```
导入模块
In [ ]: !!pip install sklearn
                 Requirement already satisfied: sklearn in /usr/local/lib/python3.7/dist-packages (0.0)
                 Requirement already satisfied: scikit-learn in /usr/local/lib/python3.7/dist-packages (from s
                 klearn) (1.0.2)
                 Requirement already satisfied: threadpoolctl>=2.0.0 in /usr/local/lib/python3.7/dist-packages
                  (from scikit-learn->sklearn) (3.1.0)
                 Requirement already satisfied: joblib>=0.11 in /usr/local/lib/python3.7/dist-packages (from s
                 cikit-learn->sklearn) (1.1.0)
                 Requirement already satisfied: numpy>=1.14.6 in /usr/local/lib/python3.7/dist-packages (from
                 scikit-learn->sklearn) (1.19.5)
                 Requirement already satisfied: scipy>=1.1.0 in /usr/local/lib/python3.7/dist-packages (from s
                 cikit-learn->sklearn) (1.4.1)
In [ ]: from sklearn import tree
                  from sklearn.datasets import load_wine
                  from sklearn.model_selection import train_test_split
                  探索数据
In [ ]: wine = load_wine()
In [ ]: wine.data.shape
Out[]: (178, 13)
In [ ]: wine.target
0, 0, 0, 0, 0, 0, 0,
                                                                            Θ,
                                                                                   0, 0, 0, 0, 0, 0, 0, 0, 0,
                                                                                                                                          Θ,
                                                                                                                                                 Θ,
                                0, 0, 0, 0, 0,
                                                                            Θ,
                                                               Θ,
                                                                     Θ,
                                                                                   Θ,
                                                                                         0, 0, 0,
                                                                                                           Θ,
                                                                                                                  Θ,
                                                                                                                         Θ,
                                                                                                                               1,
                                                                                                                                     1,
                                                                                                                                           1,
                                                                      1,
                                                                            1,
                                                                                   1,
                                                                                         1,
                                                                                               1,
                                                                                                     1,
                                                                                                            1,
                                                                                                                  1,
                                                  1,
                                                         1,
                                                               1,
                                                                                                                         1,
                                                                                                                              1,
                                                                                                                                    1,
                                                                                                                                           1,
                                                                                                                                    1,
                                     1, 1, 1,
                                                        1,
                                                               1,
                                                                      1,
                                                                            1,
                                                                                  1,
                                                                                         1,
                                                                                              1,
                                                                                                     1,
                                                                                                            1,
                                                                                                                  1, 1,
                                                                                                                              1,
                                                                                                                                           1,
                                     2, 2])
In [ ]:
                 import pandas as pd
                  pd.concat([pd.DataFrame(wine.data), pd.DataFrame(wine.target)], axis=1)
Out[]:
                                 0
                                          1
                                                   2
                                                            3
                                                                       4
                                                                                5
                                                                                                  7
                                                                                                           8
                                                                                                                              10
                                                                                                                                       11
                                                                                                                                                   12
                                                                                                                                                          0
                                                                                         6
                                                                                                                      9
                      0 14.23 1.71 2.43 15.6 127.0 2.80 3.06 0.28 2.29
                                                                                                                  5.64 1.04 3.92 1065.0
                      1 13.20 1.78 2.14 11.2 100.0 2.65 2.76 0.26 1.28
                                                                                                                  4.38 1.05 3.40 1050.0
