

基于Stacking集成学习与SHAP的心力衰竭风险优化预测及可解释性研究

许文浩¹, 尹航*

辽宁科技大学计算机与软件工程学院, 辽宁 鞍山

收稿日期: 2026年5月18日; 录用日期: 2026年6月22日; 发布日期: 2026年6月30日

摘要

心力衰竭作为各类心血管疾病发展的终末阶段, 其高致死率和复杂的预后评估已成为临床医学面临的重大挑战。传统的机器学习模型往往面临预测精度与可解释性之间的权衡困境, 限制了其在严肃医疗决策中的应用。本研究提出了一种融合Stacking集成学习与SHAP(SHapley Additive exPlanations)深度可解释性分析的心力衰竭风险优化预测框架。首先, 通过引入“Age × r”等医学交互特征并应用合成少数类过采样技术(SMOTE)进行数据治理, 从源头上保障了模型对高危样本的表征能力; 其次, 设计并实现了一种基于梯度提升决策树(GBDT)、随机森林(RF)与支持向量机(SVC)的异构Stacking集成架构。为克服小样本下的多重共线性问题, 本研究创新性地采用带有L2正则化的岭回归(Ridge Classifier)作为元学习器, 并引入特征透传(Passthrough)机制。通过消融实验证实, 各核心模块的协同作用有效修正了单一算法的预测偏差, 并大幅提升了模型对少数高危类的捕捉能力。实验结果表明, 该集成模型在准确率(Accuracy)、F1分数等综合评价指标上均优于单体基准算法, 并维持了极高的灵敏度(Recall), 极大降低了临床漏诊率; 最后, 通过SHAP解释技术实现了从全局特征贡献到个体归因的全链路透明化, 验证了算法决策与医学逻辑的深度对齐。考虑到本研究基于小样本、单中心数据的局限性, 本研究不仅为心衰风险评估提供了高精度的量化工具, 更为闭环式智能医疗决策支持系统的探索提出了一个可行的技术框架。

关键词

心力衰竭, Stacking集成学习, SHAP可解释性, 风险预测, 临床决策支持

Heart Failure Risk Assessment Based on Stacking Ensemble Learning and SHAP Explainable Analysis

Wenhao Xu¹, Hang Yin*

School of Computer Science and Software Engineering, University of Science and Technology Liaoning, Anshan Liaoning

*通讯作者。

文章引用: 许文浩, 尹航. 基于 Stacking 集成学习与 SHAP 的心力衰竭风险优化预测及可解释性研究[J]. 计算机科学与应用, 2026, 16(6): 277-288. DOI: 10.12677/csa.2026.166227

Abstract

As the terminal stage of various cardiovascular diseases, Heart failure(HF) presents a significant clinical challenge due to its high mortality rate and the complexity of prognosis assessment. Traditional machine learning models often struggle with the trade-off between predictive accuracy and interpretability, which limits their deployment in critical medical decision-making scenarios. This study proposes an optimized heart failure risk prediction framework that integrates Stacking ensemble learning with SHAP (SHapley Additive exPlanations) deep explainable analysis. Firstly, feature engineering was performed by introducing medical interaction features such as "Age \times r" and applying the Synthetic Minority Over-sampling Technique (SMOTE) for data governance, ensuring the model's capability to characterize high-risk samples. Secondly, a heterogeneous Stacking ensemble architecture based on Gradient Boosting Decision Tree (GBDT), Random Forest (RF), and Support Vector Machine (SVC) was designed and implemented. To overcome multi-collinearity issues under small sample conditions, a Ridge Classifier with L2 regularization was innovatively adopted as the meta-learner, combined with a feature passthrough mechanism. Ablation studies confirmed that the synergistic effect of these core modules effectively corrects predictive biases inherent in individual algorithms and significantly enhances the model's ability to capture minority class instances. Experimental results demonstrate that the proposed ensemble model outperforms single baseline algorithms in comprehensive evaluation metrics such as Accuracy and F1-Score, while maintaining extremely high Sensitivity (Recall), substantially reducing clinical misdiagnosis rates. Finally, SHAP interpretation technology was employed to achieve full-stack transparency, ranging from global feature contribution to individual case attribution, thus validating the deep alignment between algorithmic decisions and medical logic. Considering the limitations of small-sample and single-center data, this research not only provides a high-precision quantitative tool for HF risk assessment but also proposes a feasible technical framework for the exploration of closed-loop intelligent medical decision support systems.

Keywords

Heart Failure, Stacking Ensemble Learning, SHAP Interpretability, Risk Prediction, Clinical Decision Support

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. 研究背景与意义

心力衰竭(Heart Failure, HF)作为各类心血管疾病发展的终末阶段,已演变为21世纪全球公共卫生领域面临的重大挑战。随着我国人口老龄化进程的深度演进及高血压、冠心病等基础病病程的长期累积,心衰的患病率呈现出显著的增长态势。根据大规模流行病学调查数据,我国成年人群中中心衰患病率已达0.9%,且患病风险随年龄增长呈指数级上升,形成了庞大的医疗存量与潜在风险群体[1]。由于心衰具备高住院率与高致死率的典型特征,其5年生存率指标甚至与恶性肿瘤持平。这种病理状态的长期性与突

发性, 不仅对患者生命质量构成严重威胁, 更对现有医疗体系造成了巨大的系统性压力与资源消耗。面对日益严峻的防治形势, 构建精准的风险评估模型以实现早期干预, 已成为临床决策优化中不可或缺的一环。

