**111601517 大氣二 邱皓群**

1.比賽簡介:太空郵輪泰坦尼克號，載著近13,000名乘客的飛船啟程，旅途中不慎撞上了一個時空異常，造成幾乎一半的乘客被傳送到了另一個維度。為了幫助救援隊伍並檢索失踪的乘客，您被挑戰要使用從太空郵輪損壞的計算機系統中恢復的記錄來預測哪些乘客被該異常傳送。/旅客資料:姓名、年紀、出發地、目的地等等，其中VIP跟睡眠艙有關聯，而要從這些資料中找出跟Transported之間的關係/目標是找出錯誤地方並成功預測出被傳送的乘客。

2.程式發想:首先載入了一份訓練資料和一份測試資料，這兩份資料都包含有關乘客的資訊。然後進行了一些簡單的資料分析和模型訓練，最終將測試資料的預測結果保存到了名為 'submission.csv' 的文件中。

程式主要的步驟和發想可以總結如下：

****(1)載入資料：**** 使用 Pandas 載入訓練資料和測試資料。

****(2)資料分析：**** 進行一些簡單的資料分析，比如計算平均年齡，以及根據不同條件計算一些比例。

****(3)訓練模型：**** 使用隨機森林分類器進行模型訓練。

****(4)保存結果：**** 將測試資料的預測結果保存到 'submission.csv' 文件中。

****(5)寫出結果、信息和當前儲存位置：**寫出**模型的預測結果以及一條成功保存的消息， 並提供**當前儲存位置**，這是為了查看保存文件的位置。

/我會選擇這一個比賽，因為題目內容跟上課的作業相似，因此操作上比較好理解，可)以較快上手，而我程式中也是運用上課給的方式做出模型，但是數據是使用我網路上參考的資料，而想出來的運算方式。

# (https://blog.csdn.net/zed\_faker1/article/details/123374344 **Kaggle系列（一）：Spaceship Titanic（太空飛船泰坦尼克號)、**<https://www.wpgdadatong.com/blog/detail/70687> **Pytorch 數據分析3-太空飛船泰坦尼克號預測分析**)

1. 結果:目前有成功做出預測被傳送乘客名單，但我覺得應該有錯誤，因為我假設的算法是比較簡單的，因此正確率可能不會很高，而之後改善方式，可以參考上面網站上比較難的假設，並學習他如何想到，使正確率上升，再來，目前所使用的方式是1對1比較，但以這題的變數來說，有可能不是使用單變數，而是要使用多變數，這樣就能降低誤差，提高預測準度。







