

基于数据分析的久坐人群健康问题调查研究

杜梦瑶¹, 吴 燕¹, 李毓洁¹, 高仕龙^{1,2}

¹乐山师范学院数学与统计学院, 四川 乐山

²乐山师范学院应用数学研究中心, 四川 乐山

收稿日期: 2025年12月21日; 录用日期: 2026年1月12日; 发布日期: 2026年1月23日

摘 要

久坐等不良生活方式引发的健康问题日益严峻, 全球每年约160万人因此死亡。本文基于对久坐人群的健康调查数据, 系统分析久坐行为与健康指标间的关联, 并构建融合K-means聚类、分段函数建模、随机森林与梯度提升树的集成模型, 实现人群健康状态分型、风险阈值判定与亚健康预测。研究结果为构建面向久坐人群的主动健康干预体系提供了方法与数据支撑, 推动健康管理模式从“被动医疗”向“主动预防”转型, 对提升公众健康水平具有积极意义。

关键词

K-Means聚类分析, 分段函数模型, 随机森林模型, 梯度提升树, 久坐行为, 健康风险评估

A Survey on Health Issues of Sedentary Populations Based on Data Analysis

Mengyao Du¹, Yan Wu¹, Yujie Li¹, Shilong Gao^{1,2}

¹College of Mathematics and Statistics, Leshan Normal University, Leshan Sichuan

²Research Center of Applied Mathematics, Leshan Normal University, Leshan Sichuan

Received: December 21, 2025; accepted: January 12, 2026; published: January 23, 2026

Abstract

Health issues stemming from unhealthy lifestyles such as prolonged sitting are becoming increasingly severe, causing approximately 1.6 million deaths globally each year. Based on health survey data from sedentary populations, this study systematically analyzes the relationship between sedentary behavior and health indicators, and constructs an integrated model combining K-means clustering,

文章引用: 杜梦瑶, 吴燕, 李毓洁, 高仕龙. 基于数据分析的久坐人群健康问题调查研究[J]. 统计学与应用, 2026, 15(1): 255-264. DOI: 10.12677/sa.2026.151024

piecewise function modeling, Random Forest, and Extreme Gradient Boosting to achieve health status classification, risk threshold identification, and sub-health prediction. The research outcomes provide methodological and data support for establishing an active health intervention system for sedentary populations, promoting a transition in health management from “passive treatment” to “active prevention” and contributing to the improvement of public health.

Keywords

K-Means Clustering Analysis, Piecewise Function Modeling, Random Forest Modeling, Extreme Gradient Boosting, Sedentary Behavior, Health Risk Assessment

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

2016 年,《“健康中国 2030”规划纲要》明确提出要促进全民健康生活方式[1],其中将“久坐行为”纳入慢性病防控重点。随着城市化与数字化进程加速,久坐已成为全球公共卫生领域的核心挑战之一。世界卫生组织(WHO)数据显示,全球每年因久坐相关健康问题导致的死亡人数约 160 万,我国超 60% 职场人群与学生群体日均久坐时长超过 8 小时[2],肩颈疼痛、腰椎退行性病变等问题发生率较 2010 年上升 37%。在此背景下,开展久坐人群健康问题调查研究,不仅能增强大众对久坐危害的认知,更能为健康管理模式创新提供实证支撑,推动全社会形成关注健康的良好风尚。

从国内外研究现状来看,久坐与健康的关联研究已形成多学科交叉格局。在国外,Katzmarzyk 等[3]通过美国行为风险因素监测系统(BRFSS)数据验证,发现日均久坐超 8 小时人群的心血管疾病风险较久坐不足 8 小时的人群高 2.3 倍,且该关联在不同年龄、性别群体中均达到统计学显著性水平($P < 0.05$);Smith 团队[4]则聚焦久坐与骨骼肌肉系统的关联。研究指出,长期久坐人群椎间盘退行性疾病的患病率达 70%,且发病年龄较 2000 年提前 4.2 岁。在技术应用层面,Jocher 团队提出的 YOLOv8 目标检测模型,已实现对坐姿异常行为的毫秒级识别,为久坐行为实时监测提供了技术基础,但现有研究多聚焦单一健康结局或技术模块,缺乏“风险评估-干预方案-效果反馈”的全链条体系整合。

国内研究虽起步稍晚,但在人群特异性分析方面成果显著。钟卫宁等[5]较早指出我国办公人群因伏案工作导致的颈腰痛发生率达 43%,且与久坐时长呈正相关。国家统计局 2024 年发布的第三次全国时间利用调查公报显示,我国居民日均上网时长 6 小时 3 分钟,线上办公、在线学习等场景进一步延长了久坐时间,加剧了用眼疲劳、下肢水肿等久坐相关的健康问题[6]。