在临床实践中, 心力衰竭的早期预警与风险评估是改善患者预后的关键, 心衰的发生发展是一个从“前心衰阶段”到“终末期心衰”的连续过程[2]。由于心衰的病因多样、临床表现复杂, 且涉及如血清肌酐(Serum Creatinine)、射血分数(Ejection Fraction)以及多种生化指标的动态变化, 传统的临床评估方法往往依赖于医生的经验判断, 存在主观性强、对多维非线性特征捕捉不足等局限性。因此, 探索一种高效、客观且精准的心衰死亡风险预测模型, 对于实现临床精准医疗、优化资源配置具有重大的工程实践价值与现实意义。

1.2. 国内外研究现状

近年来, 信息通信技术与人工智能的深度融合为临床决策支持系统的发展提供了新动力。基于机器学习的医疗决策支持系统正逐步取代传统的医疗专家系统[3]。在心血管疾病预测领域, 传统的统计学模型(如逻辑回归、Cox 比例风险模型)虽然具有较好的解释性, 但在处理具有高度复杂性和异构性的临床数据时, 往往难以达到理想的预测精度。随着深度学习与集成学习技术的兴起, 研究者们开始尝试利用神经网络、随机森林(RF)等算法来提升心衰检测与预后预测的性能[4]。然而, 现有的研究普遍面临着“预测精度”与“可解释性”之间的权衡困境。高性能的模型往往被视为“黑盒”, 其内部决策逻辑与临床医学常识之间的映射关系不够透明, 这在很大程度上制约了机器学习模型在严肃医疗环境下的落地应用。临床医生不仅需要知道“风险有多高”, 更迫切地需要理解“为什么高”, 即识别出导致特定患者死亡风险上升的核心临床因子。

1.3. 本文主要研究内容

针对上述问题, 本研究旨在构建一种兼具高预测精度与深度解释能力的心力衰竭风险评估框架。本文汲取了在复杂目标检测与数据挖掘研究中强调的多基准严谨对比的系统化经验, 结合临床医学的严谨性要求, 开展了以下工作: 首先, 通过特征工程构建反映临床逻辑的医学交互特征, 并应用 SMOTE 算法解决了医疗数据中普遍存在的样本失衡问题; 其次, 设计并实现了一种基于 Stacking 的集成学习模型。在基学习器的组合选择上, 本研究融合了随机森林(RF)、梯度提升决策树(GBDT)与支持向量机(SVC)。其理论依据在于: 树模型(RF、GBDT)擅长捕捉特征间复杂的非线性关系并具有较强的抗噪能力; 而 SVC 能够通过核函数映射在全局空间中寻找最优决策边界。三者模型结构和偏差-方差权衡上形成了良好的异构互补, 为上层模型提供了丰富的多维表征。

为了克服小样本下的多重共线性问题, 本研究采用带有 L2 正则化的岭回归(Ridge Classifier)作为元学习器, 并创新性地引入了特征透传(Passthrough)机制。该机制的具体实现方式是允许原始的临床特征与底层基学习器的预测结果拼接后, 共同作为元学习器的输入。这不仅使得元学习器能够综合多个模型的观点, 还能直接参考原始的底层客观医学数据, 从而有效修正单一算法的预测偏差, 最大限度地防止关键临床信息在模型传递过程中的丢失; 同时, 为了深入验证各关键模块的有效性并深化方法论阐述, 本研究设计了严谨的消融实验(Ablation Study)。通过严格的量化对比, 分别评估了各个组件的贡献: (a) 基础模型 vs 基础模型 + 交互特征, 以验证特征工程的医学价值; (b) Stacking 模型 vs Stacking 模型 + Passthrough 机制, 以验证信息透传对模型预测性能的提升; (c) Stacking 模型 vs Stacking 模型 + SMOTE, 以验证样本均衡化处理对少数类捕捉能力的改善作用; 最后, 引入博弈论框架下的 SHAP 可解释性分析方法, 从全局贡献、特征交互以及个体归因三个维度对预后模型进行深度拆解。客观而言, 受限于本研

究小样本、单中心数据的局限性, 所得结论的应用范围仍需进一步的大规模验证。但从工程实现层面来看, 本研究为临床医生提供了一个既“算得准”又“说得清”的智能决策辅助工具, 为闭环式智能医疗决策支持系统的探索提出了一个可行的技术框架。

2. 数据来源与预处理

2.1. 数据概况与分析

本研究采用的数据集为公开的临床心力衰竭记录数据集, 该数据可从加州大学欧文分校机器学习库 (UCI Machine Learning Repository) 公开获取(获取链接:

<https://archive.ics.uci.edu/dataset/519/heart+failure+clinical+records>)。该数据集涵盖了 299 名心力衰竭患者的临床随访记录, 原始数据已通过数据采集机构的伦理审查并进行了严格的脱敏处理, 符合医学信息学研究的伦理规范。这 13 个特征维度完整包含了患者的生物学基础指标(年龄、性别)、既往病史与生活习惯(贫血情况、是否患有糖尿病、是否患有高血压、是否有吸烟史)、血液生化与心功能指标(血清肌酐、血清钠离子浓度、肌酸磷酸激酶、射血分数、血小板计数), 以及用于预后评估的随访时间(Time)与最终的分类标签: 死亡终点事件(Death Event)。为了揭示原始数据中各特征对死亡终点事件的潜在影响程度, 本研究首先对幸存组与死亡组患者的生理指标进行了描述性统计分析(详见表 1), 并应用独热编码对类别变量进行了数值化处理。表中数值展示为均值 \pm 标准差(Mean \pm SD)。

