务效果的改善提供了思路^[98]。

基于分布式表示的知识推理是将知识图谱中的知识三元组映射到低维空间,得到三元组的低维向量表示,然后将推理任务转换为简单的向量操作。该方法能够实现知识三元组到低维向量空间的映射,简化了推理过程。但是,该方法在建模时通常只考虑知识三元组的约束,未结合先验知识以融合更多的语义信息,推理能力受限。目前,有研究者针对生物学科知识之间关系表示不唯一且难以挖掘的问题,利用基于张量模型的 Tucker 分解算法实现了面向学科知识图谱的知识推理^[24]。在利用 Tucker 分解算法进行知识推理时,需要事先对学科知识三元组等数据做出符号定义。用希腊字母 χ 表示三维知识张量, χ 中的前两个维度由知识三元组的头尾实体构成,第三维度由实体间的关系构成;用大写英文字母 A 代表矩阵,表示实体集合和关系类型集合;用小写英文字母 a 代表向量, a_i 表示向量 a 中的第 i 项;用 X_n 表示三维知识张量与矩阵的 Tucker 积;用“ \circ ”表示向量的乘积。针对三维知识张量 χ ,利用 Tucker 分解可以得到核心张量 G 和三个知识矩阵 A 、 B 、 C ,核心张量 G 代表实体与关系之间语义联系的强弱水平^[99]。具体求解过程如式(4)所示。

$$\chi \approx [G; A, B, C] \equiv G \times_1 A \times_2 B \times_3 C = \sum_{p=1}^P \sum_{q=1}^Q \sum_{r=1}^R g_{pqr} \circ a_p \circ b_q \circ c_r \quad (4)$$

由式(4)可知,Tucker分解的本质是对知识进行降维处理,挖掘出知识之间的潜在关系,进而完成学科知识图谱的知识推理与知识补全。上述学科知识推理方法的总结与分析如表7所示。

面向学科知识图谱的知识推理还有很大的发展空间,在该领域还需要继续深入探索。虽然上述的知识推理方法能有效提高准确率,但是还存在计算资源消耗大、人工成本高、泛化能力低以及过度依赖先验知识等局限性。当学科知识的数量以及种类逐渐增多、学科知识图谱的规模逐渐增大时,知识之间隐藏的语义信息以及知识三元组的数量也会随之增多,同时也会导致样本的缺陷增多,例如样本数据缺少头尾实体,样本数据错误以及查询路径过长等,此时,上述的学科知识推理方法将不再适用。面向大规模学科知识图谱的知识推理,可以从数据量大、关系复杂的特点出发,通过对深度学习、表示学习、增量学习等技术的利用,建立推理能力更强以及泛化能力更高的知识推理模型,实现知识图谱中缺失实体的补全以及实体间关系多样性的挖掘,从而丰富知识的表示形式,提高推理速度,保证推理的时效性。

4.4.2 质量评估

通过质量评估来对知识图谱中的知识进行筛选和审查,是知识图谱构建过程中极其重要的环节^[100]。对学科知识图谱进行质量评估可以直观地判断其中知识的可靠性。

学科知识图谱的构建不仅应该追求构建环节的自动化程度,还应该保证构建结果的准确率和可靠性。严谨的构建方式是学科知识图谱在个性化学习中充分发挥作用的先决条件,只有高质量的学科知识图谱才能尽可能地满足学习者多方面的学习需求。

在个性化学习的背景下,针对学科知识图谱的质量评估需要在充分考虑学习者需求的基础上进行整体质量评估。学科知识图谱整体质量评估流程如图9所示。

从整体上对学科知识图谱进行质量评估需要从学

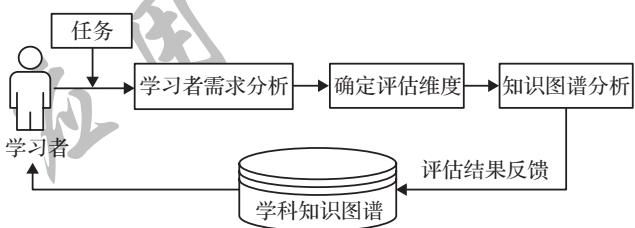


图9 学科知识图谱整体质量评估流程
Fig.9 Overall quality assessment process of discipline knowledge graph

习者的具体需求出发,在不同学习者的任务情境中理解学习者的需求。在获取到学习者需求的基础上,便可以确定学科知识图谱需要评估的维度,在相应维度下对学科知识图谱展开分析,最终将评估结果反馈至学科知识图谱的构建过程中,以实现学科知识的质量评估,从而为学习者提供高质量的个性化服务。

目前,还有研究者从学科知识实体链接的角度出发,利用 MRR 、 $Hits@N$ 等评价指标对学科知识图谱是否满足多模态资源的实体关系动态嵌入进行检测,实现了学科知识图谱与学科课程标准知识体系匹配度的计算^[24]。 MRR 表示实体动态嵌入的平均倒数排名, $Hits@N$ 表示学科知识实体链接预测中排名小于 N 的三元组平均占比,两个指标的值越大说明匹配度越高。 MRR 与 $Hits@N$ 计算公式如式(5)、(6)所示,其中 $|S|$ 代表学科知识三元组的个数。

$$MRR = \frac{1}{|S|} \sum_{i=1}^{|S|} \frac{1}{rank_i} = \frac{1}{|S|} \left(\frac{1}{rank_1} + \frac{1}{rank_2} + \dots + \frac{1}{rank_{|S|}} \right) \tag{5}$$

$$Hits@N = \frac{1}{|S|} \sum_{i=1}^{|S|} \mathbb{I}(rank_i \leq n) \tag{6}$$

在大数据、智慧赋能教育改革的背景下,教育场景中对学科知识的质量要求愈加严格,对学科知识图谱进行质量评估是构建过程中至关重要的一步,对完善知识

表7 学科知识推理方法总结
Table 7 Summary of reasoning methods of discipline knowledge

知识推理方法	主要思想	优点	缺点	相关工作
基于规则的推理	通过人工制定的推理规则实现知识的补全和纠错	准确率高	需要人工制定规则,费时费力	利用推理规则实现知识点的低阶推理和知识图谱的补全,并构建了知识答疑系统 ^[95]
基于图结构的推理	将整个知识图谱视为一个有向图,利用相关算法实现知识推理	将知识图谱视为有向图,简化推理过程	可扩展性差,计算资源消耗大	将知识图谱视为有向图,利用PR算法在图上随机游走,实现了教育知识图谱中学习路径的挖掘 ^[97]
基于本体的推理	将本体中蕴含的语义信息和逻辑提取出来,从而实现知识图谱的补全	可以利用本体中蕴含的语义信息和逻辑	过分依赖已有数据,泛化能力低,可计算性差	通过构建课程内容本体和课程习题本体,实现多种类型的知识推理,有效解决了课程知识推理中知识类型难以界定的问题 ^[98]
基于分布式表示的推理	将知识图谱中的知识三元组映射到低维空间,利用语义表达式实现知识推理	将推理任务转换为简单的向量操作	建立推理模型时没有充分考虑先验知识,推理能力受限	利用Tucker分解算法实现了面向学科知识图谱的知识推理,有效解决了学科知识之间关系表示不唯一且难以挖掘的问题 ^[24]

