

科普资源个性化智能推荐算法研究与实现*

文 ◆ 湖南食品药品职业学院 阳 晖

引言

实现科学普及是党和国家的重要国策。科普资源个性化智能推荐是进行全民科普教育和在校高校学生课外辅助获取知识的重要手段。

随着大数据、云计算等新型互联网技术在科学普及教育中的深度应用，高校教育信息化的发展催生了大量的科普知识在线学习平台，使受教育者突破了时空的限制，让科普资源的获取方式变得更加灵活，解决了科普资源共享问题，优化了科普资源的配置手段，形成了课堂教育和获取知识的辅助手段的格局。但随着网络科普资源的指数式增长，受教育者所面临的“信息过载”和“学习迷航”等问题也日益加剧。

在“人工智能+教育”全新背景下，科普资源个性化智能推荐系统^[1]是解决“信息过载”和实现“个性化服务”的一种有效方法。该系统通过对采集的受教育者信息和学习行为数据进行分析，得到受教育者个性化学习表现以及学习特征，以可视化方式呈现符合自身学习表现的学习活动方案。而基于机器学习构

建推荐引擎，为数据处理构建组合函数、学习管道，构建 KNN 分类器和回归器，计算距离分数和相关系数，查找给定数据库中的相似用户，生成科普资源推荐。在推荐系统中，采用协同过滤算法，从大量没有差别区分过的知识中检索出关于用户喜好、适合用户（受教育者）特征和需求的预测性信息，从而获得合适的科普资源，解决“信息过载”的问题。在推荐过程中，还可通过可视化方式展示和评价推荐成效。通过可视化技术为受教育者展示科普资源各知识点间的关系以及学习路径，给予受教育者指导和学习方案，缓解“知识迷航”的问题。系统一般包括受教育者模型、领域知识模型、个性化推荐引擎和接口模型 4 个部分。

1 科普资源个性化智能推荐算法研究具有重要意义

信息技术促进科普资源个性化智能推荐算法的研究与实践，是当前科普教育领域亟须探索的重要课题。其研究内容包括受教育者个性挖掘、受教育者模型构建、领域知识模型构建以及个性化智能推荐策略等，不仅为科普资源个性化智能推荐系统的设计与实现提供了具体的理论模型支持，还进一步丰富了相关理论和方法，具有重要的学术研究价值。

科普资源个性化智能推荐算法研究以基于机器学习的科普资源个性化推荐算法的基本理论为基础，构建科普资源个性化推荐算法系统。对受教育者而言，能够提高科普资源推荐的精准性、提升在线科普教育质量和个性化学习效果；对科普资源而言，个性化的主动推荐服务可有效促进资源最大化利用，提高科普资源质量和利用率。项目研究为提升在线科普教育质量和个性化科普知识学习效果提供了一种新的实践思路。

2 项目的主要研究内容、研究目标

2.1 主要研究内容

(1) 科普资源推荐系统中受教育者模型的构建

基于我国发布的受教育者模型构建标准，依据学习风格理论、认知行为理论，研究构建基于受教育者个性特征、学习行为、知识需求和情

*【基金项目】2023 年度湖南省教育厅科学研究优秀青年项目“面向人工智能发展湖南食品药品类专业课程集群建设研究”（23B1008）的科研成果之一

【作者简介】阳晖（1983—），男，湖南长沙人，硕士研究生，副教授，研究方向：计算机与信息技术。

境感知的受教育者模型，为推荐系统动态地组织合适的科普资源提供决策依据。

(2) 科普资源推荐系统中领域知识模型的构建

通过本体建模、数据挖掘、知识抽取、知识存储等研究，构建科普资源的领域知识模型，准确地表征和有效地组织学科领域的知识结构，包括知识点、知识点属性以及知识点之间的联系，为系统根据受教育者模型进行领域知识的智能推荐提供前提条件（这里科普资源仅限于医药科普资源）。

