基于多维度常规指标的机器学习模型早期预测急性冠脉综合征

于　越, 何雪梦*

(吉林市化工医院-北华大学附属第二医院, 吉林　132013)

摘　要: 目的　基于检验、影像及临床特征等多维度常规指标, 构建机器学习模型早期预测急性冠脉综合征。**方法**　回顾性分析吉林市化工医院 2022 年 1 月—2025 年 1 月期间 578 例急性冠脉综合征(acute coronary syndrome, ACS)确诊患者电子病例数据, 提取入院 24 h 内常规检验指标(血脂、心肌损伤标志物、炎症标志物等)、影像学特征(超声心动图、胸部 X 线)及临床资料共 103 项变量。通过单变量分析、最小绝对收缩和选择算子(least absolute shrinkage and selection operator, LASSO)-随机森林重要性评估、Bootstrap 稳定性验证筛选确定 18 项核心预测因子, 采用极端梯度提升(eXtreme Gradient Boosting, XGBoost)等 5 种机器学习算法构建模型, 用独立测试集(n=174)评估性能, 并通过沙普利加性解释(shapley additive explanations, SHAP)解析预测机制。**结果**　XGBoost 模型预测 ACS 的效能最优, 测试集受试者工作特征曲线下面积(area under the curve, AUC)为 0.901 [95%置信区间(confidence interval, CI): 0.872~0.927], 敏感性 83.2%、特异性 88.7%, 显著优于其他机器学习模型及传统评分(DeLong 检验 P<0.05)。SHAP 可解释性分析显示: 左心室射血分数(left ventricular ejection fraction, LVEF)是最关键预测因子(SHAP 均值绝对值|0.32|), 当 LVEF<50%时风险显著增加(SHAP 值 Δ=–0.41); 超敏 C 反应蛋白(high-sensitivity C-reactive protein, hs-CRP)>8 mg/L 与肌钙蛋白 I (cardiac troponin I, cTnI)>120 ng/L 存在协同作用, 共同显著增加 ACS 风险; 高密度脂蛋白胆固醇(high-density lipoprotein cholesterol, HDL-C)>1.42 mmol/L 联合估算肾小球滤过率(estimated glomerular filtration rate, eGFR)>75 mL/min, 可降低 35%~48%的 ACS 风险。**结论**　本研究构建的 18 项常规指标 XGBoost 模型具有优异判别能力(AUC>0.90)和临床可解释性, 并首次量化 HDL-C 与 eGFR 协同保护效应, 为 ACS 早期风险分层提供实用工具。

关键词: 急性冠脉综合征; 对维度; 常规检测; 机器学习; 早期; 风险预测; 可解释人工智能

0　引　言

急性冠脉综合征(acute coronary syndrome, ACS)是心血管疾病中最具威胁性的急症之一, 以冠状动脉粥样硬化斑块破裂或侵蚀继发血栓形成为主要病理基础, 涵盖 ST 段抬高型心肌梗死(st-segment elevation myocardial infarction, STEMI)、非 ST 段抬高型心肌梗死(non-st-segment elevation myocardial infarction, NSTEMI)及不稳定型心绞痛(unstable angina, UA)[1], 其起病急骤、进展迅速, 若未能早期识别和干预, 极易导致心肌坏死、心力衰竭、恶性心律失常甚至猝死等严重不良心血管事件, 造成重大健康负担。因此, 实现 ACS 的早期精准风险分层, 对优化临床决策和改善预后至关重要。当前广泛应用的血栓溶解心肌梗死(thrombolysis in myocardial infarction, TIMI)和全球急性冠状动脉事件注册(global registry of acute coronary events, GRACE)风险评分模型, 虽具临床价值, 但主要依赖有限临床变量, 未能充分利用入院时即可获取的多维度客观数据—包括检验科常规检测项目及常规影像学检查中的量化特征, 其预测效能存在局限[2-3]。近年来, 机器学习(machine learning, ML)技术的突破为挖掘复杂医疗数据中的潜在规律提供了新范式[4], 尽管已有研究探索单一维度的预测模型, 但综合利用常规检验指标、基础影像学特征及核心临床资料,

第一作者: 于　越, 主管技师, 主要研究方向为医学检验。E-mail: yuyue200816@163.com

*通信作者: 何雪梦, 主管技师, 主要研究方向为医学检验。E-mail: hxm200816@163.com

通过ML算法构建多指标融合风险预测模型的研究仍显不足[5]。为此, 本研究基于578例ACS患者的电子病历数据, 整合上述3类常规指标, 旨在开发并验证一种高性能的ML融合模型。该研究强调模型的区分度、校准度与临床可解释性, 致力于为急诊及基层医疗提供实用、高效的ACS早期预警工具。

