

# 机器学习实验教学中的思政元素融入： 探索与实践

张平凤<sup>1\*</sup>, 聂方彦<sup>2</sup>, 郑流云<sup>1</sup>

<sup>1</sup>贵州商学院马克思主义学院, 贵州 贵阳

<sup>2</sup>贵州商学院计算机与信息工程学院, 贵州 贵阳

收稿日期: 2025年2月16日; 录用日期: 2025年3月11日; 发布日期: 2025年3月26日

## 摘要

人工智能伦理教育与技术实践的深度融合已成为新工科建设的重要议题。针对传统机器学习实验教学“重技能轻价值”的困境, 研究构建“知识-能力-价值”三维目标体系, 创新提出“技术解构-社会关联-批判实践”的递进路径。通过双轨案例设计, 将数据偏见修正、算法可解释性等思政要素转化为可编程的模型约束条件, 结合仿真情境、跨学科协作及开源伦理实践, 形成“技术规范与价值判断”协同的教学范式。研究揭示了理工科课程思政的核心机制, 即依托技术实践的价值负载特征, 在模型调试中实现“伦理认知-价值内化”的转化。为实现教育创新制度化, 建议构建校企联动的动态案例库、开发嵌入式伦理评估工具包, 并建立涵盖技术性能与伦理敏感度的双重评价体系, 推动思政教育从课程改革向系统建构转型, 培养兼具技术理性与价值判断力的新型AI人才。

## 关键词

机器学习实验教学, 课程思政, 科技伦理, 批判性实践, 新工科

# Integrating Ideological and Political Elements into Machine Learning Experimental Teaching: Exploration and Practice

Pingfeng Zhang<sup>1\*</sup>, Fangyan Nie<sup>2</sup>, Liuyun Zheng<sup>1</sup>

<sup>1</sup>College of Marxism, Guizhou University of Commerce, Guiyang Guizhou

<sup>2</sup>Computer and Information Engineering College, Guizhou University of Commerce, Guiyang Guizhou

\*通讯作者。

文章引用: 张平凤, 聂方彦, 郑流云. 机器学习实验教学中的思政元素融入: 探索与实践[J]. 创新教育研究, 2025, 13(3): 305-315. DOI: 10.12677/ces.2025.133189

## Abstract

The deep integration of AI ethics education with technical practices has become a critical topic in the development of emerging engineering education. Addressing the challenge that traditional machine learning experimental teaching tends to emphasize skills over values, this study constructs a three-dimensional objective system encompassing knowledge, capability, and values. It innovatively proposes a progressive approach through “technical deconstruction-social association-critical practice”. By designing dual-track cases, it transforms ideological and political elements, such as bias correction in data and algorithm explainability, into programmable model constraints. Combining simulated scenarios, interdisciplinary collaboration, and open-source ethical practices, it forms a teaching paradigm where “technical norms and value judgments” are synergized. The research elucidates the core mechanism of ideological and political work in science and technology courses, highlighting the value-laden characteristics of technical practices to achieve the transformation from “ethical awareness to value internalization” during model tuning. To institutionalize educational innovation, it recommends establishing dynamic case libraries involving school-enterprise interaction, developing embedded ethical assessment toolkits, and creating a dual evaluation system covering both technical performance and ethical sensitivity, thereby fostering the transition from course reform to systematic construction in ideological and political education, and nurturing new types of AI talents who possess both technological rationality and value judgment abilities.

## Keywords

Machine Learning Experimental Teaching, Curriculum-Based Ideological and Political Education, Technology Ethics, Critical Practice, Emerging Engineering Education

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

人工智能技术的迅猛发展正在重塑全球产业格局与社会治理模式，而机器学习作为其核心驱动力，已成为高等教育中理工科人才培养的关键领域。然而，技术迭代的“加速度”与伦理规约的“滞后性”之间的矛盾日益凸显：算法偏见、数据隐私滥用、自动化决策的社会公平性争议等问题频发，暴露出单纯技术能力培养的局限性。在此背景下，中国高等教育提出“课程思政”建设要求，强调专业教育与价值观引导的深度融合[1]-[5]。

对于机器学习实验教学而言，其技术特性与思政教育存在天然契合点：从数据采集的伦理边界到模型优化的社会责任，从算法可解释性的人文关怀到技术应用的可持续发展目标，每一个实验环节均可成为价值塑造的载体。然而，当前教学实践中普遍存在“重技能轻伦理、重结果轻过程”的倾向，思政元素的融入往往流于形式，或陷入“技术 + 思政”的机械叠加困境。如何将社会主义核心价值观、科技伦理意识等思政元素有机嵌入实验教学目标、内容与评价体系，构建“价值塑造 - 知识传授 - 能力培养”三位一体的教学闭环，成为亟待突破的命题。

本文以机器学习实验课程为研究对象，聚焦以下核心问题：① 如何挖掘机器学习技术链(数据 - 算

法-应用)中隐含的思政教育触点? ② 如何在实验任务设计中实现专业技术训练与价值观引导的双向增强? ③ 如何建立可量化、可持续的课程思政教学评价机制? 通过理论建构与教学实证, 旨在为理工科实验类课程的思政融入提供可复制的方法论框架。

## 2. 融入逻辑与原则

### 2.1. 理论基础

课程思政的本质是通过专业知识的载体实现价值引领, 其核心在于打破传统思政教育与专业教育的“二元割裂”, 构建“知识传授”与“价值塑造”的共生关系[1]-[5]。对于机器学习实验教学而言, 其思政元素的融入逻辑需立足于两大理论维度: 一是课程思政的普遍性要求与理工科教学的特殊性矛盾, 二是机器学习技术的社会属性与伦理责任的内在关联性。

#### (1) 课程思政的内涵与理工科教学适配性

教育部《高等学校课程思政建设指导纲要》[6]明确提出, “所有高校、所有学科专业全面推进课程思政建设”, 但理工科课程因其强技术属性, 往往面临“思政融入难”的困境。传统思政教育以显性理论灌输为主, 而理工科实验教学更强调实践操作与逻辑验证, 二者在方法论上存在显著差异。因此, 课程思政在理工科领域的实践需遵循“隐性渗透”原则, 即通过技术实践中的伦理反思、社会问题映射和职业责任唤醒, 实现价值观的“无痕植入”。

以机器学习实验为例, 其教学过程天然包含多重思政触点: 数据采集阶段涉及隐私保护与知情权伦理, 模型训练阶段隐含算法公平性验证, 应用部署阶段则需考量技术的社会风险(如自动化决策对弱势群体的潜在排斥)。这种技术流程与伦理议题的深度交织, 为思政教育提供了“实践-反思”的天然场景, 使学生在解决技术问题的同时, 同步完成对科技社会价值的认知重构。

#### (2) 机器学习实验的伦理属性与社会价值

机器学习作为人工智能的核心技术, 具有显著的“双刃剑”特征: 一方面, 其通过数据驱动优化为医疗、教育、环保等领域带来革命性进步; 另一方面, 算法黑箱、数据偏见、技术滥用等问题可能加剧社会不平等。技术哲学家唐娜·哈拉维提出的“情境化知识”理论[7][8]指出, 技术开发必须嵌入具体的社会语境, 开发者需对技术后果承担“具身责任”。这一观点为机器学习实验教学的思政融入提供了哲学依据-实验教学不应局限于代码实现与模型调参, 而应引导学生建立“技术-社会”的系统认知框架。

从教育目标分类学视角, 布鲁姆的认知目标金字塔中“评价”与“创造”层级[9]-[12], 恰与思政教育中的价值判断能力形成呼应。例如, 在图像分类实验中, 学生不仅需要掌握卷积神经网络的原理, 更需思考训练数据集中是否存在性别、种族刻板印象, 并通过数据清洗、标签修正等技术手段主动消解偏见。这种“技术纠偏”过程本质上是一种价值观的实践表达, 使抽象的伦理原则转化为可操作的实验环节, 符合杜威“做中学”[13][14]的教育理念。

#### (3) 教育目标的三维统合模型

基于上述理论, 可构建如图1所示包含知识维度、能力维度以及价值维度的机器学习实验教学“三维目标统合模型”。该模型中, 知识维度指向掌握机器学习算法原理、编程工具与实验方法; 能力维度指向培养问题建模、工程实现与结果分析能力; 价值维度指向形成数据伦理意识、算法社会责任观与技术人文关怀。

