

# 融合大数据与人工智能的工程教育 精准教学与改进机制研究\*

于志强 蒋作\*\* 谭学文

云南民族大学数学与计算机科学学院, 昆明 650504

**摘要** 工程教育认证作为国际认可的质量保障机制, 强调以学生为中心、成果为导向, 旨在培养毕业生解决复杂工程问题的能力。然而, 当前教学实践中存在数据收集不全、分析能力欠缺、模式单一及与产业脱节等问题。为此, 文本探讨在构建教师数据素养培训框架的同时, 通过数据驱动优化教学质量和学习路径, 在工程教育中融入大数据和人工智能技术, 构建精准教学和持续改进体系。该体系展现了技术推动教育信息化、智能化转型的潜力, 为工程教育创新发展提供了新思路。

**关键词:** 工程教育, 大数据, 人工智能

## Research on Precision Teaching and Improvement Mechanisms in Engineering Education Integrating Big Data and Artificial Intelligence

Zhiqiang Yu Zuo Jiang Xuewen Tan

School of Mathematics and Computer Science, Yunnan Minzu University,  
Kunming 650504, China;

**Abstract**—Engineering education accreditation, as an internationally recognized quality assurance mechanism, emphasizes a student-centered, outcome-oriented approach aimed at developing graduates' ability to solve complex engineering problems. However, current teaching practices face issues such as incomplete data collection, inadequate analytical capabilities, single-mode teaching, and a disconnect from industry needs. Therefore, this study explores the construction of a teacher data literacy training framework, while utilizing data-driven methods to optimize teaching quality and learning paths. It incorporates big data and artificial intelligence technologies into engineering education to build a precision teaching and continuous improvement system. This approach demonstrates the potential of technology in driving the informatization and intelligent transformation of education, offering new ideas for the innovative development of engineering education.

**Keywords**—Engineering Education, Big Data, Artificial Intelligence

### 1 引言

在当前的技术变革中, 工程教育认证不仅关注课堂教学, 还积极拥抱大数据与人工智能技术, 推动教学模式向精准化、个性化发展。通过分析学生的学习数据, 教育者能够更科学地调整教学策略, 为学生提供定制化的学习支持<sup>[4]</sup>。

随着工程教育认证要求的深入推进, 大数据与人工智能的技术优势逐渐在教育领域显现出强大的应用

潜力。大数据为教育决策提供了坚实的科学基础, 通过对海量数据的采集与分析, 高校能够精准掌握学生的学习需求和痛点, 从而优化教学资源配置并设计更高效的教学方案<sup>[5]</sup>。与此同时, 人工智能技术则为教育智能化和个性化的发展开辟了全新路径。通过 AI 算法, 智能学习平台能够分析学生的学习轨迹和知识掌握情况, 提供实时的个性化学习建议, 并生成适应个人需求的学习路径<sup>[6]</sup>。

教学精准性的重要性不仅体现在教学实施环节, 也贯穿于教学改进的全过程。在教学评价方面, 大数据和人工智能技术的结合推动了动态化、个性化和智能化的评价体系<sup>[7]</sup>。教育者可以实时掌握学生的学习行为数据, 通过对其知识掌握度、学习效率和学习态度的综合分析, 提供精准的教学反馈。这种基于数据驱动的教学改进方式, 不仅提高了教学效率, 还显著

\* **基金资助:** 本文得到国家民委高等教育教学改革研究项目“工程教育认证背景下基于数据驱动的教学过程精准分析及持续改进研究”(编号: 23120); 云南省高等学校计算机基础教学指导委员会和云南省高等学校计算机教学研究会教学研究项目“基于大语言模型的计算机实践教学体系构建研究”(编号: 云高计教 202409) 资助。

