基于stacking模型融合与优化算法的ICU脓毒血症患者死亡风险预测研究*

山西医科大学公共卫生学院卫生统计学教研室(030001) 娄一明 赵瑞青 崔 宇 仇丽霞△

【提 要】 **目的** 基于MIMIC-Ⅲ(medical information mart for intensive care Ⅲ)数据库,采用stacking异质集成融合算法构建脓毒血症患者死亡风险预测模型,有助于早期识别和准确预测脓毒血症患者的死亡风险。**方法** 利用MIMIC-Ⅲ数据库中5163例脓毒血症患者数据,采用stacking模型预测死亡风险。研究分为五步:通过miceforest处理缺失值,随机森林提取特征,使用Tomek links欠采样、合成少数类过采样技术(synthetic minority over-sampling technique, SMOTE)以及SMOTETomek组合算法3种方法处理数据不平衡,比较极端梯度提升、分类梯度提升和轻量梯度提升机模型,优化参数。最后,以优化后的3个模型为基础,logistic回归为元分类器构建stacking模型,预测脓毒血症患者死亡风险。**结果** 通过随机森林特征选择确定简化急性生理评分Ⅱ、年龄等14个关键变量,使用网格搜索、灰狼优化和哈里斯鹰优化(Harris hawks optimization, HHO)3种优化算法提升模型性能,其中SMOTETomek-lightGBM+HHO组合最优。最终,使用SMOTETomek处理的初级学习器和logistic回归元分类器的stacking模型准确率为0.9516、精确度为0.9382、召回率为0.9025、AUC为0.9380、F_1分数为0.9516,显示stacking融合策略有效。**结论** 本研究构建的stacking模型在预测准确性和训练效率上均有优势,有助于重症监护室医师早期识别高风险患者,为临床诊治提供科学依据。

【关键词】 MIMIC-Ⅲ 脓毒血症 机器学习 优化算法 stacking集成
【中图分类号】 R195.1 【文献标志码】 A **DOI** 10.11783/j.issn.1002-3674.2026.03.003

Prediction of Death Risk of Patients with Sepsis in ICU based on the Fusion of Stacking Model and Optimization Algorithm

Lou Yiming, Zhao Ruiqing, Cui Yu, et al (*Department of Health Statistics, School of Public Health, Shanxi Medical University, Taiyuan* 030001)

【**Abstract**】 **Objective** Based on the medical information mart for intensive care Ⅲ (MIMIC-Ⅲ) database, constructing a sepsis patient mortality risk prediction model using the stacking heterogeneous ensemble fusion algorithm can help in the early identification and accurate prediction of mortality risks in sepsis patients. **Methods** Utilizing data from 5163 sepsis cases in the MIMIC-Ⅲ database, the stacking model was employed to predict the risk of mortality. The study was divided into five steps: handling missing values with miceforest, feature extraction with random forest, addressing data imbalance with Tomek links, synthetic minority over-sampling technique (SMOTE) and SMOTETomek, comparing extreme gradient boosting, categorical boosting, light gradient boosting machine models, and optimizing parameters. Finally, based on the optimized three models, logistic regression was used as the meta-classifier to construct the stacking model for predicting the mortality risk of sepsis patients. **Results** Through random forest feature selection, 14 key variables including simplified acute physiology score Ⅱ and age were identified. The model performance was enhanced using three optimization algorithms: grid search, grey wolf optimizer, and Harris hawks optimization (HHO) algorithm, with the SMOTETomek-lightGBM+HHO combination proving to be the most optimal. Ultimately, the stacking model with a primary learner processed by SMOTETomek and a logistic regression meta-classifier achieved an accuracy of 0.9516, precision of 0.9382, recall of 0.9025, AUC of 0.9380, and an F_1 score of 0.9516, demonstrating the effectiveness of the stacking ensemble strategy. **Conclusion** The stacking model constructed in this study has advantages in both predictive accuracy and training efficiency, which helps intensive care unit physicians to early identify high-risk patients and provides a scientific basis for clinical diagnosis and treatment.

【**Key words**】 MIMIC-Ⅲ; Sepsis; Machine learning; Optimization algorithm; Stacking integration

重症监护室(intensive care unit, ICU)中的脓毒血症是一种严重的感染性疾病,可能导致多器官衰竭和高死亡率[1]。早期识别和准确预测脓毒血症患者的死亡风险对降低死亡率至关重要[2]。

随着数据挖掘技术的进步,机器学习和人工智能在医疗领域的应用越来越广泛,特别是在脓毒血症的预后评估中[3]。研究者们利用计算机建模和数据分析技术构建预测模型,以提高预后评估的准确性。大量大数据分析研究验证了机器学习模型在医疗预测中的有效性[4]。例如,张平等[5]利用MIMIC-Ⅳ(medical information mart for intensive care Ⅳ)数据库和LASSO算法筛选特征,再基于梯度提升决策树(gradient boosting decision tree, GBDT)算法构建ICU住院时间预测模型,效果良好,有助于早期干预和分配医疗资源。王羽等[6]从MIMIC-Ⅲ数据库中提取心衰患者数据,用LASSO-logistic回归分析死亡预测因子,并构建列线图模型,能准确预测心衰患者院内死亡率。

脓毒症的发病机制复杂,涉及炎症反应、免疫紊乱和凝血功能障碍等多个方面[7]。风险预测模型能够帮助临床医生筛选出危险因素,识别高危患者群体。

*基金项目:国家自然科学基金(81973155)
△通信作者:仇丽霞,E-mail: qlx_1126@163.com

本研究采用回顾性研究方法，基于MIMIC-Ⅲ数据库，针对脓毒血症患者数据的高维、异质性和冗杂性特点，首先使用随机森林（random forest，RF）算法进行特征选择，然后采用托梅克链接（Tomek links）欠采样、合成少数类过采样技术（synthetic minority over-sampling technique，SMOTE）以及SMOTETomek三种方法处理数据不平衡问题。接着，结合极端梯度提升（extreme gradient boosting，XGBoost）、分类梯度提升（categorical boosting，CatBoost）和轻量梯度提升机（light gradient boosting machine，lightGBM）集成算法构建脓毒血症患者死亡风险预测模型。为了提高模型性能，本研究还使用了网格搜索（grid search，GS）、灰狼优化（grey wolf optimizer，GWO）和哈里斯鹰优化（Harris hawks optimization，HHO）等算法对模型进行超参数优化。最终，通过stacking异质集成融合方法，将优化后的XGBoost、CatBoost和lightGBM模型作为初级学习器，逻辑回归作为元分类器，构建了脓毒血症患者的死亡风险预测模型。本模型为临床早期识别高死亡风险脓毒血症患者提供了科学依据，有助于指导临床治疗和干预，降低患者死亡率。

对象与方法

1. 研究对象来源

本研究是一项基于MIMIC-Ⅲ数据库的单中心回顾性临床研究。MIMIC-Ⅲ是一个包含一家大型三级护理医院ICU患者详细信息的大型数据库[8]，涵盖实验室检测、生命体征、医护记录和影像报告等。该数据库的建立和使用已得到美国机构伦理审查委员会的批准，个人身份信息已经被隐匿。研究纳入了MIMIC-Ⅲ数据库中ICD-9编码为脓毒血症的6347例患者，排除重复入院、年龄小于18岁和无实验室相关指标或数据缺失大于30%的患者后，最终纳入5163例患者。研究提取了包括人口统计学、实验室检查、生命体征、合并症和评分信息在内的34个变量，其中14个离散型和20个连续型变量。

2. 研究方法

（1）数据预处理

鉴于MIMIC-Ⅲ数据库数据缺失问题，本研究采用miceforest方法[9]以提高插值的稳健性和准确性。在模型训练前，通过随机森林方法对候选变量进行特征重要性排序[10]，根据Gini系数从大到小选择变量，并通过相关性分析保留研究变量，进一步通过方差膨胀因子（variance inflation factor，VIF）评估变量间的多重共线性。对于VIF值大于5的变量，根据临床意义和特征重要性排序，优先保留对预测目标贡献更大的变量，以降低共线性对模型稳定性的影响。

（2）死亡风险预测模型的构建

Hansen等[11]在1990年提出了集成学习，这是一种通过组合多个弱学习算法来提升模型性能的方法。本研究通过SMOTETomek[12]混合采样算法进行不平衡处理，改善分类器在处理类别不平衡问题时的性能[13]，并采用XGBoost、CatBoost和lightGBM作为初级学习器构建stacking集成学习模型[14-15]。另外，本文以AUC值最大化为目标函数寻找最优超参数，使用网格搜索、灰狼优化和哈里斯鹰优化三种优化算法调整参数，在验证集中寻找到最优超参数值后代入最优参数至模型，进而在测试集中得到基于优化的各组合模型的脓毒血症死亡风险预测结果，之后将各模型的预测性能进行对比。最后使用logistic回归作为次级学习器以防止过拟合，从而提高预测精度。

（3）死亡风险预测模型的效果评价

本文将原始数据集按7∶3比例划分为训练集和测试集，其中测试集全程保持独立以验证模型泛化能力。在训练阶段，进一步将训练集以7∶3比例动态划分为训练子集和验证集，其中特征选择和数据平衡操作仅作用于训练子集，验证集用于模型调参和内部验证。为消除随机划分的影响，训练集划分过程重复执行100次并取测试集指标的平均值作为最终性能评价。模型评估采用多维度指标体系，包括准确率、精确率、召回率、F_1分数及AUC值。为深入考察模型鲁棒性，额外进行敏感性分析，在总体数据划分框架下，对训练集实施100次独立随机子划分，每次重新训练模型并记录测试集表现，最终通过各指标的平均值与标准差量化模型对数据划分波动的敏感性。

3. 统计方法

数据收集整理使用R 4.4.1与IBM SPSS 26软件实现随机森林模型的建立、缺失值填补，Python 3.8进行不平衡处理以及分类模型的建立。本文算法使用Python编程语言基于sklearn机器学习工具包实现。

