

基于 LDA 模型的乡村医学教育领域研究热点及发展导向研判

武颂文

西安医学院公共卫生学院

DOI:

[摘要] 目的：探究1990–2023年我国乡村医学教育领域的研究主题与热点。方法：以1990–2023年中国知网有关乡村医学教育有关的709篇文献为研究对象,以标题和摘要为文档语料,通过一致性指标和pyLDAvis可视化建模结果确定最佳主题个数,采用LDA主题模型实现主题聚类,并通过主题强度和阈值确定研究热点。结果：基于LDA主题模型研究发现1990–2023年乡村医学教育领域研究主题有4大类,分别是农村医学人才培养模式、乡村医生继续医学教育、乡村振兴与高职医学生培养、农村医学专业实践教学改革；研究热点为农村医学人才培养模式、乡村医生继续医学教育和农村医学专业实践教学改革。结论：4大类主题间分类清晰,代表我国乡村医学教育领域研究的方向；目前对农村医学人才培养模式、乡村医生继续医学教育、农村医学专业实践教学改革主题关注较多,取得了一些成效,但在新医科和人工智能普及的背景下,人才培养方案、实践教学和继续医学教育将会迎来新的挑战。

[关键词] 乡村医学；医学教育；LDA模型；热点主题；主题演化

中图分类号：R19 文献标识码：A

Research on the Hot Topics and Future Research Orientation of Rural Medical Education Based on LDA Model

Songwen Wu

Department of Public Health, Xi'an Medical University

[Abstract] Objective To explore the research themes and the hotspots in the field of rural medical education in China from 1990 to 2023. Methods 709 documents related to rural medical education in CNKI from 1990 to 2023 were used as the research objects, and the titles and abstracts were used as the corpus to determine the optimal number of topics through the coherence value and the visual modelling results of pyLDAvis, and the topic clustering was achieved by using the LDA topic model, and the hotspots were determined by the topic intensity and threshold value. Results Based on the LDA theme model, it was found that there were four major categories of research themes in the field of rural medical education from 1990 to 2023, which were rural medical personnel training model, rural doctors' continuing medical education, rural revitalization and senior medical students' training, and practice teaching reform of rural medical specialties. The research hotspots are rural medical personnel training mode, rural doctors' continuing medical education and rural medical professional practice teaching reform. Conclusion The four themes are clearly classified and represent the direction of research in the field of rural medical education in China; at present, more attention has been paid to the themes of rural medical talent cultivation mode, rural doctors' continuing medical education, and practice teaching reform of rural medical specialties, and some results have been achieved. However, in the context of the new medical science, as well as the popularity of artificial intelligence, the talent training program, practical teaching and continuing medical education will encounter new challenges.

[Key words] Rural Medicine; Medical Education; LDA Model; Hot Topics; Topic Evolution

引言

我国卫生工作的一贯方针是将重点放到农村,而农村卫生工作的突出问题之一是乡村医生数量短缺且其岗位胜任力

欠佳^[1]。我国一直致力于对这一突出问题的解决,在乡村医学教育领域进行了大量探索,在为基层培养“下得去、留得住、用得上、干得好”的应用型高素质医学人才方面取得了一定成绩,

但在新医科背景下,又存在新的挑战^[2,3]。对乡村医学教育领域进行主题分析,挖掘其研究热点,把握乡村医学教育研究的发展现状并分析未来研究导向,能够丰富和发展乡村医学教育理论,为开展新医科背景下乡村医学教育发展提供理论支持。本文应用隐性狄利克雷分布(Latent Dirichlet Allocation,简称LDA)模型,以国内现有的研究成果为原始数据,用计算主题强度的形式分析乡村医学教育领域研究的热点主题,并进行研究导向预判,以期为后续相关研究提供参考。

