

# 大数据驱动的高校计算机专业学生学习行为分析与个性化教学策略设计

武丹丹

新疆天山职业技术大学 新疆 乌鲁木齐 830000

**摘要:** 针对高校计算机专业教学中知识碎片化、学生差异大与教学同质化的矛盾, 本文构建大数据驱动的学习行为分析与教学优化体系。整合编程平台、课程系统等多源信息, 运用聚类算法与关联分析模型, 精准识别学生在编程实践、理论学习中的能力短板与行为偏好, 构建学生能力画像。基于此设计分层教学、动态干预等个性化策略, 并借鉴湖南大学经验强化实践教学。实验显示, 该体系有效提升学生成绩, 增强知识掌握度, 助力专业从“批量培养”向“精准育人”转型。

**关键词:** 大数据; 计算机专业; 学习行为分析; 个性化教学

引言: 在高校计算机专业教学中, 知识碎片化、学生差异大与教学同质化的矛盾日益凸显, 传统教学模式难以满足多样化学习需求。大数据技术为破解这一难题提供了新路径, 通过整合多源数据、深度挖掘学习行为规律, 可精准把握学生能力短板与行为偏好, 进而设计个性化教学策略, 推动专业教学从“批量培养”向“精准育人”转变。

## 1 大数据驱动的学习行为分析框架

### 1.1 数据采集与整合

以“全场景覆盖、多维度关联”为目标构建数据源体系。参照湖南大学计算机专业教学实践, 从三方面采集核心数据: 编程实践数据, 通过 EduCoder 平台、Git 代码仓库及“口袋猫”实验板, 获取代码提交频次等细粒度及硬件调试信息; 理论学习数据, 借助 MOOC 平台与智慧课堂系统, 收集视频观看进度等行为记录; 综合反馈数据, 整合课程作业成绩等多方面信息。采用分布式存储架构, 以学生唯一学号建立关联索引, 实现跨平台数据融合, 为后续分析提供完整基础, 保障行为分析全面且有关联性。

### 1.2 数据预处理与特征工程

数据预处理聚焦“去噪提纯、标准化转换”, 解决多源数据格式异构、质量参差的问题。首先通过异常值检测算法剔除无效数据, 如编程平台中误操作产生的重复提交记录、理论课程中偶然点击导致的异常学习时长; 对缺失数据采用针对性填充策略, 实践类数据采用同水平学生均值填充, 理论类数据采用邻域插值法补充。随后进行数据标准化处理, 将代码调试时长、测验分数等不同量纲指标转换至[0,1]区间, 消除数值差异带来

的分析偏差。特征工程阶段重点构建计算机专业特色特征集, 从编程数据中提取代码复用率、错误修复效率等衍生特征, 从理论学习数据中提炼知识点停留时长、错题关联度等特征, 最终形成包含参与度、专注度、实践能力等5个维度32项特征的结构化数据集, 为模型构建提供高质量的输入特征, 提升行为分析的精准度<sup>[1]</sup>。

### 1.3 行为分析模型构建

行为分析模型构建采用“分类聚类结合、特征关联挖掘”的复合架构, 契合计算机专业学习行为的复杂性。基础层运用K-means聚类算法对学生行为数据进行群体划分, 结合肘部法则确定最优聚类数, 精准识别主动探究型、任务驱动型等差异化学习群体, 为分层教学提供依据。核心层构建随机森林算法模型, 以32项行为特征为输入, 以课程综合成绩、项目完成质量为输出目标, 挖掘代码调试效率、讨论互动深度等关键特征与学习效果的关联机制, 特征重要性排序明确核心影响因素。决策层整合聚类结果与关联分析结论, 开发可视化分析平台, 实时呈现学生个体行为轨迹与群体特征分布, 如编程实践中的高频错误类型、理论学习的薄弱知识点模块。该模型经湖南大学教学数据验证, 行为特征识别准确率达89%, 为教学策略设计提供可靠的数据支撑。

## 2 大数据驱动的高校计算机专业学生行为特征分析

### 2.1 编程实践行为分析

编程实践行为分析聚焦“过程特征提取、能力差异识别”, 揭示不同水平学生的实践行为规律。通过对编程平台数据的深度挖掘发现, 优秀学生的编程行为呈现“规划先行、迭代优化”特征, 在“口袋猫”实验板开发等实践任务中, 会提前构建代码框架, 调试次数虽多

但错误修复效率高,代码复用率达65%以上,且善于利用讨论区分享调试思路。中等水平学生多呈现“任务导向、稳步推进”特点,能完成基础功能开发,但代码冗余度较高,错误修复依赖参考案例,拓展功能实现率不足40%。基础薄弱学生则存在“盲目试错、缺乏规划”问题,代码提交频次高但有效率低,常因语法错误反复修改,硬件调试中对电路连接逻辑理解模糊。此外,时序分析显示编程学习的“黄金时段”为19:00-22:00,此时间段提交的代码正确率比其他时段高17%,为实践教学时间安排提供数据依据。