\*\* 通讯作者: 蒋作 jiangzuo@ymu.edu.cn

改善了教学质量<sup>[8]</sup>。

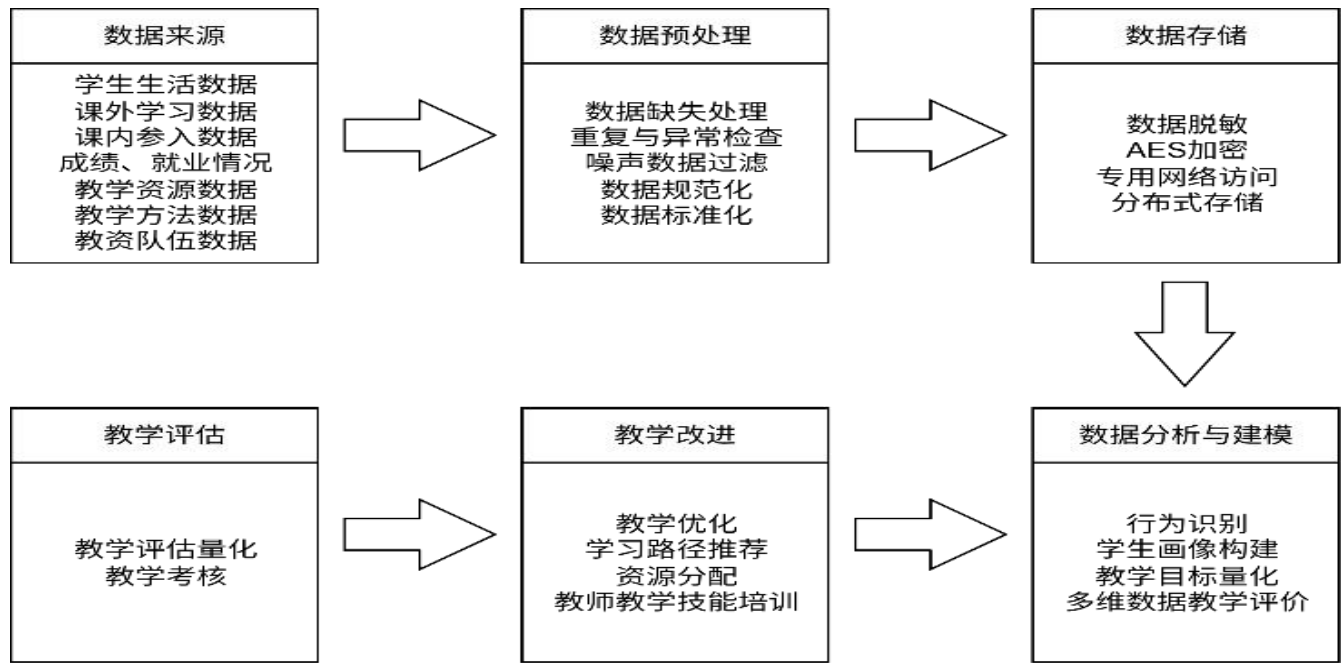


图 1 精准教学改革框架结构图

尽管工程教育在大数据和人工智能的助推下取得了显著进步，但教学实践中依然存在一系列亟待解决的问题<sup>[9]</sup>。这些问题主要集中在数据收集与整合、数据分析能力不足、教学模式单一以及与产业需求脱节等方面。同时教师信息化能力不足，也是影响教学改革的重要因素。鉴于此，本文通过大数据建模，人工智能辅助教师动态掌握学生学习实时状况，以学生为主，教师为辅的精准教学改革研究。

图 1 展示了基于大数据和人工智能技术的数据驱动教育推荐与学习资源分配框架，其核心目标在于实现个性化教学和学习路径优化。首先通过多维度采集与深度清洗学生成绩轨迹、课堂互动频次及作业质量等教育数据，构建标准化数据库；在此基础上生成融合知识掌握度、思维模式与兴趣偏好的三维学生能力画像，运用贝叶斯预测模型实时解析学习轨迹，动态优化个性化学习路径；最终通过自适应资源调度网络精准匹配微课视频、智能习题及拓展资料组合，依托学习反馈强化机制持续迭代推送策略，实现教学资源供给与个体认知节奏的深度契合，形成“分析-推荐-验证-优化”的智能教育闭环。

## 2 教育数据来源与处理

根据工程教育认证一些关键指标及数据需求，如课程的培养目标和达标情况，学生的学业成绩就业率，课程内容课程设置和相关的匹配程度。基于调研法调查工程教学的基本特征，调查基本工程专业(以计算机工程为例)的教学现状和教师、学生等基本情况。

本文通过某高校教务系统和学生在校时使用的“学习通”、“钉钉”、“雨课堂”等学习平台获取学生的数据包括基本信息数据、在校成绩和就业情况数据、学习生活行为数据。通过观察法查看教师的基本情况，包括师资队伍、教师评分（学生对教师的打分）、教师的备课情况。工程专业的教学现状（以计算机工程为例），包括学生反馈、教育部要求、行业需求等。教育大数据总体信息介绍如图 2 所示。

从以上数据集中可以提取出教师、教学现状、学生的基本情况。三个部分相互呼应让数据部分更为准确，其中主要以学生部分为主体。侧重点数据包含学生的上网时长、访问校内资源内容、网络费用、消费地点、消费记录、消费项目、消费金额、消费频率、选修课类型、选修课成绩、选修课与本专业相关程度、年龄、性别、课程成绩、课堂互动记录、挂科情况、是否就业、就业是否与所学相关的，的相关信息。