(3) 科普资源推荐系统中推荐引擎模块的构建

科普资源推荐系统通常指应用数据分析技术找出用户最可能喜欢的内容，并将相应的内容推荐给用户。从技术上来说，推荐系统通常指应用数据分析技术从海量数据中根据一定条件筛选数据，并优先提供最匹配的数据。因此，在推荐系统中包括以下内容。第一，如何在海量（千万级别以上）数据中进行快速筛选；第二，如何对数据的匹配程度进行判断（打分）；第三，如何在挑选出来的匹配结果中进行再次处理，让最匹配的结果优先展示。

在构建推荐引擎模块中，基于关联规则挖掘算法、Hbase 数据库，大数据分析框架 Hadoop，使用 Python 实现文本主题抽取，计算主题相似度，研究受教育者模型和领域知识模型之间的匹配映射关系，设计推荐策略。采用协同过滤算法过滤信息，为数据处理构建组合函数、学习管道，构建 KNN 分类器和回归器，计算距离分数和相关系数，查找给定数据库中的相似用户，生成科普资源推荐，从而获得合适的科普资源。在推荐过程中，还可通过可视化方式展示和评价推荐成效。通过可视化技术为受教育者展示科普资源各知识点间的关系以及学习路径，给予受教育者指导和学习方案^[2]。

2.2 研究目标

本项目的研究目标是针对目前科普资源个性化推荐系统存在的主要问题，对受教育者模型、领域知识模型以及推荐策略模型等进行优化研究，设计并研发基于机器学习的科普资源智能推荐系统，实现根据用户（受教育者）个体的需求、喜好、知识层次等特征差异推荐科普资源、学习效果诊断和学习路径，缓解在线学习中的“信息过载”和“学习迷茫”问题。将研究成果应用于高校校园内外的科普教育的在线服务中，进行学习成效的实证研究，以有效促进全民的科普知识的个性化教育。

3 研究方法与技术路线

3.1 研究方法

(1) 文献研究法

项目组对国内外相关的文献资料进行收集、整理，充分把握研究的主题、范围、方法以及需完善和突破之处，掌握与科普资源推荐系统密切相关的受教育者模型、领域知识模型、推荐策略的基本论点，用于构建研究的理论基础和框架。

(2) 比较研究法

比较常见的科普资源个性化推荐系统模型，分析对比其优势和不足，寻找适合本项目的研究路径，

构建科普资源个性化智能推荐系统的参考模型^[3]。对比分析受教育者使用科普资源个性化智能推荐系统前后的行为数据与效果，实证系统提高科普资源个性化推荐的学习效率、满意度与质量的有效性。

(3) 实验研究法

①受教育者模型的构建。

②基于科普资源的领域知识模型的构建。

③科普资源个性化智能推荐引擎的构建。

④科普资源个性化智能推荐系统软件产品的需求分析、概要设计、详细设计、数据库设计、测试以及部署等^[4]。

(4) 调查研究法、评价研究法、实证研究法

通过调查、收集、汇总、分析和呈现与受教育者相关的数据，包括基本信息、学习风格、认知水平和学习历史过程（如受教育者访问科普资源的媒体类型、访学时间、访问次数）等。将研究成果应用于科普教育的在线学习中，受教育者通过调查问卷对学习效率、学习效果、学习满意度等进行评价，进行学习成效的实证研究。

