CRITERION & MAX_LEAF_NODES

Classification

กลุ่มกลุ้ม

หอยหลอดดกรุ๊ป

กลุ่ม กลุ้ม X หอยหลอดด

Member

643020501-6	นายตะวัน เบ้าหล่อเพชร
643021260-7	นางสาวกิติลักษณ์ ลาดโฮม
643021261-5	นางสาวจารุพร การร้อย
643021263-1	นางสาวชนม์ชนก อังคุระษี
643021265-7	นายธนาธิป อินทรคีรี
643021266-5	นางสาวธิติพร ใจเอื้อ
643021268-1	นายพุทธิพงศ์ ยางนอก
643021273-8	นายศตวรรษ มูลสันเทียะ

Tawan Industries

CRITERION

gini

entropy

log_loss

ฟังก์ชันที่ใช้วัดคุณภาพ
ของการ split
โดย default = ‘gini’

CRITERION

GINI

Define

```
[ ]  Dtree_gini = DecisionTreeClassifier(random_state=0, criterion='gini' )
```

Train

```
Dtree_gini.fit(X_train,y_train)
```

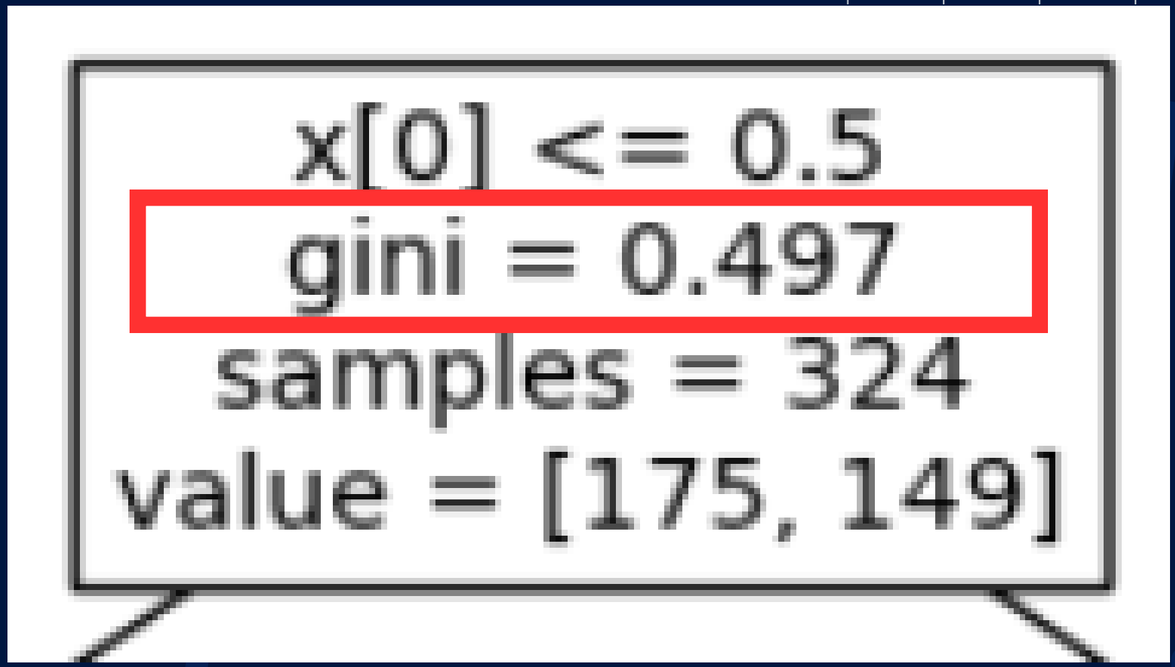
```
DecisionTreeClassifier(random_state=0)
```

Test

```
[ ]  y_predict_gini = Dtree_gini.predict(X_test)
```

```
[ ]  data1_score = accuracy_score(y_test, y_predict_gini)
     data1_score
```

```
0.8387978142076503
```



Define

```
[ ]  Dtree_entropy = DecisionTreeClassifier(random_state=0, criterion='entropy')
```

Train

```
Dtree_entropy.fit(X_train,y_train)
```

```
DecisionTreeClassifier(criterion='entropy', random_state=0)
```

