

基于 LDA-RF 的中职工科课程思政评价研究

梁芷欣¹ 刘文艺^{1*} 李自然¹ 姚成伟^{2,3}

1. 江苏师范大学机电工程学院, 江苏 徐州 221116

2. 江苏省铜山中等专业学校, 江苏 徐州 221100

3. 江苏省徐州市张集中等专业学校, 江苏 徐州 221100

[摘要]职业院校实现立德树人根本任务的关键路径在于深入实施课程思政建设, 科学、恰当的课程思政评价体系是提高课程思政教学质量的重要抓手。鉴于过往评估方式在客观性和复现性方面存在不足的情况, 本研究运用 LDA-RF 算法构建中职工科课程思政评价指标体系及模型。首先, 运用 LDA (潜在狄利克雷分配) 对职校课程思政相关政策文本与研究文献进行主题提取, 并以主题信息为指标构建基础; 其次, 通过研读文献和政策文本, 结合主题分析结果构建中职工科课程思政的评价指标体系; 最后, 通过对 RF (随机森林) 的训练和测试生成中职工科课程思政评价模型。研究表明, 中职工科教师的岗位分析与创新能力、课程思政教学水平以及学生学习投入与成效是影响课程思政教学效果的关键因素。

[关键词]课程思政; LDA 模型; 随机森林; 评价

DOI: 10.33142/fme.v6i2.15406

中图分类号: G711

文献标识码: A

Research on Ideological and Political Evaluation of Secondary Vocational Engineering Courses Based on LDA-RF

LIANG Zhixin¹, LIU Wenyi^{1*}, LI Ziran¹, YAO Chengwei^{2,3}

1. School of Mechatronic Engineering, Jiangsu Normal University, Xuzhou, Jiangsu, 221116, China

2. Jiangsu Tongshan Vocational School, Xuzhou, Jiangsu, 221100, China

3. Jiangsu Xuzhou Zhangji Secondary Specialized School, Xuzhou, Jiangsu, 221100, China

Abstract: The key pathway for vocational colleges to fulfill the fundamental task of fostering virtue through education lies in the thorough implementation of ideological and political education within courses. A scientific and appropriate evaluation system for ideological and political education within courses is a crucial means to enhance the quality of teaching in this area. Given the deficiencies in objectivity and reproducibility of previous evaluation methods, this study employs the LDA-RF algorithm to construct an evaluation index system and model for ideological and political education within secondary vocational engineering courses. Firstly, LDA (Latent Dirichlet Allocation) is used to extract themes from policy texts and research literature related to ideological and political education within vocational college courses, and these theme information serves as the foundation for constructing evaluation indicators. Secondly, by studying literature and policy texts, combined with the results of thematic analysis, an evaluation index system for ideological and political education within secondary vocational engineering courses is constructed. Finally, an evaluation model for ideological and political education within these courses is generated through the training and testing of RF (Random Forest). The research shows that the key factors influencing the effectiveness of ideological and political education within courses are secondary vocational engineering teachers' job analysis and innovation abilities, their teaching proficiency in ideological and political education, as well as students' engagement and learning outcomes.

Keywords: ideological and political education within courses; LDA Model; random forest; evaluation

在大数据背景下, “重视人工智能与教育教学实践的深度融合, 以促进教育创新与变革^[1]”已成为我国教育事业发展的关键趋势。中职院校亦需抓住数智化转型机遇, 运用数据深化课程思政建设, 实现学生专业知识技能与道德素养的全面提升^[2]。课程思政^[3]是落实立德树人任务的关键, 其评价体系的构建尤为重要。目前关于课程思政评价体系的研究主要基于学者的教学实践经验^[4], 针对职校课程思政效果的数智化测评较少。据此, 本文拟运用 LDA-RF 方法构建中职工科课程思政评价指标体系及模型,

以系统化评价思政教学效果, 助力教育工作者摆脱对传统教学经验的依赖^[5], 明确在大数据时代下所需的核心知识和素养。

1 中职工科课程思政效果评价体系

本文拟采用 LDA 模型深度剖析相关政策及学术文献^[6], 在此基础上将提炼的核心主题及关键词作为构建指标体系的依据。构建步骤如下: 一是收集并预处理政策要闻及文献数据, 确定最为适宜的主题数量; 二是对文本数据进行分词处理, 构建 LDA 模型选取主题与主题词; 最后通过

