

基于知识图谱的自适应学习推荐系统的构建与应用

付成芳

(上海城建职业学院,上海 201415)

[摘要] 面临大数据时代的信息过载和知识迷航,教育资源的个性化、精准化推荐成了教育领域的研究热点。本文设计一种结合学习者特征和教育资源知识图谱的自适应学习推荐系统。本系统包含学习者模型、知识领域模型及自适应学习推荐三个模块。该系统通过分析学生的学习需求与兴趣,对学习者的画像,结合知识图谱的精准定位能力,评估学生对于知识点的掌握程度,结合自适应学习推荐算法为学生提供个性化、精准的学习资源推荐,从而帮助他们高效地进行网络学习,克服信息过载与学知迷茫的问题,提高自主学习的质量和效率。

[关键词] 知识图谱;教育资源推荐;自适应学习;智能推荐系统

[中图分类号] G642

[文献标识码] A

[文章编号] 2096-711X(2025)12-0160-04

doi:10.3969/j.issn.2096-711X.2025.12.055

[本刊网址] <http://www.hbxb.net>

众所周知,移动网络时代的“互联网+”学习以其独特的便利性和实时性优势,为学习者提供了随时随地的学习体验。然而,学习者在面对网络上海量且多样的课程学习资源时,常常感到信息过载与学知迷茫,无法有效筛选和定位适合自己的学习资源。尽管信息过载的问题可以通过应用搜索引擎技术来得到有效缓解,但搜索引擎所获取的信息往往千篇一律,难以满足学生个性化知识的需求。

随着5G、大数据和人工智能技术的发展,智慧教育在教育领域引起了深刻变革。2021年12月,由中国教育部主办的国际人工智能与教育会议提出,利用人工智能技术增进全民终身学习,加快教育数字转型和智能升级的步伐,支持引领建设高质量教育体系。近年来,知识图谱技术和自适应诊断测评等人工智能技术的突破,让在线学习环境下的自适应学习成为可能。与搜索引擎技术相比,自适应推荐系统能够从海量知识数据网络中筛选出学生掌握薄弱的重难点、可能感兴趣的内容,从而为学生提供更加多样化、个性化的学习方案。这种个性化的推荐系统不仅提升了学生的学习效率,同时也大大提高学生的学习积极性。

自适应推荐算法是推荐系统的核心,其基本理念在于通过深入分析学生的行为模式、知识点的属性信息以及学生与知识点对应的题目之间的历史互动记录,从而挖掘出学生和知识点各自的特征。基于这些特征,算法能够为不同的学生提供定制化的推荐服务。然而,传统的推荐系统在处理过程中,仅仅将学生与知识点之间的交互数据作为输入,这种做法可能会引发两个主要问题:首先,如果交互数据过于稀疏,即数据量不足,可能会导致模型出现过拟合现象,影响推荐质量;其次,所谓的冷启动问题,指的是当面对新用户或新知识时,由于缺乏足够的历史交互数据,算法难以做出有效的推荐。

数据稀疏和冷启动的问题曾长期存在,为了这些问题,部分学者将知识图谱(Knowledge Graph, KG)作为辅助信息

整合到推荐系统中的解决方案。知识图谱的引入能够丰富学生与知识点之间的交互数据,并且增强学生之间、知识点之间的关联信息。通过这种方式,推荐系统能够更准确地捕捉到学生的偏好,从而提升推荐的相关性和准确性。将知识图谱与推荐系统结合,以提升推荐质量,确实是目前推荐系统研究领域的一个热点。

因此,在人工智能和大数据技术应用的基础上,本文提出并设计一种基于知识图谱的教育资源自适应智能学习推荐系统,为满足以学生为主的个性化差异的未来发展带来了新的契机。

一、知识图谱在教育推荐系统中的应用

作为一种描述客观存在的实体、概念及相互关系,知识图谱凭借信息的可理解性、易用性和互动性,能够以更贴近人类思维方式的方式呈现互联网中的信息。知识图谱是由于一些相关技术相辅相成形成的,包含语义网、知识表示、网络、自然语言处理技术等,以及知识图谱等技术在众多领域的广泛应用。知识图谱在教育领域的应用亦呈现出卓越的成效。在推动教育革命方面,知识图谱的实际应用大部分聚集于教育领域中的知识库创设、特征建模、自适应引擎等核心领域。

