

基于XGBoost-SHAP框架的我国农村居民生活质量测度与影响因素研究

房可昱, 徐晓慧, 赵映朝, 邱子懿

浙江理工大学经济管理学院, 浙江 杭州

收稿日期: 2026年3月15日; 录用日期: 2026年4月5日; 发布日期: 2026年4月22日

摘要

本文基于全国性微观调查数据CFPS (2020), 在多维生活质量框架下构建涵盖经济能力、教育水平、健康状况、社会支持、公共服务、数字参与与主观满意度等多个领域的农村居民生活质量指标体系。通过熵权法与因子分析分别构建客观与主观生活质量指数, 并采用XGBoost模型对多维指标进行非线性融合, 形成综合生活质量指数。同时, 引入SHAP方法对模型进行解释, 识别不同指标对生活质量的边际贡献及其影响机制。研究结果表明: 第一, 农村居民整体生活质量水平偏低, 呈现显著右偏分布, 中低生活质量群体占比大, 整体表现出“低位集中、尾部有限”的结构特征。第二, 从区域视角看, 东、中、西部的生活质量水平差异并不显著, 三大区域分布高度重叠, 区域经济发展的差距尚未充分传导至农村家庭生活层面。第三, 农村生活质量呈现“公共服务质量-社会支持网络-经济能力”三元结构, 其中教育公共服务和社会互助网络贡献度最高, 收入与教育能力发挥重要基础作用, 而健康状况、医疗体验和数字参与的边际贡献相对较弱。研究揭示了农村生活质量的结构性特征与关键驱动机制, 为精准改善农村居民福祉和优化公共服务供给提供了重要政策参考。

关键词

农村居民, 生活质量, 熵权法, XGBoost, SHAP

Research on Measuring and Influencing Factors of Rural Residents' Quality of Life in China Based on XGBoost SHAP Framework

Keyu Fang, Xiaohui Xu, Yingchao Zhao, Ziyi Qiu

School of Economics and Management, Zhejiang Sci-Tech University, Hangzhou Zhejiang

Received: *****

文章引用: 房可昱, 徐晓慧, 赵映朝, 邱子懿. 基于 XGBoost-SHAP 框架的我国农村居民生活质量测度与影响因素研究[J]. 统计学与应用, 2026, 15(4): 214-225. DOI: 10.12677/sa.2026.154084

Abstract

Drawing on microdata from the China Family Panel Studies (CFPS 2020), this study measures the quality of life of rural residents in China using a multidimensional framework that incorporates economic capacity, education, health status, social support, public services, digital participation, and subjective life satisfaction. To capture both objective conditions and subjective perceptions, the entropy weight method and factor analysis are first used to construct objective and subjective quality-of-life indices. These indicators are subsequently integrated through the XGBoost algorithm to generate a comprehensive index, and the SHAP method is employed to interpret the model and quantify the marginal contribution of each factor. The results show that the overall quality of life of rural residents remains relatively low and displays a pronounced right-skewed distribution, indicating that a large share of the population is concentrated at lower levels of well-being, with a relatively limited upper tail. Regional comparisons suggest that differences among eastern, central, and western China are not statistically significant, and the distributions across the three regions largely overlap. This implies that macro-level regional economic disparities have not yet been fully translated into differences in rural household living conditions. Further analysis indicates that rural quality of life is primarily shaped by a structural combination of public service quality, social support networks, and economic capacity. Among these factors, educational public services and social support show the strongest contributions, while income and educational attainment provide an important foundational basis for improvement. In contrast, health status, healthcare experience, and digital participation exhibit relatively weaker marginal contributions. These findings provide important policy implications for improving rural well-being and optimizing the provision of public services.

Keywords

Rural Residents, Quality of Life, Entropy Weight Method, XGBoost, SHAP

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

近年来,随着乡村振兴战略的深入推进以及共同富裕目标的持续落实,农村居民生活质量成为衡量农村发展成效与民生改善水平的重要指标[1]。新技术、新产业和新公共服务不断向农村延伸,为农村生活质量提升创造了新的机遇[2];但与此同时,区域发展不平衡、公共服务短板、医疗教育资源不足、数字鸿沟扩大、人口老龄化加速等问题依然突出[3],不同地区和不同群体之间的生活质量差异呈现结构性特征[4]。这种“改善与不足并存”的现实,使得准确识别农村生活质量的短板所在和关键驱动力变得尤为重要。

在学术层面,生活质量研究长期强调经济条件、健康状况、社会参与、公共服务、心理福祉等多维因素的综合影响[5]。国际上形成了多个成熟的福祉理论框架,如“能力方法”“主客观融合法”等,但这些框架在农村情境中的适用性仍有限。原因在于,农村地区受产业结构单一、基础设施薄弱、公共服务差距大、家庭结构复杂和数字化水平不足等因素影响,其生活质量的形成机制比城市更加复杂多样[6][7]。现有相关研究大多聚焦城乡差距或单维度指标,缺乏能够覆盖农村生活全景的指标体系;同时,传统线性模型难以处理多维指标间的非线性关系和交互作用,使得生活质量的核心影响因素难以被精确识

别,也限制了研究结论的政策价值。

基于此,本文拟构建一个更符合农村的生活质量测度框架。为此,本文以 CFPS (2020)全国性微观数据为基础,围绕农村家庭的经济能力、居住条件、健康状况、教育文化、社会参与、社会支持、公共服务治理和主观幸福等方面,构建了包含 8 个准则层、8 个指标的生活质量指标体系[8][9]。随后采用熵权法与因子分析生成客观与主观生活质量指数,再通过 XGBoost 模型对多维指标进行融合,构建综合生活质量指数。最后,引入 SHAP 方法揭示各指标对生活质量的边际贡献及其重要性排序,从而识别影响生活质量的关键因素及其作用路径。研究结果显示:我国农村居民生活质量整体呈持续改善趋势,但在经济、医疗、教育、基础设施和数字融入等方面仍存在显著差异,呈现“西弱东强、北低南高”的区域格局;生活质量的影响机制呈现明显的层次结构,其中经济能力发挥核心提升作用,医疗可及性与健康风险是主要制约因素,而数字参与则成为近年来推动生活质量提升的重要力量[10];此外,主观幸福、安全感等心理因素对生活质量具有稳定的正向影响,表明主观体验在农村生活质量中不可忽视。