1 文献综述与研究逻辑框架

1.1 文献综述

2010年3月,卫生部(现为“卫健委”)印发了《关于印发以全科医生为重点的基层医疗卫生队伍建设规划的通知》(发改社会[2010]561号),提出要适应深化医药卫生体制改革、保障基层群众全面健康,加强对基层医疗卫生人才的培养力度,通过十年的努力利用多种途径培养30万名全科医生,形成一支人数充足且业务精干的基层医疗队伍^[4]。同年6月,国家发展改革委等部门发布了《开展农村订单定向医学生免费培养工作实施意见的通知》(发改社会[2010]1198号),提出要连续三年在医学类院校开展定向免费医学生培养工作计划,努力为乡镇卫生院及以下的医疗卫生机构培养适用于基层的医疗卫生工作人员^[4]。2015年5月,教育部印发了《关于进一步做好农村订单定向医学生免费培养工作的意见》(教高[2015]6号),提出了要对目前基层农村医疗人才的培养模式进行调整,完善培养目标、优化课程体系、培养临床能力,使之更实用于基层;也要加强对学生职业责任感的培养,严格控制毕业生的就业履约情况,结合实地制定具体的实施办法,切实有效的对基层医疗人员的能力和服务水平进行提升^[4]。从目前我国的政策支持方面来看,对于农村医疗卫生人员队伍建设方面是具有可操作性的^[4]。

虽然有政策支持,但由于基层农村地区相对于较发达的市级、县级地区,其医疗环境受限,医学高校毕业生愿意优先选择到基层乡镇卫生院和村卫生室工作的极少,而目前现有的基层农村医疗工作者存在专业知识较匮乏陈旧、流动性大等问题,使得基层农村的群众依然存在“看病难”的状况。因此,研究者为了推动农村医学专业定向培养计划的顺利进行,更好更快地提升农村医务工作者的基层服务能力,分析了定向农村医学专业学生培养现状、就业意向与毕业生的就业状况,探索了新的人才培养模式^[5],也对现有乡村医生状况进行了分析^[6]。纵观研究成果,未见对我国乡村医学教育领域的研究热点进行系统性地概述和总结的工作。

LDA最早由Blei等人提出,该模型通过计算“文档—主题”概率分布,进而计算当前主题强度,是目前使用较为广泛的主题模型^[7]。自LDA模型提出以来,许多学者对其进行了扩展和应用研究,其中不乏应用在热点主题分析方面。Paul等^[8]利用LDA模型将推特平台超过150万条关于健康相关的信息进行主题识别,最终确定识别十余种常见疾病的有关信息,为跟踪流行病动向、

分析症状、药物使用状况等提供一定参考;Godin等^[9]提出将LDA模型应用于社交媒体的搜索引擎推荐当中,应用LDA主题建模的方式将推特短文本进行主题识别,从而提高搜索引擎的分类和推荐精度。国内学者将LDA模型引入国内后,在原有模型基础上进行了适应中文语料的改良,使其更好地应用于中文文本主题挖掘和识别研究中。李博^[10]等利用LDA模型对教育元宇宙的130篇核心期刊文献进行系统梳理后发现,我国教育元宇宙研究主要围绕基本理论、批判审视、思政教育应用、学习赋能、场景创设、技术支撑、模型构建等7个主题展开。孙瑞英^[11]等将从CNKI得到的313篇文献摘要作为LDA主题模型的原始语料,通过模型训练得到国内智慧阅读研究热点,即智慧阅读技术研究、智慧阅读平台、智慧阅读用户服务、智慧化读物供给、智慧图书馆建设、智慧阅读评价、智慧阅读推广。综上可知,LDA模型可以较准确提取一个研究领域文献的热门研究主题,并且比传统词频聚类统计等方法所提取的结果更具区分度,有助于研究人员更加合理地把握该领域的研究发展状态,探寻未来研究导向^[11]。

1.2 研究逻辑框架

本研究对国内乡村医学教育领域相关文献进行文本预处理后,进行LDA主题建模,运用一致性指标与pyLDAvis库的可视化建模结果综合确定模型最优主题数,根据主题词的概率分布对研究主题进行标识,通过计算主题强度的方式确认热点主题,探究当前乡村医学教育的研究热点,本文所设计的研究框架如图1所示。

2 数据来源与方法

2.1 数据来源

本研究采用八爪鱼爬取中国知网CNKI中1990—2023年标题或关键词中含有乡村医学/农村医学/乡村医生/农村医生教育相关文献,共检索获得数据1195条,收集每篇文献的标题、摘要、发表时间等信息。为保证用于主题建模的语料信息的准确性,剔除无摘要、不相关文献,最终得到709条数据,将每篇文献的标题和摘要作为LDA模型的原始语料。

2.2 文本预处理

对选定的709篇文献摘要进行预处理,运用jieba分词和自建词典对原始语料进行切分,设置停用词表,将与主题无关的数字、英文、标点符号和中文词汇剔除,保留与主题相关词。文本预处理是LDA建模的重要部分,会直接影响后续分析的准确性和可靠性。

