

# 新工科导向的大数据专业工程实践能力融合培养模式研究

刘京昕<sup>1</sup>, 余海霞<sup>1</sup>, 廖铃<sup>2\*</sup>, 彭军<sup>1</sup>, 涂静雯<sup>1</sup>, 李茜<sup>1</sup>

1. 重庆科技大学数理科学学院, 重庆;

2. 重庆三峡职业学院人工智能学院, 重庆

DOI:10.62836/jer.v4n2.1018

**摘要:** 当前大数据专业教育在工程实践能力培养方面存在结构性失配、资源性匮乏与目标性偏离等关键问题, 本文以学生为中心, 基于新工科导向、持续改进的核心理念, 构建一个涵盖“目标体系-课程结构-教学模式-评价机制”四位一体的融合培养模式。通过重构各阶段阶梯式能力目标, 设计模块化和项目化的课程体系, 以AIGC赋能多元化教学方法, 建立基于大数据的全过程、多维度评价与反馈闭环。

**关键词:** 大数据专业; 工程实践能力; 新工科; 融合培养模式

---

## Research on an Integrated Training Model for Engineering Practical Ability in Big Data Majors Oriented by Emerging Engineering Education

Jingxin Liu<sup>1</sup>, Haixia Yu<sup>1</sup>, Ling Liao<sup>2\*</sup>, Jun Peng<sup>1</sup>, Jingwen Tu<sup>1</sup>, Xi Li<sup>1</sup>

1. School of Mathematics and Physics Sciences, Chongqing University of Science and Technology, Chongqing

2. School of Artificial Intelligence, Chongqing Three Gorges Vocational College, Chongqing

**Abstract:** Currently, the education in the big data major suffers from key problems such as structural mismatch, resource scarcity, and deviation from goals in the cultivation of engineering practical abilities. Centered on students, this paper is based on the core concepts of Emerging Engineering Education orientation and continuous improvement, and constructs a comprehensive training model integrating “target system-course structure-teaching mode-evaluation mechanism” in a four-in-one manner. By reconfiguring the step-by-step ability goals at each stage, this paper designs a modular and project-based curriculum system. Furthermore, by empowering diversified teaching methods with AIGC, this paper establishes a closed loop of all-process and multi-dimensional evaluation and feedback based on big data.

**Keywords:** big data majors; engineering practical abilities; Emerging Engineering Education; integrated training model

---

\* 基金项目: 重庆市高等教育教学改革研究项目(232138); 重庆市高等教育学会高等教育科学研究课题(CQGJ23118C); 重庆市教育科学规划青年课题(K24YY2150009); 重庆科技大学本科教育教学改革研究项目(2025063)。

作者简介: 刘京昕, 男, 讲师, 研究方向为人工智能、大数据、机器学习, Jingxinliu9@126.com。廖铃, 女, 讲师, 研究方向为计算机视觉、大数据, liaoling4123@163.com。

## 1 引言

在应对新一轮科技革命与产业变革中，新工科建设对我国高等教育人才培养体系提出了战略性重构的要求。大数据技术作为汇聚算力、算法与场景的核心载体，已深度渗透至智能制造、智慧金融、精准医疗和数字治理等国民经济关键领域，成为重塑产业竞争格局、提升国家治理效能的战略制高点[1]。这一宏大背景催生了市场对高素质、复合型大数据专业人才的迫切需求。然而，与产业界火热期盼形成一定反差的是，当前部分高校大数据专业毕业生普遍面临知识结构与企业需求脱节、工程实践能力薄弱以及创新素养不足的严峻挑战，人才供给侧与需求侧的结构性矛盾仍然存在。

新工科建设的提出与深入推进是国家对这方面挑战作出的战略性回应，要求高等工程教育从理念、范式、结构与质量四个维度进行系统性和深层次变革[2]。其核心任务在于培养能够适应并引领未来产业发展的、具备卓越工程实践能力和跨学科整合能力的创新型人才。对于数据科学与大数据此类应用性极强的交叉学科而言，新工科的内涵更为具象。大数据专业学生不仅应具备坚实的数理与计算机科学理论基础，更需具备面向真实产业场景，设计和运行维护复杂大数据系统以解决非结构化、开放性工程问题的综合素养[3]。探索并构建一个能够深度融合理论教学、工程实践与科研训练，并具备动态自适应技术演进能力的系统性人才培养新模式，已成为推动大数据专业新工科建设内涵式发展

