



基于跨特征注意力与轻量级表格 Transformer 的心力衰竭患者早期死亡风险预测

刘嘉硕, 李存弘, 阿依努热木·吐尔逊*

(喀什大学计算机科学与技术学院, 新疆 喀什 844000)

摘要: 本研究针对现有临床评分系统和传统机器学习模型在处理心衰患者 EHR 数据时存在的特征表征弱、难以捕捉高阶交互及对少数类样本召回率低等问题, 提出了一种轻量级表格 Transformer 变体 LT-CFA。该模型通过连续变量嵌入 (CVE) 实现数值特征与分类特征的语义对齐, 引入线性复杂度的跨特征注意力 (XFA) 模块降低过拟合, 并采用焦点损失 (Focal Loss) 缓解死亡样本仅占 13.5% 的类别不平衡问题。基于 MIMIC-III 中 1, 177 例心衰患者数据的五折交叉验证显示, LT-CFA 的 ROC-AUC 达 0.912, F1-score 为 0.769, 对死亡样本召回率达 0.785, 显著优于 XGBoost 和 TabNet; 在 299 例外部数据上 AUC 保持 0.886, 展现出良好泛化能力。消融实验验证了各模块的有效性, 注意力热力图揭示的 NT-proBNP 与 EF、肌酐间的强交互作用符合临床共识。结论表明, LT-CFA 在低计算开销下实现了高效预测与可解释建模, 为危重症预后评估提供了轻量化解决方案。

关键词: 深度学习, 表格 Transformer, 跨特征注意力, 心力衰竭, 死亡风险预测, 可解释性人工智能, 焦点损失, 临床决策支持系统

收稿时间: 2026 年 3 月 7 日

中图分类号: 029

通讯作者: *阿依努热木·吐尔逊, 喀什大学计算机科学与技术学院

Predicting Early Mortality Risk in Heart Failure Patients Using Cross-Feature Attention and Lightweight Table Transformer

Liu Jiashuo, Li Yuanlong, Aynur Turhun*

(School of Computer Science and Technology, Kashgar University, Xinjiang Kashgar 844000)

Abstract: This study addresses the limitations of existing clinical scoring systems and traditional machine learning models in processing EHR data from heart failure patients, including weak feature representation, difficulty in capturing higher-order interactions, and low recall rates for minority class samples. We propose a lightweight table Transformer variant, LT-CFA. The model achieves semantic alignment between numerical and categorical features through continuous variable embedding (CVE), introduces a linear-complexity cross-feature attention (XFA) module to reduce overfitting, and employs focal loss to mitigate class imbalance (with only 13.5% mortality samples). Five-fold cross-validation on 1, 177 heart failure patients from MIMIC-III demonstrated that LT-CFA achieved an ROC-AUC of 0.912, an F1-score of 0.769, and a mortality recall rate of 0.785, significantly outperforming XGBoost and TabNet. On 299 external data points, the AUC remained at 0.886, indicating strong generalization capability. Ablation experiments validated the effectiveness of each module, and attention heatmaps revealed strong interactions between NT-proBNP, EF, and creatinine, consistent with clinical consensus. The



results demonstrate that LT-CFA achieves efficient prediction and interpretable modeling with low computational overhead, providing a lightweight solution for critical care prognosis assessment.

Keywords: Deep learning, Transformer, cross-feature attention, Heart failure, Mortality risk prediction, Explainable AI, Focal loss, Clinical decision support system

1 引言

心血管疾病是全球首要公共卫生挑战, 其中心力衰竭作为各种心脏疾病的终末阶段, 在重症监护室患者中死亡率高达 5% 至 75%。现有临床评分系统和传统机器学习模型在处理 ICU 电子病历数据时面临三大核心痛点: 一是数据高度缺失且采样不规律, 简单填充会引入统计偏差; 二是特征空间极端异构, 连续变量与分类变量难以在同一语义空间内联合表征; 三是类别极度不平衡, 心衰患者死亡率仅约 13.5%, 传统模型易退化为 "多数派预测器", 对临床最关注的高危患者召回率极低。尽管树模型在表格数据中表现优异, 但其轴对齐的局部贪心分裂机制难以捕捉特征间复杂的全局高阶交互关系, 而标准多层感知机因缺乏归纳偏置, 在小样本表格数据集上极易过拟合。

