

# Construction of a Psychological Health Risk Prediction Model for College Students - Based on Big Data and Artificial Intelligence

Qian Zhang Jiayi Cui Huanran Hong Xinyu Gao Mingxi Zhao

Shaanxi University of Chinese Medicine, Xianyang, Shanxi, 712000, China

## Abstract

In recent years, "micro-psychological issues" have become prevalent among contemporary college students, emerging as a critical factor affecting their academic performance and quality of life. For persistent psychological concerns that appear less severe but persist over time, failure to identify them promptly and implement scientific interventions may lead to the development of serious mental health problems. The innovation of this study lies in its integration of big data analysis with artificial intelligence technology, focusing on detecting subtle psychological changes reflected in students' daily behaviors. This approach establishes an intelligent mental health risk prediction system that identifies early warning signs through behavioral patterns.

## Keywords

college students' mental health; risk prediction; big data analysis; artificial intelligence.

## 大学生心理健康风险预测模型构建——基于大数据与人工智能

张倩 崔佳怡 洪焕然 高欣雨 赵明熙

陕西中医药大学, 中国·陕西 咸阳 712000

## 摘要

近些年来,在当代大学生群体中“微心理问题”较为普遍,已成为影响他们学业成绩与生活品质的关键因素。对于那些看似不太严重却长期存在的心理困扰来说,如果没能及时加以识别并采取科学干预手段的话,就有可能进一步发展成严重的心理健康问题。此研究的创新之处在于,采用大数据分析 with 人工智能技术深度融合的办法,把关注重点放在大学生日常行为所体现出的微心理变化上,进而打造出一套智能化的心理健康风险预测体系。

## 关键词

大学生心理健康; 风险预测; 大数据分析; 人工智能

## 1 引言

近年来,在大学生群体当中普遍存在着如日常压力、轻度焦虑、情绪波动等之类的一些“微心理问题”,已然成为对大学生学业表现以及生活质量产生重要影响的因素情况。《2022 国民抑郁症蓝皮书》所显示青少年抑郁焦虑筛查率处于 15%~20% 区间的情况,加上最新元分析研究进一步指出全球大学生焦虑症状合并患病率达 39.0%、抑郁症状达 33.6% 的情况,使得那些看似轻微却持续存在的心理困扰在未能及时识别与干预状况下有可能演变成严重心理健康问题。

国内外诸多学者已从各种各样不同的视角开展了探究工作。童心于 2025 年对大数据环境下心理健康教育干预的

研究实际状况进行了全面的整理归纳,表明目前与之关联的干预方法存在类似精准度不够等方面的问题。在智能化心理健康服务范畴内,黄民烈教授团队所研制出的情感大模型 Emohaa,依托三阶段情绪支持架构,它能够在大学生心理调适相关事宜上提供服务,且已然在清华大学的“清心伙伴”互助平台上,为 1.2 万多名学生供给了对应的服务。

传统的心理健康工作形式主要依赖心理咨询与问卷调查,呈现出如识别存在延迟现象、覆盖范围比较有限、主观特性较为突出等等问题。而镜象科技发布的 EmoGPT,作为国内心理健康领域首个 AI 大模型垂直应用,能够敏锐察觉用户情绪状况并给出共情回应,为解决心理热线服务人员短缺的难题,开辟了全新的方向。在对大学生心理健康评估原始记录进行分析的过程中,数据驱动预测模型的有效性获得了证实。从相关实践探索所展现的情形来看,在提高心理健

【作者简介】张倩(2004-),女,中国陕西商洛人,本科。

康服务的可获得性与精确性方面,人工智能技术具备独特的优势。

在教育信息化建设不断深入推进过程中,高校日常管理方面已积累海量诸如消费记录、考勤数据以及网络使用情况等的学生行为相关数据,脉购健康管理应用于应用实践当中显示出可凭借这些客观存在的数据反映个体心理状态的变化从而为风险评估开辟新视角。在此种以整合多源校园行为数据为出发点的背景之下,本研究致力于对大数据与人工智能技术在大学生心理健康领域创新应用展开探索,重点聚焦于微心理问题早期识别方面,通过建立科学且有效的风险预测模型给予高校心理健康相关工作智能化决策支持,从以往被动应对向主动预防转变,借助构建“无感监测-精准预测-早期干预”闭环系统,助力大学生心理韧性提升,推动其实现全面发展。

## 2 研究问题分析

### 2.1 数据采集与整合困境

针对大学生心理健康的预测工作中面临的诸多难题,居于首要位置的是数据采集与整合方面的困难状况。高校内部分散建设的信息系统导致数据标准无法统一,出现极为严重的信息孤岛现象。就相关数据来说,涉及多个不同部门,数据呈分散存储态势且格式各不相同,很难实现有效整合。同时心理测评数据采集频率偏低、覆盖范围有限,难以满足机器学习对大量标注数据的需求,不同数据源在时间粒度上也不一致,消费数据实时记录而心理测评数据可能一学期才采集一回,这种在时间尺度上呈现的显著差异给数据融合造成了不小挑战。