相应地,本文的主要贡献体现在三个方面:本文的主要贡献体现在三个方面:第一,构建了一个契合中国农村情境的多维度生活质量指标体系。相较于以往研究侧重城乡对比或单一维度指标,本文基于 CFPS (2020)构建了涵盖经济能力、教育水平、健康状况、社会支持、公共服务、数字参与与主观满意度等八大领域的生活质量体系,实现了对农村生活全景的系统刻画,从指标体系构建层面弥补了现有研究的不足。第二,将 XGBoost 与 SHAP 方法引入农村生活质量测度研究,提高了评价体系的精度与解释力。XGBoost 能够捕捉多维指标间复杂的非线性关系,而 SHAP 提供了透明的贡献度解释,使综合指数既具有预测优势,又能揭示驱动机制,有效克服传统线性方法难以识别非线性影响与交互效应的缺陷。第三,采用全国性大样本的农村居民“个体层面”数据开展分析,显著拓展了现有研究的分析层级。当前相关研究多基于省级宏观数据或家庭层面数据,存在指标粗糙、异质性不足、个体差异被掩盖等局限。本文使用农村居民的个体级微观数据进行测度与机制识别,不仅提高了指标稳定性和估计精度,也使得生活质量特征、区域差异和影响路径的识别更加细致,为理解农村生活质量提供了微观证据支持。

本文结构安排如下:第二部分介绍数据来源与指标体系构建方法;第三部分基于综合生活质量指数分析农村居民生活质量的总体水平、区域格局与主要影响因素;第四部分结合研究结果提出相应的政策建议。

2. 农村居民生活质量指标体系构建与测度

2.1. 数据来源

本文选取中国家庭追踪调查(China Family Panel Studies, CFPS)作为主要数据来源。CFPS 由北京大学中国社会科学调查中心自 2010 年起持续开展,覆盖全国 25 个省份,数据包括家庭、个人、社区与村级社会经济信息,具有样本量大、结构完备、时序连续、农村样本丰富等显著优势,能够系统刻画农村家庭在经济状况、公共服务、健康水平、教育资源、生活环境与主观满意度等多维度的生活状态[11]。

由于 CFPS 在家庭与个体两个层面同时提供客观指标与主观评价,特别适合构建“主客观融合”的农村生活质量评价体系。鉴于数据质量与可得性,本文选取最新可与家庭信息充分匹配的 CFPS 202 年数据作为主样本,并通过以下步骤对农村样本进行筛选:首先根据 CFPS 官方农村户籍划分,提取全部农村居民样本;其次,对关键变量缺失、主观评价空缺的样本进行剔除;最后,按居民个人为单位构建生活质量指标体系。最终获得 11949 个农村居民的有效样本。

2.2. 指标构建与测度

2.2.1. 农村居民生活质量指标体系

农村居民生活质量具有多维性与结构性特征,涵盖经济资源、健康状况、公共服务、社会参与、幸

福感等多个方面。为全面反映农村家庭在生产、生活、发展、保障等方面的真实状况,本文遵循多维性、可测性与情境适合性原则,构建由8个准则层、8个指标层组成的农村居民生活质量指标体系。首先,从多维性原则出发,以“经济、社会、健康、服务、主观福祉”五大核心领域为基准框架,进一步细化为经济生活、住房环境、健康状况、教育文化、社会参与、社会支持、公共服务治理及主观幸福八大维度,系统覆盖农村生活质量的客观条件与主观体验。其次,遵循可测性原则,所有指标均来自全国性家庭追踪调查(CFPS)可获取的数据,包含家庭、个人与村级公共服务三级信息,确保指标的可操作性、可复现性与统计稳定性。最后,基于中国农村特有的发展阶段与治理情境,本研究强调基础设施、医疗可及性、教育资源、数字融入、邻里互助等乡村特征变量,以增强指标体系对农村现实问题的适配度。在此基础上,形成如下农村居民生活质量综合指标体系表(见表1)。

Table 1. Rural residents' quality of life index system

表 1. 农村居民生活质量指标体系

维度	指标	变量名	方向	指标说明
经济能力	年收入	yincomea_a_1	+	反映家庭经济资源,是生活质量最重要的客观基础指标。
教育能力	受教育程度	edu_last	+	衡量人力资本水平,与就业能力和长期发展密切相关。
健康状况	自评健康	qp201	+	国际通行的整体健康度量,能反映身体功能状态与健康感知。
社会支持	是否获得他人帮助	qg1303	+	对应农村互助网络强度,是社会支持与信任的重要体现。
公共服务(教育)	学校满意度	qs701_b_2	+	反映教育服务质量,是农村家庭最核心的公服体验指标。
公共服务(医疗)	医疗点满意度	qp602	+	代表农村医疗服务条件与可及性。
数字参与	是否移动上网	qu201	+	反映农户的数字融入能力,是数字鸿沟和现代生活方式的重要指标。
主观满意度	对生活满意度	qn12012	+	衡量居民心理福祉,是主观生活质量的核心指标。

2.2.2. 指标赋权方法与综合生活质量指数构建

农村居民生活质量指标体系包含8个准则层和8个指标层,指标维度多、类型复杂,为避免主观赋权带来的偏差,本文采用熵权法和XGBoost模型对生活质量指数进行科学赋权与综合测度[12]。整个过程包括三步:指标标准化、熵权赋权、机器学习融合生成综合指数。