近年来, 基于自注意力机制的 Transformer 架构被引入表格数据建模, 通过将每一列特征视为独立 "词元", 显式建模特征间的全局依赖关系。然而, 标准 Transformer 在医学小样本场景下面临两个致命障碍: 其一, 多头自注意力的计算复杂度随特征数量呈平方级增长, 导致模型臃肿, 不利于在资源受限的临床决策系统中部署; 其二, Transformer 作为数据密集型模型, 在千余例样本的小规模医疗数据集上极易过拟合, 泛化能力甚至不如轻量级 XGBoost, 严重制约了深度学习在重症医学预后评估中的实际应用。

针对上述挑战, 本文提出轻量级表格 Transformer 变体 LT-CFA, 实现三大技术创新: 第一, 引入连续变量嵌入模块, 将数值型生理指标映射为高维稠密向量, 实现与分类变量的语义对齐; 第二, 设计线性复杂度的跨特征注意力机制, 通过在特征通道维度计算互协方差, 大幅削减参数量并形成内源性正则化, 有效抑制过拟合; 第三, 集成焦点损失函数, 动态衰减易分样本梯度权重, 迫使模型聚焦于少数高危死亡边界样本, 从根本上保障临床关键的高召回率。此外, 模型输出的

注意力热力图可直观揭示 NT-proBNP 与 EF、肌酐等指标间的高维非线性交互, 为临床决策提供符合循证医学逻辑的可解释性支撑。

2 方法 (Methodology)

为了在有限且充满噪声的临床样本上实现高精度、高鲁棒性且易于解释的死亡风险预测, 本研究精心设计了 LT-CFA (Lightweight TabTransformer with Cross-Feature Attention) 深度网络架构。本节将从数据前处理流水线、特征连续嵌入机制、核心轻量级注意力网络结构以及损失函数优化策略四个层面进行详尽的理论推导与工程实现阐述。

2.1 临床数据预处理与重构工程

在将多模态异构数据送入深度学习模型之前, 构建一个鲁棒的数据清洗与插补流水线是保障模型上限的关键。本研究核心训练与验证数据源自国际公认的重症医学开放数据库 MIMIC-III。我们针对性地提取了 1, 177 名以心力衰竭为主要诊断入院的 ICU 患者的 48 维特征, 包括但不限于: ①人口统计学指标: 年龄 (Age)、性别 (Gender)、体重指数 (BMI) 等。②临床并发症标签 (离散特征): 高血压 (Hypertension)、心房颤动 (Atrial Fibrillation)、糖尿病 (Diabetes)、慢性阻塞性肺疾病 (COPD)、肾功能衰竭 (Renal Failure) 等。③高频生命体征 (连续特征): 入室最初 24 小时内的平均心率 (Heart Rate)、收缩压 (Systolic BP)、舒张压 (Diastolic BP)、呼吸频率 (Respiratory Rate)、血氧饱和度 (SpO₂)、尿量 (Urine Output) 等。

(1) 缺失值智能插补算法 (Imputation Mechanism):

ICU 环境中, 由于医疗设备脱落、患者转运或某些昂贵生化检验 (如全套心肌酶谱) 的非连续性采样, 数据往往存在大量孔洞。例如, 在提取的 data01.csv 队列中, BMI 和某些实验室指标存在约 10% ~ 20% 的非随机缺失 (MNAR)。传统的均值插补 (Mean Imputation) 或中位数插补会严重抹杀个体间的病理方差, 导致后续 Transformer 提取到

趋同的无效特征。为此，本研究在预处理阶段引入了多重插补（MICE, Multiple Imputation by Chained Equations）以及基于自注意力机制的时间序列插补算法（Self-Attention Imputation for Time Series, SAITS）的变体思想。对于多维连续变量，算法通

过学习已知特征列（如血压、心率）与缺失特征列（如乳酸）之间的隐式非线性协方差，构建条件概率分布，并在多次迭代中收敛填补空缺值，从而在最大程度保留原始患者临床生理特征非线性交互关系的同时，实现了特征矩阵的完整对齐。