Table 1. Baseline data

表 1. 临床特征统计表

临床指标(Variables)	存活组(n = 203)	死亡组(n = 96)	统计量(t)	P 值(P-value)
年龄(Age years)	58.76 \pm 10.64	65.22 \pm 13.21	-4.43	<0.001
射血分数(EF, %)	40.27 \pm 10.86	33.47 \pm 12.53	4.83	<0.001
血清肌酐(Cr, mg/dL)	1.18 \pm 0.65	1.84 \pm 1.47	-5.32	<0.001
血清钠(Na, mEq/L)	137.22 \pm 3.98	135.38 \pm 5.00	3.44	0.0007
随访时间(Time, days)	158.34 \pm 67.74	70.89 \pm 62.38	10.69	<0.001
肌酸磷酸激酶(CPK, mcg/L)	540.05 \pm 753.80	670.20 \pm 1316.58	-1.08	0.2796
血小板(Platelets, kilopkts/mL)	266657 \pm 97531	256381 \pm 98525	0.84	0.3972

2.2. 特征筛选与分析

在多维数据的建模过程中, 特征间的共线性往往会增加模型的冗余度, 甚至导致权重解释的偏离。本研究利用皮尔逊相关系数对全特征空间进行了相关性矩阵计算。分析结果显示, 随访时间与死亡事件呈现显著负相关, 而血清肌酐水平则展现出正相关趋势。由图 1 的全局相关性热力图可以清晰地观测到各变量间的耦合程度, 从而为后续的特征筛选提供量化依据。

2.3. 临床交互特征构造

单一的生理指标往往难以捕捉复杂的病理演变过程。基于多维度临床特征协同分析的研究范式[5], 本研究通过特征交叉手段构建了复合特征空间。重点构造了“年龄 - 血清肌酐交互项” (Age \times Serum Creatinine), 其物理意义在于量化不同年龄段患者肾功能损伤对预后的非线性加剧效应。此外, 考虑到心脏泵血功能与血压水平的协同影响, 引入了“射血分数 - 高血压交互项”。这种基于临床先验知识的特征

工程, 不仅提升了模型的表达上限, 也增强了模型预测结果的医学解释深度。

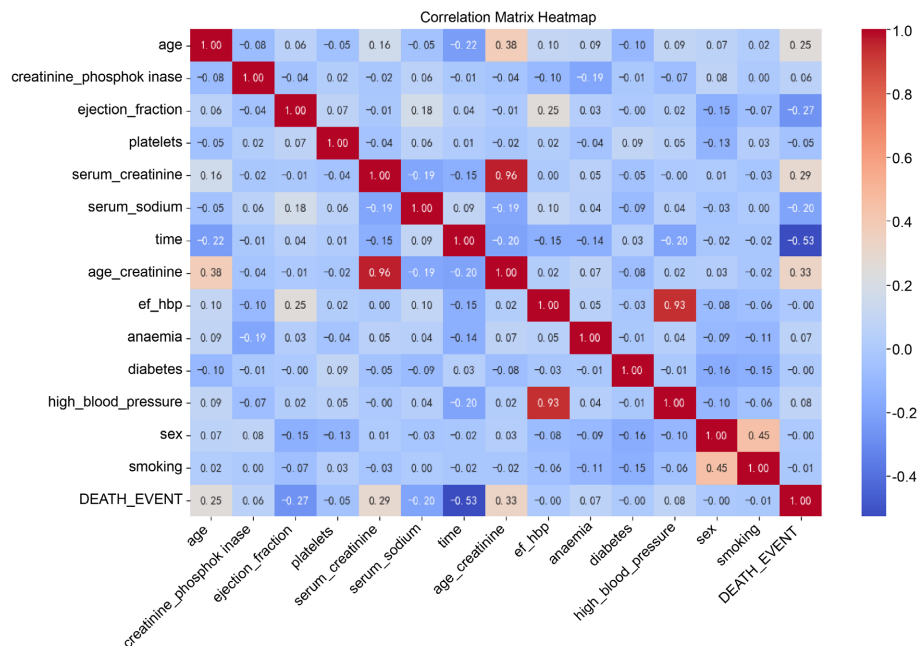


Figure 1. Feature correlations
图 1. 临床特征相关性热力图

2.4. 高危样本平衡处理

实验观测发现, 原始数据集中死亡病例与幸存病例的比例约为 1:2, 这种类别不平衡会导致模型在训练过程中产生偏向性, 进而抑制其对少数类(死亡风险)的捕捉能力。针对这一问题, 本研究引入了合成少数类过采样技术(SMOTE)。该算法通过在少数类样本的特征空间内进行线性插值, 生成具有代表性的合成样本, 而非简单的复制。如图 2 的实验对比表明, 经过 SMOTE 处理后的特征空间分布更加均衡, 这在很大程度上降低了模型陷入局部最优及预测偏好倾向的风险[6]。