然而,国内现有研究仍存在两方面不足:一是模型应用多局限于单一算法,比如仅用随机森林进行风险预测,未充分发挥多模型协同的优势;二是干预方案多采用“一刀切”模式,缺乏针对不同职业、不同健康状态人群的精准化设计。为弥补研究缺口,本文构建多模型协同的久坐人群健康问题分析体系与实践平台。首先,用 K-means 聚类依据亚健康多维度指标划分人群亚型,解决异质性以提升分析精准度;其次,通过分段函数模型拟合不同亚型“久坐时长-亚健康风险”的关联,定位个性化风险阈值;然后,借助随机森林构建决策树集成学习,捕捉非线性关系与特征交互、降低过拟合;最后,再以梯度提升树迭代优化风险预测,结合 SHAP 解析生成个性化建议。通过多模型协同和支撑软件的开发,实现久坐监测、健康报告生成、个性化运动方案推送的全流程服务,进而推动健康管理向“主动预防”转变,最终为降低久坐健康风险和提升公众健康水平提供可行方案。

2. 基本理论

在实际应用中, 我们选择模型融合的方式, 先用 K-means 聚类分析对用户健康状态分型, 解决人群异质性问题, 再使用分段函数模型, 刻画风险阈值效应, 然后用随机森林(Random Forest)筛选关键特征降维, 最后输入梯度提升树(XGBoost)迭代优化预测, 四层模型层层递进, 为久坐人群的健康风险提供科学的风险预测支撑。

2.1. K-Means 聚类分析

K-means 聚类分析可基于疲劳程度、睡眠质量、生理指标等亚健康多维度指标, 将久坐人群划分为不同健康状态亚型, 避免直接分析整体关系掩盖亚型特异性, 使后续分析更精准。在运用 K-means 聚类分析研究“久坐人群健康问题”的关系时, 流程可分为以下步骤如图 1。

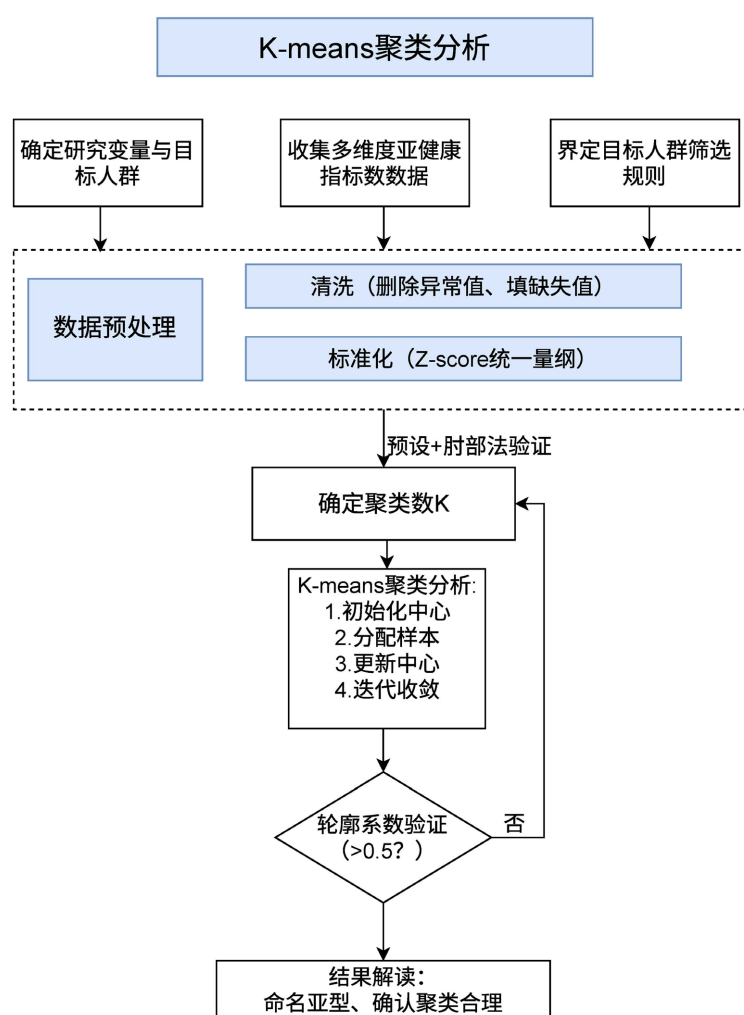


Figure 1. K-means clustering analysis flowchart

图 1. K-means 聚类分析流程图

一是确定研究变量与样本, 选取亚健康相关多维度指标, 并收集目标人群数据。二是数据预处理, 清洗数据, 删除极端异常值、用均值/中位数填充缺失值, 再标准化处理以消除变量数值范围差异的影响。三是确定聚类数量 K , 结合研究假设或文献预设 K 值, 用“肘部法”验证确定最终 K 值, 计算不同 K 的

