

# 利用注意力時序機制滾動式預測急性腎衰竭風險

Rolling prediction of acute kidney injury using

Hierarchical attention mechanism and Time series model

指導教授: 蔣榮先

專題成員: 郭冠良、蔡宇暉

開發工具: Anaconda、Google Colab

Django

測試環境: Python 3.7

## 一、簡介:

根據台北榮民總醫院學訓的資料顯示，以重症加護病房治療的病人為例，發生急性腎衰竭的比率為 30~40%，需要洗腎的比例高達 50% 以上，不僅伴隨高昂醫療支出，高達 50~80% 以上的死亡率更是令人無法接受。急性腎損傷後的倖存者常常腎功能無法完全恢復，使得慢性腎臟病與末期腎病需長期洗腎者之人數越來越多。(資料來源：台北榮民總醫院學訓)

而雖然在學界已有利用諸如 XGBoost、LSTM 等等模型預測病人急性腎衰竭風險的紀錄，但此類模型在遇到醫療現場不同檢驗項目量測時間和數量不穩定的狀況時，往往需要以補差值的方式進行處理，其真實性及合理性一直是醫師所擔心的問題。因此我們希望能在不對原始資料進行補值的狀況下，由機器學習模型，對病人發生急性腎衰竭的風險進行預測，藉此協助醫師提早對病患進行治療或任何緊急處理，達到預警的目的。

本次使用的 Dataset 採用成大醫院重症醫學科所提供的 ICU 臨床檢驗資料，並選用了包含(1)病人基本資料: 性別、年齡等等(2)血壓資料共 2 項(3)血液檢測項目共 41 項(4)尿液檢測項目共 4 項(5)尿量值等特徵進行預測，預測方式基本是由使用者提供病人 11 個小時的檢測資料，由模型預測病人接下來 6 小時後發生急性腎衰竭的機率，最終的滾動式預測結果則是以每 11 小時為單位滑動，且不斷預測 6 小時後的機率所形成的機率分佈圖。

而在模型方面，我們在嘗試過許多模型後選擇以基於 attention 機制所建構的 LSAN 模型作為基礎，並修改其結構使其符合本次任務目標，其特點是能以階層式 attention 機制的方式以單一病人不同時間的資料和單一病人同一時間不同檢測資料兩種階層對病況進行分析預測。

## 二、測試結果:

以下為最終選用之模型和其他模型比較之結果，其中 LSAN 基於 attention 機制，也是本次所使用之模型；我們另外採用另一不需要補值的 transformer-based 和 LSTM-based 的模型比較不同機制在此次目標上之表現，以下為比較之結果:

模型	AUC	Precision	Recall	F1 Score
LSAN	0.78	0.84	0.86	0.82
Transformer-based	0.65	0.76	0.86	0.81
LSTM-based	0.91	0.86	0.66	0.75

從結果中可看出 LSTM-based 的模型雖在 AUC 單一指標上取得不錯表現，但其他指標皆顯著低於其他兩者，而 LSAN 和 transformer-based 的模型兩者各指標表現平均，但 LSAN 仍略勝一籌。

而下圖則為將病人資料滾動 12 小時之機率分布圖，左圖為最終無發生 AKI 之病人，右圖則為最終有發生 AKI 之病人，兩圖之縱軸為病人發病機率，橫軸則為時間，可看出不會發生 AKI 的病人模型皆會預測出較低的機率，而會發生 AKI 的病人模型預測出的機率會隨著時間越來越高:

