

```

import pandas as pd
from sklearn.tree import DecisionTreeClassifier
from sklearn.model_selection import train_test_split

df=pd.read_csv('/content/drive/MyDrive/Iris.csv')

df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 150 entries, 0 to 149
Data columns (total 6 columns):
 #   Column      Non-Null Count  Dtype  
--- 
 0   Id          150 non-null    int64  
 1   SepalLengthCm 150 non-null    float64 
 2   SepalWidthCm  150 non-null    float64 
 3   PetalLengthCm 150 non-null    float64 
 4   PetalWidthCm  150 non-null    float64 
 5   Species      150 non-null    object  
dtypes: float64(4), int64(1), object(1)
memory usage: 7.2+ KB

df.head()

{"summary": "{\n  \"name\": \"df\",\n  \"rows\": 150,\n  \"fields\": [\n    {\n      \"column\": \"Id\",\n      \"properties\": {\n        \"dtype\": \"number\",\n        \"std\": 43,\n        \"min\": 1,\n        \"max\": 150,\n        \"num_unique_values\": 150,\n        \"samples\": [\n          74,\n          19,\n          119\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n      },\n      \"column\": \"SepalLengthCm\",\n      \"properties\": {\n        \"dtype\": \"float64\",\n        \"number\": 43,\n        \"std\": 0.8280661279778629,\n        \"min\": 4.3,\n        \"max\": 7.9,\n        \"num_unique_values\": 35,\n        \"samples\": [\n          6.2,\n          4.5,\n          5.6\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n      },\n      \"column\": \"SepalWidthCm\",\n      \"properties\": {\n        \"dtype\": \"number\",\n        \"std\": 0.4335943113621737,\n        \"min\": 2.0,\n        \"max\": 4.4,\n        \"num_unique_values\": 23,\n        \"samples\": [\n          2.3,\n          4.0,\n          3.5\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n      },\n      \"column\": \"PetalLengthCm\",\n      \"properties\": {\n        \"dtype\": \"float64\",\n        \"number\": 43,\n        \"std\": 1.7644204199522617,\n        \"min\": 1.0,\n        \"max\": 6.9,\n        \"num_unique_values\": 43,\n        \"samples\": [\n          6.7,\n          3.8,\n          3.7\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n      },\n      \"column\": \"PetalWidthCm\",\n      \"properties\": {\n        \"dtype\": \"number\",\n        \"std\": 0.7631607417008414,\n        \"min\": 0.1,\n        \"max\": 2.5,\n        \"samples\": [\n          1.7,\n          1.0,\n          0.9\n        ],\n        \"semantic_type\": \"\",\n        \"description\": \"\"\n      }\n    }\n  ]\n}
```

```

    "num_unique_values": 22, "samples": [0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2], "semantic_type": "\\", "column": "Species", "properties": {"dtype": "category", "num_unique_values": 3, "samples": [{"Iris-setosa", "Iris-versicolor", "Iris-virginica"}]}, "description": "\n\n"}, "variable_name": "df"}}

from sklearn.preprocessing import LabelEncoder
LE=LabelEncoder()
df["Species"]=LE.fit_transform(df["Species"])

x_train,x_test,y_train,y_test=train_test_split(
    df.drop("Species",axis=1),
    df["Species"],
    test_size=0.2
)

Dt=DecisionTreeClassifier()
Dt.fit(x_train,y_train)

DecisionTreeClassifier()

from sklearn.tree import plot_tree
plot_tree(Dt,filled=True,max_depth=5)

[Text(0.6, 0.8333333333333334, 'x[0] <= 100.5\ngini = 0.665\nsamples = 120\nvalue = [37, 39, 44]'), Text(0.4, 0.5, 'x[4] <= 0.7\ngini = 0.5\nsamples = 76\nvalue = [37, 39, 0]'), Text(0.5, 0.6666666666666667, 'True '),
Text(0.2, 0.1666666666666666, 'gini = 0.0\nsamples = 37\nvalue = [37, 0, 0]'), Text(0.6, 0.1666666666666666, 'gini = 0.0\nsamples = 39\nvalue = [0, 39, 0]'), Text(0.8, 0.5, 'gini = 0.0\nsamples = 44\nvalue = [0, 0, 44]'), Text(0.7, 0.6666666666666667, ' False')]
```

