

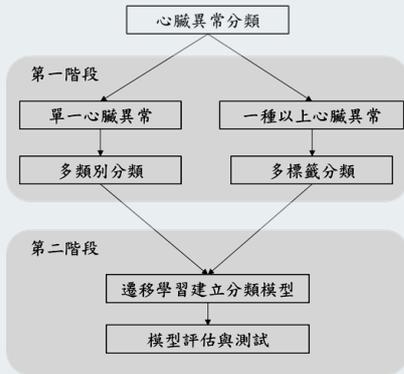
# 使用不同取樣率心電圖資料預測心血管疾病症狀

指導教授：林真如 博士 學生：王政方、陳昱婷

## 研究動機與目的

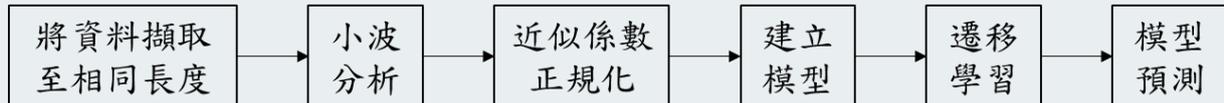
根據台灣衛生福利部統計，國人 2021 年因心臟疾病死亡人數占所有死亡人數的 11.9%，死因排名第二，因此心血管疾病檢測對醫界和民眾而言都是持續關注的議題。心血管疾病可以透過心電圖檢查，藉由分析波形判斷有何種心臟異常，隨著人工智慧技術的發展，機器學習和深度學習有許多在醫學上的應用，可協助醫生的醫療決策，提升醫療品質和效率。

## 研究方法



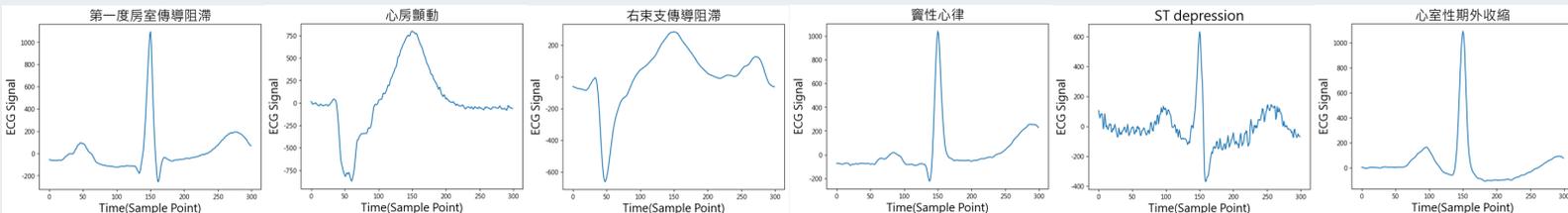
一位患者可同時有多種心臟異常症狀，本研究使用的資料集多數患者僅有一種心臟異常症狀，因此本研究分別研究兩個子題：(1)多類別分類：預測病患的單一心臟異常、(2)多標籤分類：預測病患是否有多種心臟異常。

本研究分為兩階段，第一階段：針對取樣率 500Hz 的心電圖資料建立預測心臟異常症狀的分類模型，第二階段：利用遷移學習將已建立的預測模型運用在其他取樣率的心電圖資料上進行預測。



## 研究結果

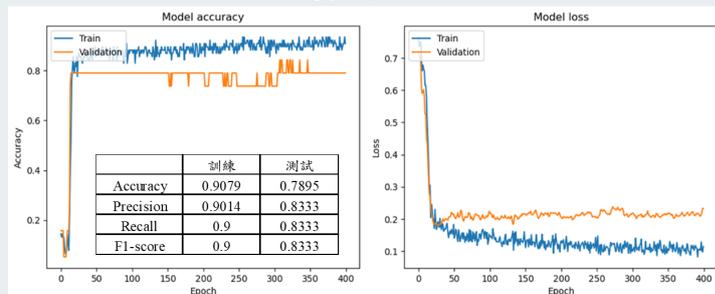
本研究使用 George B. Moody PhysioNet Challenge (2021)提供的資料集：CPSC 2018 和 PTB。兩個資料集來自不同醫院及不同國家，取樣率也不同，CPSC 2018 為 500Hz，PTB 為 1000Hz。



### 多標籤分類

5次K-fold交叉驗證結果	訓練(平均)	預測(平均)
Loss	0.2218	0.2949
Accuracy	0.7197	0.6467
Precision	0.8236	0.7396
Recall	0.6438	0.5638
F1-score	0.7157	0.6329

### 遷移學習結果



### 多類別分類

#### LDA (K=7) 混淆矩陣準確度

```
[[ 88 29 19 4 3 12]
 [ 23 119 26 13 18 13]
 [ 17 24 203 42 15 12]
 [ 2 10 35 99 20 2]
 [ 2 27 25 18 89 3]
 [ 1 18 7 10 9 72]]
Accuracy 0.5934455270150576
```

#### LDA及Random Forest 10次

加總後平均的結果

	LDA	RF
accuracy	0.39814	0.39736
precision	0.45198	0.44719
recall	0.39814	0.39736
F1	0.38356	0.3821

#### Random Forest (K=3) 混淆矩陣準確度

```
[[ 77 21 27 4 3 9]
 [ 10 124 37 9 9 3]
 [ 10 21 247 31 11 4]
 [ 1 6 37 118 22 0]
 [ 0 19 42 9 90 1]
 [ 5 14 16 13 14 65]]
Accuracy 0.6386182462356067
```

## 結論

- 使用小波分析進行特徵萃取，使模型準確率達到約 0.7。
- 使用數值型的心電圖資料進行分類，減少使用圖片資料所花費的成本和資源。
- 使用遷移學習將預訓練模型應用在不同取樣率的心電圖資料上，可以提高分類的準確度、減少過擬合問題及重新訓練模型所花費的時間，也可以使用較少的資料得到好的結果。