1　材料与方法

1.1　一般资料

本研究采用回顾性队列设计, 数据来源于吉林市化工医院心血管中心2022年1月—2025年1月期间收治的ACS确诊患者的电子病历系统。研究通过医院伦理委员会审批(批准号: JH-2025-01-01), 免除患者知情同意要求。

入组标准为: (1)年龄大于等于18周岁; (2)符合《急性冠脉综合征急诊快速诊治指南》诊断标准[6], 经冠状动脉造影(coronary angiography, CAG)或典型临床症状结合心电图动态演变及心肌损伤标志物(肌钙蛋白I/T)序列变化确诊为ACS(涵盖STEMI、NSTEMI及UA); (3)入院24 h内完成检验科常规检测(包括心肌酶谱、血脂、血糖、肝肾功能、电解质、血常规、凝血功能及炎症标志物)及常规影像学检查(胸部X线摄影与超声心动图); (4)电子病历中具备完整的基线临床特征记录。

排除标准包括: (1)关键数据缺失率大于20%(检验/影像/临床数据); (2)合并严重非心血管系统疾病(如晚期恶性肿瘤、终末期肾病)或外伤; (3)妊娠期或哺乳期妇女; (4)既往心脏移植史。经严格筛选, 最终纳入578例符合条件的ACS患者作为研究队列, 其中男性389例(67.3%), 女性189例(32.7%), 年龄分布为42~85岁(中位数64)。所有病例的诊断均经两名心血管专科医师独立确认。

1.2　实验方法

1.2.1　数据采集

研究数据通过结构化电子病历系统与医院信息管理系统(hospital information system, HIS)进行标准化提取, 涵盖人口学特征、基线临床资料、检验科常规检测结果及影像学检查报告四大维度。具体采集流程如下。(1)人口学与临床资料。包括年龄、性别、体质量指数(body mass index, BMI)、吸烟史、高血压病、糖尿病、血脂异常、既往心肌梗死/冠状动脉介入治疗(percutaneous coronary intervention, PCI)/冠状动脉旁路移植术(coronary artery bypass grafting, CABG)史等心血管危险因素与疾病史; (2)检验科指标。采集患者入院24 h内完成的常规检测项目, 涉及心肌损伤标志物[肌钙蛋白I (cardiac troponin I, cTnI)和肌酸激酶同工酶MB (creatine kinase isoenzyme MB, CK-MB)]、血脂谱[总胆固醇、甘油三酯、低密度脂蛋白胆固醇(low-density lipoprotein cholesterol, LDL-C)、高密度脂蛋白胆固醇(high-density lipoprotein cholesterol, HDL-C)]、肾功能[血肌酐、估算肾小球滤过率(estimated glomerular filtration rate, eGFR)]、肝功能(谷丙转氨酶、谷草转氨酶)、血糖、电解质(钾、钠)、血常规(白细胞计数、中性粒细胞百分比、血红蛋白、血小板计数)及炎症标志物(超敏C反应蛋白); (3)影像学特征。由两名经过培训的心血管影像医师采用双盲法解读胸部X线摄影与超声心动图报告, 提取可量化/半量化特征—胸部X线重点关注肺淤血程度(Kerley B线、肺门模糊征)、心胸比率; 超声心动图记录左心室射血分数(left ventricular ejection fraction, LVEF)、室壁运动异常指数、左心室舒张末期内径(left ventricular end-diastolic diameter, LVEDD)及心包积液等关键参数。所有数据经脱敏处理后录入数据库, 采用双人独立录入与交叉验证机制确保准确性, 对缺失值采用多重插补法处理, 要求连续变量缺失率小于5%, 分类变量缺失率小于2%。

1.2.2　特征筛选

为构建高泛化性的预测模型并降低维度灾难风险, 本研究采用三阶段特征筛选策略: (1)单变量预筛选[7]。通过Mann-Whitney *U* 检验(非正态连续变量)或卡方检验(分类变量)分析578例患者的全部103个初始特征与终点事件MACE的关联性, 保留$P<0.10$且效应量显著的特征[如超敏C反应蛋白大于10 mg/L的比值比(odds ratio, OR)=3.21, 95%置信区间(confidence interval, CI): 1.87~5.52]; (2)多变量重要性评估[8]。运用带L1惩罚项的最小绝对收缩和选择算子回归(least absolute shrinkage and selection operator, LASSO)与随机森林(random forest, RF)算法并行计算特征重要性—LASSO回归通过10折交叉验证优化正则化强度λ (λ=0.032时特征压缩至28个), RF则基于基尼不纯度下降量的量化特征贡献度(前15%特征包括左心室射血分数、肌钙蛋白I峰值、中性粒细胞/淋巴细胞比值); (3)稳定性验证[9]。采用Bootstrap重抽样(n=1000次)计算特征入选频率, 最终保留在大于80%抽样中稳定出现的核心特征子集(18个), 其筛选一致性经科恩Kappa系数验证(Kappa=0.86, $P<0.001$)。所有分析均在预处理后的训练集(n=404)独立进行, 最终入模特征经临床可解释性评估后被归类为炎症反应、心肌损伤、心功能及代谢四大病理生理维度。