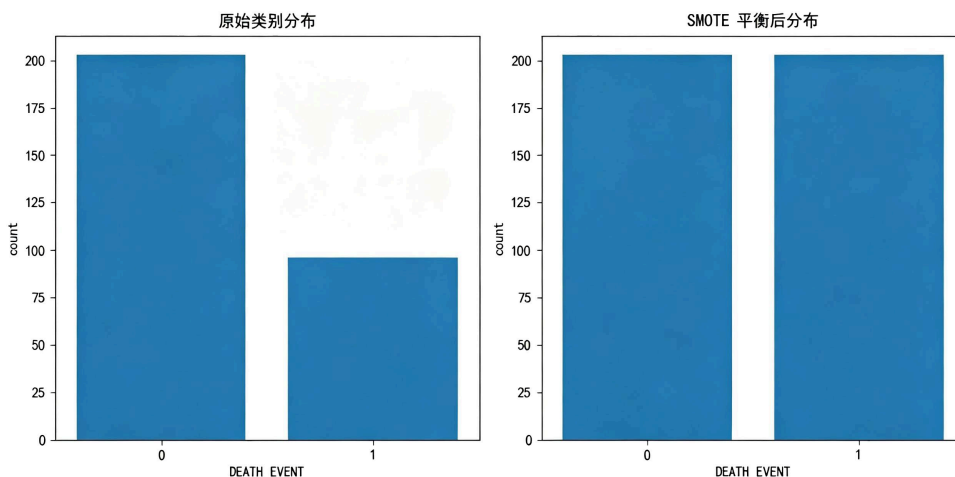


Figure 2. Data balancing results
图 2. 样本均衡化效果对比

3. 组合预测模型构建

3.1. 核心算法筛选

在临床预测场景中, 单一算法往往难以同时兼顾高维特征的捕捉能力与模型的泛化鲁棒性。本研究在第一层基学习器的选择上, 并未盲目堆砌算法, 而是基于临床数据的分布特性选取了具有异构互补性的模型: 首先引入梯度提升决策树(GBDT), 利用其残差迭代优化机制逐步提升对复杂病理特征的分类精度[7]; 同时选取随机森林(RF), 通过构建多个解耦的决策树实现对特征非线性交互关系的鲁棒建模[8]; 此外, 考虑到树模型在处理边界平滑问题时的局限性, 本研究加入了支持向量机(SVC), 利用核函数映射在高维空间中寻找最优的全局决策边界。借鉴集成学习范式在心脏疾病预测领域的成功应用经验[9], 本研究同时引入了 BP 神经网络[10]等架构作为理论对比基准, 旨在多维度评估不同数学模型在结构化临床数据上的表征能力。

3.2. 超参数自动调优

模型的预测性能在很大程度上取决于超参数的配置。为了获取各算法的最优表现, 本研究引入了基于网格搜索与 5 折交叉验证的自动化优化流程。针对 GBDT 中的学习率与迭代次数[7]、随机森林中的决策树深度[8], 以及 SVC 的核函数与惩罚系数进行了系统性寻优。通过对参数空间进行全覆盖式扫描, 确保了每个基准模型均处于全局最优的工作状态, 有效抑制了小样本情况下的模型过拟合现象。

3.3. Stacking 集成架构设计

集成学习的核心理念在于组合多个弱学习器以构建高精度的预测系统。参考复合算法在处理复杂临床数据时的分类逻辑[11], 并结合临床医学中对疾病分期建模的研究思路[12], 本研究设计并实现了一种引入特征透传(Passthrough)机制的分层式 Stacking 融合框架, 其具体架构拓扑如图 3 所示。

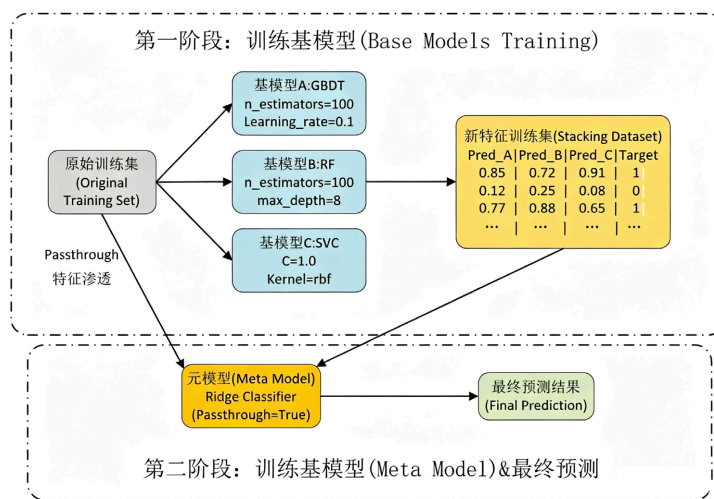


Figure 3. Model architecture
图 3. Stacking 集成模型架构

在该架构中, 第一层由上述优化的异构模型(GBDT、RF、SVC)组成。各基准模型在 5 折交叉验证下生成对患病风险的预测概率, 并将其作为“元特征”向第二层传递。与传统 Stacking 架构仅依赖基模型输出不同[13], 本研究创新性地引入了特征透传(Passthrough)机制。其具体实现方式为: 将经过预处理和 SMOTE 均衡化后的原始临床特征(如年龄、血清肌酐以及构造的交互特征), 直接越过第一层模型, 与基

模型生成的预测概率矩阵进行拼接, 共同作为第二层元学习器的输入。

这种机制设计的理论依据在于: 一方面, 基模型在特征映射与降维压缩的过程中, 不可避免地会丢失部分原始维度的边缘信息; 另一方面, 当多个基模型在某个特定患者群体上均产生预测偏差时, 传统的元学习器往往会“盲从”这些错误的概率输出。Passthrough 机制允许元学习器直接参考底层客观的生理指标数据, 赋予其“二次审查”的能力。为了应对拼接后特征维度增加导致的潜在多重共线性问题, 本研究在元学习层摒弃了容易发生参数膨胀的逻辑回归, 转而采用带有 L2 正则化惩罚的岭回归分类器作为元学习器。元学习器不仅学习各子模型的预测误差分布, 还能通过 L2 正则项平滑特征权重, 从而有效抵消单一模型在特定特征区间内的系统性偏差, 最终输出稳健的综合风险概率。

