

這兩次的助教實習課，教了前處理、分類器和評估方法的部分，課程也進行到了一半，可以嘗試著將這些應用於實際資料上。因此這份作業選擇機器學習分析中最常拿來當範例使用的『鐵達尼號資料集』，讓學生們試著實作預測看看。

Titanic: Machine Learning from Disaster


<https://www.kaggle.com/c/titanic>

(需註冊 kaggle，並將預測結果上傳，
一天上傳次數上限為 10 次，請同學盡早開始做)

繳交內容：

一份報告(.pdf)，包含

1. 資料集特徵資料說明、屬性特性說明
2. 對特徵做甚麼樣的分析? 哪些前處理? 採用哪些特徵? 原因?
3. 基於什麼理由選擇哪個分類器?
4. 採用的評估指標結果與觀察
5. 將預測結果上傳至 kaggle 並截圖測試的分數
(在 leaderboard 上的整列截圖，如下圖)

8533	yfav92386		0.77511	3	2m
------	-----------	--	---------	---	----

還有程式碼(.py)、上傳到 kaggle 的預測檔案(.csv)，
一同壓縮至一個資料夾(組別_學號姓名.zip)上傳至 moodle

※ 網路上也有很多分析此資料集的文章，可以參考但也請加入自己的看法

繳交期限 1/02(日) 12:00pm