

面向创新素养培育的医学生画像系统的构建

张红伟¹ 许萍¹ 邢丹² 郭莹¹ 陈玲¹

(¹ 济宁医学院图书馆, 济宁 272067; ² 济宁医学院医学信息工程学院, 日照 276826)

摘 要 用户画像技术在刻画医学生特征方面具有独特优势, 本文将用户画像技术应用于创新素养培育研究领域, 结合教育大数据, 构建数据资源层、数据处理层、数据服务层 3 层架构的医学生创新素养画像系统。医学生创新素养画像系统可全面展示培育水平, 有效预测培育效果, 精准推荐教育资源, 提供培育机制决策参考, 为医学生创新素养培育提供数据支撑。

关键词 创新素养; 用户画像; 教育大数据

中图分类号: G640 文献标识码: A 文章编号: 1000-9760(2023)10-349-04

Construction of a medical student portrait system for cultivating innovative literacy

ZHANG Hongwei¹, XU Ping¹, XING Dan², GUO Ying¹, CHEN Ling¹

(¹ Library of Jining Medical University, Jining 272067, China;

² School of Medical Information Engineering, Jining Medical University, Rizhao 276826, China)

Abstract: User profiling technology has unique advantages in depicting the characteristics of medical students. This article applies user profiling technology to the field of innovation literacy cultivation research, and combines it with educational big data to build a medical student innovation literacy profiling system model with a three-layer architecture of data resource layer, data processing layer, and data service layer. The medical student innovation literacy portrait system can comprehensively demonstrate the cultivation level, effectively predict the cultivation effect, accurately recommend educational resources, and provide decision-making references for cultivation mechanisms. By utilizing user profiling technology, a smart service system for medical students' innovation literacy can be constructed, providing data support for the cultivation of medical students' innovation literacy.

Keywords: Innovation literacy; User portrait; Education big data

创新人才是国家自主发展的根本资源, 是人才培养的长期战略目标。培养创新人才是高等学校的根本任务, 提升大学生创新素养成为高等学校的重要课题。要大力培养高精尖紧缺人才, 需建立高层次复合型人才培养新机制, 将创新创业能力和实践能力培养融入课程体系。医学教育是国家卫生与健康事业的基石, 在医学教育需求与国家政策的驱动下, 如何提升医学生创新素养逐渐成为医学教育领域的研究热点。以教育大数据为基础, 用户画像技术在深入挖掘、全面展示学生特征等方面展现出独特优势, 为学生管理、决策与预警带来了极大

便利。本文以医学高校各业务系统数据为基础, 结合用户画像理论, 构建医学生创新素养画像系统, 动态展示医学生创新素养培育水平, 依据系统反馈结果不断优化培育方案, 进一步促进因材施教, 提升创新素养培育效果。

1 创新素养用户画像研究概述

随着创造力研究不断深入, 逐渐延伸出了创新素养概念。创新素养研究主要分为创新素养概念及要素解析、创新素养培育路径及评测研究。关于创造力的研究最早可追溯到 1863 年, 高尔顿在大量实验的基础上, 从遗传学的角度指出创造力来源于潜能和天赋^[1]。吉尔福特和斯腾伯格分别提出著名的三维智力结构理论^[2]和创造力三侧面理

[基金项目] 山东省教育科学规划创新素养专项课题 (2022CYB208)

论^[3]。国内创新素养的概念于 1999 年提出,林崇德认为创造力是根据一定目的,运用一切已知信息,产生出某种新颖、独特、有社会价值或个人价值产品的智力品质^[4]。关于创新素养培育路径的研究,一些学者运用不同理论方法,探索出形式多样的创新素养培育路径。Zhang 研究了基于机器人竞赛平台的大学生创新素养培育机制^[5]。林如诗结合信息素养教育,探讨了元认知视角下高校图书馆创新素养培育路径^[6]。创新素养测验主要有量表测验、作品分析和主管评估 3 种方法,其中量表测验是最常用的一种。吉尔福特为测验发散思维的构成而设计了行为测验试题,这种测验为创新素养评测奠定了基础^[7]。赵苹探讨了创新素养评价指标,设计了相应的调查问卷,构建了一套创新素养评价体系^[8]。郭珊珊将协同创新理论引入创新素养评价领域,运用层次分析法计算各指标的权重,搭建了开放性评价体系^[9]。创新素养评测建立了厚实的理论和实践基础。但是,如何充分展示医学生创新素养培育水平,动态呈现医学生创新素养发展变化,如何根据医学生培育效果不断优化培育方案,这些研究内容学界还很少涉及。在教育大数据作用日益凸显的背景下,构建医学生创新素养服务平台已成为亟待解决的问题。

