



# Construction of Depression Prediction Model Based on Machine Learning and Its Interpretability

Juan Wang, Man Cui, Miao Deng, Yanshuai Fan, Zhiguang Ping\*

College of Public Health, Zhengzhou University, Zhengzhou, China

## Email address:

wjuanzb@163.com (Juan Wang), cuimana@163.com (Man Cui), ymsyz8557@163.com (Miao Deng), fyszzu@163.com (Yanshuai Fan), pingzhg@zzu.edu.cn (Zhiguang Ping)

\*Corresponding author

## To cite this article:

Jang Wang, Man Cui, Miao Deng, Yanshuai Fan, Zhiguang Ping. (2025). Construction of Depression Prediction Model Based on Machine Learning and Its Interpretability. *Science Discovery*, 13(2), 25-32. <https://doi.org/10.11648/j.sd.20251302.12>

**Received:** 10 March 2025; **Accepted:** 8 April 2025; **Published:** 14 April 2025

**Abstract:** *Objectives:* The aim of this study was to construct depression prediction models based on machine learning algorithms, compared the performance of different machine learning models on depression risk prediction, and interpreted the model. *Methods:* A total of 2573 participants from the CHARLS database. LASSO and stepwise regression were used to screen for variables. The dataset is randomly divided into training set, validation set and test set according to 6:2:2. SMOTE resampling was used to balance the training set when fitted the model. Nine machine learning algorithms were used to construct the prediction model, including Decision Tree (DT), Random Forest (RF), Extreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (LightGBM), Elastic Network Regression (Enet), Support Vector Machine (SVM), Logistic Regression, Multilayer Perceptron (MLP), and K-Nearest Neighbor (KNN). The prediction ability of each machine learning classifier was evaluated on the test set according to the evaluation index, and the "optimal" model of this study was selected. Subsequently, SHapley Additive exPlanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME) were used to analyze the interpretability of the optimal model. *Results:* The XGBoost model predicted the best performance among the 9 models. Its AUC value reached 0.908 and the clinical net benefit is the highest. The Delong test showed that there was a significant difference between the ROC curves of XGBoost and the other models ( $P < 0.05$ ). The global interpretation based on SHAP showed that life satisfaction, self-rated health status, sleep duration, and cognitive score were inversely proportional to the SHAP value. Female, rural residents, body aches and pains in any area, non-retirement, and limited Instrumental Activities of Daily Living (IADL) have a positive effect on depression. The local interpretation diagram based on SHAP and LIME showed the personalized risk prediction of a single sample. *Conclusions:* Machine learning models are an effectively tool for predict the risk of depression. The use of SHapley Additive exPlanations and Local Interpretable Model-agnostic Explanations can maximize the clinical advantages of machine learning, which is helpful to predict or detect patients at high risk of depression as early as possible, and to take comprehensive evaluation and early prevention and treatment of depression.

**Keywords:** Machine Learning, Depression, Predictive Models, SHapley Additive exPlanations, Local Interpretable Model-agnostic Explanations

---

## 基于机器学习的抑郁症预测模型构建及其可解释性研究

王娟, 崔曼, 邓淼, 樊炎帅, 平智广\*

郑州大学公共卫生学院, 郑州, 中国

## 邮箱

wjuanzb@163.com (王娟), cuimana@163.com (崔曼), ymsyz8557@163.com (邓淼), fyszzu@163.com (樊炎帅), pingzhg@zzu.edu.cn (平智广)

**摘要:** 目的: 旨在利用机器学习算法构建抑郁症预测模型, 比较不同机器学习模型对抑郁症风险预测的性能, 并进行模型解释。方法: 共纳入2573名来自CHARLS数据库中的研究对象。采用LASSO和逐步回归筛选变量。数据集按6:2:2

随机分为训练集、验证集和测试集。拟合模型时采用SMOTE重采样来平衡训练集。采用目前常用的9种机器学习算法构建预测模型，分别为决策树（DT）、随机森林（RF）、极限梯度提升（XGBoost）、轻量级梯度提升（LightGBM）、弹性网络回归（Enet）、支持向量机（SVM）、逻辑回归（Logistic）、多层感知机（MLP）、K最邻近算法（KNN）。根据评价指标在测试集上评估每个机器学习分类器的预测能力，选择本研究的“最优”模型。随后采用Shapley加性解释（SHAP）和局部代理模型（LIME）对最优模型进行可解释性分析。结果：经比较，9种机器学习模型中XGBoost表现最佳，其AUC值达到0.908，并且临床净收益最高。Delong检验显示XGBoost与其余各模型ROC曲线间的差异有统计学意义（ $P<0.05$ ）。基于SHAP的全局解释表明，生活满意度、自评健康状况，睡眠时长，认知评分等均与SHAP值成反比。女性、居住地为农村、有任何部位的身体疼痛、未退休以及IADL受限等都对抑郁有正向作用。基于SHAP与LIME的局部解释图对单个样本进行了个性化风险预测展示。结论：机器学习模型是预测抑郁症风险的有效工具。利用Shapley加性解释和局部代理模型可以最大限度发挥机器学习的临床优势，有助于尽早预测或发现抑郁症高危患者，对其采取全面评估和早期防治。

**关键词：**机器学习，预测模型，抑郁症，Shapley加性解释，局部代理模型

## 1. 引言

抑郁症是一种常见的心理疾病，据世界卫生组织统计，约有3.5亿人患有抑郁症[1]。2020年，抑郁症造成的全球疾病负担增加约5.7%，成为仅次于缺血性心脏病的第二大常见疾病[2]。已经成为全球公共卫生的主要问题之一[3]。抑郁症大多通过医患沟通和填写抑郁量表来诊断和筛查。但世界上几乎有一半的人口生活在每10万人中只有两名精神科医生的国家[1]。且大多数患者对抑郁症的认识不足，不能及时得到诊断。此外，患者在自我报告时往往受主观意识影响较大，导致症状描述模棱两可，临床医生无法做出客观正确的诊断[4, 5]。有研究表明，抑郁症在很多国家都存在大范围的漏诊，从而错失了对高危人群的早期预防和干预，使得抑郁症更加严重[6-8]。目前，亟需一种更灵敏、客观、适用性更广、预测性更强的抑郁症筛查工具，帮助识别抑郁症高危人群，进行早期干预和治疗。