然而数据智能转化面临三重壁垒：多源异构数据存在格式混乱、噪声干扰及隐私安全风险，叠加教育从业者的技术短板，形成数据处理瓶颈。解决这些问题的关键需建立“标准-采集-存储”三维治理体系。包括制定统一采集规范，明确字段格式，兼容 LTI/Ed-Fi 国际标准；通过 LMS 和智能设备实现自动化采集，拓展移动端多场景覆盖；采用 Hadoop/Spark 分布式架构，分离处理结构化（MySQL）与非结构化数据（S3 对象存储），构建分级安全防护网络。

如何构建一个高质量数据集，需要采取多项处理

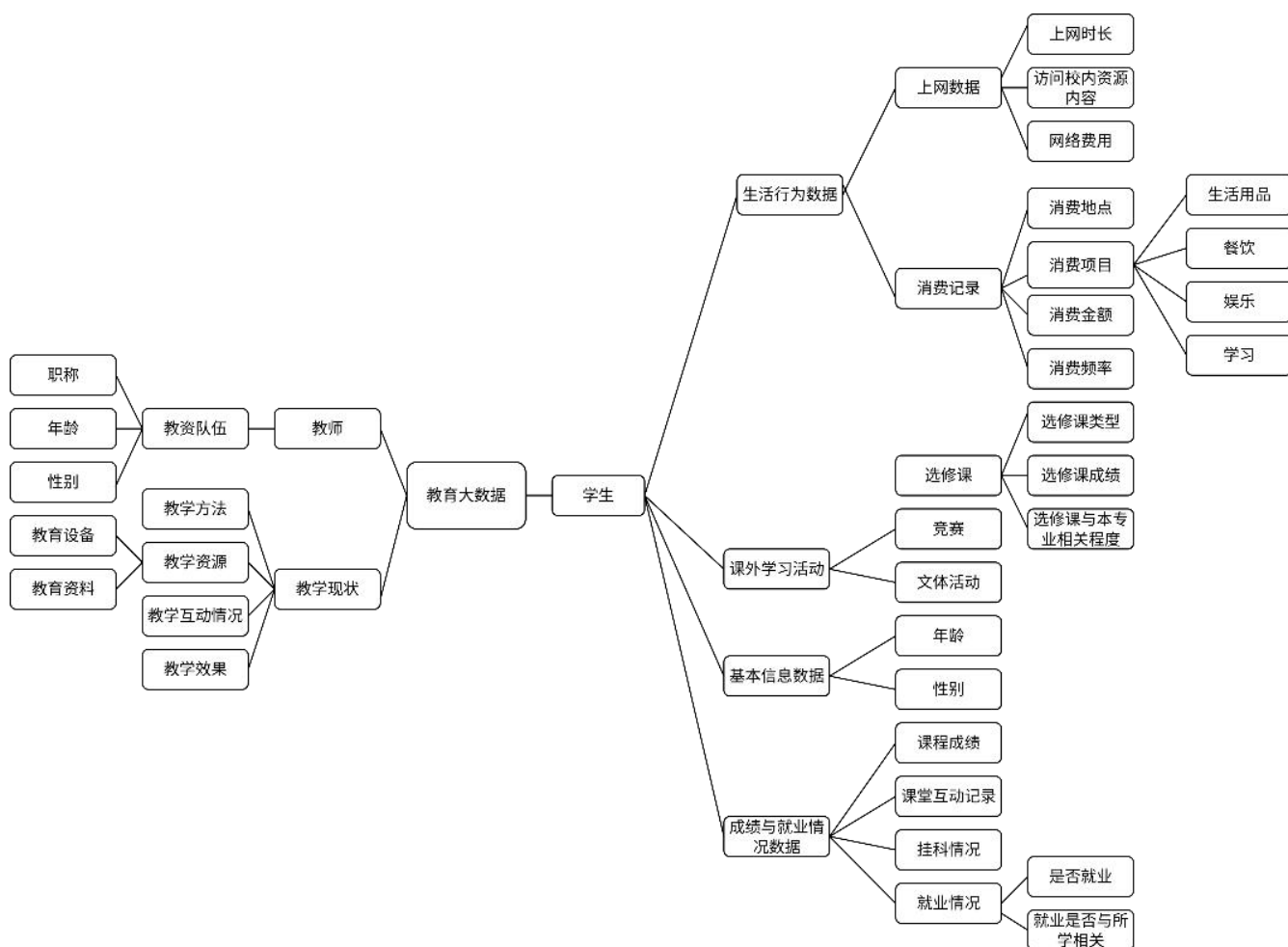
操作。首先,全面检查原始数据中是否存在缺失、重复和错误信息,并通过数据清洗提高数据质量,从而确保后续分析的准确性。接着,对数据进行二次提取、规范化处理和有意义的转换,根据信息价值对数据进行分类,提取出关键部分,并将复杂的数据结构简化或聚合,使其更易于理解和处理。同时,删除与分析无关的数据,防止无关信息干扰结果。随后,通过统计分析可视化技术,挖掘数据中隐藏的趋势和关联性,将部分数据以柱状图、饼图、折线图直观展示;对于文本信息,则进行个性化分析以提炼有效结论;而对于时间数据,则通过计算结束时间与上网开始时间的差值,提取上网时长等特征,为数据驱动的深入

