动态预测脓毒症患者诱发PICS的机器学习方法

白栋栋 成昊南 黄海骅 霍松泽

钱夕元 教授

**摘要：**PICS会导致脓毒症患者长期死亡率增加，因而适当合理地预测患者诱发PICS的几率尤为重要。本研究通过开发和验证机器学习模型，以动态预测重症脓毒症患者的PICS风险为目的。基于eICU协作研究数据库（eICU-CRD），开发并验证机器学习模型。通过数据分析在最终提取的条数据中，有条数据为正例，即的数据所对应的患者诱发了PICS。我们训练了完整模型和紧凑模型，能够动态预测脓毒症患者诱发PICS的风险，优于传统的logistic回归并且其中紧凑模型明显在临床上更加可行、更加易用。为了方便使用，随即开发了H5网页用于便捷预测患者发生PICS的概率，是本研究的创新之处。

**关键词：**脓毒症；PICS；eICU；机器学习；CatBoost模型

# **项目背景**

脓毒症是宿主应对感染出现炎症反应失调引起的器官功能损伤，表现为危及生命的一组临床综合征。脓毒症患者的病死率较高，是感染引起患者死亡的主要原因[1]。而脓毒症将会诱发PICS（持续炎症-免疫抑制-分解代谢综合征），这是部分患者在度过脓毒症急性期乃至出院后，仍处于慢性持续炎症、免疫功能抑制以及高分解代谢状态的症状。PICS早期即具有明显的炎症反应和免疫抑制反应，后来转化为持续性炎症和免疫抑制状态，出现进行性的器官损伤以及持续的肌肉质量丧失、伤口愈合不良和呼吸机依赖等，且还会导致器官功能障碍，生活质量下降，可能需要长期的医疗监护[2]。所以，若可以帮助临床上及时、准确地预测可能发生的PICS状况，并及时采取相应治疗手段，可减轻病患痛苦，增加患者痊愈的概率。

PICS概念自2012年提出以后,其诊断标准经过2015年及2017年的改进，目前包括以下4个方面：

1. ICU住院天数 天 ；
2. 持续的炎症反应：C反应蛋白 mg/L，视黄醇结合蛋白 mg/L ；
3. 免疫抑制：淋巴细胞计数 /L ；
4. 分解代谢：血清白蛋白 g/L，肌酐/身高指数 ，住院期间体重下降 或身体质量指数 kg/m2 。

尽管这些临床标志物不是对炎症、免疫抑制或分解代谢的直接测量，但其方便获取，可以作为重症监护室医务人员评估患者临床状况的依据[2]。除了以降钙素原、C反应蛋白和白介素6等传统生物标志物作为疾病评估和预后的指标外，近年关于microRNAs作为生物标志物用于组织损伤、炎症等多种疾病的研究也成为热点。microRNA参与了脓毒症后炎症反应及内皮细胞功能调节，可用于判断脓毒症后所处阶段，进而有利于评估PICS 患者的预后。目前已被证实，可用于评估预后的microRNAs包括miR-146a、miR-125b及miRNA-155等[3]。同时，亦有研究采用APACHE-Ⅱ评分、住ICU时间、CD4+/CD8+比值、机械通气4项独立危险因素构建的预测模型，预测脓毒症患者PICS的发生情况[4]。

# **材料和方法**

## 数据来源

我们主要基于一个大型重症监护数据库进行了这项研究，这个数据库是eICU协作研究数据库（eICU-Collaborative Research Database, eICU-CRD)。该数据库由美国麻省理工学院和飞利浦医疗保健公司联合提供，是包含了多家医院的个监护室中，多万条患者医学信息的数据集[5]。而它的2.0版本则涵盖了2014年至2015年全美家医院入住重症监护病房的名患者的常规数据，共条住院记录[6]。两位作者访问了这个数据库，并负责数据提取处理，为后续开展研究做准备。