的核心议题与紧迫任务。构建能够深度融合理论教学、工程实践与科研训练，并具备动态演进能力的系统性人才培养新模式，是推动大数据专业新工科建设内涵式发展的重要任务。本文基于这一现实需求，旨在提出一个具有普适意义的融合培养模式理论框架，并对其构成要素、运行机制及初步实践成效进行深入探讨。

## 2 现状及矛盾

部分高校现有的大数据专业课程体系多呈现为计算机科学、统计学、应用数学等学科知识的简单线性叠加，缺乏以解决复杂工程问题为导向的有机融合。学生习得大量离散的理论知识点，缺乏将其系统性整合并应用于完整项目生命周期的实践经验。现行评价体系仍过度依赖期末笔试与标准化实验报告，单一化、终结性的评价方式难以有效评价学生在项目设计、团队协作及技术创新等关键软实力与工程素养上的表现，不能对学生的全面发展形成正向牵引[4]。当前大数据专业工程实践能力培养方面的主要矛盾及表现如图1所示。

### 2.1 结构性矛盾

大数据技术生态是一个以高速迭代为显著特征的复杂巨系统。从以Hadoop、Spark为代表的批处理计算框架，到以Flink、Kafka为核心的流式计算技术，再到以TensorFlow、PyTorch为主流的深度学习模型库，其技术版本、编程范式与应用架构的演进周期日益缩短。反观当下部分地方院校，从人才

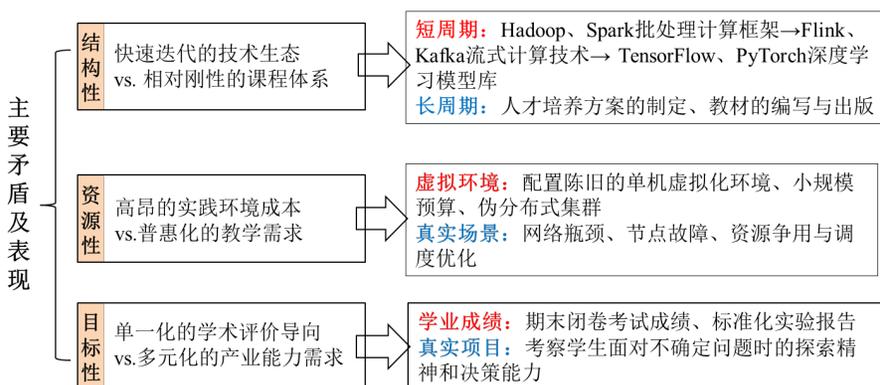


图 1. 主要矛盾及表现

培养方案的制定、课程标准的确立、教材的编写与出版到最终课堂实施往往需要一个长达数年的长周期。这种程序和节奏上的反差使得学生从课堂中系统习得的知识技能在他们毕业之际可能面临滞后甚至被淘汰的风险。快速迭代的技术生态与相对刚性的课程体系始终存在互斥性。当高校课程仍深陷于MapReduce编程模型内部机理的详尽剖析时，产业界的焦点可能早已转向基于Spark SQL的交互式查询与基于Flink的实时数仓构建。

另一关键问题在于，传统按学科分类的课程设置人为地割裂了大数据工程项目的完整工作流。高校学生分别在《分布式系统原理》课程中学习底层架构，在《数据库系统》课程中掌握数据存储，在《机器学习》课程中钻研建模算法，却鲜有综合性课程能够将这些离散的知识单元串联起来，指导学生完成一个从业务理解、数据采集、预处理、存储管理、分析建模到可视化呈现与应用部署的端到端项目实践。这种知识孤岛的效应，一定程度阻碍了学生解决复杂工程问题所必需的系统性思维培养。

