

2019 ECT 作業五

1. 請用 python 依照步驟對 voice.csv 進行 SVM 分析, 過程中對所有重要程式步驟進行截圖並加以說明, 越詳盡越好。

使用 Python:(60%)

- (a) 請檢查資料集是否有空值, 如有空值即去掉該筆資料
- (b) 將最後一個屬性值"label"切分為 Target, 其餘屬型切分為 Feature
- (c) 將 Target 進行 encoded, 用 LabelEncoder 將 male 轉為 1, female 轉為 0
- (d) 將 Feature 用 sklearn.preprocessing 的 StandardScaler 進行標準化
- (e) 切分資料集與測試集, 設 test_size=0.33, random_state=1
- (f) 最後, 使用 sklearn.svm 裡的 SVC 進行分析, kernel 設為' linear ', 並印出模型最終的準確度

使用 Weka 軟體:(40%)

- (a) 使用 weka 裡的 SMO function 對 voice.csv 進行 SVM 分析, kernel 設為' linear ', Percentage spilt 設為 66%, 截圖並附上過程及準確率

- 繳交期限:4/3 (三) 中午 12:00
- 請轉檔為 PDF 格式, 檔名為: ECT_HW5_學號_版本.pdf
- 上傳至 LMS 作業區,遲交一天扣該次作業成績 5%
- 補交請上傳至 HW5 補交區, 僅開放補交至 4/10(三) 中午 12:00