

說明：

這次的作業是做鐵達尼號的乘客生存分析(Predict survival on the Titanic)，在做預測模型之前需要先做好資料的前處理分析，先將原本的 raw data 進行資料特徵萃取，把座艙等、票價、年紀等等因素萃取出來，知道那些特徵對於存活率的影響最大(相關性最高)，最後在將這些處理好的資料做隨機森林(Random Forest)，進行生存率預測，最後預測出來的分數是 0.78229。

截圖：