结　果

1. 数据预处理

（1）描述性统计

在本研究中，我们对5163例受试者进行了描述性统计分析，其中包括3571例存活者和1592例死亡者。经数据缺失值填补后患者的基本信息及变量统计描述如表1所示。年龄中位数为68.25岁，存活组年龄中位数为66.96岁，死亡组为71.03岁。性别分布较为均衡，男性占55.47%，女性占44.53%。大多数受试者没有吸烟史（89.08%），糖尿病和肺炎的患病率分别为36.69%和32.75%，高血压的患病率为35.51%。

（2）随机森林特征筛选

本研究使用随机森林对候选变量进行特征重要性排序，如图1所示，根据Gini系数从大到小选择前20个变量。

表 1 基本信息及变量统计描述

变量	总体(n=5163)	存活组(n=3571)	死亡组(n=1592)	Z/χ^2	P
年龄[岁, $M(Q_1,Q_3)$]	68.25(56.03,80.09)	66.96(54.91,79.10)	71.03(59.03,82.00)	−6.963	<0.001
保险类型[n(%)]				16.295	0.003
政府	99(1.92)	74(2.07)	25(1.57)		
医疗补助	437(8.46)	323(9.03)	114(7.16)		
医疗保险	3341(64.66)	2263(63.30)	1078(67.71)		
私人	1261(24.40)	900(25.17)	361(22.68)		
自费	29(0.56)	15(0.42)	14(0.88)		
婚姻状态[n(%)]				1.840	0.175
无配偶	2825(54.67)	1977(55.30)	848(53.27)		
有配偶	2342(45.33)	1598(44.70)	744(46.73)		
种族[n(%)]				15.502	0.001
亚裔	171(3.31)	121(3.38)	50(3.14)		
黑人	534(10.33)	394(11.02)	140(8.79)		
白人	3798(73.50)	2638(73.79)	1160(72.86)		
其他	664(12.85)	422(11.80)	242(15.20)		
性别[n(%)]				1.937	0.164
男性	2866(55.47)	1960(54.83)	906(56.91)		
女性	2301(44.53)	1615(45.17)	686(43.09)		
吸烟史[n(%)]				0.072	0.789
无	4603(89.08)	3182(89.01)	1421(89.26)		
有	564(10.92)	393(10.99)	171(10.74)		
糖尿病[n(%)]				2.099	0.147
无	3271(63.31)	2240(62.66)	1031(64.76)		
有	1896(36.69)	1335(37.34)	561(35.24)		
肺炎[n(%)]				18.882	<0.001
无	3475(67.25)	2472(69.15)	1003(63.00)		
有	1692(32.75)	1103(30.85)	589(37.00)		
高血压[n(%)]				4.962	0.026
无	3332(64.49)	2270(63.50)	1062(66.71)		
有	1835(35.51)	1305(36.50)	530(33.29)		
高脂血症[n(%)]				31.334	<0.001
无	4379(84.75)	2963(82.88)	1416(88.94)		
有	788(15.25)	612(17.12)	176(11.06)		
慢性阻塞性肺疾病[n(%)]				0.009	0.924
无	5042(97.58)	3489(97.59)	1553(97.55)		
有	125(2.42)	86(2.41)	39(2.45)		
冠状动脉粥样硬化性心脏病[n(%)]				0.400	0.527
无	4255(82.35)	2952(82.57)	1303(81.85)		
有	912(17.65)	623(17.43)	289(18.15)		
慢性肾脏病[n(%)]				6.529	0.011
无	3931(76.08)	2756(77.09)	1175(73.81)		
有	1236(23.92)	819(22.91)	417(26.19)		
慢性心力衰竭[n(%)]				12.841	<0.001
无	3428(66.34)	2428(67.92)	1000(62.81)		