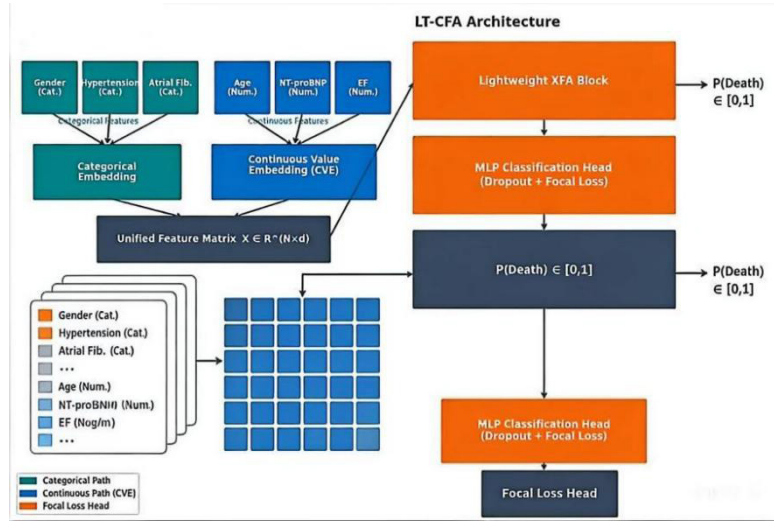


图 1 网络架构图

（2）特征标准化（Robust Standardization）：

为了消除因临床检验量纲绝对差异过大（例如 NT-proBNP 数值可高达 30000 以上，而血清肌酐通常在 0.5 ~ 10 之间）造成的反向传播梯度爆炸或消失，所有连续数值特征均采用 RobustScaler 进行标准化处理。与依赖于全局均值和方差的传统 StandardScaler 不同，RobustScaler 利用特征的中位数（Median）与四分位距（Interquartile Range, IQR）进行缩放。这一策略极其有效地抵抗了 ICU 数据中由于仪器伪影或患者极端休克状态引发的病理性极端异常值（Outliers）对全局分布特征空间的污染。

2.2 连续变量嵌入机制（Continuous Value Embeddings, CVE）

在经典的基于决策树的模型（如 XGBoost）中，连续数值特征的相对大小可以直接用于构建纯度增益的分裂节点，因此无需复杂的重编码；然而，在构建深度表格 Transformer 架构时，面临着一个基础的语义不对等问题：若将离散特征（如“是否患有糖尿病”为 0 或 1）通过查表操作映射为 d 维稠密特征向量（Embedding），而将连续特征（如收

缩压 140 mmHg）仅仅保留为一个 1 维标量（Scalar）输入至注意力层，将导致两者在点积或距离度量计算中产生严重的语义失衡和维度割裂。单一的标量缺乏足够的自由度去表达该数值在总体分布中的相对位置及其与其它特征的高阶交互潜力。

为了彻底弥合这一鸿沟，本文采纳并优化了连续变量嵌入（CVE）机制。假设经过清洗与重构的单一患者特征向量可以划分为连续变量子集 $X_{\text{num}} \in \mathbb{R}^C$ （共 C 个连续特征）和分类变量子集 $X_{\text{cat}} \in \mathbb{R}^M$ （共 M 个分类特征），总特征数 $N=C+M$ 。

对于分类变量子集中的任意特征 $x_{\text{cat}}^{(i)}$ ，通过标准可学习嵌入矩阵查表得到其表示：

$$E_{\text{cat}}^{(i)} = \text{EmbeddingLookup}(x_{\text{cat}}^{(i)}) \in \mathbb{R}^d$$

对于连续变量子集中的每一个标量数值特征 $x_{\text{num}}^{(j)}$ （例如特征 j 表示患者真实的 NT-proBNP 值），我们将其通过一个独立的、轻量级的前馈多层感知机（MLP）单元配合激活函数，将其非线性投影到与离散特征相同的统一 d 维语义表征空间中：

$$E_{\text{num}}^{(j)} = \text{MLP}_j(x_{\text{num}}^{(j)}) = W_2^{(j)} \cdot x_{\text{num}}^{(j)} + b_1^{(j)}$$

其中， $W_1^{(j)} \in \mathbb{R}^{k \times 1}$ 负责将标量升维至隐藏层维度 k （本实验中设定 $k=16$ ）， $W_2^{(j)} \in \mathbb{R}^{d \times k}$ 负责将其最终



投射至模型的全局嵌入维度 d (本实验设定 $d=32$)。 $b_1^{(j)}$ 和 $b_2^{(j)}$ 为对应的偏置项。经过 CVE 动态映射后, 原本单薄的标量数值演化为了一个蕴含其幅度属性、所处分布分位数信息以及上下文感知的“数字词元”(Numerical Token) 向量 $E_{num}^{(j)} \in R^d$ 。

