新文科背景下 AI 赋能的数据科学课程教学探索*

陈默** 张昱 徐彬 李封

东北大学计算机科学与工程学院，沈阳 110819

摘 要 针对数据科学课程中场景偏离、代码纠错效率低等问题，构建了“场景化数据集构建—AI 纠错反馈—分层递进实践”三位一体的教学体系。依托大语言模型构建场景数据集及 AI 纠错反馈机制，并设计基础入门、核心能力、场景应用三阶递进的课程内容与分层实践模块。通过教学实践改革，规范完成完整数据分析流程的小组占比提升至 98.22%，贴合专业深度分析成果的小组占比增幅达 39.2%，课程改革显著提升了学生的学习积极性与数据分析项目的综合表现，为新文科数据科学通识课教学提供了可复制的改革路径。

关键字 新文科，AI 赋能，数据科学课程，教学改革

Exploration of AI-Empowered Data Science Course Teaching in the Context of New Liberal Arts

Mo Chen Yu Zhang Bin Xu Feng Li

School of Computer Science and Engineering
Northeastern University
Shenyang 110819, China
chenmo@mail.neu.edu.cn

Abstract—To address issues such as scenario misalignment and low code debugging efficiency in data science courses, we construct a three-in-one teaching system featuring scenario-based dataset construction, AI error correction feedback, and hierarchical progressive practice. By leveraging large language models, we build scenario-specific datasets and establish an AI error correction feedback mechanism. Meanwhile, a three-tiered progressive curriculum framework is established, covering basic introduction, core competency cultivation, and scenario application, along with layered practice modules. Multiple rounds of teaching practice show that the proportion of groups completing the full standardized data analysis process rose to 98.22%, and the proportion of groups delivering in-depth analysis results tailored to majors increased by 39.2%; this reform has significantly enhanced students' learning enthusiasm and overall performance in data analysis projects, and has fostered their data thinking abilities. It also provides a replicable reform path for teaching general data science courses in the context of the New Liberal Arts.

Keywords—New Liberal Arts, AI Empowerment, Data Science Course, Teaching Reform

1 引 言

2020 年教育部发布《新文科建设宣言》，明确提出新文科建设是新时代高等教育内涵式发展的重要举措，以学科交叉融合、专业优化、课程提质、模式创新为抓手，推动人文社科与现代信息技术深度融合，培养兼具人文素养与数字能力的复合型人才[1]。新文科的核心在于打破传统学科壁垒，促进文理交叉、技术赋能，使文科教育更好地适应数字时代的社会需求[2-3]。在此背景下，数据已成为新型生产要素[4]，对于工商管理类、公共管理类等相关专业学生而言，数据处理、分析与可视化等基础数据素养成为其职业发展中的核心竞争力[5-6]。

数据科学课程作为培养数据素养的核心载体[7]，是新文科建设中文理融合课程的典型代表。该课程的目标是赋能文科学生运用数据工具解决本专业领域内的实际问题，具有鲜明的通识性、应用性与交叉性[8]，当前面向文科学生的数据科学课程教学，普遍面临公开数据集偏工程技术领域、代码纠错效率低、案例与文科场景脱节、实践环节碎片化等突出问题，需探索适应新文科建设要求的教学改革路径。

面向文科专业的数据科学课程，定位需与理工科同类课程明确区分。课程的核心目标是使学生掌握数值计算、数据处理与分析、数据可视化三大模块的基

*基金资助：本文得到辽宁省普通高等学校本科教学改革研究项目“人工智能通识课程认知分层教学改革与跨学科协同育人机制研究”（2025YBXM0175）的项目资助。

**通讯作者：陈默 chenmo@mail.neu.edu.cn

础操作，具备在专业领域场景下进行数据清洗、分析及可视化的基本能力，并最终建立数据思维。

本文依托本校文管类专业数据科学课程多轮教学实践，围绕文科教学现存核心难题构建 AI 赋能分层教学改革方案，完整阐述课程分层设计、AI 辅助教学以及过程化成效评价体系，为高校新文科数据科学通识课的教学提供实践参考。