深入解读政策要闻和文献，增删调整备选指标，形成完整的评价指标体系。

1.1 数据采集路径

在构建本研究评价指标体系前，数据采集路径是其重要前提。本研究将中国知网、教育部、地方政策文件库及其相关标准作为数据来源，据此构建文献与政策数据集。通过文本分析，人工剔除无关失效文件，同时保留虽标题不直接相关，但内容对课程思政评价及职业教育发展具有紧密联系的文件，最终精选出 889 篇政策文本作为数据集（表 1）。

表 1 课程思政政策文献数据收集

效力位阶	发布来源	数量
部门规范性文件、新闻	中国教育部	183
地方工作文件、新闻	广东省教育厅等	603
行业标准及规范	中国职业技术教育学会	103
学术文献	中国知网	335

1.2 LDA 模型原理

LDA 模型是由 Blei 等人在 PLSA 模型基础上提出的一种无监督聚类算法^[7]，常用于快速识别非结构化复杂文本库中隐藏的主题信息。LDA 模型包含文档、主题和特征词三层结构：如图 1 所示，首先依据参数 α 的狄利克雷分布，为每篇文档抽取其主题分布 θ_m （ m 为文档数）；接着从对应文档的主题多项式分布中随机选择其所属主题 $Z_{m,n}$ （ n 为该文档中的词汇数）；然后利用参数 β 的狄利克雷分布，为每个主题生成一个特征词分布 φ_k 。最后，基于这些特征词的多项式分布 φ_k ，我们逐一采样，生成文档具体特征词 $w_{m,n}$ ，完成从主题到具体词汇的生成过程。

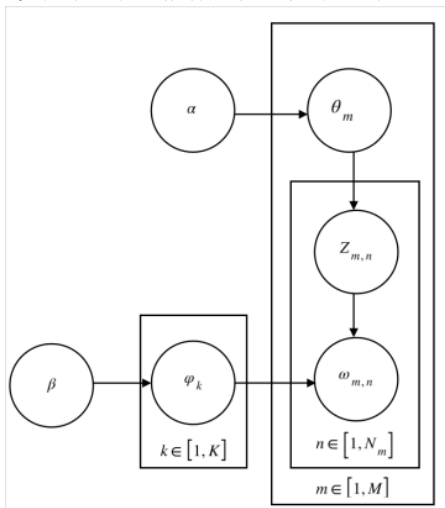


图 1 LDA 模型“文档—主题—特征词”生成流程

1.3 主题识别与结果分析

鉴于 LDA 模型的主题数目 k 无法直接确定，本研究将通过一致性变化情况来确定最优主题数目^[8]。图 2 为数据集在

迭代十四次后主题一致性指标变化情况。结果显示：主题数为 6 时一致性最佳。基于上述分析，将最优主题数设定为 6。

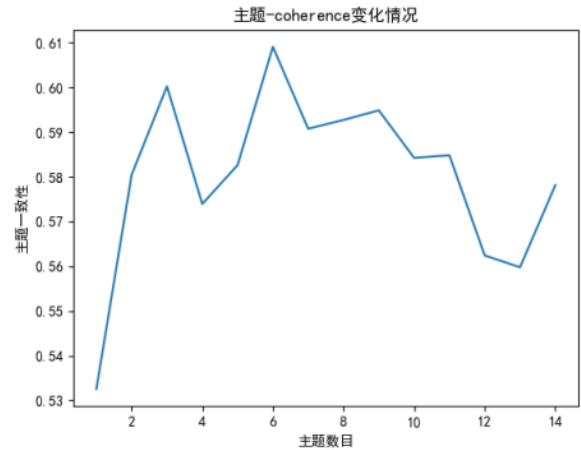


图 2 主题一致性

本研究基于数据集提炼出以下主题词，如表 2 所示，经深入理解和归纳，最终形成了以下六大主题：思政教学能力强调教学方法的运用与职业精神的培养，以及课程、教材、体系的合理选用；人才创新培养突出校企合作和产教融合模式，注重工匠精神的培养和企业项目实操；思政素养引领以立德树人为目标，注重学习新时代中国特色社会主义思想、党的二十大等重要讲话；思政知识基础强调中职工科教师应具备深厚的党史、红色文化等思政知识支撑；思政考核意识则关注党建工作机制的优化和考核制度的完善，以切实提高教师思政素养；学生就业发展重视对毕业生就业能力的指导，强调落实帮扶政策。这六大主题共同构成了评价体系的核心要素。