框架和提高知识质量具有重要意义。但是,当现实的知识体系进行变化时,知识图谱的结构也必须进行相应的改变,从而保持一致性;而且,对知识图谱进行质量评估必须考虑其应用场景,从而更有针对性地进行质量提升。由于知识体系具有动态性以及知识图谱应用场景的限制,目前还缺少一种针对学科知识图谱的质量评估框架。针对这一问题,可以从准确性、完整性以及简洁性等不同维度出发,通过计算实体链接完整率、实体以及关系冗余度、数据简洁性等指标,对学科知识图谱中实体以及关系对现实世界中事物描述的正确程度、知识图谱的完整度以及知识图谱中知识的可理解性进行全面检测,从而提高学科知识图谱的质量及其在教学过程中的可用性。

5 面向个性化学习的学科知识图谱应用

基于学科知识图谱的内涵及其对个性化学习的支持分析,可将学科知识图谱在个性化学习中的应用总结为五方面:学科知识点检索与可视化、学科知识智能问答、学习者画像模型构建、个性化学习资源推荐以及教育机器人构建,如图 10 所示。

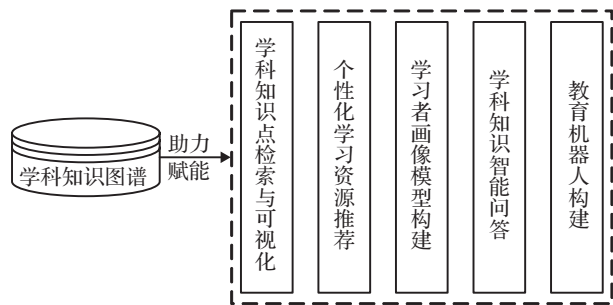


图 10 基于学科知识图谱的个性化学习应用
Fig.10 Personalized learning application based on discipline knowledge graph

(1)学科知识点检索与可视化

学科知识图谱的一个典型应用是实现学科知识资源的检索与可视化,在学习者检索某一个特定的学科知识时,知识点的详细内容及其相关的学科资源会被直观地展示出来。

文献[101]构建了一个多模态内容的检索框架,方便学习者对不同模态的教学资源进行检索与查询。文献[102]构建了一个中文情感知识图谱,并在此基础上实现了智能化的情感信息检索。文献[103]提出了一种基于知识图谱的文本观点检索方法,解决了由于用户查询时输入通常很短而导致的难以准确表达查询需求的问题。文献[104]利用知识图谱来组织文本、图像等多模态学习资源,极大程度上方便学习者进行学习资源的检索。文献[105]在 Freebase 的基础上构建了查询扩展模型,可以更好地分析用户的需求。

基于学科知识图谱的学科知识点的检索与可视化

旨在提供给学习者一个直观的学科知识网络,学习者可依据个体需求实现高效率检索,以减轻学习者在海量的学习资源中人工筛选的负担。

(2)助力学科知识智能问答

基于学科知识图谱的智能问答系统可以依据学习者的提问来提供答案,从而提高学习者的学习效率并助力个性化学习体系的发展。

智能问答最早追溯到人工智能诞生时期,阿兰·图灵^[106]提出通过观察机器是否具备正确回答问题的能力,从而验证机器是否具有智能。在教育领域,基于知识图谱的智能问答研究也备受关注。文献[107]根据构建的知识图谱,实现了英语课程知识的问答,并有效解决了当前基于知识图谱的问答方法中存在的问题答案遗漏问题。余胜泉等人^[108]通过研发基于知识图谱的育人助理系统实现了教育知识的智能问答以及个性化辅导等功能。

基于学科知识图谱的学科知识智能问答的实现是个性化学习体系构建的重要组成部分,问答结果的准确程度直接影响到个性化学习体系的质量,因此,学科知识的智能问答还需要不断地发展。

(3)助力学习者画像模型构建

个性化学习强调“以学习者为中心”的教育理念,学习者学习效果的好坏可以直观地反映出学习者对知识的掌握情况。但是,准确地诊断学习者的认知状态是一个比较复杂的过程^[109],因为学习者在进行学习活动时往往会受到多方面因素的影响。通过构建基于知识图谱的学习者画像,可以将多方面的影响因素进行整合和分析,以“打标签”的方式对学习者的多维特征进行刻画,有助于实现对学习者知识掌握情况的量化评价。

目前,有研究者构建了基于学科知识图谱的学习者画像模型,并从多个维度验证了该模型的有效性,通过该模型的应用,有效实现了对学习者学习效果的量化评价,为学习者提供精准的学习服务奠定了基础^[28]。此外,还有研究者通过融合知识图谱以及大数据等技术构建了学习者画像模型,并实现了个性化画像的输出与迭代,有效地帮助了学习者进行知识的学习与学习路径的规划^[110]。

将知识图谱等人工智能技术运用于学习者画像构建,可以提高学习者学习过程中数据的可用性和可理解性,有助于在语义层面上对学习者的学习行为、学习状态以及情感偏好等进行建模^[50],从而促进个性化学习的发展。

(4)助力个性化学习资源推荐

为学习者推荐适合其认知状态的学习资源是个性化学习发展的主要方向。近些年来,很多研究者通过人工智能等技术实现了个性化学习资源的推荐。

文献[111]提出了一种基于知识图谱的个性化资源推荐模型,通过该模型的应用,不仅取得了很好的推荐效果,还解决了数据稀疏等问题。文献[112]通过知识图

谱融合多任务特征推荐算法的方式实现了优质课程资源推荐,解决了在线课程资源缺失的问题。文献[113]通过分析课程知识点以及学习资源之间的关系,构建了课程知识图谱,并在此基础上实现了学习路径推荐和学习资源推荐,有效解决了海量的学习资源让在线学习者身处“知识迷航”的问题。文献[114]针对现有的课程视频推荐系统忽略了课程之间知识关联的问题,提出了一种基于知识图谱的跨课程视频推荐算法,有效帮助了学习者对课程知识进行深入的学习,提升了课程的教学质量。

基于学科知识图谱的个性化学习资源推荐是人工智能等技术赋能教育的实践应用,能够很好地解决学习资源与个性化学习服务不匹配的问题,是实现个性化学习的关键^[115]。

(5) 助力教育机器人构建

英国开放大学在《创新教学报告》中指出:机器人可被用于教学,起到陪伴学习者、提供学习帮助的作用,在学习过程中,机器人可充当教师、同侪或学生等不同角色^[116]。教育机器人是以培养学习者学习能力和实践能力为目标的机器人,其构建需要知识图谱、自然语言处理、知识推理等多种人工智能技术的参与。

目前,有研究者利用知识图谱等技术对教学资源和学习者学习数据进行存储,构建了大规模高质量教学资源库。在此基础上,融合深度学习、知识追踪以及知识匹配等技术构建了面向基础教育的智能教育系统。此系统不仅可以充当人类教师的角色,还可以为学习者提供资源推荐、学习效果可视化等个性化学习服务,有效促进了智能教育的发展与教育事业的现代化^[117]。此外,还有研究者基于知识图谱等技术构建了智能教育机器人,并设计了多类典型的机器人应用场景,有效帮助了学习者完成不同阶段的学习目标^[118]。汪时冲等人^[119]通过对知识图谱等技术的运用,构建了人工智能教育机器人支持下的新型“双师课堂”,实现了人工智能教育机器人与课堂环境的整合,提高了教育教学的质量,促进了教育公平的实现和优质教育资源的共享。