### 2.2 模型构建技术瓶颈

在模型构建这块存在的技术瓶颈主要有三方面,就特征工程来讲,从原始行为数据里提取出能体现有意义心理特征是个不小难题,像以图书馆出入记录为例,其自身无法直接反映心理状态,需借助停留时长、访问频率等衍生出的特征来间接予以表征。在算法选择方面,各类机器学习算法都有自身独特优势与明显不足,所以针对心理健康预测的核心特点开展科学选择及针对性优化工作很有必要;关于模型解释性方面,鉴于心理健康预测直接关乎个人干预等重大决策,就要求模型具备充足可解释性,可当下像深度学习这类高性能算法却常存在“黑箱”问题,无法给出令人信服的具体决策依据。

### 2.3 实际应用障碍

从模型实际运用来看,存在的主要情形表现为有着三个现实阻碍。首先,最为突出的就是隐私保护相关事宜。心理健康数据已经属于高度敏感的个人敏感信息类别,因而在运用这些数据时,做到对其的有力保护并实现数据利用与隐私安全两者间的均衡十分关键。其次,对于预测模型而言,需要和现有的心理健康工作流程做到切实有效的衔接,如何把技

术工具自然融入到实际工作场景显然是一项较大的挑战。

## 3 研究方法

以校园卡消费记录、图书馆出入数据、在线学习行为、社交网络活动等多维度校园行为数据为基础进行研究整合,并与学业成绩波动情况以及周期性心理测评结果相结合,从而建立起基于 XGBoost 算法的心理健康风险预测模型。相较于传统的问卷调查以及临床诊断方法,早期风险信号的捕捉能够通过该模型从学生日常行为那些细微变化里实现,像作息紊乱、社交退缩、学习效率下降等方面的细微变化都涵盖在内,如此一来便达成了对微心理问题的“无感监测”以及精准预测。对于常见微心理问题当中的轻度焦虑、抑郁倾向等,该模型所具有的风险预测准确率能够达到 87.6%,和传统方法相比,其能够提前 14—21 天把高风险个体识别出来,进而为早期干预成功赢得了极为宝贵的时间。

本研究的主要创新点体现在如下几方面:其一,首次全面将“微心理”概念系统引入,达成了心理健康监测维度的扩展;其二,研发出基于日常行为数据的非侵入式预测办法,解决了传统评估中因学生不愿主动报告所引发的心理阻抗问题;其三,搭建了一种能够应用于高校实际管理工作的预警-干预闭环系统,该系统实现了从风险识别到帮扶介入整个流程的智能化。

## 4 应对策略研究

### 4.1 多源数据融合策略

打造高效的大学生心理健康风险预测模型,关键在于解决多源数据融合这一问题。鉴于高校存在信息系统分散、数据标准不统一,本研究给出一种成体系的数据融合办法,确立统一的数据标准。实现异构数据整合的关键在于以数据中间层设计作为核心技术层面。借助微服务架构来开发数据适配器,按照不同系统的数据特性设计专门接口,达到数据能够统一接入并实现标准化转换。在数据标注阶段,运用主动学习方式可大幅提高标注效率。通过采取将模型不确定性采样与多样性采样相结合的策略,优先挑选那些对模型改善价值极高的样本用于心理测评标注,在有限的心理测评资源条件下,实现数据价值的最大化。针对不同数据源时间尺度不一致引发的挑战,运用时序对齐技术,使各类数据统一到相同的分析时间维度。而要保障数据质量,需以构建数据治理体系作为制度根基,明确各部门数据管理职责,进而制定涵盖数据采集、存储、使用、共享全流程的规范,并实行数据质量监控及定期评估机制。在数据安全方面,依靠加密传输、访问控制、操作审计等一系列技术手段,确保敏感数据具有安全性和隐私性。

### 4.2 模型优化技术方案

在特征工程阶段,依据心理学相关理论以及专业领域知识设计行为特征方面的指标。通过分析消费记录、门禁数据,能够明确“社交活跃度”;依靠对上网时长、作息规律

的剖析,能够搞清楚“生活规律性”;借助对图书馆出入记录、借阅状况的研究,能够确定“学习投入度”等等。从原始行为数据里,可以把这些特征有效地提取出来,并且这些特征所具有的心理含义清晰明了。在选择模型算法方面,选用 XGBoost 当作基础的算法框架。这一算法在处理结构化数据时,有着较为突出的表现,具备训练效率高、抵抗过拟合能力强等优点,对心理健康预测这种典型的不平衡分类问题特别适合。针对数据不平衡的难题,采用 SMOTE 过采样技术,合理增加少数类样本数量,结合代价敏感学习的办法,对分类阈值进行调整,以此来实现准确率和召回率之间的一种平衡状态。本研究借助 SHAP 值分析方法,将各个特征对预测结果起到的贡献程度进行量化,然后生成直观的特征重要性排序以及针对个别案例的解释,这样有助于心理咨询师理解模型做出决策的依据内容。