教育资源知识库是用于存储、管理和共享各种教育资源的知识库,它包含但不限于教案、课件、视频、电子书、学习资源评论,以及各个学科和各个学段的知识点等内容。为提升准确性和可用性,知识图谱利用重新构建内容框架,把知识库内的资源划分到各个相关的知识节点,以便简洁、高效地利用库中的资源。在连接知识点后,学生将通过自适应学习方法进入下一个学习过程。特征建模是根据学习者的学习行为和知识水平构建出精准的学生画像。而自适应引擎是指利用知识图谱和智能算法诊断分析后推送所需资源给学习者,学习者凭借已然建立的自适应学习系统将得到个性化学习资源,系统为学生确定下一步应该学习的知识内容。作

收稿日期:2024-11-8

基金项目:本文系中国建设教育协会教育教学科研项目(项目编号:2023384),全国高等职业院校信息技术课程教学改革研究项目(项目编号:2024177)和上海城建职业学院校级科研项目(项目编号:cjky202401)的阶段性成果。

作者简介:付成芳(1979—),女,江苏泗阳人,上海城建职业学院教授,主要从事知识图谱、推荐系统、智能控制技术的研究。

为智能教育的核心,知识图谱可以实现建构海量学习资源的分类整理形成知识网络、学习者协同学习、自适应学习、个性化推荐等关键应用。

二、基于知识图谱的自适应学习模型框架

本文通过可视化教育工具——知识图谱,以学生个性差异为基础,以促进学生个性发展为目标,以构建知识图谱为主要方式,从而为学生的自适应性学习提供个性化学习路径。基于知识图谱的自适应学习系统的总体框架有学习者模型、领域知识模型、自适应引擎模块三大模块。

(一)构建多角度的学习者模型

在自适应学习系统中,学习者模型主要是对学习者的个性化特征的分析,以便于为学习者提供更加个性化、有效的学习资源和学习路径,是自适应学习的基础。从研究热点来看,对于自适应学习系统中用户情感和态度、社会性交互环境、教育信息安全等方面的关注度有待提高。倘若研究视野仅限学生认知能力而忽略他们的情感倾向和情感态度,毋庸置疑会极大影响个性化教育和学生全面发展。本设计是一种考虑学生的学习风格、学习水平、兴趣爱好、师生信息安全等因素的学习者模型,以此为基础构建多角度的自适应模型。

多角度的学习者模型考虑的学习者的数据资源有以下几个角度:

1. 学习者基本信息

当学习者首次登录或注册一个教育系统时,收集其基本信息是非常重要的,这有助于创建一个基本的学习者档案,为个性化学习体验打下基础。以下是一些常见的信息收集点:学号,姓名,年龄,性别,学校,专业,入学年份,联系方式,密码等。收集这些信息通常通过在线注册表单进行。重要的是,系统应当确保遵守数据保护法规,对收集的个人信息进行安全存储和隐私保护。此外,注册过程中应清晰地告知学习者其信息将如何被使用,并获得他们的同意。

2. 风格的构建

首先,学习目标、先前的学习经历、兴趣爱好、技术熟练度等信息收录。其次,视觉性信息。平台针对各个信息的点击量和学习工具实际使用情况,对学生的视频、图像、PPT等视觉资源的敏感度进行有效分析。再次,言语性信息。参照收集视觉性信息的技术原理,着眼于对学生的文本、音频信息的敏感度分析。最后,以问卷调查为主要形式来构建风格。

3. 知识储备的建构情况

利用“定级测试”的方法,对学生特定领域知识内容的学习进展进行有效评估。为了提高学习成果考核的全面性,考核内容大致设定为能力水平、访问状态、测试状态、学习状态、掌握状态等方面,以提升“判断”的准确性。

系统在详细地掌握学习者的学习风格和知识体系后,建立多角度的准确的学习者模型,能够使后续的推荐更加精确和高效。

(二)基于知识图谱和 TF-IDF 算法建立领域知识模型

在构建领域知识模型的过程中,首先收集并整理学习者的交互行为数据,综合学习者的教育水平、知识结构、爱好倾向等详细信息为学习者精准画像。其次采用命名实体识别和关系抽取分别构建学习者的社交网络和学习者知识获取