有	1739(33.66)	1147(32.08)	592(37.19)		
周围血管病[n(%)]				1.155	0.283
无	4987(96.52)	3457(96.70)	1530(96.11)		
有	180(3.48)	118(3.30)	62(3.89)		
SAPSⅡ[分, $M(Q_1,Q_3)$]	43.00(34.00,55.00)	40.00(31.00,49.00)	54.00(44.00,66.00)	−29.329	<0.001
APSⅢ[分, $M(Q_1,Q_3)$]	57.00(43.00,75.00)	52.00(40.00,65.00)	73.00(57.00,93.00)	−28.072	<0.001
OASIS[分, $M(Q_1,Q_3)$]	36.00(29.50,43.00)	34.00(28.00,41.00)	42.00(35.00,48.00)	−23.034	<0.001
SOFA[分, $M(Q_1,Q_3)$]	6.00(4.00,9.00)	5.00(3.00,8.00)	9.00(6.00,12.00)	−24.930	<0.001
住院时间[d,$M(Q_1,Q_3)$]	9.83(5.54,18.88)	10.17(6.04,19.31)	8.42(3.04,17.73)	−9.748	<0.001
白细胞计数[$\times10^9$/L,$M(Q_1,Q_3)$]	12.30(7.90,18.10)	12.40(8.00,18.10)	12.10(7.47,18.10)	−1.712	0.087
丙氨酸氨基转移酶[U/L,$M(Q_1,Q_3)$]	29.00(17.00,59.00)	28.00(16.00,54.40)	32.33(19.00,71.25)	−5.579	<0.001
天门冬氨酸氨基转移酶[U/L,$M(Q_1,Q_3)$]	41.00(24.00,81.00)	37.00(23.00,70.00)	49.00(28.00,119.25)	−10.502	<0.001
红细胞计数[$\times10^{12}$/L,$M(Q_1,Q_3)$]	3.68(3.19,4.22)	3.74(3.25,4.27)	3.54(3.09,4.08)	−7.575	<0.001
中性粒细胞比例[%,$M(Q_1,Q_3)$]	82.10(73.00,88.90)	82.50(73.25,89.00)	81.80(70.00,88.32)	−3.265	0.001
淋巴细胞比例[%,$M(Q_1,Q_3)$]	7.80(4.00,13.10)	7.80(4.05,13.00)	7.90(4.00,13.42)	−0.371	0.710
血小板计数[$\times10^9$/L,$M(Q_1,Q_3)$]	221.00(140.00,315.00)	230.00(154.00,320.00)	200.00(110.00,307.00)	−7.594	<0.001
血红蛋白[g/dL,$M(Q_1,Q_3)$]	11.00(9.60,12.60)	11.20(9.70,12.70)	10.70(9.50,12.20)	−5.855	<0.001
钠[mmol/L,$M(Q_1,Q_3)$]	138.00(134.00,141.00)	138.00(134.00,141.00)	137.00(133.00,141.00)	−1.948	0.051
钾[mmol/L,$M(Q_1,Q_3)$]	4.20(3.80,4.80)	4.20(3.70,4.70)	4.35(3.80,5.00)	−7.752	<0.001
碳酸氢盐[mmol/L,$M(Q_1,Q_3)$]	23.00(20.00,26.00)	23.00(20.00,26.00)	22.00(18.00,26.00)	−5.447	<0.001
血肌酐[mg/dL,$M(Q_1,Q_3)$]	1.40(0.90,2.40)	1.30(0.90,2.20)	1.60(1.00,2.80)	−7.390	<0.001
血清尿素氮[mg/dL,$M(Q_1,Q_3)$]	29.00(18.00,48.00)	27.00(17.00,43.00)	36.00(22.00,57.00)	−12.448	<0.001
血糖[mg/dL,$M(Q_1,Q_3)$]	128.00(103.00,171.00)	128.00(104.00,172.00)	127.00(99.00,166.00)	−2.785	0.005

