

# 基于Stacking的前列腺癌预测模型

王家琪 韩爱庆 王苹 唐燕

**摘要** **目的** 构建基于可解释性机器学习的前列腺癌风险预测模型,筛选核心预警指标,为前列腺癌早期诊断提供客观依据。**方法** 数据来源于国家人口健康数据中心的前列腺肿瘤预警数据集,共纳入3 000例样本,按病理诊断结果划分为患癌组与非患癌组;结合临床病理意义构建游离PSA/总PSA、载脂蛋白B/载脂蛋白A1两项核心比值特征;采用10轮独立重复实验设计,以随机森林(random forest, RF)、轻量级梯度提升机(light gradient boosting machine, LightGBM)、CatBoost和极端梯度提升(extreme gradient boosting, XGBoost)为基模型,通过贝叶斯超参数优化与5折分层交叉验证完成各基模型的参数寻优,在此基础上构建堆叠集成学习(Stacking)模型;采用ROC曲线下面积(area under the curve, AUC)、准确率(accuracy, ACC)、F1分数等多维度指标评估模型性能,同时结合沙普利可加性解释(Shapley additive explanations, SHAP),量化核心特征的重要性及对预测结果的影响程度。**结果** 构建的Stacking集成模型综合性能表现最优,在独立测试集上AUC达 $0.7584 \pm 0.0239$ ,整体性能优于各单一基模型;SHAP分析显示,游离PSA/总PSA比值为最重要的预测特征,对模型输出贡献最大,且在多轮实验中表现稳定。**结论** 基于Stacking的集成学习模型具有良好的预测性能与稳健性,结合SHAP可解释性分析能够清晰揭示模型决策依据与关键风险指标,可为前列腺癌临床早期筛查及穿刺决策提供重要参考。

**关键词** 前列腺癌;疾病预测;机器学习;Stacking;SHAP可解释性分析

**DOI:** 10.3969/j.issn.1002-3208.2026.04.003.

**中图分类号** R318 **文献标志码** A **文章编号** 1002-3208(2026)04-0348-08

**本文著录格式** 王家琪,韩爱庆,王苹,等. 基于Stacking的前列腺癌预测模型[J]. 北京生物医学工程, 2026, 45(4): 348-355. WANG Jiaqi, HAN Aiqing, WANG Ping, et al. A Stacking-based prostate cancer prediction model[J]. Beijing Biomedical Engineering, 2026, 45(4): 348-355.

## A Stacking-based prostate cancer prediction model

WANG Jiaqi, HAN Aiqing, WANG Ping, TANG Yan

School of Management, Beijing University of Chinese Medicine, Beijing 102488

Corresponding author: TANG Yan (E-mail: tangyan97\_1017@sina.com)

**【Abstract】 Objective** We build an interpretable machine learning model to predict prostate cancer risk. The model screens out core warning indicators. It provides an objective basis for early diagnosis of prostate cancer. **Methods** Data came from the prostate tumor warning dataset of the National Population Health Data Center. The dataset contained 3 000 samples. Based on pathological diagnosis, we divided the samples into a cancer group and a non-cancer group. From clinical pathological considerations, we constructed two

core ratio features: free PSA to total PSA ratio, and apolipoprotein B to apolipoprotein A1 ratio. We then used a 10-round independent repeated experiment design. Four base models entered the study: random forest, LightGBM, CatBoost, and

**基金项目:** 国家卫生健康委能力建设和继续教育中心慢病管理研究课题(GWJJMB202510010047)、北京中医药大学社科培育项目(2025-JYB-PY-001)资助

**作者单位:** 北京中医药大学管理学院(北京 102488)

**通信作者:** 唐燕。E-mail: tangyan97\_1017@sina.com

XGBoost. Bayesian hyperparameter optimization together with 5-fold stratified cross-validation tuned each base model. On this basis, we built a Stacking ensemble model. We evaluated model performance using multiple metrics, including AUC, accuracy, and F1 score. We also used the SHAP interpretation framework. SHAP quantified the importance of core features and their influence on prediction results. **Results** The Stacking ensemble model showed the best overall performance. On the independent test set, its AUC reached  $0.7584 \pm 0.0239$ . It outperformed all the single base models. SHAP analysis revealed that the free PSA to total PSA ratio was the most important predictor. It contributed the most to the model output. This finding remained stable across multiple experiments. **Conclusions** The Stacking-based ensemble learning model has good predictive performance and robustness. Combined with SHAP interpretability, it clearly reveals the model's decision basis and key risk indicators. This provides an important reference for early clinical screening and biopsy decisions in prostate cancer.