误差平方和 SSE，选 SSE 下降变缓的 K 值。四是执行 K-means 聚类，先初始化簇中心，再计算样本与中心的欧氏距离分配样本、更新簇中心，迭代至中心稳定。 $x_i = (x_{i1}, x_{i2}, \dots, x_{in})$ 是样本的特征向量， $\mu_j = (\mu_{j1}, \mu_{j2}, \dots, \mu_{jn})$ 是簇中心的特征向量，其中 i 为样本编号， j 为簇编号，欧氏距离 $d(x_i, \mu_j)$ 如下：

$$d(x_i, \mu_j) = \sqrt{\sum_{k=1}^n (x_{ik} - \mu_{jk})^2}.$$

五是结果解读与验证，分析簇特征命名亚型，用轮廓系数(>0.5 表示合理)验证，确保同簇相似、异簇差异显著[7]。

2.2. 分段函数模型

在通过聚类分析明确健康状态分型后，分段函数模型再针对每一类细分人群，将其拟合为“久坐时长”与“亚健康风险”的关联模式，从而精准定位不同人群中久坐行为的“风险阈值”，例如可发现某一亚健康亚型人群每天久坐超过 6 小时后，亚健康风险会出现显著升高的拐点。这种基于聚类结果的分段分析，最大优势在于摆脱了“一刀切”式的统一阈值结论，让研究结论更具针对性和实践指导意义。流程可分为以下步骤如图 2：

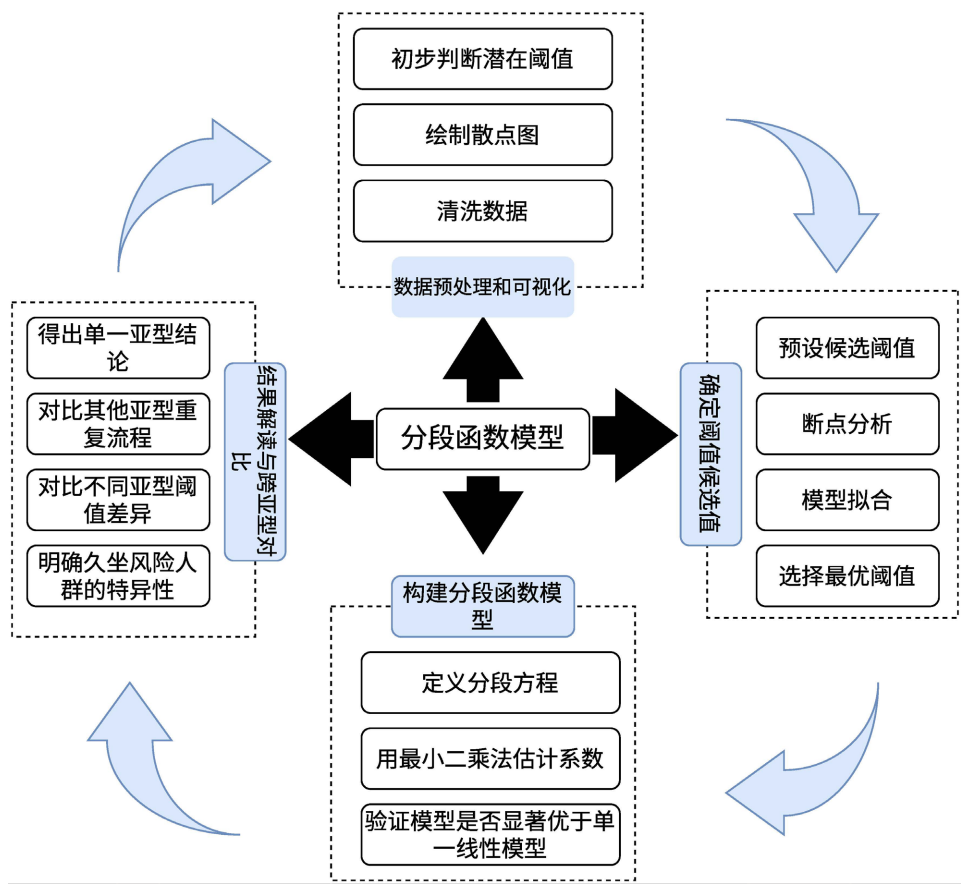


Figure 2. Piecewise function flowchart
图 2. 分段函数流程图

第一，确定分析对象与核心变量：基于 K-means 聚类结果，选取某一亚健康亚型人群，提取其“久坐时长”与“亚健康风险指标”，并预设二者存在分段关系的假设。第二，数据预处理与可视化：清洗