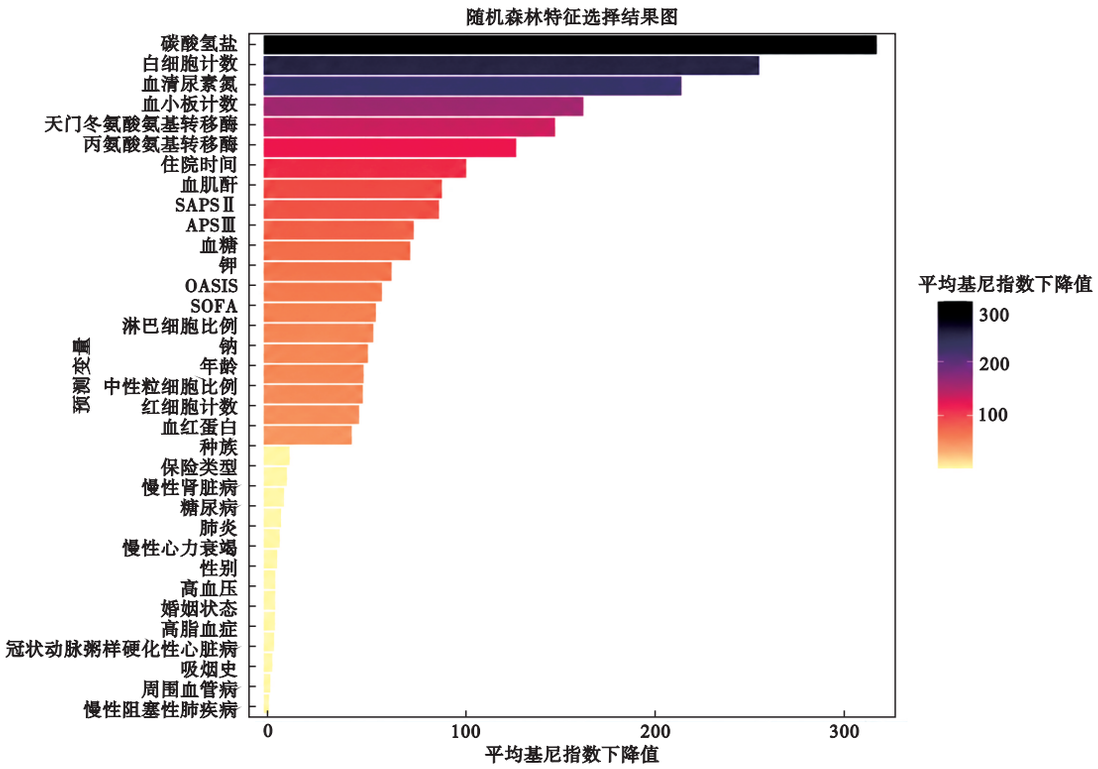


图1　随机森林特征选择结果图

对随机森林特征筛选后得到的20个变量进行相关性分析，判断变量间共线性的问题，相关系数热力图颜色越深说明相关性越大。对于随机森林特征筛选后的丙氨酸氨基转移酶与天门冬氨酸氨基转移酶、中性粒细胞与红细胞与血清尿素氮与血肌酐等相关系数大于0.5的变量，根据随机森林特征变量重要性排序，保留丙氨酸氨基转移酶、中性粒细胞、血清尿素氮、SAPSⅡ时，剩余变量间相关系数均低于0.35。故随机森林特征选择最终保留了14个变量，如图2所示。

对随机森林筛选后的14个变量进行多重共线性诊断，计算各变量的VIF值（表2）。结果显示，SAPSⅡ（VIF=1.262）、年龄（VIF=1.104）等变量的VIF值均低于阈值5，表明变量间不存在显著的多重共线性。

2. 死亡风险预测模型的构建结果

由于患者的死亡风险是我们的研究重点，因此死亡结局被归类为正样本，非死亡患者被归类为负样本，其中正样本为少数类，负样本为多数类。3种不平衡算法处理后的测试集样本数见表3。

在未进行不平衡处理的情况下，验证集和训练集各模型性能比较的实验结果中可以得到，测试集中lightGBM的准确率为0.8689，实现了约87%的分类正确率，其召回率为0.7573，即在脓毒血症死亡风险中大约有76%的样本被预测正确，其AUC值为0.8380，高于XGBoost和CatBoost，模型的整体性能较好，其在验证集中模型整体表现也最优，如表4所示。

不平衡算法处理后，可以看出SMOTETomek组合算法相比Tomek欠采样算法和SMOTE过采样算法在任何一个模型整体上都表现出了最优的性能。此外，SMOTETomek-lightGBM模型相对其他模型在所有指标上均有一定的提升。综上，相对于传统模型，SMOTETomek-lightGBM组合模型性能最优，准确率为0.8748、精确度为0.8021、AUC为0.8509、F_1分数为0.8748，具有较高的精度以及很好的泛化能力，可以帮助更好地预测脓毒血症患者的死亡风险（表4）。

3. 死亡风险预测模型的优化结果

为了进一步优化上述组合模型的性能，本文使用3种优化算法分别对SMOTETomek处理的模型进行优化，如表5。可知，经3种优化算法调参后的模型整体性能都得到了提升，相对于不同的机器学习分类模型，各个优化方法的调参效果也不相同。其中，SMOTETomek-lightGBM组合模型在经过HHO优化调参后得到了最优的预测性能，且经过HHO优化调参后的SMOTETomek-lightGBM组合模型相对于其他组合模型整体性能表现最优。

4. stacking异质集成模型融合结果

为了进一步提升模型的整体效果，我们将经过优化算法调参后的SMOTETomek-XGBoost+GWO、SMOTETomek-CatBoost+GS以及SMOTETomek-lightGBM+HHO三个模型进行融合。为简化模型并提高预测的准确性，我们选择较为简单的logistic回归作为元学习器。通过stacking模型的融合，模型在各项评价指标上均表现出了显著的提升。如表6所示，融合后的模型AUC为0.9380、准确率为0.9516、精确度为

