

```
In [1]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.metrics import accuracy_score, f1_score, precision_score, recall_score
from sklearn.tree import DecisionTreeClassifier
```

```
In [11]: df = sns.load_dataset("iris")
```

```
In [12]: df.head(2)
```

```
Out[12]:
```

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	setosa
1	4.9	3.0	1.4	0.2	setosa

```
In [13]: df.describe()
```

```
Out[13]:
```

	sepal_length	sepal_width	petal_length	petal_width
count	150.000000	150.000000	150.000000	150.000000
mean	5.843333	3.057333	3.758000	1.199333
std	0.828066	0.435866	1.765298	0.762238
min	4.300000	2.000000	1.000000	0.100000
25%	5.100000	2.800000	1.600000	0.300000
50%	5.800000	3.000000	4.350000	1.300000
75%	6.400000	3.300000	5.100000	1.800000
max	7.900000	4.400000	6.900000	2.500000

```
In [14]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 150 entries, 0 to 149
Data columns (total 5 columns):
#   Column          Non-Null Count  Dtype
---  -
0   sepal_length    150 non-null   float64
1   sepal_width     150 non-null   float64
2   petal_length    150 non-null   float64
3   petal_width     150 non-null   float64
4   species         150 non-null   object
dtypes: float64(4), object(1)
memory usage: 6.0+ KB
```

```
In [15]: df.shape
```

```
Out[15]: (150, 5)
```

```
In [17]: X = df.iloc[:,0:4]
        Y = df.iloc[:, -1]
```

```
In [18]: X
```

```
Out[18]:
```

	sepal_length	sepal_width	petal_length	petal_width
0	5.1	3.5	1.4	0.2
1	4.9	3.0	1.4	0.2
2	4.7	3.2	1.3	0.2
3	4.6	3.1	1.5	0.2
4	5.0	3.6	1.4	0.2
...
145	6.7	3.0	5.2	2.3
146	6.3	2.5	5.0	1.9
147	6.5	3.0	5.2	2.0
148	6.2	3.4	5.4	2.3
149	5.9	3.0	5.1	1.8

150 rows × 4 columns

```
In [19]: Y
```

```
Out[19]: 0      setosa
        1      setosa
        2      setosa
        3      setosa
        4      setosa
        ...
        145    virginica
        146    virginica
        147    virginica
        148    virginica
        149    virginica
        Name: species, Length: 150, dtype: object
```

```
In [20]: #train test split
```

```
In [21]: from sklearn.model_selection import train_test_split
```

```
In [22]: X_train,X_test,Y_train,Y_test = train_test_split(X , Y,test_size = 0.33,random
```

```
In [23]: from sklearn.tree import DecisionTreeClassifier
```

```
In [24]: #postpruning  
treemodel = DecisionTreeClassifier()
```

```
In [25]: treemodel.fit(X_train,Y_train)
```

Out[25]: DecisionTreeClassifier()

In a Jupyter environment, please rerun this cell to show the HTML representation or trust the notebook.

On GitHub, the HTML representation is unable to render, please try loading this page with nbviewer.org.

```
In [ ]:
```