数据以剔除久坐时长、亚健康指标的极端值，再绘制散点图观察数据分布趋势，初步判断潜在阈值。第三，确定分段点：结合散点图趋势设定若干候选阈值，同时通过“阶跃函数回归”或“断点分析”计算不同候选阈值的模型拟合优度，选择误差最小的阈值。第四，构建分段函数模型：针对选定阈值定义分段方程，用最小二乘法估计系数，再通过 F 检验验证分段模型是否显著优于单一线性模型($P < 0.05$ 则分段关系成立)。第五，结果解读与跨亚型对比：先输出单一亚型的风险阈值及分段系数，再对 K-means 得到的其他亚型重复上述步骤，对比不同亚型的阈值差异，明确久坐风险的人群特异性。

2.3. 随机森林

该模型的核心逻辑：构建多棵决策树，通过集成学习降低过拟合风险。每棵树基于 Bootstrap 抽样数据训练，通过基尼指数或均方误差划分特征，最终风险预测为多棵树的投票/均值结果[8]。随机森林能捕捉非线性关系和特征交互，对缺失值和异常值具有较强鲁棒性。在运用随机森林模型研究久坐人群的健康问题时，流程可分为以下步骤如图 3：

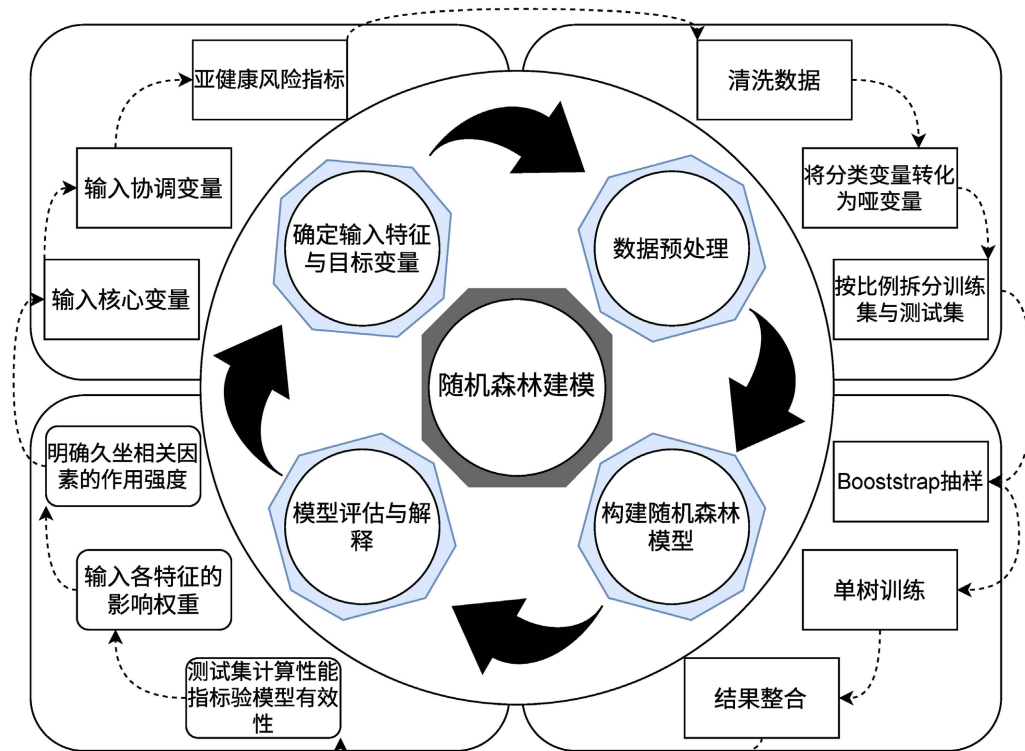


Figure 3. Random forest flowchart

图 3. 随机森林流程图

首先，确定输入特征与目标变量：输入特征含核心变量与协变量，目标变量为亚健康风险指标。其次，进行数据的预处理：清洗数据，处理缺失值，剔除异常值，将分类变量转换为哑变量，再按一定比例拆分训练集与测试集。接着，构建随机森林模型：先设参数，树数量 100 棵以平衡效率与复杂度，每棵树最大特征数取总特征数平方根。再训练模型——对训练集 Bootstrap 抽样生成子样本集，每子集单独训练决策树，最终整合多棵树结果。模型评估与解释：在测试集上计算性能指标验证模型有效性，同时输出各特征影响权重，明确久坐相关多因素的作用强度。最后是结果应用：一方面预测个体亚健康风险，另一方面识别影响久坐人群健康的关键因素，为干预提供方向。