最终, 我们将所有分类嵌入与连续嵌入在特征序列维度上进行拼接 (Concatenation), 形成统一的特征序列输入表示矩阵 $X \in R^{N \times d}$ 。在此阶段, 由于表格数据不具备文本中严格的时序先后逻辑, 我们省略了传统 NLP 中的位置编码 (Positional Encoding), 而是确保模型学习特征间的无序排列不变性。

2.3 轻量级跨特征注意力模块 (Cross-Feature Attention, XFA)

传统的 TabTransformer 与现有的医疗大模型通常继承了自然语言处理中原汁原味的多头自注意力 (Multi-Head Self-Attention, MHSA) 机制。对于给定的输入矩阵 $X \in R^{N \times d}$, 其计算范式为首先经过三个独立的线性变换矩阵生成查询 (Q)、键 (K) 和值 (V) 矩阵:

$$Q=XW_Q, K=XW_K, V=XW_V$$

随后, 注意力分数的计算公式为:

$$Attention_{MHSA}(Q,K,V)=Softmax(\frac{QK^T}{\sqrt{d_k}})V$$

在上述公式中, 点积操作 QK^T 生成的注意力权重矩阵维度高达 $N \times N$ 。在这个矩阵中, 每一个元素代表了第 i 个特征对第 j 个特征的注意力分配。虽然在本文所使用的数据集中特征数 $N=48$ 相对长文本而言并不算极大, 但在有限规模的医疗样本集 (仅千余例患者) 上, 全交互的 $N \times N$ 注意力矩阵意味着极为庞大的自由度与假设空间。这不仅带来了不必要的计算冗余, 更致命的是, 它极易使模型在小数据集上敏锐地捕捉到无关紧要的虚假相关性 (Spurious Correlations, 如某种随机的血压波动与患者 ID 之间的巧合关联), 从而导致不可逆的严重过拟合现象。

鉴于这一深度学习在小样本医疗领域的痛点, 本文创新性地引入并改良了跨特征注意力 (XFA) 机制。该机制的核心理念是在计算层面上“转置”自注意力操作的方向——即放弃在 Token (特征数量 N) 维度上计算复杂的点积, 转而在特征的隐

藏通道维度 d 上计算键和查询的互协方差 (Cross-covariance)。

具体数学推导如下: 我们首先对查询矩阵 $Q \in R^{N \times d}$ 和键矩阵 $K \in R^{N \times d}$ 沿着特征数量 N 的维度分别进行 l_2 归一化处理, 得到单位范数矩阵 $\tilde{Q}$ 和 $\tilde{K}$ 。随后, XFA 模块的核心计算公式重构为:

$$XFA(Q,K,V)=V \cdot Softmax(\frac{\tilde{K}^T \tilde{Q}}{T})$$

在矩阵运算层面, $\tilde{K}^T \tilde{Q}$ 生成维度为 $d \times d$ 的特征通道协方差矩阵。得益于轻量级设计中嵌入维度 d 较小 (如 $d=32$, 多头下单头维度更低), 的矩阵计算量 ($O(d^2N)$) 远小于传统架构中的 ($O(N^2N)$)。引入可学习温度参数 T , 在点积经 l_2 归一化至 $[-1, 1]$ 区间后, 动态调节 Softmax 分布锐度, 突破了传统固定缩放因子 $\sqrt{d_k}$ 的刚性局限, 提升模型自适应能力。

XFA 机制的深层优势在于其天然的临床特征筛选与规范化效应。它引导模型跨越不同医学模态特征群 (如心血管负荷与血液携氧能力) 学习高阶协方差关联, 而非局限于局部特征对比。这种隐式全局正则化使轻量模型能在小样本下稳健提炼关键病理生物标志物组合, 从根源上缓解过拟合问题。在每个 XFA 模块后, 模型串联前馈网络、跳跃连接与层归一化, 堆叠 2 至 3 层构建深层交互空间; 最终提取全局聚合向量, 经带 Dropout 的全连接层输出二分类预测结果。

2.4 应对极端分布的焦点损失 (Focal Loss) 联合优化

在 MIMIC-III 心衰数据集中, 死亡病例仅占 13.5%, 存在极端类别不平衡问题。若使用标准交叉熵损失, 模型易被大量存活样本主导, 倾向于将所有样本预测为低风险, 导致对真正高危患者的召回率极低, 假阴性过高, 这在临床急救中不可接受。