1. 熵权法赋权客观生活指数

(1) 指标标准化

由于各指标量纲不同,首先需采用区间标准化对指标进行无量纲化处理:

① 正向指标

$$X_{ij}^{*} = \frac{X_{ij} - \min(X_i)}{\max(X_i) - \min(X_i)}$$

② 负向指标

$$X_{ij}^* = 1 - \frac{X_{ij} - \min(X_i)}{\max(X_i) - \min(X_i)}$$

这样得到的所有指标均被转换到[0,1]区间。

(2) 计算指标的信息熵

熵反映指标的不确定性，变异越大，熵越小，其权重越高。

① 指标占比计算

$$p_{ij} = \frac{X_{ij}^*}{\sum_{j=1}^n X_{ij}^*}$$

② 信息熵计算

$$e_i = -k \sum_{j=1}^n p_{ij} \ln p_{ij}, k = \frac{1}{\ln n}$$

③ 差异度(冗余度)计算

$$d_i = 1 - e_i$$

差异度越大，该指标提供的信息量越大。

(3) 指标客观权重计算

$$w_i = \frac{d_i}{\sum_{i=1}^m d_i}$$

最终得到每一个指标的客观权重向量

$$w = \{w_1, w_2, \dots, w_{37}\}$$

(4) 构建客观生活质量指数 LQ_{obj}

对每一户农村家庭 j :

$$LQ_{obj,j} = \sum_{i=1}^{37} w_i \cdot X_{ij}^*$$

LQ_{obj} 是一个完全基于客观指标的生活质量指数。

2. 因子分析提取主观生活质量指数

为了体现居民对生活的主观体验，本文将幸福感、满意度、安全感、获得感等心理指标通过因子分析提取为主观生活质量指数：

$$LQ_{subj,j} = b_1 SWB1_j + b_2 SWB2_j + \dots + b_k SWBk_j$$

其中 b_k 为因子得分系数。

至此，得到了客观生活质量指数 LQ_{obj} 和主观生活质量指数 LQ_{subj} 。

3. 基于 XGBoost 的综合生活质量指数构建

为提高生活质量测度的精度，避免传统线性模型无法捕捉高维指标间非线性关系与交互效应的局限，本文采用 XGBoost 模型对客观生活质量指数 LQ_{obj} 、主观生活质量指数 LQ_{subj} 以及 8 个准则层的多维指标进行融合，从而构建农村居民综合生活质量指数[13]。具体方法如下。

XGBoost 模型通过多个梯度提升树学习生活质量的最优加权方式，能够在高维、复杂数据结构下实现非线性拟合。其基本形式为：

$$LQ_j = f(LQ_{obj,j}, LQ_{subj,j}, X_j)$$

其中, LQ_j 表示第 j 个农村家庭的综合生活质量指数; $LQ_{obj,j}$ 与 $LQ_{subj,j}$ 分别代表客观与主观生活质量指数; X_j 包含经济水平、健康状况、教育文化、住房条件、公共服务、数字融入、社会参与与社会支持等多维指标; $f(\cdot)$ 为 XGBoost 模型在训练过程中自动学习得到的非线性映射函数。模型以多维指标作为输入, 通过学习多个弱学习器的加权叠加方式生成最终预测值, 有效提升生活质量测度的准确性与鲁棒性。

为了保证模型的可靠性, 本文对 XGBoost 模型进行了规范的训练与验证流程: 首先, 将所有样本按 7:3 比例划分为训练集与测试集, 确保模型训练与验证均具有代表性。其次, 采用网格搜索(Grid Search)对 \max_depth 、 η 等核心超参数进行调优, 以获得最佳模型结构。接着, 使用均方根误差与平均绝对误差对模型进行拟合性能评估, 确保模型在训练集与测试集均具有较好的拟合精度。最后, 通过交叉验证、设置正则化参数(λ 、 α)以及进行多次训练等方式, 提高模型的泛化能力, 避免过拟合问题。

XGBoost 模型训练完成后, 将每个农村家庭的特征向量输入模型, 即可得到其生活质量预测值:

$$\widehat{LQ}_j = f(\cdot)$$

为便于不同家庭间的横向比较, 本文对预测值进行区间归一化处理, 将其缩放至 $[0|1]$ 区间:

$$LQ_j = \frac{\widehat{LQ}_j - \min(\widehat{LQ})}{\max(\widehat{LQ}) - \min(\widehat{LQ})}$$

最终得到的 LQ_j 即为本文所构建的农村居民综合生活质量指数。该指数不仅融合了客观与主观信息, 还充分利用了机器学习模型对复杂结构的刻画能力, 能够更准确、全面地反映农村居民生活质量的真实状况。

3. 农村居民生活质量特征及影响因素分析

3.1. 农村居民生活质量特征分析

3.1.1. 整体特征

根据综合生活质量指数的描述统计结果, 农村居民生活质量总体水平偏低, 呈现显著的右偏分布特征。样本量为 11,949 个, 指数均值为 0.1769, 中位数仅为 0.0649, 显著低于均值, 表明大多数农村居民处于较低生活质量水平。标准差为 0.1939, 显示不同群体之间生活质量差异较大。从分布形态看, 偏度系数为 1.132, 峰度为 2.754, 说明指数呈现明显的长右尾分布, 高生活质量群体比例较小, 而低生活质量群体占比显著。仅有少数家庭的指数接近 1, 且第 90 百分位仅为 0.5061, 进一步佐证高生活质量农村家庭数量有限。从分位数角度来看, 第 25 百分位数为 0.0562, 第 50 百分位数为 0.0649, 第 75 百分位数也仅为 0.3973, 表明 75% 的农村家庭生活指数低于 0.4, 生活质量结构呈现明显的“低位集中”特征。整体来看, 农村居民生活质量呈现“总体偏低、差距较大、两极不明显、右尾群体有限”的结构特征, 与我国城乡发展差异、区域发展不均衡和农村收入增长乏力的宏观环境高度一致[12]。