2.3 最优主题数确定

研究显示,LDA模型的主题聚类效果与潜在的主题数量直接相关^[12],因此在建立最优LDA模型前,需要确定最优主题数目。本研究采用一致性(Coherence)指标并结合pyLDAvis库的可视化建模结果来判定最优主题数目。

主题一致性是指一个主题中各个词语间的相似度。主题一致性数值越大,得到的主题越好^[13]。本研究采用的是基于滑动窗口的CV法^[14],这是计算一致性最好的方法之一,公式为:

$$coherence(V) = \sum_{(v_i, v_j \in V)} score(v_i, v_j, \epsilon) \quad (1)$$

公式(1)中, V 是描述主题的一组词汇, ϵ 是保证得分为实数的平滑因子。

借助Python中pyLDAvis库对LDA主题模型进行可视化展示, 各主题之间界限越分明, 分类效果较好, 模型越优。

2.4 LDA模型训练

LDA模型主要包含特征词、主题和文档三层结构^[11], 并假设所有文档都由隐含主题集合组成, 每个主题都有对每个文档的支持概率, 形成“文档—主题”概率分布, 概率越大说明该主题与某一文档的关联越大; 而每个主题又由一系列相互关联的特征词集合而成, 每个特征词都以一定概率选择某个主题, 形成“主题—词”概率分布, 概率越大说明该特征词与这一主题的关联性越强, 从而拥有更强的语义信息以解释主题的含义。LDA主题模型的生成过程^[11]: ①将文档集中的每一篇文档抽取一个主题; ②从抽取的主题所对应的特征词分布中抽取一个词; ③重复上述过程, 直到遍历整个文档集中每一个特征词为止。

本研究主要借助Python中的Gensim库调用doc2bow生成文档稀疏向量, 调用models.TfidfModel进行TF-IDF计算, 调用LDAmodel实现建模, 提取最优主题生成的“主题—词”和“文档—主题”概率分布。

2.5 热点主题识别与演化

本研究采用主题强度指标来识别热点主题。主题强度是对主题在一定时间窗口内受关注程度的描述指标, 若在某时间窗口内包含该主题的文档数目越多, 则其主题强度越大, 则越有可能是热点主题^[11]。主题强度公式如下:

$$\theta_z^t = \frac{\sum_{d=1}^{D_t} \theta_z^d}{D_t} \quad (3)$$

按公式(3)得到所有主题的主题强度后, 对其求均值, 并设定此均值为主题强度阈值, 若一个主题强度大于此阈值, 则说明此主题为热点主题^[14]。

3 结果

3.1 各年度文献发表情况

1990—2023年发文量整体呈上升趋势(图2)。1990—2007年发文量平缓, 年均发文量4篇; 2008—2014年上升迅速, 年均发文量达到35篇, 且2014年达到峰值(年发文量72篇); 2015—2023年有所下降。

3.2 “主题—词”和“文档—主题”概率分布

3.2.1 最优主题数。如图3所示, 一致性随主题数的增加呈递增趋势, 但主题数过大, 组成主题的词汇可能同时出现在多个主题中, 导致过拟合现象, 所以综合考虑pyLDAvis库的可视化结果, 如图4所示, 最终确定, 当主题数为4时, 一致性得分高, 各主题之间界限分明。因此LDA模型的最优主题数设定为4。

3.2.2 “主题—词”概率分布。LDA模型训练完毕后, “主题—词”概率分布如表1所示, 抽取每个主题前10个高概率特征词,

作为描述主题含义的代表, 优先以排在前列的词汇为判定依据, 归纳后得到的四大类主题分别为: 农村医学人才培养模式(T1), 乡村医生继续医学教育(T2), 乡村振兴与高职医学生培养(T3), 农村医学专业实践教学改革(T4)。将四大主题所指向研究成果的特征词进行整理, 同时从“文档—主题”分布中选取部分指向各主题的代表文献展开阐述。

3.2.3 “文档—主题”概率分布。LDA模型训练所得“文档—主题”概率分布如表2所示。该分布显示了每篇文档中四大主题的概率分布, 概率值越大则代表相关性越大。根据这一特性, 可以快速了解到每篇文档最重要的主题内容。如文档1, 由表2中概率分布可得出是关于T3“乡村振兴与高职医学生培养”的文献, 文档5是关于T2“乡村医生继续医学教育”的文献。

