

## FingerPrint classification

姓名: 陳澤昕

科系: 生醫產專

學號: 311356003

### I. Dataset:

本次訓練的資料集採用指紋辨識的資料集，Training set 的資料量有 500 筆，Testing set 的資料有 100 筆。

Training	Testing
500	100

### II. Feature selection:

#### Step 1:

本次訓練使用 Gabor filter 擷取圖片中的特徵，擷取的特徵有

Angles: 每 30 度為單位旋轉

Sigma: [ 3, 7, 10]

Lambdas: [ 3, 5, 7 ]

Gamma: [ 0.1, 1 ]

總共會擷取出 108 個圖片特徵

#### Step2:

對擷取出來得 108 種特徵進行標準化，避免資料數據差異過大應想 Model 效能。

#### Step3:

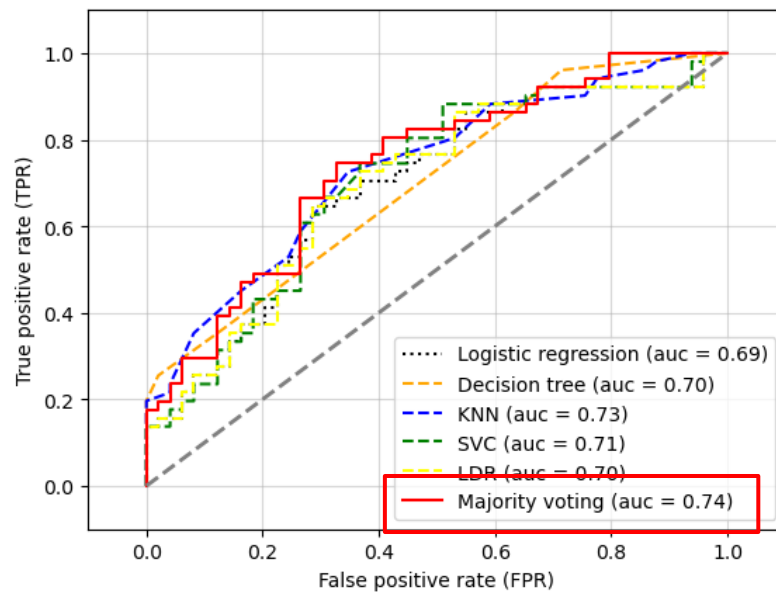
對 108 項特徵進行統計分析，採用 Mann-Whitney U Test 進行無母數分析，篩選出  $p\text{-value} < 0.05$  的特徵。

### III. Built Model:

使用 10-fold Majority vote classifier 演算法，參與投票的模型有 KNN、SVC、DecisionTreeClassifier、LinearDiscriminantAnalysis、LogisticRegression 將這些模型利用 GridSearch(5-fold) 調整出最佳參數後放入演算法中做預測。

#### IV. Validation:

將 Training set 分成 9 : 1(10-fold)做驗證可以看出 Majority vote classifier 有較好的 auc 因此本次專題採用的模型就是 Majority vote classifier。



#### V. Independent test:

將全部 Training set (500)放入 Majority vote classifier 中做訓練後，測試助教給出的 Independent set。

左右手預測檔名: final\_ans\_fingerprint.csv

男女預測檔名: final\_ans\_fingerprintGender.csv