

大数据分析在高职学生学习行为诊断与教学策略优化中的应用

黄雅茹

(南京大学, 江苏 南京 210093)

摘要: 随着职业教育数字化转型的不断推进, 大数据技术为高职学生学习行为诊断与教学策略优化提供了全新路径。高职教育以培养技术技能人才为核心, 学生学习行为具有实践性、职业性、差异性等特点, 传统教学模式难以精准捕捉学生学习过程中的问题与需求。本文基于大数据分析技术, 系统探讨高职学生学习行为数据的收集、预处理流程, 构建科学的学习行为诊断指标体系与分析模型, 精准识别学生学习异常行为与个性化需求, 进而提出针对性的教学策略优化方案, 为高职教育教学质量提升、个性化教学实施提供理论支撑与实践参考, 助力实现高职教育内涵式发展。

关键词: 大数据分析; 高职学生; 学习行为诊断; 教学策略优化; 数据预处理

中图分类号: G712

文献标识码: A

文章编号: 3106-2709 (2025) 11-0005-05

DOI: 10.62022/NCAR.issn3106-2709.2025.11.002

The Application of Big Data Analysis in Diagnosing Learning Behavior and Optimizing Teaching Strategies of Vocational College Students

Huang Yaru

(Nanjing University, Nanjing, Jiangsu 210093)

Abstract: With the continuous promotion of digital transformation in vocational education, big data technology provides a new path for diagnosing learning behavior and optimizing teaching strategies for vocational college students. Vocational education focuses on cultivating technical and skilled talents, and students' learning behavior is characterized by practicality, professionalism, and differentiation. Traditional teaching models are difficult to accurately capture the problems and needs of students in the learning process. This article is based on big data analysis technology, systematically exploring the collection and preprocessing process of learning behavior data of vocational college students, constructing a scientific learning behavior diagnosis index system and analysis model, accurately identifying abnormal learning behaviors and personalized needs of students, and proposing targeted teaching strategy optimization plans, providing theoretical support and practical reference for improving the quality of vocational education and implementing personalized teaching, and helping to achieve connotative development of vocational education.

Keywords: big data analysis; Vocational college students; Diagnosis of learning behavior; Optimization of teaching strategies; data preprocessing

一、高职学生学习行为数据的收集与预处理

高职学生学习行为数据是开展诊断与教学优化的基础, 其收集的全面性、准确性与预处理的规范性, 直接决定后续分析结果的可靠性。结合高职教育“理论+实践”的教学特点, 需构建多维度数据收集体系, 并通过科学的预处理技术, 剔除无效数据、整合有效信息, 为后续分析奠定基础。

(一) 学习行为数据收集范围与类型

高职学生学习行为数据的收集需覆盖学习全流程, 结合高职教学场景(课堂教学、实践实训、线上学习、课后拓展),

分为四大类核心数据, 确保数据的全面性与针对性。

一是课堂学习行为数据, 主要包括课堂出勤情况(迟到、早退、旷课次数及时长)、课堂互动表现(举手发言、小组讨论参与度、课堂测验成绩、课堂专注度)、课堂笔记完成质量等, 聚焦学生在理论课堂中的学习状态; 二是实践实训行为数据, 作为高职教育的核心环节, 此类数据包括实训项目参与率、实训操作规范性、实训任务完成进度、实训考核成绩、实训过程中的错误操作记录等, 反映学生技术技能掌握情况; 三是线上学习行为数据, 依托线上教学平台(如超星泛雅、学习通、职教云), 收集学生线上课程观看时长、视频暂停/回放次数、线上作业完成情况

况(提交时间、正确率、修改次数)、线上讨论发言内容、线上测试成绩、资源下载次数等,捕捉学生自主学习行为特征;四是课后拓展行为数据,包括课后作业完成质量、课外阅读资源浏览记录、职业技能竞赛参与情况、考证学习进度、校企合作实践表现等,体现学生的自主学习意愿与职业发展潜力。

