



# 如何帮助学习者走出学习资源迷航\*

## ——基于学习者画像的个性化学习资源推荐

成亚玲 谭爱平

(湖南工业职业技术学院, 湖南 长沙 410208)

**摘要:**以MOOCs为代表的在线课程为学习者提供了多模态海量的学习资源,但学习者在选择资源时也经常面临信息过载、资源内容和学习需求不匹配等问题,造成学习资源迷航的困境。如何高效精准地向学习者推荐合适的学习资源与自适应的学习路径,已成为当前教育信息化亟待解决的问题。基于学习者画像的个性化学习资源推荐模型,从个性化资源推荐的现实需求出发,以现代教育理论和数字技术为支撑,探究个性化学习资源推荐过程中,学习者的状态与学习需求等个性化特征诊断、学习资源自身属性特征参数挖掘与表征、学习资源与学习需求精准匹配等问题,旨在向学习个体和学习群体推荐个性化学习资源。将基于学习者画像的个性化学习资源推荐模型应用于教学实践,结果表明,与基于协同过滤推荐、DINA认知诊断推荐相比,该学习资源推荐模型具有更高的精准度和更优的可解释性,能够为不同学习需求的学习者推荐适切的学习资源;同时该模型也在一定程度上缓解了冷启动与数据稀疏性等问题。

**关键词:**在线开放课程;学习者画像;学习资源;深度学习;推荐模型;数字化

**中图分类号:**G720 **文献标识码:**A **文章编号:**1674-9154(2023)02-0103-10

**DOI:**10.16851/j.cnki.51-1728/g4.2023.02.019

### 一、研究背景

随着教育信息化进程的不断推进,国家不断从顶层架构层面出台系列政策文件,支持个性化的自适性服务,如《教育信息化十年发展规划(2011—2020年)》提出“努力为每一名学生和学习者提供个性化学习、终身学习的信息化环境和服务”;<sup>[1]</sup>《教育信息化2.0行动计划》提出要“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”;<sup>[2]</sup>《中国教育现代化2035》提出要“利用现代化技术加快推动人才培养模式改革,实现规模化教育与个性化培养的有机结合”。<sup>[3]</sup>提供个性化的教育服务是我国现阶段

教育现代化的重要任务之一。近年来,随着新一代信息技术与教育行业的深度融合发展,以慕课为代表的在线开放课程为学习者提供了海量的学习资源。然而,学习者在海量学习资源中也常常面临信息过载、资源内容和学习需求不匹配等问题,造成资源迷航的困境。因此,如何高效精准地向学习者推荐合适的学习资源与自适应的学习路径,已成为当前教育信息化亟待解决的问题。本研究以教育大数据、数据挖掘等技术为支撑,从学习者基本信息、知识掌握情况、学习行为等维度,构建学习者画像并对其知识掌握情况进行诊断,同时结合学习者的个性化特征和学习行为特点向其推荐适切的学

\*基金项目:2019年湖南省“十三五”教育科学规划课题“人工智能支持下的个性化学习资源推荐与学习路径规划研究”(编号:XJK19BXX007)。

作者简介:成亚玲(1981—),湖南工业职业技术学院信息工程学院,副教授,硕士,研究方向:课程与教学论,数据挖掘;谭爱平(1979—),湖南工业职业技术学院信息工程学院,副教授,硕士,研究方向:数据挖掘。

收稿日期:2023-01-12



习资源,促进其学习效率和学习成效双提升。

## 二、个性化学习资源推荐研究现状

### (一) 概念界定

“个性化推荐”概念最早由 Resnick 和 Varian 提出,认为个性化推荐是根据用户的个性化需求,利用计算机系统向其推荐相关信息。<sup>[4]</sup>姜强等认为个性化学习资源推荐是基于对学习者的认知风格等个性化特征和知识掌握情况等综合分析后,将学习资源进行重组排序,从而使学习资源更适合学习者的认知能力与学习风格,帮助学习者进行知识构建。<sup>[5]</sup>赵晋认为个性化学习资源推荐是在明确学习者学习需求和偏好的前提下,根据用户对资源的评价信息,对学习资源进行多维度优先级排序,再向学习者进行推荐。<sup>[6]</sup>赵蔚等认为个性化学习资源推荐是根据学习者特性、需求向其呈现适切的学习资源与学习路径,从而促进其学习目标达成。<sup>[7]</sup>尽管国内外学者对个性化学习资源推荐的内涵进行了多种描述,但其核心要义是一致的,即个性化学习资源推荐主要包括三个核心内容:学习者特征、推荐对象和推荐算法。学习者特征是根据学习者的基本信息、学习行为挖掘出其学习风格、偏好和学习需求并进行参数化表示;推荐对象包含 MOOCs 视频、文本等学习资源和学习路径,同时需要将其进行参数化表示;推荐算法是指根据学习者的个性化需求特征和推荐对象之间的关系,向学习者推送适切的学习资源,以提高其学习成效。因此,推荐算法是衔接两者的重要桥梁。