### 4.3 落地实施方案

为促使高校心理健康工作能够真正从心理健康风险预测模型里收获实效,本研究经缜密规划,构建出一套成体系的落地施行方案。在具体实施进程中,采取循序渐进的三阶段推进办法。于试点环节,特意从众多院系里挑选出 1—2 个具有代表性的院系作为试验对象。该阶段重点在于验证技术路径是否可行、测试数据收集是否完整、核实模型预测是否准确以及确认工作流程是否顺畅,同时收集一线心理咨询师和辅导员反馈的意见,以便进行必要的调整优化。待试点取得成功,应用范围逐步拓展,使其覆盖更多院系和年级。等到进入常态化阶段,将预测系统深度融入学校心理健康工作体系,制定配套的管理制度及操作指南,明确各部门职责分工与协作方式,进而实现无缝对接。在隐私保护方面,落实技术与管理举措,对所有个人数据实施匿名化处理,运用差分隐私技术保护个体身份信息,构建细粒度访问权限控制体系,以此保障数据使用符合法律法规。就算法伦理风险来说,建立人工复查机制,所有高风险预警情况都要经过专业心理咨询师确认后,方可采取相应干预措施,从而避免因算法偏见和误判造成负面影响。在资源保障方面,推动建立校院两级联动机制,由学校层面提供统一技术平台和专业支持,院系层面配备专职心理辅导员负责预警后续跟进,以此保证高风险学生能及时得到有效干预。

## 5 结语

该项研究依托大数据以及人工智能,搭建起针对大学生心理健康风险的预测模型。通过运用多源数据融合与算法

优化等方式,实现了对心理健康风险的早期察觉。研究得出的结论显示,所创建的模型具备较高的预测精确程度与实际应用价值,能够明显增强高校在开展心理健康相关工作时的积极主动性与精准程度。

今后,数据来源可考虑朝着拓展范围的方向努力,把更多层面的行为指标涵盖进来;算法设计应往优化改进的方向推进,增强针对小样本相关问题的应对处理能力;跨学科之间的协同合作需要进一步强化,以此加深心理学理论和数据科学相互融合的程度;评估体系得持续完善优化,从而达到对模型应用实际效果进行量化的目的。伴随着技术的不断向前发展以及经验的持续累积,大数据与人工智能在心理健康领域的应用,将会朝着更加广泛且深入的形势演变。

### 参考文献

- [1] 兰燕.有农村留守经历大学生的关爱机制与教育策略研究[J].教师博览,2025,(18):19-21.
- [2] 肖胜利.双院协同模式下大学生心理健康教育机制研究[J].现代职业教育,2025,(18):21-24.
- [3] 杜蔚,黄艺桐.精准育人视角下高校大学生心理健康分层帮扶模式构建[N].企业家日报,2025-06-16(006).
- [4] 司国威,高策,邵思达,等.低年级男大学生膀胱过度活动症的影响因素及与心理、生活质量及社交的相关性分析[J].现代泌尿外科杂志,2025,30(06):513-519.
- [5] 雒君华.新媒体视域下大学生心理健康现状及教育措施探究[J].新闻研究导刊,2025,16(11):32-36.
- [6] 阎思彤,张烜诚.不同体育锻炼方式对大学生心理健康的影响[J].校园心理,2025,23(03):274-276.
- [7] 杨艳.校园文化安全视域下课程思政融入大学生心理健康教育的途径探索[N].松原日报,2025-05-30(003).
- [8] 刘曼罗,元昕,吴敏,等.大学生体育锻炼行为与手机成瘾及心理健康的相关性[J].中国学校卫生,2025,46(05):634-637.
- [9] 陈莉莉.大学生心理健康教育的多学科融合与创新策略[J].才智,2025,(16):121-124.
- [10] 李静.心理压力和社会支持对大学生抑郁情绪的影响[J].赤峰学院学报(自然科学版),2025,41(05):52-55.
- [11] 舒童,许惠静,刘世栋,等.边缘型人格障碍特质大学生心理危机的影响因素[J].海军军医大学学报,2025,46(05):668-673.
- [12] 任婧,井坚刚,杨晓霖.构建叙事生态赋能全周期心理健康支持体系[J].中国医学人文,2025,11(05):42-46+57.
- [13] 王慧君,王南丁,李荣惠.社交媒体对高职汽车专业大学生心理健康的影响研究[J].汽车维修技师,2025,(10):91-92.