此模型强调三者的动态平衡, 也即价值维度并非独立于技术训练之外, 而是通过“批判性实验设计”融入技术实践。例如, 在自然语言处理实验中, 要求学生对比不同语料库训练的情感分析模型, 分析语言偏见如何通过数据选择被算法固化, 进而讨论技术中立性迷思的局限性。这种设计使思政教育从“外部附加”转向“内生驱动”, 契合建构主义学习理论[15]-[17]中“意义协商”的核心机制。

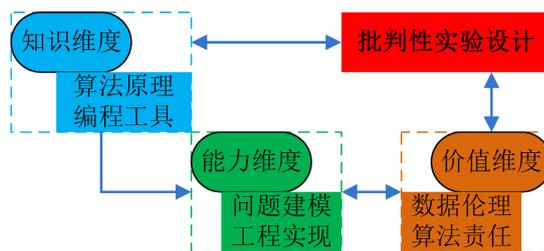


Figure 1. Diagram of 3D integrated model of educational objectives

图 1. 三维目标统合模型图

上述分析表明，机器学习实验教学的思政融入并非简单的“技术 + 思政”拼贴，而是基于技术伦理属性、教育目标分类与学习科学理论的系统性重构。其根本逻辑在于揭示技术实践背后的价值负载性，使学生在解决“如何实现” (How) 问题的同时，主动追问“为何实现” (Why) 与“为谁实现” (For Whom) 的深层命题，最终实现工具理性与价值理性的辩证统一。

## 2.2. 设计原则

机器学习实验教学的思政元素融入需遵循三大核心原则，确保价值引导与专业技术训练的深度融合。

### (1) 隐性渗透原则

避免思政内容的显性说教，通过技术实践中的伦理反思触点实现价值观的“无痕渗透”。例如，在数据预处理环节引入《个人信息保护法》条款，要求学生标注数据时主动识别敏感信息并匿名化处理；在模型评估阶段增设“公平性指标” (如不同性别、年龄群体的预测结果差异分析)，引导学生从技术维度验证算法正义性。此原则的关键在于将思政元素转化为可操作的技术规范，使学生在实验流程中自然形成伦理敏感性。

### (2) 问题导向原则

以技术痛点映射社会痛点，设计真实性思政场景。例如，针对推荐算法实验中常见的“信息茧房”现象，设置对比实验：一组基于用户历史行为数据训练模型，另一组引入多样性增强策略。通过 A/B 测试结果分析，学生可直观理解技术优化与信息平权之间的辩证关系。此类设计将抽象价值观具象化为可解决的技术问题，激发学生通过工程实践实现社会价值的双重驱动。

### (3) 知行合一原则

建立“实验任务 - 社会责任”的闭环链路，推动伦理认知向行动转化。例如，在完成人脸识别实验后，要求学生撰写《技术应用风险评估报告》，明确标注模型适用场景与禁用边界 (如禁止用于种族歧视或非法监控)。同时，通过校企合作引入真实案例 (如医疗影像算法的城乡诊断准确率差异)，组织学生团队设计“公平性增强”技术方案并提交至开源社区。此举强化责任担当意识，使思政教育从课堂认知升维至社会实践。

三大原则构成“意识唤醒 - 问题解决 - 行动落地”的递进链条，既呼应布鲁姆教育目标的高阶思维培养，又契合工程教育认证中的“复杂问题解决能力”要求，最终实现价值理性与工具理性的有机统一。

## 3. 思政元素融入路径

### 3.1. 目标重构

传统机器学习实验教学目标多聚焦于技术能力培养与知识验证，而缺乏对技术伦理、社会责任等价值维度的系统性规划。目标重构的核心在于构建“知识 - 能力 - 价值”三维目标体系，将思政元素从“附

加项”转化为“必选项”，形成可量化、可评估的教学导向。

#### (1) 传统目标的局限性

当前实验教学目标设计普遍存在两大盲区：① 价值目标缺位：实验指导书常以“准确率提升 0.5%”为成功标准，却忽视算法偏见、隐私泄露等伦理风险；② 评价维度单一：学生成绩多基于代码运行结果打分，缺乏对实验设计合理性、社会影响评估等软性指标的考察。