(二) 数据收集方法与工具

结合高职教学场景的多样性,采用“线上+线下”融合的收集方法,搭配多元化工具,确保数据收集的高效性与准确性,避免人工收集的主观性与误差。

线上数据收集主要依托现有教学平台与数字化工具,实现自动化采集:通过线上教学平台(职教云、学习通)自动记录学生线上学习行为数据,无需人工干预;借助智慧校园系统,采集学生出勤、校园消费、图书馆借阅等关联数据,间接反映学生学习状态;通过职业技能实训平台,实时捕捉学生实训操作数据,包括操作步骤、完成时间、错误类型等。线下数据收集采用人工辅助与智能设备结合的方式:课堂出勤数据可通过智慧考勤设备(人脸识别、刷卡考勤)自动采集,辅以教师人工核对;课堂互动与专注度数据可通过课堂观察记录表、智能课堂分析系统(如智慧黑板配套分析工具)采集;实践实训数据可通过实训指导教师记录、实训设备传感器数据采集,确保实训行为数据的精准性;课后拓展数据可通过学生自主申报、教师统计、校企合作单位反馈等方式收集,补充线上数据的不足。

此外,需建立数据收集规范,明确各数据的收集频率、责任主体与存储格式,确保数据的连续性与统一性,为后续预处理与分析提供便利。

(三) 数据预处理流程与关键技术

收集到的原始学习行为数据存在大量冗余、缺失、异常等问题(如学生误操作产生的无效数据、设备故障导致的缺失数据、极端异常的成绩数据),需通过科学的预处理流程,将原始数据转化为干净、规范、可用的分析数据,预处理流程主要包括数据清洗、数据集成、数据转换、数据规约四个环节。

数据清洗是预处理的核心环节,主要采用缺失值处理、异常值处理、冗余数据删除三种关键技术:针对缺失数据,根据数据类型采用不同处理方式,如数值型数据(成绩、时长)采用均值、中位数填充,分类数据(出勤状态、实训等级)采用众数填充,对于关键缺失数据(如核心课程

成绩),结合学生其他学习行为数据进行合理推断;针对异常数据(如极端高分/低分、超出合理范围的学习时长),采用 3σ 原则、箱线图法识别,结合实际教学场景判断是否保留,对于误操作产生的异常数据予以删除;针对冗余数据(如重复记录、无关数据),通过数据比对、去重算法,剔除重复信息,删除与学习行为无关的数据(如学生个人无关信息)。

数据集成是将来自不同平台、不同类型的学习行为数据(如线上教学平台数据、实训平台数据、智慧校园数据)整合到统一的数据仓库中,解决数据异构问题,通过数据关联技术(如学号关联),实现同一学生多维度数据的对应匹配,确保数据的完整性。数据转换是将集成后的数据转换为适合分析的格式,如将定性数据(课堂表现等级)量化为数值型数据,将时间数据(学习时长)标准化,通过归一化、标准化处理,消除不同数据维度的量纲影响,便于后续分析模型的应用。数据规约是在保证数据核心信息不丢失的前提下,通过降维技术(如主成分分析PCA)、数据抽样等方式,减少数据量,提高分析效率,避免数据冗余导致的分析偏差。

二、基于大数据分析的高职学生学习行为诊断

基于预处理后的学习行为数据,构建科学的诊断指标体系,选择合适的大数据分析模型,对学生学习行为进行全面诊断,精准识别学生学习特征、存在的问题及异常行为,为教学策略优化提供精准依据。

(一) 学习行为诊断指标体系构建

结合高职学生学习行为特点与教学目标,遵循科学性、系统性、可操作性、针对性原则,构建“4个一级指标、12个二级指标、30个三级指标”的学习行为诊断指标体系,覆盖学习全流程,确保诊断的全面性与精准性。

一级指标包括课堂学习行为、实践实训行为、线上自主学习行为、课后拓展行为,与数据收集类型一一对应,确保指标体系与数据来源的一致性。二级指标在一级指标基础上细化,如课堂学习行为包含出勤情况、课堂互动、课堂成绩3个二级指标;实践实训行为包含实训参与、实训操作、实训成绩3个二级指标;线上自主学习行为包含线上学习时长、作业完成、线上测试3个二级指标;课后拓展行为包含课后作业、职业拓展、自主学习3个二级指标。三级指标为具体可量化的指标,如出勤情况包含旷课次数、迟

