

# 基于专业知识图谱的大数据专业课程体系优化与实践研究

周 勇 陈 欣 廖 宁 司利红

重庆工程学院 重庆 400056

**摘要:** 本研究针对大数据专业课程体系存在的教学内容重叠、产业适应性不足等问题,提出了一种基于专业知识图谱的课程优化方法。通过产业需求调研、专业知识图谱整合优化、目标达成度体系设计,实现了数据科学与大数据技术专业核心课程的系统性优化。实践表明,该方法能有效提升课程与产业需求的匹配度,为应用型本科高校专业建设提供了可操作的实践路径。

**关键词:** 专业知识图谱; 大数据专业; 课程体系优化; 教学目标达成度

## 引言

随着国家数字中国战略持续推进和新工科建设不断深化,数据科学与大数据技术专业已成为支撑产业转型升级的关键领域。然而,在高校新工科建设实践中,数据科学与大数据技术专业的发展仍面临诸多挑战。首先,课程内容体系性不足,多门核心课程之间存在内容交叉,导致资源分散重复、学生知识结构断裂。其次,教学内容与快速迭代的产业需求脱节,AIGC等新兴技术未能系统融入课程体系,毕业生技能结构与市场岗位要求存在错配。再者,人才培养质量的评价机制亟待革新,传统定性描述和主观评价方式,难以对毕业要求达成度实现情况进行量化诊断与监控。

针对上述问题,本研究聚焦课程体系如何对接动态产业需求和人才培养质量如何实现科学量化评价问题,以重庆工程学院数据科学与大数据技术专业为研究对象,探索基于专业知识图谱系统化解决方案。

**基金资助:** 重庆市高等教育教学改革研究项目基于专业知识图谱的大数据专业课程体系优化与应用研究(项目编号:243296);重庆市教育科学规划课题知识图谱驱动的个性化学习路径研究-以《大数据采集与处理》课程为例(项目编号:K24ZG2190318)

## 1 课程体系重构

为解决与产业需求脱节的核心矛盾,对课程体系进行了系统性重构。重构工作严格遵循需求分析、能力定义、课程映射的路径,其核心产出是以实证调研为基础的《大数据专业能力培养与课程支撑矩阵》。

### 1.1 基于岗位数据的技能需求画像

为确保科学性,项目组采用了线上数据挖掘、线下实地访谈相结合的方法,对大数据人才需求进行深入挖掘。线上数据挖掘主要通过Python爬虫技术,对主流招聘网站发布的、与大数据相关的岗位信息进行采集与清洗,累计获得有效数据58,600余条。利用文本挖掘方法,重点分析了金融科技、数据分析、数据开发等领域岗位描述中的核心技能关键词。线下实地访谈,深入走访了成渝地区10余家合作企业(包括金融机构及多家科技公司),与业务部门负责人进行了深度交流。同时,项目组还对重庆市8所兄弟院校进行了调研,了解其课程设置与产业对接经验。

上述调研形成了《大数据专业岗位技能需求矩阵》,如表1所示。该矩阵清晰地揭示了开发、分析、治理三大核心岗位方向对技能的不同要求:开发岗位强调分布式系统(Hadoop/Spark)与实时处理(Flink);分析岗位聚焦于数据分析(SQL/Python)与机器学习建模;而数据治理岗位则对数据安全、质量管理及行业法规表现出极强的刚性需求。