网络,形成在线自适应学习系统的知识图谱,最后对该图谱进行向量化表示,得到实体向量,构建自适应学习系统的领域知识模型。

1. 知识图谱技术

知识图谱的构建是一个技术驱动的过程,它涉及从不同类型的数据源中提取有用信息,并将这些信息以结构化的形式存储于数据库中。这个过程包括处理结构化数据、半结构化数据和非结构化数据,以抽取有价值的知识。

知识图谱的构建不仅仅是一次性的活动,而是一个持续的、动态更新的过程。每一次的迭代更新都是对现有知识图谱的补充和完善,确保其反映最新的信息和知识状态。

信息抽取 (Information Extraction: IE) 是从非结构化文本信息中提取出结构化信息的技术。主要基于 D2R 技术,通过数据预处理、实体识别、关系抽取、形成三元组等步骤,将存储在数据库的结构化数据转化成为若干富有语义关系的三元组,进而构建出详细的知识图谱;使用爬虫技术配合 XPath 解析器将 Web 环境里半结构化的数据转化成想要的目标数据。

2. 知识图谱建构

(1) 数据源获取与处理

本系统的数据来源可归纳为三大类:第一类为存储于数据库中的结构化数据;第二类为散布于网页中的半结构化数据;第三类为从书籍、文本中提取的非结构化数据。本系统利用 Scrapy 框架从学习通、MOOC、智慧职教等网络资源平台中得到课程的资源信息。爬取教材的目录以及有关链接,爬取知识信息并加以数据清洗生成 csv 的格式,为之后导入图数据库构建知识图谱做准备工作。

(2) 抽取实体

在本系统中,TF-IDF 算法主要应用于评估知识点关键词在资源获取过程中的重要性。以一个实例来阐释:若某个知识点关键词 K 在文档 Q 中出现频率较高,这表明关键词 K 对于文档 Q 具有极高的重要性,并且对文档 Q 具有很强的代表性。

TF-IDF 算法的核心理念在于,如果一个词语在特定文档中出现频率较高,而在其他文档中出现频率较低,则该词语具有较强的区分能力,有助于识别关键信息。TF (Term Frequency, 词频) 指的是一个词在文档中出现的次数,通过统计得出。IDF (w) 用于表示逆文档频率。TF-IDF 算法通过将 TF (词频) 与 IDF (逆文档频率) 相乘来计算一个词在文档中的重要性。如果一个知识点关键词的 TF-IDF 值相对较高,这通常意味着此知识点对于该文档具有显著的重要性。

(3) 抽取实体关系

在针对《智能控制系统与工程》课程的学习过程中,各个知识点之间存在多种联系和属性,本文主要探讨了相关性、前置性和后置性、包含性、相关性和同义性这五种属性。任意两个知识点之间只能存在一种关系,新导入的关系将覆盖旧关系,关系设置总数上限为 2000 个。相关性是指已掌握的知识点 A 与新知识点 B 之间的联系,这种联系有助于学生在学习 B 时回忆和联系 A 的内容,从而促进复习和理解。后置性则描述了知识点 A 和 B 之间的逻辑顺序,如果学习 B 之前必须先掌握 A,那么 A 就被视为 B 的先决条件,两者之间存

在后继关系。相关的关系主要用来描述各个知识点之间的相关性。包含关系用来描述知识章节和知识点之间的层级关系。同义关系用于表示具有相同含义的知识点。

(4) 抽取知识点的属性和特征

平台中的每个知识点都被标注相应的属性和特征,以便学生更好地理解。标注的内容包含各知识点的中英文名称、定义、关键词、实际案例、运用情境等。通过详尽地描述各个知识点的属性及其特征,有利于学生透彻理解并牢固掌握知识。

本图谱的知识点还可以定义标签,固定标签包括:重点、难点、考点、思政知识点,可根据需要自定义标签。

对知识点认定认知维度,包括:记忆、理解、应用、分析、评价、创造,每个知识点填入一个认知维度。

同时对知识点进行分类,包括:事实性、概念性、程序性、元认知,每个知识点只能填入一个分类。

(5) 知识图谱可视化

本系统将提取的知识点及其相互关系进行存储,采用的是 Neo4j 图数据库。Neo4j 的优势在于其能够创建复杂的关系,实现数据之间的深度关联。与传统的关系型数据库相比较,使用 Neo4j 无需提前定义表结构,可以方便、灵活地添加学生、试题等字段信息。在 Neo4j 中,一般使用 Cypher 作为查询语言,它允许用户以声明式的方式查询和操作图数据。《智能控制系统与工程》课程的知识图谱存储结构,可以通过 Neo4j 进行直观展示。

