

運用深度學習建構 ICU 急性心肌梗塞患者死亡率之 預測模型

摘要

根據衛生福利部統計，在 111 年死因結果分析中心臟疾病在國人十大死因中位居第二，僅次於癌症，在眾多心血管疾病中，冠狀動脈疾病、心律不整、心臟衰竭、心肌梗塞、動脈硬化為較常見的心血管疾病，其中心肌梗塞是常見的冠狀動脈心臟病，心肌梗塞對於心血管疾病的死亡率有重要的密切關係，而近年來心血管疾病的死亡率更是呈現逐年上升的趨勢，因此要瞭解導致心肌梗塞死亡的關鍵因素以及如何降低心肌梗塞死亡率是一項重要的課題。本研究採用重症醫學資料庫(Medical Information Mart for Intensive Care, MIMIC)的 MIMIC-III 數據集，患者人數總共有 3985 位，其中存活人數有 3433 位，死亡人數有 552 位，前處理分別利用均值插值法和 K 最近鄰插值法以及多重插值法進行缺失值處理，後續利用三種特徵選擇方法(XGBoost、LightGBM、CatBoost)找出對於真正影響急性心肌梗塞死亡率預測的特徵，最後使用深度學習模型(DNN、1D-CNN)進行對急性心肌梗塞死亡率的預測，以提升模型精確度。本研究透過模型搭配共有 18 種組合進行比較，最後找出最佳的組合為均值插值搭配 XGBoost 以及 1D-CNN 預測模型，模型績效 AUC 則為 97%，並找出影響急性心肌梗塞患者死亡風險的 10 個重要因子為平均動脈氧分壓、收縮壓最大值、收縮壓最小值、膽紅素最大值、體溫最小值、白血球計數最小值、舒張壓最大值、年紀、舒張壓最小值及心率最大值，希望此結果能提供給醫療人員在診斷方面有個參考依據，對於患者能更迅速的診斷，以降低急性心肌梗塞患者的死亡風險。

關鍵詞:急性心肌梗塞、深度學習、均值插值法、K 最近鄰插值法、多重插值法、XGBoost、LightGBM、CatBoost