                      2 13.16 2.36 2.67
                                                      18.6 101.0
                                                                           2.80
                                                                                    3.24
                                                                                              0.30 2.81
                                                                                                                  5.68 1.03
                                                                                                                                    3.17 1185.0
                      3 14.37 1.95 2.50
                                                       16.8 113.0 3.85 3.49
                                                                                                                                  3.45 1480.0
                                                                                            0.24 2.18
                                                                                                                  7.80 0.86
                      4 13.24 2.59 2.87
                                                      21.0
                                                                118.0 2.80 2.69
                                                                                              0.39 1.82
                                                                                                                  4.32 1.04
                                                                                                                                    2.93
                                                                                                                                               735.0
                                                                   95.0 1.68 0.61 0.52 1.06
                   173 13.71 5.65 2.45 20.5
                                                                                                                  7.70 0.64 1.74
                                                                                                                                               740.0
                                                                                                                                                          2
                                                       23.0 102.0 1.80
                                                                                                                                                          2
                   174 13.40 3.91 2.48
                                                                                    0.75 0.43 1.41
                                                                                                                  7.30 0.70
                                                                                                                                   1.56
                                                                                                                                               750.0
                   175 13.27 4.28 2.26
                                                                                                                10.20 0.59
                                                       20.0 120.0 1.59 0.69
                                                                                              0.43 1.35
                                                                                                                                  1.56
                                                                                                                                               835.0
                                                                                                                                                          2
                   176 13.17 2.59 2.37
                                                      20.0 120.0 1.65 0.68
                                                                                            0.53 1.46
                                                                                                                  9.30 0.60 1.62
                                                                                                                                               840.0
                                                                                                                                                          2
                   177 14.13 4.10 2.74 24.5
                                                                   96.0 2.05 0.76 0.56 1.35
                                                                                                                  9.20 0.61 1.60
                                                                                                                                               560.0
                  1/8 rows × 14 columns
In [ ]: wine.feature_names
Out[]: ['alcohol',
                    'malic_acid',
                    'ash',
                    'alcalinity_of_ash',
                    'magnesium',
                    'total_phenols',
                    'flavanoids',
                    'nonflavanoid_phenols',
                    'proanthocyanins',
                    'color_intensity',
                    'hue',
                    'od280/od315_of_diluted_wines',
                    'proline']
In [ ]: wine.target_names
Out[]: array(['class_0', 'class_1', 'class_2'], dtype='<U7')</pre>
                 分训练集和测试集
In [ ]: Xtrain, Xtest, Ytrain, Ytest = train_test_split(wine.data, wine.target, test_size=0.3)
                 建立模型
In [ ]: clf = tree.DecisionTreeClassifier(criterion="gini", random_state=20, splitter="best")
                  clf = clf.fit(Xtrain, Ytrain)
                  score = clf.score(Xtest, Ytest)
                  score
Out[]: 0.9259259259259259
                  绘图
In [ ]: tree.plot_tree(clf, filled = True, rounded = True)
Out[]: [Text(0.4230769230769231, 0.9, 'X[9] \le 3.945 \cdot 1000 = 0.663 \cdot 1000 = 124 \cdot
                 37]'),
                   Text(0.15384615384615385, 0.7, 'X[6] \le 0.52 = 0.126 = 45 = 45 = [2, 42, 42]
                 1]'),
                   Text(0.07692307692307693, 0.5, 'gini = 0.0 \nsamples = 1 \nvalue = [0, 0, 1]'),
                   Text(0.23076923076923078, 0.5, 'X[12] \le 953.0 = 0.087 = 44 = [2, 42, 42]
                 0]'),
                   Text(0.15384615384615385, 0.3, 'X[11] \le 3.73 \cdot gini = 0.045 \cdot gles = 43 \cdot gles = [1, 42, 6]
                 0]'),
                    Text(0.07692307692307693, 0.1, 'gini = 0.0 \nsamples = 42 \nvalue = [0, 42, 0]'),
                   Text(0.23076923076923078, 0.1, 'gini = 0.0 \nsamples = 1 \nvalue = [1, 0, 0]'),
                   Text(0.3076923076923077, 0.3, 'gini = 0.0 \nsamples = 1 \nvalue = [1, 0, 0]'),
                    Text(0.6923076923076923, 0.7, 'X[6] \le 1.58 \cdot gini = 0.557 \cdot samples = 79 \cdot nvalue = [38, 5, 3]
                  6]'),
                   Text(0.5384615384615384, 0.5, 'X[2] \le 2.06 \cdot gini = 0.053 \cdot samples = 37 \cdot nvalue = [0, 1, 3]
                 6]'),
                   Text(0.46153846153846156, 0.3, 'gini = 0.0 \nsamples = 1 \nvalue = [0, 1, 0]'),
                   Text(0.6153846153846154, 0.3, 'gini = 0.0 \nsamples = 36 \nvalue = [0, 0, 36]'),