研究提供支持<sup>[10]</sup>。

### 3 融合大数据与人工智能的精准教学分析方法

#### 3.1 学习行为模式识别模型

学生学习行为解析是破解个性化教育密码的关键路径。教育大数据挖掘技术通过解码知识掌握度与认知偏好,赋能教学范式转型——从经验驱动型授课转向数据支撑的精准学情画像。教学决策坐标由此完成从教师本位到学生本征的维度跃迁,实现教与学的认知同频共振。



对于多维度的学生特征数据,本文采用 k-means 遗传算法对学生行为特征进行提取。实现步骤如下:

首先,对清洗后的学生数据(包括性别、年龄、学历、每日学习时长、互动次数、视频时长、竞赛次数等)进行编码,构建数据矩阵;接着将数据归一化到[0,1],并采用浮点数编码生成染色体,即将 K 个聚类中心的 j 维向量连接,基因值随机取自[0,1]。最终,

随机生成 100 个染色体构成初始种群  $p(0)$ , 并以此构建适应度函数:

$$f_i = \frac{1}{1+J}$$

$$J = \sum_{i=1}^k \sum_{x \in c_i} w_i \text{dist}(c_i, x)^2$$

通过遗传算子（选择/交叉/变异）实施种群迭代，将适应度最低的染色体淘汰并以初始种群  $p(0)$  中适应度最高的染色体替换，从而形成新一代种群  $p(1)$ ；这一过程不断重复直至达到设定迭代终止条件，最终输出的种群中选取适应度最高的染色体，其三维坐标映射即为加权 K-means 最优聚类中心解集。据此，采用 k-means 遗传算法实现对学生行为特征进行提取

### 3.2 教学目标达成度的量化分析方法

为探究本文提出的方法对教学产生的影响，本文提出教学目标达成度量化分析模型。通过该模型可以判断学生完成度，并且通过回归分析方法确定教学与目标达成的因果关系。

评价学生在某方面的能力强度，可以泛化为一个具体的数字，采用单项指标评价，整体分为三类：知识目标大程度、能力目标大程度、素养目标达成度。定义如下：

知识目标达成度：

$$D_k = \frac{\text{正确回答的知识数量}}{\text{总知识点数量}} \times 100\%$$

能力目标达成度：

$$D_a = \frac{\text{项目总分}}{\text{满分}} \times 100\%$$

素养目标达成度：

$$D_s = \frac{\text{参与率}}{\text{目标参与率}} \times 100\%$$

通过单项指标，能够看出该学生在某一方面的达成度，通过这个达成度，可以判断该学生在该方面知识掌握程度或者能力体现。但是单项指标并不能体现一个学生的总体能力。因此本文采用如下综合达成度计算方法判断该学生整体素养。定义如下：

$$D_c = W_k \cdot D_k + W_a \cdot D_a + W_s \cdot D_s$$

其中  $D_k$ ,  $D_a$ ,  $D_s$  分别为知识、能力、素养目标达成度， $W_k$ ,  $W_a$ ,  $W_s$  相应的达成度指标的权重。

教学达成度分析基于师生信息、教学过程数据和量化目标，通过数据清洗、特征工程及数据挖掘进行评估。它涵盖课程目标、毕业要求和培养目标等层级，用可量化指标全面衡量学生的知识掌握与能力发展。借助分析结果，教师能够迅速发现学生问题，提供个性化辅导，同时为优化教学策略提供依据。教学达成度分析具体实施流程如图 3 所示。