以往的研究大多使用传统统计方法来推断抑郁与特定变量之间的关系[9, 10]。然而，传统统计方法由于假定变量之间存在线性关系，因此在表现现实世界的复杂性和预测未来数据方面存在局限性[11, 12]。而机器学习（Machine Learning, ML）则不同，它侧重于发现特定数据集中隐藏的相互作用，从而进行预测[9, 10, 13]，其独特的优势在于能够从大量复杂的数据中选择最合适的算法[14]。ML方法可以通过预测疾病风险来促进早期检测[15-17]。利用ML的强大预测能力，有可能开发出在某些情况下优于传统统计建模的预测工具，从而更好地预测抑郁症的发生。因此，本研究旨在利用ML算法构建抑郁症预测模型，探讨影响抑郁症的关键因素，比较不同ML模型的预测的性能，并采用Shapley加性解释（SHapley Additive exPlanations, SHAP）和局部代理模型（Local Interpretable Model-agnostic Explanations, LIME）进行模型解释，为ML算法更好地应用于抑郁症的临床诊断及预测提供理论依据。

## 2. 对象与方法

### 2.1. 研究对象及变量

本研究的数据来源于北京大学国家发展研究院于2023年11月16日发布的2020年中国健康与养老服务纵向

调查（CHARLS）数据库。根据以往关于抑郁症的相关研究并结合CHARLS数据库的可用信息，共提取出32个潜在预测变量，将纳入研究的变量初步分为以下4类：

- 1) 人口统计学特征变量，包括性别、年龄、户籍、受教育程度、退休状况、居住地、婚姻状况、生活满意度和自评健康状况；
- 2) 健康状况变量，包括吸烟、饮酒、睡眠时间、身体疼痛、基本（Basic Activities of Daily Living, BADL）和工具性（Instrumental Activities of Daily Living, IADL）日常生活活动、社交活动、认知评分；
- 3) 家庭状况变量，包括家庭年消费、人均年消费、父母对子女的经济支持、子女的数量、子女对父母的经济支持以及家庭年收入；
- 4) 疾病史，包括高血压、血脂异常、糖尿病/血糖升高、心脏病、慢性肺病、肝脏疾病、胃部/消化系统疾病、肾脏疾病以及风湿/关节炎。

本研究的结局变量为抑郁症。采用Radloff[18]开发的CESD-10量表进行评估，该量表在中国人中具有良好的心理测量特性[19]。总分在0-30分之间，国际上通常使用15分作为是否具有抑郁症的分界点，当CES-D10<15时表示心理症状良好，当CES-D10≥15时，表示具有临床意义的抑郁症状。

### 2.2. 数据预处理

本研究分别采用LASSO和逐步回归两种分析方法筛选变量。筛选后的变量纳入模型构建。将处理后的数据按6:2:2的比例随机分为模型训练集、验证集和测试集。使用训练集来拟合模型，验证集来超参数调优，测试集来评估模型的性能和泛化能力。在拟合模型时采用合成少数过采样技术（Synthetic Minority Over-Sampling Technique, SMOTE）方法生成与原样本相似但略有差异的合成样本来平衡训练集。将数据中的非数值型分类变量（如“否”，“是”）用0、1等数字表示。

### 2.3. 模型构建与评估

在ML预测模型构建过程中采用5折交叉验证，模型参数的选择均采用网格搜索的超参数寻优。采用目前常用的9种ML算法构建预测模型，分别为决策树（DT）、随机森林（RF）、极限梯度提升算法（XGBoost）、轻量级梯度提升机（LightGBM）、弹性网络回归（Enet）、支持

向量机(SVM)、逻辑回归(Logistic)、多层感知机(MLP)、K最邻近算法(KNN)。在测试集上评估每个ML分类器的预测能力,并计算了接受者操作特征下面积(AUC)以及相应的灵敏度、特异度、准确率、精确率、F1评分和Kappa值等评价指标。此外,还进行了决策曲线分析(DCA)。结合ML评估结果,选择本研究的“最优”模型。随后采用SHAP和局部代理模型LIME对最优模型进行可解释性分析。

2.4. 统计分析

本研究使用Excel 2021和SPSS 27.0管理数据,统计分析采用R软件(版本4.2.1)实现,连续变量以均数±标准差中表示,组间比较采用t检验或Mann-Whitney U检验。分类变量以频数(百分比)表示,组间比较采用 $\chi^2$ 检验。

采用Delong检验来评估不同模型之间AUC值的差异。所有检验均为双侧检验, $P<0.05$ 为差异有统计学意义。

3. 结果

3.1. 变量筛选

首先使用LASSO筛选潜在预测变量。采用10折交叉验证确定最佳惩罚参数 $\lambda$ ,二项偏差曲线与 $\lambda$ 的关系见图1。选择代表最小二项偏差的 $\lambda$ 值,此时的 $\lambda$ 值0.0079,对应的自变量个数为16。之后采用向后逐步回归法进一步筛选,结果显示至第4步时AIC值达到最小,之后再剔除任一个变量AIC值都会增大,此时对应的特征变量数为12。不同特征人群中抑郁和非抑郁的差异比较见表1。