### (二) 推荐算法及模型构建

#### 1. 推荐算法

常用个性化学习资源推荐算法包括内容推荐、<sup>[8]</sup>协同过滤推荐、<sup>[9]</sup>混合推荐三种,<sup>[10]</sup>其中协同过滤推荐算法应用最为广泛。协同过滤推荐算法是通过分析与目标用户相似的用户进行推荐,即找到与目标用户相似的用户群组,将相似用户偏好的学习资源、学习路径等推荐给目标用户。EduRank 利用协同过滤推荐算法筛选出与目标学生相似度最高的学生,然后将相似学生的学习内容、练习试题、学习路径等推荐给目标学习者。<sup>[11]</sup>姜强等将具有相

同或相近知识结构水平、学习偏好的学习者进行分类,并进行学习者特征与学习对象媒体类型、理解等级、难度级别的匹配计算,从而生成精准个性化学习路径,为差异化教学提供了新思路。<sup>[12]</sup>赵宁对协同过滤推荐算法进行了优化,充分利用学习者的用户属性信息,计算学习者之间的相似性并形成最佳的邻居集合,进而向目标用户推荐与其兴趣爱好相近的邻居所评价过的学习资源。<sup>[13]</sup>近年来,随着云计算和大数据技术的发展,基于深度学习的推荐算法逐渐兴起。基于深度学习的推荐算法主要是通过数据样本进行训练、提取特征属性来提升预测或分类的准确性。兰明祥利用深度神经网络技术设计了基于相似度排序的资源过滤算法,从而向学习者进行个性化推荐。<sup>[14]</sup>文孟飞等利用深度学习和支持向量机(Support Vector Machine, SVM)相结合的方式对视频学习资源的个性化推荐,从而大大提高了视频资源的获取率和利用率。<sup>[15]</sup>任蓓蓓等人通过深度学习挖掘用户特征、时间关系、资源参数之间的关系,大大提高了学习者与资源间的匹配度,从而实现了数字图书馆资源的个性化推荐。<sup>[16]</sup>

#### 2. 模型或系统构建

目前个性化学习资源推荐模型主要有基于协同过滤算法的传统模型、矩阵分解模型、逻辑回归模型、特征交叉模型和深度神经网络推荐模型等。JieLu 等人开发了在线学习资料个性化推荐系统,通过学习者历史浏览数据,挖掘其偏好模型并引入内容过滤,优化了推荐效果。<sup>[17]</sup>王晓东等将学习者的学习水平和学习风格等特征融入推荐过程,构建了学习者和学习资源知识相融合的学习资源推荐模型,其个性化与精准度超过了传统的协同过滤算法。<sup>[18]</sup>李浩等针对移动学习时间碎片化、学习环境动态性等特点,设计了基于学习者时空特征的移动学习资源推荐模型,有效解决了由于学习用户数据量过大而导致的用户相似度计算过慢的问题,提高了移动学习情境下学习资源推荐的精准度。<sup>[19]</sup>王奕利用概率矩阵分解模型,从用户图书借阅历史记录中学习低维的近似矩阵,从而实现个性化的图书推荐。<sup>[20]</sup>常诗卉在内容推荐和协同过滤推荐策略的基





础上,融合逻辑回归精排序模型,并添加交叉特征,以此提升推荐结果的个性化程度。<sup>[21]</sup>字云飞等提出了一种基于多用户—项目结合的深度神经网络个性化推荐模型,该模型利用深度神经网络学习、抽取数据特征,再融合协同过滤中的广泛个性化产生候选集,然后通过二次模型学习产生排序集,实现精准、实时、个性化推荐。<sup>[22]</sup>近年来,基于认知诊断模型(Cognitive Diagnosis Models, CDMs)的个性化学习资源推荐备受关注。认知诊断模型是基于项目反应理论(Item Response Theory, IRT)框架并融入认知变量,对学习者的认知过程、知识加工技能或知识结构等方面进行诊断与评价。当前常用的认知诊断模型主要有规则空间模型(Rules Space Model)、决定型输入、噪声与门模型(Deterministic Inputs, Noisy and Gate Model, DINA)、融合模型(Fusion Model, FM)、多维项目反应理论模型(Multidimensional Item Response Theory, MIRT)等。<sup>[23]</sup>其中, DINA 模型最为经典,应用也最为广泛。DINA 模型利用学习者的答题历史记录,引入试题—知识点关联矩阵(Question Knowledge matrix, 简称Q矩阵),将学生的认知状态建模为一个多维的知识点掌握向量,建模得到的认知状态具有极高的可解释性。赵宇丹提出基于DINA模型的教师认知诊断模型,通过引入潜变量和滑动矩阵,使DINA模型能够实现多级评分,从而为学习者提供个性化学习资源。<sup>[24]</sup>李全等基于学生知识点掌握信息,利用DINA认知诊断模型对学生、试题和知识点三者信息进行联合概率矩阵分解,最后根据难度范围进行试题推荐。<sup>[25]</sup>另外,基于图模型的个性化推荐是近期业界和学界的研究热点。<sup>[26][27]</sup>

综上,基于协同过滤推荐算法的推荐能够有效降低模型构建的复杂性,但存在忽略学习个体的知识点掌握情况和知识点之间存在的关系,且存在矩阵稀疏和冷启动等问题。基于深度学习的推荐能够有效降低学习资源推荐过程中对模型的依赖性,但存在对推荐结果的解释性不强等问题。基于DINA认知诊断的推荐能够以多维向量准确反映学习者的知识掌握情况,具有较好的可解释性,但由于只关

