使用人工智能技术预测股骨脆性骨折住院患者院内死亡率的模型

1. 数据来源与患者选择

数据来源为数据表1-6中的数据。

1. 输入特征

模型使用的变量包括年龄、性别、诊断、住院天数、检查报告、检验信息、手术时长或护理信息。

就诊流水号：指的是每一个独立的患者，同一个患者的就诊流水号保持不变。

EMPI：是指每一个患者的每一次检查的特征码，如果一个患者的某一项检查了两次，以第一次检查的结果（即EMPI数值小的这一次检查）为准。

以下表格中的项目是需要纳入研究的内容数据：

数据表1：EMPI、就诊流水号、转归情况、就诊年龄、住院天数、植入材料费用、性别、出生日期、ABO血型、Rh血型

数据表2：EMPI、就诊流水号、手术时长、手术部位、手术方式、是否关节置换、是否下腔静脉滤器植入

数据表3：EMPI、就诊流水号、体重指数、肥胖等级、疼痛评分、Apache评、VTE评分、跌倒评分

数据表4：EMPI、就诊流水号、检查部位、检查结论

数据表5：所有内容

数据表6：所有内容

1. 模型开发与验证

3.1 数据划分

数据集以8:2的比例分为训练集和测试集。该测试集仅用于独立测试我们开发的AI模型，而不是用于训练器内部验证。首先使用训练数据数据进行了5折的交叉验证，以确认其泛化能力。训练数据集被随机洗牌并分层为5均等组，其中4组是从训练模型中选择的，其余组用于内部验证。通过移动内部验证组重复此过程5次。

3.2 数据平衡

由于存活患者数量远高于死亡患者数量，使用合成少数过采样技术(SMOTE)对存活患者数据进行上采样，以防止模型偏向存活患者。

3.3 模型开发

开发了一个基于深度神经网络(DNN)的AI模型，使用上面提到的数据特征作为输入层。最终输出层通过sigmoid层提供死亡率概率。

3.4超参数优化

通过搜索超参数，例如全连接层的深度和宽度，找到最佳模型。

3.5 集成方法

采用集成方法来组合5折交叉验证的5个模型，使用软投票法对概率进行平均。

1. 性能评估与比较

4.1 评估指标

性能评估基于敏感性、特异性、准确性、平衡准确性和ROC曲线下面积(AUROC)。

4.2模型比较

将本研究的DNN模型与其他机器学习模型进行比较，包括AdaBoost、XGBoost、LightGBM等。

通过这些方法，研究成功开发了一个能够有效预测股骨脆性骨折住院患者院内死亡率的AI模型。