图 3 教学达成度实施具体流程

### 3.3 构建基于多维数据的教学评价模型

传统方法主要关注学生个体，而教学改革质量却受教师、环境、设备等多重因素影响。为此，本文提出了一种基于多维数据的教学评价模型。模型整合了课程、学生学习、行为、反馈、教师教学、资源、设备、效果活动及竞赛等信息，全面评估教学效果。依据 SMART 原则（具体、可衡量、可实现、相关性、时限性），设计了量化指标：教学达成度指标涵盖课程目标完成率、考试通过率和平均分；学生行为指标包括学习参与指数（互动次数与课后完成率之和）和行为一致性（轨迹模式与目标路径匹配程度）；教师行为指标则关注课堂时间分配合理性和反馈覆盖率。为构建多维模型，采用随机森林方法，通过整合多棵决策树，不仅提升了预测准确性，还增强了抗过拟合能力，即使在高维特征下也能高效运。构建如下：

首先样本抽取：从原始数据集中有放回地随机抽取多个样本，生成多个子数据集，定义如下：

$$D^{(b)} = \{(X_i, y_i) | i \in S_b\}$$

其次决策树的训练：对于每棵决策树，利用随机抽取的特征集合进行分裂，生成回归树。定义如下：

$$Split\ Feature = argmin_{j \in Random\ Subset} \sum_{i=1}^m (y_i - \hat{y}_i)^2$$

其中  $j$  是特征索引， $m$  是当前节点中的样本数， $y_i$  是第  $i$  个样本的真实值， $\hat{y}_i$  是根据特征  $i$  进行分裂后的预测值。

最后对于新数据点的预测结果是所有决策树预测结果的平均值：

$$\hat{y}_i = \frac{1}{B} \sum_{b=1}^B \hat{y}^{(b)}(X)$$

其中  $\hat{y}_i$  是输入样本  $X$  的最终预测值， $B$  是决策树的数量， $\hat{y}^{(b)}$  是第  $b$  棵树的预测值。

对于多维度数据集，需要考虑综合问题，本文采用数据聚合与评分公式，综合评价函数如下：

$$E = \sum_{i=1}^n w_i \cdot S_i$$

其中  $E$  表示教学评价得分， $w_i$  表示第  $i$  个指标的权重， $S_i$  表示第  $i$  个指标的标准化得分。

## 4 数据驱动的教学过程持续改进机制

### 4.1 实时数据驱动的教学优化闭环

实时数据驱动的教学优化闭环是一种基于数据分

析的动态教学调整机制。通过持续采集教学过程中的多维数据（如学生学习行为、测试结果、课堂参与情况等），实时分析学生的学习效果和存在的问题，将结果反馈给教师 and 系统，用于即时优化教学策略和学习路径，形成“数据采集—分析—反馈—优化”的闭环流程。实时数据驱动的教学优化闭环框架图如图 4 所示：

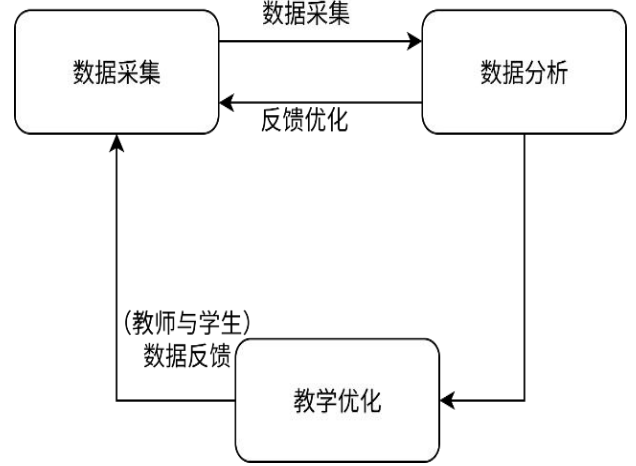


图 4 实时数据驱动的教学优化闭环框架图

从多种渠道实时采集学生学习数据，包括学习轨迹、作业表现、考试成绩以及课堂参与度。通过大数据分析与人工智能技术，诊断学生的学习状态，识别知识盲点与行为模式。将分析结果以可视化方式呈现，供教师参考，或通过智能推荐直接对学生提供资源建议。教师根据反馈调整教学内容、方法或节奏；学生接收到调整后的学习路径或推荐资源。采集优化后的数据，验证调整的效果并为下一轮优化提供依据。教学优化闭环能在教学中实现动态监控和灵活调整，使教学更加精准、个性化，最终提升学生的学习效率和教师的教学效果。

### 4.2 个性化教学推荐与自适应学习资源分配

个性化教学推荐与自适应资源分配是一种利用大数据和人工智能的动态教学机制。该机制根据学生的学习特征和需求，动态调整学习路径并推送合适的学习资源，实现教学内容与个性化需求的高效匹配。它不仅提升学生的学习效率和效果，还为教师优化教学设计提供数据支持。通过分析学生行为、学习偏好和知识掌握情况，系统能够为每位学生打造“千人千面”的个性化学习体验。