1.2.3　机器学习模型构建

本研究采用RF、支持向量机(support vector machine, SVM)、轻量级梯度提升机(light gradient boosting machine, Light GBM)、K邻近算法(k-nearest neighbor, KNN)及极限梯度提升(eXtreme Gradient Boosting, XGBoost)五类机器学习算法[10-12], 基于筛选的18维核心特征子集构建预测模型。通过Optuna框架(v3.2.0)执行贝叶斯超参数优化: 针对各算法关键参数(如RF的树数量与最大深度、SVM的惩罚系数C、Light GBM的叶子节点数)在预设空间进行100轮搜索, 以五折交叉验证的准确率最大化为目标函数确定最优配置(如RF: 树

数量=180, 最大深度=8; KNN: 邻点数=17, 曼哈顿距离度量)。采用优化后的超参数在训练集(n=404)完成模型训练后, 于严格隔离的测试集(n=174)评估性能, 主要指标包括准确率、召回率及受试者工作特征(receiver operating characteristic, ROC)曲线下面积(area under the curve, AUC)。最终选取验证性能最佳的模型, 利用沙普利加性解释(shapley additive explanations, SHAP)(SHAP v0.42.0)实现特征贡献度的全局分析与个体预测决策的可视化解析, 重点阐释 LVEF、超敏 C 反应蛋白等核心特征的生物学意义[12]。

1.3 统计学方法

研究采用约登指数(Youden index)确定机器学习模型的最佳分类阈值, 应用该阈值将预测概率转换为二元分类结果(阳性/阴性), 进而计算真阳性(true positives, TP)、假阳性(false positives, FP)、真阴性(true negatives, TN)和假阴性(false negatives, FN)样本数[13]。其中 TP 指正确识别的主要不良心血管事件发生者, TN 代表正确判别的未发生者, FP 表示将非 MACE 患者错误识别为阳性, FN 则为漏诊的 MACE 患者。通过绘制 ROC 曲线并计算 AUC, 评估模型在独立测试集上的整体判别效能。为验证不同算法间预测性能的统计差异, 采用 DeLong 检验进行两两比较, 同时计算以下量化指标综合评估模型表现:

$$F_1=2\times(\text{精密度}\times\text{召回率})/(\text{精密度}+\text{召回率}) \quad (1)$$

$$\text{准确率}=(TP+TN)/(TP+TN+FP+FN) \quad (2)$$

$$\text{特异性}=TN/(TN+FP) \quad (3)$$

$$\text{敏感性}=TP/(TP+FN) \quad (4)$$

$$\text{阳性预测值}=TP/(TP+FP) \quad (5)$$

$$\text{阴性预测值}=TN/(TN+FN) \quad (6)$$

所有指标的 95% CI 通过威尔逊方法(Wilson's method)计算, 统计显著性阈值设为双尾 P<0.05。数据分析在 PMedCalc (v22.001)和 SPSS (v25.0)中完成, 包含阈值优化、假设检验及性能指标计算流程[13]。

2 结果与分析

2.1 训练集与测试集基线特征比较

训练集(n=404)与测试集(n=174)的基线特征表现出高度一致性: 两组患者的年龄中位数均为 64 岁[四分位距(interquartile range, IQR): 训练集 56~72 岁, 测试集 55~71 岁], 男性比例相近(67.1% *vs* 67.8%), 心血管危险因素如高血压(65.1% *vs* 63.2%)和糖尿病(33.2% *vs* 34.5%)的分布差异均小于 2%。在关键预后指标方面, 心肌损伤标志物肌钙蛋白 I 峰值中位数(152 ng/L *vs* 146 ng/L)和炎症指标超敏 C 反应蛋白(high-sensitivity C-reactive protein, hs-CRP)中位数(8.2 mg/L *vs* 8.5 mg/L)高度匹配, LVEF 均值也基本一致(52.3% *vs* 53.1%)。影像学特征中室壁运动异常指数(1.8 *vs* 1.7)和心胸比率(0.52 *vs* 0.51)存在轻微数值差异, 但 Bootstrap 重抽样验证显示无统计学意义(P>0.05), 疾病亚型分布同样保持稳定(STEMI 占比 42.3% *vs* 41.4%)。所有 27 项指标的标准化均值差(standardized mean difference, SMD)均低于 0.20 的临界阈值(最大 SMD=0.180), 其中 91% 的指标 SMD≤0.10, 核心变量 SMD<0.05, 表明两组数据在人口学、临床病史、实验室及影像学特征上均满足机器学习建模的同质性要求(表 1)。