注个体学习状态诊断,忽视相似学习者的共性特征,对数据隐含的信息挖掘不够等,容易造成建模模型过大和误差偏大等问题。为克服上述不足,借鉴协同过滤推荐算法和DINA认知诊断推荐模型,本研究构建了基于学习者画像的个性化学习资源推荐模型,该模型基于学习者画像动态获取学习需求,根据学习者个性化参数、学习资源参数等信息匹配,计算出适合学习者当前学习需求和偏好的资源推送给学习者,大大提高了推荐结果的准确性与可解释性。

### 三、基于学习者画像的个性化学习资源推荐模型构建

在学习者画像个性化学习资源推荐模型中,学习者画像是个性化推荐模型的基础,通过画像诊断出学习者的认知水平和未掌握的知识点;通过画像挖掘提取出学习者学习风格与偏好。学习资源匹配计算模块是个性学习资源推荐模型的核心,通过多个参数(未掌握的知识点、“知识点—学习资源”关系、学习风格与偏好等)精准匹配,计算并筛选出适合学习者当前认知水平与学习偏好的各类学习资源。学习资源推荐模块根据相应的策略向学习者个体、学习群体推荐适切的学习资源(见图1)。

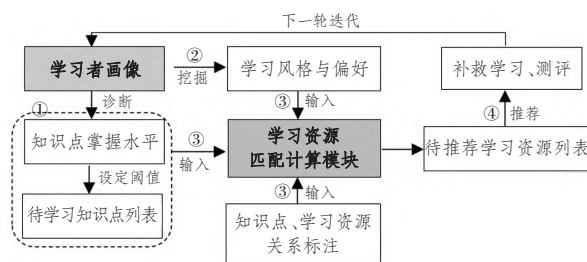


图1 基于学习者画像的个性化学习资源推荐模型

#### (一) 学习者画像构建

学习者画像概念源于商业领域的“用户画像”。用户画像最初由库珀(Alan Cooper)提出,是对用户的基本属性、偏好特征、行为特点等用户数据进行语义化的标签表示,其对用户全貌进行刻画、细化用户特征,能够为企业实现精准化服务提供管理决策参考。<sup>[28]</sup>用户画像技术最初在医疗、商业等领域有着广泛的应用,改进了产品服务质量和提升用户体验。近年来,随着学习分析、人工智能



等技术在教育领域的深度应用,用户画像概念与技术也被引入教育领域。学习者画像的本质在于收集学习者基本属性数据、学习过程全样本数据,并对这些数据进行处理分析,从而生成数字化标签和知识体系。

学习者画像构建流程通常包含明确画像目标、数据采集与处理、构建标签体系、生成学习者画像库、画像服务输出五个阶段(见图2)。当然,针对不同的画像目标,画像建模的维度也不尽相同。本研究构建学习者画像的目标是对学习者学习状态进行精准识别、对学习者的认知水平进行精准诊断,以便后期对学习者的个性化学习资源推荐。本研究主要从学习者基本属性、知识掌握情况、学习行为三个维度构建学习者画像,由于本研究团队

在《基于学习者画像的在线开放课程学习预警研究》一文中,曾对学习者的构建流程做过详细表述,<sup>[29]</sup>因此,这里主要介绍学习者画像标签体系的构建流程(见图3)。

### 1.原始数据采集

全面持续采集学习者全样本数据是学习者画像构建的基础。原始数据采集主要是采集学习者基本信息、历史学业数据、学习过程数据和学习结果数据等。考虑到线下实体课堂中师生教学互动等行为数据的采集存在诸多条件限制,本研究暂且仅对线上课堂数据进行全面采集。其中,基本信息数据包括学习者的人口属性、部分学习风格等数据,主要通过在线开放课程用户注册、调查问卷等进行采集;历史学业数据主要通过学校学业综合成绩管理