个性化推荐内容包括：课件、视频、练习题目、拓展阅读材料等。基于知识点关联关系，为学生定制学习顺序。针对学生薄弱环节提供复习或巩固资源。通过学生画像技术和机器学习模型，预测学生的学习需求，并根据学习目标推送相应的资源。

自适应学习资源分配以动态调整为核心,通过实时监测学生学习进展,对资源分配和学习路径进行优化,保证学生在不同学习阶段获得匹配其需求的内容。

自适应资源分配根据学生实时学习表现动态调整资源,优先分配学生尚未掌握或理解较差的知识点资源,学习压力过大或进展顺利时动态调整学习任务负荷,通过知识点掌握度、行为模式分析等数据,判断学生当前的学习需求,合理分配学习资源。

根据行为模型识别算法对每个学生特征进行识别,对识别出来的结果本文采用个性化教学推荐与自适应学习资源分配算法对每个学生进行资源分配。资源分配框架如图 5 所示:

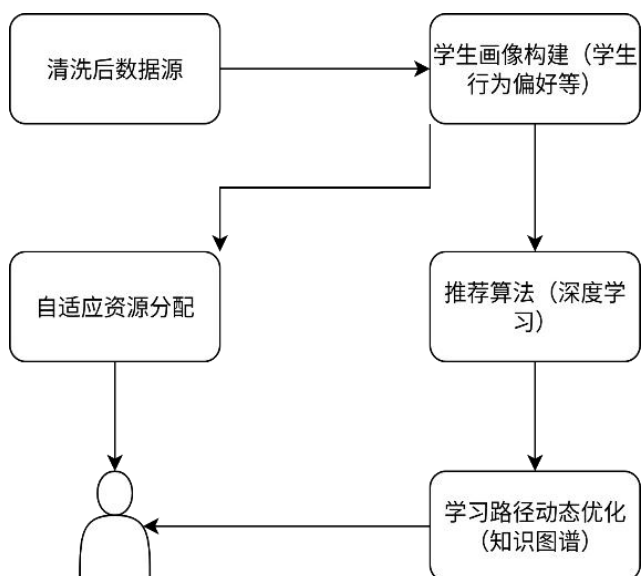


图 5 根据学生行为、偏好等进行资源分配整体流程

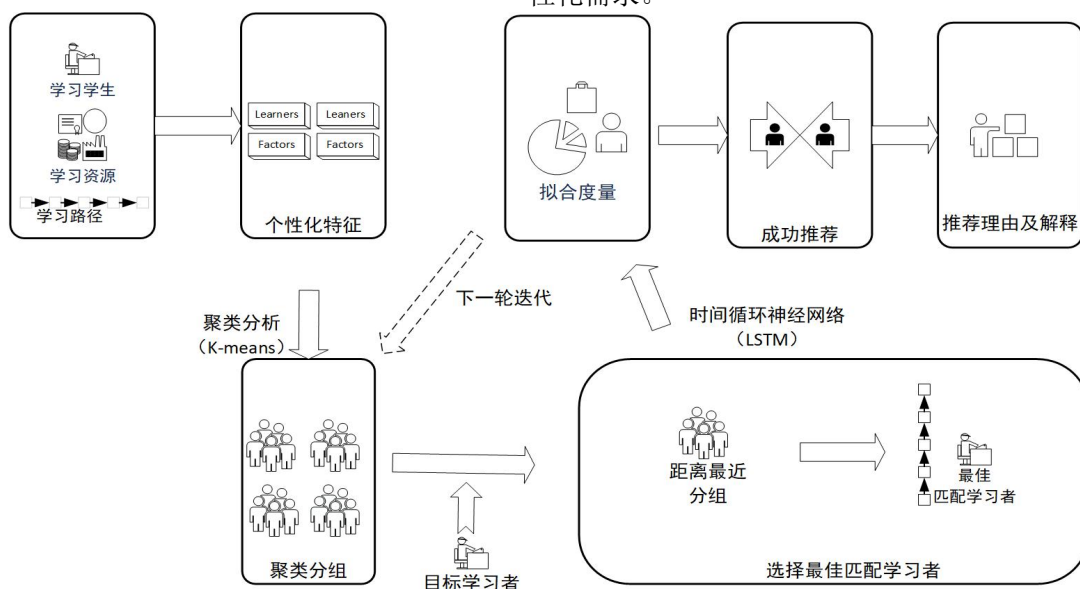


图 6 个性化教学推荐与自适应学习资源分配具体流程

如图 6 所示的基于大数据与人工智能的个性化教学自适应资源分配流程:先对多源教育数据(学习平台、课堂行为、考试成绩等)进行清洗标准化,构建分析基础;其次利用行为数据挖掘和模式识别技术,解析学习偏好、知识掌握度等特征形成学生画像;随后通过深度学习推荐算法,结合实时学习状态与历史数据预测适配资源,同时运用知识图谱技术分析知识点关联性,动态优化学习路径;最终由自适应分配模块整合推荐内容与优化路径,通过动态调整机制持续匹配学习进度与需求变化,实现个性化课程、习题等资源的精准分发。

