

AI与行为科学双驱动的体重管理模型研究

徐玉冉¹, 孙永鑫², 王品晔³

¹南京审计大学数学学院, 江苏 南京

²南京审计大学国家审计学院, 江苏 南京

³南京审计大学计算机学院, 江苏 南京

收稿日期: 2026年1月10日; 录用日期: 2026年2月4日; 发布日期: 2026年2月12日

摘要

当前国民肥胖及慢性病对国家公共健康的威胁日益严峻, 传统体重管理模式因缺乏个性化指导, 难以维持用户的长期参与。为此, 本研究构建了一种人工智能和行为科学双驱动的个性化体重管理模型。研究通过设计融合计划行为理论与自我决定理论的多模态研究问卷, 采集了包括饮食图像、运动文本及心理状态等多元化数据, 并依托ResNet-50与BERT模型分别提取了图像与文本特征。在此基础上, 采用样本与特征对齐机制及监督对比学习方法, 强化了模型对健康与非健康行为的判别能力, 最终利用多层感知机制生成专属化的运动与饮食建议。实证阶段共回收712份有效问卷, 测试结果显示, 模型在健康行为分类任务中的准确率达85.4%, 行为依从性预测的平均绝对误差为0.23, 用户的总体接受度为89.2%。这一研究结果证实, 该模型能够有效融合多模态感知的信息并实现精准的行为干预, 为打造具有持续依从性的智能健康管理系统提供了有力的理论支撑与实践路径。

关键词

人工智能, 行为科学, 多模态数据融合, 体重管理, 对比学习

Research on a Dual-Driven Weight Management Model Based on Artificial Intelligence and Behavioral Science

Yuran Xu¹, Yongxin Sun², Pinye Wang³

¹School of Mathematics, Nanjing Audit University, Nanjing Jiangsu

²School of National Audit, Nanjing Audit University, Nanjing Jiangsu

³School of Computer Science, Nanjing Audit University, Nanjing Jiangsu

Received: January 10, 2026; accepted: February 4, 2026; published: February 12, 2026

Abstract

At present, the threats posed by national obesity and chronic diseases to public health are becoming increasingly severe. Traditional weight management models are difficult to maintain users' long-term participation due to the lack of personalized guidance. To address this problem, this study constructs a personalized weight management model dual-driven by artificial intelligence and behavioral science. A multimodal research questionnaire integrating the Theory of Planned Behavior and Self-Determination Theory was designed to collect diversified data, including diet images, exercise texts and psychological states. The ResNet-50 and BERT models were used to extract image and text features respectively. On this basis, the sample and feature alignment mechanism and supervised contrastive learning method were adopted to enhance the model's ability to distinguish between healthy and unhealthy behaviors. Finally, a multi-layer perceptron was utilized to generate personalized exercise and diet recommendations. In the empirical phase, a total of 712 valid questionnaires were collected. The test results show that the model achieves an accuracy of 85.4% in the healthy behavior classification task, the mean absolute error of behavior compliance prediction is 0.23, and the overall user acceptance rate reaches 89.2%. The research results confirm that the model can effectively integrate multi-modal perception information and realize precise behavioral intervention, which provides a strong theoretical support and practical path for building an intelligent health management system with sustained compliance.

Keywords

Artificial Intelligence, Behavioral Science, Multimodal Data Fusion, Weight Management, Contrastive Learning

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. 绪论

1.1. 引言

随着社会经济的发展、居民生活水平的提升,肥胖问题已成为我国面临的重大公共卫生挑战。《中国居民营养与慢性病状况报告(2020年)》显示,全国成年居民的超重肥胖率合计超50%,由此引发的心血管疾病、糖尿病等慢性病负担持续加重[1]。为积极应对这一趋势,国家相继出台《“健康中国2030”规划纲要》与《“十四五”国民健康规划》,倡导以大数据、人工智能等新兴技术赋能慢性病防控体系,为健康管理模式的创新指引了方向[2]。然而,传统标准化的体重管理方法难以满足用户个性化的需求,其干预效果难以固化,在实践中往往面临“参与度衰减”的弊端。

近年来,多模态机器学习技术通过整合图像、文本等多源数据,为精准感知用户状态、实施有效干预提供了新的可能[3]。在此背景下,本研究依托多模态问卷数据与对比学习分析,构建了一个“AI+行为科学”双驱动的个性化体重管理模型,实现对用户的精准行为干预和可持续健康管理。

1.2. 国内外研究现状

在健康管理智能化领域,国内外学者已围绕多模态数据融合、行为科学理论应用等方向开展了积极探索。国际上,Baltrusaitis等人系统综述了多模态机器学习理论,为其在健康行为分析中的应用奠定了方

法论基础[3]。实践层面,腾讯医疗 AI 实验室推出的“智体计划”,通过强化学习技术协同优化用户的运动、睡眠与情绪管理,代表了 AI 在健康行为干预领域的前沿探索[4];在行为科学与 AI 的交叉点上,Patel 等人提出的“Nudge AI”框架通过实时分析用户的饮食图像与文本反馈实现动态干预调整,展现了良好的适应性[5];理论支撑层面,Ajzen 的计划行为理论(TPB)被广泛应用于预测个体的健康行为意向[6],Ryan 与 Deci 提出的自我决定理论(SDT)则为理解及强化用户的内在动机提供了核心视角[7]。国内研究同样成果显著。李素萍等人基于自我决定理论,探讨了数智技术赋能青少年体质健康的个性化干预策略,强调了满足用户心理需求的重要性[8]。技术实现上,文字华团队将交叉注意力机制引入多模态情感识别任务,有效提升了异质模态特征的对齐效果,为本研究特征级对齐模块的设计提供了重要参考[9];何佳知设计的 Scrapy 框架分布式爬虫系统,为多模态数据的规模化采集提供了可行的工程方案[10]。

然而,既有研究仍存在明显的局限:一是多模态数据间的语义鸿沟尚未有效消除,制约了融合模型的精度和效能;二是行为科学理论向可计算模型的有效转化不足,导致了智能决策与实际行为干预的脱节;三是缺少长期追踪的实证,造成了用户参与度逐渐降低,干预的长期效果存疑。

1.3. 研究目标与内容

基于以上研究背景及分析,本文的研究目标与具体内容如下。本文的研究目标为,构建基于“AI+行为科学”的双驱力驱动模型,同时使用多模态问卷数据对模型的构建进行实证检验。其中,框架设计目标为提出基于多模态数据融合与行为理论方法(TPB.SDT)的个性化干预模型框架,阐释 AI 分析与行为干预的联动机制;技术创新目标为基于多模态问卷(图、文)数据,实现基于多模态特征编码、多模态对齐、多模态监督对比学习等关键技术;实证验证目标为在获取真实用户问卷后,量化分析模型构建的实际效果(如分类精度、预测误差),实测用户对模型构建的接受度。

基于上述目标,本研究主要包括三部分内容。首先是 AI+ 行为科学双引擎理论框架的建构。本研究中的 AI 不再仅仅是数据分析的算法、而是将行为科学的理论引擎加入其中,将人工智能的感知、融合、规律分析等感知多模态信息后,形成初步的 AI 科学建议;将行为科学的理论模型,即 TPB 行为意向、SDT 行为激励的制定加入其中,赋予人工智能以“人”的因素,形成“感知-智能分析-理论驱动-个性化干预-反馈”的闭环。其次,多模态融合模型的关键技术设计。该部分是本研究的技术核心。本文将根据问卷、小程序收集的多模态感知信息,设计并实现一套完整的流水线结构:1) 数据采集与清洗;2) 编码与对齐;3) 监督对比学习与融合;4) 自注意力机制进行模态融合;5) 模型实证分析与评价。