表 2 中职工科课程思政评价相关主题提取结果

序号	主题名称	主题词	主题强度
1	思政教学能力	教学方法 理论 实践 职业精神 课程思政 思政课 教师 课堂 教材 设计	0.25
2	人才创新培养	校企合作 产教融合 工匠精神 模式 职业教育 项目 企业 人才培养 产业 职业院校	0.21
3	思政素养引领	习近平 新时代 立德树人 中国特色社会主义 现代化 党 二十大 重要讲话 高质量 学习	0.17
4	思政知识基础	党史 红色文化 青年 师生 技术创新 故事 爱国主义 精神 历史 百年	0.163
5	思政考核意识	考核 机制 责任 党建 评估 师德 干部 党组织 指导 完善	0.135
6	学生就业发展	就业指导 毕业生 就业 职业技能 帮扶 创业 学生 职业技能 地区 乡村振兴	0.073

1.4 指标构建

中职工科课程思政效果评估的关键在于构建科学恰当的指标体系。由于课程思政建设与教学教育过程相融相通，具有全员全方位全过程育人特点，故本研究将中职工科课程

表 3 中职工科课程思政评价指标体系

一级指标目标层 (A)	二级指标准则层 (B)	指标说明
问题辨识与目标提出 (A1)	思政教学的问题识别力 (B1)	能敏锐发现、分析和判断教学过程中的问题、挑战
	明确思政教学目标 (B2)	能从专业课程性质、企业时代要求等角度设置思政目标
	教学总结意识 (B3)	能及时反思教学活动并做详实总结
关键信息的收集 (A2)	学生学情收集 (B4)	能对学生学习态度、基础、兴趣等做全面了解分析
	专业课思政元素收集 (B5)	能挖掘思政元素并结合时事热点收集思政素材
	对教学资源的洞察力 (B6)	能及时利用除教材外的教学资源, 更新教学资源库
	专业就业情况收集 (B7)	能多渠道了解当前就业动态, 灵活调整教学内容
课程思政教学内容 (A3)	合理性 (B8)	能准确恰当地选取思政元素、传达思政观点
	启发性 (B9)	能激发学习兴趣、增强情感体验, 形成正确价值观。
	实用性 (B10)	能贴合学生未来工作岗位, 融入新技术、新工艺
课程思政教学过程 (A4)	政治理念引领 (B11)	思政知识基础扎实, 能培养学生的政治认同感
	思政元素融入能力 (B12)	思政元素融入自然, 可促进知识内容的理解掌握
	学生参与情况 (B13)	学生专注认真, 积极发言、课后认真完成思政学习任务
	思政教学方法设计 (B14)	能结合新教学法和现代教育技术创新思政观点传达方式
课程思政教学效果及改进 (A5)	知识技能掌握情况 (B15)	学生课后任务完成情况良好, 学业成绩有进步
	思政育人情况 (B16)	学生学习满意度高, 在职业精神多方面显著提升
	思政教学改进能力 (B17)	教师能根据教学情况做改进举措, 及时更新教学案例库

思政教学流程作为依据, 遵循过程性导向设计思路^[9], 将该过程细分为问题辨识与目标提出、关键信息的收集、课程思政教学内容、课程思政教学过程、课程思政教学效果及改进五大维度并以此为一级指标。随后通过整合前期 LDA 主题模型提取到的关键信息, 基于各一级指标进行指标筛选和细化, 最终提炼出 17 个二级指标, 如表 3 所示。

2 中职工科课程思政效果评价模型

本研究采用随机森林多分类模型建立中职工科课程思政效果评价模型。步骤如下: 一是根据评价指标体系设计中职工科课程思政教学效果的测评问卷; 二是初步建立多分类随机森林模型^[10]并通过使用 GridSearchCV(网格搜索算法)获取最佳参数组合来训练模型; 三是将训练好的模型应用于测试样本集, 检验模型预测的准确性; 四是总结分析评价模型, 得出评价结果。

2.1 样本数据集构建及信效度评估

本研究基于评价指标设计调查评分问卷, 通过匿名网络形式将问卷发放到中职院校工科专业教师、学生群体中, 收集他们在各指标上的量化评分, 并将教师表现划分为四个等级。调研共计回收 201 份有效问卷, 有效回收率高达 93.05%。表 4 显示, 问卷信度系数高达 0.985, 效度评估显示 KMO 值为 0.980 且 P 值极低 (0.000), 证明了问卷的可靠性和一致性。基于上述检验结果, 本文将样本划分为训练数据集 (140 份) 与测试数据集 (61 份), 分别用于模型的构建和验证。