这种目标导向易导致学生陷入“技术至上主义”误区，将机器学习简化为数学优化问题，忽视其作为社会技术系统的复杂性。

#### (2) 三维目标体系的构建

基于 OBE 理念，重构后的目标体系需明确各维度的具体内涵与关联逻辑，即：

知识维度：掌握核心算法原理(如监督学习与无监督学习的适用场景)、工具链操作(如 PyTorch 框架的公平性评估库)。

能力维度：培养数据建模、实验设计、结果可视化等工程技术能力，以及算法可解释性分析、伦理风险评估等复合能力。

价值维度：树立三项核心意识，即① 数据正义观：理解数据采集的代表性偏差(如城乡数字鸿沟对训练集的影响)，主动规避歧视性数据标注；② 算法责任感：在模型开发中贯彻“预防性原则”，对高风险应用场景设置技术防火墙；③ 技术人文性：在性能优化之外，关注算法对用户尊严、文化多样性的潜在影响(如表情识别算法对少数族裔的误判率)。

#### (3) 目标落地的实施策略

分层设计：依据实验难度阶梯化设置价值目标。例如：① 基础实验(线性回归)：重点培养数据来源溯源性意识，要求标注数据集采集方式与授权协议；② 高阶实验(神经网络)：增加模型可解释性报告撰写，分析特征权重是否符合社会公平预期。

量化工具：开发思政融合度评价量化准则，从“伦理敏感度”“社会影响分析深度”“解决方案创新性”三个指标进行 5 级评分。例如，在聚类分析实验中，若学生主动提出通过“差分隐私”技术保护用户轨迹数据，则在“解决方案创新性”维度给予加分。

案例驱动：以医疗诊断模型实验为例，除要求达到临床级准确率外，增设“弱势群体覆盖度”考核项(如农村患者数据的纳入比例)，迫使学生在技术优化中平衡效率与公平。

### 3.2. 内容设计

机器学习实验的思政元素融入需以技术链解构与社会问题映射为核心策略，通过重构实验内容载体，将价值观引导嵌入技术实践全流程。以下从案例选择、实验环节、方法论三个层面展开设计：

#### (1) 案例选择：从技术痛点映射社会痛点

以现实争议性场景为切入点，构建“技术 - 伦理”双主线实验案例库：

数据层：选用医疗健康数据集(如糖尿病预测)，要求学生分析城乡患者数据分布的差异，探讨“数据偏见如何导致医疗资源分配算法加剧地域不平等”。

算法层：设计基于用户行为数据的个性化推荐实验，对比“点击率优化模型”与“多样性增强模型”的差异，揭示“信息茧房”的技术生成机制及破解路径。

应用层：引入人脸识别门禁系统案例，模拟不同光照条件与种族特征的识别准确率测试，触发对“技术滥用边缘”的边界讨论。

此类案例将抽象伦理议题转化为可编程、可验证的技术问题，使学生在模型调参中自然触碰价值观冲突点。

## (2) 实验环节：植入伦理审查与技术纠偏机制

在标准实验流程中增设思政增强节点，形成“技术实现 - 伦理反思 - 行动优化”闭环：

数据预处理阶段：强制要求标注数据来源合法性声明；增加“偏见检测”环节，利用工具包分析数据集的性别、年龄分布均衡性。

模型训练阶段：引入公平性约束项，对比约束前后模型在不同群体的性能差异；要求可视化特征权重，分析“高权重特征是否隐含歧视性关联”（如贷款预测模型中邮政编码与种族的相关性）。

结果评估阶段：除准确率、F1 值外，增设公平性指标；撰写《技术应用社会风险评估报告》，明确模型适用场景与禁用红线。

## (3) 方法论：构建“批判性实验”教学模式

对比实验法：例如在自然语言处理实验中，分别使用包含刻板印象的语料库（如“护士 - 女性”“工程师 - 男性”）与去偏见语料库训练情感分析模型，通过结果差异直观呈现语言偏见的技术传导路径。

角色扮演法：分组模拟“算法开发团队”“受影响社群代表”“政府监管部门”，就自动驾驶伦理决策规则展开辩论，形成兼顾效率与公平的算法改进方案。

开源贡献任务：鼓励学生将实验中开发的公平性增强代码（如数据脱敏工具）提交至 GitHub 开源社区，推动伦理实践从课堂向行业延伸。

### 3.3. 方法创新

为实现思政元素与机器学习实验的深度耦合，需突破传统“教师演示 - 学生复现”的单一教学模式，构建“价值引导 - 技术实践 - 社会行动”三位一体的方法论体系，具体通过以下四类创新方法实现。