到次数、早退次数3个三级指标；实训操作包含操作规范度、错误操作次数、实训完成效率3个三级指标；线上学习时长包含课程视频观看时长、资源浏览时长、线上讨论时长3个三级指标。

同时，采用层次分析法（AHP）确定各指标的权重，结合高职教育“重实践、强技能”的特点，适当提高实践实训行为指标的权重，确保诊断结果贴合高职教学核心目标；通过标准化处理，将各指标转化为统一量级，避免权重分配不合理导致的诊断偏差。

（二）大数据分析模型的选择与应用

结合高职学生学习行为数据的特点（多维度、海量性、关联性），选择适合的大数据分析模型，实现对学习行为的精准诊断，主要采用描述性分析、预测性分析、聚类分析三种核心模型，三者结合形成完整的诊断体系。

描述性分析模型主要用于梳理学生学习行为的整体特征，通过统计分析（均值、中位数、标准差、频率分布），呈现学生在各指标上的表现，如整体出勤率、线上学习时长分布、实训成绩等级分布等，明确学生学习行为的整体现状与共性问题，为后续诊断提供基础。例如，通过描述性分析发现，某班级学生线上课程平均观看时长不足规定时长的60%，说明该班级学生自主学习积极性不足，需重点关注。

预测性分析模型用于预测学生学习效果与发展趋势，结合预处理后的多维度数据，选择逻辑回归、随机森林等算法，构建学习效果预测模型，通过学生前期学习行为数据（如课堂表现、线上学习情况、实训成绩），预测学生后续课程成绩、技能掌握程度，提前识别学习困难学生，为针对性干预提供依据。例如，通过随机森林算法，结合学生前半学期的学习行为数据，预测学生期末实训成绩，对预测成绩较低的学生，提前制定帮扶计划。

聚类分析模型用于识别学生学习行为的差异，采用K-means聚类算法，根据学生在各诊断指标上的表现，将学生划分为不同的学习类型，如自主型、被动型、实践型、理论薄弱型等，明确不同类型学生的学习特征与需求。例如，通过聚类分析，将学生分为“高参与高成绩型”“低参与低成绩型”“高参与低成绩型”“低参与高成绩型”四类，针对不同类型学生制定差异化教学策略。

此外，结合异常检测算法（如孤立森林算法），识别学生学习异常行为，如长期旷课、线上学习数据异常、实

训操作频繁出错等，及时发现学生学习过程中的问题，避免学习问题积累。

（三）学习行为诊断结果分析（含异常行为识别）

基于上述指标体系与分析模型，对高职学生学习行为进行全面诊断，从整体特征、个体差异、异常行为三个维度进行分析，形成完整的诊断报告，为教学策略优化提供精准支撑。

整体特征分析主要呈现学生学习行为的共性规律与问题：从课堂学习来看，部分高职学生出勤情况较差，迟到、旷课现象较为普遍，课堂互动积极性不足，主动发言、参与小组讨论的学生占比偏低；从实践实训来看，学生实训操作规范性参差不齐，部分学生存在畏难情绪，实训任务完成效率不高，核心技能掌握不扎实；从线上学习来看，学生线上学习主动性差异较大，部分学生仅完成最低要求，线上作业敷衍了事，资源利用效率不高；从课后拓展来看，学生自主学习意识薄弱，参与职业技能竞赛、考证学习的学生占比不高，课后作业完成质量有待提升。

个体差异分析基于聚类分析结果，明确不同学习类型学生的特征：自主型学生（约占20%），课堂出勤良好，线上线下学习积极，实训操作规范，成绩优异，具备较强的自主学习能力与职业素养；被动型学生（约占40%），学习积极性不高，课堂参与度低，线上学习多为完成任务，实训操作被动，成绩中等，需教师引导与督促；实践型学生（约占25%），理论学习薄弱，课堂表现一般，但实训操作能力较强，动手能力突出，需加强理论知识补充；理论薄弱型学生（约占15%），理论基础差，实训操作也存在困难，学习态度不端正，存在较多学习问题，需重点帮扶。