**【Keywords】** prostate cancer; disease prediction; machine learning; Stacking; SHAP interpretability analysis

## 0 引言

前列腺癌是男性泌尿系统最常见的恶性肿瘤,国家癌症中心数据<sup>[1]</sup>显示,2022年中国前列腺癌新发病例约13.42万例,死亡病例约4.75万例,发病率持续上升,前列腺癌的早诊早治工作非常重要。目前临床筛查主要依靠前列腺特异性抗原(prostate specific antigen, PSA)检测,但是PSA特异性不高会造成过度诊断<sup>[2]</sup>。穿刺活检虽为诊断的金标准,但是会带来血尿、感染等并发症的风险<sup>[3]</sup>,增加患者的负担。因此,建立准确的可解释预测模型来发现高危人群、减少不必要的穿刺,成了本领域研究的主要方向。

近年来,机器学习在前列腺癌预测中应用广泛。集成学习凭借融合多个弱学习器优势、降低过拟合风险的特性,泛化性能显著优于单一模型<sup>[4]</sup>,但现有研究仍存在特征挖掘不充分、多源信息整合不足等问题<sup>[5]</sup>。XGBoost、LightGBM等梯度提升模型虽被广泛应用,可是性能易受训练数据分布影响,在面对临床数据常见的不平衡、高噪声问题时,泛化能力存在局限<sup>[4,6]</sup>。Stacking作为典型集成策略,通过初级学习器特征变换与次级学习器结果融合在疾病预测中表现突出<sup>[7]</sup>,但相关研究或临床特征工程薄弱<sup>[8]</sup>,或缺乏决策机制解析<sup>[9]</sup>。随着复杂模型在医疗领域的普及,可解释性机器学习的研究日益紧迫。沙普利可加性解释(Shapley additive explanations, SHAP)作为后验解释框架,可以量化特征重要性及对预测结果的影响方向,已在疾病预测

模型解释中得到验证<sup>[10-14]</sup>。然而,现有研究多聚焦单一模型与SHAP的结合,将SHAP与Stacking集成相结合,同时解析临床特征意义的研究仍较为匮乏,为此本研究构建融合可解释性分析的前列腺癌风险预测模型。

综上,现有研究存在特征挖掘不够深入与集成模型可解释性较弱两大核心局限。针对上述问题,本研究结合集成学习与可解释建模思想,提出一种融合SHAP解释框架的Stacking集成预测方案,通过构建临床相关衍生特征、优化基模型超参数、融合多模型优势并揭示决策逻辑,为临床医生提供直观、可信的决策支撑。

## 1 资料与方法

本研究围绕前列腺癌风险预测,设计并构建了五阶段研究框架(如图1所示),包括数据准备、特征工程、模型构建、模型评估以及模型可解释性分析五个模块。

### 1.1 数据处理与特征工程

#### 1.1.1 数据来源

数据来源于国家人口健康科学数据中心数据仓储PHDA,由中国人民解放军总医院提供的对外开放的前列腺肿瘤预警数据集。原始数据包含3000例样本,涵盖基本信息(年龄、身高、体质量)、生化检查指标(载脂蛋白、PSA、无机磷、钾等)、病理诊断结果等共32项变量。其中病理诊断标签分为前列腺增生、前列腺癌、同时有前列腺癌和前列腺增生。为降低单中心数据带来的偏倚、提升模型外

