

# 多模态分析视角下基于数据挖掘的职业院校 教学模型构建与教学行为诊断分析

张 芳 池光胜

山东工程职业技术大学, 中国·山东 济南 250200

**【摘要】**随着云计算等新兴技术在教育领域深度融入与广泛应用,亟需将海量数据提炼转化为有实际效用的信息,动态再现整体趋势和个体差异,从而提供可靠的教学反馈。教育数据挖掘ED (Education Data Mining) 是一项至关重要的技术与研究范畴。学习表现使得学习历程中所展现的行为模式、动态反应及应对策略展现出来,深刻映射其态度取向、学习动机、能力层级及学习习惯等内在心理特征。通过对教育数据的测量、收集、分析和报告,运用不同的技术对学习数据进行分析,能发现学生学习表征,能有效反馈教学支持个性化教学。

**【关键词】**多模态分析; 数据挖掘; 教学模型构建; 教学行为诊断

**【基金课题】**2023年度山东省智慧教育研究项目《多模态分析视角下基于数据挖掘的职业院校教学模型构建与教学行为诊断分析》,项目编号202359; 2023年度山东工程职业技术大学教学改革项目《职业院校“一引领两驱动双三化四融合”课程构建与数字化评价改革与实践》,项目编号JG202323。

## 1 国内外研究现状

2022年,我国学者顾小清针对技术密集型环境,深入探究了技术的精细分析与诊断方法,倡导运用序列解析策略来剖析图像化学习历程的数据,旨在挖掘促进知识与能力发展的图像化学习行为模式,通过数据整合与洞察、信息检索与挖掘、预测与干预等手段,整合多元化的数据资源,确保教育实践形式的精准实施。曹佩声(2022)基于翻转课堂的ESP课程混合式教学模式构建与效果研究中认为评价维度可以从学生参与度、师生互动性、学生对教学模式的适应性以及教学的效果、学生学习满意度等来评价。赵志、刘卫东、肖鹏等也做了大数据背景下的学生学业表现预测模型等进行了研究,学业评价模型的建立与修正对教学行为存在有效反馈。以上做出的学生学习行为模型构建,从数据清洗、数据分析、模型建立等方面都进行了探索和实践,而基于关于线上在线数据的模型构建能有效指导课程改革,后续也将是促进职业院校课程蓬勃发展的有效途径。

## 2 基于数据挖掘的职业院校教学模型构建与教学行为诊断分析

### 2.1 基于因子分析的学习行为数据采集与处理

利用“超星学习通”平台搭建在线开放课程《高等数学》,将职业院校学习行为划分为线上学习行为和线下学习行为。本研究择取了2023-2024学年参与混合式教学的1614名学生作为研究样本,课程周期共计16周。研究主要

依据每位学生自初次登录至最终登录平台期间的数据进行分析,设立了八项教学成效评估指标。 $X_1, X_2$ 是平台登录时长与内容观看时长, $X_3, X_4, X_5, X_6$ 分别是登录次数、观看次数、讨论课堂次数,参与论坛讨论次数, $X_7, X_8$ 分别为联系完成度与测试完成度。

在此过程中,对学习行为采用因子分析法(亦称维度降维技术),其核心策略在于依据变量间的相关性强度进行分组聚合,确保组内变量高度相关,而组间则保持相对独立性。通过因子分析,任一变量均可表达为公共因子的线性组合形式。

共因子的线性函数与特殊因子之和。其中

$$X_i = a_{i1}F_1 + a_{i2}F_2 + \dots + a_{im}F_m + \varepsilon_i$$

通过矩阵间转化关系

$$\begin{bmatrix} X_1 \\ X_2 \\ \dots \\ X_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix} \begin{bmatrix} F_1 \\ F_2 \\ \dots \\ F_m \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon \\ \dots \\ \varepsilon_m \end{bmatrix}$$

可以筛选具备较强载荷的指标,这就是最终影响课程建设与影响的重点因素。

经过KMO检验统计量与Bartlett's球形检验同济,适合进行因子分析。检验结果表明, $X_1, X_2, X_3, X_6, X_4, X_3$ 六项指标呈现出显著的载荷强度,彼此间具有高度相关