表1 大数据专业岗位技能需求矩阵

岗位	核心技能项	需求热度	技能类型	在金融领域的典型应用场景
大数据开发	Python/Java	☆☆☆☆☆	核心基础	金融数据接口开发、量化交易系统
	Hadoop/Spark	☆☆☆☆☆	核心基础	海量交易数据、用户行为数据的分布式处理
	数据仓库/Hive	☆☆☆☆	核心基础	构建金融主题数据仓库，支撑分析与报表
	Flink/Kafka	☆☆☆	新兴趋势	实时反欺诈、实时风险监控、流式数据处理
大数据分析	SQL	☆☆☆☆☆	核心基础	核心金融业务数据查询与提取
	Python	☆☆☆☆☆	核心基础	数据分析、建模脚本编写
	数据可视化	☆☆☆☆	核心基础	企业经营报表、风险看板、客户画像可视化
	机器学习	☆☆☆☆	转型关键	信用评分模型、欺诈检测、智能投顾
	大模型应用 / Prompt工程	☆☆	新兴趋势	智能金融客服、合规文档审阅、对话式BI
数据治理	数据安全与隐私	☆☆☆☆	核心基础	客户信息脱敏、数据访问权限控制
	数据质量管理	☆☆☆☆	核心基础	确保金融报表数据、风控数据准确性
	数据标准与法规	☆☆☆☆☆	行业刚需	遵循《个人信息保护法》、金融行业数据规范

表2 大数据专业能力培养与课程支撑矩阵

目标岗位	技术核心课程	金融应用课程
大数据开发 (培养金融数据平台与系统构建能力)	《分布式存储技术》	《数据智能体开发》
	《分布式计算》	《大数据应用开发项目实训》 (实时风控、量化交易系统)
大数据分析 (培养金融数据建模与决策支持能力)	《数据分析与可视化》	《商业智能分析》 (经营分析看板、营销效果评估)
	《机器学习》	《财务数据挖掘实训》 (企业信用评级、欺诈交易识别)
数据治理 (培养金融数据资产管理与合规能力)	《数据仓库》(内含数据集成、质量管理)	《数据质量评估技术》
		《数据发现与血缘分析技术》 (金融审计与监管合规要求)

## 1.2 形成技术-应用双轨支撑的课程矩阵

基于需求分析结论，以成果导向理念为核心，将抽象的岗位技能转化为具体、可落地的课程培养目标，并设计了全新的课程体系方案。核心是构建了纵向技术进阶、横向应用融合的课程支撑矩阵，如表2所示。

基于矩阵，进一步完成了课程内容的优化设计。依据能力映射关系，对《分布式存储技术》《数据分析与可视化》《数据仓库》等专业核心课程的教学大纲进行了修订，规划了金融领域典型案例与AI工具的融入，并明确了重复知识点的整合方案。同时，为强化金融大数据培养特色与前瞻性，在课程体系中增设了《金融风控建模》等特色模块，并将大模型应用等新兴技术趋势设置为独立教学单元。

## 2 专业知识图谱整合与优化

在课程体系重构的基础上，本研究对原有分散的课程知识图谱进行了系统性整合、优化与升级。依托超星学习通平台，以《大数据专业能力培养与课程支撑矩阵》为统一框架，对《大数据采集与处理》《分布式存储技术》《机器学习》等核心课程的课程级知识图谱进行深度集成，消除信息孤岛，构建了服务于新课程体系的专业级知识图谱。

### 2.1 以重构矩阵为纲领的图谱对齐

整合工作以课程矩阵为顶层指导，对原分散图谱进行目标与结构对齐。在目标层面，将各课程教学目标统一映射至开发、分析、治理三大岗位能力方向；在内容层面，依据技术核心课与金融应用课两大课程群，重组知识模块归属，使图谱结构反映新的课程逻辑。

### 2.2 基于大模型的跨图谱知识点治理

在对齐后的课程结构基础上,引入DeepSeek大模型作为智能治理引擎,对多源课程图谱中存在的语义不一致与内容重叠问题进行深层处理。该引擎通过语义理解与上下文分析,智能识别并处理同名异义知识点的概念消歧,实现同义异名知识点之间的自动合并;同时,结合产业需求矩阵,对冗余或过时的孤立知识点进行智能剪裁。