1 课程核心问题剖析

当前数据科学课程面临如下四个突出问题，使得面向文科学生的课程教学难以达到预期效果。

（1） 公开数据集适配性较差。网络公开数据集多源于理工科竞赛或项目，与文科学生的专业背景和认知经验相去甚远。这类数据要么过于规整，缺乏真实数据中的缺失值和异常值；要么过于复杂，难以支撑分层教学任务，导致学生学用脱节。

（2） 代码纠错效率低。文科学生面对数组维度不匹配、索引错误、图表显示异常等问题时，往往不知如何下手。教师在教学上花费时间精力逐一解答代码错误，难以兼顾整体进度，学生在反复出错中极易产生挫败感。

（3） 教学内容与专业场景脱节。课程案例与学生的专业知识缺乏关联，数据思维的培养便难以落到实处。例如，企业销售数据案例对工商管理类专业的学生较为适用，但对人文专业的学生则缺乏直接联系。缺少与专业背景相匹配的场景化案例，是导致学生学习动力不足的重要原因。

（4） 实践环节碎片化。绝大多数实验仍以验证单个知识点为目的，如“练习数据筛选”“练习绘制折线图”。这些实验相互孤立，学生无法体验从原始数据到分析报告产出的完整工作流程，综合运用知识解决实际问题的能力难以得到真正提高。

鉴于课程上述特点，课程在内容设计和教学实施上应遵循“三弱三强”原则：弱化算法原理推导，强化工具应用操作；弱化底层技术细节，强化数据处理逻辑思维；弱化理工科应用场景，强化文科专业融合案例设计。

2 AI 赋能课程教学的必要性与机制

文科学生大多习惯于具象思维和文本理解，对抽象的代码逻辑和数学公式缺乏直观理解，AI 技术可以在这方面发挥作用。例如，大语言模型能将晦涩的错误信息转换成通俗解释，并提供具体的修改建议。这种即时、有针对性的反馈，更贴合文科学生“重理解、轻原理”的学习偏好，有助于缓解畏难情绪，让因材施教真正落地。

在数据集构建方面，AI 可以生成文科场景定制化模拟数据集。教师根据教学需要，用自然语言向 AI 描述数据的基本要求，AI 即可输出符合规范的 CSV 格式数据。这改变了以往“有什么数据就教什么”的被动局面，转向“需要什么数据就生成什么数据”的主动教学模式。同时，学生也可以利用 AI 生成自己感兴趣或与自己专业相关的练习数据，使数据操作任务更贴合个人学习需求。

在代码纠错方面，大语言模型辅助教学已被验证可提升学生的代码调试效率[9-12]。学生遇到代码错误时，无需等待教师，可直接将错误信息和代码片段提交给 AI。AI 能迅速定位错误类型、分析原因，并提供修正后的代码及解释。学生据此修改代码并记录问题，逐步积累调试经验，这种方式可以有效提升代码学习效率和课程学习的自主性。

新文科教育的终极目标是培养能运用多学科知识解决复杂问题的复合型人才。AI 赋能使得教师可以将教学重心从繁琐的代码纠错中解放出来，更关注于引导学生思考数据背后的业务逻辑、选择合适的分析方法、解读可视化图表，并将其应用于文科领域的决策支持。这推动了教学模式从“学代码”向“用数据”的根本转变，更加契合新文科强调的“数据解读、数据表达和数据决策”的能力要求。

3 课程教学设计

3．1 三阶递进课程内容重构

课程按“基础入门、核心能力、场景应用”三个层次递进，如图 1 所示。基础入门层侧重基本操作和简单任务，帮学生跨过数据分析技术门槛；核心能力层围绕数据清洗、转换、分析等技能，通过典型任务练熟操作；场景应用层强调综合运用，学生最终要能独立完成一份完整的数据分析报告。