2.4. 梯度提升树

该模型的核心逻辑：通过迭代训练弱分类器(决策树)，每次聚焦前一轮模型的预测误差，逐步优化风险预测。以 XGBoost 为例，目标函数包含损失项和正则项，通过泰勒展开近似求解最优树结构为

$$Obj = \sum_{i=1}^n l(y_i, \hat{y}_i^{(t)}) + \sum_{k=1}^t \Omega(f_k).$$

其中 l 为损失函数，衡量模型对久坐人群亚健康风险的预测值与真实值的偏差； Ω 为正则项，控制决策树复杂度，避免过拟合，确保模型对学生、职员等不同职业人群的预测通用稳定； y_i 为真实标签，表示第 i 个调查对象的真实健康状态； $\hat{y}_i^{(t)}$ 为第 t 轮迭代预测值，经 t 轮决策树训练后，模型对第 i 个样本的亚健康风险预测结果，迭代中逐步接近真实值 y_i ； f_k 为第 k 轮决策树，每轮新增的弱分类器，专门修正前 $k-1$ 轮模型的预测误差，比如优化对久坐且运动少人群的风险预测； n 为样本总数，参与训练的有效调查样本数(544 人)，损失函数求和范围为所有 544 个样本，整体衡量模型在全量数据上的偏差。

梯度提升树预测精度高，能挖掘细微特征影响，适合处理高维、非线性数据。该流程如图 4：

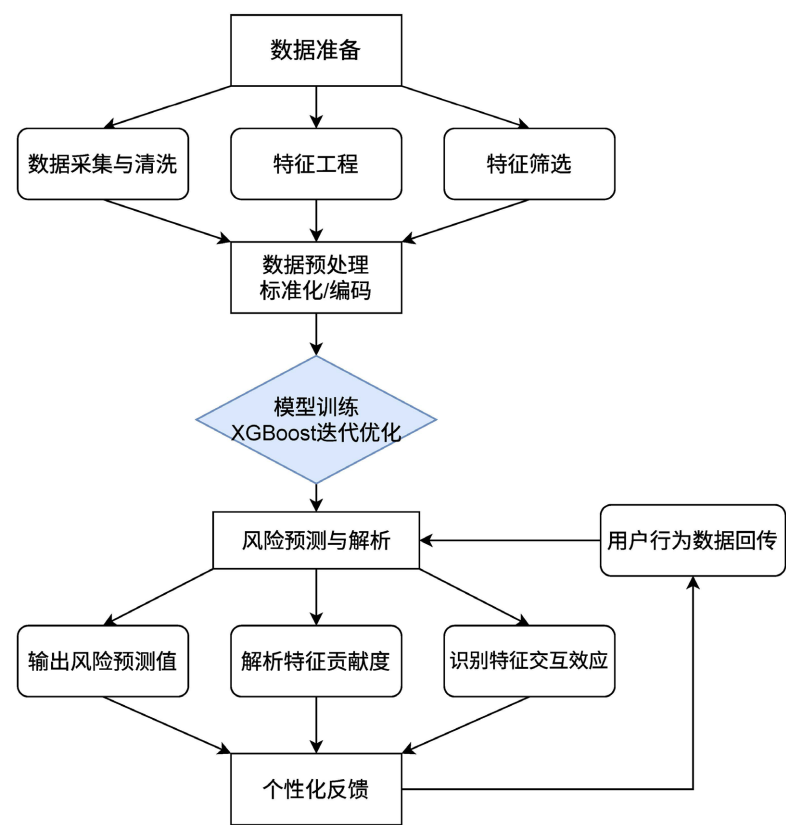


Figure 4. Gradient boosting tree (XGBoost) flowchart
图 4. 梯度提升树(XGBoost)流程图

数据采集好后，先以随机森林筛选出“久坐间隔时长”等与亚健康强相关的核心特征，再经预处理生成适配梯度提升树的衍生特征。XGBoost 从初始模型开始，每轮训练决策树拟合前序误差，通过目标函数控复杂度防过拟合，借助泰勒展开提升训练效率。模型输出实时风险值后，用 SHAP 值量化久坐行为细微影响并识别多因素交互，再转化为易懂提示与个性化建议，结合用户后续数据持续迭代，以适配应用“实时监测 - 精准干预”需求，在特征筛选基础上进一步提升预测精度与建议针对性。

3. 数据来源及模型应用

3.1. 数据来源

本文数据来源于问卷星平台的在线问卷调查,调查时间为 2025 年 6~8 月,有效样本量 544 人。样本的人口统计学特征详细信息如下:

3.1.1. 年龄结构

年龄跨度涵盖青少年至中老年多个阶段,其中 18 岁以下 79 人(占比 14.52%)、18~25 岁 100 人(占比 18.38%)、26~35 岁群体规模最大,达 153 人(占比 28.13%)、36~45 岁 121 人(占比 22.24%)、46 岁以上 91 人(占比 16.73%)。样本分布均衡全面,既囊括久坐行为高发群体,亦纳入中老年群体,能够客观反映不同年龄段人群的久坐相关特征,与调研主题高度契合。

3.1.2. 职业类型

本次久坐健康效应调研中,职业聚焦久坐高发群体,其中办公室职员 265 人(48.71%,占比最高),学生 140 人(25.74%),自由职业者 135 人(24.82%),其他职业 4 人(0.74%)。样本职业分布与调研主题高度契合,代表性强,可有效支撑调研结论。

3.1.3. 地域分布

所有样本主要来自四川省乐山市及周边区域,其中乐山市主城区样本 435 人,占比 79.96%;周边区县样本 109 人,占比 20.04%,地域集中性较强。

不同职业类型日均久坐时长的时段占比见表 1:

Table 1. Statistical results of questionnaire survey
表 1. 问卷调查统计结果

时长/类型	学生	办公室职员	自由职业	其他
<4 h	22.86%	44.91%	25.19%	0%
4 h~6 h	29.29%	31.32%	25.93%	50%
6 h~8 h	32.14%	18.11%	29.63%	25%
>8 h	15.71%	5.66%	19.26%	25%

通过 544 人的问卷我们可以知道学生与自由职业在不同时段的占比先升后降,多数在 6~8 h 这个时间段(学生 6~8 h 占比 32.14%,自由职业 6~8 h 占比 29.63%)。办公室职员在小于 4 h 占比最高之后逐渐减少(小于 4 h 占比 44.91%)。其他职业则是多数集中在 4~6 h (占比 50%)。不同职业由于工作内容和场景的影响久坐时长差异较大。

根据表 2 可知,久坐人群仅 3% 暂无明显问题,97% 出现不良健康结果,导致亚健康人群增多。当下人群的现状如上班上学伏案工作、下班久坐娱乐和出行依赖交通工具,这些行为导致运动量少,久坐使肌肉、韧带功能弱化,容易引起椎间盘突出及颈椎、腰椎疼痛[9]。通过中国互联网络信息中心我们可以知道在 2025 年 6 月之前我国网民有 11.23 亿人,互联网普及率达 79.7%。由于互联网的普及率较高导致线上办公和上网课的现象盛行,进一步加剧了久坐问题。经调查网民平均每日上网时间为 6 小时 3 分钟。长时间上网会导致长期用眼压力,造成眼睛疲劳视力衰退。久坐会减少下肢肌肉活动,静脉和淋巴回流受阻,导致液体淤积,因此造成下肢水肿。世界卫生组织(WHO)的《关于身体活动和久坐行为指南》中提到久坐使身体代谢减缓、脂肪堆积,增加肥胖风险,其实不止表 2 中提到的这几种病症,该指南指出

久坐行为还有可能导致睡眠时间减少、心血管疾病以及癌症等不良健康结果。我们可以根据以上分析来判断生病类型，再据表 1 统计的不同的职业在不同久坐时间段的占比，来制定健康管理计划，使运动方案贴合日常生活节奏，提高使用率，提升大家的健康水平。

Table 2. Proportion of sedentary-related health problems
表 2. 久坐相关健康问题占比表

久坐相关健康问题	颈椎/腰椎疼痛	眼睛疲劳	下肢水肿	肥胖/体重增加	暂无明显问题
占比	43%	52%	48%	57%	3%

3.2. 模型应用

基于上述理论与数据分析，为有效预防久坐引发的健康问题，我们依托“K-means 聚类分析、分段函数模型、随机森林、梯度提升树及 YOLOv8 目标检测模型”，构建数据处理与功能实现核心技术体系。结合涵盖学生、办公室职员等多职业群体的 544 份问卷数据，通过“K-means 聚类划分用户健康分型→分段函数拟合‘久坐时长 - 健康风险’关联阈值→随机森林与 XGBoost 构建风险预测模型→YOLOv8 实现实时坐姿检测→多模型协同输出个性化干预方案”的全流程，完成从数据输入到功能落地的转化：先对问卷调查的数据进行清洗与特征提取；再通过聚类、预测、检测类模型的协同运算，解析久坐行为与健康风险的关联规律；最终实现用户久坐行为的实时监测、健康风险预警及个性化干预，并通过数据验证了模型与功能的有效性，构建精准健康管理功能，深度融合模型预测与数学逻辑，为用户打造个性化健康守护方案，以下为模型应用功能层——用户端落地场景。