3.3 四大主题分析

根据各主题的高频特征词以及代表文献对其进行展开阐述。

3.3.1 农村医学人才培养模式(T1)。农村订单定向免费医学人才培养是构建全科医师培养体系的重要组成部分, 如何培养“用得上、下得去、干得好、留得住”的基层全科医学医学人才是一直以来需要迫切解决的问题。程梦琳^[15]等从思想道德培养、课程设置、实践能力培养、全科医学实验实训教学中心网络平台构建与应用、考核评价体系改革等方面进行了农村医学人才培养模式的探索与实践, 初步构建了基于岗位胜任力的农村订单定向免费医学人才培养模式。蔡瑜^[5]等以国内西部全科医生主要培养模式改革为出发点, 结合国家卓越医生教育培养计划实施政策, 介绍了西安医学院构建农村订单定向免费医学人才培养模式的具体做法, 包括人才培养方案制定、课程设置、教学团队建设、教学方法改革、实践教学及基地建设、技能考核等方面, 推行“3.5+1.5”的培养模式。任风云等总结了牡丹江医学院的经验和做法, 从修订培养方案、完善课程机构等九个方面进行了介绍, 形成了以岗位胜任能力为核心的新型“3.5+1+0.5”全科医学人才培养模式。廖纪元^[3]等探讨了在新医科建设背景下, 通过增设全科医学特色课以及大数据医疗、智能医学选修课完善教学体系等方式创新农村订单定向免费医学人才培养模式。

3.3.2 乡村医生继续医学教育(T2)。研究表明^[16], 我国卫生工作的重点在农村, 拥有一支较高素质的农村卫生人才队伍, 是农村卫生工作搞好的关键。而从乡村医生的产生和发展历史来看, 整个队伍的素质相对较低, 难以满足农村居民对健康的需求。因此, 必须重视这支队伍的建设, 关注乡村医生的继续医学教育工作。郑宇东^[16]等从我国农村卫生队伍的建设出发, 探讨了乡村医生的继续教育和培训现状, 调查结果显示乡村医生年轻人偏少, 学历层次偏低, 以中专为主, 学医途径多样, 县卫校居多, 培训内容主要为临床技能、用药知识和预防保健知识等, 并建议对于不同年龄、学历的乡村医生采用不同的培训方式。力晓蓉^[6]等通过对四川省农村卫生人员继续医学教育学分获得情况的调查, 发现农村卫生人员学历、职称和技术水平偏低, 多数人认为科研立项、发表学术论文和出版医学著作是获得学分

最困难的继续医学教育活动,并建议要针对不同的继续医学教育对象确定合理和可行的学分项目结构。罗汝珍^[17]综合阐述了美国、英国、日本、澳大利亚等国农村医生培养的途径和经验,包括三级培养模式,即通过全科医学生的学院教育、毕业后教育和继续教育三方面结合进行培养,面向农村医学生的招聘和留用政策等,为我国农村医生培养带来了启示。

3.3.3 乡村振兴与高职医学生培养(T3)。基层医疗建设是实现乡村振兴的重要方面,目前有很多阻碍基层医疗发展的因素,其中人才缺乏问题亟需解决。医学高职院校培养的学生是建设乡村医疗的主力军,承担着推动基层医疗建设的重要责任^[18]。然而,现实情况是很多学生毕业之后并不愿意服务基层,这也是导致基层医疗人才匮乏的一个重要原因^[18]。石沙沙^[18]等从医学高职院校人才培养机制角度出发,剖析高校毕业生不愿服务基层医疗的原因,并通过调整人才培养方案、引导学生树立正确的职业定位、加强基层实践活动等措施,助推基层医疗建设。张瑾钰^[19]等要围绕乡村振兴战略融入高职医学生培养的意义、困境以及创新路径展开论述,提出采用信息化工具转变学生刻板印象、拓展实践培养路径以及加强教师队伍建设的措施来创新高职医学生培养之路。廖德君^[20]等以问题为导向,发现目前存在思政课缺乏相关核心价值引领、通识课中缺乏浓厚乡土情怀培育等问题,建议医学高职院校应加强医学特色的乡村振兴战略思政教育、厚植乡村振兴战略价值引领的校园文化、构建乡村振兴战略价值引领的实践教学体系等路径,培养一支“爱农村、护农村、兴农村”的后备军,为乡村振兴保驾护航。