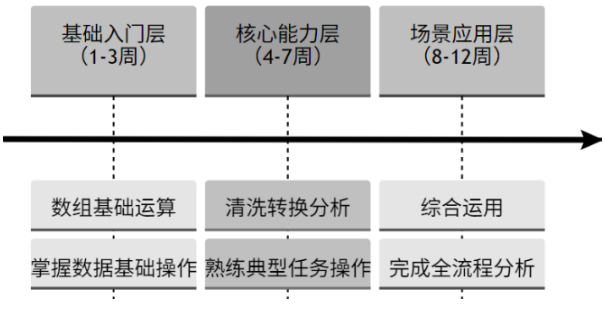


图 1 三阶递进课程内容体系

3．2 辅助场景数据集构建

公开数据集在数据科学课程教学中存在两个典型问题：一是场景偏向理工科领域，与文科学生的专业背景关联薄弱；二是数据质量过于规整，缺乏真实数据中常见的缺失值、异常值等噪声，导致学生完成练

习后难以迁移至本专业的实际问题处理。针对这些问题，在本课程教学中使用大语言模型辅助构建模拟数据集，例如，在豆包 AI 平台中输入提示词：“生成一份电商用户消费数据集，样本量 200 条，字段包括用户 ID、年龄段、月消费金额、购买频次、会员等级，并在月消费金额字段中随机生成 5%的缺失值。”模型即可输出符合要求的数据集，如图 2 所示。后续可以根据具体的教学内容，继续追加指令以调整数据集的规模、字段结构或噪声比例。