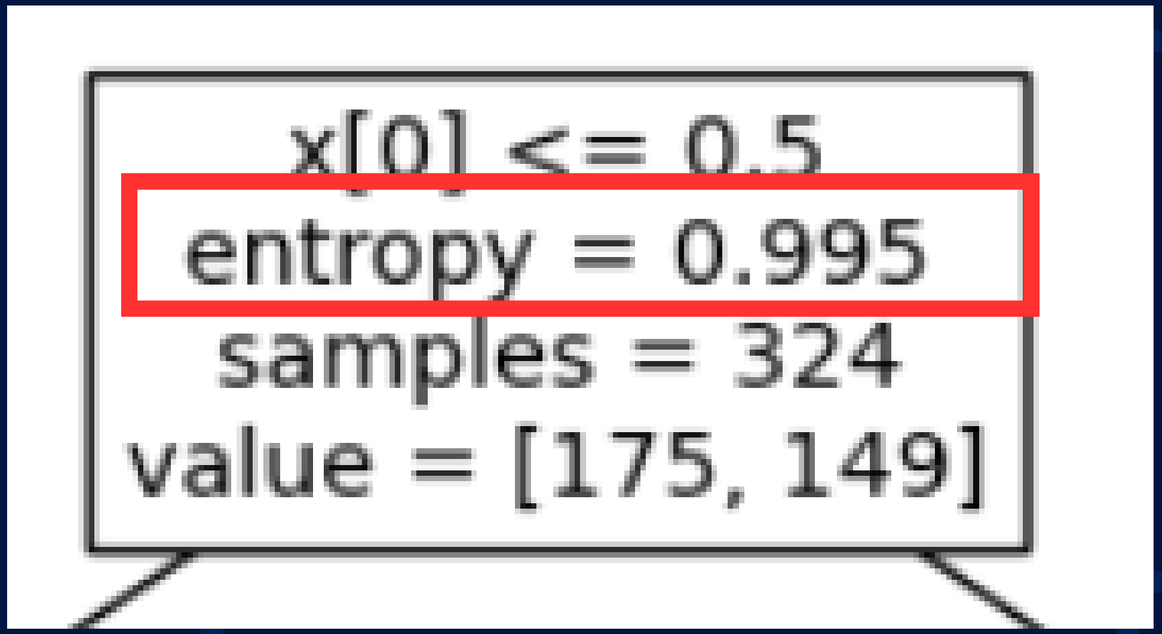
```
[ ]  _, ax = plt.subplots(figsize=(15,10))
     tree.plot_tree(Dtree_entropy, ax= ax);
```

Test

```
[ ]  y_predict_entropy = Dtree_entropy.predict(X_test)
```

```
[ ]  data1_score = accuracy_score(y_test, y_predict_entropy)
     data1_score
```

```
0.8387978142076503
```



CRITERION
ENTROPY

CRITERION

LOG LOSS

Define

```
[ ]  Dtree_log_loss = DecisionTreeClassifier(random_state=0, criterion='log_loss' )
```

Train

```
Dtree_log_loss.fit(X_train,y_train)
```

```
DecisionTreeClassifier(criterion='log_loss', random_state=0)
```

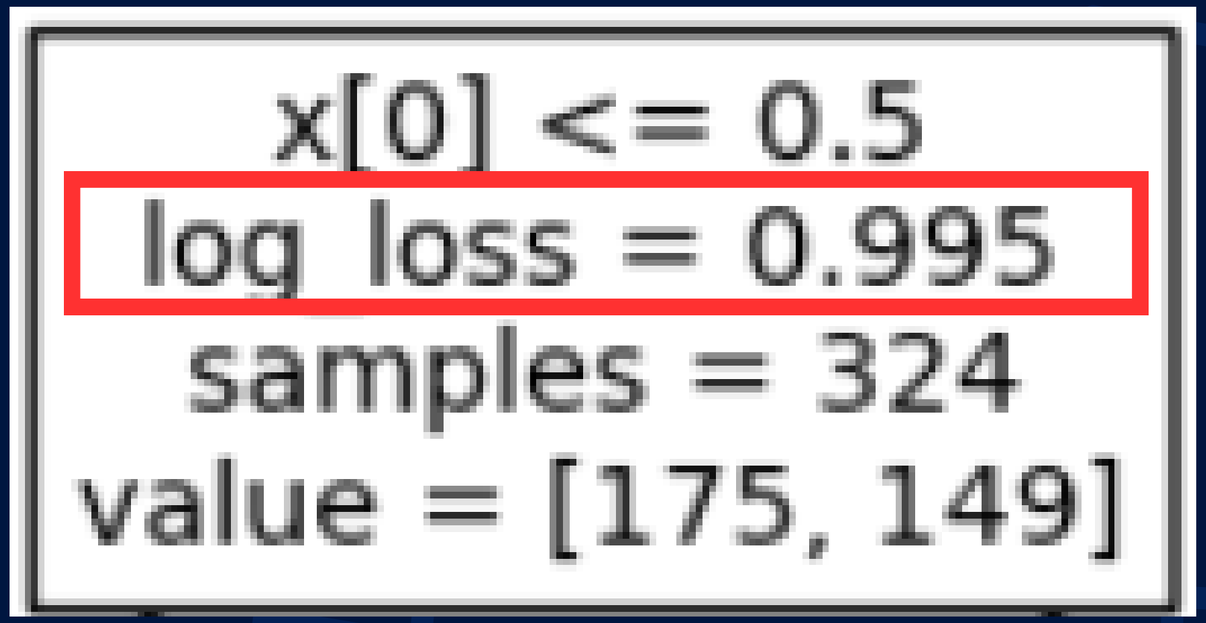
```
[ ]  _, ax = plt.subplots(figsize=(15,10))
     tree.plot_tree(Dtree_log_loss, ax= ax);
```