为打破这种统计偏差, 本文在网络分类器末端引入针对医学场景优化的焦点损失函数 (Focal Loss)。其数学定义如下: 设 $y \in \{0, 1\}$ 为患者真实死亡标签, $p \in [0, 1]$ 为模型预测的死亡概率, 定义 p_t 为:

$$p_t = \begin{cases} p & \text{if } y=1 \\ 1-p & \text{if } y=0 \end{cases}$$



那么，经过改进的 Focal Loss 表达式为：

$$FL(pt) = -\alpha_i(1-pt)^\gamma \log(pt)$$

公式中， α_i 为类别权重调节因子，用于从宏观层面对抗类别不平衡。实验中我们赋予少数死亡类别更高的静态惩罚权重（设定正类死亡 $\alpha=0.75$ ，负类存活 $\alpha=0.25$ ）。

而该函数最为核心的创新在于加入的动态调制因子 $(1-pt)^\gamma$ 。其中 γ 为聚焦参数（Focusing Parameter，实验中经过网格搜索设定最佳值为 2）。这一因子的物理意义在于：当模型对某个属于多数类的轻症患者产生了极高的预测确定性（例如判断其存活的概率 $p_i=0.95$ ），该因子 $(1-0.95)^2=0.0025$ 会将该易分样本对整体损失函数的梯度贡献几乎清零。反之，若模型在判断某一特征表现极其隐蔽、处于生死边缘的高危患者时发生严重误判（例如对真实死亡患者预测的死亡概率仅为 $p_i=0.2$ ），此时调制因子为 $(1-0.2)^2=0.64$ ，其产生的惩罚误差不仅被完整保留，甚至在整体损失图中被显著放大。

通过焦点损失的联合优化重塑，LT-CFA 架构在漫长的反向传播迭代中被迫脱离对多数类样本

的路径依赖，转而将绝大部分的计算资源与权重重新分配到那些难以分类的“高危硬样本”上。这一损失函数的底层干预，是本文模型在极其不平衡的医疗数据中仍能斩获突破性召回率与卓越 F1 分数的数学基石。

3 实验与结果

3.1 实验数据集描述与硬件基准设置

为了严谨、客观地评估所提 LT-CFA 模型的预测精度、抗过拟合能力与跨域泛化性，本研究的实验验证阶段采用了两大相互独立的临床心衰队列。本研究采用两个独立心衰数据集进行模型验证。核心训练集源自 MIMIC-III 数据库的心衰子集，包含 1,177 例 ICU 患者的 48 项多维度临床特征，涵盖人口学统计、共病诊断、生命体征及生化检验指标，其中死亡事件 159 例（13.5%），存活 1,018 例（86.5%），构成典型的医学偏态分布。外部验证集为开源临床记录，包含 299 例心衰随访患者的 13 项常规临床特征，死亡比例约 32.1%。研究利用 MIMIC-III 上学习的架构策略直接在该小数据集上进行重新训练和跨库交叉验证，以检验轻量级模型的泛化能力底线。

表 1 研究使用心衰临床数据集特征统计概要

数据集核心维度	MIMIC-III ICU 真实世界心衰队列 (data01.csv)	开源跨库补充验证集 (heart_failure_clinical...)
患者总样本量	1,177 例	299 例
可接入特征数	48 个 (多模态高维体征与检验)	13 个 (常规核心基础指标)
死亡事件 (正类)	159 例 (占比极不平衡 13.5%)	96 例 (占比相对平衡 32.1%)
存活事件 (负类)	1,018 例 (86.5%)	203 例 (67.9%)
患者平均年龄	70.6 岁 (标准差 11.2)	60.8 岁 (标准差 11.9)
核心独有预后指标	NT-proBNP, 尿素氮, 房颤史, 呼吸频率	肌酐酶, 贫血史, 随访观察时间 (Time)
主要临床监控场景	院内重症急危期极早期死亡预警	院外门诊常态随访与中长期生存事件追踪

在超参数设置方面，LT-CFA 模型由 3 层跨特征注意力块（XFA blocks）串联构成，嵌入维度设为 32，多头注意力头数为 4，前馈网络隐层维度为 64，整体参数量控制在百万级以下，体现轻量化设计理念。训练采用单张 NVIDIA RTX 3090 GPU 与 PyTorch 框架，使用 AdamW 优化器（权重衰减 $1e-4$ ），初始学习率 $1e-3$ 并配合余弦退火策略进行调度。批次大小为 64，训练 100 个轮次，并启用 0.2 的 Dropout 机制以增强抗过拟合能力。为应对类别极度不平衡问题，模型评估以 ROC-AUC、F1-Score