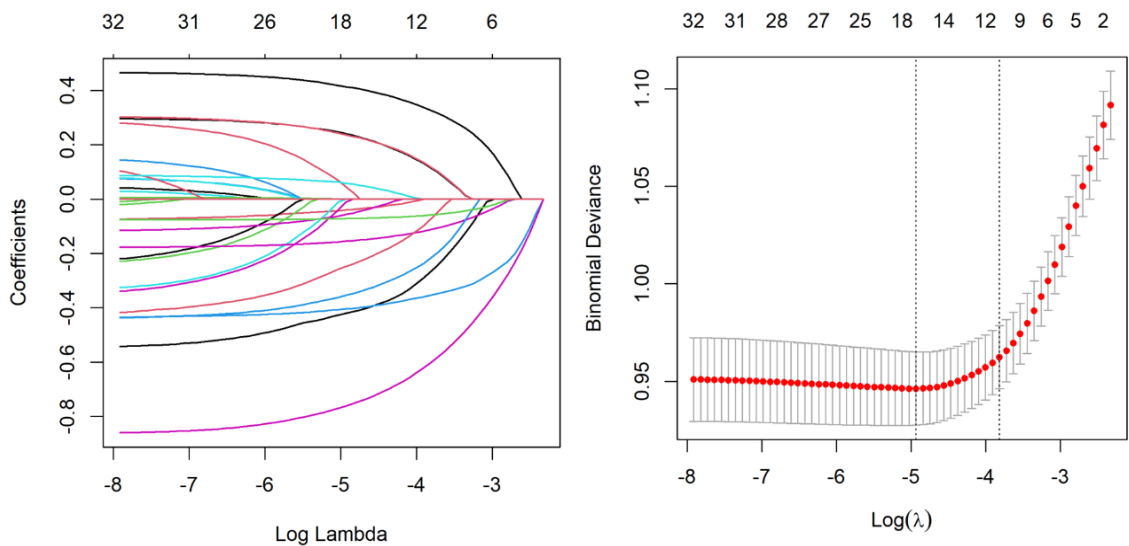


图1 LASSO回归变量筛选示意图。

表1 不同特征人群抑郁和非抑郁组间比较。

特征	分组	抑郁 (n=607)	非抑郁 (n=1966)	t/ $\chi^2$	P
性别	女	359 (59.1)	872 (44.4)	40.65	<0.001
	男	248 (40.9)	1094 (55.6)		
退休状况	否	542 (89.3)	1603 (81.5)	20.12	<0.001
	是	65 (10.7)	363 (18.5)		
家庭住址	农村	412 (67.9)	1100 (56.0)	27.21	<0.001
	城镇	195 (32.1)	866 (44.0)		
自评健康状况	差	91 (15.0)	103 (5.24)	13.05	<0.001
	一般	356 (58.6)	887 (45.1)		
生活满意度	好	160 (26.4)	976 (49.6)	19.98	<0.001
	不满意	110 (18.1)	62 (3.15)		
	一般	350 (57.7)	1098 (55.8)		
身体疼痛	满意	147 (24.2)	806 (41.0)	93.99	<0.001
	否	265 (43.7)	1291 (65.7)		
	是	342 (56.3)	675 (34.3)		
BADL	不受限	524 (86.3)	1865 (94.9)	50.91	<0.001
	受限	83 (13.7)	101 (5.14)		
IADL	不受限	533 (87.8)	1879 (95.6)	47.69	<0.001
	受限	74 (12.2)	87 (4.43)		
睡眠时长		5.96±1.68	6.56 ±1.45	8.49	<0.001
认知评分		12.33±3.27	13.28±3.13	7.84	<0.001
子女数量		2.24±1.06	2.08 ±1.06	-3.35	<0.001
经济支持 (千元)		4.95±1.47	6.04±2.31	2.28	0.013

注: BADL: 基本日常生活活动; IADL: 工具性日常生活活动;

3.2. 模型预测性能评估

在预测抑郁症时通过约登指数（灵敏度+特异度-1）的最大值来确定最佳截断值，以AUC值作为主要评价指标，且通过Delong检验评估不同模型间AUC值差异的显著性。以灵敏度、特异度、准确率、精确率、F1评分和Kappa值为次要辅助评价指标。比较9种ML算法构建的模型预测价值，XGBoost表现最佳，其AUC值达到0.908，其余各指

标也相对较高（表2）。各模型ROC曲线见图2。Delong检验对XGBoost与其余各预测模型ROC曲线间的差异进行检验，各模型间均有统计学差异（ $P<0.05$ ）。决策曲线显示（图2），所有模型在一定阈值概率范围内均具有一定的临床净收益，其中XGBoost模型的净收益明显优于其他预测模型，这也证明了XGBoost模型具有最佳的临床适用性。

表2 预测模型的性能比较。

模型	AUC	准确率	灵敏度	特异度	精确率	F1	Kappa
Logistic	0.773	0.701	0.684	0.718	0.708	0.696	0.402
Enet	0.738	0.672	0.649	0.695	0.680	0.664	0.344
DT	0.780	0.738	0.720	0.756	0.747	0.733	0.476
RF	0.901	0.827	0.880	0.756	0.786	0.823	0.621
XGBoost	0.908	0.810	0.898	0.740	0.818	0.838	0.654
SVM	0.881	0.798	0.766	0.830	0.772	0.791	0.595
MLP	0.788	0.714	0.634	0.794	0.755	0.689	0.427
LightGBM	0.821	0.726	0.763	0.690	0.711	0.736	0.453
KNN	0.801	0.730	0.710	0.751	0.740	0.725	0.461