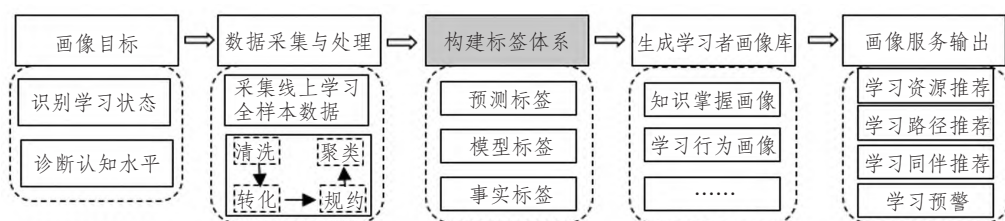


图2 学习者画像构建流程

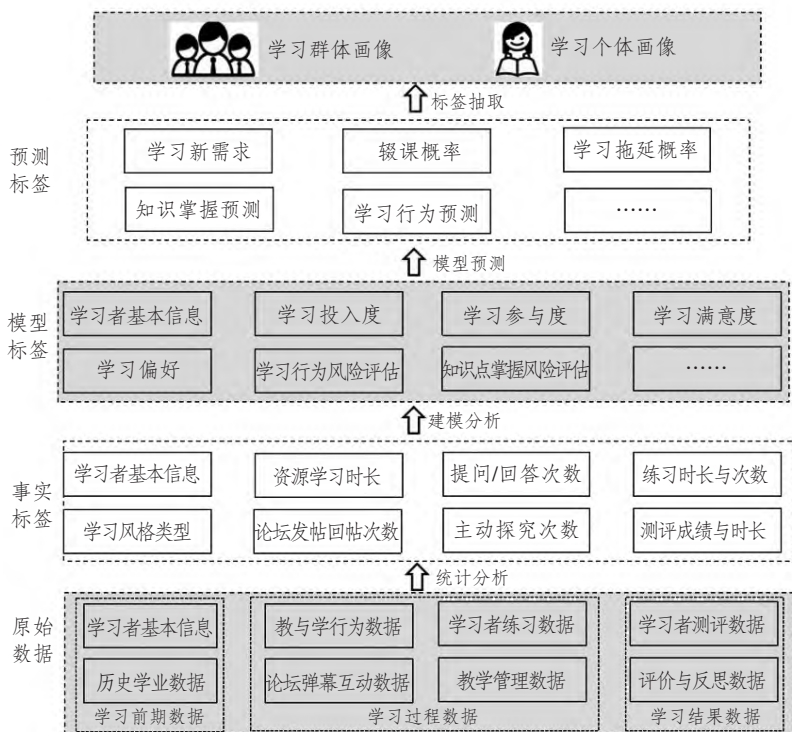


图3 构建学习者画像标签体系流程



信息系统进行采集;学习过程数据包括在线开放课程学习过程中教与学行为数据、教学管理数据等,即师生或生生互动数据、论坛交流、探究过程等数据;学习结果数据包括学习者在线做练习、在线即时测评和期末终结性考核评价等数据。

### 2.事实标签处理

事实标签处理是对原始数据进行统计分析。这里的事实标签主要指静态事实标签,包括学习者基本信息、学习风格类型等,具有稳定性。比如,学习风格反映学习者如何感知学习环境并做出交互行为,可以使用 Felder-Silverman 量表从信息感知(感悟型、直觉型)、信息输入(视觉型、言语型)、信息加工(活跃型、沉思型)和内容理解(序列型、综合型)四个方面对学习者的学习风格进行量化与统计,<sup>[30]</sup>也可以通过对学习者查阅学习资源类型、学习时长、浏览资源次数、论坛发帖或回帖数量、做练习次数与时长、测评成绩与时长等方面的学习行为数据进行统计分析。

### 3.模型标签处理

模型标签处理是在静态事实标签处理的基础上进行建模分析。可以从学习者的课堂出勤、互动、资源利用、探究过程等方面,对学习者的学习投入度、参与度进行标签化处理;可以从学习行为轨迹、课堂互动、作业按时完成等方面,对学习者的学习拖延与辍课风险进行标签化处理;可以通过练习效果与完成时间、即时测评成绩等方面,对学习者知识掌握情况进行标签化处理。值得注意的是,人口属性、学习偏好等标签指标基本保持稳定,不需要做过多的处理。

### 4.预测标签处理

预测标签处理是从模型标签库中根据具体需要抽取某些标签对学习者某一方面情况进行预测,如抽取学习投入度、参与度、学习满意度等模型标签并进行分析,可以对学习者的辍课情况进行预测;抽取知识点掌握风险评估、学习满意度、学习偏好等标签分析,可以预测学习者的掌握情况和学习新需求;抽取学习行为风险评估、学习投入度等标签可以预测学习者的学习拖延概率;抽取学习投入

度、参与度、学习偏好等标签对学习者的群体聚类分析,可以预测不同学习群体的学习新需求等。通过上述标签体系构建流程,本研究最终根据画像目标构建出学习群体画像和学习个体画像。

### (二)基于学习者画像确定未掌握的知识点

学习者画像中全程记录了学习者的学习过程和学习结果,特别是对学习者做练习、测评、资源学习、探究互动等学习行为和学习结果做了详细记录并进行了深度剖析。通过抽取学习者参与度、投入度、知识掌握等标签生成画像,挖掘学习者做练习、即时测验和终结性测验等数据,诊断出学习者知识点的掌握水平。学习者对某个知识点的掌握水平可以分为基本认知(L1)、夯实基础(L2)和拓展应用(L3)三个级别。其中总分设定为1,基本认知水平表示学习者对该知识点有了基本的理解,其评分为(0-0.6)之间;夯实基础水平表示学习者对该知识点有了较好的掌握,其评分为[0.6-0.8)之间;拓展应用水平表示学习者已经很好地掌握了该知识点,并且可以将知识进行迁移应用,其评分为[0.8-1)之间。根据学习者对某个知识点掌握水平(L1、L2、L3),下游推荐学习资源时对应匹配资源的内容类型(R1:概念理解,R2:知识巩固,R3:综合应用)。通过学习者画像可以将学习者未掌握的所有知识点筛选出来,形成待学习知识点列表。

### (三)学习资源匹配计算

在学习资源与学习需求匹配计算前,首先需要构建课程知识体系,然后标注“知识点—学习资源”间的关联关系。知识点是课程知识体系中最基本的单元,也是教与学过程中传递教学信息的最小单元。每个知识点都有相关属性,且知识点之间也存在一定的关联。以数据结构课程为例,该课程学习内容逻辑性较强,且偏重实践应用,同时还需面向社会人群提供学习服务,所以首先应确定数据结构课程知识构建的目标和知识点范围,并对知识点进行分类;然后根据知识管理理论和数学领域知识层次结构思想,采用自上向下的设计方法,在课程开发专家指导下反复论证构建出树状层级结构课程