基于知识图谱技术构建的智能教育机器人的发展仍处于初级阶段,还面临着感知能力差、功能简单、交互性不足等诸多问题。因此,加强学科知识图谱的专业构建,持续探索学科知识图谱的创新应用,都是智能教育机器人走向规模化、专业化的重要保障。

6 问题与挑战

学科知识图谱在个性化学习体系的构建与发展中能发挥十分重要的作用,也特别符合当今教育的发展趋势,但是随着学科知识资源的体系和覆盖范围不断扩大,学科知识图谱与个性化学习相结合的过程中仍面临以下问题和挑战。

(1) 缺乏高质量公开数据集

因为学科知识图谱构建主要依赖机器学习、深度学习等人工智能技术,所以需要大量的数据来训练模型。目前,研究者多依赖手工构建的数据集来完成学科知识图谱构建任务,这样不仅会消耗大量的人力和财力,还不能保证数据质量。同时,由于学科知识图谱构建涉及多个领域,一些关系到个人、行业隐私的敏感数据是不易被挖掘的,导致高质量的公开数据集难以获得。针对上述问题,可以在实际的教学场景和网络中挖掘更多的学科知识资源,并结合学科专家对数据进行审核和筛选,从而提高数据质量,还可以将远程监督思想融入到数据集构建任务中,实现数据的自动标注,进而减少人力和财力的消耗。此外,针对敏感数据,可以利用数据脱敏技术进行处理,从而在消除原始数据中的敏感信息的基础上,保留目标任务所需的数据内容与数据特征,进而促进高质量数据集的构建与共享。

(2) 自顶向下的构建方式存在本体难以修订的问题

高质量的学科知识图谱是构建智能化学习应用以及完善个性化学习体系的基础。学科知识图谱的质量在很大程度上取决于构建方式的选取。自顶向下的构建方式需要事先构建好本体,用来对学科知识进行建模和规范。这种构建方式能够很好地体现概念之间的层次关系,并保证数据权威性。但是,随着学科知识数量和种类的不断增加,有很多数据类型可能没有包含在构建好的本体中,这时就需要根据数据特点对本体进行修订,这样会增加额外的人工成本,不利于碎片化知识到结构化知识的转换,以及学科知识图谱的高度自动化构建。因此,将自顶向下和自底向上的构建方式进行有效结合,在知识抽取的基础上构建学科知识图谱的模式层,之后再对新得到的知识进行总结归纳,实现模式层的迭代更新,并基于更新后的模式层进行新一轮的实体、关系以及属性等信息的填充,从而实现数据层的完善,进而在保证准确率和权威性的同时提升自动化构建程度,减轻人工负担。

(3) 学科知识图谱构建方法中缺乏对多模态资源融合的研究

在教育教学过程中,单纯依赖文本形式的知识资源难以帮助学习者深化对知识的理解和运用,多模态资源能够调动学习者多种感官对知识进行学习,同时还可以避免因知识资源模态单一导致的思维定式以及学习方式固化等问题。但是,因为多模态资源之间具有异质性差异,难以对特征进行统一提取,所以在学科知识图谱构建过程中,对模态资源进行融合并形成统一的知识表示具有较大难度。针对这一问题,可以将多模态资源融合任务转换为实体属性添加任务。首先,利用实体抽取技术对文本中的实体进行抽取;其次,将多模态资源的描述信息与实体名称进行匹配,从而获得多模态资源与实体的对应关系;最后,将匹配出的多模态资源作为实

体属性添加至知识图谱中,从而实现多模态资源的融合,进而丰富知识类型。

(4)深度学习模型存在灾难性遗忘的局限性

学科知识图谱的构建流程可以划分为知识抽取、知识融合以及知识加工等步骤。但是,由于不同学科之间的知识内容以及知识体系具有较大差异,在不同学科的知识图谱构建过程中针对同一任务所应用的数据集是不同的,同时也导致在某一学科数据上训练好的模型很难迁移到其他学科上直接应用,还需要重新进行模型的训练,建立相应的权重,具有灾难性遗忘的缺陷,不利于多学科知识图谱的统一构建。针对这一问题,可以将增量学习策略引入到模型训练的过程中,从而使模型能够不断地处理连续的学科知识信息流,使其获得吸收新知识的同时保留、整合以及优化旧知识的能力,从而为多学科知识图谱的统一构建提供思路。

(5)构建过程中所应用的人工智能技术缺乏可解释性

在学科知识图谱自动化构建过程中,多数任务均依赖人工智能技术以及深度学习模型进行实现。但是,现有的人工智能技术均缺乏可解释性。人工智能技术的可解释性是其能否被广泛应用的关键因素,充分理解模型在运行过程中如何决策、决策了什么以及决策是否合理是构建高质量学科知识图谱的重要保障。针对这一问题,可以在建模前、建模中以及建模后对模型进行解释。在建模之前,可以利用数据预处理或者数据可视化的方法对模型进行初步解释,当数据规模很大或者维度很高时,对数据进行预处理和可视化可以帮助人们全方位、多层次地了解数据分布,以判断模型选择是否准确合理。在建模过程中,需要建立本身具备可解释性的模型。例如,在多数深度学习任务中,一些输入和输出存在正相关或者负相关的单调性关系,如果在模型训练过程中这种单调性关系是可被挖掘的,那么就可使模型具备更高的可解释性。在建模后对模型进行解释主要针对的是深度学习模型的黑盒性质,其黑盒性质主要来自于人们没办法用人类可理解的方式对模型的决策行为和具体含义进行理解。但是,神经网络中的神经元具有分层组合的性质,可以利用一些可视化方法对神经网络中的隐藏层进行分析,将神经网络中每层学习到的概念以人类可理解的图像的形式展示出来,进而促进人工智能技术的可解释性研究。

(6)学科知识图谱构建过程中存在多学科交叉融合不足的问题

学科知识图谱作为领域知识图谱的一种,可以将学科知识资源进行有效组织和优化,进而映射学习者的思维模式和高阶认知活动,为教育的发展提供支持。与其他领域知识图谱构建不同的是,学科知识图谱的构建需要信息学科、教育学以及认知科学等多种学科的交叉融合。但是,由于学科之间具有相对独立性,学科壁垒是

客观存在的,这就导致学科知识图谱构建过程中会出现多学科交叉融合不足,学科壁垒难以突破的问题,不利于学科知识图谱的高质量构建。因此,在构建过程中,不仅要重视人工智能技术的合理运用,还需要结合多学科专家的经验和建议,对学习者的学习行为以及教育过程进行深入挖掘,从而促进学科之间的交叉融合,打破现有的学科壁垒,进而提高学科知识图谱的构建质量。

(7)学科知识图谱中缺少对学习者的学习行为等数据的融入

个性化学习的目的是根据学习者的认知水平等个体特征推荐合适的学习资源与学习方式。在此过程中,学科知识图谱通常作为一种学科知识存储的工具,被用来向学习者提供相关学习资源。但是,学习者在过程中所产生的学习行为等数据并没有被存储在学科知识图谱中,这就导致为学习者提供个性化学习服务时需要先依赖外部数据对学习者的个体特征进行挖掘,才能推荐相关学习资源,不利于教育过程的连贯性。针对这一问题,可以在学科知识图谱中构建学习行为等数据与学科知识资源之间的联系,将学习行为等数据融入到学科知识图谱中,从而建立起融合学习者个体特征与学科知识资源等信息的异构表征网络,实现认知主体、对象以及状态的一致性表达和复杂关系的建模,进而促进教育资源的精准适配。

