

Decision Tree Report

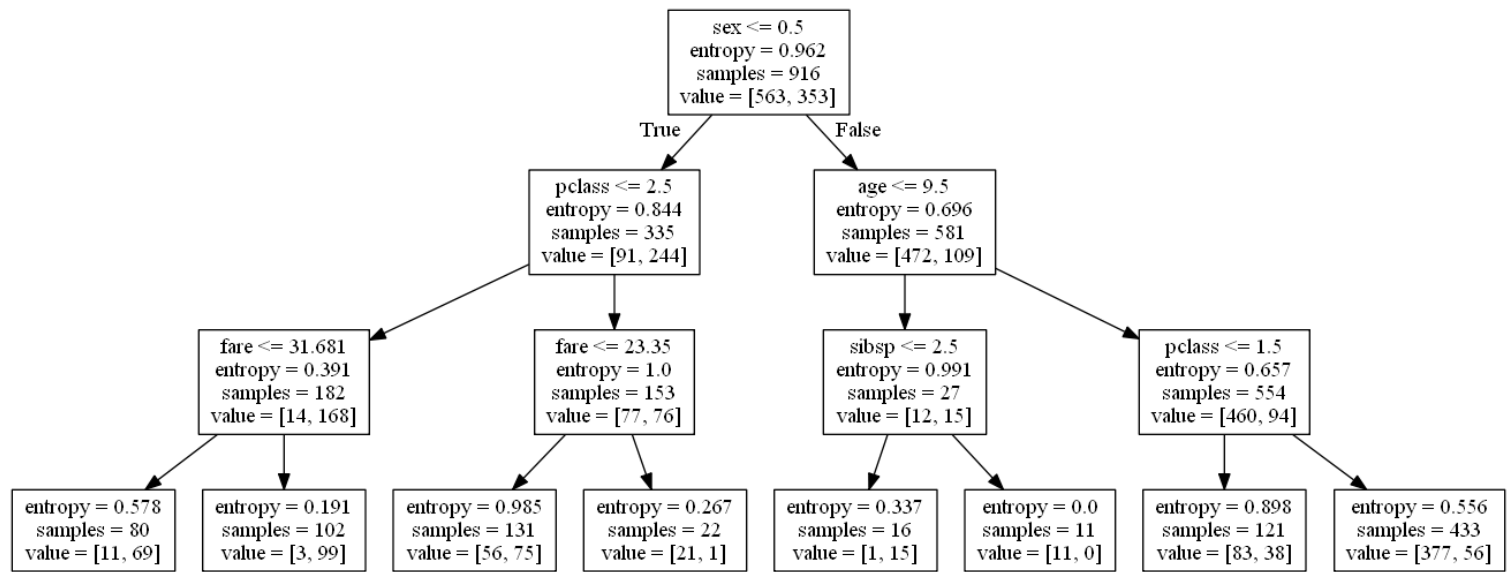
以鐵達尼號資料來做分析,主要 features 有 pclass, sex, age, sibsp, parch, fare,並以這些 features 判斷鐵達尼號乘客是否存活

My Rule

- 當 sex==1 時
 - ◆ age>=50
 - ✧ parch>=2
 - 活
 - ✧ sibsp>=2
 - 活
 - ✧ 此外
 - 死
 - ◆ age<=15
 - ✧ parch>=2
 - 活
 - ✧ sibsp>=2
 - 活
 - ✧ 此外
 - 死
 - ◆ 15<age<50
 - 活
- 當 sex==0 時
 - ◆ fare>=50
 - 活
 - ◆ fare<50
 - ✧ parch>=2
 - 活
 - ✧ sibsp>=2
 - 活
 - ✧ 此外
 - 死

結果: 與鐵達尼號資料比較,準確度僅 34.6%

Decision Tree



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,此層是以兄弟姊妹+老婆丈夫數量和艙等做分類

結果: 與鐵達尼號資料比較,準確度為 81.17%

Decision tree 以 sex 和 pclass 為影響比較大的因素,而我所設定的 rule 並沒有將 pclass 放進篩選的條件中