知识体系。

通常情况下,知识点与学习资源之间是一对多的关系。由于数据结构课程知识点以及学习资源比较丰富且类型多样,本研究采用人工标注+机器学习标注相结合的方式,对知识点与学习资源之间的关系进行标注。同时,学习资源本身也包含了难度水平、关联程度(与知识点关联度、与其他学习资源关联程度)、媒介类型(文本、图片、音视频)和内容类型(概念理解、知识巩固、综合应用)等信息。通过标注,在知识点与学习资源两者之间建立基于语义的深层映射关系,为后续根据学习者某个知识点的掌握程度进行个性化学习资源推荐做好铺垫。将学习者的知识掌握水平、学习风格、知识点—学习资源关联关系三个变量因素输入到学习资源匹配计算模块中,与课程资源池中的学习资源进行匹配计算,进而将筛选出来的学习资源推荐给学习者。

#### (四) 个性化学习资源推荐

##### 1. 学习个体个性化学习资源推荐

具体推荐流程如下:步骤一是诊断个体知识点掌握水平,确定未掌握知识点。基于学习者画像从学习者做练习、即时测评、学习行为轨迹等数据中精确诊断出学习者对知识点的掌握水平,明确每个知识点的综合评分(0-1之间,每个知识点总分设为1,综合评分越接近1表示该知识点掌握越好);然后根据知识点类型及难易程度,设定知识点达标阈值,当学习者某知识点综合评分小于该阈值时,则认为学习者未掌握该知识点,应将其加入待学习知识列表中(如K1、K2、K3……)。步骤二是基于画像从学习者信息感知、信息输入、信息加工和内容理解等方面,综合挖掘提取学习者的学习风格与偏好。步骤三是需要将步骤一和步骤二的结果和“知识点—学习资源关系”作为三个参数输入到学习资源匹配模块中,与学习资源池中的资源进行匹配计算。具体计算过程如下:首先,根据待学习知识点与学习资源之间的关联程度,从学习资源池中筛选出各类学习资源S1。其次,根据学习者对知识点的掌握水平(L1:基本认知,L2:夯实基础,

L3:拓展应用),从学习资源S1中筛选出同水平层次的学习资源S2,如学习者的掌握水平为L2(夯实基础级别),则筛选出S2资源的难度级别为R2(知识巩固级别的各类学习资源)。最后,根据学习者的学习风格与偏好等信息,从S2中筛选出适合其学习偏好的学习资源S3,并将S3加入待推荐学习资源列表中。按照上述步骤将与所有未掌握知识点关联度高且与学习者学习风格相近的学习资源更新到待推荐学习资源列表中。步骤四是将待推荐学习资源列表中的资源向学习者推荐,学习者根据学习资源进行补救性学习后进行学习成效测评;学习平台对学习过程和学习结果进行记录,及时更新画像并进入下一轮诊断与推荐。此外,如果学习者对某个知识点考核的综合评分高于设定知识点达标阈值,则认为学习者对该知识点掌握较好,此时学习资源匹配计算模块就会根据学习者的学习偏好从课程学习资源池中筛选出难度级别为R3(综合应用)类型的学习资源推荐给该学习者。

##### 2. 学习群体个性化学习资源推荐

在课程学习初期,学习者的学习行为数据和结果数据比较少的时候,学习者画像不能够全面清晰地描述学习者学习全貌,此时需要根据学习者的课程注册信息、调查问卷信息以及少部分学习行为数据来挖掘其学习偏好,并根据学习偏好将学习者分为合作学习型、自主探究性和被动接受型三类群体。合作学习型群体具有较强的学习能力和较强的团队协作能力,但其学习目标性较差,对教师引导或学习同伴引导的依赖性较高,针对此类学习群体,可以从学习资源池中向其推荐具有任务导向和明确分工的学习资源;自主探究型群体具有较强的自学能力和自我管控能力,学习效率高,针对此类学习者可以从学习资源池中向其推荐知识巩固型、稍微有点难度的学习资源,并引导其在规定时间内独立完成知识学习和内化;被动接受型群体自主学习能力强、缺乏学习动机与毅力、学习进程易中断,针对此类群体可以向其推荐趣味性强的慕课视频资源、游戏闯关式的练习题与测试题等学习资源,这样可以“诱导”学习者持续学习,尽可能地







达成学习目标。

如果学习平台有学习者丰富的学习行为和学习结果数据,就可以根据学习者知识掌握情况进行群组划分,将知识点综合评分在0.6以下的学习者划分为基本认知群体;将在[0.6-0.8)区间的学习者划分为夯实基础群体;将在[0.8-1)区间的学习者划分为拓展迁移群体。向基本认知群体推荐趣味性强、概念理解型学习资源;向夯实基础群体推荐具有一定探究能力、闯关式学习资源和活动任务;向拓展迁移群体推荐具有一定挑战难度、综合型的学习资源和活动任务,这样才能促进不同学习群体达成学习目标。