Test

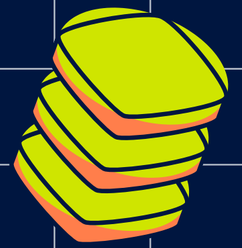
```
[ ]  y_predict_gini = Dtree_log_loss.predict(X_test)
```

```
accuracy_score(y_test, y_predict_gini)
```

```
0.8387978142076503
```



MAX _LEAF_ NODES

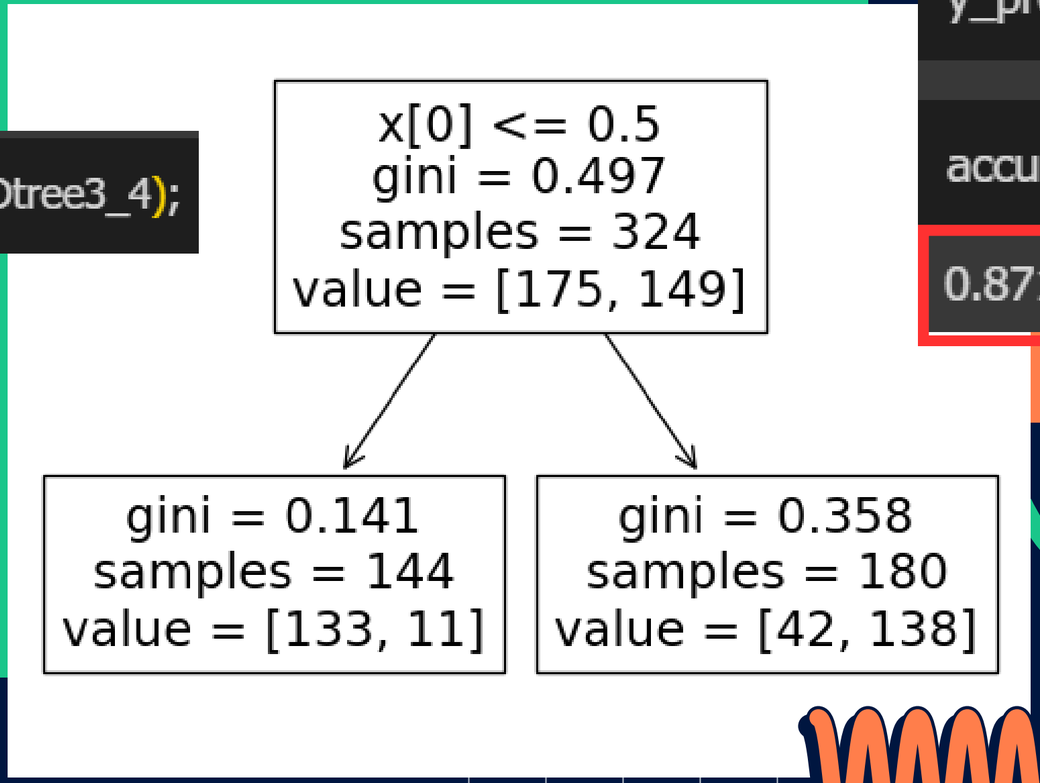


```
Dtree3_4 = DecisionTreeClassifier(max_leaf_nodes=2)

Dtree3_4.fit(X_train,y_train)
```

```
DecisionTreeClassifier
DecisionTreeClassifier(max_leaf_nodes=2)
```