3.3.4 农村医学专业实践教学改革(T4)。通过对农村订单定向毕业生和定向单位的访谈和调查发现,定向医学生存在实践机会较少,实践操作能力较弱等问题^[21]。陈又佳^[21]对农村订单定向专科医学生实践教学进行了调查研究,提出完善实践教学基地建设、合理安排实践教学内容;科学化设定师生实践教学评价体系;结合定向专科医学生短学制的实际,差别化开展实践教学;注重实践教学环节的思想教育等建议。陆永良^[22]等以“能力本位教育”理念为指导,确立了定向农村社区全科医学人才必备的社区综合卫生服务能力,构建了人文素质模块、岗位技能模块、科研创新模块等三模块的实践教学内容体系,以及临床与社区一体化综合实践教学平台,创建了学校教育、医院教育、社区实践三位一体、三年融通的实践教学新模式。蒋培余^[23]等以胜任力培养为核心、构建第一课堂与第二课堂有效对接的实践教学体系,积极探索“医教协同、校地共育”实践教学模式。白纯^[24]等针对所在地区农村特点,临床医学(乡村医生)专业实践课程打破了学科界限,教学内容的设置突出职业的定向性和实用性,增加了在乡镇卫生院、社区卫生服务站和疾病控制中心的见习,实践教学过程渗透执业准入标准,毕业后与执业助理医生培训合理衔接。

3.4 热点主题与演化分析

图5结果显示,超出主题强度阈值(红色虚线)的主题为热点

主题,农村医学人才培养模式(T1)、乡村医生继续医学教育(T2)、农村医学专业实践教学改革(T4)是关注较高的3个热点主题。从各主题的发文量来看,乡村医生继续医学教育(T2)发文量最多(204篇)。

图6结果显示,主题T3“乡村振兴与高职医学生培养”的热度值从1990年的最高,持续下降,到2023年热度降为最低;主题T4“农村医学专业实践教学改革”的热度值从1990年的倒数第二,逐年上升为2023年第二大热度主题;主题T2“乡村医生继续医学教育”从1990年第二大热度,经过“M”样起伏变化,于2023年成为第一大热度主题;主题T1“农村医学人才培养模式”从1990年的热度最低值,逐年上升,于1999年达到四大主题热度之首,之后逐年下降,于2023年降为热度值第三。

4 结论与建议

本研究以我国1990-2023年间,CNKI发表的乡村医学教育领域文献为数据源,以摘要为语料,采用LDA模型进行主题挖掘。研究发现,该领域研究主题分为四大类,分别是农村医学人才培养模式(T1),乡村医生继续医学教育(T2),乡村振兴与高职医学生培养(T3),农村医学专业实践教学改革(T4),各主题之间分类清晰,代表我国乡村医学教育研究的方向。

4.1 乡村医生数量不足问题得到一定解决

自2010年起,我国就在高等医科类院校实施农村订单定向医学生免费培养项目,初衷是为了解决城乡卫生人力资源配置不均衡、乡村医生数量不足的问题。此项目早期效果并不明显,据统计^[25],2016年底全国乡村医生93.3万人,而2017年底却减少至90.1万人。乡村医生数量不足成为农村卫生服务能力较低的主要原因之一,既严重制约了县乡卫生服务体系一体化、分级诊疗、家庭医生签约服务等一系列改革的推进,也加剧了城乡居民在获得基本卫生服务方面的不平等。面对如此难题,我国一方面总结自身经验,一方面积极向英国、巴西等国学习,开展了一系列研究和实践,2022年底,我国执业(助理)医师和持乡村医生证的人员已升至114.1万人^[26],取得了成效。

4.2 农村医学专业实践教学改革有成效

本研究挖掘的四大主题就代表了我国探索的主要内容,对四大主题进行主题强度和文本聚类分析发现,农村医学专业实践教学改革(T4)是热度最高的主题,从T4主题热度逐年变化趋势来看,1990年我国大部分地区对实践教学是不太重视的,但随着调查研究和访谈的开展,各地均发现实践教学的重要性,之后既开展了针对农村医学教育的教师专业实践教学能力建设的思考,也开展了大量的农村定向医学生实践教学体系的构建与实践,发展到2023年T4已成为第二大热点主题。研究显示,各地均在实践基地建设、实践内容重构、实践评价体系构建等方面探索了实践教学新模式,在农村基层全科医学生岗位胜任力培养方面取得了初步成效。

