

● Confusion Matrix:

- 左至右為分類 1(not survived)及分類 0(survived)，Positive 為 not survived，Negative 為 survived。從 Confusion matrix 當中可以看出 True Positive 有 46 個，False Negative 有 31 個，False Positive 有 25 個，True Negative 有 128 個。

```
Confusion Matrix:
array([[ 46,  31],
       [ 25, 128]], dtype=int64)
```

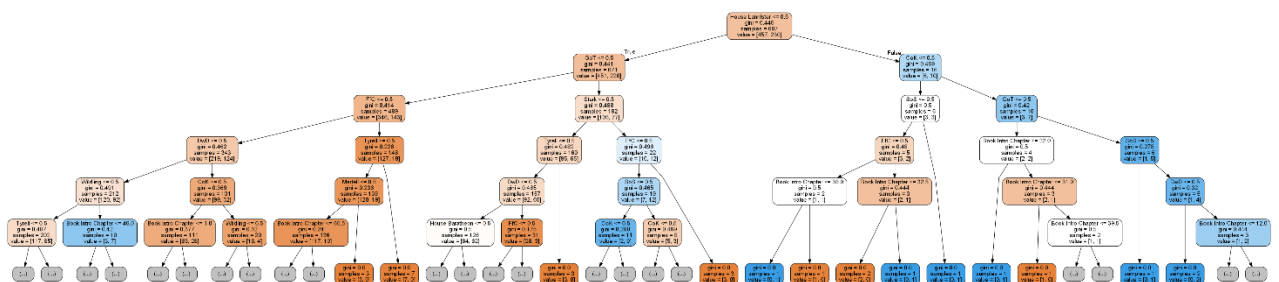
● Precision, Recall, Accuracy:

- 利用 Confusion Matrix 可以計算
 1. Accuracy=(TP+TN)/ALL=(46+128)/230=75.65%
 2. Precision=46/(46+25)=64.79%
 3. Recall=46/(46+31)=59.74%

其中 Precision 代表，在模型預測為死亡的情況下，有 64.79%的可能性是真的死亡。Recall 代表，在真實是死亡的情況下，有 59.74%的可能性被預測為死亡。而模型的正確度為 75.65%。

```
accuracy score is:
0.7565217391304347
precision score is:
0.647887323943662
recall score is:
0.5974025974025974
```

● 決策樹(Decision Tree)



在建立決策樹的參數當中，我將 max_feature 設為 sqrt，代表每一次在分裂節點的時候，都會隨機從全部特徵個數開根號後的特徵個數當中找一個最好的 split point。節點當中代表在分裂中所挑到的變數、在該點的 Gini 係數、落在該節點的資料個數、預測的存活與死亡的個數。