## PICS判断标准

我们按照医学上判断PICS发生的标准，以患者每天的身体状态参数判断并标记是否有PICS发生。具体而言，判断标准有：

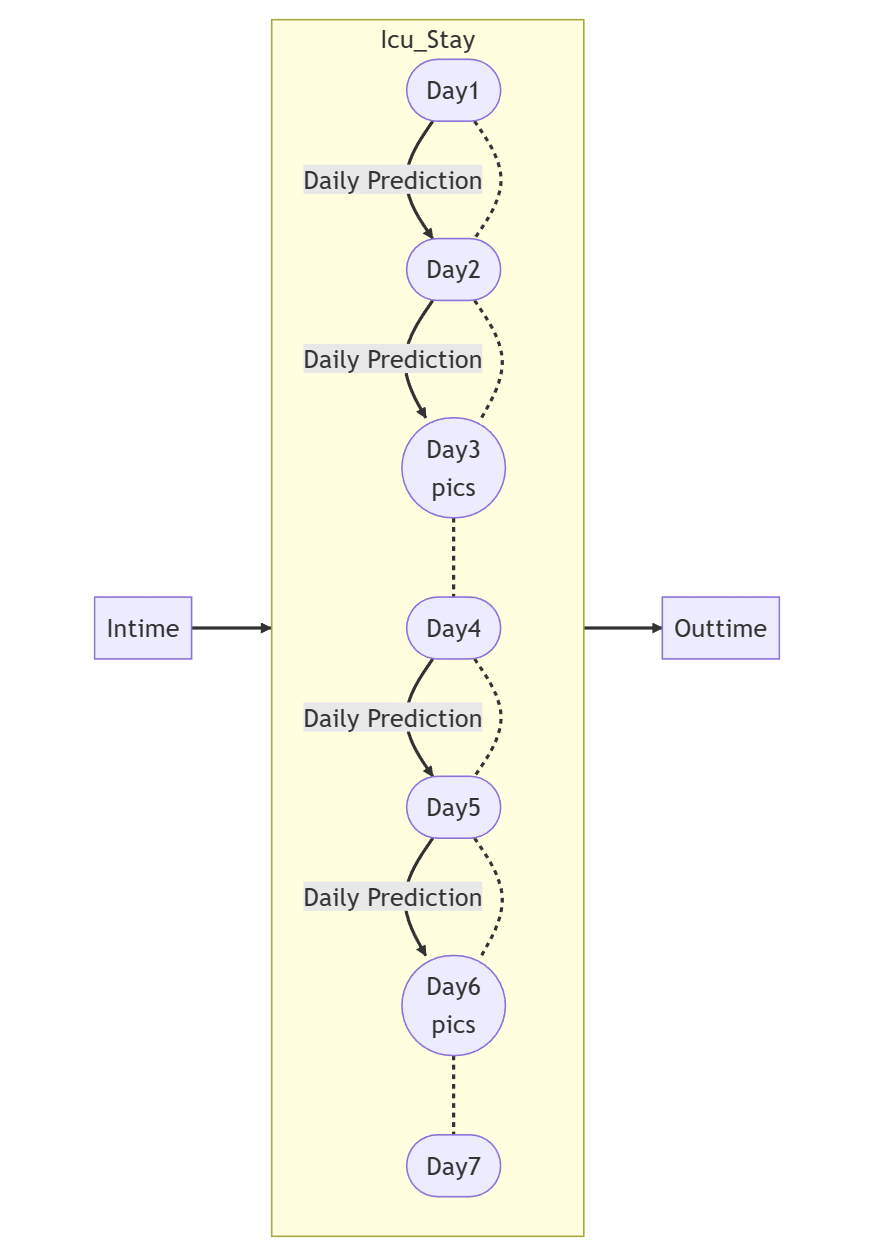
1. ICU住院天数 天 ；
2. 持续的炎症反应：C反应蛋白 mg/L，视黄醇结合蛋白 mg/L ；
3. 免疫抑制：淋巴细胞计数 /L ；
4. 分解代谢：血清白蛋白 g/L，肌酐/身高指数 ，住院期间体重下降 或身体质量指数 kg/m2 。