                   Text(0.8461538461538461, 0.5, 'X[12] \le 597.5 \cdot end{0.172 \cdot end{0
                 0]'),
                    Text(0.7692307692307693, 0.3, 'gini = 0.0 \nsamples = 4 \nvalue = [0, 4, 0]'),
                   Text(0.9230769230769231, 0.3, 'gini = 0.0 \nsamples = 38 \nvalue = [38, 0, 0]')]
                                                X[9] <= 3.945
gini = 0.663
samples = 124
                                                                          X[6] <= 1.58
gini = 0.557
In [ ]:
                 import graphviz
                  dot_data = tree.export_graphviz(clf
                                                                                         out_file = None
                                                                                         filled = True
                                                                                         rounded = True)
                  graphviz.Source(dot_data)
Out[]:
                                                                                     X[9] <= 3.945
                                                                                      gini = 0.663
                                                                                    samples = 124
                                                                                 value = [40, 47, 37]
                                                                      X[6] \le 0.52
                                                                                                                   X[6] <= 1.58
                                                                      gini = 0.126
                                                                                                                    gini = 0.557
                                                                     samples = 45
                                                                                                                   samples = 79
                                                                   value = [2, 42, 1]
                                                                                                               value = [38, 5, 36]
                                                                    X[12] <= 953.0
gini = 0.087
                                                                                                                    X[2] <= 2.06
                                                                                                                                                               X[12] <= 597.5
                                         gini = 0.0
                                                                                                                    gini = 0.053
                                                                                                                                                                 gini = 0.172
                                       samples = 1
                                                                     samples = 44
                                                                                                                   samples = 37
                                                                                                                                                                samples = 42
                                     value = [0, 0, 1]
                                                                   value = [2, 42, 0]
                                                                                                                                                               value = [38, 4, 0]
                                                                                                                 value = [0, 1, 36]
                                      X[11] <= 3.73
                                                                                                     gini = 0.0
                                                                       gini = 0.0
                                                                                                                                    gini = 0.0
                                                                                                                                                                   gini = 0.0
                                                                                                                                                                                                  gini = 0.0
                                       gini = 0.045
                                                                                                                                                                                              samples = 38
                                                                      samples = 1
                                                                                                   samples = 1
                                                                                                                                 samples = 36
                                                                                                                                                                 samples = 4
                                      samples = 43
                                                                    value = [1, 0, 0]
                                                                                                                                                                                             value = [38, 0, 0]
                                                                                                 value = [0, 1, 0]
                                                                                                                               value = [0, 0, 36]
                                                                                                                                                               value = [0, 4, 0]
                                    value = [1, 42, 0]
                                                        gini = 0.0
                          gini = 0.0
                       samples = 42
                                                       samples = 1
                     value = [0, 42, 0]
                                                    value = [1, 0, 0]
                 探索决策树
In [ ]: clf.feature_importances_
                                                                               , 0.02366124, 0.
Out[ ]: array([0.
                                                                                                                                  , 0.39582053,
                                Θ.
                                                         0.44608223, 0.
                                                                                                        , 0.
                                Θ.
                                                         0.02375295, 0.11068306])
```

In []: [*zip(wine.feature_names, clf.feature_importances_)]

Out[]: [('alcohol', 0.0),

('malic_acid', 0.0),

('magnesium', 0.0), ('total_phenols', 0.0),

('hue', 0.0),

('ash', 0.023661237232525747), ('alcalinity_of_ash', 0.0),

('nonflavanoid_phenols', 0.0),

('proline', 0.1106830551910983)]

('proanthocyanins', 0.0),

('flavanoids', 0.44608222886316634),

('color_intensity', 0.3958205312588524),

('od280/od315_of_diluted_wines', 0.02375294745435724),