#### (1) 情景模拟教学法：技术决策的伦理沙盘

以真实社会争议为背景，设计角色扮演实验框架，例如在自动驾驶伦理决策实验中：

角色分配：学生分组扮演算法工程师、交通管理部门、弱势道路使用者（如儿童、残障人士）等利益相关方。

冲突模拟：针对“电车难题”场景（如紧急避让时优先保护乘客还是行人），要求基于不同价值观（功利主义/义务论）编写决策规则代码。

共识构建：通过多轮 ABM (Agent-Based Modeling) 仿真，观察不同规则下的系统级社会影响，最终形成兼顾效率与公平的技术方案。

该方法通过具身化参与与激发同理心，使学生理解技术决策的复杂性远超代码逻辑本身。

#### (2) 反思性实践法：实验报告的价值升维

在传统技术报告模板中嵌入伦理反思模块，形成结构化输出框架。

技术层：模型准确率、损失函数曲线等常规分析。

伦理层：使用“伦理影响矩阵”评估技术应用的四个象限，即受益群体（如老年人通过医疗 AI 获得诊断便利）、受损风险（如算法误诊导致的法律责任）、权力关系（如平台通过推荐算法操控用户注意力）、长期效应（如自动化加剧职业替代危机）。

行动层：提出“最小化伤害”改进方案（如加入人工复核接口）。

通过该框架，学生的反思焦点从“模型为何失效”转向“技术应为何服务”，契合批判性教育学的核心主张。

#### (3) 跨学科协同实验法：打破技术孤岛

联合法学、社会学等专业共建复合型实验项目，例如：

隐私计算与法律合规实验：与法学院合作，设计联邦学习框架下的医疗数据共享实验，要求同时满

足《个人信息保护法》的“最小必要原则”与模型性能要求。

算法审计与社会调研实验：在社会学家指导下，对校园招聘算法进行逆向工程，分析筛选规则是否隐含性别偏好，并基于田野调查数据提出去偏策略。

此类实验通过学科视角碰撞，迫使学生直面技术的社会嵌入性，培养系统性思维。

(4) 开源伦理行动法：从课堂到社会的责任延伸

将实验成果转化为可落地的伦理工具包，例如：在 GitHub 发布模型时，强制要求附《技术社会影响说明书》，明确使用场景限制与公平性测试结果；与 NGO 合作，将学生设计的“低收入群体信用评估模型”应用于小微贷款试点项目。通过连接开源生态与社会需求，使伦理实践从课堂认知升维至公民行动，呼应“责任式创新”的全球共识。

## 4. 实践案例与效果

### 4.1. 典型实验设计

案例名称：基于用户画像聚类的数据伦理实验。

技术框架：K-means 聚类算法(Python/scikit-learn)。

思政结合点：隐私保护意识、群体标签化风险、算法社会责任。

(1) 实验设计背景

用户画像聚类是电商、社交平台等场景的通用技术，但其滥用可能导致“数据暴力”(如过度收集、标签歧视)。本实验以某电商用户行为数据集为基础，通过技术实现与伦理反思的双路径，引导学生直面算法背后的权力关系与社会责任。

(2) 实验流程与思政融入节点

数据预处理阶段：① 技术任务：清洗原始数据(用户 ID、购买记录、浏览时长等)，标准化处理。② 思政设计：引入《个人信息保护法》条款，要求对用户 ID 进行哈希脱敏，并标注数据采集的知情同意状态(如“匿名化处理”或“需二次授权”)。③ 冲突点：学生发现原始数据包含用户地理位置(精确至区县)，需自主决定是否保留该字段以提升聚类效果，或删除以避免地域歧视风险。

模型训练阶段：① 技术任务：通过肘部法则确定聚类数，生成用户群体标签(如“高消费活跃青年”“低频折扣敏感群体”)。② 思政设计：使用 Fairness Indicators 工具包分析聚类结果，检测敏感属性(如性别、年龄)与群体标签的关联强度。例如，若“高消费”群体中女性占比显著高于男性，需重新评估特征选择是否隐含性别偏见。

结果应用阶段：① 技术任务：基于聚类结果设计精准营销策略。② 思政设计：模拟企业决策会，学生需分别扮演“算法工程师”“消费者权益代表”“法务顾问”，辩论“是否向‘低频消费’群体推送高利贷广告”。通过多利益方博弈，揭示技术方案的社会外部性。

