12/25/23, 6:30 AM Decision_tree - Colaboratory

```
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline
from sklearn.datasets import load_iris
iris=load_iris()
iris
          [6.3, 2.9, 5.6, 1.8],
          [6.5, 3., 5.8, 2.2],
          [7.6, 3., 6.6, 2.1],
          [4.9, 2.5, 4.5, 1.7],
          [7.3, 2.9, 6.3, 1.8],
          [6.7, 2.5, 5.8, 1.8],
          [7.2, 3.6, 6.1, 2.5],
          [6.5, 3.2, 5.1, 2.],
          [6.4, 2.7, 5.3, 1.9],
          [6.8, 3., 5.5, 2.1],
          [5.7, 2.5, 5., 2.],
          [5.8, 2.8, 5.1, 2.4],
          [6.4, 3.2, 5.3, 2.3],
          [6.5, 3., 5.5, 1.8],
          [7.7, 3.8, 6.7, 2.2],
          [7.7, 2.6, 6.9, 2.3],
          [6., 2.2, 5., 1.5],
          [6.9, 3.2, 5.7, 2.3],
          [5.6, 2.8, 4.9, 2.],
          [7.7, 2.8, 6.7, 2.],
          [6.3, 2.7, 4.9, 1.8],
          [6.7, 3.3, 5.7, 2.1],
          [7.2, 3.2, 6., 1.8],
          [6.2, 2.8, 4.8, 1.8],
          [6.1, 3., 4.9, 1.8],
          [6.4, 2.8, 5.6, 2.1],
          [7.2, 3., 5.8, 1.6],
          [7.4, 2.8, 6.1, 1.9],
          [7.9, 3.8, 6.4, 2.],
          [6.4, 2.8, 5.6, 2.2],
          [6.3, 2.8, 5.1, 1.5],
          [6.1, 2.6, 5.6, 1.4],
          [7.7, 3., 6.1, 2.3],
          [6.3, 3.4, 5.6, 2.4],
          [6.4, 3.1, 5.5, 1.8],
          [6., 3., 4.8, 1.8],
          [6.9, 3.1, 5.4, 2.1],
          [6.7, 3.1, 5.6, 2.4],
          [6.9, 3.1, 5.1, 2.3],
          [5.8, 2.7, 5.1, 1.9],
          [6.8, 3.2, 5.9, 2.3],
          [6.7, 3.3, 5.7, 2.5],
          [6.7, 3., 5.2, 2.3],
          [6.3, 2.5, 5., 1.9],
          [6.5, 3., 5.2, 2.],
          [6.2, 3.4, 5.4, 2.3],
          [5.9, 3., 5.1, 1.8]]),
     1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,
          'frame': None,
     'target_names': array(['setosa', 'versicolor', 'virginica'], dtype='<U10'),</pre>
     'DESCR': '.. _iris_dataset:\n\nIris plants dataset\n-----\n\n**Data Set Characteristics:**\n\n :Number of Instances: 150 (50 in each of three classes)\n
                                                                                                                                               :Number of
    Attributes: 4 numeric, predictive attributes and the class\n :Attribute Information:\n

    sepal length in cm\n

                                                                                                                                petal length in cm\n

    sepal width in cm\n

    petal width in cm\n
                           - class:\n
                                                - Iris-Setosa∖n
                                                                                                     - Iris-Virginica\n
                                                                        Iris-Versicolour\n
                                                                                                                                \n :Summary
```

iris.data

```
Decision_tree - Colaboratory
            [6.9, 3.1, 5.1, 2.3],
            [5.8, 2.7, 5.1, 1.9],
            [6.8, 3.2, 5.9, 2.3],
            [6.7, 3.3, 5.7, 2.5],
            [6.7, 3., 5.2, 2.3],
            [6.3, 2.5, 5., 1.9],
            [6.5, 3., 5.2, 2.],
            [6.2, 3.4, 5.4, 2.3],
            [5.9, 3., 5.1, 1.8]])
iris.target
```

1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,

x=df.iloc[:,:-1]

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test=train_test_split(x,y,test_size=0.33,random_state=42)

X_train

y=df.iloc[:,-1:]

	sepal_length	sepal_width	petal_length	petal_width
96	5.7	2.9	4.2	1.3
105	7.6	3.0	6.6	2.1
66	5.6	3.0	4.5	1.5
0	5.1	3.5	1.4	0.2
122	7.7	2.8	6.7	2.0
71	6.1	2.8	4.0	1.3
106	4.9	2.5	4.5	1.7
14	5.8	4.0	1.2	0.2
92	5.8	2.6	4.0	1.2
102	7.1	3.0	5.9	2.1

100 rows × 4 columns

from sklearn.tree import DecisionTreeClassifier

model=DecisionTreeClassifier()

model.fit(X_train,y_train)