3.1.2. 区域比较

基于划分为东部、中部和西部三大区域的综合生活质量指数, 表 3 呈现了不同区域的生活质量水平比较结果。从均值来看, 东部地区(0.1784)、中部地区(0.1739)与西部地区(0.1820)的指数水平十分接近, 整体差异较小, 三大区域均值均低于 0.20, 反映出我国农村居民生活质量水平整体偏低且区域分布相对均衡。所有区域的中位数均集中在 0.0648 附近, 进一步说明低生活质量群体在全国范围内占比较大。从

波动情况看,各区域标准差在 0.189~0.197 区间,表明三个区域内部的生活质量差异均较显著,而区域间差异相对有限。东部西部的最大值分别接近 0.97 和 0.98,中部达到 1,这意味着三大区域均存在少量生活质量较高的家庭,但比例极小。

Table 2. Descriptive statistical analysis of comprehensive quality of life in the three major regions: eastern, central, and western
表 2. 东部、中部和西部三大区域的综合生活质量描述性统计分析

地区	Mean	SD	Median	Min	Max
东部	0.1784	0.1969	0.0649	0.0019	0.9723
中部	0.1739	0.1894	0.0648	0.0038	1.0000
西部	0.1820	0.1952	0.0649	0.0000	0.9753
总体	0.1785	0.1941	0.0649	0.0000	1.0000

为了进一步判断地区差异是否具有统计显著性,本文采用 ANOVA 方差分析。在 11,748 个样本上得到的 F 值为 1.67, p 值为 0.1886,未通过显著性检验($p > 0.1$),说明整体上东、中、西部农村地区的综合生活质量差异不显著。Bartlett 等方差齐性检验的 p 值为 0.055,接近 0.05 的显著性水平,提示不同区域内部差异可能存在轻微不均衡。进一步采用 Bonferroni 进行两两比较发现,东部与中部($p = 0.965$)、东部与西部($p = 1.000$)、中部与西部($p = 0.203$)均未表现出显著差异。

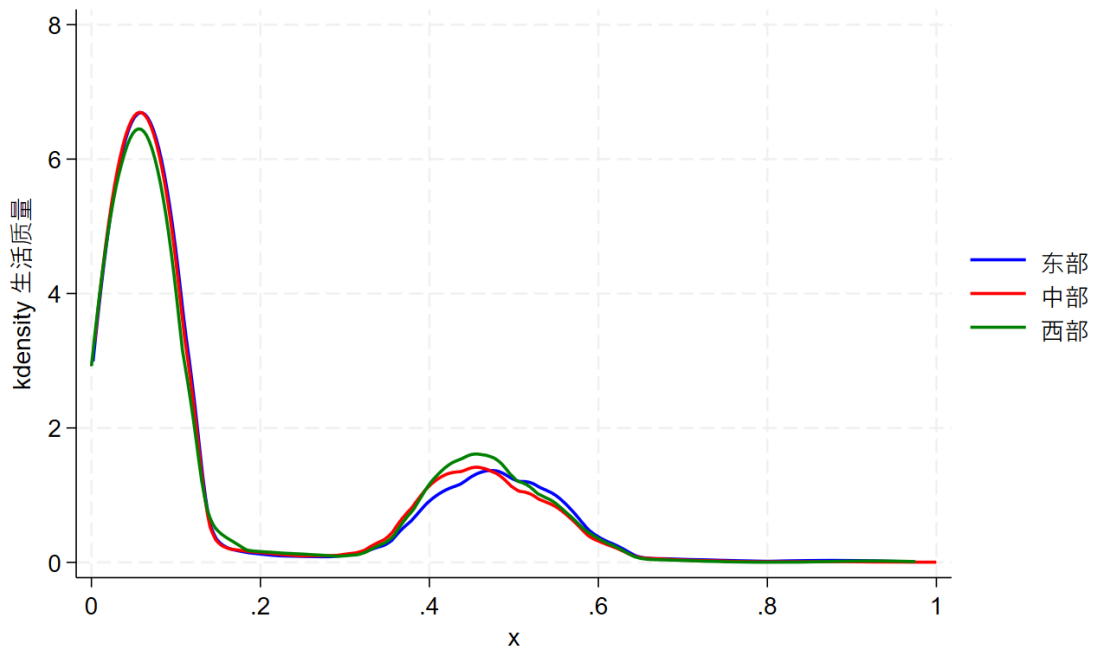


Figure 1. Kernel density map of rural residents' quality of life in the eastern, central and western regions
图 1. 东部、中部和西部农村居民生活质量核密度图

图 1 展示了东部、中部和西部农村地区综合生活质量指数的核密度分布。从曲线形态看,三大区域呈现出明显的“两峰结构”,即低生活质量群体集中分布在 0.05 附近,高一些的群体集中在 0.4~0.5 区间。这说明农村居民生活质量总体呈现典型的“双峰”特征,反映出生活质量提升存在“分层结构”。从区域比较来看,东部(蓝色)、中部(红色)和西部(绿色)三条曲线高度重叠,说明地区间生活质量差异相

对较小。三个区域的分布峰值位置接近，低生活质量峰均集中在 0 附近，高生活质量峰均位于约 0.45 左右，表明区域经济发展差异并未显著转化为农村居民生活质量的显著差距。在低端分布部分，中部地区的密度略高，说明中部农村低生活质量人口占比略大；而在高值区(0.4~0.6 附近)，西部地区的密度略高于中部与东部，意味着西部存在一小部分生活质量较高的家庭群体。尽管如此整体差异依然不显著，与 ANOVA 的统计结果一致。