## 2.2 资源性矛盾

当前部分地方高校还不具备真实的实践环境，大数据专业学生们的工程实践能力必须是在与之匹配的高仿真环境中不断锻炼形成的。一个能够稳定、高效处理TB/PB级工业数据，并完整部署主流大数据组件生态（Hadoop/Spark集群及其周边工具）的实验平台，对于服务器硬件采购、数据中心电力能耗、网络带宽及专业运维人力均构成高昂的成本压力。诸多高校受限于预算，仅能搭建小规模伪分布式集群或依赖于配置陈旧的单机虚拟化环境。在此类受限环境中，学生们无法亲身体验分布式系统在数据规模与并发负载提升时所暴露出的网络瓶颈、节点故障、资源争用与调度优化等真实挑战，他们获得的实践经验存在严重的片面性与失真度。优质实验资源的稀缺性与学生规模的庞大性之间亦存在尖锐矛盾，应该确保每一位学生都能获得充分且深度的实践机会，而非少数拔尖的学生，这是保障实践教学质量普惠性的关键[5]。云计算资源

以其弹性伸缩和按需付费的特性，为破解此难题提供了理论上的可行路径，但要求其长期、大规模应用于教学实践的成本控制与资源管理，这对于高校而言仍是一个待优化的现实课题。

## 2.3 目标性矛盾

社会产业界对理想大数据专业人才的能力期望是全方位的，不光是要求其具备扎实的技术功底，同时也高度看重其界定问题、分解问题、方案设计、项目管理、跨领域沟通协作以及对特定行业业务逻辑的深刻洞察力，这是一种综合性的数据思维。但是，当前部分高校内部现行的学生评价体系受路径依赖与管理便利性等因素影响，仍呈现出强烈的学术化与标准化倾向。期末闭卷考试成绩和基于预设数据集的标准化实验报告，依然是衡量学生学业成就的主导性指标。这种评价范式只是操作上简便易行，其效度在评估学生在开放式和复杂性项目中的真实表现时依旧存在很大的局限性[6]。闭卷考试成绩和标准化实验报告无法有效甄别大数据专业学生在面对模糊和不确定性问题时的探索精神和权衡决策能力，以及向非技术背景利益相关者清晰阐释分析结论的沟通能力。通过考核的“指挥棒”效应直接塑造了学生的学习策略与精力分配，这种目标导向的偏差，使得学生可能为追逐高分而沉溺于概念知识、框架命令等的机械背诵，容易忽视在更具挑战性的真实项目环境中锤炼自身综合能力[7]。因此，构建一套与产业能力需求矩阵紧密对接、过程性评价与终结性评价相结合的评价体系，是驱动整个培养模式成功转型的关键杠杆。

## 3 “四位一体”融合培养模式

### 3.1 阶梯式能力目标

在目标体系中，本文摒弃了传统的以“知识点覆盖”为中心的目标陈述，转而构建以“能力产出”为导向的、清晰可测的阶梯式目标体系。该体系深度融合《工程教育认证标准》毕业要求素质特征与大数据专业特色，确立了从基础认知到创新研究的四个递进层次（见表1），为整个培养过程提

供了精准的导航与评估基准。

### 3.2 模块化与项目化的重构设计

围绕上述能力目标，本文对课程体系进行了系统性重构，其核心特征是横向模块化与纵向项目化的结合。横向模块化目的是为了打破原有学科壁垒，将关联知识重组为四大逻辑清晰的课程模块，构建厚基础、强专业、重交叉的知识体系。纵向项目化是设计一条贯穿整个本科学习阶段的项目链。从低年级的验证性、设计性项目，到高年级的综合性、创新性项目，乃至毕业设计，项目难度、复杂度与开放性逐级递增，确保学生在四年中持续浸润于工程实践环境，实现能力的螺旋式上升（见表2）。

### 3.3 多元化教学模式

从教学实施层面出发，本文大力推动从教师主导到学生主体的范式转移，采用多元化的教学方法激发学生内在动力与创造潜能。通过项目来驱动学生学习，将所有大数据专业核心课程围绕一个或多个核心项目展开，使教师从知识的传授者转变为学习教练、项目顾问等角色，引导学生团队经历从项目立项、需求分析、方案设计、开发实现、测试部署到总结答辩的全过程。同时，本文正视生成式人工智能的技术浪潮，致力于将其从潜在的“学术诚信挑战”转化为提升教学效能的认知工具。鼓励学生利用 ChatGPT、DeepSeek、元宝、Mixtral、文心一言