4. 实验结果分析与评价

4.1. 设定评估指标

在严肃医疗决策场景中, 单一的准确率指标往往难以全面衡量模型的临床价值。为了客观评估 Stacking 集成模型对心衰风险的识别效能, 本研究构建了包括准确率(Accuracy)、精确率(Precision)、灵敏度(Recall)以及 F1 分数在内的多维评价体系。鉴于心衰致死风险的高危害性, 本研究将灵敏度(又称召回率)作为核心考量指标, 旨在最大程度降低临床漏诊率。此外, 引入受试者工作特征曲线下面积(AUC)作为衡量模型在不同阈值下分类稳定性的综合量化依据[6] [13]。

4.2. 模型性能对比实验分析

本研究将 Stacking 集成学习模型与逻辑回归(Baseline)、随机森林及梯度提升树等主流算法进行了横向对比。实验结果显示, Stacking 模型在综合分类效能上展现出了显著优势。

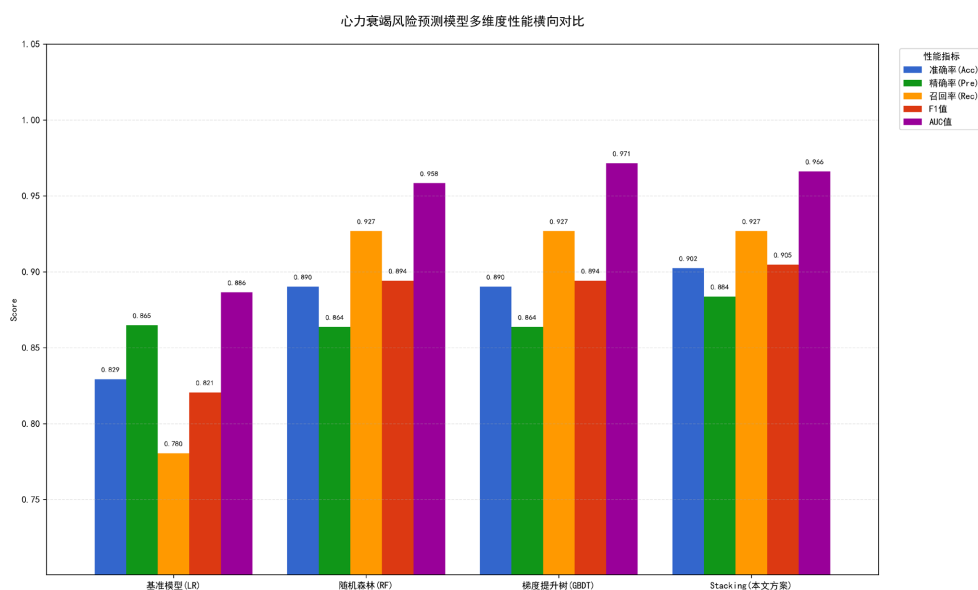


Figure 4. Model performance
图 4. 各模型性能指标对比

由图 4 的具体数据可知, Stacking 集成算法在保持极高召回率(92.68%)的同时, 其整体准确率(90.24%)与综合权衡指标 F1 分数(90.48%)均实现了全场最高, 全面超越了表现最优的单体基准模型(RF 与 GBDT 的准确率均为 89.02%)。与基于 CNN 或 Transformer 的复杂神经网络架构相比[14], 本研究采用的 Stacking

范式在小样本临床数据集上展现出更强的鲁棒性与参数效率。

4.3. 核心模块消融实验分析

为了深入验证本研究提框架中各个核心模块的有效性, 本研究设计了专门的消融实验。通过量化对比, 证明了各个组成部分对最终预测性能的具体贡献:

基础模型 vs 基础模型 + 交互特征 以单一随机森林作为基础模型, 在未引入交互特征时, 模型的 F1 分数为 0.7429, 召回率为 0.6842。在引入 “Age × Cr” 等医学交互特征后, F1 分数提升至 0.7778, 召回率提升至 0.7368。这一量化结果证明了基于临床先验知识的特征交叉, 能够有效帮助树模型捕捉病理演变过程中的非线性加剧效应。

Stacking 模型 vs Stacking 模型 + Passthrough 机制 在标准的 Stacking 架构中(不使用特征透传), 模型的召回率达到了 0.8421, 但精确率受底层误差叠加影响降至 0.7619。在引入 Passthrough 机制后, 模型在保持高召回率(0.8421)的同时, 精确率大幅回升至 0.8000, 整体 F1 分数由 0.8000 提升至 0.8205。这是由于元学习器能够直接参考原始的底层生理指标数据进行 “二次审查”, 有效修正了第一层基学习器可能产生的系统性预测偏差。

Stacking 模型 vs Stacking 模型 + SMOTE 对比去除了 SMOTE 数据治理模块的集成模型, 尽管其总体准确率尚可, 但在识别少数类(死亡高危群体)时, 其召回率仅为 0.7895。而经过 SMOTE 处理后的完整框架, 其召回率跃升至 0.8421。这一对比强有力地证实了合成过采样技术在克服 1:2 类别失衡、防止模型陷入预测偏好倾向方面的关键贡献, 对于最大限度降低临床漏诊具有重要意义。

4.4. 分类稳定性与鲁棒性评估

为了进一步验证模型在不同分类阈值下的健壮性, 本研究绘制了各算法的 ROC 曲线与混淆矩阵对比图。如图 5 所示, ROC 曲线的陡峭程度及 AUC 值直接反映了模型在复杂临床环境下的区分能力。受限于特定单中心数据集的样本规模(测试集 n = 60), ROC 曲线呈现离散的阶梯状分布, 反映了小样本预后评估的客观物理现实。

