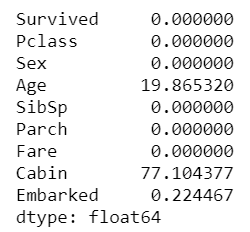
**【Lab1】Kaggle Titanic - Machine Learning from Disaster**

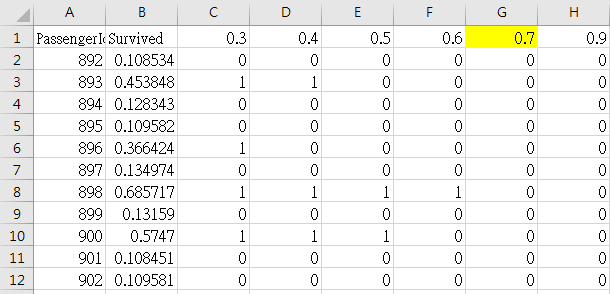
**Deep 柴**

1. 資料前處理
   1. 將 Unique 的資料行去掉: PassengerId、Name、Ticket
   2. 處理NA的資料

右方圖片說明有NA資料的資料欄位，依照不同狀況做處理

* 1. Age: 約20% 的資料有NA，數量不算多、且為數值型態，所以用中位數做填補
  2. Cabin: 約77%的資料是NA，所以直接Drop掉這一個Feature
  3. Embarked: 有NA的資料很少(兩筆)，所以直接刪去有NA的Row
  4. 女性存活之機率高達74%，而男性存活率只有18%，相差甚鉅。
  5. 將非數值Feature進行One-hot encoding
     1. Sex: 拆成Sex\_female、Sex\_male
     2. Embarked: 拆成 Embarked\_C、Embarked\_Q、Embarked\_S
  6. StandardScaler
     1. 進行標準化

1. 程式碼: 使用 Jupyter Notebook 運行，下圖是模型的訓練程式碼
   1. 使用三層Dense Layer搭建模型，測試結果最好的是三層都使用sigmoid作為activation function
2. 模型訓練: K-Fold、嘗試不同 Threshold
   1. K-Fold: K = 5
   2. Threshold: 用 Excel 進行嘗試

如下圖，將輸出的機率以0.3~0.9的Threshold嘗試，最後的最佳結果是0.7

1. 結果