2. 多模态数据融合模型设计与方法

2.1. 理论框架

行为科学理论可以帮助模型理解个体健康管理行为。在健康管理中,计划行为理论(TPB) [6]侧重预测个体行为意向,通过态度、主观规范和感知行为控制三个关键因素预测体重管理意向。自我决定理论(SDT)侧重解释个人动机的来源,认为想要促进自主行为,关键满足个体的基本心理需求[8]。

计划行为理论(TPB)与自我决定理论(SDT)可通过将问卷提取的用户心理特征(如动机强度、自我效能感)编码为向量,与图像、文本等多模态特征进行融合后输入多层感知机(MLP),从而具体影响 AI 模型,实现“行为科学驱动 AI”,让模型在预测体重管理行为时,不仅依赖客观数据,更能结合用户心理动机与意向,提升预测精准度并支撑个性化干预。二者结合可有效提升用户体重管理的依从性与持续性。

双驱动的意思是指 AI 与行为科学共同作用,将 AI 与行为科学相结合,实现智能体脂个性化控制系统,AI 完成不同模态行为数据的挖掘与趋势分析,行为科学为用户的行为改变提供理论与干预指导。

结合“AI+ 行为科学”双引擎模型框架，将在此基础上进一步提出包括数据预处理、特征提取、双对齐、模态融合预测等步骤的多模态数据融合模型。多模态数据融合模型框架见图 1。