(3) 教学实施与效果

课前导引：播放纪录片《社交媒体陷阱》，讨论“大数据杀熟”案例，激发伦理认知冲突。

课中实践：① 技术组：完成聚类模型构建与可视化；② 伦理组：撰写《数据使用合规性审查表》，标注潜在风险点。

课后延伸：提交《负责任用户画像白皮书》，要求从技术优化(如差分隐私)、政策建议(如标签透明度披露)双维度提出改进方案。

本实验通过“技术 - 伦理”双主线设计，证明机器学习实验可成为思政教育的天然载体。其成功关键在于将法律条款、社会争议转化为可编程的技术约束，使价值观引导从抽象说教落地为工程实践。

## 4.2. 教学实施流程

以“用户画像聚类实验”为例，教学实施分为三阶段九环节，通过“认知冲突 - 实践反思 - 行动延伸”的螺旋式设计，实现思政元素的全流程渗透。

### (1) 课前：情境创设与伦理认知激活

社会事件导入：发布预习材料《大数据杀熟背后的算法逻辑》，结合央视报道的“外卖平台差异化定价”案例，引导学生思考“技术效率与社会公平的边界”。

伦理冲突预判：要求学生提交《初始立场声明》，回答“你是否愿意为更高的聚类准确率使用未经完全授权的用户数据？”通过匿名投票统计，暴露学生初始价值观分布。

技术 - 伦理任务包下发：分组领取双路径任务，即① 技术组：学习 scikit-learn 聚类 API 与可视化工具；② 伦理组：研读《个人信息保护法》与《互联网用户画像安全规范》。

### (2) 课中：双轨并行与冲突协商

技术实践线(120分钟)：① 数据清洗：处理含用户性别、地域、消费记录的原始数据；② 聚类建模：通过轮廓系数评估确定最佳聚类数；③ 标签生成：定义“高端消费者”“价格敏感群体”等画像标签。

伦理反思线(60分钟)：① 数据合规审查：使用开源工具检测数据匿名化强度；② 偏见诊断：分析“高端消费者”标签中一线城市用户占比；③ 角色辩论：模拟“平台算法部”与“消费者协会”就标签歧视风险展开质询。

协同决策(30分钟)：投票表决是否调整聚类特征(如删除“地域”字段)，重新训练模型并对比社会影响评估指标。

### (3) 课后：行动延伸与社会连接

技术优化：根据课堂共识修改模型，如对农村用户数据过采样，或添加地域公平性约束项。

伦理倡导：制作《负责任的用户画像指南》信息图，发布至校园自媒体，吸引校内外不同受众群体关注。

跨界对话：邀请法律顾问与算法工程师开展线上研讨会，学生就实验中的合规争议进行跨界质询(如“用户画像的知情权如何落地”)。

## 4.3. 成效评估

为全面验证思政元素融入机器学习实验教学的效果，采用“过程 - 结果 - 影响”三维评估框架，结合量化数据与质性分析，形成闭环反馈机制。

### (1) 多维评估体系构建

技术能力指标：模型准确率、代码规范度等传统技术评分(占比 60%)。

伦理实践指标：基于《思政融合度评价量规》评分(占比 30%)，包括数据合规性审查、社会影响分析深度、改进方案创新性。

价值观内化指标：通过反思日志文本分析、课堂辩论参与度、课外伦理倡导行动等质性评价(占比 10%)。

### (2) 量化结果分析

伦理敏感度提升：实验报告显示，学生主动识别数据隐私风险的比例从初期的 28% 提升至 89%，使用公平性评估工具(如 Fairlearn、AI Fairness 360)的覆盖率从 12% 增至 76%。

技术方案优化：在聚类实验的最终提交代码中，73% 的团队自主加入差分隐私或去偏约束项，对照组(未接受思政融合教学)仅为 9%。

长期行为改变：跟踪调查表明，参与课程的学生在后续毕业设计中，65% 的项目包含技术伦理论证章

节，远高于全校平均水平。

### (3) 质性反馈洞察

反思日志高频词分析(基于 Python/NLTK 情感分析工具): Top5 关键词(① 知情同意(词频 0.32), ② 算法权力(0.25), ③ 技术谦抑性(0.19), ④ 社会脆弱群体(0.15), ⑤ 预防性原则(0.11))。