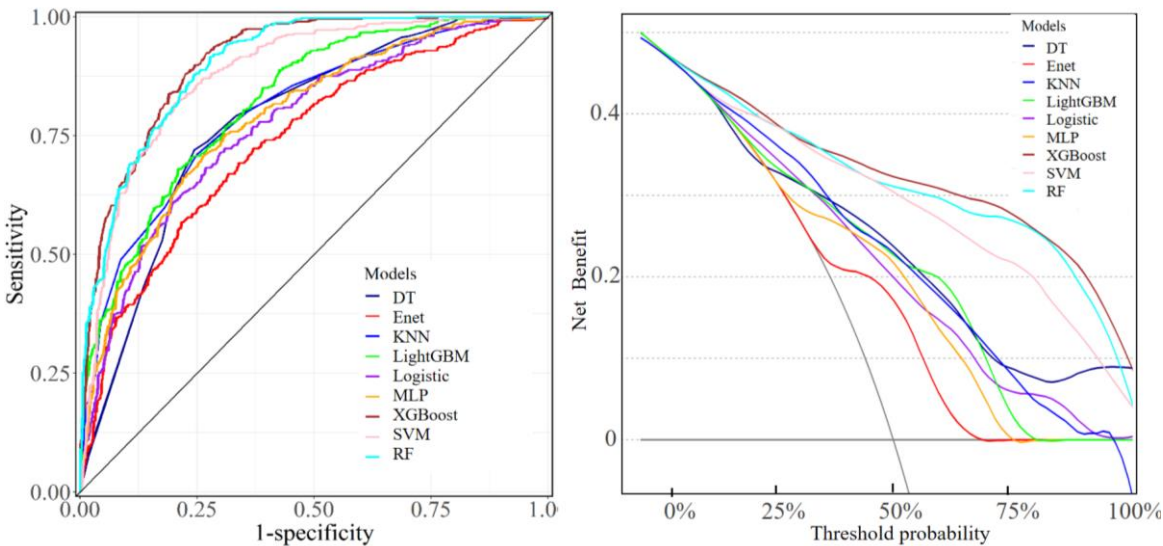


图2 预测模型的ROC曲线和决策曲线。

3.3. 最佳模型的解释

本研究选择XGBoost模型作为抑郁症预测的最优模型。采用SHAP和LIME对XGBoost模型进行可解释性分析。图3为SHAP全局解释图，直观地呈现了模型中各个变量对预测结果的影响程度。可以发现图中生活满意度的样本分布最宽，且与SHAP值成反比。同样地，自评健康状况，睡眠时长，认知评分均与SHAP值成反比。除此之外，女性、居住地为农村、有任何部位的身体疼痛、未退休以及

IADL受限等都对抑郁有正向影响。SHAP依赖图显示了分类变量生活满意度和性别以及连续变量睡眠时长和认知评分对抑郁症的影响情况（图4）。图中可以看出生活满意度越低，SHAP值越高。相比男性，女性SHAP值更高。连续变量睡眠时长和认知评分与预测结果的线性关系相对较明确，和SHAP值为反比关系。依赖图还显示了两个变量间的相关作用，如当生活满意度一般时，居住地为城镇的个体更容易出现抑郁症状。而生活满意度较高时，居住地为农村的个体更容易出现抑郁症状。

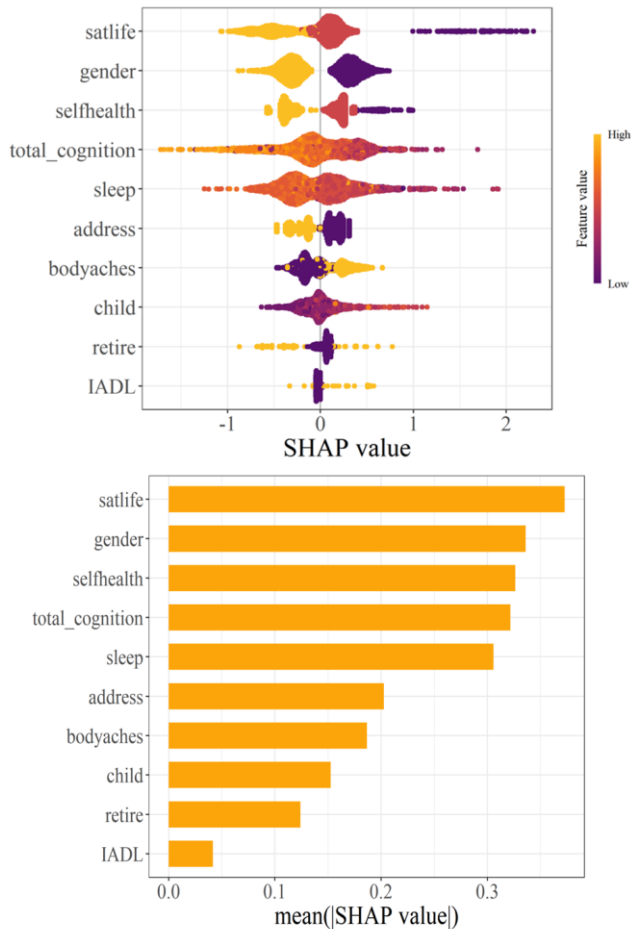


图3 基于SHAP的概括图和特征重要性排序图。

注: satlife: 生活满意度; gender: 性别; selfhealth: 自评健康状况; total\_cognition: 认知评分; sleep: 睡眠时长; address: 居住地; bodyaches: 身体疼痛; child: 子女数量; retire: 退休状况; IADL: 工具性日常生活活动;