3.1.3. 机制解释

针对本文发现“东、中、西部农村居民生活质量差异不显著”的结果，可从国家制度环境与社会结构变迁两个层面进行进一步解释。

首先，从国家政策层面来看，近年来我国持续推进以乡村振兴战略和区域协调发展战略为代表的一系列重大战略，通过财政转移支付、基本公共服务均等化、基础设施下沉等制度安排，在很大程度上弱化了区域间公共服务供给差距。尤其是在义务教育、基本医疗保障、农村低保等领域，全国范围内逐步形成“底线公平”的制度格局，使得不同区域农村居民在基础生活条件上的差异被显著压缩。因此，宏观经济发展水平的区域差距，并未完全传导至微观个体的生活质量层面。

其次，从社会变迁视角来看，人口流动与家庭结构变化也在重塑农村生活质量格局。一方面，伴随城乡要素流动加速，大量农村劳动力外出务工并通过汇款机制反哺家庭，形成“空间分离-经济补偿”的生活模式；另一方面，近年来出现的“人口回流”现象，使部分中西部地区通过返乡创业、电商发展等路径获得新的发展机会。这种“双向流动”在一定程度上对冲了区域发展差异，使不同地区农村家庭在收入来源、消费能力和生活方式上趋于收敛。

最后，需要强调的是，本文基于微观个体数据进行分析，结果显示“区域内部差异大于区域之间差异”。这意味着当前农村生活质量的不平等更多体现为个体层面的结构性差异(如家庭资源禀赋、社会支持网络等)，而非传统意义上的区域差距。

3.1.4. “三元结构”的深层机制解释

进一步地，SHAP 分析所揭示的“公共服务-社会支持-经济能力”三元结构，本质上反映了农村生活质量形成机制由“物质主导”向“结构与预期驱动”转变的过程。

首先，教育满意度对生活质量的贡献超过收入水平，体现出农村家庭对“代际流动”的高度重视。在资源约束条件下，教育不仅被视为公共服务，更被赋予“改变家庭命运”的功能。根据人力资本理论与社会流动理论，教育是农村居民突破阶层固化、实现向上流动的核心路径。因此，学校满意度不仅反映教育质量本身，更隐含了家庭对未来机会的预期，其对生活质量的贡献具有“预期效用放大效应”，从而在模型中表现出高于收入的边际贡献。

其次，社会支持网络的重要性，体现出农村社会仍具有典型的“熟人社会”特征。基于关系网络的互助机制，在风险分担、信息获取和情感支持等方面发挥着重要作用，能够在正式制度之外提供“非正式保障”，从而显著提升个体生活质量。这也解释了为何社会支持在模型中具有较高解释力。

再次，收入虽仍为基础性变量，但其边际作用相对弱化，说明农村生活质量已不再单纯由经济因素决定，而是呈现出明显的多维结构特征。

最后，数字参与变量贡献较低，则反映出当前农村数字化发展仍处于“接入阶段向应用阶段过渡”的过程。尽管基础设施不断完善，但数字技术尚未充分嵌入农村生产生活体系，其对生活质量的提升仍停留在潜在阶段，尚未形成显著的结构影响。

3.2. 农村居民生活质量影响因素分析

为深入识别农村居民生活质量的核心驱动因素，本文在构建 XGBoost 非线性融合模型的基础上，引

入 SHAP 方法对模型输出进行解释[14]。SHAP 通过计算每个变量对预测结果的边际贡献，揭示各指标在综合生活质量指数形成过程中的重要性与影响方向。该方法既继承了博弈论的可解释性优势，又兼具机器学习模型的高精度与稳定性，是当前指标体系评价中最具前沿性的解释工具。

3.2.1. 指标变量重要性排序

基于 XGBoost 模型输出的 Shapley 值贡献度，本文对 8 个指标的重要性进行了排序。结果表明，农村居民生活质量的关键驱动因素集中在就业稳定性与住房条件两个方面。其中，qg1303 (就业状态)与 qs701_b_2 (住房质量)贡献度最高，说明经济基础与居住设施仍是提升农村生活质量的核心。收入水平(yincomea_a_1)虽具有显著影响，但其相对重要性低于结构性因素。家庭结构变量(qu201)居于中等水平，而教育类指标(qn12012、edu_last)及主观类因素(qp201、qp602)权重偏低，反映农村地区生活质量更多由客观经济-物质条件决定，而教育与主观感受的影响具有滞后性与边际效用递减特征。