(8)学科知识图谱与传统育人模式融合程度低

学科知识图谱作为人工智能等技术在教育领域的杰出成果,可以在很大程度上促进智慧教育的发展,加速个性化学习模式的形成。但是,随之涌现的还有全方位育人程度不足的问题,教育技术的飞速发展使得传统的育人模式受到质疑和威胁。不可否定的是,传统的育人模式至今在教育领域仍发挥着极大的作用,能够很好地展现教育情感,丰富教育过程。但是,学科知识图谱更侧重于对学科知识的总结与表述,缺少对课程思政以及教育情感等方面内容的融合,以致没有很好地完成“育人”这一教育本质对教育模式的要求。因此,亟需促进传统育人模式和学科知识图谱等新兴教育技术的深度融合,实现教育主体可理解、教育情境可计算、教育规律可解释以及教育服务可制定,使二者相得益彰,进而更好地服务于教育模式的创新与发展。

7 结束语

学科知识图谱作为智能教育从“感知智能”向“认知智能”演变的基础,能够为教育大数据的处理、个性化学习体系的构建以及智能教育应用产品的研发提供技术与理论支撑。本文从个性化学习的角度出发,对学科知识图谱的构建技术、应用方向以及目前存在的问题与挑战等方面进行综述,得出了以下结论:学科知识图谱对个性化学习的发展具有重要的指导意义,是教育变革过

程中应该重点关注的内容。虽然学科知识图谱与个性化学习具有天然的适配性,但是目前还存在构建技术不够成熟、基于学科知识图谱的个性化学习应用不够完善、与传统育人模式融合程度不足等问题。未来的研究应该重点关注学科知识图谱构建技术的创新发展,从而更好地解决当前教育领域资源呈现方式多样化、资源标注不清晰、资源之间逻辑关系不明确等问题,进而推进人工智能技术和教育环境的深度融合,使得综合性、创造性、前瞻性以及普适性的个性化学习成为可能。

参考文献:

- [1] BIZER C, HEATH T, BERNERS-LEE T. Linked data: the story so far[M]//Semantic services, interoperability and web applications: emerging concepts. [S.l.]: IGI Global, 2011: 205-227.
- [2] 马昂, 于艳华, 杨胜利, 等. 基于强化学习的知识图谱综述[J]. 计算机研究与发展, 2022, 59(8): 1694-1722.
- MA A, YU Y H, YANG S L, et al. A survey of knowledge graph based on reinforcement learning[J]. Journal of Computer Research and Development, 2022, 59(8): 1694-1722.
- [3] 李振, 周东岱, 刘娜, 等. 人工智能应用背景下的教育人工智能研究[J]. 现代教育技术, 2018, 28(9): 19-25.
- LI Z, ZHOU D D, LIU N, et al. Research on the artificial intelligence in education under the background of artificial intelligence application[J]. Modern Educational Technology, 2018, 28(9): 19-25.
- [4] 国务院. 国务院关于印发新一代人工智能发展规划的通知[Z]. 中华人民共和国国务院公报, 2017(22): 7-21.
- PRC. Notice of the state council on issuing the development plan on the new generation of artificial intelligence[Z]. Gazette of the State Council of the People's Republic of China, 2017(22): 7-21.
- [5] XIE H, CHU H C, HWANG G J, et al. Trends and development in technology-enhanced adaptive/personalized learning: a systematic review of journal publications from 2007 to 2017[J]. Computers & Education, 2019, 140: 103599.
- [6] MALAISE Y, SIGNER B. Personalised learning environments based on knowledge graphs and the zone of proximal development[C]//Proceedings of the 14th International Conference on Computer Supported Education, 2022: 199-206.
- [7] JI S, PAN S, CAMBRIA E, et al. A survey on knowledge graphs: representation, acquisition, and applications[J]. IEEE Transactions on Neural Networks and Learning Systems, 2022, 33(2): 494-514.
- [8] 付雷杰, 曹岩, 白瑀, 等. 国内垂直领域知识图谱发展现状与展望[J]. 计算机应用研究, 2021, 38(11): 3201-3214.
- FU L J, CAO Y, BAI Y, et al. Development status and prospect of vertical domain knowledge graph in China[J]. Application Research of Computers, 2021, 38(11): 3201-3214.
- [9] 张吉祥, 张祥森, 武长旭, 等. 知识图谱构建技术综述[J]. 计算机工程, 2022, 48(3): 23-37.
- ZHANG J X, ZHANG X S, WU C X, et al. Survey of knowledge graph construction techniques[J]. Computer Engineering, 2022, 48(3): 23-27.
- [10] AUER S, BIZER C, KOBILAROV G, et al. DBpedia: a nucleus for a web of open data[M]//The semantic web. Berlin, Heidelberg: Springer, 2007: 722-735.
- [11] SUCHANEK F M, KASNECI G, WEIKUM G. YAGO: a core of semantic knowledge[C]//Proceedings of the 16th International Conference on World Wide Web, 2007: 697-706.
- [12] BOLLACKER K, EVANS C, PARITOSH P, et al. Freebase: a collaboratively created graph database for structuring human knowledge[C]//Proceedings of the 2008 ACM SIGMOD International Conference on Management of Data, 2008: 1247-1250.
- [13] WU W, LI H, WANG H, et al. Probase: a probabilistic taxonomy for text understanding[C]//Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data, 2012: 481-492.
- [14] 搜狗百科. 搜狗知立方[EB/OL]. (2022-06-09)[2022-11-29]. <https://baike.sogou.com/v66616234.html>.
- Sougou Baike. Sogou known cube[EB/OL]. (2022-06-09)[2022-11-29]. <https://baike.sogou.com/v66616234.html>.
- [15] 百度百科. 百度知心[EB/OL]. (2021-12-01)[2022-11-29]. <https://baike.baidu.com/item/%E7%99%BE%E5%BA%A6%E7%9F%A5%E5%BF%83/12015195?fr=aladdin>.
- Baidu Baike. Baidu's bosom[EB/OL]. (2021-12-01)[2022-11-29]. <https://baike.baidu.com/item/%E7%99%BE%E5%BA%A6%E7%9F%A5%E5%BF%83/12015195?fr=aladdin>.
- [16] NIU X, SUN X, WANG H, et al. Zhishi.me-weaving Chinese linking open data[C]//International Semantic Web Conference. Berlin, Heidelberg: Springer, 2011: 205-220.
- [17] XU B, XU Y, LIANG J, et al. CN-DBpedia: a never-ending Chinese knowledge extraction system[C]//International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems. Cham: Springer, 2017: 428-438.
- [18] 阿里云. 电商知识图谱[EB/OL]. (2020-04-11)[2022-11-29]. <https://developer.aliyun.com/article/754652>.
- Alibaba Cloud Computing Co.Ltd. E-commerce knowledge graph[EB/OL]. (2020-04-11)[2022-11-29]. <https://developer.aliyun.com/article/754652>.
- [19] 于彤, 陈华钧, 姜晓红. 中医药知识工程[M]. 北京: 科学出版社, 2017.
- YU T, CHEN H J, JIANG X H. Knowledge engineering for traditional Chinese medicine[M]. Beijing: Science Press, 2017.
- [20] 2021 全国知识图谱与语义计算大会. 工业界论坛[EB/OL]. (2021-01-05)[2022-11-29]. http://sigkg.cn/ccks2021/?page_id=342.
- China Conference on Knowledge Graph and Semantic Computing. Industry forum[EB/OL]. (2021-01-05)[2022-