3.2.1. 实时监测与提醒

当连续久坐超过 1 小时(可自定时间)，设备震动提醒“起身活动”，实时显示当日久坐时长、起身次数，同时对标 WHO 健康建议(每小时活动 5 分钟)展示行为达标情况。

我们将使用 YOLOv8 模型，该模型有高效的目标检测能力，可用于识别用户的坐姿状态。如基于 YOLOv8 的智能坐姿检测系统，能识别正常坐姿、高低肩、距离过近、驼背、翘二郎腿及用眼距离过近等不良状态，检测精度可达 95.1%，每张图像检测时间为毫秒级，可实时监测并提醒用户调整坐姿。

3.2.2. 风险预警报告

该功能将结合随机森林和梯度提升树，每日生成“久坐健康风险评分”，每周推送风险趋势分析报告，精准标注高风险时段及对应的健康影响维度。

3.2.3. 个性化干预建议

我们将根据用户行为数据与健康档案，针对不同风险程度的人群将推送定制化干预方案：对轻度风险人群，推荐“动态调整提醒间隔”“3 分钟站立办公”；对中度风险人群，推荐“办公瑜伽动作”、“20 分钟步行计划”，同时支持与智能手表联动，在久坐超限时自动触发震动激励；针对重度风险人群，建议其预约专业康复评估，并推送适配的康复训练方案。

3.2.4. 数据安全

在数据安全方面，我们严格采用匿名化处理机制，全面契合 GDPR、《个人信息保护法》等全球及国内隐私保护法规要求。这一做法并非形式化设置，而是基于问卷调查数据的用户关切——个人隐私保护在用户关切因素中占比极高，既体现了用户对数据安全的重视，也印证了我们将隐私保护置于核心位置的必要性与合理性。

3.2.5. 用户反馈

采集用户对于干预建议的执行率和评分数据,用于优化干预建议的可行性。模型可构建“数据监测→风险预测→干预反馈”闭环,基于久坐时长、坐姿状态等数据开展统计与建模,利用概率分布、数据分析等数学工具预测健康风险,精准量化干预效果,把技术融入日常健康管理场景,助力用户从被动补救久坐亚健康问题,转向主动预防,让健康守护更具前瞻性。

4. 伦理与隐私保护

健康监测技术与个人敏感数据的深度融合,使久坐健康管理系统在发挥健康干预价值的同时,不可避免地面临伦理争议与隐私泄露风险。本次调查显示,多数受访者将“个人隐私保护”列为影响其使用相关 APP 的核心因素。基于这一关键,需构建系统化的伦理与隐私保护体系,在推动健康管理技术发展的同时保障用户合法权益,实现技术创新与权益保障的动态平衡。

4.1. 伦理与隐私风险识别

- 1) 数据泄露风险:用户久坐时长等敏感数据,若存储、传输防护不足,易被非法获取,衍生身份冒用、精准营销骚扰等问题。
- 2) 过度监测与隐私侵犯风险:YOLOv8 坐姿检测需持续采集数据,若监测超健康管理范围或频率过密,易侵犯隐私,引发用户心理不适与信任危机。
- 3) 数据滥用风险:未获用户授权,健康数据用于商业、科研外用途(如向第三方出售),背离采集目的,损害用户知情权与控制权。
- 4) 算法伦理风险:风险预测模型数据偏差,易对高龄、残障久坐者评估不公,致干预资源失衡,加剧健康不平等。

4.2. 风险缓解策略

- 1) 数据最小化原则:仅采集久坐时长、关键坐姿状态等健康管理核心数据,不采集姓名、身份证号等非必要标识信息;对已采数据脱敏,剥离可关联个人身份的特征。
- 2) 用户主导数据控制权:设清晰隐私权限模块,保障用户数据访问、修改、删除及授权的权利;明确告知数据用途、范围与保存期,遵循“明示同意”,用户可随时撤回授权且立即终止数据采集使用。
- 3) 差分隐私与加密技术:存储分析阶段用差分隐私(加微小噪声、数据聚合),保障分析有效且避免用户数据精准定位;传输 SSL/TLS 协议,存储用 AES-256 加密库,定期漏洞扫描与渗透测试防泄露。
- 4) 透明化披露与合规:公开数据处理流程、算法原理及隐私政策,通俗阐释技术逻辑;恪守《个人信息保护法》《GDPR》,主动申请 ISO/IEC 27001 等认证,强化用户信任。
- 5) 算法公平性优化:拓宽样本覆盖(含不同年龄、职业、地域、身体状况)以降模型偏差;定期审计算法公平性,核查弱势群体风险评估准确性,修正参数确保干预方案公平包容。