深度访谈摘录: 例如, “以前我只关心让模型跑得更快, 现在会本能地质问: 这个特征选择会不会伤害某些人?” (计算机科学与技术专业学生 A); “角色扮演让我意识到, 算法工程师的一行代码可能影响千万人的生存境遇”(数据科学与大数据技术专业学生 B)。

### (4) 横向对比与纵向追踪

跨班级对比: 与未实施课程改革的平行班级相比, 实验班学生在“科技伦理敏感性测试”中的得分高出 41%。

行业反馈: 合作企业评价, 实习生的技术方案中“法律合规性审查完整性”提升 55%, 极大促进了实习学员在构建技术方案中考虑科技伦理因素的主动性。

## 5. 反思与改进

### 5.1. 关键经验

#### (1) 思政触点与技术痛点的精准映射

实验证明, 思政元素的渗透需以技术链关键节点为锚点, 避免泛化说教。例如, 在数据预处理阶段聚焦隐私合规性, 在模型优化阶段绑定算法公平性, 使伦理反思与工程实践形成逻辑闭环。某医疗影像分类实验中, 通过对比“城乡患者数据集偏差对诊断准确率的影响”, 学生自发提出“数据补偿采样 + 区域公平性约束”方案, 验证了技术痛点与价值观引导的共振效应。

#### (2) 灰度决策空间中的价值观淬炼

刻意保留技术伦理的矛盾张力(如“模型效率与隐私保护的取舍”), 能有效激发深层次价值思辨。例如, 在自然语言处理实验中, 允许学生在 0%~20%敏感数据脱敏强度间自主选择, 并通过 AB 测试分析不同选择的社会代价。此类灰度设计促使学生从“非黑即白”的伦理判断转向复杂情境下的责任权衡, 84%的学生在反思日志中承认“技术决策需要妥协智慧”。

#### (3) 轻量化融合策略降低认知负荷

通过模块化思政工具包(如数据偏见检测插件、伦理影响矩阵模板), 将价值观实践转化为可复用的技术动作。轻量化设计确保思政融入不增加额外课业负担, 契合理工科学学习者的效率偏好。

#### (4) 跨界资源整合提升实践真实感

校企合作案例库与跨学科实验项目的引入, 打破了课堂与社会的认知壁垒。真实场景的冲击力远胜理论说教, 是价值观内化的催化剂。

### 5.2. 持续改进方向

为实现思政教育与机器学习实验教学的深度协同, 可以从以下方向持续优化。

#### (1) 校企合作伦理案例库建设

联合科技企业与法律机构, 开发行业级技术伦理案例库, 涵盖金融风控、医疗诊断、智慧城市等场景。例如, 与互联网法院合作, 将“算法价格歧视诉讼案”转化为实验任务, 要求学生从技术逻辑与法律条款双维度提出合规方案。案例库需配备动态更新机制, 实时纳入新兴技术争议(如生成式 AI 版权问题), 确保教学内容的前沿性与实践性。

#### (2) 完善多维度思政评价体系

构建“过程-结果-影响”三级评价框架,完善框架内以下几方面内容。

过程性评价:利用学习分析工具(如 Jupyter Notebook 插件),自动追踪学生代码中的伦理实践行为(如数据脱敏函数调用频次)。

结果性评价:引入伦理认证标准,对实验报告中的社会影响分析进行星级评分。

长效性评价:建立毕业生追踪机制,通过雇主反馈评估职业场景中的伦理决策能力。

### (3) 提升教师跨学科能力

设计“技术-伦理”双轨师资培训体系提升教师跨学科能力,为达此目的,可从以下几方面着手。

构建技术伦理工作坊,邀请科技哲学家、法律专家解析 AI 伦理前沿问题;教学法转型培训,推广“批判性实验设计”方法,帮助教师从“技术传授者”转向“价值引导者”;建立跨专业教研共同体,建立开放共享的思政实验教案库,鼓励教师协作开发模块化教学资源(如可插拔的伦理审查代码模块)。

## 6. 结语

人工智能技术的深度社会化,使机器学习教育必然承载价值塑造使命。本文通过重构目标体系、创新实验设计、开发评估工具,验证了思政元素与专业技术训练的融合路径:以技术链痛点映射社会伦理议题,将价值观引导转化为可编程的工程实践。实践表明,当数据偏见检测、算法公平性约束等伦理行动内化为实验环节时,学生既能提升复杂问题解决能力,又可形成“技术向善”的责任自觉。