表1. 大数据专业阶梯式能力目标

能力层级	核心能力描述	教学活动	毕业要求
L1: 基础认知层	掌握数学、计算机、统计学等学科基础理论与方法；了解大数据技术生态的整体轮廓与发展趋势；具备基本的编程实现与数据操作能力。	基础理论课程 编程实验 认知实习	1.工程知识 2.问题分析 5.使用现代工具
L2: 技术应用层	能熟练运用一种以上主流大数据处理工具（Hadoop、Spark、Flink）完成特定数据处理任务；能够对中小规模数据集进行完整的清洗、转换、分析与可视化；具备在教师或导师指导下参与项目模块开发的能力。	专业核心课程 课程设计 基础性实验项目	2.问题分析 3.设计/开发解决方案 5.使用现代工具 9.个人与团队
L3: 系统集成层	能够针对特定应用领域的需求，设计端到端的大数据解决方案架构；具备系统性能调优、故障诊断与排除能力；能在跨学科项目团队中有效协作，管理项目进程，解决具有一定复杂度的工程问题。	综合性项目实训 跨学科团队项目 企业实习 高级专题实验	3.设计/开发解决方案 4.研究 7.环境与可持续发展 9.个人和团队 10.沟通 11.项目管理
L4: 创新研究层	能够持续追踪国际技术前沿，批判性评估现有技术局限，并提出创新性的算法、模型或系统优化方案；具备独立开展科学研究、撰写高质量学术或技术论文及进行技术成果转化的潜力。	毕业设计 科研导师项目 学科竞赛 学术论文撰写	4.研究 6.工程与社会 8.职业规范 12.终身学习

表2. 课程体系结构

课程模块	代表性课程	实践项目链（示例）	支撑的能力层级
通识与数理基础	高等数学、线性代数、概率论与数理统计、程序设计基础、数据结构	大一：城市公共自行车使用情况数据分析报告	L1
数据工程与管理系统	数据采集与预处理技术、分布式系统原理、NoSQL数据库、数据仓库与ETL技术、数据治理与质量	大二：社交媒体热点话题实时追踪与存储系统构建	L1, L2
智能分析与建模算法	机器学习、数据挖掘、数据可视化技术、自然语言处理、时间序列分析	大三：基于多源数据的金融风险模型或电商智能推荐系统设计与实现	L2, L3
系统架构与交叉应用	大数据平台架构与运维、云计算与服务计算、IT项目管理与伦理、大数据工业应用	大四：面向智慧城市的交通流量预测与调度平台（毕业设计）	L3, L4

等工具进行概念澄清和技术方案的初步调研，并借助这些工具对学习资料进行归纳整理。

AIGC能为学生的自主学习和编程训练提供有用的信息。允许专业学生在编程实践中使用AIGC辅助生成基础代码框架、单元测试用例、API文档草稿及进行代码重构建议。要求学生必须深刻理解、逐行验证并记录优化所有AI生成的代码，并在项目报告中明确说明AI的贡献与个人的创造性工作。同时，鼓励学生向AIGC提出开放式、探索性问题，获取多样化的解决方案思路，进而通过批判性分析与实验筛选出真正价值的信息，整合完善最终的决策方案。制定明确的内部《AIGC教学使用公约》，并纳入评价标准，将AIGC定位为提升学习深度、广度与效率的战略合作伙伴，而非替代思考的单纯工具。

建立“导师-研究生-本科生”的传帮携带机制，鼓励教师将国家级、省部级纵向科研项目或横向企业课题分解为适合研究生、高年级本科生的微课题或子任务。通过设立科研学分、开放科研实验室和举办“学术沙龙”等方式，吸引优秀本科生早期进入科研团队，接触学术前沿，培养其科学探究精神。在案例选取和项目设计方面，需将“基于大数据的公共卫生事件态势分析”、“乡村振兴背景下的农产品电商数据分析”、“智能制造中的工业数据赋能”等体现国家战略的真实案例融入，引导大数据专业学生理解工程技术的社会价值，培育他们的家国情怀和工匠精神。