表 1 训练集与测试集基线特征比较

具体指标	训练集(n=404)	测试集(n=174)	SMD
年龄/岁[中位数(IQR)]	[64 (56~72)]	[64 (55~71)]	0.042
男性[n (%)]	271 (67.1%)	118 (67.8%)	0.016
BMI/(kg/m^2)	24.3±3.1	24.1±3.3	0.063
吸烟史[n (%)]	182 (45.0%)	81 (46.6%)	0.032
高血压[n (%)]	263 (65.1%)	110 (63.2%)	0.039
糖尿病[n (%)]	134 (33.2%)	60 (34.5%)	0.027
血脂异常[n (%)]	241 (59.7%)	101 (58.0%)	0.034
既往心肌梗死[n (%)]	87 (21.5%)	38 (21.8%)	0.007
PCI 史[n (%)]	73 (18.1%)	30 (17.2%)	0.023
CABG 史[n (%)]	29 (7.2%)	11 (6.3%)	0.036
肌钙蛋白峰值/(ng/L)[中位数(IQR)]	[152 (48~423)]	[146 (51~410)]	0.031
CK-MB/(U/L)	38.6±24.1	36.9±23.4	0.071
LDL-C/(mmol/L)	2.81±0.79	2.76±0.82	0.063
HDL-C/(mmol/L)	1.02±0.31	1.05±0.29	0.099
eGFR/[mL/(min·1.73m^2)]	78.5±18.3	80.2±17.6	0.095
hs-CRP/(mg/L) [中位数(IQR)]	[8.2 (3.5~15.1)]	[8.5 (3.8~14.9)]	0.057

表 1(续)

具体指标	训练集(n=404)	测试集(n=174)	SMD
白细胞计数(×10^9/L)	9.8±3.2	10.1±3.0	0.096
中性粒细胞百分比/%	74.6±9.1	73.8±8.7	0.090
血红蛋白/(g/L)	132±18	134±17	0.115
血糖/(mmol/L)	7.6±2.3	7.4±2.1	0.091
LVEF/%	52.3±8.7	53.1±8.4	0.093
室壁运动异常指数	1.8±0.6	1.7±0.5	0.178
LVEDD/mm	52.1±5.7	51.8±5.9	0.052
心胸比率	0.52±0.06	0.51±0.05	0.180
Kerley B 线阳性[n (%)]	121 (30.0%)	55 (31.6%)	0.035
肺门模糊征阳性[n (%)]	89 (22.0%)	36 (20.7%)	0.031
心包积液[n (%)]	68 (16.8%)	27 (15.5%)	0.035
STEMI [n (%)]	171 (42.3%)	72 (41.4%)	0.018
NSTEMI [n (%)]	147 (36.4%)	64 (36.8%)	0.008
UA [n (%)]	86 (21.3%)	38 (21.8%)	0.012

2.2　机器学习模型构建结果

2.2.1　特征筛选结果

通过两独立样本非参数检验(Mann-Whitney U 检验/卡方检验)对初始 103 个特征进行单变量预筛选, 排除与终点事件无关的特征(P≥0.10), 保留部分潜在预测因子; 在采用 SMOTE 算法对训练集进行类别平衡后, 通过 LASSO 回归(λ=0.032 时特征压缩至 28 个)和 RF 基尼重要性评估并行计算, 再结合递归特征消除交叉验证与 Bootstrap 重抽样(1000 次)稳定性验证, 从剩余特征中最终识别出 18 个高频稳定出现(入选率大于 80%)的核心特征子集。该子集涵盖四大病理生理维度: 炎症反应维度的 hs-CRP、中性粒细胞百分比和中性粒细胞/淋巴细胞比值(neutrophil-to-lymphocyte ratio, NLR); 心肌损伤维度的肌钙蛋白 I 峰值、CK-MB 和白细胞(white blood cell, WBC); 心功能维度的 LVEF、室壁运动异常指数、左心室舒张末期内径和心包积液; 代谢维度的估算 eGFR、血糖、HDL-C 和 LDL-C; 临床背景维度的年龄、既往心肌梗死史、糖尿病史和吸烟史。该特征子集涵盖了炎症、心肌损伤、心功能、代谢及临床背景等多个病理生理维度, 反映了 ACS 发生发展的多因素性。其中, 炎症反应维度以 hs-CRP 和 NLR 为代表, 提示免疫炎症在斑块不稳定中的核心作用[14]; 心肌损伤维度以 cTnI 和 CK-MB 为核心, 直接反映了心肌细胞坏死程度[15]; 心功能维度以 LVEF 和室壁运动异常指数为主, 体现了心脏泵功能及局部室壁运动的改变[16]; 代谢维度中 eGFR 和 HDL-C 提示肾功能和脂代谢状态对心血管风险的调控作用[17]。超声心动图参数(LVEF、室壁运动指数)的贡献率达 34.7%, 表明心功能状态是 ACS 风险分层的重要维度, 与检验指标形成互补。图 1 展示了该特征子集的皮尔逊相关系数热力图。