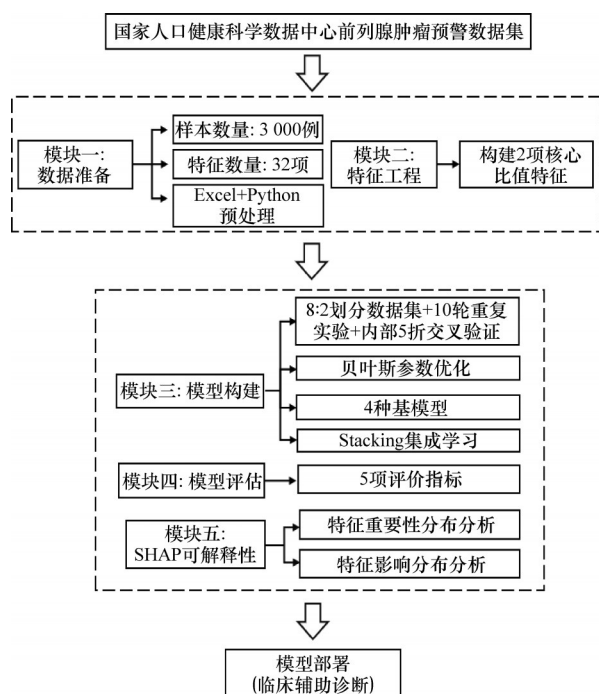


图1 研究框架总体示意图

Figure 1 Overall schematic diagram of the research framework

推性,本研究设计并开展10轮独立重复实验。

### 1.1.2 数据预处理

使用Excel完成基础清洗:仅保留数值型信息,删除含“已复查”“溶血”“乳糜”等无效标识的记录,并剔除含“<”“>”符号的异常数值记录;对病理诊断标签重新赋值,将前列腺增生赋值为0(不患前列腺癌),前列腺癌和同时有前列腺癌、前列腺增生统一赋值为1(患前列腺癌);计算32项变量的缺失率,剔除缺失率超过50%的特征,对剩余特征的缺失值采用中位数进行填充。

### 1.1.3 特征工程

结合临床病理意义,构建游离PSA/总PSA(fPSA\_ratio)、载脂蛋白B/载脂蛋白A1(ApoB/A1)2项核心比值特征,同时删除原始PSA列避免多重共线性;采用鲁棒缩放器(RobustScaler)对特征进行标准化处理,提升模型收敛效率与泛化能力。

## 1.2 模型构建与优化

### 1.2.1 数据集划分

将数据集按8:2比例划分为训练集(2 400例)与独立测试集(600例),训练集进行模型构建与5折交叉验证调优,测试集仅用于评估泛化性能;在全

流程中固定随机种子,确保实验可重复、结果稳定可靠。

### 1.2.2 基模型选择

基学习器包括RF、LightGBM、CatBoost和XGBoost。RF通过特征子采样处理高维共线性;LightGBM和XGBoost分别侧重效率与精度;CatBoost优化类别特征处理。四类模型兼顾预测精度、训练效率及可解释性,充分保证了基学习器的多样性,为后续集成模型构建提供了良好基础。

### 1.2.3 超参数优化

超参数优化采用Optuna贝叶斯优化策略,以内部验证集中5折交叉验证AUC最大化为目标。每轮独立重复实验中,各基模型基于训练集的5折分层交叉验证独立完成超参数寻优,每折迭代10次参数搜索;其中梯度提升类模型均设置早停策略,有效抑制过拟合。超参数空间(表1)依据样本量、特征维度和模型特性设定,涵盖树数量、深度、学习率及正则化系数等。

### 1.2.4 集成模型构建

针对单一模型性能受限的问题,构建Stacking集成模型:第一层为4个优化后基模型输出概率特征,第二层以逻辑回归为元模型完成决策融合,进一步提升整体预测性能与稳健性。

## 1.3 模型评价与可解释性分析

### 1.3.1 模型评价指标

为全面评估模型性能,本实验采用AUC、准确率、精确率(precision)、灵敏度(sensitivity, SE)和F1分数共5项指标进行评价,各指标计算公式见式(1)~式(4)。所有指标结果为10次独立重复实验在测试集上的均值±标准差。

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$SE = \frac{TP}{TP + FN} \quad (3)$$

$$F1 = \frac{2Precision \times SE}{Precision + SE} \quad (4)$$

式中:TP为真阳性样本数;TN为真阴性样本数;FP为假阳性样本数;FN为假阴性样本数。

### 1.3.2 可解释性分析

针对Stacking模型“黑箱”问题,本研究引入SHAP解释框架,从全局特征重要性和指标与患病

表1 基模型超参数优化空间  
Table 1 Hyper parameter optimization space of base models

| 模型名称     | 参数名称                         | 参数范围             |
|----------|------------------------------|------------------|
| RF       | n_estimators(决策树数量)          | [50,120]         |
|          | max_depth(树最大深度)             | [4,8]            |
|          | min_samples_leaf(叶节点最小样本数)   | [5,15]           |
|          | min_samples_split(节点分裂最小样本数) | [20,40]          |
|          | max_features(单棵树使用特征比例)      | [0.5,0.8]        |
| LightGBM | n_estimators(弱学习器数量)         | [50,150]         |
|          | max_depth(树最大深度)             | [4,8]            |
|          | learning_rate(学习率)           | [0.01,0.05](对数域) |
|          | num_leaves(叶节点数量)            | [16,23]          |
|          | lambda_l2(L2正则化系数)           | [10,50]          |
|          | min_child_samples            | [10,20]          |
|          | subsample(样本采样比例)            | [0.7,0.9]        |
|          | colsample_bytree(特征采样比例)     | [0.7,0.9]        |
|          | n_estimators(弱学习器数量)         | [50,150]         |
| CatBoost | depth(树深度)                   | [4,8]            |
|          | learning_rate(学习率)           | [0.01,0.05](对数域) |
|          | l2_leaf_reg(L2正则化系数)         | [10,50]          |
|          | min_child_samples            | [10,20]          |
|          | subsample(样本采样比例)            | [0.7,0.9]        |
|          | colsample_bylevel(层级特征采样比例)  | [0.7,0.9]        |
|          | n_estimators(弱学习器数量)         | [50,150]         |
| XGBoost  | max_depth(树最大深度)             | [4,8]            |
|          | learning_rate(学习率)           | [0.01,0.05](对数域) |
|          | lambda(L2正则化系数)              | [10,50]          |
|          | alpha(L1正则化系数)               | [5,20]           |
|          | min_child_weight             | [5,20]           |
|          | subsample(样本采样比例)            | [0.7,0.9]        |
|          | colsample_bytree(特征采样比例)     | [0.7,0.9]        |
|          | gamma(节点分裂阈值)                | [0,5]            |
|          |                              |                  |