和正类召回率为核心指标，避免单纯依赖准确率带来的误导。实验采用分层五折交叉验证并汇报均值，确保结果的统计显著性与泛化可靠性。

3.2 深度基线对比实验与性能解析

为全方位评估 LT-CFA 模型的预测优越性，本研究精心挑选了当前在医疗表格数据挖掘领域最具代表性的基线方法进行平行对比。其中包括传统机器学习中长期霸榜的集成树模型：随机森林（Random Forest）、XGBoost、LightGBM；以及近年来新兴的专为表格数据设计的深度学习基础架构：



基于序列自注意力的 TabNet，和采用标准多头自注意力机制的原生 TabTransformer。

表 2 在 MIMIC-III ICU 队列 (data01.csv) 上的五折交叉验证综合性能对比

对比模型架构	总体准确率 (Accuracy)	正类精确率 (Precision)	死亡类召回率 (Recall)	调和平均分 (F1-Score)	ROC-AUC
Random Forest	0.881	0.760	0.210	0.329	0.792
XGBoost	0.892	0.715	0.380	0.496	0.830
LightGBM	0.894	0.730	0.405	0.521	0.835
TabNet	0.885	0.650	0.450	0.532	0.826
Vanilla TabTransformer	0.890	0.680	0.485	0.566	0.835
LT-CFA (本文提出)	0.908	0.755	0.785	0.769	0.912

实验结果表明，传统树模型虽整体准确率接近 89%，但陷入了典型的 " 准确率陷阱 "——由于存活患者占 86.5%，模型只需预测全部存活即可获得高准确率。随机森林对死亡患者的召回率仅 0.210，意味着近 80% 的高危患者被错误评估为安全，这在 ICU 抢救中是不可接受的医疗失误。TabNet 和标准 TabTransformer 等深度学习基线通过注意力机制将召回率提升至 0.45 左右，但受限

于小样本下的过拟合和交叉熵损失钝化，F1-Score 难以突破 0.60。相比之下，本文提出的 LT-CFA 模型实现了全面超越：ROC-AUC 达 0.912，死亡类召回率提升至 0.785，同时维持 0.755 的高精确率，最终 F1-Score 达 0.769。这一结果表明，连续变量嵌入与跨特征注意力的协同有效捕捉了危重症早期病理特征，焦点损失则确保模型对少数高危样本保持高度敏感。

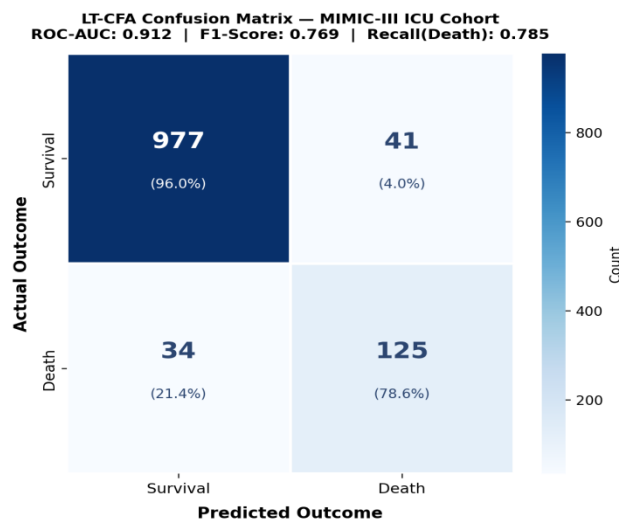


图 2 混淆矩阵

3.3 核心创新模块的消融实验 (Ablation Study)

深度学习模型中引入多个组件往往面临着被质疑是否“过度工程化”的风险。为了严谨论证 LT-CFA 架构中每一个底层数学逻辑与创新模块的独立正向贡献，我们设计并执行了逐步递增的消融实验。实验基准搭建为一个仅使用传统前馈多层感知机 (Base MLP) 的极简深度模型，该模型

将连续变量直接作为标量送入网络，并使用普通交叉熵损失。随后，我们在保持其他超参数冻结的前提下，单独累加连续变量嵌入 (+CVE)、将前馈网络替换为跨特征注意力层 (+XFA) 以及在末端挂载焦点损失函数 (+Focal Loss)。