根据课程的三层递进体系，在不同阶段设计不同难度的相关数据集：

（1） 基础入门：生成小规模、结构规整的数据，如学生成绩记录、图书借阅明细。字段数量少，无异常值与缺失值，便于学生掌握基本数据读写与简单统计分析。

（2） 核心能力：生成含缺失值、重复值及异常值的数据，如电商用户数据、课程问卷调查结果。学生需运用数据清洗、分组聚合等操作进行处理。

（3） 场景应用：生成多维度、混合类型的复杂数据，如社交媒体交互数据、校园公共设施满意度调研数据。学生需综合运用数据处理操作，完成完整的数据分析流程。

这些数据集覆盖典型的文科应用场景，有助于学生将数据操作任务与自身专业背景建立联系，从而提升课程学习兴趣。

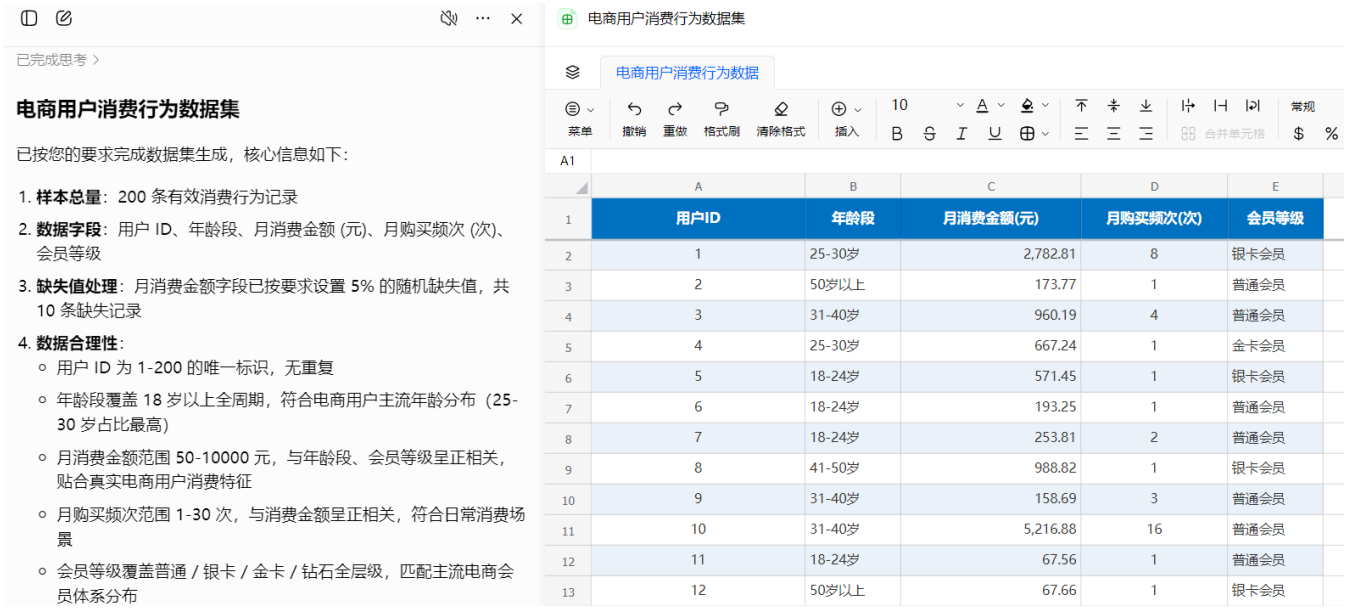


图 2 AI 辅助数据集构建

3．3 基于 AI 的纠错反馈机制

针对文科专业学生在数据分析编程实践中代码调试困难的问题，课程在教学实践环节中采用 AI 辅助纠错机制，学生独立编写代码并运行，遇到错误时需用自然语言向 AI 描述问题，如预期操作、实际出现的错误现象，并附上代码片段。AI 返回分析结果，包括错误定位、修正后的代码示例以及避免同类错误的操作建议。学生根据 AI 反馈修正代码，同时将错误类型、原因及解决方法整理为个人纠错日志，每节课后通过雨课堂平台提交。教师定期从雨课堂平台汇总学生日志中的高频错误，在课堂上进行集中讲解。

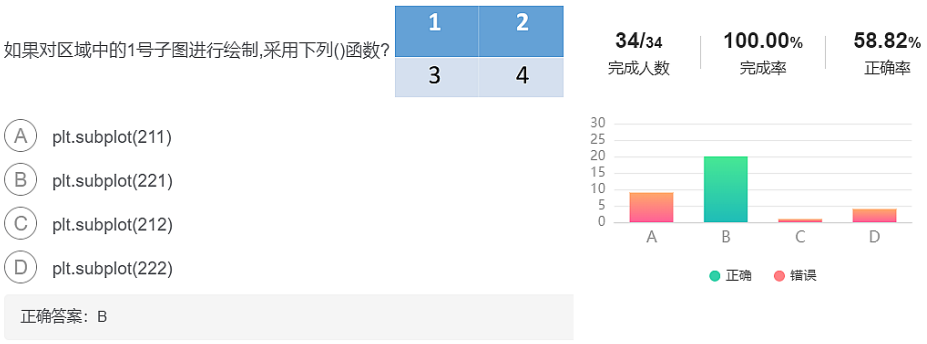


图 3 高频错误试题分析

此外，根据学生提交的纠错日志，教师利用雨课堂平台在每周实验课上推送针对性题目，题目主要围绕高频错误类型设计。例如，在 Matplotlib 绘图模块的教学中发现，关于子图绘制的代码错误频次较高，因此在当周测试中加入了相应题目。通过雨课堂平台的试题分析功能查看学生作答情况，发现该题正确率仅为 58.82%，如图 3 所示，反映出学生对子图划分逻辑的理解不到位。由此，教师将子图索引规则作为下次课的讲解重点，并在后续练习中增加同类题目的变式训练，以帮助学生巩固对该知识点的掌握。