4.3 乡村医生继续医学教育工作热度持续上升

本研究发现,乡村医生继续医学教育(T2)与农村医学专业实践教学改革(T4)的热度发展方向相反。1990年,T4热度最低,

而T2处于较高的第二热度;到了1993年,T4热度持续降低,而T2热度持续上升为第一热度;三年后,1996年发生反转,T4热度快速增长为第一热度,而T2热度持续下降为倒数第二热度;2008-2011年也观察到了两主题相反的发展趋势;2023年,T2再次成为第一热度主题。对1990-2023年文献梳理后发现,我国乡村医生文化程度以中专为主,学历层次偏低、业务能力不强是制约乡村医生队伍执业水平、农村卫生事业发展的瓶颈。郑宇东比较了2016年和2008年河南省平顶山市乡村医生持有执业资格证比例,发现比例虽有增加,但转化速度缓慢,执业程度亟待提高。由于医学院校是师资力量与实践基地资源均较丰富的单位,在乡村医生进修教育、更新教育、学历教育中能发挥主导作用,适宜地开展一些针对性强、周期短、形式灵活的培训是非常有必要的。

4.4 农村医学人才培养模式因地制宜

全国各地对农村医学人才培养模式都进行了有益探索,西部地区探索了“3.5+1.5”模式^[5],通过整体延长实践教学时间来提升学生的岗位胜任力,为学生毕业后尽快适应基层医疗卫生工作奠定基础;鄂南地区在新医科背景下,探索了“2.5+1+1+0.5”模式^[15],通过增设全科医学特色课以及大数据医疗、智能医学选修课等方式创新人才培养模式。“新医科”是国家应对新时代产业革命提出的战略规划,新医科建设需适应信息时代的医学研究和实践,服务于大健康的社会需求^[3]。因此,在前期有益探索基础上,要继续完善各地农村医学人才培养体系,使之与时代发展相适应,尤其是人工智能技术的发展将会对人才培养提出新的挑战。

[参考文献]

- [1]吕繁.全科医学与乡村医生教育[J].医学教育,1993(12):1-3.
- [2]赵茜阳.新医科背景下农村订单定向免费本科医学教育人才培养模式探析[J].就业与保障,2024(01):25-27.
- [3]廖纪元,梅武轩,胡振武,等.新医科建设背景下农村订单定向免费医学生人才培养模式研究[J].中国高等医学教育,2023(08):8-9.
- [4]王蓉.XZ卫校农村医学专业定向培养学生基层就业研究[D].西安电子科技大学,2020.
- [5]蔡瑜,弥曼,李亚军.农村订单定向免费本科医学教育人才培养模式的实践与探讨[J].中华全科医学,2018,16(10):1734-1736+1772.
- [6]力晓蓉,易易,陈运,等.四川省农村卫生人员继续医学教育学分情况调查[J].中国卫生事业管理,2008,25(12):835-837.
- [7]BLEI D M,NG A Y,JORDAN M I.Latent dirichlet allocation[J].Journal of machine learning research,2003,3(1):12-13.
- [8]PAUL M J,DREDZE M.You are what your tweet: analyzing twitter for public health[J].Artificial Intelligence,2011,38(2):265-272.
- [9]GODIN F,SLAVKOVIKJ V,DE NEVE W,etal.Using topic models for twitter hashtag recommendation[C].Proceedings of the 22nd international conference on world wide web.2013:593-596.
- [10]李博,汪基德,陈文峰,等.我国教育元宇宙研究的热点透视与未来路向——基于核心文献的LDA主题模型分析[J].现代远程教育研究,2024,36(03):103-112.
- [11]孙瑞英,陈宜泓.基于LDA主题模型的国内智慧阅读研究热点及发展导向研判[J].图书馆建设,2023,(03):82-93+103.
- [12]Zhao W Z,Chen J J,Perkins R,et al. A heuristic approach to determine an appropriate number of topics in topic modeling.BMC Bioinformatics,2015,16(Suppl13):S8.
- [13]Stevens K,Kegelemeyer P,Andrzejewski D,etal.Exploring topic coherence over many models and many topics. Proceedings of the 2012 Joint conference on empirical methods in natural language processing and computational natural language learning.New York:ACM,2012:952-961.
- [14]Röder M,Both A,Hinneburg A. Exploring the space of topic coherence measures. proceedings of the eighth acm international conference on web search and data mining. Shanghai, China.New York:ACM,2015:399-408.
- [15]程梦琳,孟巍,梅武轩,等.鄂南某高校实施农村订单定向免费医学生培养的探索与实践[J].中国高等医学教育,2022(06):19-20+71.
- [16]郑宇东.河南省平顶山市乡村医生继续医学教育的现状研究[J].中国卫生事业管理,2018,35(06):452-454.
- [17]罗汝珍.发达国家农村医生培养概况及其对我国的启示[J].医学与哲学(A),2015,36(04):71-74.
- [18]石沙沙,高谨,翟哲,等.乡村振兴背景下医学高职院校人才培养机制助推基层医疗建设研究[J].黑龙江科学,2021,12(17):18-21.
- [19]张瑾钰,靳力,杨红梅.乡村振兴战略融入高职医学生培养的创新分析[J].继续医学教育,2023,37(09):129-132.
- [20]廖德君,曹冰莹,杨茜,等.乡村振兴战略融入高职医学生培养的创新思考[J].卫生职业教育,2022,40(16):20-23.
- [21]陈又佳.安徽农村订单定向专科医学生实践教学调查研究[J].九江学院学报(自然科学版),2019,34(04):12-16.
- [22]陆永良,蒋培余,沈志坤,等.基于“能力本位教育”的农村社区全科医学专业实践教学体系研究[J].中国全科医学,2012,15(13):1501-1503.
- [23]蒋培余,卢东民,沈志坤,等.基于胜任力培养的农村基层卓越全科医学人才实践教学体系的构建与实践[J].中国高等医学教育,2019(04):71-72.
- [24]白纯,陈芳,许有华.面向农村服务的临床医学(乡村医生)专业实训课程建设[J].中国高等医学教育,2013(06):66-67.
- [25]中华人民共和国国家卫生健康委员会.2017年我国卫生健康事业发展统计公报[EB/OL].(2018-06-12)[2018-06-26].<http://www.nhfp.gov.cn/guihuaxxs/s10743/201806/44e3cdf11fa4c7f928c879d435b6a18.shtml>.

