



第一層

1. 以性別來做分類
2. Sex=0 為女生
3. Sex=1 為男生
4. Sex<=0.5 會跑到左邊分支, Sex>0.5 則跑到右邊

第二層

1. 左邊以艙等做分類(以 2.5 為界)
2. 右邊以年齡做分類(以 9.5 歲為界)

第三層

1. 如果第二層以艙等為分類,此層都是以票價做分類
2. 如果第二層是以年齡分類,此層曾是以兄弟姊妹+老婆丈夫數量和艙等做分類

結果

1. 根據 Decision tree 的預測,393 個次是資料中,有 74 個預測錯誤,準確度為 81.17%
2. "pclass", "sex", "age", "sibsp", "parch", "fare", "embarked_C", "embarked_Q", "embarked_S" 的 feature_importances_ 是 [0.2247905, 0.59449224, 0.04272083, 0.0655803, 0., 0.07241612, 0., 0., 0.]
3. 以第二點的 feature_importances_ 來看, "embarked_C", "embarked_Q", "embarked_S" 和 "parch" 的重要性較低, 而 "sex" 的重要性最高