2.2.2　特征构建机器学习模型性能

在测试集(n=174)的验证中, XGBoost 模型的预测效能最优, 其 AUC 达 0.901 (95% CI: 0.872~0.927), 显著优于传统风险评分(DeLong 检验 Z=5.47, P<0.001)。与其他机器学习模型对比显示: Light GBM (AUC 0.873, 95% CI: 0.823~0.923)、SVM (0.841, 95% CI: 0.786~0.896)、KNN (0.832, 95% CI: 0.775~0.889) 和 RF (0.827, 95% CI: 0.769~0.885)的判别效能均低于 XGBoost (DeLong 检验, P<0.05)。基于 18 维特征子集(涵盖炎症反应、心肌损伤、心功能及代谢四大维度)的分类指标进一步证实: XGBoost 的准确性(86.4%)、敏感性(83.2%)和特异性(88.7%)均居首位, 其 F_1 分数达 0.834(详见表 2)。所有模型均通过校准度验证(Brier 评分: XGBoost 0.089, Light GBM 0.101, SVM 0.113, RF 0.124, KNN 0.119; Hosmer-Leme show 检验均为 P>0.05)。经 1000 次 Bootstrap 重抽样, XGBoost 的 AUC 标准差最低(±0.018), 显著优于其他模型(Light GBM±0.020, SVM±0.023, KNN±0.024, RF±0.025), 表明其具有最稳定的泛化能力(图 2)。XGBoost 模型在各项评价指标上均表现最优, 显示出其在处理混合类型数据和捕捉特征非线性关系方面的优势[16]。该模型基于入院 24 h 内可完成的常规检测及心功能参数, 通过极简指标组合实现了高危人群的快速识别, 具备良好的临床可操作性。

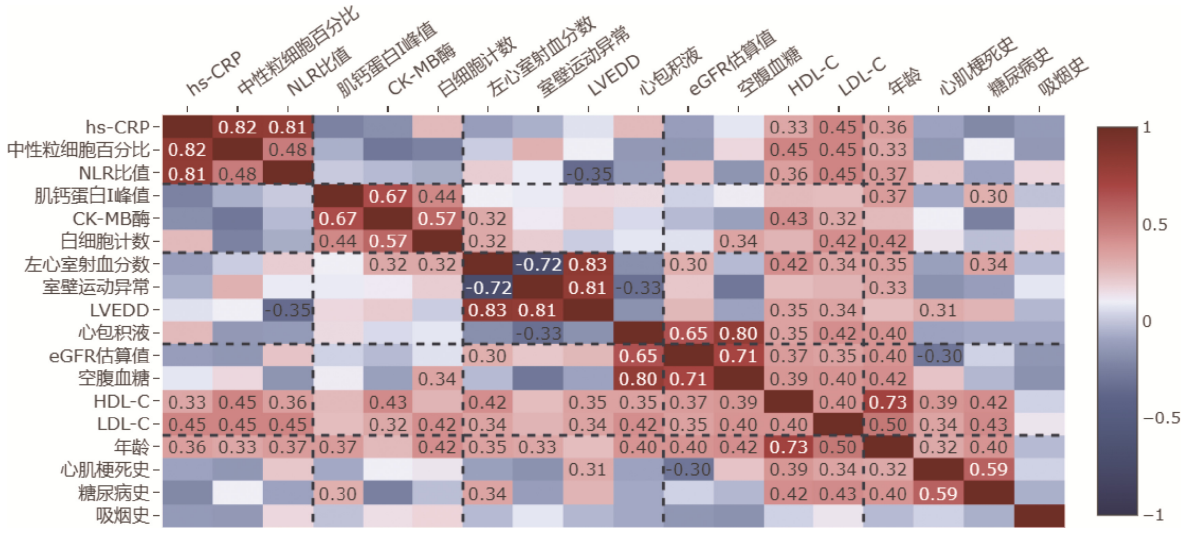

图 1　特征子集的皮尔逊相关系数热力图

表 2　特征构建的 5 种机器学习模型性能(%)