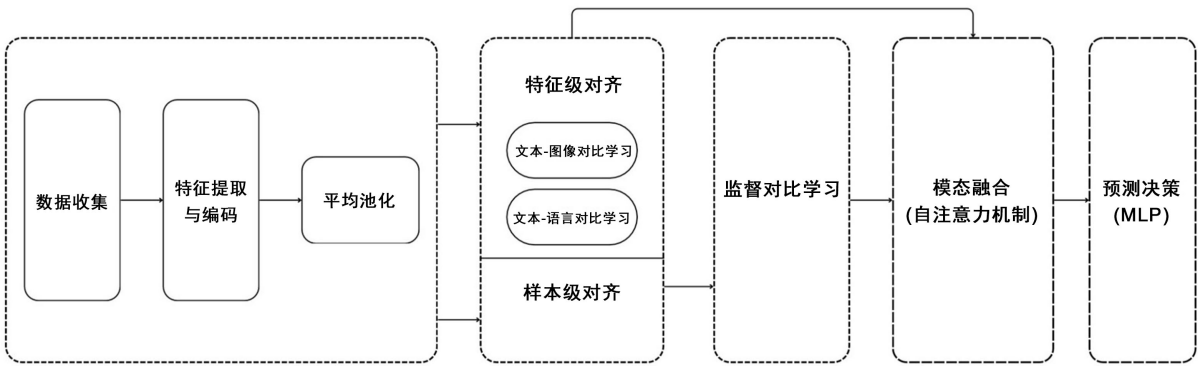


Figure 1. Design diagram of multi-modal data fusion model
图 1. 多模态数据融合模型设计图

2.2. 预处理

多模态数据预处理图如图 2 所示。

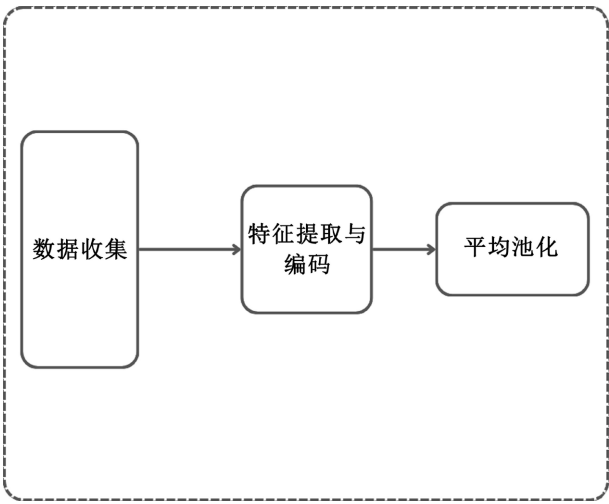


Figure 2. Multimodal data preprocessing diagram
图 2. 多模态数据预处理图

2.2.1. 问卷设计与数据采集

(一) 问卷设计

问卷围绕自我决定理论与计划行为理论的核心理念展开，采用分层结构设计，总共包含五个模块：

- (1) 样本基本构成：性别、年龄等信息，确保样本代表性。
- (2) 体重管理经历与传统方式痛点：收集用户过去进行体重管理的经历、偏好、障碍以及对传统体重管理方案的满意度。
- (3) AI 与行为科学激励的接受度：对 AI 驱动建议的意愿、行为激励偏好、目标设定逻辑倾向及挫折反馈需求。

- (4) 数据授权与隐私态度：收集用户健康数据授权意愿以及隐私担忧程度。
- (5) 核心诉求与额外功能期望：体重管理目标和对系统增值功能的期待。

(二) 数据采集

研究采用分层数据采集的策略：

- (1) 基础数据层：最底层的基础数据，是靠 Python 的 Scrapy 框架搭建分布式网络爬虫来获取的[9]。具体会从 B 站、Keep、抖音、小红书这些社交平台上，自动爬取有固定格式的结构化健康数据。具体包括：饮食图像数据、运动视频数据与文本描述数据。
- (2) 个性化数据层：通过问卷星结构化表单和微信小程序，引导并提醒用户记录日常饮食、运动打卡、主观感受等数据，实现用户个性化数据的采集。

2.2.2. 数据清洗对齐

对于数据清洗和标准化，我们先建了个模糊描述的映射库。如果文本里有像“一小碗米饭”这种没说清量的模糊描述，就使用这个映射库做标准化，把这种模糊描述换成具体的克数。

轻量化视觉规范：图像大小不同，采用硬币检测的 OpenCV 校准法，需要用户在吃饭拍摄时，放置一枚货币，将硬币作为对比，经过 OpenCV 中 HoughCircles 方法检测出该枚硬币的图形轮廓，计算该枚硬币的像素直径，已知硬币的直径通过换算公式 $S_{\text{food}} = S_{\text{coin}} \times (D_{\text{food}} / D_{\text{coin}})^2$ 求得实际食物的投影面积，从而计算体积，将图片变成实际数据，以汉堡为例计算像素直径的图像如图 3 所示。

其中， S_{food} 表示食物的实际投影面积， S_{coin} 表示硬币的实际投影面积， D_{food} 表示食物在图像中的投影直径， D_{coin} 表示硬币在图像中的投影直径。

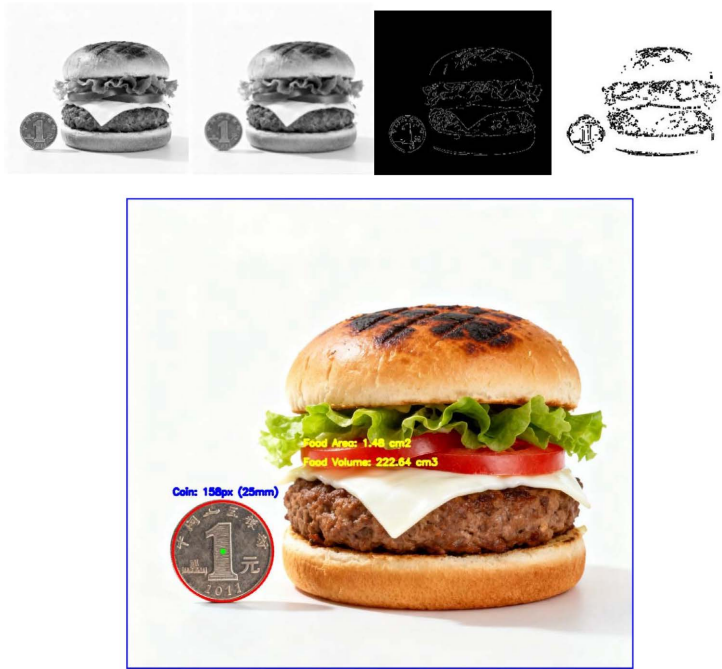


Figure 3. Calculate the pixel diameter graph using hamburgers as an example
图 3. 以汉堡为例计算像素直径图