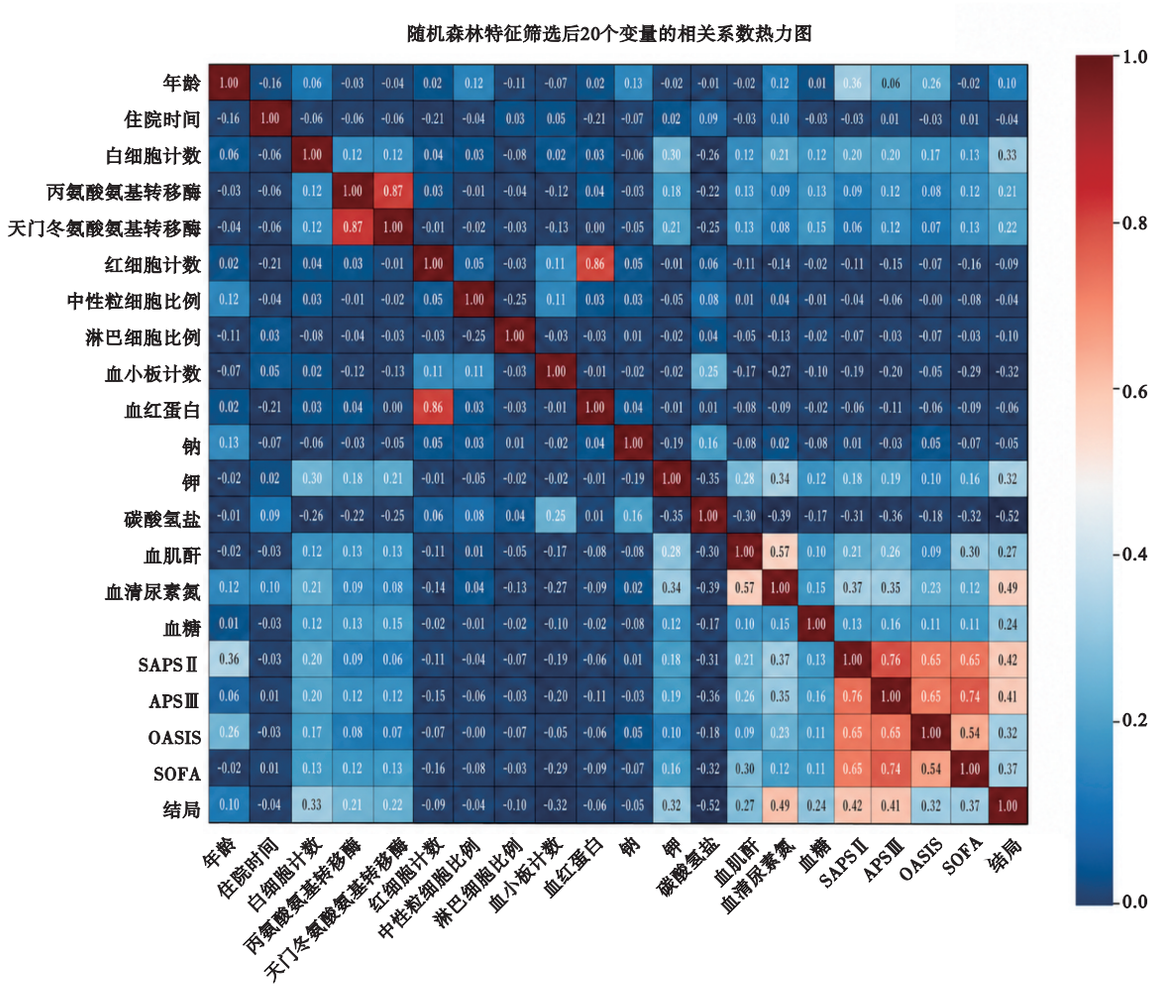


图 2 随机森林特征筛选后 20 个变量的相关系数热力图

表 2 最终变量的 VIF 值分析

变量	VIF
碳酸氢盐	1. 155
白细胞计数	1. 124
血清尿素氮	1. 406
血小板计数	1. 138
住院时间	1. 035
SAPS Ⅱ	1. 262
血糖	1. 042
钾	1. 262
丙氨酸氨基转移酶	1. 013
淋巴细胞	2. 060
钠	1. 101
年龄	1. 104
中性粒细胞	2. 052
血红蛋白	1. 060

表 3 测试集中 3 种不平衡算法处理后的样本数

方法	训练集	Tomek	SMOTE	SMOTETomek
正样本	1133	1133	2481	2348
负样本	2481	1133	2481	2348
总计	3614	2266	4962	4696

0. 9382、召回率为 0. 9025、F_1 分数为 0. 9516，均优于单一模型。这表明，通过 stacking 模型的融合，我们能够有效地提升模型的预测性能，为脓毒血症患者的死亡风险预测提供了一个更为准确和可靠的工具。

5. 数据划分稳定性分析

对 stacking 模型的稳定性分析显示（表 7），100 次随机划分后测试集的各性能指标的标准差均低于 0. 01，表明模型对数据划分的随机性不敏感，具有较高的稳定性。

讨 论

本研究基于 MIMIC−Ⅲ数据库，采用 stacking 异质集成融合算法构建脓毒血症患者死亡风险预测模型，旨在为临床提供早期识别和治疗依据。通过随机森林特征选择确定 SAPS Ⅱ、年龄等 14 个关键变量，结合 SMOTETomek 处理数据不平衡问题，以及对 XGBoost、CatBoost 和 lightGBM 模型的优化，最终构建的 stacking 模型在预测准确性和训练效率上均有显著优势。在模型优化方面，本研究采用了网格搜索、GWO 和 HHO 三种算法对模型进行超参数优化。结果显示，SMOTETomek−lightGBM 组合模型在经过 HHO 优化调参后得到了最优的预测性能，这可能归因于 HHO 算法在搜索空间中具有更好的全局搜索能力和