## 2.2 理论课程学习行为分析

理论课程学习行为分析围绕“知识吸收规律、互动参与特征”展开,结合计算机专业理论课程的抽象性特点挖掘行为价值。数据显示,系统类课程如操作系统、计算机组成原理的学习中,学生行为差异显著:主动型学生在视频学习时会频繁暂停标记,章节测验正确率达90%以上,且会主动发起“软硬件协同设计”等深度讨论;被动型学生多以倍速播放完成视频学习,测验错题集中在内存管理、指令集架构等核心模块,互动参与多为“点赞”等浅层行为<sup>[2]</sup>。知识点关联分析发现,数字电路课程中“逻辑门设计”的掌握程度,与后续计算机系统课程的“汇编器实现”成绩呈强正相关(相关系数0.78)。同时,理论学习存在“前松后紧”现象,课程初期学习时长不足,临近考核时突击学习,这种行为导致的知识遗忘率比规律学习高45%,为教学节奏调整提供明确方向。

## 2.3 学生能力画像构建

学生能力画像构建采用“多维度标签化、动态更新”模式,整合编程实践与理论学习的分析结果,形成全面立体的学生认知模型。画像核心包含四大维度标签:一是基础能力标签,涵盖编程语言掌握程度、数据结构应用能力等,如“Java熟练+Python入门”“链表应用薄弱”;二是实践创新标签,基于“口袋猫”实验、原型机开发等任务表现,标注“硬件调试能力强”“系统设计创新不足”等特征;三是学习习惯标签,如“规律学习型”“考前突击型”“讨论互动积极”;四是潜力发展标签,结合拓展任务参与度预测发展方向,如“嵌入式开发潜力大”“算法研究兴趣浓厚”。画像采用可视化仪表盘呈现,教师可直观掌握学生能力短板,学生能清晰认知自身优势。该画像与济南三中数字画像实践理念一致,且每两周基于最新学习数据更新一次,确保画像的时效性,为个性化教学提供精准靶向。

## 3 基于学习行为分析的个性化教学策略设计

### 3.1 分层教学策略

分层教学策略依据学生能力画像与聚类结果,构建“基础-进阶-拓展”三级教学体系,打破传统班级授课的同质化局限。针对基础薄弱学生,聚焦编程语法、理论概念等核心知识点,采用“微视频+小任务”模式,如将C语言指针知识点拆解为5个微任务,配套即时练习与错题解析,同时安排一对一帮扶;中等水平学生以能力提升为目标,设置“基础功能+优化任务”的编程项目,如在原型机开发中要求完成基础指令集后,尝试优化代码执行效率,理论课程中增加案例分析环节;优秀学生侧重创新培养,提供科研课题、开源项目参与机会,如加入华为鲲鹏开发板相关研发任务,组织技术沙龙分享研究成果。教学内容通过智能平台精准推送,考核评价实行“分层标准”,基础层侧重知识掌握,拓展层强调创新成果,确保各层次学生均能获得适配的学习挑战与成长空间。

### 3.2 动态干预策略

动态干预策略依托实时行为数据监测,构建“预警-干预-反馈”的闭环机制,及时解决学生学习困境。基于行为分析模型设置预警阈值,当出现编程错误率连续3天超60%、理论课程学习时长骤降50%等异常情况时,系统自动触发预警并推送干预方案。干预方式实行“分类施策”:针对学习态度问题,推送个性化激励消息与学习规划建议;针对能力短板,提供定制化学习资源,如代码调试薄弱学生推送典型错误案例库;针对心理压力问题,联动辅导员开展疏导沟通。干预效果通过后续行为数据实时反馈,如干预后代码正确率提升至45%以上则解除预警,若未达预期则调整干预策略<sup>[3]</sup>。该策略在操作系统课程应用中,使学习困难学生的课程通过率从58%提升至82%,有效降低了学业风险,体现了教学干预的精准性与时效性。

### 3.3 实践环节强化策略

实践环节强化策略借鉴湖南大学“虚实融合”教学经验,构建“螺旋递进+项目驱动”的实践体系,破解学做“两张皮”问题。以“程序设计→系统设计→系统工程”能力链为核心,设计阶梯式实践任务:大一通过“口袋猫”实验板掌握硬件基础,完成简单电路焊接与代码编写;大二利用HNU原型机进行汇编器、链接器开发,深化系统理解;大三参与物联网应用开发等综合项目,培养复杂问题解决能力;大四结合企业需求完成毕业设计。同时搭建在线实践平台,实时采集实践过程数据,自动分析调试漏洞并推送解决方案,如检测到“口袋猫”串口通信错误时,推送电路连接示意图与代码修