### 3.4 多维度反馈闭环的评价机制

大数据专业学生的行为数据包括代码提交频率与质量、文献查阅记录、在线讨论参与度、项目进度日志等。构建可持续的项目平台对学生们的过程性成果进行存储，这些成果包括设计文档、阶段性报告、原型系统，并与最终的项目报告作比较，进行关联度分析。为了体现评估标准的全面性，将代码复杂度、系统性能指标、算法准确性与过程表现融合考虑，其中的过程表现涉及项目执行、问题解决日志、贡献度自评与互评等内容。此外，还应考虑专业学生们的协作沟通能力和创新思维维度，创

新思维维度可从方案设计的新颖性和技术路线的深度去构建合理性指标。

全流程评价的反馈闭环机制对大数据专业学生工程实践能力的培养规划是具有科学意义的。整个系统还应帮助学生们生成个性化的能力发展雷达图与学习诊断报告，为他们提供精准的改进建议与发展指南。聚合的、脱敏后的群体评估数据，可为专业学科负责人和授课教师在优化课程目标、更新教学内容和改进教学方法方面提供科学的决策依据，从而形成一个有效的“评估-反馈-干预-改进”教学质量保障闭环。

## 4 综合案例

### 4.1 智能物联网设备预测性维护平台项目

为了将上述模式具象化，本文设计了一个关于智能物联网设备预测性维护平台的综合项目。该项目的实施计划安排贯穿三个学期（从大三上学期延续至大四上学期），覆盖多门核心课程。各个阶段和对应的任务模块表3所示。

### 4.2 讨论分析

通过这个从数据感知到智能决策的完整工业应用项目，是大数据专业学生不再局限于学习孤立的技术碎片，而是亲身经历一个工业互联网应用中典型的从0到1再到100的全过程，实现基本理论、工程技能与创新思维的深度融合。为初步验证该融合培养模式的有效性，本文对应用该模式的相邻两届学生（实施前对照组n=65，实施后实验组n=68，n表示学生数）进行了为期两年的跟踪调查。参考数据来源于课程成绩、项目答辩评分、标准化能力量表、学生自评报告和企业导师外部评价等核心能力认证得分，其中标准化能力量表采用5分制的Likert scale。实验结果见表4。

根据实验结果可以看出，在平台工具熟练度和系统构建能力两个硬核指标上，实验组表现出极其显著的进步，其中 $p < 0.001$ 。效果量分别高达1.12和1.24，远超0.8 Cohen's d，这有力地证明项目驱动的课程与多层次平台对夯实学生工程基础的决定性作用。在问题分析和创新方案设计能力等衡量高

表3. 智能物联网设备预测性维护平台项目实施计划

时间阶段/模块	主要工作	任务部署	AIGC赋能点	能力聚焦
大三上学期/数据工程与管理系统模块	数据采集与存储层构建	1. 学生在边缘端使用振动传感器、温度传感器与嵌入式开发板采集模拟工业设备的运行状态数据； 2. 利用Kafka构建实时数据流管道，将数据稳定传输至云端大数据平台； 3. 设计并实现基于HBase/HDFS的分布式存储方案，确保数据的高效写入与查询。	使用AIGC工具辅助理解Kafka生产者/消费者API、生成HBase表结构设计的示例代码，并学习数据序列化Avro、Parquet的最佳实践。 借助AIGC工具进行	重点训练大规模、高吞吐数据采集、传输与分布式存储的工程实现能力（L2→L3）。
大三下学期/智能分析与建模算法模块	数据分析与模型开发	1. 学生使用Spark MLlib或Flink ML对历史设备状态数据进行特征工程，提取与设备退化相关的关键指标； 2. 训练基于时序分析的故障预测模型（LSTM）或剩余使用寿命回归模型； 3. 对模型进行评估与优化，并完成初步的可视化展示。	Matplotlib、Plotly等数据可视化脚本的快速生成，获取不同特征选择方法与模型调参策略的对比分析，以辅助决策。	重点训练数据清洗、特征工程、机器学习模型构建、评估与优化的全流程数据分析能力（L3）。
大四上学期/系统架构与交叉应用模块	系统集成与应用部署	1. 将训练好的模型通过PMML/TensorFlow Serving等方式部署为微服务RESTful API； 2. 开发一个集设备监控、实时数据展示、预警通知与报告生成于一体的Web管理看板； 3. 对整个平台的性能、可靠性进行测试，并撰写详尽的技术文档、部署手册与项目总结报告。	使用AIGC工具辅助生成API接口代码（Spring Boot框架）、前端组件代码、技术文档初稿以及进行文档的润色与校对。	全面锤炼系统架构设计、服务部署、运维监控、项目管理、文档撰写与跨领域结合的综合能力（L3→L4）。