- 11-29].http://sigkg.cn/ckcs2021/?page_id=342.
- [21] 刘政昊,钱宇星,衣天龙,等.知识关联视角下金融证券知识图谱构建与相关股票发现[J].数据分析与知识发现,2022,6(2):184-201.
- LIU Z H, QIAN Y X, YI T L, et al. Constructing knowledge graph for financial securities and discovering related stocks with knowledge association[J]. Data Analysis and Knowledge Discovery, 2022, 6(2): 184-201.
- [22] 汪文生,张静静.基于CiteSpace知识图谱的能源安全研究进展与展望[J].矿业科学学报,2021,6(4):497-508.
- WANG W S, ZHANG J J. Research progress and prospect of energy security based on citespace knowledge graph[J]. Journal of Mining Science and Technology, 2021, 6(4): 497-508.
- [23] 中华人民共和国国家质量监督检验检疫总局,中国国家标准化管理委员会.学科分类与代码:GB/T 13745—2009[S].北京:中国标准出版社,2009.
- General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, Standardization Administration. Classification and code of disciplines; GB/T 13745—2009[S]. Beijing: Standards Press of China, 2009.
- [24] 林健,柯清超,黄正华,等.学科知识图谱的动态生成及其在资源智能组织中的应用[J].远程教育杂志,2022,40(4):23-34.
- LIN J, KE Q C, HUANG Z H, et al. Dynamic generation of subject knowledge graph and its application in resources intelligent organization[J]. Journal of Distance Education, 2022, 40(4): 23-34.
- [25] 李艳燕,张香玲,李新,等.面向智慧教育的学科知识图谱构建与创新应用[J].电化教育研究,2019,40(8):60-69.
- LI Y Y, ZHANG X L, LI X, et al. Construction and innovative application of discipline knowledge graph oriented to smart education[J]. e-Education Research, 2019, 40(8): 60-69.
- [26] 李龙飞,张国良.算法时代“信息茧房”效应生成机理与治理路径——基于信息生态理论视角[J].电子政务,2022(9):51-62.
- LI L F, ZHANG G L. Generation mechanism and treatment path of “information cocoon house” effect in the algorithmic era—from the perspective of information ecology theory[J]. E-Government, 2022(9): 51-62.
- [27] 周炫余,唐祯,唐丽蓉,等.基于多源异构数据融合的初中数学知识图谱构建[J].武汉大学学报(理学版),2021,67(2):118-126.
- ZHOU X Y, TANG Z, TANG L R, et al. Construction of junior high school mathematics knowledge graph based on multi-source heterogeneous data fusion[J]. Journal of Wuhan University(Natural Science Edition), 2021, 67(2): 118-126.
- [28] 赵玲朗,范佳荣,赵一婷,等.基于知识图谱的学习者画像模型设计与应用——以“高中物理”课程为例[J].现代教育技术,2021,31(2):95-101.
- ZHAO L L, FAN J R, ZHAO Y T, et al. The design and application of the learners' portrait model based on knowledge mapping—taking the “high school physics” course as an example[J]. Modern Educational Technology, 2021, 31(2): 95-101.
- [29] 杨云飞,穗志方.面向医学知识图谱的可视化方法设计与实现[J].中文信息学报,2022,36(2):40-48.
- YANG Y F, SUI Z F. Design and implementation of visualization for medical knowledge graph[J]. Journal of Chinese Information Processing, 2022, 36(2): 40-48.
- [30] CUI J, YU S. Fostering deeper learning in a flipped classroom: effects of knowledge graphs versus concept maps[J]. British Journal of Educational Technology, 2019, 50(5): 2308-2328.
- [31] BOUCHARD P. Network promises and their implications[J]. RUSC Universities and Knowledge Society Journal, 2011, 8(1): 288-302.
- [32] 赵学孔,徐晓东,龙世荣.协同推荐:一种个性化学习路径生成的新视角[J].中国远程教育,2017(5):24-34.
- ZHAO X K, XU X D, LONG S R. Collaborative recommendation: a new perspective on the generation of personalized learning path[J]. Distance Education in China, 2017(5): 24-34.
- [33] 牟智佳.“人工智能+”时代的个性化学习理论重思与开解[J].远程教育杂志,2017,35(3):22-30.
- MOU Z J. The reconsideration and solution of personalized learning theory in the era of “artificial intelligence plus”[J]. Journal of Distance Education, 2017, 35(3): 22-30.
- [34] 祝智庭,贺斌.智慧教育:教育信息化的新境界[J].电化教育研究,2012,33(12):5-13.
- ZHU Z T, HE B. Intelligence education: a new realm of educational informationization[J]. e-Education Research, 2012, 33(12): 5-13.
- [35] MODRITSCHER F, WILD F. Personalized e-learning through environment design and collaborative activities[C]//LNCS 5298: HCI and Usability for Education and Work, 4th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society, 2008: 377-390.
- [36] 钟启泉.为了中华民族的复兴 为了每位学生的发展[Z].基础教育课程改革纲要(试行)解读,2001.
- ZHONG Q Q. For the rejuvenation of the Chinese nation and the development of every student[Z]. Interpretation of the Outline of Basic Education Curriculum Reform (Trial), 2001.
- [37] 范佳荣,钟绍春.学科知识图谱研究:由知识学习走向思维发展[J].电化教育研究,2022,43(1):32-38.