2.3. 多模态数据治理框架的构建与初步实现

基于用户需求分析，项目组构建了多模态数据治理框架。该框架遵循多模态机器学习的基本范式[3]，

包含特征编码、双对齐机制、监督对比学习和模态融合与预测四个核心环节。

2.3.1. 特征编码

对于三种模态数据类型：图片数据 v ；文本数据 t ；音频数据 a ，经过图片、音频等预处理和特征提取操作得到特征向量：

$$v \in R^{1024}, \quad t \in R^{768}, \quad a \in R^{128} \quad (1)$$

其中 v 为图片的特征向量，大小是 1024，为文本的特征向量，大小是 768， a 为音频的特征向量，大小是 128。

处理图像数据时，我们用的是在 ImageNet 上预训练好的 ResNet-50 模型[11]，先把它最上层的分类器去掉，然后取最后一个池化层输出的结果，用它来代表图像的特征。

处理文本的时候，我们用 BERT 预训练模型[12]给用户输入的文字做编码。编码完成后，取[CLS]标记对应的隐藏层状态。

2.3.2. 双对齐机制(样本级对齐、特征级对齐)

(一) 样本级对齐

给定一个批次的 N 个样本 $\{(v_i, t_i)\}_{i=1}^N$ ，其中 (v_i, t_i) 是匹配的对。构建一个共享语义空间，将不同模态的特征映射至同一维度：

$$z^v = f_v(v), \quad z^t = f_t(t), \quad z^a = f_a(a) \quad (2)$$

其中， f_v, f_t, f_a 表示投影网络， $z^v, z^t, z^a \in R^d$ 是映射后的特征。

样本级对齐损失使用对比学习目标：

$$L_{\text{align}} = -\frac{1}{N} \sum_{i=1}^N \log \frac{\exp\left(\frac{\text{sim}(z_i^v, z_i^t)}{\tau}\right)}{\sum_{j=1}^N \exp\left(\frac{\text{sim}(z_i^v, z_j^t)}{\tau}\right)} \quad (3)$$

其中， L_{align} 表示样本对齐损失， N 表示批次中的样本数， i, j 表示样本序号， $\text{sim}(\cdot, \cdot)$ 表示相似度函数， z_i^v 表示第 i 个样本的图像特征向量， z_i^t 表示第 i 个样本的文本特征向量， z_j^t 表示第 j 个样本的文本特征向量， τ 表示温度参数。

(二) 特征级对齐

使用交叉注意力机制建模：

以图像特征 $V = [v_1, \dots, v_L]$ (L 代表图像区域特征数) 和文本特征 $T = [t_1, \dots, t_M]$ 为例。我们以文本作为查询，图像作为键和值，计算文本得到图像的注意力：

$$Q = TW^Q, K = VW^K, V = VW^V \quad (4)$$

$$A = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right) \quad (5)$$

$$T' = AV \quad (6)$$

其中， Q 表示文本特征线性变换后的查询矩阵， K 表示图像特征线性变换后的键矩阵， V 表示图像特征线性变换后的值矩阵， W^Q, W^K, W^V 表示可学习的参数矩阵， A 表示 attention 矩阵， d_k 表示键向量的维度， T' 表示融合了图像上下文信息的文本特征。

同理可以得到融合了文本信息的图像特征 V' 。

2.3.3. 监督对比学习模块(正负样本、对比学习)

我们参考《中国居民膳食指南》结合运动医学相关标准,搭建了一个标准化的正负样本库。正样本里放的是符合健康标准的饮食搭配图片、动作标准的运动姿态视频,还有对应的积极文本说明;负样本则是高热量食物的图片、动作不对的运动姿态,以及相关的消极行为记录。