风险之间的关联两个层面解析模型决策机制。

$$\phi_i = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!(|F| - |S| - 1)!}{|F|! [f(S \cup \{i\}) - f(S)]} \quad (5)$$

式中: $\phi_i$ 为特征*i*的SHAP值,代表其对模型预测结果的贡献程度; $F$ 为全部特征集合; $S$ 为不含特征*i*的特征子集; $f(\cdot)$ 为定义在特征子集上的预测函数, $f(S \cup \{i\}) - f(S)$ 为特征*i*对特征子集*S*的边际贡献。

## 2 结果

### 2.1 数据基本特征

数据预处理后纳入3 000例样本,其中前列腺

癌患者占比53.13%,非前列腺癌患者占比46.87%,样本标签分布相对平衡。经特征工程处理后保留23项特征,包括年龄、身高、体质量、球蛋白、碱性磷酸酶、肌酸激酶同工酶、钠、钙、氯化物、无机磷、游离钙、乳酸脱氢酶、肌酸激酶、肌酐、血清尿酸、甘油三酯、载脂蛋白B、载脂蛋白A1、高密度脂蛋白胆固醇、低密度脂蛋白胆固醇、钾21项原始指标,以及基于临床意义构建的fPSA\_ratio、ApoB/A1。

### 2.2 超参数优化结果

基于前述超参数空间,采用Optuna贝叶斯优化完成寻优后,各基模型最优超参数在10次独立重复实验中分布集中,且均处于预设范围内,寻优结果

稳定性良好。第 1 次重复实验的各基模型最优超参数(保留 4 位小数)详见表 2。

### 2.3 模型性能对比

测试集评估结果显示(表 3),Stacking 集成模型综合性能最优,测试集 AUC 达 0.758 4±0.023 9、F1 为 0.693 7±0.020 1、ACC 为 0.694 2,通过融合多个基模型的输出特征,有效整合各模型优势,进一步提升了模型的整体判别能力。

单一基模型中,CatBoost 表现最为突出,各项指标均处于前列,整体性能最接近 Stacking。LightGBM 的 ACC 达 0.689 0,为单一模型中最高,AUC(0.752 6±0.021 7)与 F1(0.692 2±0.021 3)亦保持较高水平,综合性能仅次于 CatBoost。XGBoost 与 RF 性能表现相近,AUC 分别为 0.749 1±0.021 4、0.751 1±0.023 1,F1 值与 ACC 略低于前述模型,但整体分类效果稳定,未出现明显波动。

从指标稳定性来看,所有模型的 AUC、F1 值标准差均处于较低水平,表明各模型在 10 轮独立重复实验中均展现出良好的稳健性,本研究的实验结果具备可靠性与可重复性。

### 2.4 SHAP 可解释性分析

鉴于 Stacking 的稳定性已在 10 次独立重复实验中得到验证(AUC 变异系数<5%),实验结果具备良好的一致性与可靠性,因此本研究基于最后一轮收敛稳定的重复实验结果,开展 SHAP 可解释性分析,解析模型决策机制和提升临床可解释性。

#### 2.4.1 特征重要性分布分析

特征重要性采用各特征平均 SHAP 值进行量化,用以反映特征对模型输出的整体影响强度。结果显示(图 2),fPSA\_ratio 以 0.341 的重要性值位居首位,显著高于其他特征;其次为无机磷(0.142)、低密度脂蛋白胆固醇(0.041)、球蛋白(0.038)、肌酸激酶同工酶(0.030)分别位列第四、第五。