表 4 问卷信效度检验结果

检验类别	检验系数	系数值	
信度检验	克隆巴赫 Alpha	0.985	
	基于标准化项的克隆巴赫 Alpha	0.987	
	KMO 取样适切性度量	0.980	
效度检验	Bartlett 球形检验	近似卡方	5749.607
		自由度	153
		显著性	0.000

2.2 评价模型的训练优化和验证

本研究基于有效样本数据集, 在 Python3.8 环境下借助 Random Forest Classifier 工具包构建随机森林模型, 并使用 GridSearchCV(网格搜索交叉验证)工具对随机森林模型进行参数调优^[11]。表 5 为 4 个关键超参数的取值范围 and 对应优化结果。研究结果显示, 在随机森林中, 当决策树数量设定为 100 棵, 每棵树的最大深度限制为 5, 拆分内部节点所需的最小样本量设为 2, 且叶节点所需的最小样本数达到 1 时, 模型的性能分数达到最优值 0.9654。为验证优化后模型的准确性, 采用五折交叉验证方法, 结果显示, 模型的平均误判率为 3.45%, 证明该模型具备较高的预测可靠性。

表 5 随机森林超参数优化结果

超参数	取值范围	最佳值
决策树数量	100, 200, 300	100
最大深度	5, 10, 15	5
节点最小样本数	2, 5, 10	2
叶节点最小样本数	1, 2, 4	1

2.3 评价模型的测试和评估

在模型评价的测试环节,我们将经过优化的训练模型应用于数据库中的 61 份样本测试集,并采用随机森林算法的 OOB (袋外数据) 验证机制来衡量模型的性能。如图 3 所示,通过混淆矩阵结果,我们得出模型在 OOB 验证下的误判率为 2.52%,根据混淆矩阵可得出测试集的准确率 (Accuracy) 为 95.08%、精确率 (Precision) 为 95.4%、召回率 (Recall) 为 95.1%和 F1 值为 95.2%。

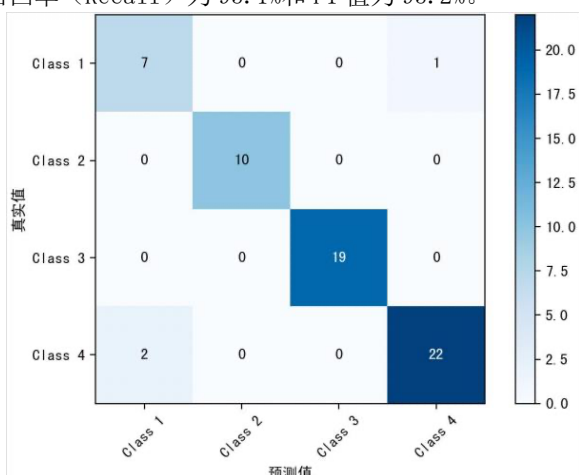


图 3 随机森林模型的袋外验证混淆矩阵

为验证随机森林评价模型是否具备优势,本研究使用训练样本集构建 KNN(K 最近邻)和 Adaboost (自适应提升算法)模型,最后通过三种模型在测试样本集中的表现做出判断。基于训练集训练后 3 个集成模型在测试集的性能表现见表 6。结果显示,检验样本经过随机森林模型处理后,显示出较高的分类准确率。同时精确率、召回率和 F1 值分别为 95.7%、95.1%和 95.2%,说明本研究构建的随机森林评价模型不仅整体分类准确率高,还表现出高精度性和高召回率,展现了该模型在处理复杂分类问题时的稳定性。

表 6 分类器性能指标

算法	准确率	精确率	召回率	F1 值
随机森林	95.08%	95.7%	95.1%	95.2%
KNN	93.44%	93.7%	93.4%	93.5%
Adaboost	85.71%	76.2%	85.7%	80.0%

2.4 模型结果分析

通过随机森林算法构建的多分类模型对本研究评价指标进行重要性排序,见图 4。

研究结果显示,在评价体系中,专业就业情况收集 (B7)、思政教学方法设计(B14)、思政元素融入能力(B12)、课程思政教学内容的实用性(B10)等指标占据关键位置。具体来说, B7 和 B10 指标体现了中职工科专业的职业导向特征,在问卷中,总体表现达到优秀的教师在 B7 指标的得分均大于 7,这说明教师岗位分析与创新能力是课程