- FAN J R, ZHONG S C. Research on subject knowledge mapping: from knowledge learning to thinking development[J]. *e-Education Research*, 2022, 43(1): 32-38.
- [38] 杨现民, 余胜泉. 智慧教育体系架构与关键支撑技术[J]. *中国电化教育*, 2015(1): 77-84.
- YANG X M, YU S Q. The architecture and key support technologies of smart education[J]. *China Educational Technology*, 2015(1): 77-84.
- [39] 高茂, 张丽萍. 融合多模态资源的教育知识图谱的内涵、技术与应用研究[J]. *计算机应用研究*, 2022, 39(8): 2257-2267.
- GAO M, ZHANG L P. Research on connotation, technology and application of educational knowledge graph based on multimodal resources[J]. *Application Research of Computers*, 2022, 39(8): 2257-2267.
- [40] 胡钦太, 刘丽清, 郑凯. 工业革命 4.0 背景下的智慧教育新格局[J]. *中国电化教育*, 2019(3): 1-8.
- HU Q T, LIU L Q, ZHENG K. The new structure of smart education under the background of the fourth industrial revolution[J]. *China Educational Technology*, 2019(3): 1-8.
- [41] 李振, 周东岱. 教育知识图谱的概念模型与构建方法研究[J]. *电化教育研究*, 2019, 40(8): 78-86.
- LI Z, ZHOU D D. Research on conceptual model and construction method of educational knowledge graph[J]. *e-Education Research*, 2019, 40(8): 78-86.
- [42] 刘凤娟, 赵蔚, 姜强, 等. 基于知识图谱的个性化学习模型与支持机制研究[J]. *中国电化教育*, 2022(5): 75-81.
- LIU F J, ZHAO W, JIANG Q, et al. Research on personalized learning model and support mechanism based on knowledge graph[J]. *China Educational Technology*, 2022(5): 75-81.
- [43] 杨宗凯. 个性化学习的挑战与应对[J]. *科学通报*, 2019, 64(5): 493-498.
- YANG Z K. The challenges of personalized learning and their solutions[J]. *Chinese Science Bulletin*, 2019, 64(5): 493-498.
- [44] 刘三女牙, 刘盛英杰, 孙建文, 等. 智能教育发展中的若干关键问题[J]. *中国远程教育(综合版)*, 2021(4): 1-7.
- LIU S Y, LIU S Y J, SUN J W, et al. Key issues concerning the development of intelligent education[J]. *Distance Education in China*, 2021(4): 1-7.
- [45] STUDER R, BENJAMINS V R, FENSEL D. Knowledge engineering: principles and methods[J]. *Data & Knowledge Engineering*, 1998, 25(1/2): 161-197.
- [46] WONG W, LIU W, BENNAMOUN M. Ontology learning from text: a look back and into the future[J]. *ACM Computing Surveys*, 2012, 44(4): 1-36.
- [47] RUBIN D L, NOY N F, MUSEN M A. Protege: a tool for managing and using terminology in radiology applications[J]. *Journal of Digital Imaging*, 2007, 20(1): 34-46.
- [48] 王向前, 张宝隆, 李慧宗. 本体研究综述[J]. *情报杂志*, 2016, 35(6): 163-170.
- WANG X Q, ZHANG B L, LI H Z. Overview of ontology research[J]. *Journal of Intelligence*, 2016, 35(6): 163-170.
- [49] SURE Y, ANGELE J, STAAB S. OntoEdit: multifaceted inferencing for ontology engineering[J]. *Journal on Data Semantics*, 2003, 1: 128-152.
- [50] 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.
- LI Z, ZHOU D D, WANG Y. Research of educational knowledge graph from the perspective of “artificial intelligence+”: connotation, technical framework and application[J]. *Journal of Distance Education*, 2019, 37(4): 42-53.
- [51] 童名文, 牛琳, 杨琳, 等. 课程本体自动构建技术研究[J]. *计算机科学*, 2016, 43(S2): 108-112.
- TONG M W, NIU L, YANG L, et al. Research on technique of course ontology automatically constructing[J]. *Computer Science*, 2016, 34(S2): 108-112.
- [52] FAWEI B, PAN J Z, KOLLINGBAUM M, et al. A semi-automated ontology construction for legal question answering[J]. *New Generation Computing*, 2019, 37(4): 453-478.
- [53] 高劲松, 梁艳琪, 王学东, 等. 学科知识地图的本体构建方法研究[J]. *情报科学*, 2013, 31(7): 72-77.
- GAO J S, LIANG Y Q, WANG X D, et al. Study of the ontology construction method for subject knowledge map[J]. *Information Science*, 2013, 31(7): 72-77.
- [54] 李冬梅, 张扬, 李东远, 等. 实体关系抽取方法研究综述[J]. *计算机研究与发展*, 2020, 57(7): 1424-1448.
- LI D M, ZHANG Y, LI D Y, et al. Review of entity relation extraction methods[J]. *Journal of Computer Research and Development*, 2020, 57(7): 1424-1448.
- [55] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. *Proceedings of the IEEE*, 1998, 86(11): 2278-2324.
- [56] MIKOLOV T, KARAFIÁT M, BURGET L, et al. Recurrent neural network based language model[C]//11th Annual Conference of the International Speech Communication Association, 2010: 1045-1048.
- [57] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. *Nature*, 2015, 521(7553): 436-444.
- [58] GRISHMAN R, SUNDHEIM B M. Message understanding conference-6: a brief history[C]//16th International Conference on Computational Linguistics, 1996.
- [59] 赵山, 罗睿, 蔡志平. 中文命名实体识别综述[J]. *计算机科学与探索*, 2022, 16(2): 296-304.
- ZHAO S, LUO R, CAI Z P. Survey of Chinese named entity recognition[J]. *Journal of Frontiers of Computer Science and Technology*, 2022, 16(2): 296-304.
- [60] 温秀秀, 马超, 高原原, 等. 基于标签聚类的中文重叠命名