[26]中华人民共和国规划发展与信息化司.2022年卫生健康事业发展统计公报发布[EB/OL].(2023-10-12)[2024-06-30].
http://www.nhc.gov.cn/guihuaxxs/s3586s/202310/5d9a6423f2b74587ac9ca41ab0a75f66.shtml.

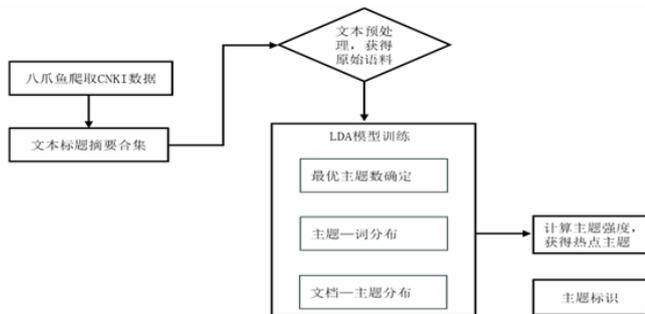


图1 基于LDA模型的热点主题识别研究框架

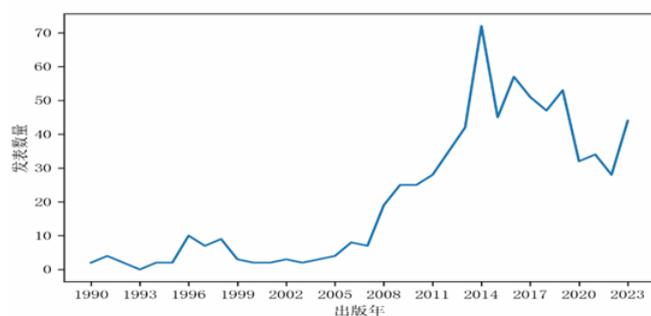


图2 文献发表数量年度分布

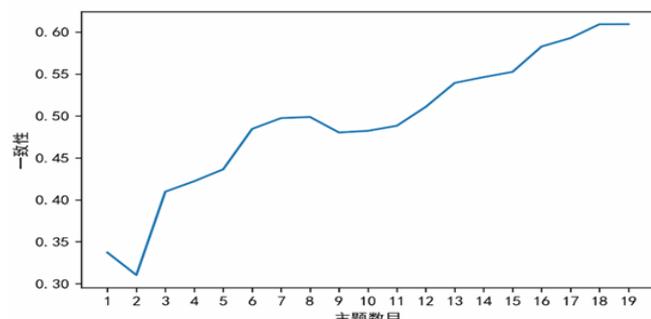


图3 主题一致性变化趋势

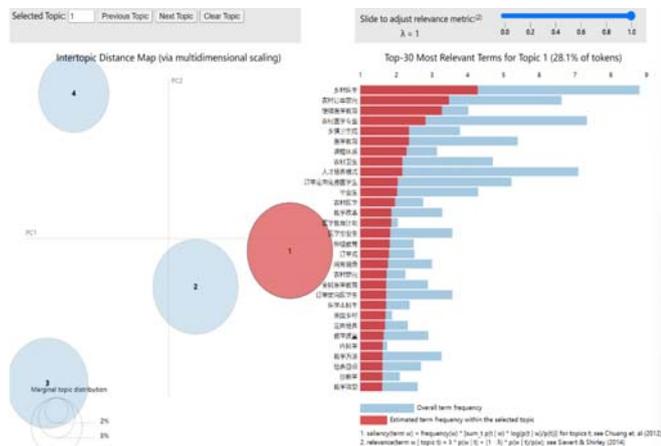


图4 pyLDAvis可视化结果

表1 乡村医学教育领域论文“主题-词”概率分布

主题	前 10 高概率特征词
农村医学人才培养模式(T1)	人才培养模式, 订单定向免费医学生, 免费医学生, 医学教育, 全科医学, 农村卫生人才, 农村基层, 医学人才, 农村医学专业, 健康教育
乡村医生继续医学教育(T2)	乡村医生, 继续医学教育, 农村医学专业, 乡镇卫生院, 毕业生, 医学教育计划, 继续教育, 美国乡村医生, 问卷调查, 教学改革
乡村振兴与高职医学生培养(T3)	高职医学生, 毕业生, 免费医学生, 乡村振兴, 农村基层, 职业认同, 医学人才培养, 学习动力, 职业道德教育, 健康乡村建设