#### 四、基于学习者画像的个性化学习资源推荐模型的实践应用

##### (一) 研究对象

本研究以国内某院校数据结构课程学习为研究对象。该课程是软件技术、移动应用开发等专业的核心基础课程,课程目标是培养学生数据抽象能力和程序设计能力。课程开设周期为1~14周,每周4学时,总共64学时。课程学习者为在校软件技术专业的127名大二学生,这些学生在大一时已提前学习了部分专业基础课程,能够适应在线开放课程混合教学模式且具备一定的探究学习能力。数据结构课程学习内容偏向逻辑思维训练和实践操作技能培养,学习者在学习过程中容易出现一些学习瓶颈,如能及时给予教学干预和推送适切的学习资源,则可以较好地帮助学习者克服学习障碍达成学习目标。本研究基于学习者画像构建了个性化学习资源推荐模型,为验证其实践效果,将127名学习者划分成实验组、对照组1、对照组2三个组别,分别应用本模型、基于协同过滤算法的推荐模型和DINA认知诊断的推荐模型,并向实验组推送各类学习资源数为891条,向对照组1推送各类学习资源数968条,向对照组2推送各类学习资源数924条。

##### (二) 选择模型对比角度

学习资源推荐有别于商品推荐、新闻推荐等,商品推荐、新闻推荐中个人兴趣与偏好基本上是稳

定的,而学习资源推荐中学习者的需求具有动态变化性,即随着学习进程的推进,学习者的认知能力、知识构建水平、学习能力等方面也会发生改变。为验证基于学习者画像的个性化学习资源推荐模型的有效性和准确度,本研究从分类准确率 $P(L)$ 和平均排序分 $RS$ 两个角度对三类模型的推荐效果进行实验论证。

一是分类准确率指推荐学习资源正确(被学习者接受的资源)的数目占所有推荐资源总数的百分比,是衡量推荐模型的主要指标,其计算公式为: $P(L) = \sum_{i=1}^N |S_i| / \sum_{i=1}^N |I_i|$ 。其中, $N$ 为学习者人数, $I_i$ 为第 $i$ 个学习者收到的推荐资源集合, $|I_i|$ 为资源集合 $I_i$ 中的资源个数(如 $|I_i|=15$ ,表示向学习者推荐了15个资源), $S_i$ 为第 $i$ 个学习者接受推荐资源的个数。比如,如果多个模型推荐的 $X$ 个资源中都有 $Y$ 个资源是用户所需的,则多个模型的推荐准确率 $P(L)$ 均为 $Y/X$ 。在每个模型中学习者所需的 $Y$ 个资源的排序并不是相同的,在相同的准确率 $P(L)$ 下,如果能够使 $Y$ 个学习资源排名靠前,则这个模型更具有优势。二是平均排序分 $RS$ 指学习者所需的推荐资源,在推荐算法所生成的推荐序列中的排名之和与推荐序列中所有资源的排名之和的百分比,是衡量推荐算法排序准确度的重要指标,其计算公式为: $RS = \sum_{j \in Y} R_j / \sum_{i \in X} R_i$ 。其中, $X$ 为推荐模型所推荐资源集合, $R_i$ 为推荐资源 $i$ 在推荐序列集合 $X$ 中的排名; $Y$ 为用户所需的推荐资源集合, $R_j$ 为用户所需的推荐资源 $j$ 在 $X$ 中的排序,平均排序分 $RS$ 越低,该模型准确度则越高。

##### (三) 研究结果

一是分类准确率 $P(L)$ 对比。数据显示,随着学习资源推荐数目逐渐增大,三类推荐模型的整体准确率均会逐渐降低。DINA认知诊断推荐模型在训练集数据较少时,对学习者的认知诊断精度较差,向学习者推荐资源的准确率也较低,但随着数据集的丰富,其准确率高于基于协同过滤算法的推荐模型,而基于学习者画像的个性化学习资源推荐模型的准确率明显高于其他两种模型。在实证研究过程中,基于学习者画像的个性化学习资源推荐模

型向实验组推送各类学资源数为891条,被学习者采纳条465条,整体平均准确率P(L)为52.21%,其中操作演示型视频资源和趣味闯关练习资源最受学习欢迎,占比分别为22.6%、31.6%。基于协同过滤算法的推荐模型向对照组1推送各类学习资源数为968条,被学习者采纳294条,平均准确率P(L)为30.36%,其中操作演示型视频资源和趣味闯关练习资源最受学习欢迎,占比分别为21.4%、30.9%。DINA认知诊断推荐模型向对照组2推送各类学习资源数为924条,被学习者采纳416条,平均准确率P(L)为44.98%,其中操作演示型视频资源和趣味闯关练习资源最受学习欢迎,占比分别为24.3%、31.9%(见图4,表1)。整体而言,三