这一特征重要性排序与现有临床认知高度一致,验证了模型对疾病相关指标的精准捕捉能力。fPSA\_ratio 降低是前列腺癌的经典预警信号,临床常以 0.16 作为穿刺决策的临床参考值<sup>[15]</sup>,模型将其识别为首要特征,体现了对病理意义明确指标的精准捕捉;无机磷作为排名第二的生化指标,它的重要性与王逸飞等<sup>[16]</sup>的研究结论一致,提示血清磷水平异常与前列腺癌风险存在潜在关联;在中国人群研

表 2 模型最优参数设置

Table 2 Optimal hyper parameter settings of each model

| 模型       | 参数设置                                                                                                                                                                 |
|----------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| RF       | n_estimators=50,max_depth=6,min_samples_leaf=10,<br>min_samples_split=33,max_features=0.793 5                                                                        |
| LightGBM | n_estimators=105,max_depth=4,learning_rate=0.047 6,num_leaves=29,<br>lambda_l2=47.578 0,min_child_samples=19,subsample=0.819 6,<br>colsample_bytree=0.884 4          |
| CatBoost | n_estimators=104,depth=4,learning_rate=0.018 0,l2_leaf_reg=28.243 0,<br>min_child_samples=18,subsample=0.740 0,colsample_bylevel=0.802 9                             |
| XGBoost  | n_estimators=121,max_depth=4,learning_rate=0.047 6,lambda=43.297 7,<br>alpha=8.185 1,min_child_weight=7,subsample=0.736 7,<br>colsample_bytree=0.760 9,gamma=2.623 8 |

表 3 各模型测试集性能对比

Table 3 Performance comparison of each model on test set

| 模型       | AUC             | ACC             | 精确度             | SE              | F1              |
|----------|-----------------|-----------------|-----------------|-----------------|-----------------|
| RF       | 0.751 1±0.023 1 | 0.686 0±0.016 4 | 0.737 7±0.017 1 | 0.635 4±0.022 4 | 0.682 6±0.018 1 |
| LightGBM | 0.752 6±0.021 7 | 0.689 0±0.020 9 | 0.730 7±0.023 8 | 0.658 0±0.025 9 | 0.692 2±0.021 3 |
| CatBoost | 0.756 3±0.024 7 | 0.690 7±0.020 6 | 0.748 7±0.023 9 | 0.629 8±0.023 1 | 0.684 0±0.021 7 |
| XGBoost  | 0.749 1±0.021 4 | 0.688 5±0.019 2 | 0.736 9±0.024 6 | 0.644 8±0.016 5 | 0.687 7±0.017 8 |
| Stacking | 0.758 4±0.023 9 | 0.694 2±0.020 4 | 0.742 2±0.024 4 | 0.651 4±0.020 8 | 0.693 7±0.020 1 |

究中,低密度脂蛋白胆固醇升高被确定为预测前列腺癌的独立风险因素<sup>[17]</sup>;肌酸激酶、肌酸激酶同工酶等酶类指标波动与肿瘤进展过程中的组织损伤、代谢紊乱相关,常作为肿瘤患者全身状态的补充参考<sup>[18-19]</sup>;此外,球蛋白水平升高则与疾病进展中的免疫炎症反应相关<sup>[20]</sup>。

#### 2.4.2 特征影响分布分析

为进一步揭示模型的决策逻辑,本研究通过SHAP摘要图(图3)可视化特征取值对患病风险的影响方向与强度。

(1)核心特征fPSA\_ratio具有显著的负向归因特征。作为贡献度最高的特征,fPSA\_ratio的效应分布呈现清晰的负向关联模式:当fPSA\_ratio指标水平较低时(蓝色),对应正SHAP值,表明该指标水平下降会增加前列腺癌患病风险;当fPSA\_ratio指标水平较高时(红色),对应负SHAP值,表明该指标水平升高会降低患病风险。该趋势与临床病理逻辑高度吻合,证实基于临床意义构建的衍生特征可有效提升模型风险识别能力。

(2)关键生化指标呈现明确的风险关联规律。无机磷指标水平较高时多对应正SHAP值,水平较低时多对应负SHAP值,表明高磷状态与前列腺癌风险呈正相关;低密度脂蛋白胆固醇水平较高时多对应正SHAP值,提示血脂水平升高与患癌风险增