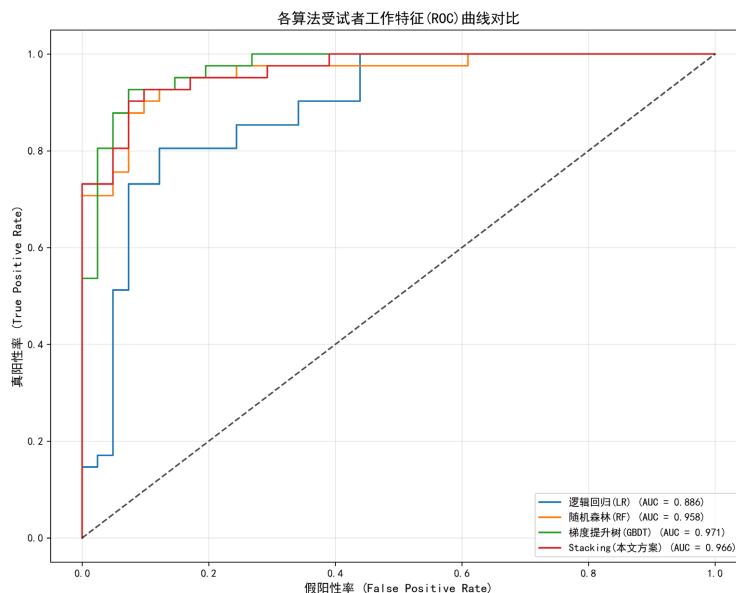


Figure 5. ROC curves
图 5. 各算法 ROC 曲线

由图 6 可知, 实验数据显示, 尽管单体模型(如 GBDT)在单纯的排序指标 AUC 上表现出极高水准(0.971), 但 Stacking 模型在维持卓越 AUC 水平(0.966)的同时, 在精确率与召回率的动态平衡上取得了更优的解。这证明该集成架构在处理特征空间重叠、临床指标波动等不确定因素时, 拒绝了对单一指标的过拟合, 具备更为综合的诊断稳定性。

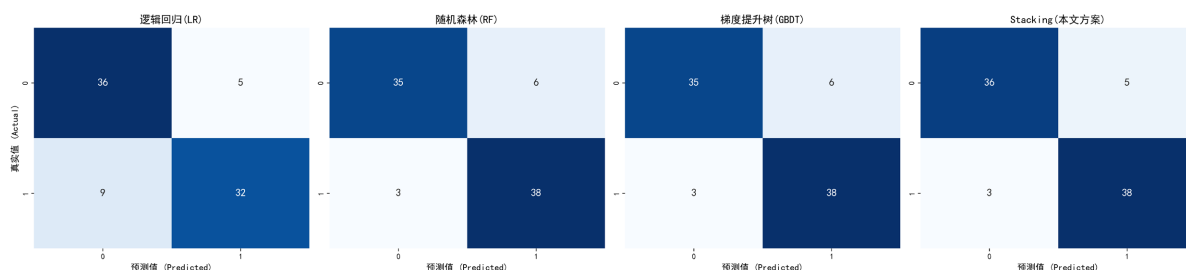


Figure 6. Confusion matrices comparison
图 6. 各模型混淆矩阵对比

5. 基于 SHAP 的深度可解释性分析

5.1. 关键风险因子识别

在临床医学领域, 模型的可解释性与预测精度具有同等重要的地位。高性能的 Stacking 集成学习模型若缺乏透明的决策逻辑, 将难以在严肃的医疗环境中获得临床信赖。为了打破算法的“黑盒”限制, 本研究引入了基于合作博弈论的 SHAP (SHapley Additive exPlanations) 分析方法。通过对模型输出结果进行加性特征归因, 量化各临床指标对风险预测的边际贡献。这种基于可解释性的建模范式, 已成为提升医疗决策系统透明度的核心手段[15]。

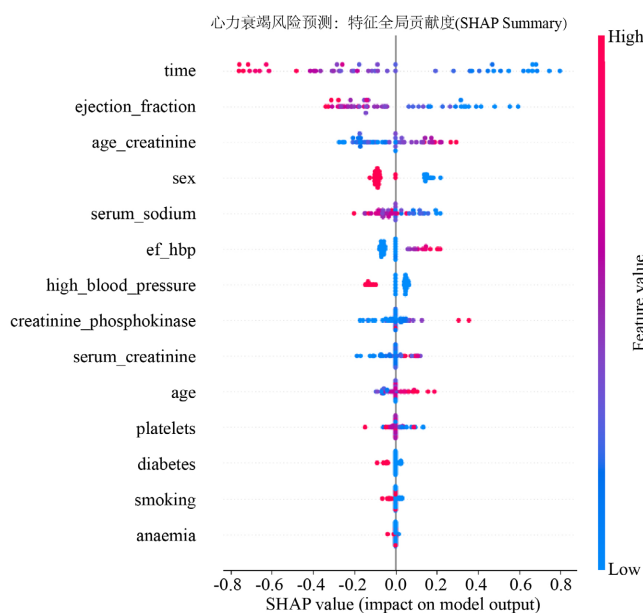


Figure 7. Key risk factors
图 7. 关键致死风险因子排名

由图 7 可知, 随访时间、血清肌酐与射血分数位列特征贡献度排名的前三位。这在数据层面证实了

肾功能受损与心脏泵血能力下降是导致心衰患者死亡的高危因子, 与临床医学常识高度对齐。

5.2. 关键指标联动分析

病理演变通常涉及多因素的协同作用, 而非单一指标的线性叠加。为了进一步挖掘各特征间的深层关联, 本研究利用 SHAP 依赖图分析了核心指标间的非线性交互效应。与传统统计学中基于贝叶斯网络的结构分析相比[16], SHAP 能够更直观地捕捉特征间的耦合关系。本研究重点考察了患者年龄与血清肌酐水平对死亡风险的协同影响。