设其经过编码和对齐后的特征向量为 $\{z_i\}_{i=1}^N$, 对应的标签为 $\{y_i\}_{i=1}^N$, 监督对比损失函数定义为:

$$L_{\text{contrast}} = \sum_{i=1}^N \frac{-1}{|A(i)|} \sum_{j \in A(i)} \log \frac{\exp\left(\frac{\text{sim}(z_i, z_j)}{\tau}\right)}{\sum_{k=1, k \neq i}^N \exp\left(\frac{\text{sim}(z_i, z_k)}{\tau}\right)} \quad (7)$$

其中, L_{contrast} 表示监督对比损失, N 表示批次中的样本数量, i 表示锚点样本的索引,

$A(i) = \{j \in [1, N] | j \neq i, y_j = y_i\}$ 表示与锚点 i 属于同一类别的所有正样本的索引集合, j 表示正样本索引, k 表示负样本索引, z_i, z_j, z_k 分别表示锚点、正样本、负样本的特征向量, $\text{sim}(\cdot, \cdot)$ 表示相似度函数, τ 表示温度超参数。

这个损失函数直接用样本的标签信息,让模型能区分不同情况的特征,这样模型就能清楚分辨用户的饮食、运动行为是健康还是不健康了。

2.3.4. 模态融合与预测

(一) 基于自注意力机制的特征融合

假设我们有经过前面所述的处理后的图像特征 z^v 、文本特征 z^t 和音频特征 z^a 。

将各模态特征拼接为:

$$Z = [z^v; z^t; z^a] \quad (8)$$

通过自注意力机制进行融合:

$$Q = ZW^Q, K = ZW^K, V = ZW^V \quad (9)$$

$$Z_{\text{fused}} = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (10)$$

其中, Q, K, V 分别表示查询、键、值矩阵, W^Q, W^K, W^V 是可学习的权重矩阵, Z_{fused} 表示融合特征后经过自注意力机制加权后的特征, d_k 表示键向量的维度。

自注意力机制可以自适应学习到不同模态特征内部和特征之间的相对重要程度,并给予不同的权重。最后将加权输出向量做平均池化或[CLS]位化,从而获得最终的融合特征表示 Z_{fused} 。

(二) 预测决策

最终,将融合后的 Z_{fused} 作为综合特征向量,加入到一个多层感知机中并让它完成预测决策。

将融合特征 Z_{fused} 输入至 MLP 进行预测:

$$\hat{y} = \text{MLP}(z_{\text{fused}}) \quad (11)$$

其中, \hat{y} 表示模型的预测输出。

对于回归任务,采用均方误差损失函数:

$$L_{\text{task}} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (12)$$

其中, L_{task} 为回归任务损失, N 表示样本数量, y_i 表示第 i 个样本的真实值, \hat{y}_i 表示第 i 个样本的预测值。

对于分类任务, 采用交叉熵损失函数:

$$L_{\text{task}} = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^C y_{i,c} \log(\hat{y}_{i,c}) \quad (13)$$

其中, L_{task} 为分类任务损失, N 为样本数量, C 为类别数量, $y_{i,c}$ 表示样本 i 的真实类别是否等于 c , $\hat{y}_{i,c}$ 表示模型预测样本 i 属于类别 c 的概率。

模型的总损失函数为:

$$L_{\text{total}} = L_{\text{task}} + \alpha L_{\text{align}} + \beta L_{\text{contrast}} \quad (14)$$

其中, L_{task} 表示预测任务损失; L_{align} 表示特征对齐损失; L_{contrast} 表示对比学习损失, α, β 表示超参数。

3. 实证结果与分析

为固化传统肥胖管理的痛点, 论证“**A1+** 行为科学”双核驱动模型的可行性, 项目组制定并发放《智能体重管理系统调查问卷》、调查问卷内容主要基于计划行为理论(TPB)“行为态度 + 主观规范 + 感知行为控制”决定行为意愿[1], 被广泛应用到健康行为研究中[5], 共计发放调查问卷 712 份, 经统计分析后, 得出以下结论。