加相关,为其作为前列腺癌潜在风险因素的假设提供了初步支持;球蛋白水平较高时多对应正SHAP值,反映了以球蛋白升高为标志的慢性炎症与前列腺癌风险的潜在关联。

此外,肌酸激酶、甘油三酯等指标也呈现出明确的效应分布模式,提示代谢紊乱、组织及细胞损伤相关的生物标志物,在前列腺癌风险评估模型中同样可能提供有价值的补充信息。

### 3 讨论

本研究构建的Stacking前列腺癌风险预测模型在测试集AUC达 $0.7584 \pm 0.0239$ ,综合预测性能优于RF、LightGBM、CatBoost、XGBoost四类单一模型,判别能力与实验稳健性良好。与现有Stacking相关研究对比,熊思伟等<sup>[8]</sup>借助Borderline-SMOTE过采样平衡样本后模型AUC可达0.982,刘佳欣<sup>[9]</sup>构建的Stacking模型AUC为0.798;本研究模型精度略低,主要原因是未采用人工重采样手段校正样本分布,完全依托原始真实临床数据建模,且通过10轮独立重复实验验证结果稳定性,所得结论更贴合临床真实筛查场景。

现有研究多侧重模型性能提升,缺乏对集成模型的可解释性分析。本研究将Stacking与SHAP相结合,可量化特征贡献、清晰揭示模型决策依据,是

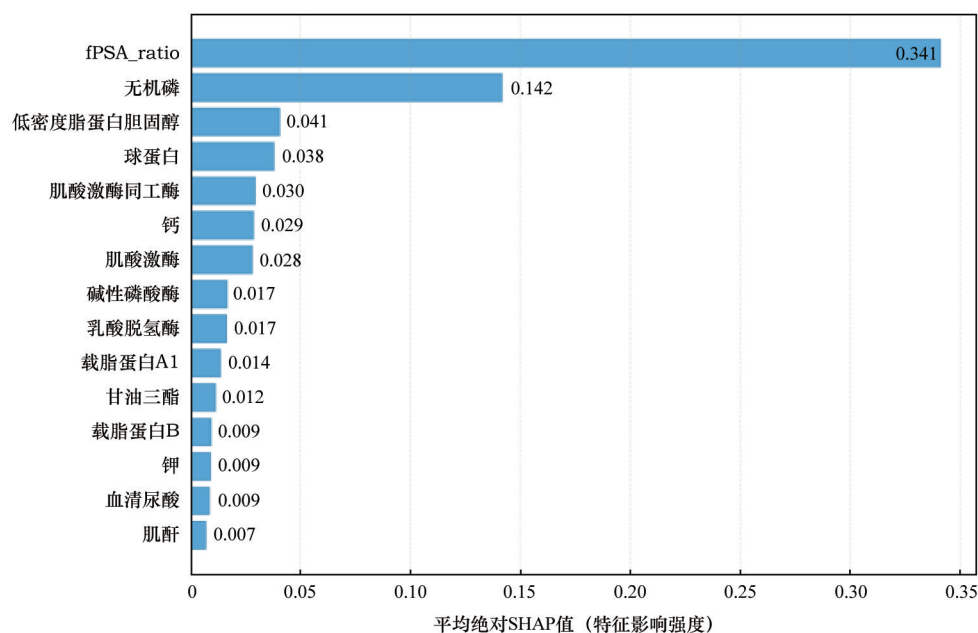


图2 Stacking模型Top15特征SHAP重要性排序

Figure 2 SHAP importance ranking of Top 15 features for the Stacking

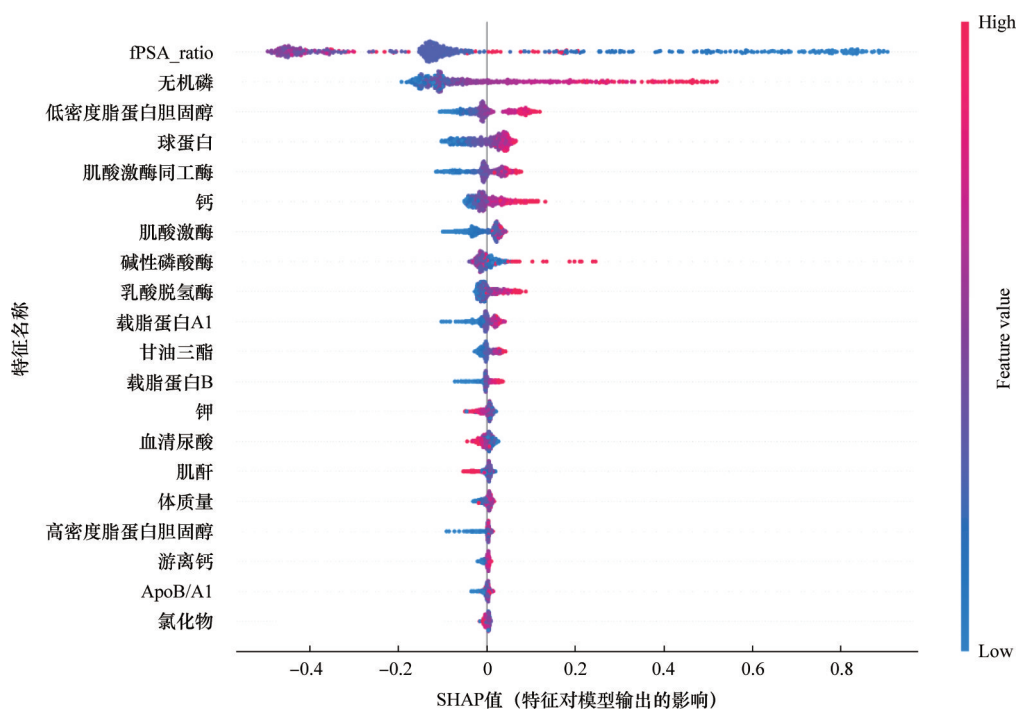


图 3 Stacking 模型 SHAP 摘要图

Figure 3 SHAP summary plot for the Stacking

本文的主要创新点。SHAP 分析表明, fPSA\_ratio、无机磷、低密度脂蛋白胆固醇为核心预测特征, 其风险关联趋势与临床认知及文献结论一致。fPSA\_ratio 作为首要预测因子, 其水平降低与患癌风险升高显著相关, 印证了临床衍生特征构建的合理性; 无机磷与血脂指标的重要性, 则符合代谢紊乱、炎症反应参与肿瘤发生发展的病理机制。

本研究仍存在若干局限性: 研究数据来源于单中心, 模型尚未在外部独立数据集上开展验证; 仅纳入生化指标, 未融合影像及生活方式等多维度信息; 未对特征间交互作用进行深入探讨, 模型暂不具备动态更新机制。未来可进一步开展多中心外部验证、整合多源异构数据、构建动态更新框架, 从而持续提升模型的可靠性与潜在临床应用价值。

