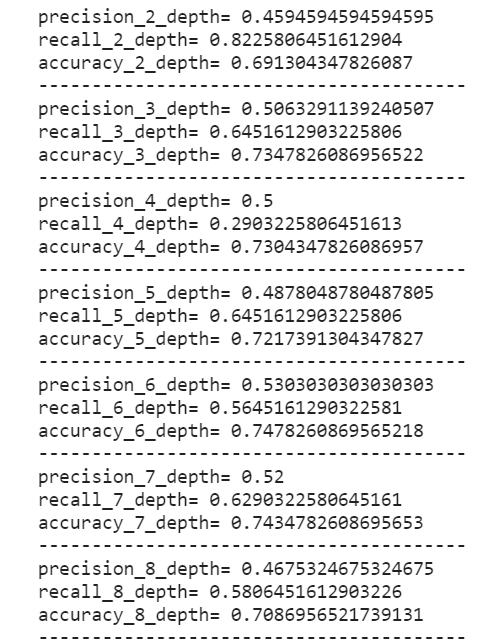
1. Precision, Recall, Accuracy

**樹深限制：3**

右圖為樹深2~8分別計算得到的Precision, Recall, Accuracy，下圖為依據右圖繪製的圖表

討論:

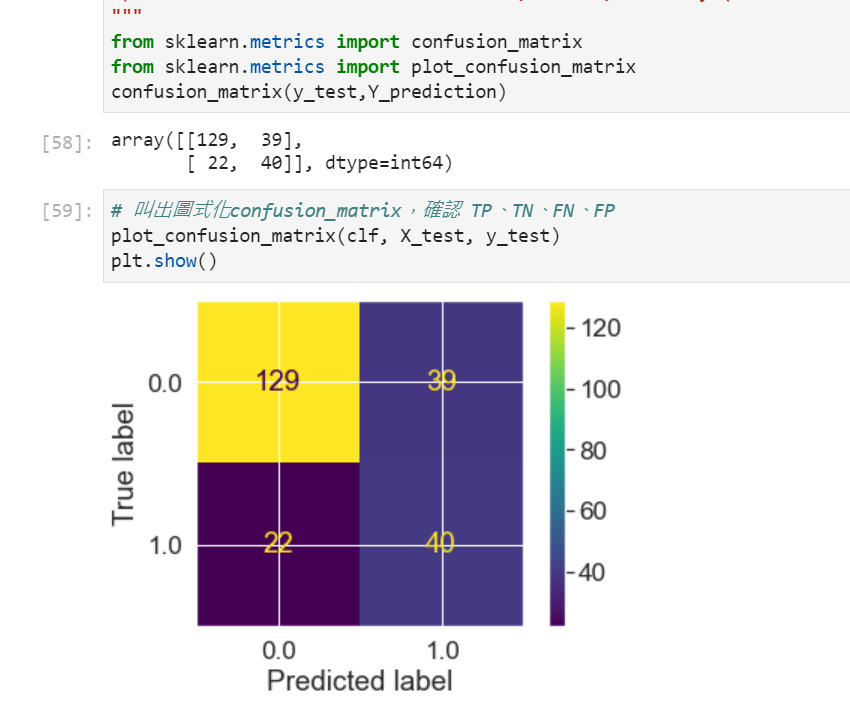
1. Precision一直在Accuracy之下：

有太多沒死亡的腳色被判斷為死亡，決策樹參考的特徵可能有問題，需要再做特徵篩選。

1. Recall在樹深=4時，劇烈的下降，但是在樹深=5時再度回升：

樹深=4時新納入的特徵應該有巨大的問題，若剔除此批特徵後再做一次決策樹也許能提升整體分數。

1. 決策樹(深度3)時採用的7個特徵中，有4個是家族姓氏([X8~X29])，可見家族陣營很大程度決定了腳色生存與否。
2. **Confusion Matrix的結果截圖**





False

X[14] <= 0.5

gini = 0.26

samples = 182

value = [154, 28]



gini = 0.48

samples = 15

value = [6, 9]



gini = 0.48

samples = 10

value = [6, 4]



True

X[7] <= 0.5

gini = 0.49

samples = 505

value = [288, 217]

X[6] <= 0.5

gini = 0.459

samples = 687

value = [442, 245]

1. **決策樹的圖**



X[23] <= 0.5

gini = 0.5

samples = 347

value = [170, 177]

gini = 0.47

samples = 98

value = [61, 37]

gini = 0.492

samples = 249

value = [109, 140]



gini = 0.34

samples = 143

value = [112, 31]

X[28] <= 0.5

gini = 0.378

samples = 158

value = [118, 40]



gini = 0.212

samples = 166

value = [146, 20]

X[22] <= 0.5

gini = 0.236

samples = 176

value = [152, 24]



X[0] <= 16.0

gini = 0.444

samples = 6

value = [2, 4]

gini = 0.0

samples = 3

value = [0, 3]

gini = 0.444

samples = 3

value = [2, 1]