

# 機器學習實務與應用

## Homework #5

Due 2020 **Oct 26** 11:00PM

(一) 針對員工離職率(left)進行預測

資料檔案：[hr-analytics.csv](#)

作業要求：

1. 讀入資料，並判斷出那些數據格式不是數字，或是有缺失值。
2. 將非數字類型的資料進行必要的編碼。
3. 若有缺失值請填補。
4. 將資料切割成訓練集 70%，預測集 30%。
5. 利用 Logistic Regression 及 SVM 模型進行預測。
6. 針對各個特徵與離職率的關係進行探討。看是否透過特徵轉換提高預測之準確率。

繳交說明：請繳交 jupyter notebook 之檔案。若有討論部分也利用 jupyter notebook 說明。

(二) 繳交相關 AI、Machine Learning 競賽，譬如 AI CUP 愛文芒果影像辨識 (<https://aidea-web.tw/topic/72f6ea6a-9300-445a-bedc-9e9f27d91b1c>) 之報名證明。