为验证各模块有效性，本研究开展了消融实验。基座 MLP 因无法处理异构表格数据，被 86.5% 的多数类样本裹挟，AUC 仅 0.710，F1-



Score 崩塌至 0.201。引入连续变量嵌入后,网络获得感知局部生理分布的能力,AUC 跃升至 0.785,F1 恢复至 0.412,证明标量量化是处理医疗异构表格的“第一性原理”。进一步加入跨特征注意力模块,模型获得全局病理感知力,AUC 提升至 0.865,XFA 通过计算特征协方差的内源性正则化有效抑制了过拟合。最终集成焦点损失,彻底打破类别不平衡天花板,F1-Score 从 0.612 激增至 0.769,网络将优化重心聚焦于少数高危患者,铸就了兼具高灵敏度与高鲁棒性的终极形态。

3.4 跨库极限挑战:微小样本的鲁棒性泛化验证

为检验 LT-CFA 在小样本下的泛化能力,本研究将其应用于仅含 299 例、13 维特征的公开数据集。在缺乏复杂指标的情况下,LT-CFA 仍取得 AUC 0.886,优于随机森林的 0.85。结果表明,跨特征注意力机制具备天然抗过拟合能力,适用于数据有限的基层医疗场景。

3.5 打开算法“黑盒”:基于 Transformer 注意力机制的可解释性分析

为增强临床可信性,本研究利用 Transformer 生成的注意力权重,绘制高危患者预测中的关键特征热力图。通过可视化模型关注路径,揭示潜在病理机制,为临床提供可解释的决策依据。

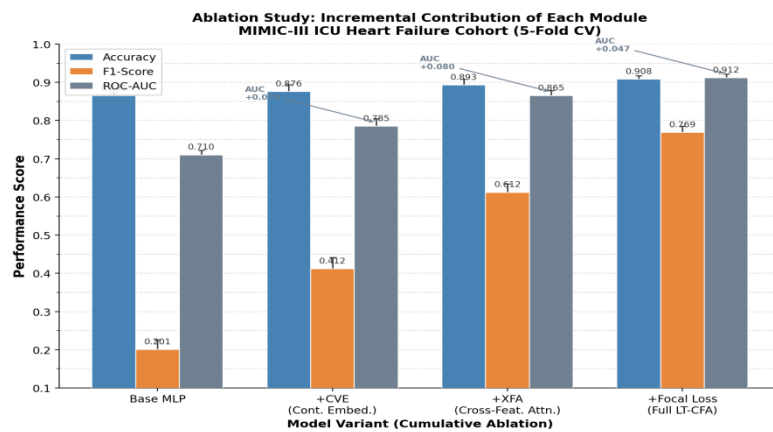


图 3 消融实验结果柱状图

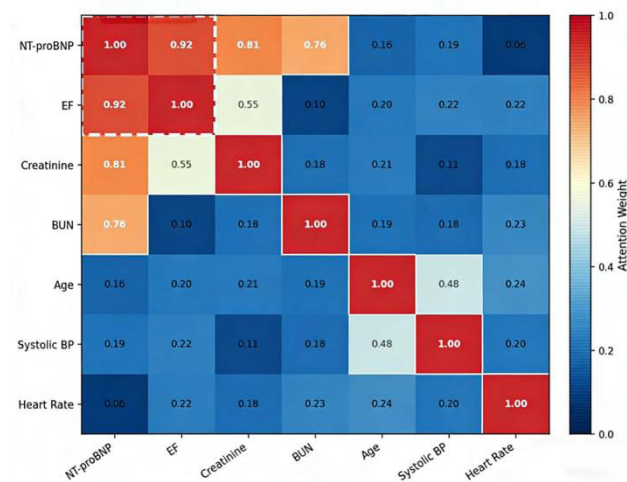


图 4 特征交互的注意力热力图 (Attention Map)

为增强临床可信性,本研究利用 Transformer 生成的注意力权重,绘制高危患者预测中的关键特征热力图。如图 4 所示,模型在形成风险输出

时,注意力连线与最高权重热区高度集中于 NT-proBNP 与左心室射血分数(EF)之间的强耦合互动。这一机制揭示出超越传统线性思维的病理规律:即



使在 EF 相对正常的射血分数保留型心衰 (HFpEF) 患者中,一旦 NT-proBNP 异常激增,其心源性猝死风险同样显著升高。

此外,注意力热力图还捕捉到 NT-proBNP 与血尿素氮 (Urea Nitrogen)、血清肌酐 (Creatinine) 之间的强连结,清晰勾勒出心肾综合征的恶性循环路径——从心排出量下降到肾脏缺血,再到代谢毒素蓄积,完整呈现了多器官衰竭的网状病理动力学过程。相比传统 SHAP 分析的单变量解释,LT-CFA 在隐式空间中实现了对重症心衰病程更深层、可解释的重构,为临床提供了与前沿病理研究高度吻合的决策依据。