(max_leaf_nodes=2)

```
tree.plot_tree(Dtree3_4);
```




```
y_predict3_4 = Dtree3_4.predict(X_test)

accuracy_score(y_test, y_predict3_4)
```

```
0.871584699453519
```

```
Dtree3_3 = DecisionTreeClassifier(max_leaf_nodes=5)

Dtree3_3.fit(X_train,y_train)
```

DecisionTreeClassifier

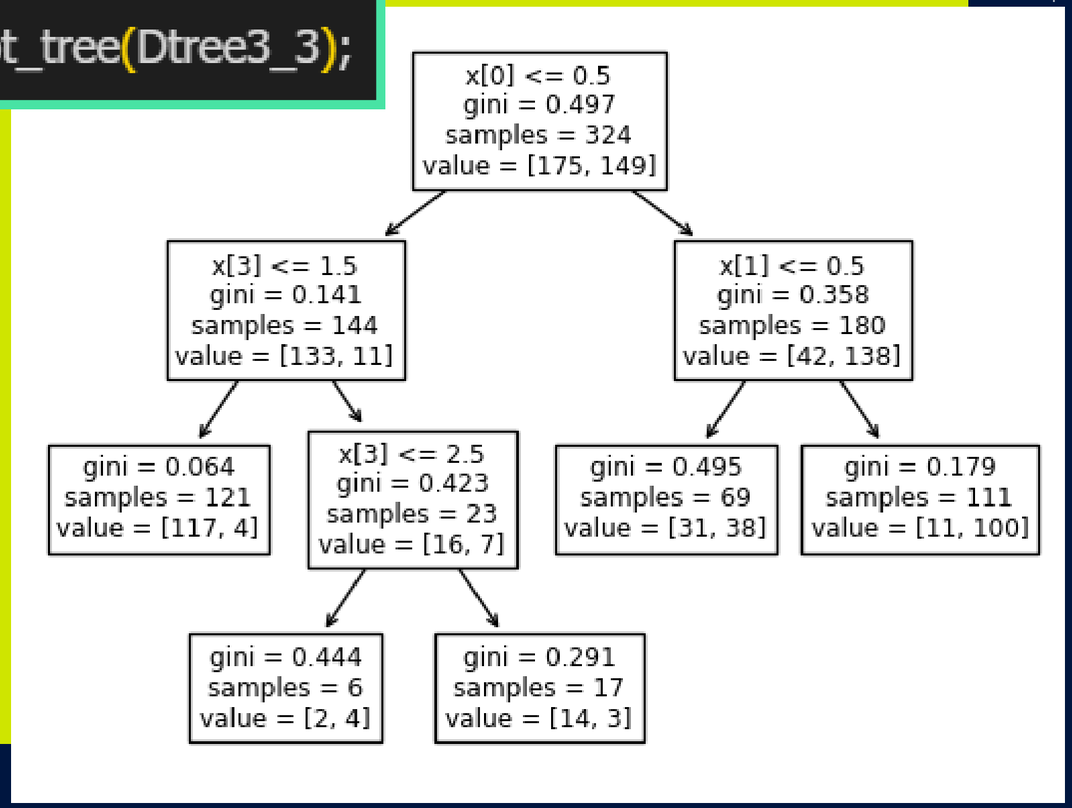
DecisionTreeClassifier(max_leaf_nodes=5)

```
y_predict3_3 = Dtree3_3.predict(X_test)

accuracy_score(y_test, y_predict3_3)
```