## 选择变量

统计病人在脓毒症住院期间的临床和测试变量，对于一些具有多个测量值的变量，估计其平均值。为了预测PICS，我们收集了个变量，包括患者的特征（年龄、性别、种族和住院类型）、生命体征（呼吸频率、血压、心率、SpO2和体温)、测试数据（血气、常规血液分析、肝功能、肾功能和凝血特征）、输血量（红细胞、血小板和新鲜冷冻血浆）和尿量。

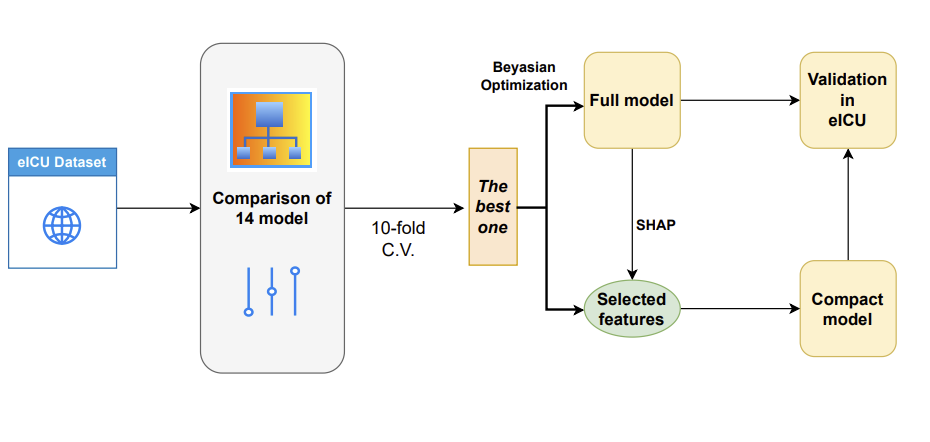
## 研究方法

如图 1所示，我们的模型在患者被诊断为脓毒症的每一天，基于上述个变量生成一个连续的预测评分。评分评估第二天发生PICS的风险。当日已达到PICS标准的病人不进行预测，当患者从PICS状态恢复后，如果患者仍然患有脓毒症，我们的模型将重新开始预测他们第二天发生PICS的风险。



**图 1 动态预测PICS流程示意图**

如图 2所示，我们初步比较了种开源的模型训练后的预测表现，训练过程采用十折交叉验证。在每折验证集上计算准确度和受试者工作特征曲线下面积（AUC），并将其汇总以评估每种模型，选择精度最高、AUC值最大的算法。随后利用贝叶斯优化算法对模型进行细粒度超参数调整，该算法是一种高效的约束全局优化工具，使用bayes\_opt开源工具包的函数实现。优化后的模型为本研究预测PICS的最佳模型，将其作为我们的完整模型。



**图 2 模型训练流程图**

变量对PICS预测的影响是使用Python的SHAP库来测量的，该开源包使用基于验证集的博弈论方法评估每个特征的重要性。我们选择了个非常重要且在临床环境中尽可能容易收集的变量，然后基于所选变量训练一个紧凑的模型用于PICS预测，虽然该模型不如完整模型准确，但在临床中可能更实用。

# **模型结果**

## 基准特征

从eICU数据库中提取出条数据，包含名不同的脓毒症患者。其中，）条数据为正例，（）条数据为反例。

如表 3所示，正例数据拥有更长的ICU入住天数、更少的血浆蛋白、更少的淋巴细胞数目、更高的心率、更高的呼吸频率、更少的血清总蛋白、更低的红细胞比容、更少的肌酸酐、更高的白细胞计数、更多的血小板、更低的平均动脉压。