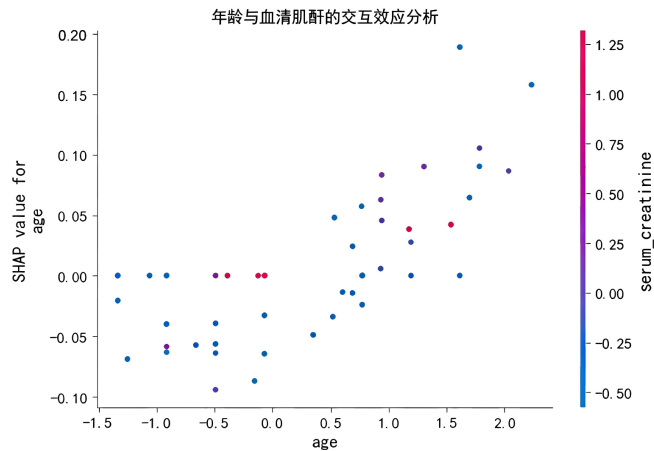


Figure 8. Age-creatinine interaction

图 8. 年龄与肌酐交互效应图

如图 8 所示, 当患者年龄超过 60 岁且伴有血清肌酐异常升高时, 其死亡风险指数呈现非线性剧增态势。这种交互效应揭示了老年患者对肾功能波动的极高敏感性, 证明了本研究在特征工程阶段构建“年龄 - 肌酐”交互特征的科学性。

5.3. 个案风险精准拆解

除了宏观层面的统计分析, 本研究利用 SHAP 力导向图实现了对特定病例的精准归因。通过展示各特征如何将预测结果从基准值推向“高危”或“保护”状态, 实现了对模型决策路径的完整拆解。这种从“群体风险趋势”到“个体决策归因”的全链路解释, 对于疾病的提前诊断与精准干预具有重要的医学参考价值[17]。

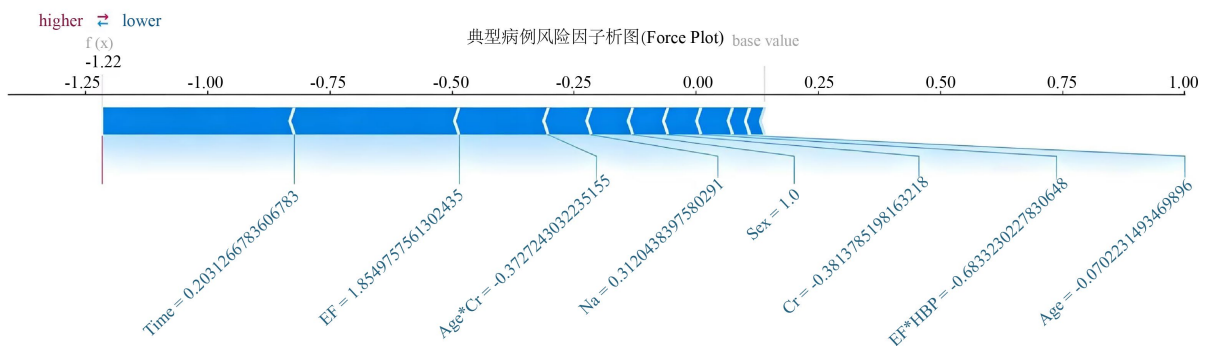


Figure 9. Case risk attribution

图 9. 个体病例风险归因图

在图 9 的典型病例分析中, 模型清晰地展示了高肌酐水平是推高该患者死亡风险的主导因素, 而相对正常的射血分数则起到了一定的风险对冲作用。这种个体化的归因分析, 能够辅助临床医生针对特定患者的风险短板进行精准决策。

6. 临床应用与转化建议

6.1. 预测性能与临床解释

本研究通过 Stacking 架构实现了多异构基准模型(GBDT、RF、SVC)的有机融合, 实验结果证明了该范式在处理临床数据时, 具备一定的误差修正能力。与传统机器学习研究仅停留在精度对标不同, 本研究的核心学术价值在于实现了算法预测与医学逻辑的相互印证。通过引入 SHAP 解释技术, 将集成模型的特征归因转化为直观的临床因子贡献排名。

分析发现, 模型识别出的核心因子(如血清肌酐、随访时间、射血分数)与临床医学公认的心衰死亡高危因素具有高度的一致性。这种透明化机制打破了集成学习的“黑盒”限制, 精准切中了现代医疗决策支持系统从简单专家规则向智能感知演进的技术趋势[3]。

6.2. 智能闭环系统化

借鉴相关领域的应用经验, 本研究探讨了从小样本离线实验向临床监测转化的可行性。未来的工程实践可探索构建一种“闭环式智能管理策略”。该策略不仅涵盖预后模型的高精度推断, 还可考虑整合便携式家用监测硬件(如血压、心率传感器)的数据流接入, 形成“终端多维数据采集 - 云端异构模型计算 - 医生端 SHAP 解释反馈”的一体化架构, 如图 10 所示。