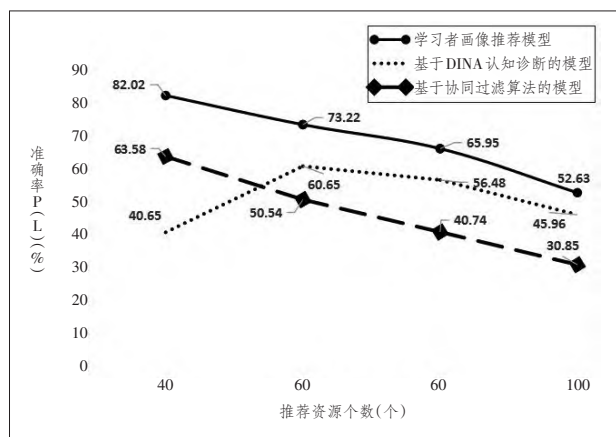


图4 三种推荐模型的学习资源推荐准确率对比

表1 三类推荐模型平均准确度对比

三类推荐模型	推荐资源总数(条)	被采纳资源数(条)	平均准确率P(L)
学习者画像推荐模型	891	465	52.21% 被采纳的讲解型视频资源78条(占比16.7%)、操作演示型视频资源105条(占比22.6%)、趣味闯关练习147条(占比31.6%)、测评试题60条(占比12.9%)、文本型资源63条(占比16.2%)
基于协同过滤算法的推荐模型	968	294	30.36% 被采纳的讲解型视频资源40条(占比13.5%)、操作演示型视频资源63条(占比21.4%)、趣味闯关练习91条(占比30.9%)、测评试题60条(占比20.4%)、文本型资源40条(占比13.8%)
DINA认知诊断推荐模型	924	416	44.98% 被采纳的讲解型视频资源58条(占比13.9%)、操作演示型视频资源101条(占比24.3%)、趣味闯关练习133条(占比31.9%)、测评试题60条(占比14.4%)、文本型资源63条(占比15.5%)

类模型推荐的个性化学习资源中,被采纳最多是趣味闯关练习资源和操作演示型视频资源,这在一定程度上也反映出学生注重自身操作技能训练和形成综合解决实际问题能力。今后教学者在教学设计、学习资源开发时应偏重开发实践操作型、趣味性、难易梯度合理的各类学习资源。

二是平均排序分RS对比。数据显示,随着I值的逐渐增大,三种推荐模型的平均排序分也会逐渐提高,基于协同过滤算法的推荐模型其RS值比DINA认知诊断模型推荐模型更大,而基于学习者画像的个性化学习资源推荐模型的RS值较小(见图5)。

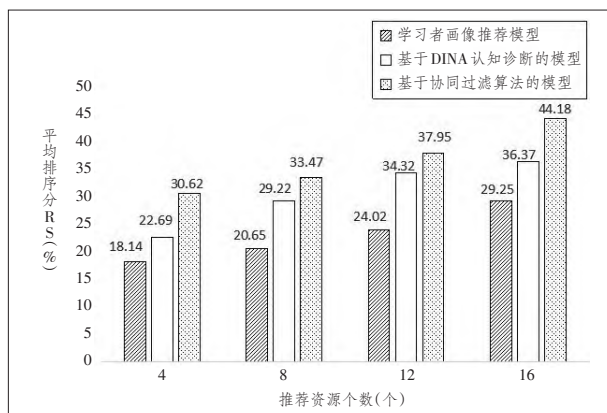


图5 三种推荐模型平均排序分对比

#### (四) 研究结论

一是基于学习者画像的个性化学习资源推荐模型比基于协同过滤算法的推荐模型和DINA认知诊断推荐模型具有更高的准确度(即较高的分类准确度和较低的排序准确度),既保证了推荐结果的精准性与可解释性,又融合了学习者个性特点,能够向学习者推荐不同学习难度、不同学习类型的适切学习资源。

二是基于学习者画像的个性化学习资源推荐模型一定程度上缓解了数据稀疏等问题。当学习者的学习行为数据较少或刚开始加入课程学习时,利用学习者画像采集学习者历史学业数据、学习者个性特征数据,挖掘出与其学习偏好相似的其他学习者,将其他相似学习者学习过的学习资源和学习路径推荐给目标学习者,在一定程度上缓解了冷启动与数据稀疏等问题。





当然, 尽管基于学习者画像的个性化学习资源推荐模型能够对学习者个体和群体精准推荐学习资源, 在实践应用中也取得了较好的效果, 但其在学习者画像及时动态更新、知识点与学习资源之间语义关系标注的准确性等方面还有待进一步优化, 另外, 该推荐模型只对静态事实标签进行了处理, 并未涉及动态学习路径规划推荐等, 在下一阶段的研究中, 研究团队将针对这些不足进行深入研究。