5. 结论与建议

久坐引发的健康问题已成为公众健康的重要挑战,本文结合问卷调查、多维度分析与技术整合取得了阶段性成果。研究核心在于进行了数据统计和分析,并构建了由 K-means 聚类分析、分段函数模型、随机森林、梯度提升树构成的四层协同模型体系:K-means 聚类破解人群异质性并精准分型健康状态;分段函数模型定位不同亚型人群的久坐风险阈值,为干预措施提供量化依据;随机森林与梯度提升树的组合则提升了亚健康风险预测精度,并结合 SHAP 值解析实现个性化风险归因[10]。基于 544 份调查数据的验证显示,该模型体系能有效捕捉不同职业人群的久坐特征,实现实时监测、风险预警、个性化干

预等功能,推动健康管理从“被动医疗”向“主动预防”转变。需要明确的是,本研究结论主要适用于青年学生、办公室职员群体,且适配以学习、办公为核心的日常久坐场景;对于其他地域群体、非办公类职业或中老年人群,结论的直接应用需结合实际场景进一步验证。

本文虽为久坐行为健康管理相关应用的研发提供了理论与实践支撑,但仍有进一步优化空间,未来可从以下方面深化:

首先,数据集存在局限性,可能影响模型普适性。当前数据仅来自乐山市 544 名受访者,样本量与地域范围有限,且职业集中于学生、办公室职员,年龄偏青年化,导致模型对其他地域、中老年及体力劳动者等群体适配性不足,结论普适性受限。未来可拓展地域覆盖,增加中老年、体力劳动者等群体样本,延长数据采集周期,提升模型适配性与长期风险预测能力,提升结论推广价值。

其次,现有分析主要聚焦久坐行为本身,未充分纳入其他关键影响因素。未来研究可引入多变量分析,综合考虑个体基础疾病史、每日饮食结构、工作强度等细节因素,以及季节变化、室内环境等外部条件,进一步提升风险预测的全面性与干预建议的精准度。

最后,该应用的健康管理功能仍有优化空间:可加强与智能穿戴设备、办公场景硬件的深度联动,提升实时监测的无缝性,同时结合用户反馈持续迭代干预方案;增加针对特殊人群的定制化模块,增强应用的实用性与包容性,未来可深化硬件联动实现无感监测,新增特殊人群定制模块;此外,也可与医疗机构、企业、学校合作推广,建立“监测-干预-康复”的闭环服务体系,推动研究成果向社会健康实践转化。

基金项目

乐山师范学院 2025 年省级大学生创新创业训练计划资助项目(S202510649206)。

参考文献

- [1] 国务院. “健康中国 2030”规划纲要[J]. 中国卫生政策研究, 2016, 9(10): 1-20.
- [2] 世界卫生组织. 全球久坐行为与健康问题数据报告[R]. 日内瓦: 世界卫生组织, 2023(更新版).
- [3] Katzmarzyk, P.T., Church, T.S., Craig, C.L., *et al.* (2023) Sedentary Behavior and Cardiovascular Disease Risk: Analysis of the BRFSS Data. *American Journal of Preventive Medicine*, **65**, 341-348.
- [4] Smith, J.D., Chen, C., Lee, Y.C., *et al.* (2024) Global Multicenter Study on Sedentary Behavior and Intervertebral Disc Degeneration. *Spine Journal*, **24**, 189-197.
- [5] 钟卫宁, 李立明, 王广增. 我国办公人群颈腰椎痛流行病学调查[J]. 中华劳动卫生职业病杂志, 2004, 22(5): 342-345.
- [6] 国家统计局. 第三次全国时间利用调查公报[R]. 北京: 国家统计局, 2024.
- [7] Rousseeuw, P.J. (1987) Silhouettes: A Graphical Aid to Cluster Analysis. *Computational Statistics*, **2**, 53-65. Thorndike, R.L. (1953) Who Belongs in the Family? *Psychometrika*, **18**, 267-276. <https://doi.org/10.1007/bf02289263>
- [8] Breiman, L. (2001) Random Forests. *Machine Learning*, **45**, 5-32. <https://doi.org/10.1023/a:1010933404324>
- [9] 钟卫宁, 褚洁芹. 久坐少动颈腰椎痛[N]. 北京日报, 2004-12-08.
- [10] Lundberg, S.M., *et al.* (2018) A Unified Approach to Interpreting Model Predictions. *Journal of Machine Learning Research*, **19**, 4765-4774.