**表 3 基准特征比较**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 指标名称 | 正例平均值 | 反例平均值 |  | 单位 |
| ICU入住天数 |  |  |  | 天 |
| 血浆蛋白 |  |  |  | g/dL |
| 淋巴细胞数目 |  |  |  | % |
| 心率 |  |  |  | 次/分钟 |
| 呼吸频率 |  |  |  | 次/分钟 |
| 血清总蛋白 |  |  |  | g/dL |
| 红细胞比容 |  |  |  | K/mcL |
| 肌酸酐 |  |  |  | mg/dL |
| 白细胞计数 |  |  |  | K/mcL |
| 血小板 |  |  |  | K/mcL |
| 平均动脉压 |  |  |  | mmHg |

## 模型比较

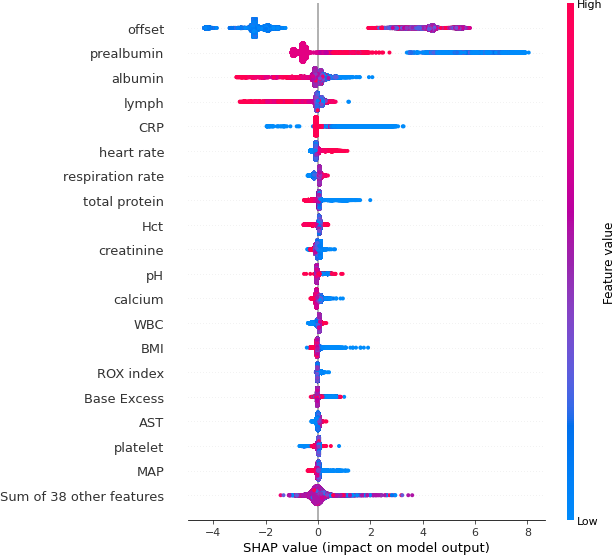
**表 4 模型比较结果（按平均准确率排序）**

|  |  |  |  |
| --- | --- | --- | --- |
| 排名 | 模型名称 | 平均准确率 | 平均AUC1 |
| 1 | CatBoost | 0.996 (±0.001) | 0.996 (±0.001) |
| 2 | Light Gradient Boosting | 0.995 (±0.001) | 0.996 (±0.001) |
| 3 | Extreme Gradient Boosting | 0.995 (±0.001) | 0.994 (±0.002) |
| 4 | Hist Gradient Boosting | 0.994 (±0.002) | 0.996 (±0.002) |
| 5 | Ada Boost | 0.993 (±0.002) | 0.995 (±0.002) |
| 6 | Decision Tree | 0.989 (±0.002) | 0.949 (±0.013) |
| 7 | Multi-Layer Perceptron | 0.982 (±0.004) | 0.975 (±0.008) |
| 8 | SVM (RBF Kernel) | 0.973 (±0.003) | 0.957 (±0.011) |
| 9 | Logistic | 0.966 (±0.007) | 0.956 (±0.012) |
| 10 | Extra Trees | 0.961 (±0.006) | 0.977 (±0.006) |
| 11 | Naive Bayes | 0.961 (±0.006) | 0.689 (±0.034) |
| 12 | Ridge | 0.961 (±0.007) | 0.952 (±0.013) |
| 13 | Linear Discriminant Analysis | 0.961 (±0.010) | 0.952 (±0.013) |
| 14 | K-Nearest Neighbours | 0.951 (±0.006) | 0.544 (±0.025) |

1 AUC：Area Under Curve，接受者操作特性曲线下与坐标轴围成的面积。

用提取出的数据训练预测模型，各种模型的交叉验证结果如表 4所示。其中，Logistic回归表现良好（平均准确率：，平均AUC：），而集成学习方法拥有更高的平均准确率和平均AUC。其中，CatBoost模型的预测结果最好（平均准确率：，平均AUC：），故选择CatBoost模型进入下一阶段。