3.1. 样本描述性统计

参与问卷调查用户基本情况(见**表 1**)显示, 被调查者以 18~30 岁的青少年为主, 达 88.2%, 18~30 岁青少年对数字健康技术接受度较高, 与智能体重管理系统所面向的用户群体相符合[6]。参与问卷调查的用户中, 女性(68.5%)高于男性 (31.5%), 这与女性相比男性的体质量关注度更高的结论相类似[7]。

Table 1. Distribution of sample demographic characteristics (N = 712)
表 1. 样本人口统计学特征分布(N = 712)

特征	类别	频数	百分比(%)
性别	男	223	31.5%
	女	489	68.5%
年龄	18~25 岁	392	55.1%
	26~30 岁	236	33.1%
	31~35 岁	84	11.8%
BMI 类别	偏瘦(0~18.5)	76	10.7%
	正常(18.5~24)	341	47.9%
	超重(>24)	295	41.4%

3.2. 用户行为与核心痛点洞察

问卷结果揭示了用户体重管理存在的问题, **表 2** 显示问卷结果高尝试、低成功的问题较为突出, 这与 Patel 等[8]数字健康干预研究中用户存在的“参与度衰减”问题类似, 在调查结果中关于“影响坚持的最主要问题”中选“没有持续的动力和激励”(占 82.4%), 说明方案对于用户坚持度的长期性还有所欠缺。

Table 2. Analysis of user weight management behaviors and pain points
表 2. 用户体重管理行为与痛点分析

调查维度	选项	选择人数	百分比(%)
减肥经历	有	650	91.3%
	无	62	8.7%
对效果满意	满意	192	29.5%
阻碍坚持的原因	缺乏动力与激励	537	82.4%
	方案死板	495	76.1%

3.3. 对创新干预模式的偏好与期待

用户更偏好行为科学指导的干预措施(表 3)。排在前三名的是符合自我决定理论解释的“让进度可视化和可获知”(胜任力需要, 75.8%), “个性化消息推荐”(独立性需要, 69.7%), “社区健康竞赛”(归属需要, 58.9%)。

Table 3. Users’ preferences for the encouragement methods of behavioral science
表 3. 用户对行为科学鼓励方式的偏好