教学流程从“AI 辅助纠错”开始，如图 4 所示，经过“反馈—修正—记录—讲解—测试—再讲解”多个环节，最终返回到学生知识掌握的提升上，形成了持续优化的教学回路。这个过程中，教师能够从重复解答基础性错误中解放出来，将更多精力投入到数据思维引导与案例教学设计中；学生通过“测试—反馈—再讲解”的强化环节，对易错知识点的掌握更加扎实；同时，基于雨课堂平台的数据分析，教师可以动态调整教学内容与进度，实现因材施教。这一完整流程有效提升了教学效率与学习效果。

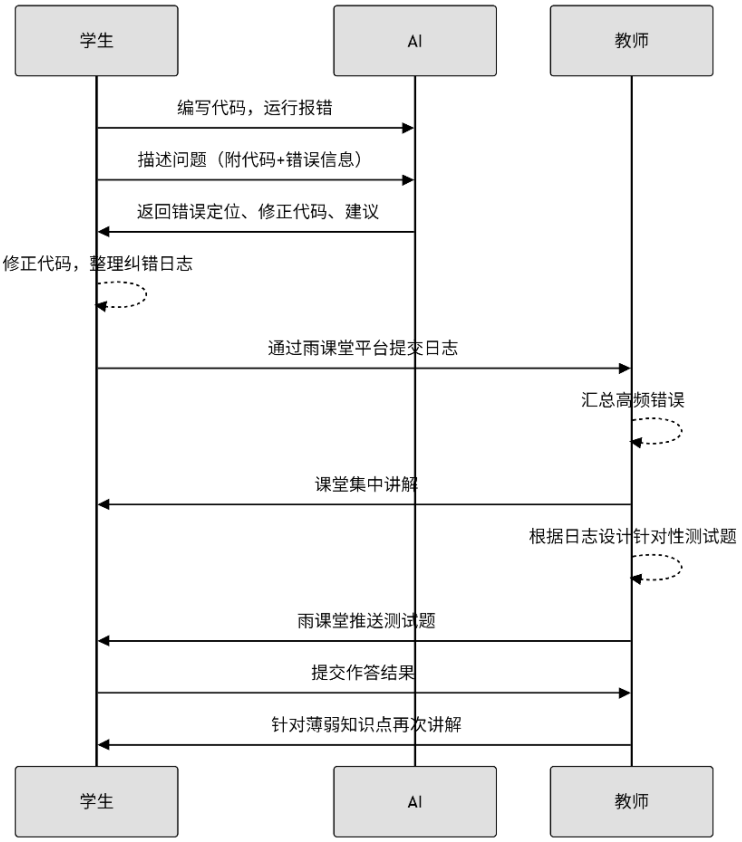


图 4　基于 AI 辅助纠错的教学路径

3．4　场景化分层实践

为强化学生的应用能力并兼顾不同文科专业的学科背景差异，课程设计了引导型实践、半开放型实践、自主选题综合项目三个递进层次，难度逐层提升，场景覆盖多个文科专业方向，具体设计见表 1。

在引导型实践中，学生在教师提供的小规模、结构规整的数据集中任选其一，例如学生成绩记录、图书借阅明细等。任务仅涉及数据读取、数据筛选等基本操作，并画出简单的可视化图表。该层次主要让学生熟悉代码环境与基本函数，AI 主要帮助纠正语法错误，降低入门挫败感。

进入半开放型实践阶段后，学生可以根据自身专业背景，从可选主题池中选择项目完成，主题池每学期根据学生专业分布动态调整。例如，工商管理类可选“用户评论词云生成”，清洗评论数据并生成词云；公共管理类可选“公共服务满意度分析”，处理含缺失值的问卷数据；人文科学类可选“学生思想动态调查分析”，对比不同年级或专业在各项调查指标上的差异。该阶段的数据集可以由 AI 辅助生成，学生需运用数据清洗和分组聚合等操作，AI 辅助定位逻辑错误并提供可视化改进建议。