异常行为识别主要通过异常检测算法，结合人工核对，识别两类异常行为：一是学习状态异常，如长期旷课（连续3次及以上）、线上学习时长为0、实训任务多次未完成，此类学生多存在学习态度不端正、厌学等问题；二是学习行为异常，如线上作业正确率骤降、实训操作错误次数异常增多、课堂专注度突然下降，此类学生可能存在学习困难、心理压力等问题。针对识别出的异常行为，建立预警机制，及时反馈给班主任与任课教师，开展针对性干预，避免学生学习状况进一步恶化。

三、基于学习行为诊断的教学策略优化

结合学习行为诊断结果，遵循“以学生为中心、因材

施教、知行合一”的原则，明确教学策略优化的目标，针对不同学习行为特征的学生，设计差异化教学策略，制定科学的实施路径，确保教学策略落地见效，提升高职教育教学质量。

（一）教学策略优化的原则与目标

教学策略优化需遵循三大核心原则：一是因材施教原则，结合学生学习行为差异，针对不同学习类型、不同学习水平的学生，制定差异化教学方案，避免“一刀切”教学；二是知行合一原则，贴合高职教育培养技术技能人才的目标，强化实践教学，将理论教学与实践实训深度融合，提升学生动手能力；三是动态调整原则，基于大数据分析，实时跟踪学生学习行为变化，及时调整教学策略，确保教学策略的针对性与有效性。

教学策略优化的核心目标分为三个层面：一是整体目标，提升学生学习积极性与主动性，改善学生学习行为，提高教学质量与人才培养质量，助力学生掌握核心职业技能；二是个体目标，针对不同学习类型学生，解决其学习过程中的问题，如帮助被动型学生提升学习积极性，帮助理论薄弱型学生夯实理论基础，帮助实践型学生补充理论知识，促进每个学生的个性化发展；三是长效目标，构建“数据收集—行为诊断—策略优化—效果反馈”的闭环教学模式，实现教学策略的动态优化，推动高职教育数字化、个性化发展。

（二）针对不同学习行为特征的教学策略设计

基于学习行为诊断得出的四种学生类型，结合各类型学生的学习特征与需求，设计差异化的教学策略，确保教学策略的针对性与实效性。

针对自主型学生，采用“引导提升型”教学策略：减少基础内容的讲解，重点讲解拓展性、综合性知识，布置具有挑战性的学习任务（如职业技能创新项目、复杂实训任务）；搭建自主学习平台，提供丰富的拓展资源（如行业前沿技术资料、考证辅导资源），鼓励学生参与职业技能竞赛、校企合作项目，培养其创新能力与职业素养；建立激励机制，对表现优异的学生给予表彰，引导其发挥示范引领作用，带动班级整体学习氛围。

针对被动型学生，采用“引导督促型”教学策略：加强课堂管理，严格考勤制度，通过课堂提问、小组讨论等方式，强迫学生参与课堂互动，提升其课堂专注度；线上学习采用“分层任务”模式，将学习任务拆解为基础任务、

提升任务，要求学生逐一完成，教师实时跟踪学习进度，对未完成任务的学生进行提醒与督促；建立一对一引导机制，由任课教师或学习委员对其进行针对性指导，帮助其树立学习信心，培养自主学习习惯。

针对实践型学生，采用“理论补充型”教学策略：优化课程设置，增加理论知识与实践操作的结合点，将理论知识融入实训教学中，通过案例教学、情境教学等方式，帮助其理解理论知识；针对其理论薄弱的环节，开展专项辅导，布置针对性的理论作业与练习，夯实其理论基础；鼓励其将实践经验转化为理论认知，参与实践心得分享、案例编写等活动，提升理论表达能力。

针对理论薄弱型学生，采用“重点帮扶型”教学策略：建立专项帮扶小组，由教师与成绩优异的学生组成帮扶团队，开展一对一辅导，重点讲解基础理论知识与核心技能要点；优化教学内容，简化复杂理论知识，采用通俗易懂的语言、直观的教学手段（如动画、视频、实物演示），降低学习难度；调整考核方式，增加过程性考核权重，重点考核学生的学习态度与进步幅度，激发其学习动力；针对其异常学习行为，及时沟通了解原因，解决其学习与生活中的困难，帮助其端正学习态度。