表4. 学生核心能力前后比较分析

评估维度	观测指标	对照组均值 (n=65)	实验组均值 (n=65)	T检验 (p值)	效果量 (Cohen's d)
技术实践能力	大数据平台与工具链熟练度	3.15	4.23	p < 0.001	1.12
	端到端系统设计与构建能力	2.81	3.95	p < 0.001	1.24
创新与决策能力	面对模糊需求的分析与界定	3.08	3.87	p < 0.01	0.85
	创新性技术解决方案的设计	2.92	3.71	p < 0.01	0.79
	团队协作与项目管理有效性	3.35	4.12	p < 0.001	0.88
职业综合素养	技术文档与报告撰写质量	3.02	3.85	p < 0.01	0.91
	AIGC工具的批判性与建设性使用	2.45	4.05	p < 0.001	1.58

阶思维的指标上，实验组也呈现出显著提升，其中  $p < 0.001$ 。Cohen's d 达到0.79-0.85的中等效应到大效应，这表明科研项目设计能够有效激活学生的探究潜能与创新意识。在团队协作、文档撰写等软技能方面，实验组亦有明显进步，能够实现职业综合素养的全面发展。值得一提的是，在对AIGC工具的批判性使用方面，Cohen's d 高至1.58，这表明实验策略成功地帮助大数据专业学生建立与AI协同工作的先进范式，使学生们对待AIGC工具的态度实现由被动禁用到主动善用的转变。另外，来自合作企业的反馈普遍认为，参与该模式培养的学生在实习与入职后，表现出更快的环境适应能力以及更扎实的系统工程观念，能够更高效地融入实际研发流程。

## 5 结语

本文系统性地提出一个面向大数据专业学生的“四位一体”新工科工程实践能力融合培养模式。该模式通过顶层设计与系统集成，将培养目标、课程体系、教学方法与评价反馈机制有机融合，旨在从根源上解决当前大数据教育中的结构性、资源性与目标性矛盾。初步的实证研究表明，该模式能有效、显著地提升学生的工程技术实践能力和职业综合素养。本文提出的培养模式的初步成功源于系统思维与协同设计，四个层面并非机械堆叠，而是构成了一个包含目标牵引、课程与平台支撑、模式落地、评价保障的动态自适应系统。AIGC的引入非但未导致学生思维惰

化，反而在精心设计的教学框架内成为拓展认知边界、和提升实践效率的倍增器。但是，模式的深入推广与持续优化仍然面临若干挑战。师资队伍需要转型，教师群体不仅需要具备深厚的学术素养，还需拥有丰富的工程实践经验和前沿的技术洞察力。平台环境的可持续运营需要稳定的经费投入和专业的技术支持团队。如何在班级授课制教学中实现个性化培养与精细化管理，仍需探索更有效的管理策略。

## 参考文献

- [1]宋莹,王兴芬,张伟.新工科背景下多学科交叉融合的大数据人才培养新模式探索[J].创新教育研究,2019,7(5):608-612.
- [2]王渝.大数据背景下教育教学高质量发展的探索与实践[J].教育科技前沿,2025,1(4).
- [3]陈继强,高林庆,马丽涛,等.新工科背景下地方本科院校数据科学与大数据技术专业课程体系建设探索与实践[J].教育进展,2022,12(3):662-666.
- [4]张莹,殷亚林,殷萍,等.基于大数据的计算思维能力培养的可视化分析[J].社会科学前沿,2024,13(10):93-102.
- [5]刘婧瑶,车丽萍.大数据技术在高校教育管理领域中的应用研究——基于CiteSpace的可视化分析[J].运筹与模糊学,2025,15(3):360-371.
- [6]胡蔓,曹利华,刘健,等.新工科背景下高校工程训练中心建设探索与实践[J].实验室研究与探索,2021,5:154-157.
- [7]朱玉平,张学军,高翔,等.工程实践创新能力融合培养研究[J].实验科学与技术,2022,20(3):44-48.