模型	AUC (95% CI)	敏感性(95% CI)	特异性(95% CI)	阳性预测值(95% CI)	阴性预测值(95% CI)	准确性(95% CI)	F_1 分数
XGBoost	0.901 (0.872~0.927)	83.2% (76.4~89.1)	88.7% (83.5~93.9)	86.1% (79.8~91.7)	86.8% (81.2~91.6)	86.4% (81.2~91.6)	0.834
Light GBM	0.873 (0.823~0.923)	79.6% (72.1~86.3)	85.2% (79.1~90.5)	82.4% (75.3~88.4)	82.9% (76.5~88.3)	84.5% (79.0~89.4)	0.798
SVM	0.841 (0.786~0.896)	76.8% (69.0~83.8)	83.9% (77.5~89.2)	80.1% (72.7~86.5)	80.9% (74.3~86.7)	82.2% (76.3~87.4)	0.772
KNN	0.832 (0.775~0.889)	75.2% (67.2~82.3)	82.5% (75.9~88.0)	78.6% (70.9~85.2)	79.3% (72.5~85.3)	81.0% (74.9~86.4)	0.761
RF	0.827 (0.769~0.885)	74.1% (66.0~81.3)	81.3% (74.5~87.1)	76.9% (69.0~83.7)	78.1% (71.1~84.3)	80.5% (74.3~85.9)	0.749

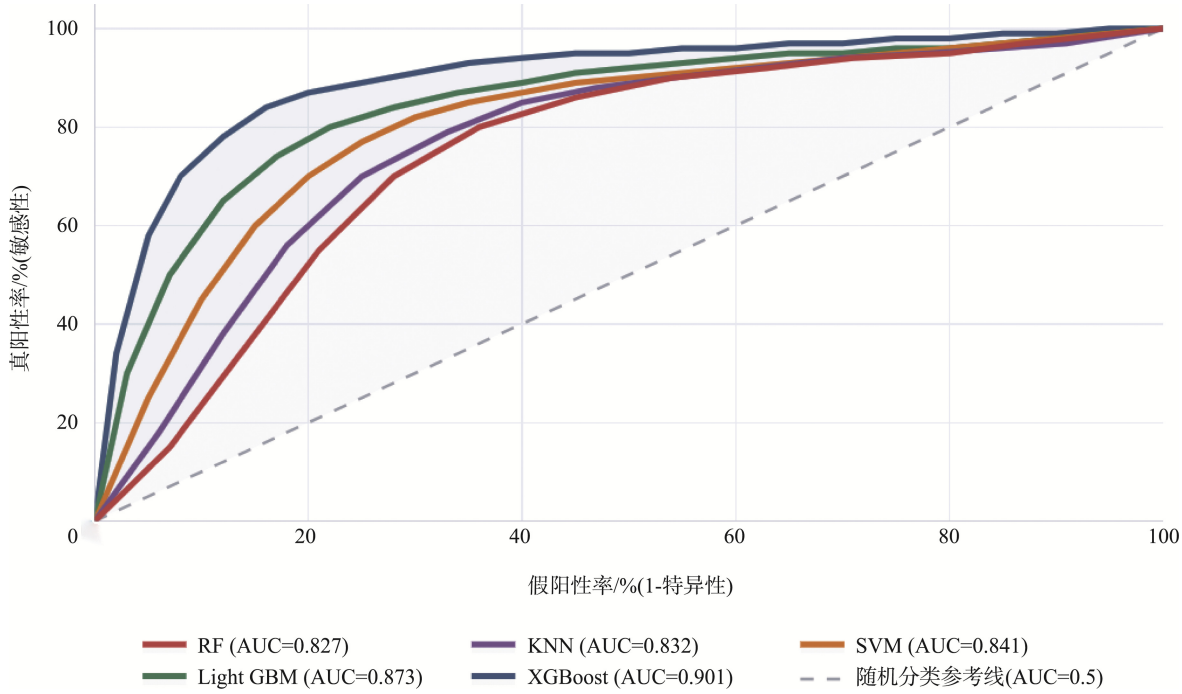

图 2　特征模型的 ROC 曲线

2.3　Shapley 加法解释可视化最佳机器学习模型

基于 SHAP (shapley additive explanations, v0.42.0)方法对 XGBoost 模型的解析显示(图 3), LVEF 以 SHAP 均值绝对值|0.32|为最具影响力的预测变量, 其依赖图揭示了关键病理转折现象: 当 LVEF 降至 50%临界值以下时, SHAP 值发生显著负向偏移(Δ=–0.41), 客观量化了心功能不全的恶化机制。hs-CRP 贡献度(|0.28|)位列第二, 在质量浓度大于 8 mg/L 区间呈现指数级风险增长(SHAP 斜率 k=0.33), 明确表征炎症反应对动脉斑块稳定性的破坏效应。定量分析证实: LVEF 每降低 5%单位对应 SHAP 风险值升高 0.12 (95% CI: 0.09~0.15), 而 hs-CRP>10 mg/L 时风险增幅达 0.51 (P<0.001)。