性,可整合为一个维度——在线学习参与度因子;此外, $X_7$ 、 $X_8$ 部分指标亦承载较大载荷,共同构成课程测试因子。

最终可以得到各个因子的得分表达式

$$F_1 = 0.208X_1 + 0.214X_2 + 0.179X_3 + 0.238X_4 + 0.182X_5 + 0.174X_6 - 0.097X_7 - 0.162X_8$$

$$F_2 = 0.002X_1 - 0.03X_2 + 0.009X_3 - 0.155X_4 - 0.027X_5 - 0.008X_6 + 0.582X_7 + 0.744X_8$$

## 2.2 基于数据挖掘的学习行为模型构建

网络学习活动聚焦于学子在在线教育平台中的互动行为,涵盖资料检索、视频内容观摩及线上论坛交流等诸多方面。通过平台内置的“任务节点追踪”、“学习频次统计”及“分数管理”等功能模块,我们得以导出基础数据。具体而言,“分数”模块囊括了章节小测完成情况、综合评分等内容;“交流概况”模块则记录了讨论篇目、主动发帖与回复帖数等数据;“章节测验概览”模块则详尽列出了各次测验的分数及提交时间等关键指标。随后,采用Python软件执行数据清洗流程,剔除冗余数据行列,并对数据予以脱敏处理,以保障学生隐私安全。为避免影响聚类结果的准确性,对数据进行了标准化处理,并遵循标准正态分布原则,以确保数据间具备良好的可比性。

计算如下所示:  $z = \frac{x - \mu}{\sigma}$ , 其中,  $z$  代表标准化分数,  $x$  代表原始数值, 平均值  $\mu$  及标准差  $\sigma$  均基于该组数据计算得出。

依据历史特征工程提炼的学生学习行为数据构建模型,模型分析结果反馈给师生,既有助于教师精进教学策略、实施个性化教学干预,也为学生调整学习策略提供科学依据。

## 2.3 对学生学习行为的模型影响维度分析聚类分析

聚类分析作为一种重要的多元统计分析方法,致力于探究指标间的分类规律,旨在揭示学生在学习行为上的共性特征。通过该方法,我们能够更为深入地洞察并剖析不同群体所独有的特性。

采用K-means聚类算法精细划分学生群体,深入探究其学习行为背后的动因机制。聚类分析成果揭示了影响职业院校学生学习历程的四大核心要素:专注力维度、绩效驱动力维度、互动参与度维度及毅力持久性维度。其中,专注维度表现在:学员在单项任务上的完成情况,以及其能否维持较长时间的学习投入、发帖内容的详尽程度与平均登录时长等。在绩效投入维度,具体表现为完成的自测题与作业数量、自测成绩优劣,以及对拓展资源的访问频率等方面。至于互动层面,则通过发帖贡献比例、帖子阅读量及回复数量等多项指标进行综合评估。至于坚持性,则

通过累积的学习节点积分值进行直观呈现。

利用Python软件卓越的可视化能力,学生们可被明确界定为四大类别群体。

第一类学生往往表现出较长的在线学习时长,且其学习内容呈现出高度的连贯性与系统性特征。

第二类学生虽在练习与测试等环节展现出极高的积极性与热忱,但在面对在线课程的学习投入方面,却显现出一定的不足。

第三类学生对师生互动环节关注点高。

第四类学生网络学习的参与度与自主性较差。

## 3 结果建议

①对专注维度低的学生,可以调整上传资料的趣味性、与内容的新颖型,提升学生学习兴趣。②对绩效努力维度低的学生,可以精简自测题与作业题目的数量,少而精,提升学生参与积极性。③对交互维度低的学生,可以增大教师与学生互动频率,实时关注,提升学生的认同感。④对于坚持维度低的学生,可以增大学习节点的积分,增加平时成绩比重。

针对各异学生群体的行为模式展开深度剖析,仍有部分学生显现出兴趣缺失之态,这迫切要求我们为其打造更高品质的沉浸式学习体验,并赋予其更为鲜明的学习成就感。需聚焦学生的学习行为特质,核心在于审视其学习历程,但不可忽视的是,学生的学习习性培育及学业成就的获取,往往深受个性化要素的影响,包括学习态度、吸收能力、情绪波动等。将这些个性化因素纳入综合考量,无疑将促进对学生群体分析的深化与精确化。

教育承载着培育英才的直接使命,其形态丰富且日新月异。在精准掌握学生学习状态的基础上,巧妙运用智能学习平台所具备的精细诊断与分析功能,辅助教师深化对学生的认知理解,进而实施针对性的教学辅助策略,稳健推进个性化教育的理想实践。

## 参考文献:

- [1] 顾小清,李世瑾.人工智能教育大脑:以数据驱动教育治理与教学创新的技术框架[J].中国电化教育,2021,(01):80-88.
- [2] 顾小清.教育信息化步入数字化转型时代[J].中小学信息技术教育,2022,(04):5-9.
- [3] 顾小清,王超.打开技术创新课堂教学的新窗:刻画AIoT课堂应用场景[J].现代远程教育研究,2021,33(02):3-12.