## 完整模型与紧凑模型



**图 3 完整模型中各变量的平均SHAP值比较**

我们使用eICU中的数据（共位患者、条数据，正例占），选取了个自变量，对第二天的PICS情况进行预测（训练数据的输出为：当天或第二天满足PICS条件）。接着，使用种模型进行了折交验证。经分析比较，选取CatBoost作为完整模型。

根据预测结果比较，选择含个输入变量的CatBoost模型为完整模型。计算完整模型中各变量的平均SHAP值（Shapley Value，即沙普利加和解释值），结果如图 3所示。此摘要图展示了各个变量对预测结果的影响情况分布。例如，ICU入住天数（offset）对结果影响明显，且ICU入住天数越长，发生PICS的概率越大。

随之，根据SHAP值和数据常见程度排序，选择前个作为紧凑模型的自变量：偏移量、白蛋白、淋巴、心率、呼吸率、总蛋白、Hct、肌酐、pH、钙、WBC、BMI、AST、血小板、MAP。使用以上选出的自变量训练紧凑模型，并使用贝叶斯优化优化超参数，得到的最优参数为：

**表 1 紧凑模型最优超参数**

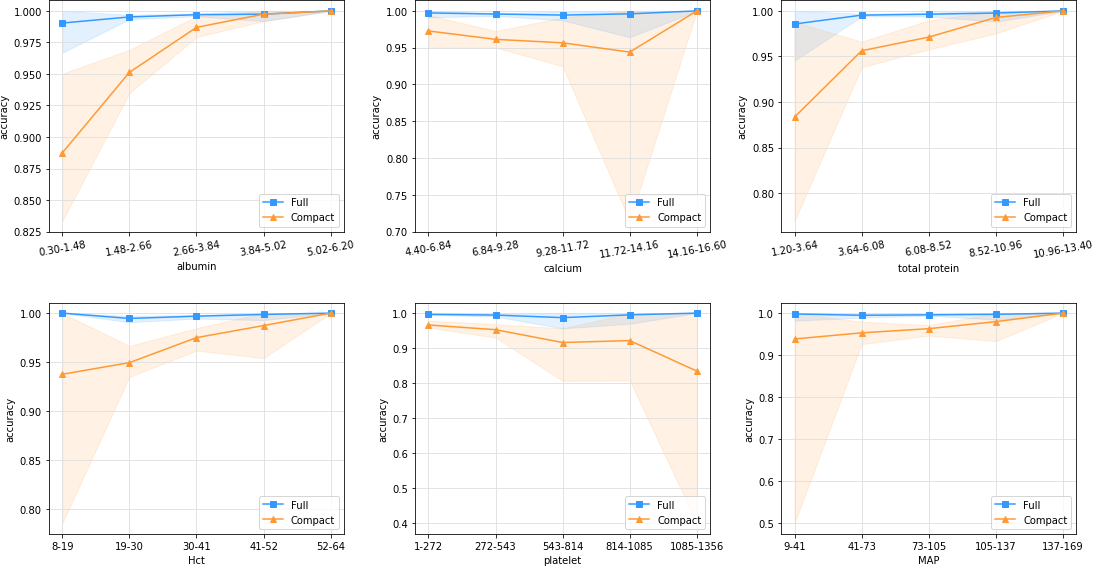
|  |  |  |  |
| --- | --- | --- | --- |
| 参数名称 | 最优值 |  | 参数说明 |
| depth | 6 |  | 决策树深度，正整数 |
| iterations | 350 |  | 迭代次数，正整数 |
| l2\_leaf\_reg | 17.6845 |  | 正则化参数，非负实数 |
| learning\_rate | 0.0282 |  | 学习率，正实数 |

而优化后紧凑模型的表现如下：

**表 2 优化后紧凑模型的得分**

|  |  |  |
| --- | --- | --- |
| 评价指标 | 平均值 | 标准差 |
| Accuracy |  |  |
| AUC |  |  |