#### 4 结论

本研究成功构建融合 SHAP 可解释性分析的 Stacking 前列腺癌风险预测模型, 该模型具有良好的预测性能与稳健性, 综合效能优于单一模型; SHAP 分析明确 fPSA\_ratio、无机磷、低密度脂蛋白胆固醇为关键风险特征, 其影响趋势与临床认知一致。本研究基于无创生化指标可实现前列腺癌风

险的定量评估, 在保证判别能力的同时提升了集成模型的可解释水平, 可为临床早期筛查提供辅助参考, 有望减少不必要的穿刺检查、减轻患者医疗负担。

#### 参考文献

- [1] Han B, Zheng R, Zeng H, et al. Cancer incidence and mortality in China, 2022. Journal of the National Cancer Center, 2024, 4(1): 47-53.
- [2] 刘洋, 王涌泉. PSA, PSAD, f/tPSA 对前列腺癌的诊断及穿刺后 Gleason 评分的预测[J]. 医学研究杂志, 2016, 45(2): 116-120.  
Liu Y, Wang YQ. PSA, PSAD, f/t PSA for diagnosis of prostate cancer and prediction of post needle biopsy Gleason score[J]. Journal of Medical Research, 2016, 45(2): 116-120.
- [3] 刘爽, 袁仁斌. 前列腺穿刺活检技术的演变过程[J]. 微创泌尿外科杂志, 2022, 11(2): 139-143.  
Liu S, Yuan RB. Evolution of prostate biopsy technology[J]. Journal of Minimally Invasive Urology, 2022, 11(2): 139-143.
- [4] 彭岩, 马铃, 张文静, 等. 基于集成学习的风险预测模型研究与应用[J]. 计算机工程与设计, 2022, 43(4): 956-961.  
Peng Y, Ma L, Zhang WJ, et al. Research and application of risk forecast model based on ensemble learning[J]. Computer Engineering and Design, 2022, 43(4): 956-961.