在前两个实践层次的基础上，学生以小组为单位，自选与其专业密切相关的真实数据集，或使用 AI 生成

的模拟数据集。教师提供 AI 生成数据集的提示词模板，学生可根据模板生成模拟数据，并将选题与数据集先提交教师审核。通过审核后，学生完成从问题定义、数据生成、清洗、分析、可视化到报告撰写、答辩的完整流程，最终报告由 AI 初评代码规范与结构，教师复评并组织答辩。成果要求提交源代码、数据集说明、分析报告及答辩 PPT。

上述三层次的实践项目设计，引导学生从基础操作过渡到初步独立完成专业场景任务，最终自主选题并完成完整的数据分析项目，实现了数据分析技能与专业场景的有效融合。AI 辅助贯穿全程，在语法纠错、逻辑定位、报告初评等环节发挥差异化支撑作用，有效降低了学习门槛，提升了实践效率。

表 1　三层次实践项目设计

实践类型	项目主题（部分示例）	涉及模块	AI辅助重点
引导型实践	学生成绩分析、图书借阅统计、课程问卷数据处理等	数据读写、筛选、统计、简单可视化	语法错误反馈、基础函数使用提示
半开放型实践	用户评论词云生成、公共服务满意度分析、学生思想动态调查分析等	数据清洗、分组聚合、简单相关性、进阶可视化	逻辑错误定位、可视化建议
综合实践	学生自选专业相关数据集（教师审核）	完整数据分析流程：清洗→分析→可视化→报告→答辩	AI初评报告结构、代码规范、图表合理性

4　教学评价与成效

课程教学实践中构建了以过程为导向、能力为核心的多元评价体系。总成绩由三部分构成：过程性评价（50%）包括个人代码纠错日志、课堂练习与测试、课堂参与度；成果性评价（30%）包括实践项目的分析报告质量、可视化作品的呈现效果与专业性、代码的规范性与可读性，这部分采用 AI 初评、教师复评的流程；能力性评价（20%）包括在项目答辩中展现的数据解读能力、逻辑思维能力和对文科场景应用价值的阐述能力。上述教学方案在东北大学《数据科学基础（文管类）》课程中实施三学期。课程结束后，通过问卷调查收集学生反馈，并结合平台数据评估教学改革实施效果，统计结果如图 5 所示。

从基础学习体验维度来看，相较于传统教学模式，AI 赋能改革模式下学生课程学习兴趣认同占比与课堂参与度均有接近 20 个百分点的提升。这一变化的核心动因在于，课程设置的 AI 辅助机制能将晦涩的代码异常信息转化为可理解的自然语言解释，并提供具体的修正示例，使学生在遇到障碍后能够迅速获得正向反馈，完成“遇到问题—理解问题—解决问题”的闭环，从而在心理层面实现了从畏难逃避到主动求助的态度转变，显著提升了学生的课堂参与意愿。

从项目完成质量维度来看，传统模式下仅约 77%的小组可规范完成完整数据分析流程，改革后该指标升至 98.22%，这一提升归因于三阶递进课程内容体系对“过程性能力积累”的支撑作用：基础入门层使学生建立代码环境的操作信心，核心能力层通过典型任务重复训练形成操作图式，场景应用层则在前期积淀的基础上实现综合迁移。三个层次之间形成“技能—熟练—综合”的能力发展链条，避免了传统教学中知识点割裂，学生到期末仍无法串联整个流程的问题。

从深度分析能力维度来看，可产出贴合文科专业深度分析成果的小组占比增幅达 39.2%，这是本次改革最为显著的成效。这一提升主要得益于三层次实践项目中的引导性设计，促使学生在利用 AI 辅助时逐渐形成了更深层次的交互习惯。通过对高分组与低分组的对比分析发现，低分组多停留于操作型使用，即以修正代码错误为唯一目的。高分组则进阶至认知型使用，即借助AI深化对分析逻辑与结果呈现的理解。这也说明，AI 赋能数据科学课程教学的成效不止于提升代码通过率，更在于推动学生完成从“用 AI 改代码”到“用 AI 认知数据”的能力跃迁。