表4　验证集和测试集各模型性能比较

分类器	准确率	精确率	召回率	AUC	F_1分数
验证集					
XGBoost	0. 8756	0. 8116	0. 7853	0. 8510	0. 8756
Tomek+XGBoost	0. 8691	0. 8680	0. 8706	0. 8691	0. 8691
SMOTE+XGBoost	0. 8912	0. 8859	0. 8978	0. 8912	0. 8912
SMOTETomek+XGBoost	0. 9086	0. 8983	0. 9214	0. 9086	0. 9086
CatBoost	0. 8829	0. 8492	0. 7618	0. 8500	0. 8829
Tomek+CatBoost	0. 8676	0. 8743	0. 8588	0. 8676	0. 8676
SMOTE+CatBoost	0. 8946	0. 9015	0. 8858	0. 8946	0. 8946
SMOTETomek+CatBoost	0. 9079	0. 8982	0. 9200	0. 9079	0. 9079
lightGBM	0. 8861	0. 8544	0. 7676	0. 8539	0. 8861
Tomek+lightGBM	0. 8921	0. 8866	0. 8996	0. 8921	0. 8921
SMOTE+lightGBM	0. 8991	0. 8999	0. 8982	0. 8991	0. 8991
SMOTETomek+lightGBM	0. 9119	0. 9111	0. 9129	0. 9119	0. 9119
测试集					
XGBoost	0. 8657	0. 8027	0. 7490	0. 8334	0. 8657
Tomek+XGBoost	0. 8677	0. 7987	0. 7636	0. 8388	0. 8677
SMOTE+XGBoost	0. 8631	0. 7703	0. 7929	0. 8437	0. 8631
SMOTETomek+XGBoost	0. 8709	0. 7837	0. 8033	0. 8522	0. 8709
CatBoost	0. 8528	0. 7828	0. 7238	0. 8171	0. 8528
Tomek+CatBoost	0. 8664	0. 7965	0. 7615	0. 8373	0. 8664
SMOTE+CatBoost	0. 8631	0. 7917	0. 7552	0. 8333	0. 8631
SMOTETomek+CatBoost	0. 8683	0. 7878	0. 7845	0. 8451	0. 8683
lightGBM	0. 8689	0. 8062	0. 7573	0. 8380	0. 8689
Tomek+lightGBM	0. 8703	0. 7974	0. 7774	0. 8446	0. 8703
SMOTE+lightGBM	0. 8696	0. 7851	0. 7950	0. 8489	0. 8696
SMOTETomek+lightGBM	0. 8748	0. 8021	0. 7887	0. 8509	0. 8748

表5　优化后模型性能比较

优化方法	准确率	精确率	召回率	AUC	F_1分数
网格搜索优化					
SMOTETomek−XGBoost+GS	0. 9419	0. 9199	0. 8891	0. 9273	0. 9419
SMOTETomek−CatBoost+GS	0. 9438	0. 9204	0. 8954	0. 9304	0. 9438
SMOTETomek−lightGBM+GS	0. 9303	0. 9445	0. 8933	0. 9303	0. 9445
GWO优化					
SMOTETomek−XGBoost+GWO	0. 9477	0. 9325	0. 8954	0. 9332	0. 9477
SMOTETomek−CatBoost+GWO	0. 9019	0. 8340	0. 8515	0. 8879	0. 9019
SMOTETomek−lightGBM+GWO	0. 9451	0. 9281	0. 8912	0. 9302	0. 9451
HHO优化					
SMOTETomek−XGBoost+HHO	0. 9451	0. 9226	0. 8975	0. 9319	0. 9451
SMOTETomek−CatBoost+HHO	0. 9406	0. 9232	0. 8808	0. 9240	0. 9406
SMOTETomek−lightGBM+HHO	0. 9490	0. 9309	0. 9017	0. 9359	0. 9490

表6　基分类器与stacking模型融合对比

优化方法	准确率	精确率	召回率	AUC	F_1分数
SMOTETomek−XGBoost+GWO	0. 9477	0. 9325	0. 8954	0. 9332	0. 9477
SMOTETomek−CatBoost+GS	0. 9438	0. 9204	0. 8954	0. 9304	0. 9438
SMOTETomek−lightGBM+HHO	0. 9490	0. 9309	0. 9017	0. 9359	0. 9490
stacking	0. 9516	0. 9382	0. 9025	0. 9380	0. 9516

表7　模型性能稳定性分析（n=100次划分）

指标	均值	标准差	最大值	最小值
准确率	0. 9516	0. 0032	0. 9584	0. 9448
精确率	0. 9382	0. 0080	0. 9613	0. 9151
召回率	0. 9025	0. 0091	0. 9302	0. 8750
AUC	0. 9380	0. 0041	0. 9501	0. 9261
F_1分数	0. 9516	0. 0035	0. 9624	0. 9398

局部搜索精度[16]。此外，SMOTETomek模型在处理数据不平衡问题时表现突出，这为后续研究提供了重要的参考价值[17]。stacking模型通过融合多个初级学习器的预测结果，利用元学习器进一步提升模型性能。本研究中，logistic回归作为元分类器，简化了模型结构，同时保持了较高的预测准确性[18]。stacking模型