▼ DecisionTreeClassifier DecisionTreeClassifier()

from sklearn import tree

plt.figure(figsize=(15,10)) tree.plot_tree(model,filled=True) 12/25/23, 6:30 AM

```
Decision tree - Colaboratory
              Text(0.458333333333333, 0.7857142857142857, 'gini = 0.0\nsamples = 31\nvalue = [31, 0, 0]'),
               Text(0.625, 0.7857142857142857, 'x[3] \leftarrow 1.75 \cdot ngini = 0.5 \cdot nsamples = 69 \cdot nvalue = [0, 35, 34]'),
               Text(0.416666666666667, 0.6428571428571429, 'x[2] <= 5.35 \setminus ngini = 0.188 \setminus nsamples = 38 \setminus nvalue = [0, 34, 4]'),
               Text(0.3333333333333333, 0.5, x[3] <= 1.65 \cdot ngini = 0.105 \cdot nsamples = 36 \cdot nvalue = [0, 34, 2]'),
               Text(0.08333333333333333, 0.21428571428571427, 'gini = 0.0\nsamples = 32\nvalue = [0, 32, 0]'),
               Text(0.25, 0.21428571428571427, 'x[1] \le 2.45 \cdot s = 0.5 \cdot s = 2 \cdot v = [0, 1, 1]'),
               Text(0.5, 0.35714285714285715, 'x[1] <= 2.75 \setminus = 0.5 \setminus = 2 \setminus = 2 \setminus = 0.5 \setminus = 2 \setminus = 0.5 \setminus = 0
               Text(0.41666666666667. 0.21428571428571427. 'gini = 0.0\nsamples = 1\nvalue = \[ \text{0.0.1} \] ').
y_pred=model.predict(X_test)
               Text(0.83333333333334. 0.6428571428571429. 'x[2] <= 4.85 | ngini = 0.062 | nsamples = 31 | nvalue = [0. 1. 30]'
from sklearn.metrics import accuracy_score,classification_report
               Text(0 8333333333333334. 0 35714285714285715. 'gini = 0 0\nsamnles = 2 \cdot nvalue = [0.0.21']
score=accuracy_score(y_pred,y_test)
score
             0.98
print(classification_report(y_pred,y_test))
                                               precision
                                                                               recall f1-score support
                                                          1.00
                                                                                   1.00
                                                                                                           1.00
                                                                                                                                         19
                           setosa
                  versicolor
                                                          1.00
                                                                                    0.94
                                                                                                            0.97
                                                                                                                                         16
                                                                                                                                         15
                    virginica
                                                           0.94
                                                                                   1.00
                                                                                                            0.97
                                                                                                                                         50
                      accuracy
                                                                                                            0.98
                                                           0.98
                                                                                    0.98
                                                                                                            0.98
                                                                                                                                         50
                    macro avg
             weighted avg
                                                           0.98
                                                                                    0.98
                                                                                                            0.98
                                                                                                                                          50
                                                                                                                         x[3] <= 1.65
                                                                                                                                                                                                                                                                                   x[0] \le 5.95
                                                                                                                                                                                           gini = 0.0
                                                                                                                                                                                                                                                                                                                                                      gini = 0.0
                                                                                                                                                                                                                                                                                    gini = 0.444
                                                                                                                          gini = 0.105
                                                                                                                                                                                                                                                                                                                                               samples = 28
                                                                                                                                                                                       samples = 2
                                                                                                                       samples = 36
                                                                                                                                                                                                                                                                                    samples = 3
                                                                                                                                                                                   value = [0, 0, 2]
                                                                                                                                                                                                                                                                                                                                           value = [0, 0, 28]
                                                                                                                   value = [0, 34, 2]
                                                                                                                                                                                                                                                                               value = [0, 1, 2]
                                                           x[2] <= 4.95
                                                                                                                                                                                      x[1] \le 2.75
                                                                                                                                                                                                                                                         gini = 0.0
                                                                                                                                                                                                                                                                                                                       gini = 0.0
                                                            gini = 0.057
                                                                                                                                                                                           gini = 0.5
                                                                                                                                                                                                                                                     samples = 1
                                                                                                                                                                                                                                                                                                                  samples = 2
                                                          samples = 34
                                                                                                                                                                                       samples = 2
                                                                                                                                                                                                                                                 value = [0, 1, 0]
                                                                                                                                                                                                                                                                                                              value = [0, 0, 2]
                                                     value = [0, 33, 1]
                                                                                                                                                                                  value = [0, 1, 1]
                                                                                           A
                                                                                         x[1] \le 2.45
                                 gini = 0.0
                                                                                                                                                                                                                          gini = 0.0
                                                                                                                                                             gini = 0.0
                                                                                              gini = 0.5
                           samples = 32
                                                                                                                                                        samples = 1
                                                                                                                                                                                                                      samples = 1
                                                                                           samples = 2
                       value = [0, 32, 0]
                                                                                                                                                   value = [0, 0, 1]
                                                                                                                                                                                                                 value = [0, 1, 0]
                                                                                     value = [0, 1, 1]
```