

# 2019 ECT 作業八

一、 利用 Weka 對 DeerHunter.arff 進行前處理，依序完成以下步驟及

問題：

(a). Replace Missing Value, 需列出補上的值為何 (10%)

(b). Outlier Detection& Remove (10%)

(c). Attribute Selection, 請篩選出適合的屬性 (10%)

二、 請用 weka 對 BreastCancer.csv 對目標 diagnosis 進行 Ensemble

learning 並與未使用的結果進行比較(請列出重要過程及適當說明)：

(a). 以 10 Folds cross-validation 進行 J48 分類(5%)

(b). 以 10 Folds cross-validation 進行 Bagging 分類並選擇 J48 classifier  
進行分類(10%)

(c). 以 10 Folds cross-validation 進行 Bagging 分類並選擇 Randomforest  
進行分類(10%)

(d). 以 10 Folds cross-validation 進行 AdaBoost 分類並選擇  
DecisionStump 進行分類(10%)

三、 請用 python 對 BreastCancer.csv 對目標 diagnosis 進行 Ensemble

learning 並與未使用的結果進行比較(請列出重要過程及適當說明)：

(a). 以 10 Folds cross-validation 進行 DecisionTreeClassifier 分類(5%)

(b). 以 10 Folds cross-validation 進行 BaggingClassifier,

n\_estimators=10 分類(10%)

(c). 以 10 Folds cross-validation 進行 RandomForestClassifier 分類

(10%)

(d). 以 10 Folds cross-validation 進行 AdaBoost, n\_estimators=10 分類

(10%)

- 繳交期限： 5/15 (三) 中午 12:00
- 請轉檔為 PDF 格式, 檔名為：ECT\_HW8\_學號\_版本.pdf,  
並同時附上 python 的 ipynb.檔,命名格式同上
- 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 HW8 補交區, 僅開放到 5/22(三)中午 12:00