- 实体识别方法[J]. 计算机工程, 2020, 46(5): 41-46.
- WEN X X, MA C, GAO Y Y, et al. Chinese overlapping named entity recognition method based on label clustering[J]. Computer Engineering, 2020, 46(5): 41-46.
- [61] ZHANG Y, XU J, CHEN H, et al. Chemical named entity recognition in patents by domain knowledge and unsupervised feature learning[J]. Database, 2016. DOI: 10.1093/database/baw049.
- [62] FINKEL J R, GRENNER T, MANNING C D. Incorporating non-local information into information extraction systems by Gibbs sampling[C]//Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics, 2005: 363-370.
- [63] LAFFERTY J D, MCCALLUM A K, PEREIRA F. Conditional random fields: probabilistic models for segmenting and labeling sequence data[C]//International Conference on Machine Learning, 2001.
- [64] WÖLLMER M, EYBEN F, GRAVES A, et al. Bidirectional LSTM networks for context-sensitive keyword detection in a cognitive virtual agent framework[J]. Cognitive Computation, 2010, 2(3): 180-190.
- [65] LUO L, YANG Z, YANG P, et al. An attention-based BiLSTM-CRF approach to document-level chemical named entity recognition[J]. Bioinformatics, 2018, 34(8): 1381-1388.
- [66] DEVLIN J, CHANG M W, LEE K, et al. BERT: pretraining of deep bidirectional transformers for language understanding[J]. arXiv: 1810.04805, 2018.
- [67] 张毅, 王爽胜, 何彬, 等. 基于BERT的初等数学文本命名实体识别方法[J]. 计算机应用, 2022, 42(2): 433-439.
- ZHANG Y, WANG S S, HE B, et al. Named entity recognition method of elementary mathematical text based on BERT[J]. Journal of Computer Applications, 2022, 42(2): 433-439.
- [68] BIRD S, KLEIN E, LOPER E. Natural language processing with Python: analyzing text with the natural language toolkit[M]. Sebastopol: O'Reilly Media, Inc., 2009.
- [69] ZHANG C, RÉ C, CAFARELLA M, et al. DeepDive: declarative knowledge base construction[J]. Communications of the ACM, 2017, 60(5): 93-102.
- [70] MANNING C D, SURDEANU M, BAUER J, et al. The Stanford CoreNLP natural language processing toolkit[C]//Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations, 2014: 55-60.
- [71] ZHANG W, LIU T, YIN Q, et al. Neural recovery machine for Chinese dropped pronoun[J]. Frontiers of Computer Science, 2019, 13(5): 1023-1033.
- [72] 谢明鸿, 冉强, 王红斌. 基于同义词词林和规则的中文远程监督人物关系抽取方法[J]. 计算机工程与科学, 2021, 43(9): 1660-1667.
- XIE M H, RAN Q, WANG H B. A Chinese distant supervised personal relationship extraction method based on tongyici cilin and rules[J]. Computer Engineering and Science, 2021, 43(9): 1660-1667.
- [73] 赵哲焕, 杨志豪, 孙聪, 等. 生物医学文献中的蛋白质关系抽取研究[J]. 中文信息学报, 2018, 32(7): 82-90.
- ZHAO Z H, YANG Z H, SUN C, et al. Protein-protein interaction extraction from biomedical literature[J]. Journal of Chinese Information Processing, 2018, 32(7): 82-90.
- [74] PAN L, LI C, LI J, et al. Prerequisite relation learning for concepts in MOOCs[C]//Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2017: 1447-1456.
- [75] 韩萌, 李蔚清. 基于特征增强的中文STEM课程知识的关系抽取[J]. 计算机应用研究, 2020, 37(S1): 40-42.
- HAN M, LI W Q. Relationship extraction of Chinese stem course knowledge based on feature enhancement[J]. Application Research of Computers, 2020, 37(S1): 40-42.
- [76] SONG M, ZHAO J, GAO X. Research on entity relation extraction in education field based on multi-feature deep learning[C]//Proceedings of the 2020 3rd International Conference on Big Data Technologies, 2020: 102-106.
- [77] 鄂海红, 张文静, 肖思琪, 等. 深度学习实体关系抽取研究综述[J]. 软件学报, 2019, 30(6): 1793-1818.
- E H H, ZHANG W J, XIAO S Q, et al. Survey of entity relationship extraction based on deep learning[J]. Journal of Software, 2019, 30(6): 1793-1818.
- [78] 杨东明, 杨大为, 顾航, 等. 面向初等数学的知识点关系提取研究[J]. 华东师范大学学报(自然科学版), 2019(5): 53-65.
- YANG D M, YANG D W, GU H, et al. Research on knowledge point relationship extraction for elementary mathematics[J]. Journal of East China Normal University (Natural Science), 2019(5): 53-65.
- [79] SU Y, ZHANG Y. Automatic construction of subject knowledge graph based on educational big data[C]//Proceedings of the 2020 3rd International Conference on Big Data and Education, 2020: 30-36.
- [80] 徐绪堪, 房道伟, 蒋勋, 等. 知识组织中知识粒度化表示和规范化研究[J]. 图书情报知识, 2014(6): 101-106.
- XU X K, FANG D W, JIANG X, et al. Research on knowledge granularity representation and standardization during knowledge organization[J]. Document, Informaiton & Knowledge, 2014(6): 101-106.
- [81] SHEN W, WANG J, HAN J. Entity linking with a knowledge base: issues, techniques, and solutions[J]. IEEE Transactions on Knowledge and Data Engineering, 2014, 27(2): 443-460.
- [82] CECCARELLI D, LUCCHESI C, ORLANDO S, et al. Dexter: an open source framework for entity linking[C]//

- Proceedings of the 6th International Workshop on Exploiting Semantic Annotations in Information Retrieval, 2013: 17-20.
- [83] FERRAGINA P, SCAIELLA U. TAGME: on-the-fly annotation of short text fragments (by Wikipedia entities)[C]// Proceedings of the 19th ACM International Conference on Information and Knowledge Management, 2010: 1625-1628.
- [84] USBECK R, NGONGA NGOMO A C, RÖDER M, et al. AGDISTIS-graph-based disambiguation of named entities using linked data[C]// International Semantic Web Conference. Cham: Springer, 2014: 457-471.
- [85] 王昊奋, 漆桂林, 陈华钧. 知识图谱: 方法、实践与应用[M]. 北京: 电子工业出版社, 2019.
- WANG H F, QI G L, CHEN H J. Knowledge graph: method, practice and application[M]. Beijing: Publishing House of Electronics Industry, 2019.
- [86] ZHANG Y, DAI H, YUN Y, et al. Meta-knowledge dictionary learning on 1-bit response data for student knowledge diagnosis[J]. Knowledge-Based Systems, 2020, 205: 106290.
- [87] SONG Y, JEONG S, KIM H. Semi-automatic construction of a named entity dictionary for entity-based sentiment analysis in social media[J]. Multimedia Tools and Applications, 2017, 76(9): 11319-11329.
- [88] DANG F R, TANG J T, PANG K Y, et al. Constructing an educational knowledge graph with concepts linked to Wikipedia[J]. Journal of Computer Science and Technology, 2021, 36(5): 1200-1211.
- [89] 陈烨, 周刚, 卢记仓. 多模态知识图谱构建与应用研究综述[J]. 计算机应用研究, 2021, 38(12): 3535-3543.
- CHEN Y, ZHOU G, LU J C. Survey on construction and application research for multi-modal knowledge graphs[J]. Application Research of Computers, 2021, 38(12): 3535-3543.
- [90] AKHTAR W, KOPECKÝ J, KRENNWALLNER T, et al. XSPARQL: traveling between the XML and RDF worlds-and avoiding the XSLT pilgrimage[C]// 5th European Semantic Web Conference. Berlin, Heidelberg: Springer, 2008: 432-447.
- [91] SCHARFFE F, BIHANIC L, KÉPÉKLIAN G, et al. Enabling linked data publication with the Datalift platform[C]// Workshops at the 26th AAAI Conference on Artificial Intelligence, 2012.
- [92] 熊晶, 焦清局, 刘运通. 基于多源异构数据的甲骨学知识图谱构建方法研究[J]. 浙江大学学报(理学版), 2020, 47(2): 131-141.
- XIONG J, JIAO Q J, LIU Y T. Oracle bone studies knowledge graph construction based on multisource heterogeneous data[J]. Journal of Zhejiang University(Science Edition), 2020, 47(2): 131-141.
- [93] 蒋秉川, 万刚, 许剑, 等. 多源异构数据的大规模地理知识图谱构建[J]. 测绘学报, 2018, 47(8): 1051-1061.
- JIANG B C, WAN G, XU J, et al. Geographic knowledge graph building extracted from multi-sourced heterogeneous data[J]. Acta Geodaetica et Cartographica Sinica, 2018, 47(8): 1051-1061.
- [94] 李涓子, 侯磊. 知识图谱研究综述[J]. 山西大学学报(自然科学版), 2017, 40(3): 454-459.
- LI J Z, HOU L. Reviews on knowledge graph research[J]. Journal of Shanxi University(Natural Science Edition), 2017, 40(3): 454-459.
- [95] 陈志云, 商月, 钱冬明. 基于知识图谱的智能答疑系统研究[J]. 计算机应用与软件, 2018, 35(2): 178-182.
- CHEN Z Y, SHANG Y, QIAN D M. Research on intelligent question answering system based on knowledge graph[J]. Computer Applications and Software, 2018, 35(2): 178-182.
- [96] LAO N, COHEN W W. Relational retrieval using a combination of path-constrained random walks[J]. Machine Learning, 2010, 81(1): 53-67.
- [97] 钟卓, 唐烨伟, 钟绍春, 等. 人工智能支持下教育知识图谱模型构建研究[J]. 电化教育研究, 2020, 41(4): 62-70.
- ZHONG Z, TANG H W, ZHONG S C, et al. Research on constructing model of educational knowledge map supported by artificial intelligence[J]. e-Education Research, 2020, 41(4): 62-70.
- [98] 张春霞, 彭成, 罗妹秋, 等. 数学课程知识图谱构建及其推理[J]. 计算机科学, 2020, 47(S2): 573-578.
- ZHANG C X, PENG C, LUO M Q, et al. Construction of mathematics course knowledge graph and its reasoning[J]. Computer Science, 2020, 47(S2): 573-578.
- [99] 杨洋, 邸一得, 刘俊晖, 等. 基于张量分解的排序学习在个性化标签推荐中的研究[J]. 计算机科学, 2020, 47(S2): 515-519.
- YANG Y, DI Y D, LIU J H, et al. Study on learning to rank based on tensor decomposition in personalized tag recommendation[J]. Computer Science, 2020, 47(S2): 515-519.
- [100] 马忠贵, 倪润宇, 余开航. 知识图谱的最新进展、关键技术和挑战[J]. 工程科学学报, 2020, 42(10): 1254-1266.
- MA Z G, NI R Y, YU K H. Recent advances, key techniques and future challenges of knowledge graph[J]. Chinese Journal of Engineering, 2020, 42(10): 1254-1266.
- [101] VIZCARRA J, NISHIMURA S, FUKUDA K. Ontology-based human behavior indexing with multimodal video data[C]// 2021 IEEE 15th International Conference on Semantic Computing, 2021: 262-267.
- [102] 由丽萍, 郎宇翔. 基于商品评论语义分析的情感知识图谱构建与查询应用[J]. 情报理论与实践, 2018, 41(8): 132-136.
- YOU L P, LANG Y X. Construction and query application of emotion knowledge graph based on semantic analysis of product reviews[J]. Information Studies: Theory