用户画像是一种刻画用户信息模型的技术,在挖掘用户真实数据的基础上,按照需求提炼出用户信息,呈现虚拟的用户信息全貌。用户画像技术最初应用于市场营销、社交媒体及数字图书馆等领域,随着大量研究的持续开展,用户画像应用范围不断拓宽,大量学者将其用于绘制大学生画像。刘漫利用用户画像描绘大学生基本属性、用户行为、用户兴趣等特征,通过聚类与关联算法为图书馆提供阅读需求数据,有效提升图书馆个性化阅读服务能力^[10]。郭顺利运用在线社区的大学生画像,获悉大学生需求,帮助平台运营商提供精准服务^[11]。刘雯通过用户画像呈现大学生网络素养特征,依据这些特征,从社会环境、媒介话语、个体心理 3 个层面给与干预策略,从而有效提升大学生网络素养^[12]。随着智慧校园和教育大数据的快速发展,用户画像在大学生研究领域应用广度和深度持续拓展。用户画像虽未直接应用于创新素养研究,但在刻画大学生特征方面展现出了无可比拟的优势,为全方位反馈大学生创新素养培育水平提供了新思路和新方法。

2 医学生创新素养画像模型的构建

在教育大数据基础上,展示医学专业知识、临床思维能力、临床实践能力、临床创新意识等是当今医学生独特创新素养培养的重点,为此我们设计了医学生创新素养画像系统模型并进行了有益的探讨。医学生创新素养画像系统模型分为 3 层,分别是数据资源层、数据处理层和数据服务层。

2.1 数据资源层

数据是医学生创新素养画像系统的物质基础,数据资源层解决了系统数据来源问题。医学生创新素养评测采用量表测验的方法,这些数据通过调查问卷的方式获取。

2.2 数据处理层

数据处理层是医学生画像系统模型的核心,原始数据在数据处理层经过数据预处理、数据挖掘等流程转化为标签数据,为数据服务层勾勒医学生数据画像奠定基础。

2.2.1 数据预处理 数据预处理是数据挖掘的前提,数据挖掘的效果很大程度上取决于数据预处理的优劣程度。系统从数据资源层提取的原始数据存在不完整、不一致、有噪声、冗余等问题,这些问题数据不宜直接进行数据挖掘,需进一步转换为高质量数据,这个转换过程就是数据预处理。数据预处理的方法主要有数据清洗、数据集成、数据变换和数据规约。数据清洗去除原始数据中的噪声数据和无关数据,清洗脏数据,处理缺失值和异常值等问题;数据集成通过实体识别和冗余数据识别等方法将不同数据源的数据整合,形成一致的数据,进行统一存储;数据变换通过规范化、离散化等处理将数据进一步转化为适合数据挖掘的格式;数据规约在保证数据完整性的前提下,最大限度地精简数据,降低数据维度和规模。

2.2.2 标签提取 用户画像的核心在于给用户打标签,标签体系是医学生创新素养画像系统的基础。吉尔福特在《创造性才能-它们的性质、用途与培育》一书中详细分析了创新素养构成体系,至今仍是创新素养研究的基础。本文以吉尔福特创新素养构成体系为基础,结合医学生职业特性,梳理出基本实性“知识结构、认知能力、创新意识、创新思维、人格特征等”6 个一级指标和 41 个二级指标的医学生创新素养指标体系。画像标签分为事实标签、模型标签和预测标签。对于学生姓名、学号、

班级等既定事实的数据,系统运用文本挖掘、自然语言处理等方法对这些数据进行直接提取,生成事实标签;对于创新素养当前成绩等一些原始数据中不存在直接对应的数据,系统需要通过定义规则和关联数据,运用机器学习等方法生成模型标签;在事实标签和模型标签的基础上,系统通过预测算法(支持向量机、决策树、随机森林等)和聚类算法等生成创新素养预测成绩等预测标签。系统不仅要完成医学生个体画像的构建,还需通过聚类分析、关联分析等完成医学生群体画像和各类教育资源画像的构建,建立医学生和教育资源画像标签数据库,为数据服务层提供数据支撑。