改建议。引入企业真实项目如香橙派开发板研发,让学生在实战中积累经验,实践能力提升效果显著,学生开发的智能监护系统已获创业基金立项。

### 3.4 教师-学生协同策略

教师-学生协同策略以“数据共享、双向互动”为关键核心,全力构建新型教学共同体。为达成这一目标,专门搭建了功能完备的师生协同平台。在教师端,平台能够实时呈现学生的能力画像与详细的行为分析报告。教师依据这些精准信息,灵活且及时地调整教学方案。例如,当发现多数学生在编译原理的“语法分析”部分掌握不扎实,便迅速增加相关案例讲解的课时,强化知识传授;学生端同样功能丰富,学生可随时查看个人学习报告,清晰了解自身优势与不足,明确改进方向。并且能通过平台向教师发起精准提问,提问时还可附带代码片段、错误截图等详细信息,方便教师精准把握问题;课堂上,建立“启发式授课+探究式讨论”的创新模式。教师结合数据反馈,抛出具有针对性的问题,如“如何运用底层系统思维优化代码性能”,引导学生展开深度思考与探究。同时,借鉴鲲鹏虚拟教研室的宝贵经验,组织跨班级学习小组,让不同能力层级的学生共同协作完成项目。在此过程中,优秀学生大方分享经验,基础薄弱的学生则能获得悉心指导。此外,建立四维评价机制,综合考量学习过程、教师评价、第三方测试以及企业反馈,确保评价全面客观,有力促进师生共同成长。

## 4 未来方向

### 4.1 引入AI大模型辅助个性化答疑

引入AI大模型辅助个性化答疑,旨在构建“即时响应、精准适配”的智能教学支撑体系,弥补传统答疑中教师精力有限的不足。基于计算机专业知识库训练专属大模型,整合编程语言手册、系统开发案例、硬件调试指南等专业资源,针对学生编程实践与理论学习中的问题提供定制化解答。该模型具备三大核心能力:一是语境理解精准,能识别代码报错信息、理论概念混淆等专业问题,如准确判断“空指针异常”的具体成因并给出修复方案;二是答疑方式适配,为基础薄弱学生提供步骤化讲解,为优秀学生推送拓展思路,契合不同能力画像需求;三是学习过程联动,将答疑数据同步至学生能力画像,更新知识薄弱点标签,为教学策略调整提供依

据<sup>[4]</sup>。同时设置人工复核环节,确保复杂问题解答的准确性,形成“AI初答+教师精评”的答疑模式,提升学习支持的及时性与专业性。

### 4.2 构建校际共享数据平台,扩大研究样本规模

构建校际共享数据平台,以“资源整合、协同研究”为目标,突破单一院校研究的局限性,提升行为分析与教学策略的普适性。借鉴教育部—华为鲲鹏虚拟教研室的经验,联合不同层次高校计算机专业共建数据共享联盟,制定统一的数据采集标准与安全规范,明确数据脱敏规则,确保学生隐私安全。平台核心功能包括:数据汇聚模块,整合各院校编程平台、课程系统的学习行为数据,形成大规模数据集;协同分析模块,联合多校科研力量优化行为分析模型,挖掘不同教学模式下的行为规律;资源共享模块,发布优秀教学案例、个性化策略方案等成果,促进校际经验交流。通过扩大样本规模,覆盖不同地域、不同层次学生的学习特征,使分析结论更具代表性,教学策略更具推广价值,推动计算机专业教学改革的整体进步,为关键领域人才培养提供协同支撑。

## 结束语

本文构建的大数据驱动教学体系,整合多源数据、精准分析行为并设计个性化策略,有效解决了高校计算机专业教学痛点。实践显示其能精准识别学生差异,提升编程实践成绩,为“精准育人”提供可行方案。体系融合先进教学经验,契合专业培养需求。未来,引入AI大模型、建设校际数据平台,将提升其智能化与普适性。高校计算机专业教学应深化大数据应用,以数据赋能创新,强化学生能力,为行业培养高质量人才,履行教育使命。

## 参考文献

- [1]王萌.基于大数据分析的高校计算机教育个性化教学策略探究[J].办公自动化,2025,30(13):46-48.
- [2]云洁,白雪.大数据驱动的计算机网络远程教学质量评价研究[J].现代电子技术,2021,44(9):166-170.
- [3]郭甜甜,樊蒙蒙.大数据技术驱动下高校计算机专业教学改革探索[J].网络游戏技术与应用,2025(2):145-147.
- [4]郝宇博.大数据背景下计算机科学与技术课程教学改革策略研究[J].E动时尚,2025(9):142-144.