

機器學習實務與應用

Homework #3

Due 2022 Oct 17 11:00PM

(一) 針對員工離職率(left)進行預測

資料檔案：[HW3_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 模型進行預測。
6. 請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) 使用 Linear Regression 預測在不同的時間，租借共享單車的人數預測(count)

資料檔案：[HW3_bike-sharing.csv](#)

作業要求：

1. 讀入資料。
2. 對日期欄位進行處理。
3. 利用 regression 推測使用人數。顯示出預測結果。

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。

(三) 繳交參與 AI、機器學習競賽證明。