从而，我们最终得到优化后的紧凑模型。虽然紧凑模型的预测结果得分略低于完整模型，但是其明显在临床上更加可行、更加易用。

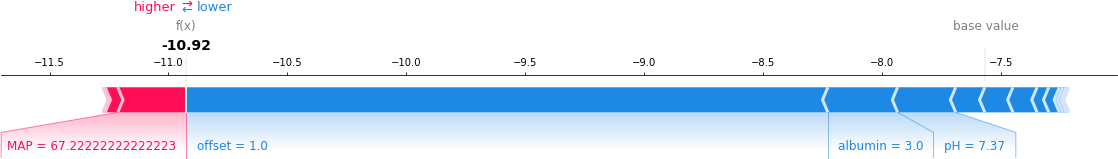


**图 4 模型性能分析**

## 性能分析

如图 4所示，完整模型和紧凑模型在各种指标的不同范围下都表现良好。当某个指标出现明显的异常值时，模型可以非常敏锐地察觉到并给出十分准确的预测结果。

## 模型解释



**图 5 个例(A)中主要变量的SHAP值**



**图 6 个例(B)中主要变量的SHAP值**

图 3从整体上展示了各个变量对于预测结果的影响情况，同时也展现了模型对输入变量变化的灵敏性。而图 5和图 6展示了两个个例中主要变量的SHAP值。图中红色条和蓝色条分别表示危险因素和安全因素，它们共同作用决定了最终的结果。如图 5，在个例(A)中，虽然患者的平均动脉压偏低，但是其ICU入住天数很短、血浆蛋白较多、pH值也良好，所以模型准确预测了患者次日无PICS风险。又如图 6，在个例(B)中，虽然患者的血浆蛋白较多，但是其ICU入住天数较长、身体质量指数（BMI）偏低，所以模型准确预测了患者次日的PICS。

## H5预测工具

为了方便临床上对上述紧凑模型的测试，开发了一款预测脓毒症患者诱发PICS的H5应用。只需在表单中输入指标数值，然后点击“提交”，就可以获得紧凑模型对患者次日发生PICS概率的预测。目前应用已部署在此网址上：<http://1.15.185.22/sepsis-pics-tool/>。

# **结论**

总之，本研究比对了种模型的交叉验证结果，认为CatBoost模型最适合本课题，且基于CatBoost训练了一个完整模型与一个紧凑模型。其中，完整模型预测相对更精确，而紧凑模型在临床上更加简易实用，二者均可以比较准确预测脓毒症患者次日诱发PICS的风险。

# **数据声明**

本研究使用了公开的数据集。这些数据可以在这里找到：https://eicu-crd.mit.edu/。

**参考文献**

1. 王锦权,金朋.脓毒症患者病情及预后评估的临床意义[J].安徽医学,2022,43(08):869-872.
2. 韩雪妹,王日兴.脓毒症相关持续性炎症-免疫抑制-分解代谢综合征的研究进展[J].感染、炎症、修复,2021,22(03):171-174.
3. 李盼,马莉.持续性炎症-免疫抑制-分解代谢综合征发病机制及诊疗的新进展[J].中华重症医学电子杂志(网络版),2020,6(03):318-321.
4. 杨蓉,王鹏,屈文静.脓毒症患者继发持续性炎症-免疫抑制-分解代谢综合征的影响因素及预测模型构建[J].海南医学,2021,32(17):2182-2185.
5. 周倩钰. 基于eICU数据库急性胰腺炎严重程度及预后预测指标的研究[D].华中科技大学,2022.DOI:10.27157/d.cnki.ghzku.2022.005907.
6. Pollard Tom J;;Johnson Alistair E W;;Raffa Jesse D;;Celi Leo A;;Mark Roger G;;Badawi Omar.The eICU Collaborative Research Database, a freely available multi-center database for critical care research.[J].Scientific data,2018(1).