- [5] 路帅,李文杰,徐紫薇,等. 前列腺癌风险预测模型的构建与验证[J]. 重庆医科大学学报, 2023, 48(3): 328-334.  
Lu S, Li WJ, Xu ZW, et al. Construction and validation of a prostate cancer risk prediction model[J]. Journal of Chongqing Medical University, 2023, 48(3): 328-334.
- [6] Wang C, Chang QX, Wang XM, et al. Prostate cancer risk prediction and online calculation based on machine learning algorithm [J]. Chinese Medical Sciences Journal, 2022, 37(3): 210-217.
- [7] Wolpert DH. Stacked generalization [J]. Neural Networks, 1992, 5(2): 241-259.
- [8] 熊思伟,刘玉琳. 基于Borderline-SMOTE算法与Stacking集成学习的前列腺肿瘤风险预测研究[J]. 现代肿瘤医学, 2023, 31(16): 3075-3081.  
Xiong SW, Liu YL. Risk prediction study of prostate tumors based on Borderline-SMOTE algorithm and Stacking ensemble learning[J]. Modern Oncology, 2023, 31(16): 3075-3081.
- [9] 刘佳欣. 基于Stacking集成学习的前列腺癌预测模型研究[D]. 温州: 温州医科大学, 2024.  
Liu JX. Research on prostate cancer prediction model based on Stacking integrated learning [D]. Wenzhou: Wenzhou Medical University, 2024.
- [10] 孙悦,陈广新,于森,等. 基于可解释性机器学习的新冠肺炎疾病风险预测研究[J]. 新一代信息技术, 2023, 6(12): 40-44.  
Sun Y, Chen GX, Yu M, et al. Disease risk prediction of COVID-19 based on interpretive machine learning [J]. New Generation of Information Technology, 2023, 6(12): 40-44.
- [11] 徐青,余冰,周佩敏,等. 基于机器学习与SHAP的全髋关节置换术患者下肢深静脉血栓可解释性预测模型构建研究[J]. 中国医院统计, 2024, 31(1): 11-18, 24.  
Xu Q, Yu B, Zhou PM, et al. Construction of interpretable prediction model for lower limb deep vein thrombosis in patients undergoing total hip arthroplasty based on machine learning and SHAP [J]. Chinese Journal of Hospital Statistics, 2024, 31(1): 11-18.
- [12] 吴佳琪,方素真,黎楷,等. 基于机器学习与SHAP可解释性住院老年2型糖尿病患者衰弱风险预测模型的构建及验证[J]. 护理学报, 2025, 32(11): 7-12.  
Wu JQ, Fang SZ, Li K, et al. Construction and validation of frailty risk prediction model for hospitalized elderly patients with type 2 diabetes based on machine learning and SHAP explainability [J]. Journal of Nursing (China), 2025, 32(11): 7-12.
- [13] 富丹,孙悦,郭金兴,等. 基于可解释性机器学习的阿尔茨海默症痴呆风险预测[J]. 新一代信息技术, 2023, 6(24): 12-16.  
Fu D, Sun Y, Guo JX, et al. Prediction of Alzheimer's dementia risk based on interpretable machine learning [J]. New Generation of Information Technology, 2023, 6(24): 12-16.
- [14] Lundberg SM, Erion G, Chen H, et al. From local explanations to global understanding with explainable AI for trees [J]. Nature Machine Intelligence, 2020, 2(1): 56-67.
- [15] Huang CG, Chen XY, Ma CQ, et al. Evaluation of serum FPSA/TPSA and hormone determination combined with nuclide bone imaging in diagnosis of prostate carcinoma and its recidivation and metastasis [J]. International Journal of Radiation Medicine and Nuclear Medicine, 2010, 34(6): 337-343.
- [16] 王逸飞,吴欢,薛万国,等. 前列腺癌与前列腺增生的分类预测及癌症风险因素分析[J]. 解放军医学院学报, 2021, 42(3): 277-281, 305.  
Wang YF, Wu H, Xue WG, et al. Differentiating prostate cancer from prostate hyperplasia and related factors analysis [J]. Academic Journal of Chinese PLA Medical School, 2021, 42(3): 277-281, 305.
- [17] Gui Q, Wu D, Xu F, et al. Exploring the positive association of blood lipid levels with prostate cancer risk and their relationship to pathological features in the Chinese population [J]. Medicine (Baltimore), 2025, 104(12): e41762.
- [18] Forman DT. The significance of creatine kinase (CKBB) in metastatic cancer of the prostate [J]. Annals of Clinical and Laboratory Science, 1979, 9(4): 333-337.
- [19] Yan YB. Creatine kinase in cell cycle regulation and cancer [J]. Amino Acids, 2016, 48(8): 1775-1784.
- [20] 李宏基,赵晓龙,胡伟,等. 前列腺癌早期筛查风险预测模型的建立与验证研究[J]. 中国全科医学, 2024, 27(20): 2483-2490.  
Li HJ, Zhao XL, Hu W, et al. Development and validation of a prediction model for prostate cancer early screening [J]. Chinese General Practice, 2024, 27(20): 2483-2490.

(2026-03-09收稿,2026-05-06修回)