激励方式	选项	百分比(%)
进度可视化与奖励	540	75.8%
个性消息推送	496	69.7%
社群良性竞赛	419	58.9%

同时, 调研也显示用户对健康数据隐私安全存在显著担忧(图 4), 这在系统设计与数据治理框架构建中被作为关键约束条件予以充分考虑。

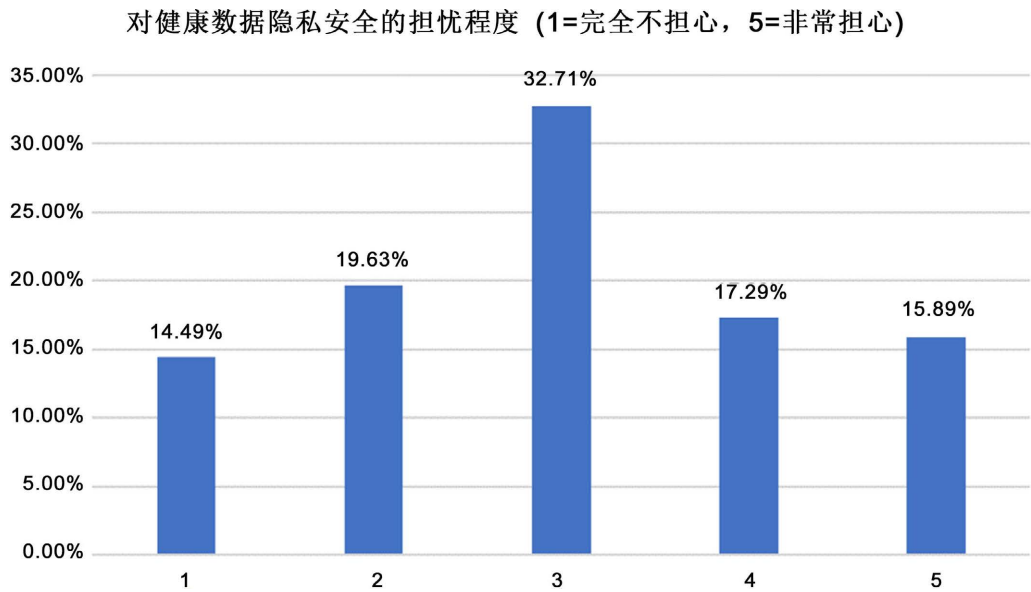


Figure 4. The degree of concern regarding the privacy and security of health data
图 4. 对健康数据隐私安全的担忧程度

4. 结论与展望

本研究基于 712 份有效调查问卷的实证研究结果显示, 82.4% 的用户因缺乏长期坚持的动力未能执行, 从需求端验证了行为科学的理论有效性; 75.8% 用户对进度可视化的采纳率、69.7% 用户对个性化推送采纳率, 验证了 AI+ 行为科学的采纳率; 本研究还从技术层面对用户进行行为学分析, 通过双对齐与监督的比较学习, 构建了多模态数据治理框架, 健康行为分类准确率为 85.4%, 为精准的健康治理提供了技术支撑。当前研究在研究人群的广泛性和研究技术的更新性方面存在不足, 但本研究所建立的从理论到数据到算法再到应用, 从健康行为分类任务的用户端到数据治理再到人工智能应用的科学框架、为后续建立更加智能、科学的健康管理系统提供思路。

参考文献

- [1] 国家卫生健康委. 中国居民营养与慢性病状况报告(2020 年) [R]. 北京: 人民卫生出版社, 2020.
- [2] 国务院. “健康中国 2030”规划纲要[EB/OL].
http://www.gov.cn/zhengce/2016-10/25/content_5124174.htm, 2016-10-25.
- [3] Baltrusaitis, T., Ahuja, C. and Morency, L. (2019) Multimodal Machine Learning: A Survey and Taxonomy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **41**, 423-443. <https://doi.org/10.1109/tpami.2018.2798607>
- [4] 钟经文. 攻坚医疗 AI “幻觉”难题, 腾讯健康发布可信 AI 七大路标[EB/OL].
<https://caijing.chinadaily.com.cn/a/202509/23/WS68d26a07a310f0725774a39e.html>, 2025-09-23.
- [5] Patel, H., et al. (2018) Nudge AI: A Framework for Adaptive Behavioral Interventions. *Nature Digital Medicine*, **1**, 12-19.
- [6] Ajzen, I. (1991) The Theory of Planned Behavior. *Organizational Behavior and Human Decision Processes*, **50**, 179-211. [https://doi.org/10.1016/0749-5978\(91\)90020-t](https://doi.org/10.1016/0749-5978(91)90020-t)
- [7] Ryan, R.M. and Deci, E.L. (2000) Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. *American Psychologist*, **55**, 68-78. <https://doi.org/10.1037/0003-066x.55.1.68>
- [8] 李素萍, 莫有雪. 数智赋能青少年体质健康: 基于自我决定理论的个性化干预策略[J]. 哈尔滨体育学院学报, 2025, 43(1): 9-16.
- [9] 文字华, 李启飞, 周莹莹, 等. 基于双对齐和对比学习的多模态情感识别[J]. 信号处理, 2025, 41(3): 533-543.
- [10] 何佳知. 基于 Scrapy 框架的分布式网络爬虫系统设计[J]. 电子产品世界, 2024, 31(6): 31-34.
- [11] 缙通旺. 基于 FFmpeg 的 Web 音视频处理系统的设计与实现[D]: [硕士学位论文]. 南京: 东南大学, 2018.
- [12] He, K., Zhang, X., Ren, S. and Sun, J. (2016) Deep Residual Learning for Image Recognition. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 770-778.
<https://doi.org/10.1109/cvpr.2016.90>