

作業七

資料集是一些國家的經濟相關資料，請幫我依據這些資料判斷該國家是否為經濟自由的國家。

資料集：economic_freedom_data.csv (target : freedom)

參數設定：random_state = 10 , 兩層全連接的隱藏層(64,128) , batch_size = 8

以下題目請使用 Python 完成

1. 將從 CSV 讀取為 DataFrame 格式，並打亂原始順序，並進行前處理。
(5%)
2. 使用 scikit-learn 建立 RandomForest 分類器，並以 10 cross validation 評估模型在此資料集的分類表現(指標：F1-score，並印出平均分數)。(20%)
3. 使用 scikit-learn 建立 MLP 分類器，並以 10 cross validation 評估模型在此資料集的分類表現 (指標：F1-score，並印出平均分數)。(20%)
4. 使用 Keras 建立 MLP 分類器，並以 10 cross validation 評估模型在此資料集的分類表現 (指標：F1-score，並印出平均分數)。(20%)
5. 請問你在上述 3 題所使用的 F1-score 是以哪一類別為依據 (freedom、not-freedom、兩者平均)。(5%)
6. 再以 t-test 比較上述三個模型的表現，並簡述結論。(10%)
7. 導入 new_data.csv，來判斷該國家是否經濟自由。(10%)
8. 分析哪一個屬性在 RandomForest 的分類中最重要?(10%)

作業繳交說明：繳交期限： 5/31 (三) 0:00

- Python 題請繳交 ipynb 檔，檔名 ECT_HW7_學號。
- 程式中請以註解或文字方塊標示題號。
- 需確保程式執行上傳至 eeclass 作業區那一版本的不會出錯。
- 上傳至 eeclass 作業區，**遲交一天扣該次作業得分 5 %，最多 50%。**