(三) 自适应引擎

自适应引擎是一种在自适应学习系统中使用的关键技术,它是自适应学习系统的核心部分,根据学习者的知识状态信息和模型匹配关联领域知识模型,自动地调整推荐相匹配的学习路径和学习内容,为每一位学习者量身定制个性化深层次的知识学习方案,最终达到个性化学习路径、个性化学习资源、个性化学习策略的自适应,使每一位学习者学有所成。

在具体实现过程中,自适应引擎先从不同维度对学习者的进行画像,收集学生的答题数据,系统能够评估学生对各个知识点的掌握程度。结合具体的推荐算法,系统能够生成一个针对性的试题推荐序列,并适时调整推荐的知识点和试题。学生可以通过点击推荐按钮来查看推荐的试题详情。当学生选择开始作答,系统便会提供这些推荐试题供学生练习。

三、基于知识图谱自适应学习推荐算法设计

(一) 算法设计

本自适应推荐系统的算法是基于知识图谱和协同过滤算法进行推荐的,其实现过程如下:

针对新用户:对于刚开始使用系统且还没有保存任何测试记录的学生,采用知识图谱中的知识体系结构来推荐热门试题。这些试题是系统根据历史数据,即那些被推荐给学生次数最多的题目来选择的。利用知识图谱的可解释性,这种方法能够在一定程度上有效解决新用户所具有的“冷启动”问题。

针对有记录用户:对于已经有一定做题记录的学生,分析他们以往的测试数据。结合知识图谱,系统能够精准判断学生对于以往所学知识点的掌握程度,并据此推荐难易度适

中、知识点匹配的试题。

相似度计算:将拟推荐的试题与学生以往的错题进行相似度计算,以找到与错题在结构和内容上相似但有所不同的题目,帮助学生从不同角度理解和掌握知识点。

考虑学生喜好:系统还会考虑学生对试题风格的个人偏好,比如题目类型、难度级别等,以生成一个符合学生个性化需求的试题推荐列表。

个性化推荐生成:最终,系统会基于上述分析生成一个个性化的试题推荐列表,旨在提供给学生最有帮助、最符合他们当前学习需求的试题。

通过这种基于知识图谱的个性化推荐方法,系统能够为每位学生提供定制化的学习资源,不仅帮助他们巩固已学知识,还能激发他们对新知识的学习兴趣。

(二) 自适应推荐系统的实现

本系统使用 neo4j 作为课程的试题资源知识图谱生成软件,并基于 Spring 框架搭建个性化学习推荐系统的基础环境,使用 Mybitias 将后台接收到的数据保存到存储设备或数据库中。由于 Maven 是项目构建管理的工具,易于项目研发过程中的管理。Spring + SpringMVC + Mybitias 已十分成熟,使用 Maven 里众多的库,消除了手动导入 jar 包的复杂性,有效地简化操作:数据库用的是 Mysql+neo4j。

采用协同过滤推荐算法,并与知识图谱相结合对于知识进行推荐。构建图形化界面,对于整个系统实现可视化,最终构建系统并实现包括推荐算法、知识查询、在线测试和用户信息管理的基础功能。

最后,将系统的各个模块综合在一起进行系统集成测试,用来检验系统的稳定性和可用性。

四、结语

自适应学习时代,学习者面临学习资源的过载和学知迷茫的挑战,为使学习者从海量的学习资源中快速找到适合自己的学习内容,满足学习个性化的需求,本文基于大数据驱动视角,结合学习者特征,提出一种基于知识图谱和自适应协同推荐算法的教育资源自适应学习模式总体框架、研究内容以及应用案例。文章通过构建基于知识图谱的语义化知识体系结构,使学生能够明确每个知识点在体系中的位置以及它们之间的相互联系。根据学生对知识点掌握的不同程度,知识图谱根据学习者的学习习惯和对知识的掌握情况。利用知识图谱的精准定位与推荐,为学生提供个性化的学习路径,指导其下一步的学习内容。此外,根据每位学生的实际情况,系统还能匹配不同难度的试题,帮助他们逐步提升学习效果。

参考文献:

[1] 中华人民共和国教育部. 创新技术服务教学进步,人工智能助力未来教育[EB/OL]. (2021-12-8) [2022-1-1]. http://www.moe.gov.cn/jyb_xwfb/gzdt_gzdt/moe_1485/202112/t20211208_585822.html.