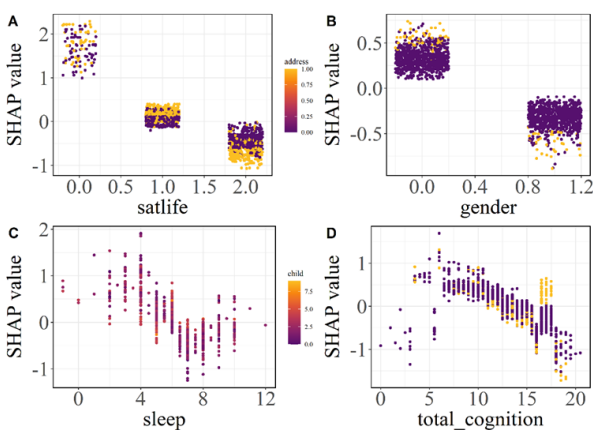


图4 SHAP依赖图。

注: satlife: 生活满意度; gender: 性别; ; total\_cognition: 认知评分; sleep: 睡眠时长; address: 居住地; IADL: 工具性日常生活活动; child: 子女数量; retire: 退休状况

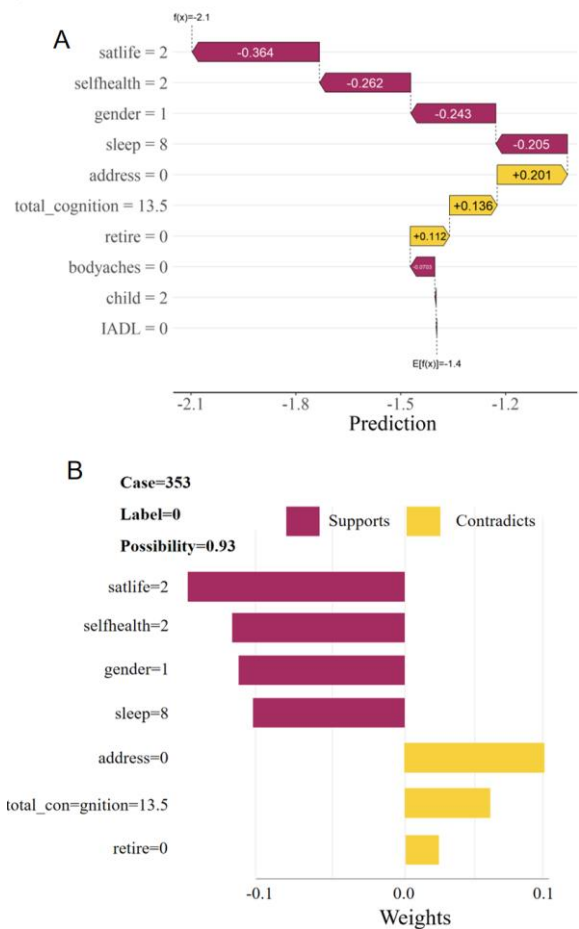


图5 基于SHAP和LIME的局部解释图。

注: satlife: 生活满意度; gender: 性别; selfhealth: 自评健康状况; total\_cognition: 认知评分; sleep: 睡眠时长; address: 居住地; bodyaches: 身体疼痛; child: 子女数量; retire: 退休状况; IADL: 工具性日常生活活动;

SHAP和LIME对单个样本进行局部解释分析, 结果如图5, 随机选择样本中的一个样本进行局部解释的个性化风险预测展示。基于SHAP的个体解释图(图5A)显示353号样本的最终预测值为-2.1, 低于平均水平-1.4, 其中生活满意度(满意)、自评健康状况(良好)、性别(男)、睡眠时长(8h)、身体疼痛(无)、子女数量(2)以及IADL(不受限)等均对预测结果产生了负向影响, 其影响值分别为0.364、0.262、0.243、0.205、0.071和接近0, 而居住地(农村)、认知评分(13.5)以及未退休对预测结果产生了正向影响, 其影响值分别为0.201、0.136和0.112。LIME解释图(图5B)表示第353号样本不出现抑郁症状的预测概率为93%。样本条形图显示, 该样本生活满意度(满意)、自评健康状况(良好)、性别(男)、睡眠时长(8h)以及未退休等特征对抑郁症的预测分类结果产生了负向作用。其余样本特征对预测结果产生正向作用。

## 4. 讨论

之前已经有研究构建了一些用于预测抑郁症的模型。其中, 有一部分使用传统统计学模型[20-22]。另一些则采用单一ML算法[23-25], 只有少数研究对多种ML算法进行比较[26, 27]。本研究基于9种常见的ML算法构建模型,

比较不同ML模型对抑郁症风险预测的性能，并对最优模型进行SHAP和LIME可解释分析。经对比，XGBoost模型预测抑郁症的价值最高，并且表现出了较高的灵敏度、特异度、准确率、精确率、F1评分和Kappa值。说明相比较其他模型，XGBoost模型预测抑郁症状具有一定优势。

XGBoost模型解释结果表明，生活满意度低、女性、自评健康状况差、认知评分低和睡眠时长不足是抑郁症的最重要高危因素。其中最突出的风险因素是生活满意度低，先前有研究已证实了生活满意度与抑郁症之间的相关性[28]。他们的研究同样发现生活满意度较低的个体更容易发展成抑郁症，特别是在缺乏有效的应对机制时。低生活满意度往往反映了个人在多个生活领域的困境，如职业困扰、人际关系不和谐、身体健康问题等，这些都可能成为引发或加重抑郁症状的诱因[29]。除此之外，相比较男性，女性更容易出现抑郁症状。国内外很多研究都表明女性的抑郁程度更高[30-32]。全球重度抑郁障碍的12个月患病率在女性中为5.8%，而在男性中为3.5%[33]。虽然这种性别差异背后的具体机制尚不清楚，但一种可能的生物学理论是与性别特异性因素有关，即雌激素水平下降可能导致女性患抑郁症的风险增加[34]。心理学理论是由于女性处于社会弱势群体，缺乏社会支持，而且受传统观念“男主外，女主内”的影响，女性需要承担照顾家庭的重担，长期从事繁重的家务劳动，容易产生较大的心理压力，久而久之容易诱发抑郁症[35]。