- & Application, 2018, 41(8): 132-136.
- [103] 马飞翔, 廖祥文, 於志勇, 等. 基于知识图谱的文本观点检索方法[J]. 山东大学学报(理学版), 2016, 51(11): 33-40.
MA F X, LIAO X W, YU Z Y, et al. A text opinion retrieval method based on knowledge graph[J]. Journal of Shandong University (Science Edition), 2016, 51(11): 33-40.
- [104] KANNAN A V, FRADKIN D, AKROTIRIANAKIS I, et al. Multimodal knowledge graph for deep learning papers and code[C]//Proceedings of the 29th ACM International Conference on Information & Knowledge Management, 2020: 3417-3420.
- [105] 郝林雪, 张鹏, 宋大为, 等. 融合知识图谱的查询扩展模型及其稳定性研究[J]. 计算机科学与探索, 2017, 11(1): 37-45.
HAO L X, ZHANG P, SONG D W, et al. Research on knowledge graph based query expansion model and its retrieval stability[J]. Journal of Frontiers of Computer Science and Technology, 2017, 11(1): 37-45.
- [106] TURING A M. Computing machinery and intelligence[J]. Mind, 1950, 59(236): 433-460.
- [107] WANG L. An improved knowledge graph question answering system for English teaching[J]. Mobile Information Systems, 2022. DOI: 10.1155/2022/3401074.
- [108] 余胜泉, 彭燕, 卢宇. 基于人工智能的育人助理系统——“AI好老师”的体系结构与功能[J]. 开放教育研究, 2019, 25(1): 25-36.
YU S Q, PENG Y, LU Y. An artificial intelligence assistant system for educating people: the structure and function of “AI educator”[J]. Education Research, 2019, 25(1): 25-36.
- [109] 祝智庭. CAI的教学策略设计(之四)[J]. 电化教育研究, 1998, 29(4): 49-52.
ZHU Z T. Teaching strategy design of CAI (Part 4)[J]. e-Education Research, 1998, 29(4): 49-52.
- [110] 唐烨伟, 茹丽娜, 范佳荣, 等. 基于学习者画像建模的个性化学习路径规划研究[J]. 电化教育研究, 2019, 40(10): 53-60.
TANG Y W, RU L N, FAN J R, et al. Research on planning of personalized learning path based on learner portrait modeling[J]. e-Education Research, 2019, 40(10): 53-60.
- [111] 沈杰, 乔少杰, 韩楠, 等. 融合多信息的个性化推荐模型[J]. 重庆理工大学学报(自然科学), 2021, 35(3): 128-138.
SHEN J, QIAO S J, HAN N, et al. Personalized recommendation model with multiple information fusion[J]. Journal of Chongqing University of Technology (Natural Science), 2021, 35(3): 128-138.
- [112] 吴昊, 徐行健, 孟繁军. 课程资源的融合知识图谱多任务特征推荐算法[J]. 计算机工程与应用, 2021, 57(21): 132-139.
WU H, XU X J, MENG F J. Knowledge graph-assisted multitask feature-based course recommendation algorithm[J]. Computer Engineering and Applications, 2021, 57(21): 132-139.
- [113] 高嘉骐, 刘千慧, 黄文彬. 基于知识图谱的学习路径自动生成研究[J]. 现代教育技术, 2021, 31(7): 88-96.
GAO J J, LIU Q H, HUANG W B. Research on automatic generation of learning paths based on knowledge graph[J]. Modern Educational Technology, 2021, 31(7): 88-96.
- [114] ZHU H, LIU Y, TIAN F, et al. A cross-curriculum video recommendation algorithm based on a video-associated knowledge map[J]. IEEE Access, 2018, 6: 57562-57571.
- [115] 师亚飞, 彭红超, 童名文. 基于学习画像的精准个性化学习路径生成性推荐策略研究[J]. 中国电化教育, 2019(5): 84-91.
SHI Y F, PENG H C, TONG M W. Research on generative paths recommendation strategies of precise personalized learning based on learning profile[J]. China Educational Technology, 2019(5): 84-91.
- [116] 李青, 闫宇. 新技术视域下的教学创新: 从趣悦学习到机器人陪伴学习——英国开放大学《创新教学报告》(2019版)解读[J]. 远程教育杂志, 2019, 37(2): 15-24.
LI Q, YAN Y. Innovating pedagogies with technology: from playful learning to learning with robots: introductions to innovating pedagogy report 2019[J]. Journal of Distance Education, 2019, 37(2): 15-24.
- [117] 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 E H, LIU Q, WANG S J, et al. Key techniques and application of intelligent education oriented adaptive learning[J]. CAAI Transactions on Intelligent Systems, 2021, 16(5): 886-898.
- [118] 卢宇, 薛天琪, 陈鹏鹤, 等. 智能教育机器人系统构建及关键技术——以“智慧学伴”机器人为例[J]. 开放教育研究, 2020, 26(2): 83-91.
LU Y, XUE T Q, CHEN P H, et al. A study on the system design and key technologies of an AI-driven educational robot: taking the “smart learning partner” as an example[J]. Open Education Research, 2020, 26(2): 83-91.
- [119] 汪时冲, 方海光, 张鸽, 等. 人工智能教育机器人支持下的新型“双师课堂”研究——兼论“人机协同”教学设计与未来展望[J]. 远程教育杂志, 2019, 37(2): 25-32.
WANG S C, FANG H G, ZHANG G, et al. Research on the new “double teacher classroom” supported by artificial intelligence educational robots: discuss about “human-machine collaboration” instructional design and future expectation[J]. Journal of Distance Education, 2019, 37(2): 25-32.