[2] LI Z, LIU H, ZHANG Z, et al. Learning Knowledge Graph Embedding with Heterogeneous Relation Attention Networks[J]. IEEE Transactions on Neural Networks and Learning Systems, 2021, 33(8): 3961-3973.

- [3]胡殿侠.数据驱动下高校图书馆智慧服务的构建策略[J].湖北开放职业学院学报,2024,37(18):177-181.
- [4]王慧欣,童向荣.融合知识图谱的推荐系统研究进展[J].浙江大学学报(工学版),2023,57(8):1527-1540.
- [5]孙嘉奇.基于知识图谱个性化学习推荐系统的研究与实现[D].沈阳:辽宁大学,2020.
- [6]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.
- [7]朱艳茹.基于知识图谱的自适应学习系统设计与实践[D].长春:吉林大学,2018.

Construction and Application of Self-adaptive Learning Recommendation System Based on Knowledge Graph

FU Cheng-fang

(Shanghai Urban Construction Vocational College, Shanghai 201415, China)

Abstract: Facing the information overload and knowledge disorientation in the era of big data, the personalization and precision recommendation of educational resources have become a research hotspot in the field of education. This paper designs an adaptive learning recommendation system that combines learner characteristics and the knowledge graph of educational resources. The system includes three modules: learner model, knowledge domain model, and adaptive learning recommendation. By analyzing students' learning needs and interests, the system creates a profile of learners, combines the precise positioning capabilities of the knowledge graph to assess students' mastery of knowledge points, and uses adaptive learning recommendation algorithms to provide personalized and precise learning resource recommendations for students. This helps them to efficiently engage in online learning, overcome the issues of information overload and knowledge confusion, and improve the quality and efficiency of self-directed learning.

Key words: knowledge graph; educational resource recommendation; adaptive learning; intelligent recommendation system

(责任编辑:章樊)

(上接第159页)

- [11]蔡雨玲,唐瑜梳,杨溢,等.大学生就业核心竞争力结构模型及评价指标体系的构建[J].现代商贸工业,2021,42(28):68-69.
- [12]路健,李鹏,康辉英.独立学院毕业生就业竞争力评价指标体系研究[J].合作经济与科技,2020(19):140-141.
- [13]高苗苗.高校音乐专业毕业生就业竞争力评价体系的构建[J].文教资料,2020(14):147-148,225.
- [14]李迅.AHP下的高职院校学生就业竞争力评价体系研究[J].创新创业理论与实践,2020,3(4):6-8.
- [15]王金龙,刘建锋,倪伟.新工科视域下大学生就业竞争力评价体系研究[J].中国大学生就业,2019(23):34-39,52.

Reconstruction of the Evaluation System for the Employment Competitiveness of New Business Students in Higher Vocational Colleges from the Perspective of Ingenuity Leading Digital Intelligence Empowerment

RUAN Zhi-hua¹, LIU Wen-hua²

(1. School of Accounting and Finance, Taizhou Vocational College of Science and Technology, Taizhou Zhejiang 318020;
2. School of Economics and Management, Yichun University, Yichun Jiangxi 336000, China)

Abstract: At present, the new business talents in higher vocational colleges have some problems, such as the mismatch between the demand of "craftsman spirit and digital intelligence" of enterprises and the ability of new business in higher vocational colleges, the lack of craftsman spirit of higher vocational colleges, the lack of team spirit, the lack of personal quality, career planning and positioning of graduates, etc. On this basis, the employment competitiveness evaluation system of new business students in higher vocational colleges from the perspective of ingenuity leading digital intelligence empowerment should be reconstructed. Using SPSS principal component factor empirical test, the evaluation system of new business employment competitiveness of higher vocational students from the perspective of ingenuity leading digital intelligence empowerment is composed of four main factors: knowledge literacy, ability literacy, application literacy and comprehensive vocational ability. On this basis, countermeasures and suggestions are put forward to enhance the employment competitiveness of higher vocational students from the perspective of ingenuity leading digital intelligence empowerment.

Key words: craftsman spirit; digital intelligence empowerment; new business science; competitiveness of employment

(责任编辑:陈思婷)