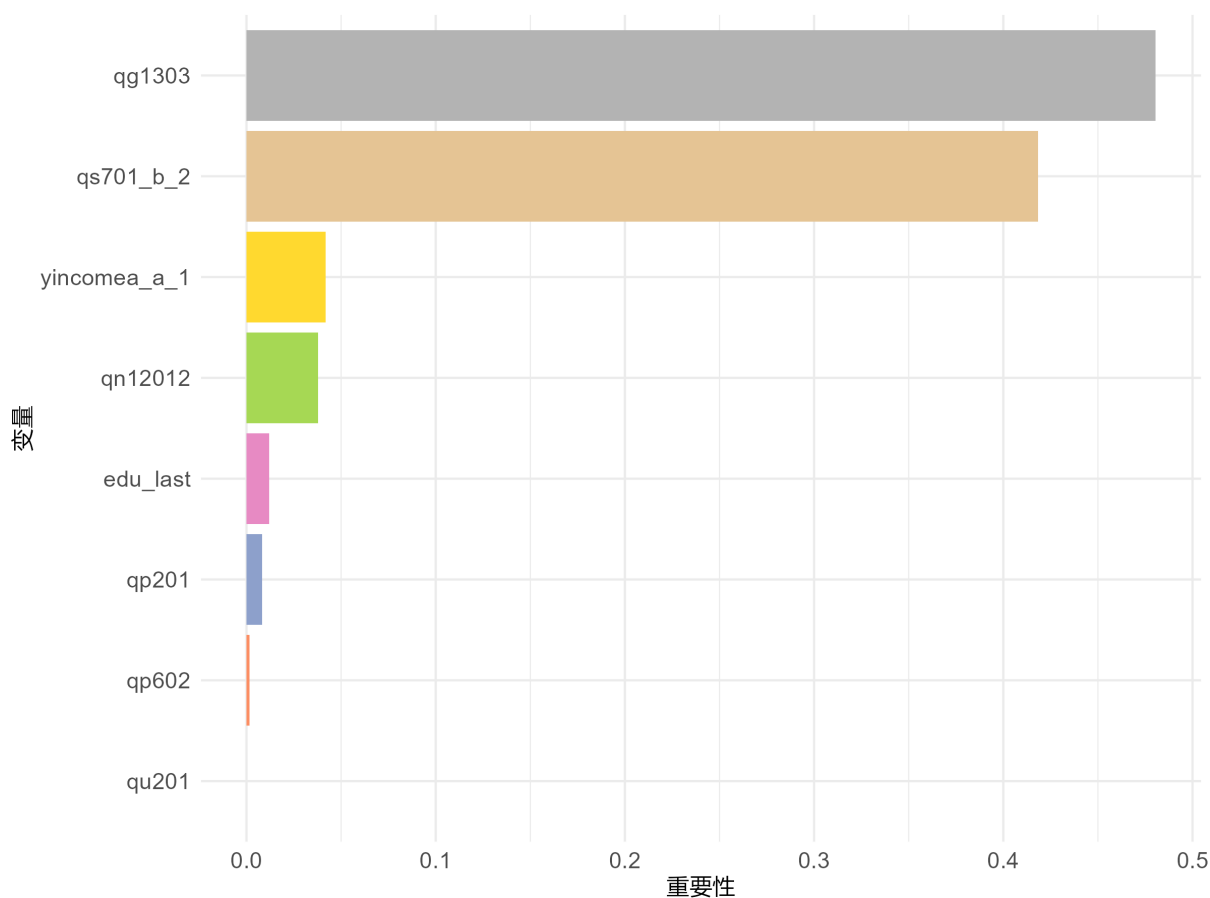


Figure 2. Bar chart of variable importance ranking

图 2. 变量重要性排序条形图

3.2.2. SHAP 蜜蜂图

基于 XGBoost 模型获得的 SHAP 蜜蜂图(见图 3)展示了八个最终指标对农村居民综合生活质量(LQ_total)的边际贡献及其影响分布特征。整体来看，各变量对模型输出的影响具有显著的正负方向和异质性，不同变量的贡献度差异明显。学校满意度(qs701_b_2)展现出最强的解释力，高满意度样本的 SHAP 值集中在正向区间，表明教育公共服务质量是推动农村生活质量提升的核心因素；这反映出农村家庭对

教育公平、教育可及性和学校办学水平的高度依赖，其重要性甚至超过家庭收入。社会支持变量(qg1303)同样表现出显著的正向影响，获得他人帮助的个体具有更高的预测生活质量，这说明农村社会资本与邻里互助网络仍是生活福祉的重要基础，体现出典型的“熟人社会”互惠机制。

家庭年收入(yincomea_a_1)对生活质量的边际贡献整体为正，收入越高，SHAP 值越正，进一步验证了经济资源仍然是改善生活质量的重要基础性因素，但其重要性略低于学校满意度与社会支持，说明农村生活质量具有明显的多维性结构。受教育程度(edu_last)同样呈现稳定的正向贡献，高教育水平促进就业质量与发展机会，从能力提升渠道改善生活质量。主观生活满意度(qn12012)虽对综合指数具有正向贡献，但其影响范围相对较窄，与公共服务与经济能力相比具有更弱的解释力，显示生活质量的主体结构主要由客观资源支持而非纯主观评价决定。

医疗点满意度(qp602)和自评健康(qp201)也呈现一定的正向作用，但总体边际贡献较弱，可能与农村地区医疗资源差异大、基层医疗能力有限有关。数字参与(qu201)影响最弱，其 SHAP 值多集中在零附近，表明农村地区的数字化使用虽已普及，但其对生活质量的提升尚未形成显著的结构作用，数字红利仍处于释放初期。总体而言，SHAP 蜜蜂图揭示了农村生活质量的主要驱动机制呈现“公共服务质量 - 社会支持网络 - 经济能力”三元结构，其中教育公共服务和社会资本的作用尤为突出，收入与教育能力作为基础资源发挥重要支撑作用，而医疗与数字参与的影响相对有限，体现出农村生活质量提升路径的结构特征。

