FingerPrint classification

姓名: 陳澤昕

科系: 生醫產專

學號: 311356003

1. 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 的資資料有100筆。

|  |  |
| --- | --- |
| Training | Testing |
| 500 | 100 |

1. 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:

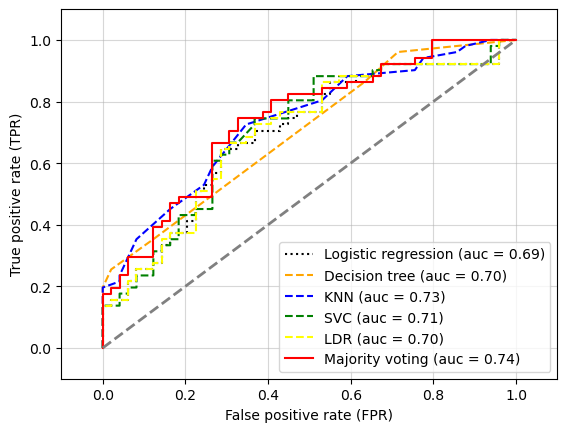
對108項特徵進行統計分析，採用Mann-Whitney U Test進行無母數分析，將資料分類為左右手，比較同特徵下左右手的資料平均值有無顯著差異。篩選出p-value < 0.05的特徵。

1. Built Model:

使用10-fold Majority vote classifier 演算法，參與投票的模型有有 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。

1. 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。



1. 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