许多研究表明，自评健康状况较差的个体更容易出现抑郁症状[36-38]。身体健康问题（如慢性疾病、疼痛、失能等）会直接影响自评健康状况，如果个体存在这些健康问题，他们可能会感到身体的限制和功能的丧失，导致个体自评健康差，进而增加抑郁的风险[39]。自评健康较差的个体也可能经历更多的社会隔离和缺乏支持，这使他们更加容易感到孤独、焦虑和无助，从而增加抑郁的风险[40]。此时，可能产生无力感、悲观情绪和对未来的负面预期，这些情绪状态与抑郁症的核心症状（如兴趣丧失、情绪低落、失望等）高度重合。抑郁症状（如情绪低落、无望感）和认知功能之间可能存在相互作用[41, 42]。一方面，认知障碍本身可能加剧抑郁症状。认知能力较差的抑郁患者更可能陷入消极情绪和负性思维的循环，难以自我调节情绪和思维。比如，记忆问题和注意力不集中可能导致个体无法有效应对日常挑战，进而增加负性情绪和焦虑，长期的认知障碍可能使抑郁患者难以有效应对生活中的压力，从而导致抑郁症的慢性化。另一方面，抑郁症也可影响个体情绪和行为，可能导致认知功能下降，还可能导致显著的认知功能损害，尤其是在注意力、记忆、执行功能等方面[43]。除此之外，睡眠时长不足可能会增加白天疲劳，从而导致疲劳引起的不良事件和情绪，最终导致抑郁[44]。睡眠健康状况越好，越不容易出现抑郁症状。一项纵向研究显示，抑郁评分与睡眠时长之间存在很强的相关性，表明睡眠不足对抑郁有很大影响[45]。此外，另一项前瞻性研究表明，较短的睡眠时间与抑郁症状的严重程度增加有关[21]。睡眠和抑郁之间存在很强的双向关系，其中睡眠时间缩短是急性抑郁症状增加的最强预测指标。

目前大部分ML模型均可以依据信息增益做出模型的特征重要性排序图。然而单纯依靠各特征变量的重要性排序还不足以清楚地解释模型的具体决策方式，也无法展示特征与目标变量之间的具体关系。而SHAP和LIME不仅在

模型选择方面展现出更高的自由度和灵活性，可以用于任何类型的ML模型。还可以直观地呈现模型中各个变量对预测结果有正向或负向的影响以及影响程度。结合SHAP与LIME，还能够提供对单个数据点的深度解释，不仅帮助了解特定预测的局部行为，还能从全局角度理解特征对模型的整体贡献。

SHAP和LIME之间的主要区别在于SHAP值与各个特征之间的依赖图分析既可以提供某个特征对预测的全局性贡献，提供更加精细的解释，也可以细致到单个样本，展示特征对预测值的正负方向影响，并量化特征取值与目标变量之间的关系。LIME则通过生成一个由扰动实例组成的新数据集并将其输入黑盒模型中获得新预测结果，针对新数据集训练一个可解释的模型[46]。仅对ML模型所做的单个预测（局部解释）提供解释。目前，国外学者对ML和模型可解释性领域进行了诸多深入探索与应用，例如Alabi[47]等使用LIME方法和SHAP技术将ML模型与可解释的人工智能相结合，通过展示模型在局部区域的行为，提高对鼻咽癌患者总体生存机会预测模型的可解释性。因此，将ML“黑盒”特质与这些可解释性方法相结合，不仅有助于提升模型应用的灵活性，还能有效弥补模型本身可能存在的解释性不足的问题，这对于从数据中精确识别出高风险个体，采取有效措施，具有重要实际意义。

本研究采用SMOTE重采样、数据拆分、交叉验证和超参数优化等方法提升模型精度，减少过拟合，确保模型的严谨性和可推广性。此外，通过SHAP和LIME分析预测因素对抑郁症的影响，这些因素可控或可逆，及时干预有助于改善抑郁状况，减轻医疗负担。但本研究也存在一些不足之处。首先，数据不包括研究对象的医疗使用情况和临床指标，这些数据已被发现可以预测抑郁情况。第二，回顾性研究的性质可能会导致选择偏差，并且所选择的变量受限于调查问卷的结构，这可能导致未包含所有潜在影响抑郁状况的因素。最后，尽管本研究采用了一定的方法保证了模型的可靠性和普适性，但仍需要在外部队列中进一步验证和确认。

## 5. 结论

综上所述，ML模型是预测抑郁症风险的有效工具。利用Shapley加性解释和局部代理模型可以最大限度发挥机器学习的临床优势，有助于尽早预测或发现抑郁症高危患者，对其采取全面评估和早期防治。

## ORCID

0009-0008-3322-7654 (Juan Wang)  
0009-0008-2405-9732 (Man Cui)  
0009-0006-0920-0983 (Miao Deng)  
0009-0000-2668-578X (Yanshuai Fan)  
0000-0001-8924-0761 (Zhiguang Ping)

## 致谢

本文是河南省自然科学基金项目《肥胖及其相关疾病与基因、环境因素关联的网络模型》(NO.182300410303)的阶段性成果之一。

## 参考文献

- [1] Smith K. Mental health: a world of depression [J]. *Nature*, 2014, 515(7526): 181.  
<https://doi.org/10.1038/515180a>
- [2] Wang P S, Aguilar-Gaxiola S, Alonso J, et al. Use of mental health services for anxiety, mood, and substance disorders in 17 countries in the WHO world mental health surveys [J]. *Lancet*, 2007, 370(9590): 841-850.  
[https://doi.org/10.1016/S0140-6736\(07\)61414-7](https://doi.org/10.1016/S0140-6736(07)61414-7)
- [3] Herrman H, Kieling C, McGorry P, et al. Reducing the global burden of depression: a Lancet-World Psychiatric Association Commission [J]. *Lancet*, 2019, 393(10189): e42-e43.  
[https://doi.org/10.1016/S0140-6736\(18\)32408-5](https://doi.org/10.1016/S0140-6736(18)32408-5)
- [4] Alexopoulos G S. Depression in the elderly [J]. *Lancet*, 2005, 365(9475): 1961-1970.  
[https://doi.org/10.1016/S0140-6736\(05\)66665-2](https://doi.org/10.1016/S0140-6736(05)66665-2)
- [5] McCarron R M, Shapiro B, Rawles J, et al. Depression [J]. *Ann Intern Med*, 2021, 174(5): ITC65-ITC80.  
<https://doi.org/10.7326/AITC202105180>
- [6] Mitchell A J, Vaze A, Rao S. Clinical diagnosis of depression in primary care: a meta-analysis [J]. *Lancet*, 2009, 374(9690): 609-619. [https://doi.org/10.1016/S0140-6736\(09\)60879-5](https://doi.org/10.1016/S0140-6736(09)60879-5)
- [7] Hu C, Jiang Q, Yuan Y, et al. Depressive symptoms among the oldest-old in China: a study on rural-urban differences [J]. *BMC Public Health*, 2024, 24(1): 3604.  
<https://doi.org/10.1186/s12889-024-21069-5>
- [8] Loechner J, Starman K, Galuschka K, et al. Preventing depression in the offspring of parents with depression: A systematic review and meta-analysis of randomized controlled trials [J]. *Clin Psychol Rev*, 2018, 60: 1-14.  
<https://doi.org/10.1016/j.cpr.2017.11.009>
- [9] Bzdok D, Altman N, Krzywinski M. Statistics versus machine learning [J]. *Nat Methods*, 2018, 15(4): 233-234.  
<https://doi.org/10.1038/nmeth.4642>
- [10] Grimmer J, Roberts M E, Stewart B M. Machine Learning for Social Science: An Agnostic Approach [J]. *Annual review of political science*, 2021, 24(1): 395-419.  
<https://doi.org/10.1146/annurev-polisci-053119-015921>
- [11] Khan K, Ahmad W, Amin M N, et al. Compressive Strength Estimation of Steel-Fiber-Reinforced Concrete and Raw Material Interactions Using Advanced Algorithms [J]. *Polymers*, 2022, 14(15): 3065.  
<https://doi.org/10.3390/polym14153065>
- [12] Ray S. A Quick Review of Machine Learning Algorithms [C], 2019. IEEE, 2019-01-01.  
<https://doi.org/10.1109/COMITCon.2019.8862451>
- [13] Li J J, Tong X. Statistical Hypothesis Testing versus Machine Learning Binary Classification: Distinctions and Guidelines [J]. *Patterns (N Y)*, 2020, 1(7): 100115.  
<https://doi.org/10.1016/j.patter.2020.100115>
- [14] Chen Q, Zhang Y, Zhang M, et al. Application of Machine Learning Algorithms to Predict Acute Kidney Injury in Elderly Orthopedic Postoperative Patients [J]. *Clin Interv Aging*, 2022, 17: 317-330.  
<https://doi.org/10.2147/CIA.S349978>
- [15] Lee Y, Ragguett R M, Mansur R B, et al. Applications of machine learning algorithms to predict therapeutic outcomes in depression: A meta-analysis and systematic review [J]. *J Affect Disord*, 2018, 241: 519-532.  
<https://doi.org/10.1016/j.jad.2018.08.073>
- [16] Shatte A, Hutchinson D M, Teague S J. Machine learning in mental health: a scoping review of methods and applications [J]. *Psychol Med*, 2019, 49(9): 1426-1448.  
<https://doi.org/10.1017/S0033291719000151>
- [17] Hatton C M, Paton L W, McMillan D, et al. Predicting persistent depressive symptoms in older adults: A machine learning approach to personalised mental healthcare [J]. *J Affect Disord*, 2019, 246: 857-860.  
<https://doi.org/10.1016/j.jad.2018.12.095>
- [18] Radloff L S. The CES-D scale: A self-report depression scale for research in the general population [J]. *Applied psychological measurement*, 1977, 1(3): 385-401.  
<https://doi.org/10.1177/014662167700100306>
- [19] Zhang B, Fokkema M, Cuijpers P, et al. Measurement invariance of the Center for Epidemiological Studies Depression Scale (CES-D) among Chinese and Dutch elderly [J]. *BMC Med Res Methodol*, 2011, 11: 74.  
<https://doi.org/10.1186/1471-2288-11-74>
- [20] Stephens A, Allardyce J, Weavers B, et al. Developing and validating a prediction model of adolescent major depressive disorder in the offspring of depressed parents [J]. *J Child Psychol Psychiatry*, 2023, 64(3): 367-375.  
<https://doi.org/10.1111/jcpp.13704>
- [21] Wang Y, Li J, Bian W, et al. Latent classes of symptom trajectories among major depressive disorder patients in China [J]. *J Affect Disord*, 2024, 350: 746-754.  
<https://doi.org/10.1016/j.jad.2024.01.144>
- [22] Wang K, Zhao Y, Nie J, et al. Higher HEI-2015 Score Is Associated with Reduced Risk of Depression: Result from NHANES 2005-2016 [J]. *Nutrients*, 2021, 13(2).  
<https://doi.org/10.3390/nu13020348>
- [23] Xiao M, Yan C, Fu B, et al. Risk prediction for postpartum depression based on random forest [J]. *Zhong Nan Da Xue Xue Bao Yi Xue Ban*, 2020, 45(10): 1215-1222.  
<https://doi.org/10.11817/j.issn.1672-7347.2020.190655>
- [24] Gu S C, Zhou J, Yuan C X, et al. Personalized prediction of depression in patients with newly diagnosed Parkinson's disease: A prospective cohort study [J]. *J Affect Disord*, 2020, 268: 118-126. <https://doi.org/10.1016/j.jad.2020.02.046>
- [25] Choi J, Choi J, Choi W J. Predicting Depression Among Community Residing Older Adults: A Use of Machine Learning Approach [J]. *Stud Health Technol Inform*, 2018, 250: 265.
- [26] Lin S, Wu Y, Fang Y. A hybrid machine learning model of depression estimation in home-based older adults: a 7-year follow-up study [J]. *BMC Psychiatry*, 2022, 22(1): 816.  
<https://doi.org/10.1186/s12888-022-04439-4>
- [27] Shin D, Lee K J, Adeluwa T, et al. Machine Learning-Based Predictive Modeling of Postpartum Depression [J]. *J Clin Med*, 2020, 9(9). <https://doi.org/10.3390/jcm9092899>
- [28] Abdoli N, Salari N, Darvishi N, et al. The global prevalence of major depressive disorder (MDD) among the elderly: A systematic review and meta-analysis [J]. *Neurosci Biobehav Rev*, 2022, 132: 1067-1073.  
<https://doi.org/10.1016/j.neubiorev.2021.10.041>