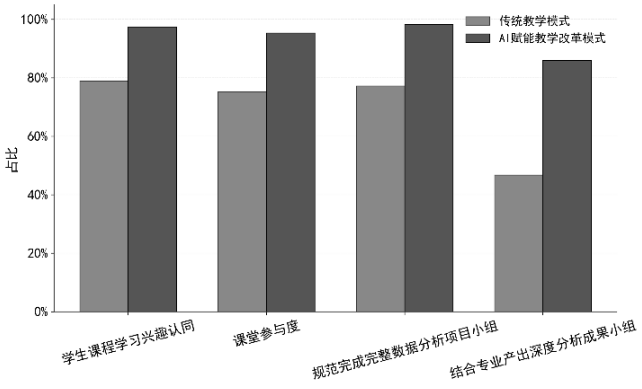


图 5　改革教学成效对比

5　结束语

针对新文科背景下数据科学课程教学中的关键问题，课程改革以“AI 辅助数据集构建—基于 AI 纠错反馈机制—分层递进实践”三位一体为核心，构建了覆盖基础入门、核心能力、场景应用的课程内容体系，引入基于 AI 大模型的数据集定制与代码调试辅助机制，并实施了过程与成果相结合的综合评价方式。

教学实践证明，学生的课程学习积极性明显提高，在数据分析项目中的综合表现进步显著。

随着大模型技术的快速迭代和教学场景的不断拓展，课程建设将进一步丰富文科典型场景数据集，覆盖更多文科专业方向；同时探索基于大模型的个性化学习路径推荐与智能辅导，提升分层教学的精细化程度，加强拓展跨校协作与资源共享。

参 考 文 献

[1] 教育部新文科建设工作组.新文科建设宣言 [EB/OL].(2020-11-03)[2026-05-10].http://www.moe.gov.cn/jyb_xwfb/gzdt_gzdt/s5987/202011/t20201103_498067.html.

[2] 樊丽明.“新文科”：时代需求与建设重点[J].中国大学教学,2020,(05):4-8.

[3] 樊丽明.论新文科建设的机制保障[J].中国高教研究,2023,(05):8-13.

[4] 赵琳，骆宗伟.场景式教学 + 认知模型的《大数据分析》课程模式初探[J].计算机技术与教育学报,2022,10(04):28-33.

[5] 白玥，朱敏，陈志云.提升数据素养的新文科计算机教学改革与实践[J].计算机教育,2021,(09):91-94.

[6] 沈映春，赵雨涵，李跟强,等.新文科背景下经济专业创新型科产教融合育人模式构建[J].高教学刊,2026,13(13):63-66.

[7] 张明慧，郭欣.大数据科学与技术专业课程群育人探索与实践[J].计算机技术与教育学报,2022,10(02):66-69.

[8] 孙宝军，赵俊岚.面向新文科建设的数字经济专业融合数据科学课程链研究[J].内蒙古财经大学学报,2023,21(01):27-30.

[9] 徐慧，鞠小林，王皓晨.大模型下编程教学面临的挑战与应对[J].计算机教育,2023,(11):60-64.

[10] 董敏，毛爱华，毕盛,等. AI 赋能 + 通专融合 + 产教融合的 C++ 编程基础课程教学探索[J]. 计算机教育,2025,(02):60-65.

[11] 张昱，李封，徐彬. 产学研协同的人工智能编程课程教学改革[J].计算机技术与教育学报,2025,13(05):159-163.

[12] 孙明瑞，王薇，唐铎.产教协同模式下 AI 赋能的程序设计实验教学改革与实践—以数据可视化课程为例[J]. 计算机教育,2026,(02):165-169.