图 4 的 SHAP 依赖分布蜂窝图深入解析了个体风险贡献模式: 垂直轴按特征整体影响力降序排列(LVEF 最高, SHAP=0.407), 水平轴以零基线划分风险驱动因子(正值域)与保护因子(负值域)。色谱映射清晰界定临床阈值——红色区域(如 LVEF<50%)标记高危状态, 绿色区域(如 HDL-C>1.42 mmol/L)指示保护效应。SHAP 分析首次量化了多项核心临床规律。LVEF<50%作为心功能恶化的临界拐点, 从功能学层面揭示了心室泵血能力下降对 ACS 风险的直接影响。hs-CRP>8 mg/L 与 cTnI>120 ng/L 的协同高危效应, 表明炎症激活与心肌损伤在病理进程中相互促进, 共同推动高危状态的形成。HDL-C>1.42 mmol/L 联合 eGFR>75 mL/min 产生显著风险缓冲效应, 提示良好的脂代谢状态和肾功能可协同发挥心血管保护作用, 这一定量关系为代谢干预提供了新的临床参考阈值。NLR>5.0 对 LVEF 降低风险的正向协同放大, 进一步印证了炎症-心功能耦联机制在 ACS 进展中的关键作用。

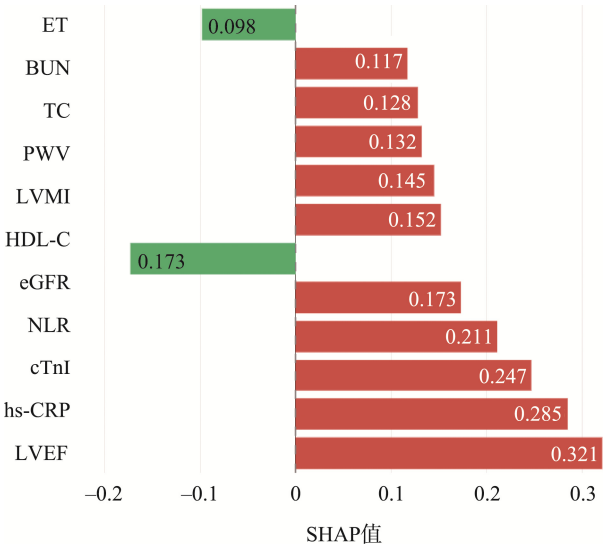

图 3　特征重要度

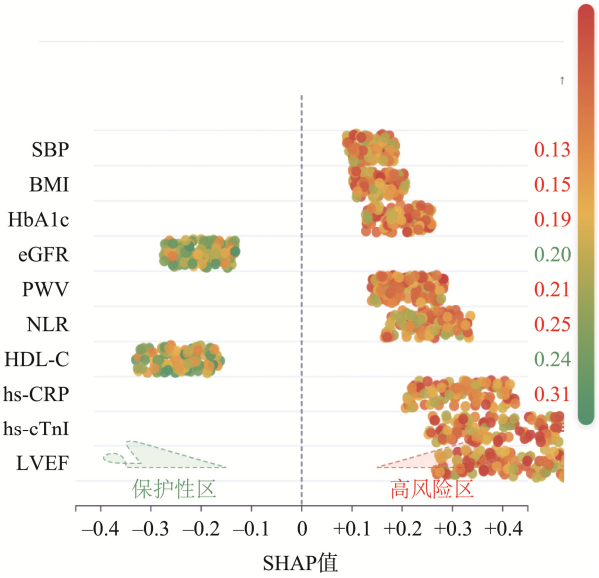

图 4　SHAP 蜂窝图

3　讨论与结论

本研究基于 18 项常规检验、影像及临床指标, 通过多阶段特征筛选和 XGBoost 算法构建了 ACS 早期风险预测模型。模型在独立测试集中表现出优异的判别效能(AUC 0.901), 显著优于传统风险评分, 与既往基于 XGBoost 的心血管预测研究[18-19]结果一致。SHAP 可解释性分析首次量化界定了 LVEF<50%、hs-CRP>8 mg/L 等关键临床阈值, 揭示了 hs-CRP 与 cTnI 的协同高危效应以及 HDL-C 联合 eGFR 的协同保护效应, 为 ACS 风险分层提供了新的量化依据。优化后的模型基于入院 24 h 内可完成的常规检测, 具有良好的临床可操作性和可解释性。然而, 本研究也存在一些局限性。首先, 作为回顾性单中心设计, 可能存在选择偏倚, 模型的泛化能力有待多中心外部数据验证。其次, 样本量相对有限(n=578), 虽通过 Bootstrap 重抽样验证了特征稳定性, 但扩大样本队列有助于进一步提升模型的稳健性。其次模型尚未充分挖掘影像原始数据(如超声心动图原始图像)的潜力, 可能损失部分预测信息。未来的研究应开展多中心前瞻性验证, 纳入更广泛的基层医疗机构人群; 整合时序数据构建动态风险监测体系; 并结合深度学习技术优化影像特征自动提取流程, 以进一步提高模型的预测效能和临床适用性。通过这些改进, 有望为急诊及基层医疗场景提供更加精准、高效的 ACS 早期预警工具。

参考文献

[1] 周仟慧, 古满平. 急性冠脉综合征不良心血管事件机器学习预测模型的研究进展[J]. 现代医药卫生, 2024, 40(17): 3023-3027.