3.2 技术路线

项目开发的基本技术路线如图 1 所示。

4 项目研究的主要创新点

本项目的研究重点是构建受教育者模型和科普资源的领域知识模型，研究难点是设计个性化推荐引擎。二者是研究的关键所在，其主要创新之处如下。

(1) 人机协同构建科普资源的领域知识模型。本项目运用知

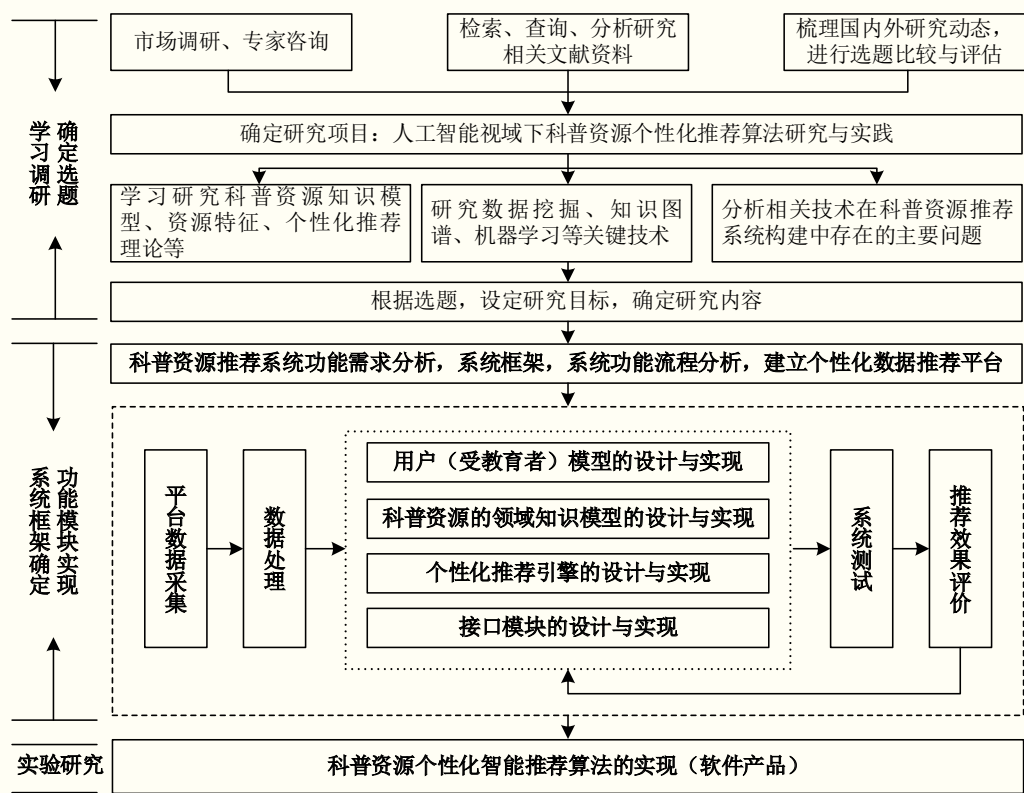


图 1 项目开发的基本技术路线

识表示和推理能力对多源异构的科普资源进行建模,为科普资源个性化智能推荐系统中的领域知识建模提供了新思路。在构建过程中,为解决人工构建主要依赖学科专家的主观认知和经验,成本较高,一致性差;自动化构建可节约构建成本,但精准度低、实用性差的问题,本项目提出了一种通过将学科专家经验与人工智能技术相融合的人机协同方式,构建领域知识模型的解决方案。

(2) 科普资源个性化智能推荐系统中科普资源、学习路径的优化推荐。本项目在受教育者模型和领域知识模型的基础上,采用资源推荐服务策略,利用关联规则挖掘算法、KNN 分类算法、距离分数和相关系数,计算、查找数据集中相似用户和协同过滤方法实现个性化引擎推荐中的科普资源、学习路径的优化推荐,将人工智能领域的机器学习思想

与教育领域相融合,是对科普资源个性化推荐系统研究的一次探索,为缓解“信息过载”和“学习迷航”,提升在线科普教育、全民科普教育质量和个性化学习效果提供了一种实践,为在校生提供课堂以外的获取知识的辅助手段。

结语

本研究针对科普教育“信息过载”与“学习迷航”问题,以机器学习为支撑,构建融合受教育者多维度特征的用户模型,创新人机协同方式打造医药科普领域知识模型,结合多种算法研发推荐引擎,形成个性化智能推荐系统。研究丰富了“人工智能+教育”理论,验证了系统提升推荐精准度、优化学习路径的成效,可为高校学生及全民科普提供支持,响应科学普及国策。未来可进一步优化系统,助力构建高效科普教育生态。■

引用

[1] 李浩君,张征,郭海东,等.深度学习视角下的个性化学习资源推荐方法[J].现代远程教育研究,2019(4):94-103.

[2] 曹先帅.基于知识图谱的机器学习算法推荐模型研究及实现[D].济南:山东大学,2022.

[3] 刘芷茵.深度学习在音频资源推荐中的应用策略[J].图书馆学刊,2021(12):78-82.

[4] 龙元媛,潘娅,代尚灵,等.智能技术驱动的专业实验室资源统一管理平台的设计[J].电子技术与软件工程,2021(11):195-198.