的融合策略不仅提高了模型的预测性能，还增强了模型的泛化能力，这对于临床实践中早期识别高风险患者具有重要意义。

在本研究中，我们探讨了脓毒症患者的死亡风险预测问题。脓毒症是全球范围内高死亡率的疾病之一，其发病率和死亡率在过去几十年中并未显著下降[19-20]。因此，早期、准确地预测脓毒血症患者的死亡风险对于临床决策和降低患者死亡风险至关重要。本研究构建的stacking模型为ICU医师早期识别高风险脓毒血症患者提供了科学依据，有助于指导临床治疗和干预，降低患者死亡率。未来研究可进一步探索模型在不同人群、不同环境下的适用性，以及模型在实时数据流中的预测能力。此外，结合更多的临床数据和生物标志物，可能会进一步提高模型的预测准确性和临床应用价值。

尽管本研究取得了一定的成果，但仍存在一些局限性。首先，研究数据来源于单一中心，可能存在选择偏差。其次，模型的外部验证尚不充分，需要在更多样本和不同人群中进行验证。未来研究可以考虑多中心、多人群的数据集，以增强模型的泛化能力和临床适用性。同时，探索更多先进的机器学习算法和优化技术，以进一步提升模型性能。

综上所述，本研究构建的stacking模型在脓毒血症患者死亡风险预测方面具有较高的准确性和效率，为临床决策提供了有力支持。未来的研究应致力于模型的进一步优化和临床应用，以实现更精准的个体化医疗。

参 考 文 献

[1] Rhodes A, Evans L, Alhazzani W, et al. Surviving Sepsis Campaign: International Guidelines for Management of Sepsis and Septic Shock: 2016 [J]. Intensive Care Medicine, 2017, 43(3): 304-377.

[2] Seymour CW, Liu VX, Iwashyna TJ, et al. Assessment of Clinical Criteria for Sepsis: For the Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3) [J]. JAMA, 2016, 315 (8): 762-774.

[3] Kooistra EJ, van der Voort PHJ, van der Heijden GJMG, et al. Critical Care Predictive Models: A Systematic Review and Analysis [J]. Critical Care Medicine, 2018, 46(11): 1768-1776.

[4] Silver D, Huang A, Maddison CJ, et al. Mastering the game of Go without human knowledge [J]. Nature, 2017, 550(7676): 354-359.

[5] 张平,吴念悦,张浩天,等. 心外科手术患者重症监护室住院时间预测模型研究[J]. 电子科技大学学报,2022,51(4): 500-505.

[6] 王羽,刘志玄,邱洪斌,等. 心力衰竭重症患者院内死亡率的预测模型：基于MIMIC-Ⅲ数据库的回顾性研究[J]. 中国动脉硬化杂志,2023,31(3):245-252.

[7] Seymour CW, Rosengard AM, Angoff GH, et al. The Role of the Microbiome in Sepsis [J]. Critical Care Medicine, 2018, 46(12): 2023-2030.

[8] Johnson AEW, Ghassemi M, Niehaus K, et al. MIMIC-Ⅳ: A Large-Scale Database for Critical Care Research [J]. Scientific Data, 2020, 7: 317.

[9] Rios-Avila L, Zhu J, Paul J, et al. Multiple Imputation by Chained Equations in R: The mice Package [J]. Journal of Statistical Software, 2019, 91(5): 1-16.

[10] Geurts P, Ernst D, Wehenkel L. Extremely Randomized Trees [J]. Machine Learning, 2006, 63(1): 3-42.

[11] Hansen LK, Salamon P. Neural network ensembles [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1990, 12 (10): 993-1001.

[12] He H, Bai Y, Garcia EA, et al. ADASYN: Adaptive Synthetic Sampling Approach for Imbalanced Learning [J]. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 2008, 39(2): 539-549.

[13] Kohavi R. A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection [C]. Reprinted from the Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI), 1995.

[14] Chen T, He T, Guestrin C. XGBoost: A Scalable Tree Boosting System [C]. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016.

[15] Ke G, Meng Q, Finley T, et al. LightGBM: A Highly Efficient Gradient Boosting Decision Tree [C]. In Advances in Neural Information Processing Systems, 2017.

[16] Heidari AA, Mirjalili S, Faris H, et al. Harris hawks optimization: Algorithm and applications [J]. Future Generation Computer Systems, 2019, 97: 849-872.

[17] Batista GE, Prati RC, Monard MC, et al. A study of the behavior of several methods for balancing machine learning training data [J]. ACM Sigkdd Explorations Newsletter, 2004, 6(1): 20-29.

[18] Kuncheva LI, Whitaker CJ. Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy [J]. Machine Learning, 2003, 51(2): 181-207.

[19] 张海红,王斌,吴晓雅. 基于多因素logistic回归及列线图模型预测脓毒症休克患者发生脑卒中风险[J]. 中国卫生统计,2022,39(4): 604-606.

[20] Singer M, Deutschman CS, Seymour CW, et al. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3) [J]. JAMA, 2016, 315(8): 801-810.

（责任编辑：张 悦）