#### 4.3 教师数据素养与技术能力的培训

大数据与人工智能推动教育变革背景下,教师数据素养与技术能力成为教学转型的核心要素。传统经验型教学正转向基于学习行为与成绩数据的精准化教学,要求教师掌握智能化工具操作技能,通过数据分析制定个性化方案、诊断学习问题并动态调整策略。教育机构需强化技术培训体系,既助力教师利用数据科学优化教学决策,提升因材施教效率,又为其适应智能化教学环境提供持续发展支撑,最终实现教育质量系统性升级。

教师数据素养与技术能力培训旨在提升教师对教育数据的理解与应用,以适应数字化转型。教师需掌握数据清洗、可视化、统计分析等基本技能,从多源动态数据中洞察学生规律、优化教学策略,同时熟悉隐私保护法规确保数据安全。掌握 LMS、在线教学工具和数据可视化软件(如 Tableau)也是必备技能。随着人工智能普及,教师还需了解 AI 原理,熟练使用智能题库和个性化推荐系统,提高教学效率,满足个性化需求。



图7展示了以教师数据素养与技术能力提升为核心的培训体系的逻辑框架,涵盖了从培训内容到实施和反馈优化的完整闭环。该框架围绕数据驱动教学能力提升,构建五大核心环节:数据素养聚焦多源教育数据的采集、处理与分析,为教学决策提供依据;技术能力强化智能工具(如LMS、AI推荐系统)的操作训练,促进数据向教学策略的高效转化;数据伦理贯穿全流程,强调隐私法规与敏感信息保护;混合培训模式融合线上课程与线下实操,依托社群协作与即时反馈深化技能内化;最终通过动态评估体系跟踪教学实践效果,建立“培训-应用-优化”反馈闭环,持续提升教师数字化胜任力。

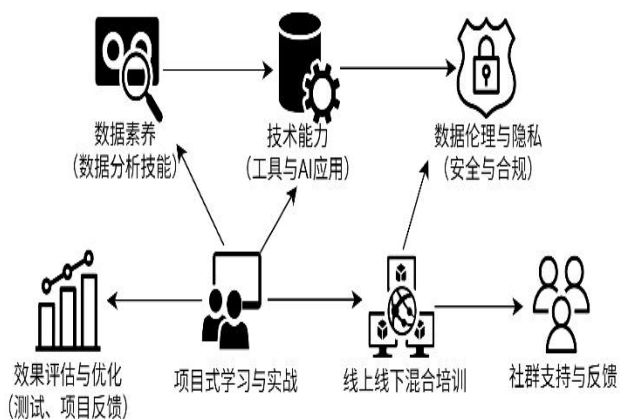


图7 教师数据素养与技术能力的培训框架图

## 5 数据安全与隐私保护

### 5.1 教育数据管理中的安全风险分析

随着信息技术的快速发展,教育领域对数据的依赖日益增加,但与此同时,教育数据管理中也面临诸多安全风险。数据泄露风险尤为突出。教育数据中包含学生个人信息、学业成绩等敏感数据,一旦遭受非法获取或泄露,可能导致隐私侵害,甚至引发身份盗用等问题。数据存储与传输中的技术漏洞增加了风险。一些教育机构采用的存储系统或网络传输未实现严格的加密保护,使得数据在未授权访问、网络攻击(如中间人攻击或DDoS攻击)中易受损害。内部管理不当也构成安全威胁。教育机构的工作人员若缺乏网络安全意识或制度执行不严,可能因不慎操作或权限滥用,导致数据丢失或滥用。外部供应商服务中的安全隐患不可忽视。教育系统广泛使用第三方平台和云服务,若供应商的安全措施不足或发生数据外包引发的权限扩散,将进一步放大风险。加强教育数据的加密保护、提升管理者与使用者的安全意识、健全法律法规,并建立完善的安全管理机制,是保障教育数据安全的关键。