2.3 数据服务层

医学生创新素养画像系统不仅能全面反应医学生创新素养培育水平,还为创新素养培育本身提供了新途径。

2.3.1 创新素养培育水平展示 医学生创新素养画像系统从个体、群体、整体等不同维度动态展示医学院校创新素养培育水平。系统将调查问卷结果中的医学生评测成绩采用 5 级评分法表示,并与医学生创新素养二级标签相匹配。系统根据医学生原始评测成绩,按照各标签相应权重,计算出医学生创新素养综合成绩。针对群体和整体医学生创新素养成绩评测,系统计算每一项二级评测指标的平均值作为群体或整体医学生创新素养二级评测指标的成绩,然后依据各标签的权重计算出群体或整体医学生创新素养综合成绩。系统利用调查问卷定期对学生的创新素养培养水平进行评测并更新数据,为了进一步展示培育效果,系统将各个检测周期的数据汇集在一起,形成医学生创新素养成绩历史走势图,便于体现医学生创新素养培育水平发展趋势。

2.3.2 创新素养培育效果预测 医学生创新素养画像系统提供创新素养培育水平预测功能。教育数据挖掘已成为促进教育发展的助推器,医学生创新素养画像系统纳入了学生和教育资源的大量数据,包括学生所学课程及成绩、医学技能竞赛,图书借阅数据、科研数据、社团活动数据等,以这些数据为基础,依据回归算法、分类算法、关联分析算法、推荐算法等构建创新素养预测模型,从个体、群体、整体等不同层面预测医学生创新素养培育水平。医学生创新素养培育水平预测为创新素养培育把脉,诊断创新素养培育中存在的问题。从学生层面

通过创新素养培育效果预测进行学前预警,有助于学生进一步调整自己的学习策略;从学校层面通过对群体或整体学生创新素养培育效果预测,有助于验证创新素养培育策略,动态调整创新素养培育方案,对医学生提供有效的干预机制,进一步提升创新素养培育效果。

2.3.3 教育资源精准推荐 医学生创新素养画像系统提供教育资源精准推荐功能。医学高校储备的图书、电子资源、教学课程、医学实验器材、实践活动、师资等教育资源数量众多,学生难以快速筛选出适合自己的资源,这种状况阻碍了创新素养培育的发展。

智能推荐以大数据技术为支撑,在用户画像领域应用广泛,有效解决了资源超载与用户无法找到合适资源之间的矛盾。主流推荐算法包括协同过滤和基于内容的推荐算法,两种算法各有优缺点。协同过滤推荐算法利用用户资源使用经验为其他用户推荐优质资源。在创新素养培育中,朋辈在使用资源后形成的经验非常重要,一个同学认为优秀的教育资源也可能适用于其他同学,这种思想恰好与协同过滤算法相吻合,利用朋辈经验提升推荐精度。但是这种算法没有考虑资源内容属性,并且当新资源进入系统后,没有使用记录,存在资源冷启动问题。基于内容的推荐算法首先根据用户画像中的历史行为标签,利用兴趣矩阵将用户需求数据化,同时将资源内容特征运用向量空间模型转化为可计算的向量,通过 TF-IDF 等方法计算出资源特征权重,用户需求和资源内容特征都转化为向量后,采用余弦距离等方法计算两者的相似性,相似度越高的资源越贴近用户需求。基于内容的推荐算法充分将医学生创新素养培育需求与教育资源特征相结合,但是,这种算法依赖学生的历史行为数据,只能发现学生以前的需求,无法发现潜在需求,并且当新生刚建立画像时,没有行为记录,无法判断需求,存在新用户冷启动问题。本研究将两种算法相结合应用于医学生创新素养教育资源智能推荐。一方面,系统采用协同过滤算法思想,针对每项医学生创新素养二级指标,得分较高的同学使用过的学习资源也值得大家借鉴,将这些同学在画像系统中形成兴趣社区,利用协同过滤推荐算法推荐出 TOP-N 的资源列表,系统定期将这些资源列表推荐给该项二级指标得分较低的同学;另一方面,系统采用基于内容的推荐算法思想,为每个教