思政教学效果的有力支撑。同时 B12、B14 指标在课程思政效果评价中占比显著,总体表现为欠佳的教师在 B12 和 B14 指标上的得分均低于 2.5,这表明教师课程思政教学水平在一定程度上决定课程思政评价效果的下限。

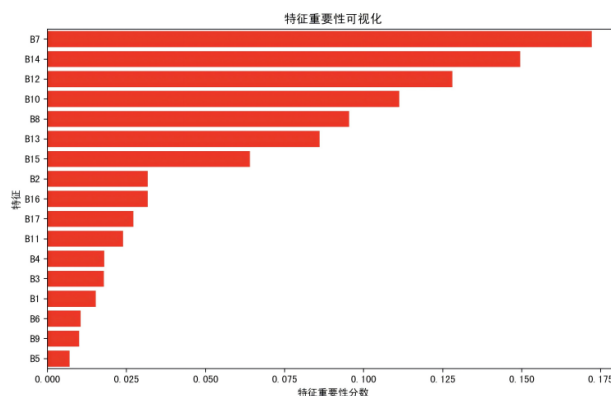


图 4 特征重要性排序

3 结论

本研究运用 LDA 主题模型提取政策与文献主题,并细分为 17 个二级指标,构建中职工科课程思政教学的评价指标体系。通过随机森林算法验证模型的科学性与适用性,显著提升了评价体系的全面性与客观性。综上所述,职业院校应抓住数智化转型机遇,利用信息技术完善评价体系,并从拓展校企合作深广度、打造教学支持平台、发挥校企协同作用等方面入手,增强中职工科课程思政教学成效。

基金项目:江苏师范大学研究生科研与创新计划项目 (2024XKT0697),江苏省现代教育技术研究 2022 年课题 (2022-R-101324),江苏师范大学研究生课程思政示范项目 (Y2021KCSZ06)。

[参考文献]

- [1]沈杰. 2024 年第二届教育创新与经验交流研讨会论文集[C]. 青海: 西宁市世纪职业技术学校, 2024.
- [2]王厥轩. 深化课程改革, 落实立德树人[J]. 上海教育, 2022(19): 20-21.
- [3]张金顺. 国外德育理论研究的新进展[J]. 广西民族大学学报(哲学社会科学版), 2006(6): 172-175.
- [4]侯雨欣, 侯建东. 课程思政协同发力中职思政教育——从大国工匠培育视角出发[J]. 教育学术月刊, 2020(10): 100-105.
- [5]齐磊磊, 李晨曦. 基于 LDA-BP 神经网络的高校思政教师数据驱动决策力评价研究[J]. 黑龙江高教研究, 2024, 42(3): 110-119.
- [6]Watanabe K, Batur A. Seeded Sequential LDA: A Semi-Supervised Algorithm for Topic-Specific Analysis of Sentences[J]. Social Science Computer Review, 2024, 42(1): 224-248.
- [7]Zhang J, Gui W, Wen J. China's policy similarity

evaluation using LDA model: An experimental analysis in Hebei province[J]. Journal of Information Science, 2024, 50(2): 515-530.

[8]朱红军,冯兰荣,李余石.基于 LDA 主题模型的林下经济政策文本分析[J].湖北林业科技,2024,53(3):49-54.

[9]胡晓晖,韩芳,董大奎.基于 CIPP 模式的高职专业教学质量评价指标体系构建研究[J].中国职业技术教育,2015(3):27-31.

[10]Ma X, Zhang F, Luo H, et al. Water shortage detection of jujube trees based on spectral technology and random forest classifier integrated

with fractal theory[J]. Industrial Crops & Products, 2024(1): 216118729.

[11]Alohali A M, Rashidy E N, Alaklabi S, et al. Swin-GA-RF: genetic algorithm-based Swin Transformer and random forest for enhancing cervical cancer classification[J]. Frontiers in Oncology, 2024(1): 141392301-1392301.

作者简介:梁芷欣(1997—),女,硕士研究生,主要研究方向为机械工程智能教育与管理;*通讯作者:刘文艺(1984—),男,教授,博士,硕士生导师,主要研究方向为机械工程智能教育与管理。