### 5.2 数据存储与传输中的隐私保护技术

在教育领域,确保学生和教师数据隐私至关重要,尤其在数据存储与传输中易遭泄露。为此,本文设计了三个数据隐私保护处理过程,首先采用AES加密敏感信息,并利用TLS或VPN保护数据传输安全,有效防止未授权访问。其次,通过强密码、多因素认证(MFA)和基于角色的访问控制(RBAC)限制访问权限,并借助详细日志追踪数据使用情况。与此同时,差分隐私技术通过引入随机噪声,在保证统计准确性的同时防止个体信息泄露;而区块链技术则利用其去中心化和不可篡改特性,为数据透明存储和共享提供新思路。通过这些处理技术,并辅以完善的管理制度与法律保障,可最大限度地确保教育数据的隐私安全。

### 5.3 教学数据安全策略与法律规范

教学数据涉及学生的个人信息、学业表现等敏感数据,其安全管理至关重要。为有效防范数据泄露与滥用,需制定全面的安全管理策略,并严格遵守相关法律规范。

#### (1) 安全管理策略

构建权责清晰的安全管理体系,同步建立风险评估机制-教育机构应定期实施网络安全评估,形成“漏洞扫描-修复验证”动态闭环。数据防护双擎驱动:存储端采用高强度加密技术,传输链路部署端到端加密保障教学数据全流程安全;访问控制端依托RBAC模型,按角色精准划分数据操作权限,遏制越权行为。意识防线筑基:针对管理人员与教师开展周期性安全培训,通过模拟演练、案例解析降低人为失误风险,筑牢“人防+技防”协同屏障。

#### (2) 法律规范

法律合规基线:严格遵循《个人信息保护法》《网络安全法》等国内法规,确保数据收集、存储及使用的合法必要;跨境协作适配:国际合作项目需对标GDPR等国际标准,规避合规风险;第三方强监管:通过协议约束第三方服务商权责,堵塞外包环节数据泄露漏洞。最终依托“技术加固防线、管理细化流程、法律兜底边界”的协同机制,实现教学数据安全与教育信息化发展的动态平衡。

## 6 结论与展望

本文以大数据与人工智能在工程教育中的应用为切入点,构建了一套融合技术与教育实践的精准教学与持续改进机制。通过建立教学数据挖掘与分析模型,采用学习行为模式识别、教学达成度分析及个性化教学推荐等方法,实现了教学过程的闭环反馈与动态调整,显著提升了教学的科学性与精准性。该机制不仅将学生学习行为数据转化为教学改进依据,为工程教育认证提供了量化分析支持,还通过多维数据分析优

化课程设计和资源自适应分配,助力培养符合新时代需求的高素质工程人才。展望未来,结合虚拟现实、增强现实、物联网及区块链等前沿技术,研究成果将进一步构建沉浸式、交互式教学场景,实现跨平台数据整合与安全管理,推动教育信息化和智能化转型,同时也为其他学科的教学改革提供了宝贵的理论基础和实践参考。

## 参考文献

- [1] 林健,杨冬.工程教育智能化:内涵、特征与挑战[J].清华大学教育研究,2023,44(06):1-11.
- [2] 卢建云,郑卉,吴焱岷,武春岭.工区块链技术与原理课程思政设计与实践[J]. 计算机技术与教育学报, 2025,13(1): 18-23.
- [3] 杨勇毅,李陶深. 无人驾驶汽车车路协同任务卸载计算研究进展[J]. 计算机技术与教育学报,2025,13(1): 24-28.
- [4] 刘昊,张英婷,陈佳,韦铸娥.基于生产过程中检测方法的研究与优化决策[J], 计算机技术与教育学报,2025,13(1): 29-35.
- [5] 杨爽,宫正,王宗跃.新工科背景下机器人学课程教学改革与实践[J]. 计算机技术与教育学报,2025,13(1): 65-70.
- [6] 刘佳荟,张思佳,胡泽元,刘明剑.智慧教育时代数字赋能高校教学的改革探讨[J]. 计算机技术与教育学报,2025,13(1): 76-81.
- [7] 丁云霏.数字化背景下我国高等教育发展趋势与实践路径[J].扬州大学学报(高教研究版),2024,28(05):21-27.
- [8] 李亚坤,颜荣恩,杨波,李冬梅,苗苗. 生成式人工智能背景下高校软件工程课程的教学改革与探索[J]. 计算机技术与教育学报,2024,12(5):8-12.
- [9] 段玲玲,马伟杰,田雪.基于教育数字化转型的高校教师信息化教学能力提升[J].管理工程师,2024,29(05):65-69.
- [10] 陈军成,王猛. “物联网感知技术课设”课程建设与探索[J]. 计算机技术与教育学报,2024,12(5): 70-74.