育资源建立画像,针对每种教育资源设立使用频率阈值,所有使用频率超过阈值的教育资源是学生最常使用的资源,将这些资源特征利用矩阵进行数字化表示,系统根据学生创新素养画像的历史行为数据,将学生需求利用矩阵进行数字化表示,计算学生需求和资源特征的相似度,为学生推荐适合本人需求的资源。协同过滤和基于内容的推荐算法相结合既发挥了学生的资源利用经验,又考虑到资源内容本身与学生需求的完美匹配,同时还规避了用户和资源冷启动问题,为医学生教育资源精准推荐提供了完善的解决方案。

2.3.4 创新素养教育机制决策参考 医学生创新素养画像系统从个体、群体、整体 3 个维度,不仅能展示当前创新素养培育水平,还可以绘制出创新素养水平历史走势图,基于海量数据挖掘形成的数据画像可为医学院校创新素养培育决策提供数据支撑。系统依据医学生创新素养数据画像,形成创新素养培育质量报告,根据创新素养培育水平变化趋势,不断探索创新素养培育的内在规律,精准分析、研判不同时期教育机制的优缺点,依据分析结果动态调整创新素养培育策略,持续积累培育经验,建立长效机制,形成“培育-展示-优化”的闭环培育模式,不断提升创新素养培育水平。

3 结语

创新素养培育水平的展示及反馈是高等教育发展的必然要求。在教育大数据作用逐渐凸显和智慧校园日趋成熟的背景下,数据驱动的创新素养培育模式逐渐得到应用,以大数据技术为基础的用户画像为医学生创新素养培育提供了新思路和新方法。凭借用户画像在刻画医学生信息特征方面的天然优势,本文将用户画像引入创新素养研究领域,构建了数据资源层、数据处理层和数据服务层 3 层架构的画像模型,建立了多维度创新素养水平展示、成绩预测、教育资源精准推荐、培育机制决策参考等服务内容为主的智慧服务体系,以期为医学生创新素养培育提供借鉴。

利益冲突:所有作者均申明不存在利益冲突。

参考文献:

- [1] Galton F. Hereditary Genius [M]. London: Macmillan, 1869:1-5.
- [2] Guilford JP. Creative talents: their nature, uses and developments [M]. NY: Bearly Ltd, 1986:36-40.
- [3] Sternberg, Robert J. Implicit theories of intelligence, creativity, and wisdom [J]. J Personal Soc Psychol, 1985(3): 607-627. DOI: 10. 1037/0022-3514. 49. 3. 607.
- [4] 林崇德. 创造性心理学的几项研究[J]. 山东师范大学学报(人文社会科学版), 2014, 59(6): 5-14, 2. DOI:10. 16456/j. cnki. 1001-5973. 2014. 06. 008.
- [5] Zhang Y, Wu C, Hao C, et al. Innovation quality cultivation of college students by scientific contest [C]//First International Workshop on Education Technology & Computer Science. IEEE Computer Society, 2009.
- [6] 林如诗, 田稷. 元认知视角下的高校图书馆创新素养教育路径探讨[J]. 图书馆学研究, 2020(22): 66-73. DOI:10. 15941/j. cnki. issn1001-0424. 2020. 22. 010.
- [7] 段继扬, 程良道, 谈仕明, 等. 关于创造力与智力的关系的测试研究[J]. 湖北师范学院学报(哲学社会科学版), 1992(5): 102-107.
- [8] 赵革. “创新创业”视域下高职学生创新素养评价与实证分析[J]. 农业图书情报学报, 2021, 33(3): 90-97. DOI:10. 13998/j. cnki. issn1002-1248. 20-0579.
- [9] 郭珊珊, 郑直. 大学生创新素养开放性评价体系的构建与实践[J]. 青海社会科学, 2022(3): 72-81. DOI: 10. 14154/j. cnki. qss. 2022. 03. 025.
- [10] 刘漫. 基于用户画像的高校图书馆阅读推广模式构建[J]. 图书馆理论与实践, 2019(7): 21-26. DOI:10. 14064/j. cnki. issn1005-8214. 20190418. 001.
- [11] 郭顺利, 张宇. 基于 VALS2 的在线健康社区大学生用户群体画像构建研究[J]. 现代情报, 2021, 41(10): 47-58. DOI:10. 3969/j. issn. 1008-0821. 2021. 10. 006.
- [12] 刘雯. 信息连接过载下大学生网络素养的表征和提升研究[J]. 传媒, 2020(21): 82-84. DOI:10. 3969/j. issn. 1009-9263. 2020. 21. 029.

(收稿日期 2023-03-18)

(本文编辑:石俊强)