0.8715846994535519

(max_leaf_nodes=5)

```
tree.plot_tree(Dtree3_3);
```

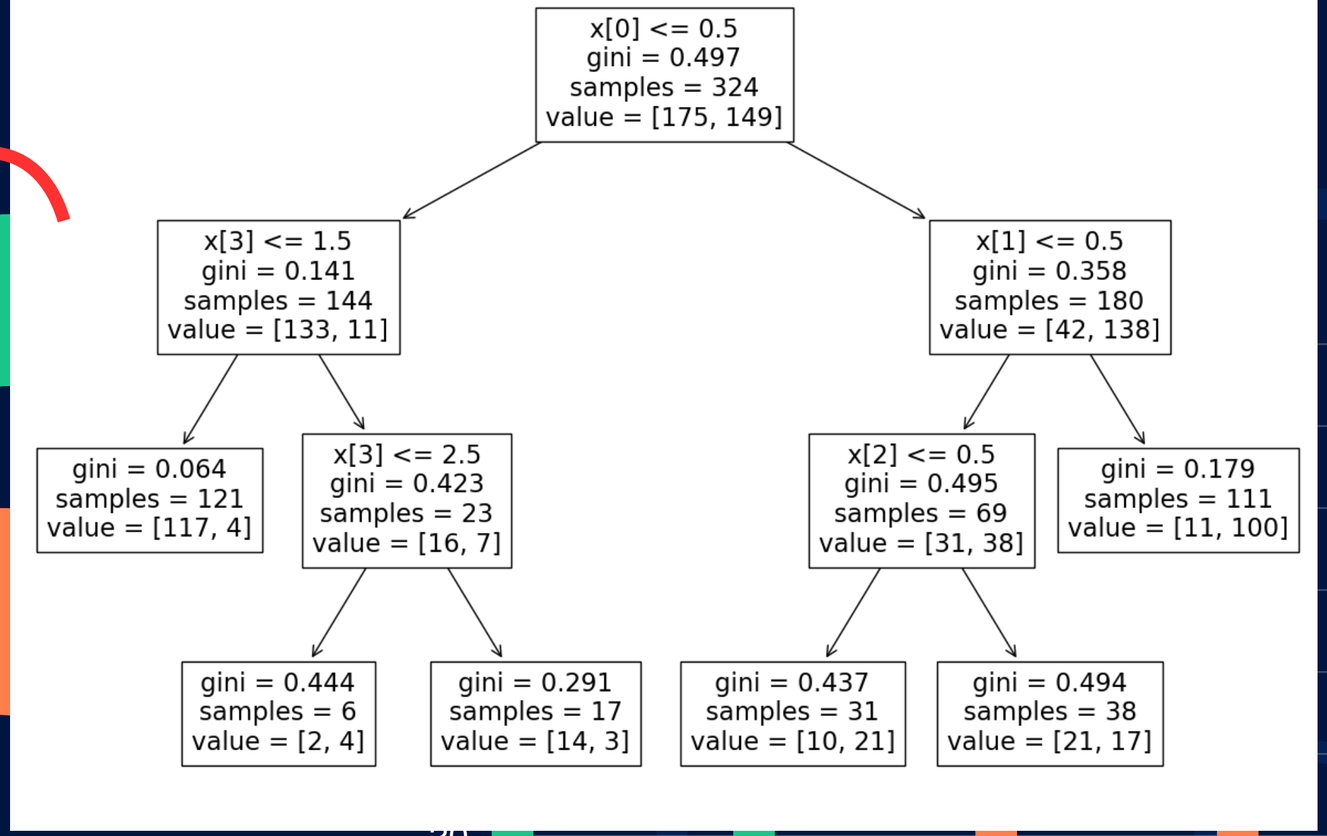


```
Dtree3_5 = DecisionTreeClassifier(random_state=0,max_leaf_nodes=6)

Dtree3_5.fit(X_train,y_train)
```

```
DecisionTreeClassifier
DecisionTreeClassifier(max_leaf_nodes=6, random_state=0)
```

```
max_leaf_nodes=6)
```

```
_, ax = plt.subplots(figsize=(15,10))
tree.plot_tree(Dtree3_5, ax = ax);
```




```
y_predict3_5 = Dtree3_5.predict(X_test)

accuracy_score(y_test, y_predict3_5)
```

```
0.83879781420076503
```