此外，针对识别出的异常行为学生，制定个性化干预策略：对厌学、旷课的学生，加强心理疏导与思想教育，了解其厌学原因，制定个性化学习计划；对学习困难的学生，开展专项辅导，帮助其解决学习中的具体问题；对学习行为突然异常的学生，及时沟通，排查心理压力、家庭问题等因素，提供针对性支持。

（三）教学策略优化的实施路径

为确保教学策略优化落地见效，结合高职教学实际，制定“四个到位”的实施路径，构建闭环教学模式，实现教学策略的动态优化与长效运行。

一是师资培训到位，提升教师大数据应用能力。开展大数据分析、数字化教学相关培训，帮助教师掌握数据收集、分析、应用的基本方法，提升教师利用大数据诊断学生学习行为、设计个性化教学策略的能力；鼓励教师参与数字化教学改革项目，探索大数据与教学融合的新模式、新方法，推动教师教学理念与教学能力的提升。

二是教学资源优化到位，支撑个性化教学实施。整合线上线下教学资源，搭建一体化教学平台，根据不同学习类型学生的需求，上传分层教学资源（基础资源、拓展资

源、专项辅导资源），方便学生自主学习；优化实训教学资源，更新实训设备，增加实训项目的针对性与实用性，贴合行业岗位需求，提升学生实践操作能力；建立资源更新机制，及时补充行业前沿技术、岗位最新需求相关的教学资源，确保教学内容的时效性。

三是过程监控到位，确保教学策略落地。依托大数据技术，实时跟踪学生学习行为变化，监测教学策略的实施效果，如学生出勤情况、线上学习进度、作业完成质量、成绩变化等；建立定期反馈机制，每月收集学生、教师对教学策略的意见与建议，结合诊断结果，及时调整教学策略，解决实施过程中存在的问题；班主任与任课教师密切配合，加强对学生的日常管理与引导，确保教学策略落实到每一位学生。

四是考核评价优化到位，强化激励导向。优化考核评价体系，改变“一考定终身”的考核方式，增加过程性考核权重（占比不低于50%），将学生学习行为、课堂参与、实训表现、线上学习情况等纳入考核范围，全面评价学生的学习过程与学习效果；建立多元化评价主体，结合教师评价、学生自评、小组互评、校企评价等方式，确保评价结果的客观性与全面性；完善激励机制，对学习进步明显、表现优异的学生给予表彰与奖励，对教师的教学改革成果给予肯定与支持，激发学生学习与教师教学改革的积极性。

四、结束语

大数据分析技术为高职学生学习行为诊断与教学策略优化提供了全新的技术支撑，打破了传统教学模式的局限，实现了从“经验教学”向“数据驱动教学”的转变。本文通过系统探讨高职学生学习行为数据的收集与预处理、学习行为诊断指标体系与分析模型的构建、基于诊断结果的教学策略优化，构建了“数据收集—行为诊断—策略优化—效果反馈”的闭环教学模式，为高职教育教学质量提升提供了理论支撑与实践参考。在实际应用中，大数据分析在高职学生学习行为诊断与教学策略优化中的应用仍存在一些不足，如部分高职院校数字化建设水平不高、教师大数据应用能力不足、数据安全保障体系不完善等。未来，需进一步加强高职教育数字化建设，加大师资培训力度，完善数据安全保障机制，深化大数据与教学的深度融合，不断优化学习行为诊断模型与教学策略，充分发挥大数据技术的优势，推动高职教育内涵式发展，培养更多适应行业岗位需求的高素质技术技能人才。

参考文献：

- [1] 鞠凤琪，吕佳璇，张君鑫.探讨自主学习能力培养的高职学生学习评价体系[J].教育教学研究前沿，2025，3（10）：25-27.
- [2] 巫昭海.借助积极心理学激发高职学生学习动机有效策略探究[J].国家通用语言文字教学与研究，2022，（12）：104-106.