经过消歧、合并、剪裁的系统化治理流程,原本分散、重复、边界模糊的知识点得以精细化提炼,形成了一套概念清晰的独立知识单元集合,最终构建出专业层面的统一知识库。典型案例如数据分析与可视化与机器学习课程边界不清,15个知识点存在语义重叠;大数据采集与处理与数据分析与可视化课程,在数据清洗模块出现内容重复,经智能治理后均得到有效整合与结构优化。

### 2.3 图谱评估

图谱评估采用了结构、一致性与有效性三维方法。通过图谱拓扑分析验证能力、课程、知识点层级结构与课程间逻辑衔接;通过语义一致性检测确保知识点定义、属性及关联全局无矛盾;并邀请课程教师与行业专家对关键知识路径进行教学合理性与岗位吻合度评审。结果表明,图谱消除18%冗余知识点,显著提升了知识连通性与教学清晰度。关键学习路径被确认符合认知规律与职业成长需要。

## 3 目标达成度体系构建

为科学量化人才培养质量并破解多课程、多数据源评价的实践难题,本研究突破传统单一评价范式的局限,构建了一套“数据驱动、分层实施、动态反馈”的目标达成度体系。该体系以专业知识图谱为结构化框架,以多源考核数据为输入,构建了覆盖从课程终结性诊断到教学过程实时监测的分级评价闭环。

### 3.1 统一框架下分层评价模型

遵循“基线统一,深度分层”的核心设计原则,平衡评价的全面性、可行性。具体由两个协同层级构成:终结性评价层面向所有专业核心课程,以期末试卷分析为统一基线,在课程结束后实施系统性教学体检,识别共性薄弱环节;过程性评价层则选取《大数据采集与处理》等建设基础好的课程进行试点,贯穿教学全周期进行数据采集与实时分析,实现学期中的教学监测与即时干预。两层评价均

以专业知识图谱为基础,确保评价结果可追溯,共同支撑教学持续改进。

### 3.2 基于试卷分析的课程级统一诊断

课程级诊断依托统一的终结性评价基线展开。每学期考核结束后,系统自动采集各课程试卷试题、以及每道试题平均得分。设定70%为达成度阈值,自动筛选出低于此阈值的试题。随后,通过大语言模型分析习题与课程知识点的映射关系,将薄弱试题关联至图谱中的知识节点。最终生成《课程教学目标达成度诊断报告》。报告重点剖析未达标知识点列表及其支撑产业能力要求,并对试题难度与区分度进行辅助分析。

### 3.3 基于全流程数据的重点课程实时监测

为深化过程性评价,选取《大数据采集与处理》作为深度实践对象,依托超星学习通、海浪智学等智慧教学平台,构建了一套教学全程实时质量监控与反馈体系。

该体系实现了对学习全链条数据的自动化采集,自动获取学生在教学视频观看、课前/课中测试、课内实验操作及阶段测试等各个环节的完成情况、耗时与表现数据。以此生成的实时学情诊断,教师可精准定位到班级共性薄弱点与个体学习困境,从而在课堂上开展高效的针对性重难点讲解,实现以学定教。

### 3.4 总结与展望

本研究基于专业知识图谱,完成了大数据专业课程体系的重构与教学评价体系的系统化建设,显著提升了课程与产业需求的对接效果。研究发现,教学质量的持续提升不仅依赖于评价方法本身,更与教学数据闭环反馈机制密切相关。未来研究将重点推进以制度为保障、以数据为驱动的课程质量持续改进体系,健全多课程协同评价规范,推动专业建设从经验驱动向制度保障、数据赋能的系统化方向发展。

### 参考文献

- [1] 李元霞. 数智时代基于专业能力图谱的应用型人才培养研究——以财务管理专业为例[J]. 质量与市场, 2025, (12): 188-190.
- [2] 单玄龙, 贺金鑫, 许中杰, 等. 知识图谱赋能吉林大学资源勘查工程本科专业课程建设[J]. 中国地质教育, 2025, 34 (05): 53-56.