[2] 朱星宇, 张凯杰, 隗蕊, 等. 急性冠状动脉综合征患者介入术后主要不

良心血管事件风险预测模型的研究进展[J]. 中华老年心脑血管病杂志, 2024, 26(7): 846-848.
[3] 施景文, 孙颖. 急性冠脉综合征预后模型研究进展[J]. 中国实用内科杂志, 2024, 44(7): 607-610.
[4] 王雅倩. 基于机器学习及可解释框架面向县级医院ACS行PCI人群的远期死亡危险因素分析及风险预测模型研究[D]. 沈阳: 中国医科大学, 2023.
[5] 赖润民. 急性冠脉综合征易损斑块病证结合预测模型构建及生物学机制研究[D]. 北京: 中国中医科学院, 2024.
[6] 钱尔堃. 急性冠脉综合征(ACS)患者 PCI 术后发生急性肾损伤的危险因素及预测模型的建立[D]. 广州: 广州医科大学, 2024.
[7] 杨贵分, 张少华, 刘卫珍, 等. 两种机器学习模型对急性冠脉综合征患者发生院内心搏骤停的预测价值比较[J]. 实用心脑肺血管病杂志, 2024, 32(3): 39-42, 48.
[8] 刘秉时. 生物信息学分析联合机器学习算法预测急性冠脉综合征相关生物标志物[D]. 承德: 承德医学院, 2024.
[9] 魏思萌, 石亚君, 陈韵岱. 远程心电人工智能技术在急性冠脉综合征诊疗中的应用进展[J]. 实用心电学杂志, 2024, 33(1): 98-103.
[10] 周乐, 王珏, 张尉华, 等. 人工智能在急性冠脉综合征诊疗中的应用[J]. 心电与循环, 2023, 42(6): 505-509.
[11] 黄萱, 木胡牙提·乌拉斯汗, 陆晨, 等. 基于机器学习的急性冠脉综合征患者急性肾损伤的预测模型[J]. 临床肾脏病杂志, 2023, 23(8): 653-662.
[12] 李汶键. 基于神经网络的冠心病诊断与预后算法研究与实现[D]. 成都: 四川大学, 2023.
[13] 方明. 基于机器学习分析铁死亡相关基因参与心力衰竭的临床研究[D]. 衡阳: 南华大学, 2023.
[14] 陈海威. 基于冠脉CTA影像特征的机器学习模型预测冠心病人群ACS事件的临床队列研究[D]. 北京: 中国人民解放军医学院, 2022.
[15] 王琼娇. 基于 CCTA 成像急性冠脉综合征斑块周围脂肪衰减指数和放射组学特征分析[D]. 大连: 大连医科大学, 2023.
[16] 景梦园. 基于冠状动脉周围脂肪组织影像组学预测急性冠状动脉综合征[D]. 兰州: 兰州大学, 2022.
[17] 文武成, 赵波洋, 盘中贤, 等. 基于临床特征和 CTA 冠周脂肪影像组学的机器学习模型预测急性冠状动脉综合征的应用价值[J]. 影像诊断与介入放射学, 2024, 33(6): 403-410.
[18] 白琳. 机器学习对急性冠脉综合征危险程度的预测研究[D]. 成都: 四川大学, 2022.
[19] 孔艺璇. 急性冠脉综合征患者 CCTA 影像组学特征及与中医证型相关性研究[D]. 南宁: 广西中医药大学, 2022.