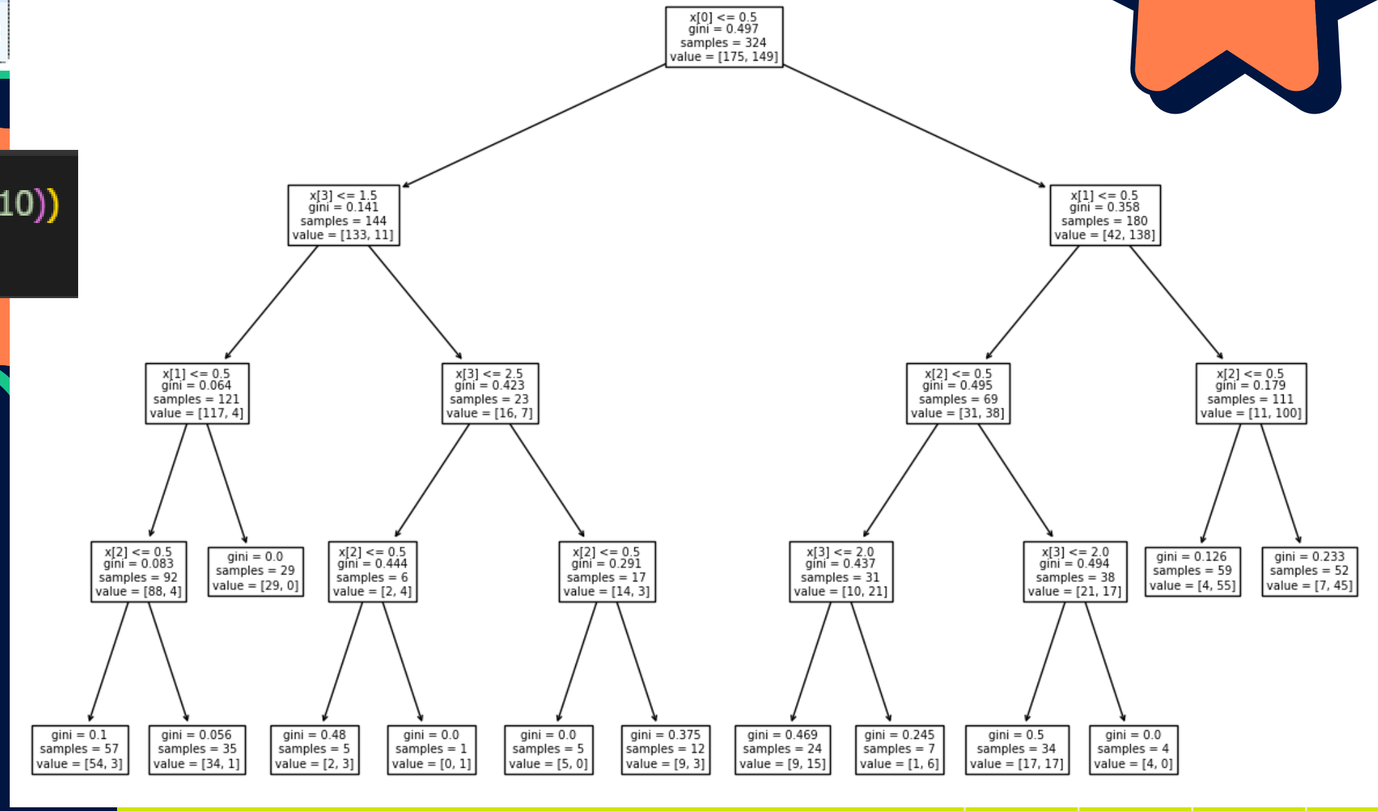
```
[54]  Dtree3_2 = DecisionTreeClassifier(max_leaf_nodes=13)

[55]  Dtree3_2.fit(X_train,y_train)
```

```
er(max_leaf_nodes=13)
```

DecisionTreeClassifier

DecisionTreeClassifier(max_leaf_nodes=13)

```
_, ax = plt.subplots(figsize=(15,10))
tree.plot_tree(Dtree3_2);
```




```
y_predict3_2 = Dtree3_2.predict(X_test)

accuracy_score(y_test, y_predict3_2)
```

```
0.8387978142076503
```

MAX LEAF_NODES

```
from sklearn.tree import DecisionTreeClassifier
from sklearn.model_selection import GridSearchCV

Dtreee= DecisionTreeClassifier()

param_grid = {'max_leaf_nodes': [2,5,6, 13,]}

grid_search = GridSearchCV(Dtreee, param_grid, cv=11)
```

```
grid_search.fit(X_train, y_train)
```

GridSearchCV

estimator: DecisionTreeClassifier

DecisionTreeClassifier

DecisionTreeClassifier()

```
best_max_leaf_nodes = grid_search.best_params_['max_leaf_nodes']

best_max_leaf_nodes
```

```
2
```

TW Industries

THANK YOU SO MUCH

Classification

กลุ่มกลุ้ม

หอยหลอดดกรุ๊ป