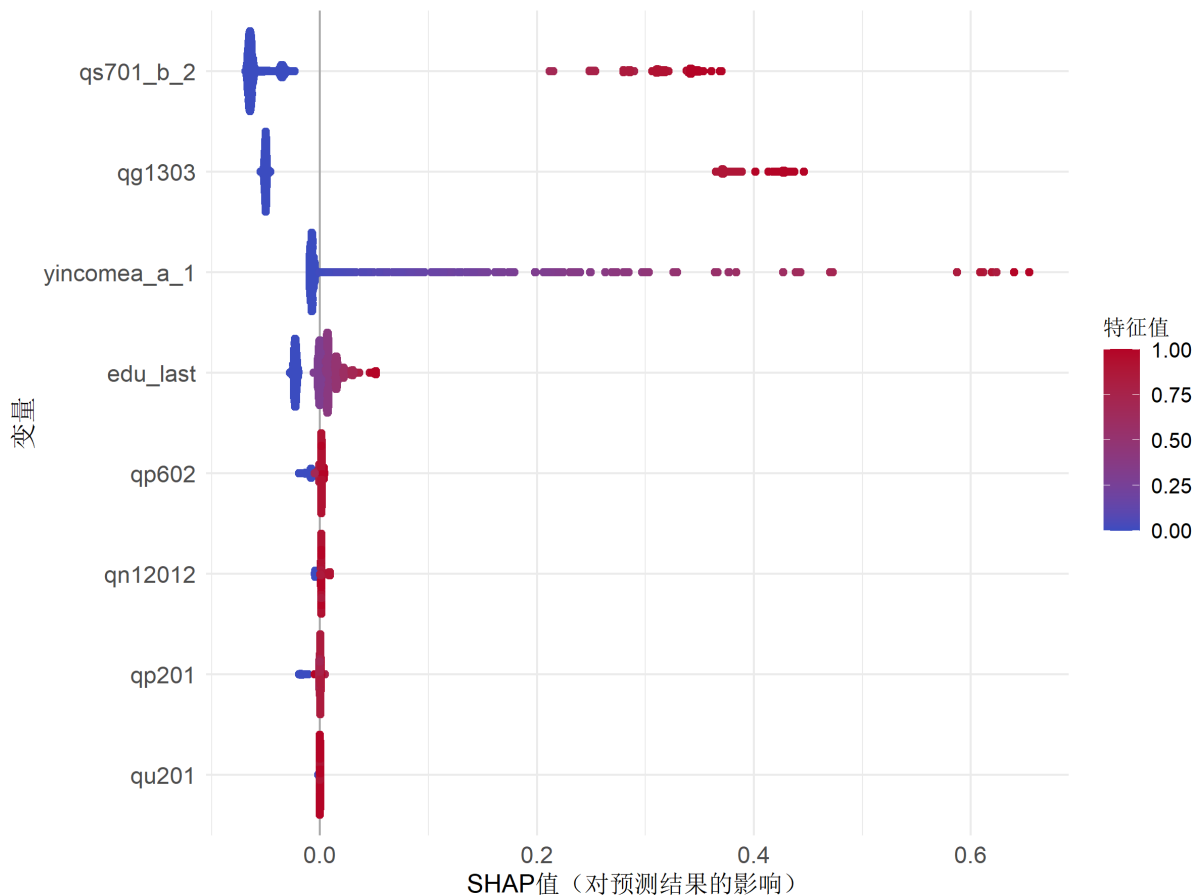


Figure 3. SHAP honey bee chart

图 3. SHAP 蜜蜂图

4. 结论与政策建议

4.1. 结论

基于 CFPS (2020)农村居民个体数据, 本文构建多维度生活质量指标体系, 采用熵权法、因子分析与 XGBoost 融合算法测度农村生活质量, 并利用 SHAP 解释方法识别关键驱动因素。主要研究结论如下: 第一, 农村居民生活质量整体水平偏低, 呈现明显“低位集中”特征。生活质量指数的均值虽有所改善, 但整体分布呈显著右偏, 中低生活质量群体占比大, 高生活质量群体比例有限, 说明农村生活质量具有总体偏低、改善缓慢的结构性特征。第二, 东、中、西三大区域生活质量差异并不显著。核密度分布与 ANOVA 检验均显示, 三大区域生活质量曲线高度重叠, 区域差异缺乏统计显著性, 说明宏观区域经济发展差距尚未充分传导至农村家庭层面的生活质量表现。地区内部差异大于地区之间差异。

第三, 农村生活质量的影响机制呈现“公共服务 - 社会支持 - 经济能力”三元结构 SHAP 分解结果表明, 学校满意度与社会支持对生活质量的提升作用最为显著, 公共服务质量与互助网络是关键因素; 收入与教育水平作为基础资源发挥重要支撑作用; 健康状况、医疗体验和数字参与的边际贡献相对较弱, 说明农村生活质量更多由结构性资源条件决定, 而数字融入与主观感受具有一定滞后性。

4.2. 政策建议

基于本文实证结果所揭示的影响机制, 应从“公共服务质量 - 社会支持网络 - 经济能力 - 数字转化”四个维度, 有针对性地优化农村生活质量提升路径[15][16]。

首先, 应从“主观体验导向”出发, 重点提升农村教育公共服务的“软实力”。研究表明, 学校满意度对生活质量的贡献度最高, 这意味着农村居民对教育质量的评价不仅取决于硬件条件, 更取决于教育体验[16]。因此, 在继续推进教育资源均衡配置的基础上, 应更加注重师生关系优化、课堂质量提升与课程多样化建设, 增强学生的学习获得感; 同时, 通过完善家校沟通机制、强化家长参与, 提升家庭对教育过程的认同感, 使教育公共服务从“供给导向”转向“体验导向”[17]。

其次, 应强化农村社会支持体系建设, 巩固生活质量的“关系型基础”[18][19]。针对社会支持变量的重要作用, 应鼓励村庄内部发展互助组织、志愿服务与社区活动, 增强社会联结; 同时, 针对老年人、留守群体等弱势群体, 完善基层照护与心理支持服务体系, 提升社会支持的可及性与稳定性, 从而形成“制度保障 + 社会网络”双重支撑结构[20]。

再次, 应持续夯实农村居民经济能力基础, 但更加注重结构优化而非单纯增收[21][22]。结合收入变量的基础性作用, 应通过发展乡村特色产业、农村电商与本地就业机会, 提升收入稳定性与可持续性[23][24]; 同时, 推动人力资本投资与技能培训, 提高农村居民的长期发展能力, 使经济能力在生活质量提升中发挥“基础支撑而非唯一驱动”的作用[25]。

最后, 应推动数字参与从“接入数量”向“应用质量”转化。针对数字参与贡献较弱的发现, 政策重点不应仅停留在基础设施建设层面, 而应加强数字技能培训与应用场景开发[26]。例如, 围绕农业生产、农产品销售、远程医疗与在线教育等领域, 开发贴合农村实际需求的数字应用; 同时, 重点提升老年群体与低学历群体的数字素养, 使数字技术真正转化为提升生活质量的有效工具[27], 实现从“有网可用”向“用网有效”的转变。

参考文献

- [1] Pan, D., Yu, Y. and Ji, K. (2024) The Impact of Rural Living Environment Improvement Programs on the Subjective Well-Being of Rural Residents in China. *Humanities and Social Sciences Communications*, **11**, 1-10. <https://doi.org/10.1057/s41599-024-03052-y>