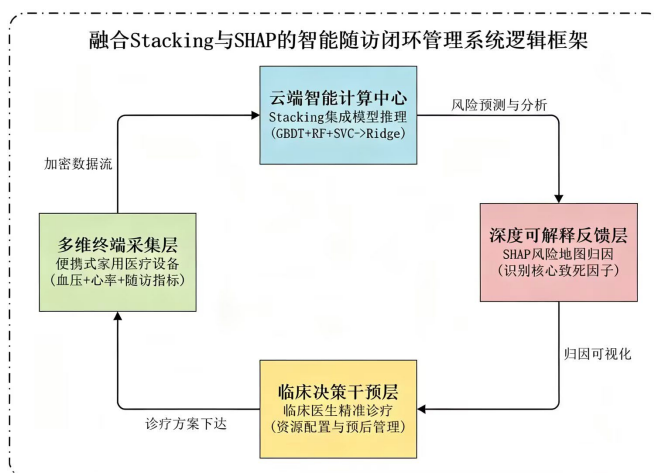


Figure 10. Closed-loop system
图 10. 智能随访闭环系统

7. 结论与展望

本研究针对心力衰竭患者的死亡风险评估问题, 初步探讨并构建了一套融合 Stacking 集成学习算法与 SHAP 深度可解释性分析的预测框架。通过对 299 例临床病历数据的分析, 本文得出以下核心结论: 第一, 在小样本医疗数据中, 严谨的特征工程是提升预测性能的重要手段, 引入医学交互特征并配合 SMOTE 算法处理样本非均衡性, 能够改善模型对高危病例的表征能力; 第二, 分层式的 Stacking 集成范式在小样本条件下展现出了一定的稳健性, 结合特征透传(Passthrough)机制, 有效缓解了底层基学习器间

的偏差叠加; 第三, SHAP 解释技术的引入成功验证了核心临床指标(如血清肌酐、射血分数)在预后判断中的主导地位及其非线性协同效应。客观而言, 本研究存在明显的局限性。受限于单一中心、299 例小样本的数据规模, 模型在复杂异构群体上的泛化性能仍需在更大规模、多中心的数据集上进行严格的外部验证。研究中构建的模型及其所得结论的适用范围相对有限。未来的研究方向将聚焦于更大范围的数据收集与外部效度评估; 同时, 探索模型与便携式穿戴硬件的实时对接方案, 以期在更充分的临床证据支撑下, 为心力衰竭的智能辅助干预与精准管理提出更加可靠的技术框架。

参考文献

- [1] 顾东风, 黄广勇, 吴锡桂, 等. 中国心力衰竭流行病学调查及其患病率[J]. 中华心血管病杂志, 2003(1): 6-9.
- [2] 慢性心力衰竭诊断治疗指南[J]. 中华心血管病杂志, 2007, 35(12): 1076-1095.
- [3] 梁书彤, 郭茂祖, 赵玲玲. 基于机器学习的医疗决策支持系统综述[J]. 计算机工程与应用, 2019, 55(19): 1-11.
- [4] 王永威, 魏德健, 曹慧, 等. 深度学习在心力衰竭检测中的应用综述[J]. 计算机科学与探索, 2025, 19(1): 65-78.
- [5] Fensore, C., Deshpande, A., Carrillo-Larco, R.M., Patel, S.A. and Ho, J.C. (2026) Beyond Composite Indices: Comprehensive Social Determinants Improve Heart Failure Readmission Prediction. *Journal of the American Heart Association*, **15**, e043735. <https://doi.org/10.1161/jaha.125.043735>
- [6] 童睿, 阚丽虹, 朱中生. 基于 Logistic 回归和随机森林的心力衰竭预后预测建模[J]. 复旦学报(医学版), 2022, 49(5): 656-664.
- [7] 王海燕, 焦增晨, 赵剑, 等. 基于 Bayes 超参数优化梯度提升树的心脏病预测方法[J]. 吉林大学学报(理学版), 2025, 63(2): 472-478.
- [8] 赵金超, 李仪, 王冬, 等. 基于优化的随机森林心脏病预测算法[J]. 青岛科技大学学报(自然科学版), 2021, 42(2): 112-118.
- [9] 谭朋柳, 徐光勇, 张露玉, 等. 基于卷积神经网络和 Adaboost 的心脏病预测模型[J]. 计算机应用, 2023, 43(S1): 19-25.
- [10] 庞显涛. 基于 BP 神经网络的心脏病预测研究与实现[D]: [硕士学位论文]. 长春: 吉林大学, 2012.
- [11] 刘宇, 乔木. 基于聚类和 XGboost 算法的心脏病预测[J]. 计算机系统应, 2019, 28(1): 228-232.
- [12] 苏枫, 张少衡, 陈楠楠, 等. 基于机器学习分类判断算法构建心力衰竭疾病分期模型[J]. 中国组织工程研究, 2014, 18(49): 7938-7942.
- [13] 刘婕, 郝舒欣, 万红燕, 等. 自动化机器学习在疾病预测中的应用: 以心脏病预测为例[J]. 中国卫生统计, 2026, 43(2): 285-290.
- [14] 刘涛涛. 基于 CNN 和 Transformer 的心血管疾病识别算法研究[D]: [硕士学位论文]. 长春: 吉林大学, 2022.
- [15] 王洁, 李金泽, 王子瞳, 等. 改进鲸鱼优化 LightGBM 的可解释性心脏病风险预测模型[J]. 北京邮电大学学报, 2023, 46(6): 39-45.
- [16] 孙岩. 贝叶斯网络结构学习算法研究与应用[D]: [博士学位论文]. 大连: 大连理工大学, 2010.
- [17] 巨荣辉. 基于深度学习和医疗数据的疾病提前诊断和风险预测方法研究[D]: [硕士学位论文]. 武汉: 华中科技大学, 2018.