#### 参考文献:

- [1]中华人民共和国教育部.教育部关于印发《教育信息化十年发展规划(2011—2020年)》的通知[EB/OL].[2022-12-30].[http://www.moe.gov.cn/srcsite/A16/s3342/201203/t20120313\\_133322.html](http://www.moe.gov.cn/srcsite/A16/s3342/201203/t20120313_133322.html).
- [2]中华人民共和国教育部.教育部关于印发《教育信息化2.0行动计划》的通知[EB/OL].[2022-12-30].[http://www.moe.gov.cn/srcsite/A16/s3342/201804/t20180425\\_334188.html](http://www.moe.gov.cn/srcsite/A16/s3342/201804/t20180425_334188.html).
- [3]中华人民共和国教育部.中共中央、国务院印发《中国教育现代化2035》[EB/OL].[2022-12-30].[http://www.moe.gov.cn/jyb\\_xwfb/s6052/moe\\_838/201902/t20190223\\_370857.html](http://www.moe.gov.cn/jyb_xwfb/s6052/moe_838/201902/t20190223_370857.html).
- [4]P. Resnick & H. R. Varian. "Recommender Systems"[J]. Communications of the ACM, 1997(3): 56-58.
- [5][30]姜强, 赵蔚, 杜欣, 等.基于用户模型的个性化本体学习资源推荐研究[J].中国电化教育, 2010(5): 106-111.
- [6]赵晋, 周苏苏, 张建军.基于在线评论的学习资源推荐方法设计与实证[J].数学的实践与认识, 2022(9): 260-270.
- [7]赵蔚, 姜强, 王朋娇, 等.本体驱动的e-Learning知识资源个性化推荐研究[J].中国电化教育, 2015(5): 84-89.
- [8]Afridi A H, Yasar A, Shakhshuki E M. Facilitating Research Through Serendipity of Recommendations[J]. Journal of Ambient Intelligence and Humanized Computing, 2020(6): 2263-2275.
- [9]梁婷婷, 李春青, 李海生.基于内容过滤PageRank的Top-k学习资源匹配推荐[J].计算机工程, 2017(2): 220-226.
- [10]Ibrahim M E, Yang Y Y, et al. Ontologybased Personalized Course Recommendation Framework[J]. IEEE Access, 2018(1): 1.
- [11]Segal A, Katzir Z, Shapira B, et al. EduRank: A Collaborative Filtering Approach to Personalization in E-learning[C]. Educational Data Mining, 2014: 1023-1025.
- [12]姜强, 赵蔚, 李松, 等.大数据背景下的精准个性化学习路径挖掘研究——基于AprioriAll的群体行为分析[J].电化教育研究, 2018(2): 45-52.
- [13]赵宁.基于协同过滤技术的学习资源个性化推荐系统[D].石家庄: 石家庄铁道大学, 2015: 30.
- [14]兰明祥.基于深度神经网络的学习资源个性化推荐模型构建与实现[D].重庆: 西南大学, 2021: 23.
- [15]文孟飞, 胡超, 于文涛, 等.一种基于特征提取的教育视频资源推送方法[J].现代远程教育研究, 2016(3): 104-112.
- [16]任蓓蓓, 徐益强, 封丽.基于深度学习的数字图书馆资源个性化推荐模型设计[J].贵阳学院学报(自然科学版), 2021(1): 32-36.
- [17]L Dai, Y Fan. Review of Personalized Recommendation learners in E-Learning Systems[J]. International Conference on Semantic Information Retrieval, 2019(3): 3230-3239.
- [18]王晓东, 时俊雅, 李淳, 等.学习资源精准推荐模型及应用研究[J].河南师范大学学报(自然科学版), 2019(1): 26-32.
- [19]李浩, 余雪, 杜旭, 等.基于学习者时空特征的移动学习资源推荐模型研究[J].现代教育技术, 2020(10): 13-19.
- [20]王奕.基于概率矩阵分解的个性化图书推荐模型研究[D].西安: 西安电子科技大学, 2020: 21.
- [21]常诗卉.基于逻辑回归精排序模型的个性化推荐系统研究[D].广州: 广东技术师范大学, 2020: 26.
- [22]字云飞, 李业丽, 孙华艳.基于深度神经网络的个性化推荐系统研究[J].电子技术应用, 2019(1): 14-18.
- [23]涂冬波, 蔡艳, 丁树良.认知诊断理论[M].北京: 北京师范大学出版社, 2021: 3.
- [24]赵宇丹.基于GP-DINA的教师继续教育的自适应学习系统设计[J].微型电脑应用, 2022(1): 74-77.
- [25]李全, 刘兴红, 许新华, 等.基于联合概率矩阵分解的个性化试题推荐方法[J].计算机应用, 2018(3): 639-643.
- [26]赵新萍.基于图模型的个性化推荐技术研究[D].沈阳: 沈阳航空航天大学, 2011: 28.
- [27]刘生昊.基于图模型的个性化推荐研究[D].武汉: 华中科技大学, 2021: 15.
- [28]莫尉.学习者画像建构及应用研究\_莫尉[J].湖南理工学院学报(自然科学版), 2021(3): 64-69.
- [29]成亚玲, 谭爱平.基于学习者画像的在线开放课程学习预警研究[J].当代职业教育, 2022(4): 102-112.

责任编辑 黎恩

## How to Help Learners Get Out of Learning Resource Maze ——Personalized Learning Resource Recommendations Based on Learner Persona

CHENG Yaling, TAN Aiping

(Hunan Industry Polytechnic, Changsha Hunan 410208)

**Abstract:** Online courses represented by MOOCs provide learners with multimodal and massive learning resources. However, learners are often faced with information overload, and mismatch between resource content and learning needs when selecting resources, creating a learning resource maze. How to efficiently and accurately recommend appropriate learning resources and adaptive learning paths to learners has become an urgent problem in current education informatization. The personalized learning resource recommendation model based on learners' persona, starting from the practical needs of personalized learning resource recommendation, and supported by modern education theory and digital technology, explores the personalized feature diagnosis of learners' status and learning needs, the mining and representation of learning resources' own attribute feature parameters, and the precise matching of learning resources and learning needs in the process of personalized learning resource recommendation. It aims to recommend personalized learning resources to learning individuals and groups. The personalized learning resource recommendation model based on learner persona is applied to teaching practice. The results show that compared with collaborative filtering recommendation and DINA cognitive diagnosis recommendation, the learning resource recommendation model has higher accuracy and better interpretability, and can recommend appropriate learning resources for learners with different learning needs. At the same time, the model also alleviates the problems of cold start and data sparsity to some extent.

**Keywords:** online open courses; learner persona; learning resources; deep learning; recommendation model; digital