- [2] 王一臣, 刘畅. 数字赋能农村人居环境改善的路径研究——以江苏省为例[J]. 中国林业经济, 2024(3): 31-35.
- [3] 何田. 基于 fsQCA 分析的农村老年人生活质量感知及影响因素研究[J]. 老龄化研究, 2024, 11(2): 525-536.
- [4] Meng, Y., Yu, R., Bai, H. and Han, J. (2024) Evidence from the China Family Panel Studies Survey on the Effect of Integrating the Basic Medical Insurance System for Urban and Rural Residents on the Health Equity of Residents: Difference-In-Differences Analysis. *JMIR Public Health and Surveillance*, **10**, e50622. <https://doi.org/10.2196/50622>
- [5] Xia, Y., Wang, G. and Yang, F. (2024) A Nationwide Study of the Impact of Social Quality Factors on Life Satisfaction among Older Adults in Rural China. *Scientific Reports*, **14**, Article No. 11614. <https://doi.org/10.1038/s41598-024-61398-4>
- [6] 张展, 周富春. 乡村振兴视角下的农村人居环境优化路径[J]. 可持续发展, 2023, 13(4): 1088-1094.
- [7] Li, J. (2023) A Scoping Review of Internal Migration and Left-Behind Children's Well-Being in China. arXiv: 2305.04348.
- [8] 丁可可, 马正兵. 数字普惠金融对农村家庭生活质量的的影响研究[J]. 重庆第二师范学院学报, 2024, 37(4): 18-26, 127.
- [9] Chen, K., Wang, W., Qiu, J., Guo, W., Du, J., Gao, B., et al. (2024) Housing Conditions, Cooking Fuels, and Health-Related Quality of Life among Rural Middle-Aged and Elderly in Northwest China: A Ten-Year Balanced Panel Study. *Preventive Medicine Reports*, **37**, Article ID: 102563. <https://doi.org/10.1016/j.pmedr.2023.102563>
- [10] Wei, X. and Li, X. (2024) Internet Use and Rural Income Inequality: Evidence from China. *Economic Research-Ekonomska Istraživanja*, **37**, Article 2186914. <https://doi.org/10.1080/1331677x.2023.2186914>
- [11] 北京大学中国社会科学调查中心. 中国家庭追踪调查(CFPS)数据说明[EB/OL]. <https://www.iss.pku.edu.cn/cfps/>, 2026-03-14.
- [12] 田果, 向东, 余济云. 基于 AHP-熵权法的国家森林公园康养资源评价[J]. 绿色科技, 2023, 25(21): 148-156.
- [13] 邵祥东, 高宇航. 数字乡村建设与农村居民生活质量——基于新结构经济学视角[J]. 经济论坛, 2024(9): 101-116.
- [14] Hernandez, J.I., van Cranenburgh, S., de Bruin, M., Stok, M. and Mouter, N. (2024) Using XGBoost and SHAP to Explain Citizens' Differences in Policy Support for Reimposing COVID-19 Measures in the Netherlands. *Quality & Quantity*, **59**, 381-409.
- [15] 向春玲. 发挥农民主体作用提升乡村治理效能[EB/OL]. 求是理论网, 2024-12-25. https://www.cccps.gov.cn/dxsy/202412/t20241225_165697.shtml, 2026-04-14.
- [16] Wan, L. and Di, X. (2025) Study on the Influencing Factors of the Demand of Rural Older Adults in China for Elderly Care Services. *Healthcare*, **13**, Article 1086. <https://doi.org/10.3390/healthcare13091086>
- [17] 蒋雯静. 长期护理保险制度是否促进了居民幸福感?——基于 fsQCA 的试点案例分析[J]. 西华大学学报(哲学社会科学版), 2024, 43(4): 86-98.
- [18] 赵周华, 姚茗曦, 宋婷. 家庭结构变动、社会支持与老年人心理健康——基于 CHARLS 数据的研究[J]. 老龄科学研究, 2025, 13(2): 60-78.
- [19] 杨月琦. 乡村振兴背景下新质生产力对农村居民生活质量的影响研究[J]. 当代农村财经, 2025(1): 31-35.
- [20] 宋东阳. 数字技术驱动农村经济高质量发展[J]. 农村经济与科技, 2024, 35(24): 34-37.
- [21] Guo, M., Zhang, X., Zhuang, Y., Chen, J., Wang, P. and Gao, Z. (2023) Exploring the Intersection of Complex Aesthetics and Generative AI for Promoting Cultural Creativity in Rural China after the Post-Pandemic Era. In: Zhao, F. and Miao, D., Eds., *AI-Generated Content*, Springer, 313-331. https://doi.org/10.1007/978-981-99-7587-7_27
- [22] 刘伟江, 刘冰琪. 农村数字基础设施建设赋能乡村振兴的路径——基于数字化与现代农业产业融合的视角[J]. 山西财经大学学报, 2024, 46(10): 72-88.
- [23] 王淳. 数字技术赋能乡村文化产业高质量发展的作用、困境与优化路径[J]. 农村经济与科技, 2024, 35(20): 41-45, 66.
- [24] 向运华, 程鹏, 黄蓉. 数字素养对农村居民生活质量的影响[J]. 东南学术, 2024(5): 136-148, 248.
- [25] 赵一凡, 刘子玉. 电子商务发展对农村居民生活满意度的影响——基于 2020 年中国乡村振兴综合调查(CRRS)数据[J]. 经济与管理, 2024, 38(4): 9-18.
- [26] 冯伯豪, 王晓红. 数字农业助推乡村振兴的影响机制及政策建议[J]. 西安财经大学学报, 2024, 37(1): 119-129.
- [27] 国家统计局“新时期农民生活质量研究”课题组, 王贵荣, 夏欢. 我国农村居民生活质量评价研究[J]. 调研世界, 2025(1): 4-13.