农村医学专业实践教学(T4)	农村医学专业, 农村医学专业学生, 教学方法, 三年制, 实践教学, 乡镇卫生院, 课程设置, PBL 教学法, 实践教学体系, 订单定向全科医学生

表2 乡村医学教育领域论文“文档-主题”概率分布

文档编号	T1	T2	T3	T4	主题	年份
1	0.068 590	0.068 555	0.789 288	0.073 567	2	2023
2	0.141 647	0.118 424	0.696 565	0.043 364	2	2023
3	0.074 837	0.097 291	0.074 033	0.753 838	3	2023
4	0.813 205	0.059 659	0.057 379	0.069 756	0	2023
5	0.066 283	0.801 193	0.066 198	0.066 326	1	2023
...
705	0.041 148	0.876 492	0.041 052	0.041 307	1	1991
706	0.732 223	0.150 024	0.058 084	0.059 669	0	1991
707	0.784 786	0.071 381	0.071 394	0.072 44	0	1991
708	0.895 52	0.034 255	0.036 400	0.033 824	0	1990
709	0.052 120	0.842 302	0.051 860	0.053 718	1	1990

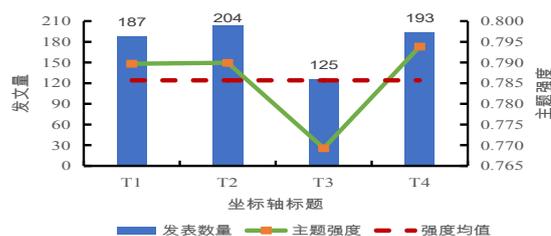


图5 各主题发文章量与主题强度

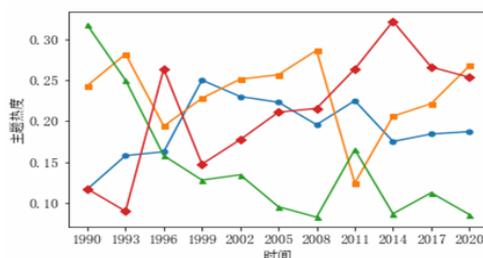


图6 四大主题热度随时间演化趋势