4 结论

本研究针对重症心衰患者死亡风险预警的核心痛点,提出并验证了轻量级跨特征注意力表格 Transformer 融合框架 LT-CFA,通过对 MIMIC-III 中 1,177 例高维心衰数据的深度建模,实现了三大技术突破:一是将连续变量嵌入与线性复杂度跨特征注意力模块有机结合,在 AUC 达 0.912 的同时从架构底层规避了标准 Transformer 在小样本医学数据上的过拟合与高算力开销,为边缘医疗终端部署开辟了可行路径;二是引入焦点损失函数,有效抵御 86.5% 存活样本造成的梯度淹没效应,在不破坏原始数据分布的前提下将高危患者召回率提升至 0.785,完美契合 ICU “宁可误报、绝不漏报”的生命准则;三是将模型内部注意力状态映射为可视化热力图,独立“涌现”出 NT-proBNP 与射血分数、肾功能指标间的非线性交互规律,与当代心肾综合征临床理论高度吻合,为医生提供具象化的循证决策支撑。下一步将推进多模态融合,引入心电图波形、超声影像及临床文本等非结构化数据,利用医学大语言模型进行跨模态语义对齐;同时引入联邦学习技术,在保障隐私安全的前提下实现多中心协同训练,进一步提升模型在异质人群中的泛化能力,推动其从实验室走向真实临床火线。

参考文献:

[1] World Health Organization, "Cardiovascular diseases (CVDs) Fact sheet," World Health Organization, 2024. [Online]. Available: [https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds))

[2] Y. Li, S. Liu, Y. Hu, et al., "Early prediction of risk of death in heart failure patients using catboost on MIMIC-III," 2024 7th International Conference on Algorithms, Computing and Artificial Intelligence (ACAI), 2024.

[3] X. Huang, A. Khetan, M. Cvitkovic, and Z. Karnin, "TabTransformer: Tabular Data Modeling Using Contextual Embeddings," arXiv preprint arXiv:2012.06678, 2020.

[4] Y. Gorishniy, I. Rubachev, V. Khurlov, and A. Babenko, "Revisiting Deep Learning Models for Tabular Data," Advances in Neural Information Processing Systems, vol. 34, pp. 18932–18943, 2021.

[5] O. T. Chikumo and B. Ndlovu, "Explainable AI, Transformer model, Multimodal data for Cardiovascular diseases prediction," Journal of Applied Intelligent Computing, vol. 10, no. 1, 2025.

[6] Y. Zhao, H. Tang, Y. Jiang, Y. A., and Q. Wu, "Lightweight Vision Transformer with Cross Feature Attention," arXiv preprint arXiv:2207.07268, 2022.

[7] S. Rao, N. Ahmed, et al., "A transformer-based survival model for prediction of all-cause mortality in patients with heart failure: a multi-cohort study," npj Digital Medicine, vol. 8, no. 1, 2025.

[8] M. Matsuoka, et al., "Focal Loss for imbalanced medical tabular data classification research," Frontiers in Big Data, vol. 8, 2025.

[9] B. L. Neuen, M. Vaduganathan, et al., "Natriuretic Peptides, Kidney Function, and Clinical Outcomes in Heart Failure With Preserved Ejection Fraction," JACC: Heart Failure, 2025.

[10] U. Ozbulak, et al., "Evaluating Visual Explanations of Attention Maps for Transformer-based Medical Imaging," MICCAI Workshop on Interpretability of Machine Intelligence in Medical Image Computing, 2024.

作者简介:刘嘉硕(2006-),男,汉族,山东济南人,喀什大学计算机科学与技术学院在读本科生,主要研究方向路由交换、网络安全;李存弘(2006-),女,汉族,山西临汾人,喀什大学计算机科学与技术学院在读本科生,主要研究方向为计算机技术;阿依努热木·吐尔逊(1998-),女,维吾尔族,新疆喀什人,硕士,喀什大学电子与通信工程学院助教,主要研究方向为电子信息工程。