本研究为理工科课程思政建设提供了可复制的范式,也即思政不是机械叠加理论说教,而是通过“批判性实验设计”揭示技术的价值负载性,在解决“如何实现”的同时追问“为何实现”。未来需进一步推动校企共建伦理案例库、完善多维度评价体系、构建跨学科教研生态,使伦理教育从课堂延伸至行业实践,为人工智能时代培育兼具创新力与责任感的“新工科”人才。

教育的本质在于唤醒人的反思性。当学生在调试代码时思考数据正义,在优化模型时警惕算法霸权,让每一次技术选择都透射人文的温度。这既是课程思政的初心,亦是技术教育的终极使命。

## 基金项目

贵州商学院 2023 年度校级一流专业《数据科学与大数据技术》建设项目(2023XJYZ01)。

## 参考文献

- [1] 施晓秋. 产出导向的工程教育类专业课程思政体系构建[J]. 高等工程教育研究, 2025(1): 75-82.
- [2] 王红雨, 闫广芬. 工科专业课教师课程思政参与行为及其影响路径——基于计划行为理论的分析[J]. 高等工程教育研究, 2025(1): 83-90.
- [3] 郑石. 重温经典助力高校课程思政提质创新——“经典进校园·文化共传承”活动评述[J]. 中国新闻传播研究, 2024(4): 239-244.
- [4] 徐曼, 郑宏宇. “课程思政”政策何以成为国家行动?——基于多源流理论的分析[J]. 河南师范大学学报(哲学社会科学版), 2024, 51(6): 42-49.
- [5] 王淑娉, 查亮亮. 高校课程思政实施: 打开“传知”与“育人”理论症结[J]. 中国高等教育, 2024(21): 60-64.
- [6] 教育部. 教育部关于印发《高等学校课程思政建设指导纲要》的通知[EB/OL]. [https://www.gov.cn/zhengce/zhengceku/2020-06/06/content\\_5517606.htm](https://www.gov.cn/zhengce/zhengceku/2020-06/06/content_5517606.htm), 2020-05-28.
- [7] 洪晓楠, 郭丽丽. 唐娜·哈拉维的情境化知识观解析[J]. 东北大学学报(社会科学版), 2012, 14(2): 95-100.
- [8] 战洪飞, 鄢益男, 余军合, 王程强. 面向业务求解的情境化知识模块建模方法研究[J]. 情报理论与实践, 2018, 41(7): 149-154, 122.
- [9] 周子琦, 高飞, 方春晖, 赵一帆. 基于布鲁姆认知分类的大学生 AI 依赖风险分析与对策研究[J/OL]. 云南民族大学学报(自然科学版): 1-7. <http://kns.cnki.net/kcms/detail/53.1192.N.20250106.1122.002.html>, 2025-02-16.

- 
- [10] 方盛汉. 基于“布鲁姆认知目标”的雨课堂教学模式实践——以“中国古代文学”为例[J]. 池州学院学报, 2024, 38(5): 140-142.
- [11] 吴本祥, 陈亮, 李耕, 张雪妍. 基于布鲁姆教学目标的船舶电气实践课程改革研究——以“电气数字控制技术及综合设计”课程为例[J]. 教育教学论坛, 2024(40): 131-134.
- [12] 李可. 基于布鲁姆认知层次理论的高职美育模块化教学策略研究[J]. 美术教育研究, 2024(17): 143-145.
- [13] 涂诗万, 朱凯. 作为社会理论的“做中学”——深化杜威研究的一个新尝试[J]. 华东师范大学学报(教育科学版), 2023, 41(6): 14-25.
- [14] 黄英杰. 杜威的“做中学”新释[J]. 课程·教材·教法, 2015, 35(6): 122-127.
- [15] 黄凌梅, 钟秉林. 建构主义视域下实习辅助课的启示[J]. 中国大学教学, 2020(6): 64-69, 81.
- [16] 孙庆国, 王佩国. 基于建构主义学习理论的服装专业实践教学体系构建[J]. 艺术工作, 2018(6): 106-107.
- [17] 叶欣. 基于建构主义理论的教学设计反思——以国防教育教学为例[J]. 教育理论与实践, 2015, 35(33): 54-55.