- [29] Ma H, Zhao M, Liu Y, et al. Network analysis of depression and anxiety symptoms and their associations with life satisfaction among Chinese hypertensive older adults: a cross-sectional study [J]. *Front Public Health*, 2024, 12: 1370359. <https://doi.org/10.3389/fpubh.2024.1370359>
- [30] Kim H R, Kim S M, Hong J S, et al. Character strengths as protective factors against depression and suicidality among male and female employees [J]. *BMC Public Health*, 2018, 18(1): 1084. <https://doi.org/10.1186/s12889-018-5997-1>
- [31] Labonté B, Engmann O, Purushothaman I, et al. Sex-specific transcriptional signatures in human depression [J]. *Nat Med*, 2017, 23(9): 1102-1111. <https://doi.org/10.1038/nm.4386>
- [32] Plaisier I, de Bruijn J G, de Graaf R, et al. The contribution of working conditions and social support to the onset of depressive and anxiety disorders among male and female employees [J]. *Soc Sci Med*, 2007, 64(2): 401-410. <https://doi.org/10.1016/j.socscimed.2006.09.008>
- [33] Ferrari A J, Somerville A J, Baxter A J, et al. Global variation in the prevalence and incidence of major depressive disorder: a systematic review of the epidemiological literature [J]. *Psychol Med*, 2013, 43(3): 471-481. <https://doi.org/10.1017/S0033291712001511>
- [34] Garde K. Depression--gender differences [J]. *Ugeskr Laeger*, 2007, 169(25): 2422-2425.
- [35] Albert P R. Why is depression more prevalent in women? [J]. *J Psychiatry Neurosci*, 2015, 40(4): 219-221. <https://doi.org/10.1503/jpn.150205>
- [36] Su Q, Fan L. Impact of caregiving on mental, self-rated, and physical health: evidence from the China health and retirement longitudinal study [J]. *Qual Life Res*, 2024, 33(7): 1-10. <https://doi.org/10.1007/s11136-024-03659-3>
- [37] Zhao H, Ma Q, Xie M, et al. Self-rated health as a predictor of hospitalizations in patients with bipolar disorder or major depressive disorder: A prospective cohort study of the UK Biobank [J]. *J Affect Disord*, 2023, 331: 200-206. <https://doi.org/10.1016/j.jad.2023.02.113>
- [38] Ambresin G, Chondros P, Dowrick C, et al. Self-rated health and long-term prognosis of depression [J]. *Ann Fam Med*, 2014, 12(1): 57-65. <https://doi.org/10.1370/afm.1562>
- [39] Maier A, Riedel-Heller S G, Pabst A, et al. Risk factors and protective factors of depression in older people 65+. A systematic review [J]. *PLoS One*, 2021, 16(5): e251326. <https://doi.org/10.1371/journal.pone.0251326>
- [40] 安适, 袁娟, 陈涛, 等. 自评健康在老年人自理能力和抑郁症状之间的中介效应 [J]. *护理学报*, 2022, 29(20): 55-59. <https://doi.org/10.16460/j.issn1008-9969>
- [41] Zhang Y, Wang S, Hermann A, et al. Development and validation of a machine learning algorithm for predicting the risk of postpartum depression among pregnant women [J]. *J Affect Disord*, 2021, 279: 1-8. <https://doi.org/10.1016/j.jad.2020.09.113>
- [42] Knight M J, Baune B T. Cognitive dysfunction in major depressive disorder [J]. *Curr Opin Psychiatry*, 2018, 31(1): 26-31. <https://doi.org/10.1097/YCO.0000000000000378>
- [43] Shimada H, Park H, Makizako H, et al. Depressive symptoms and cognitive performance in older adults [J]. *J Psychiatr Res*, 2014, 57: 149-156. <https://doi.org/10.1016/j.jpsychires.2014.06.004>
- [44] Zhai L, Zhang H, Zhang D. Sleep duration and depression among adults: a meta-analysis of prospective studies [J]. *Depress Anxiety*, 2015, 32(9): 664-670. <https://doi.org/10.1002/da.22386>
- [45] Firth-Cozens J. Individual and organizational predictors of depression in general practitioners [J]. *Br J Gen Pract*, 1998, 48(435): 1647-1651.
- [46] 史可为, 吴亚飞, 方亚. 临床预测模型的可解释性及应用进展 [J]. *现代预防医学*, 2023, 50(06): 1122-1127. <https://doi.org/10.20043/j.cnki.MPM.202208083>
- [47] Alabi R O, Elmusrati M, Leivo I, et al. Machine learning explainability in nasopharyngeal cancer survival using LIME and SHAP [J]. *Sci Rep*, 2023, 13(1): 8984. <https://doi.org/10.1038/s41598-023-35795-0>