

Muhammad Arham Adeel

Decision Tree Classifier

Practice and task given in the video on phool dataset.

```
In [1]: import pandas as pd
import seaborn as sns
```

```
In [2]: df = sns.load_dataset("iris")
df.head()
```

```
Out[2]:
```

	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
2	4.7	3.2	1.3	0.2	setosa
3	4.6	3.1	1.5	0.2	setosa
4	5.0	3.6	1.4	0.2	setosa

```
In [3]: import matplotlib.pyplot as plt
from sklearn.tree import DecisionTreeClassifier

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

```
In [4]: x.head()
```

```
Out[4]:
```

	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

```
In [5]: y.head()
```

```
Out[5]:
```

	species
0	setosa
1	setosa
2	setosa

species

3 setosa

4 setosa

In [9]:

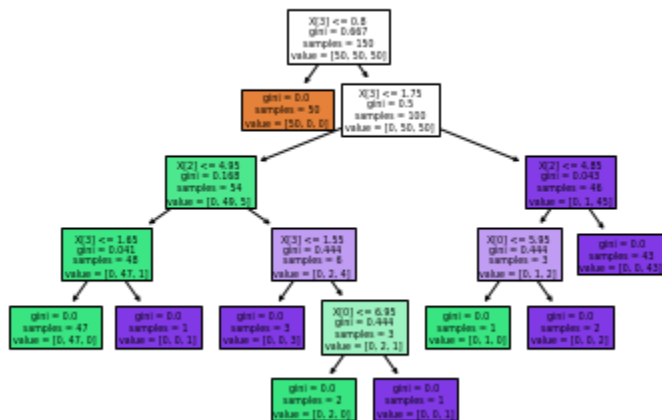
```
from sklearn.tree import DecisionTreeClassifier
from sklearn.tree import plot_tree

model = DecisionTreeClassifier().fit(x,y)
model

plot_tree(model, filled = True)
plt.title("Decision Tree Trained model of IRIS dataset")

# how to save this plot in tif, png and PDF files in HD quality?
# # plt.savefig("DecisionTree.png", dpi = 300)
# plt.savefig("tiff_compressed.tiff", dpi = 600, format = "tiff", facecolor = "white",
# #         edgcolor = None, pil_kwargs = {"compression": "tiff_lzw"})
plt.show()
```

Decision Tree Trained model of IRIS dataset



In [10]:

```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(x, y, test_size=0.2, random_state=0) #here i can s
```

In [11]:

```
reg = DecisionTreeClassifier().fit(x,y)
reg
```

Out[11]:

DecisionTreeClassifier()

In [12]:

```
model.fit(X_train, y_train)
predicted_values = model.predict(X_test)
```

In [13]:

predicted_values

Out[13]:

```
array(['virginica', 'versicolor', 'setosa', 'virginica', 'setosa',
       'virginica', 'setosa', 'versicolor', 'versicolor', 'versicolor',
       'virginica', 'versicolor', 'versicolor', 'versicolor',
       'versicolor', 'setosa', 'versicolor', 'versicolor', 'setosa',
       'setosa', 'virginica', 'versicolor', 'setosa', 'setosa',
```

```
'virginica', 'setosa', 'setosa', 'versicolor', 'versicolor',  
'setosa'], dtype=object)
```

```
In [14]: #Checking Score  
  
from sklearn.metrics import accuracy_score  
score = accuracy_score(y_test,predicted_values)  
score
```

```
Out[14]: 1.0
```

here i can check the accuracy scor of data split in to 70/30

```
In [15]: from sklearn.model_selection import train_test_split  
X_train,X_test,y_train,y_test = train_test_split(x,y, test_size=0.3, random_state=0) #here i can s
```

```
In [16]: reg = DecisionTreeClassifier().fit(x,y)  
reg
```

```
Out[16]: DecisionTreeClassifier()
```

```
In [17]: model.fit(X_train,y_train)  
predicted_values = model.predict(X_test)  
predicted_values
```

```
Out[17]: array(['virginica', 'versicolor', 'setosa', 'virginica', 'setosa',  
                'virginica', 'setosa', 'versicolor', 'versicolor', 'versicolor',  
                'virginica', 'versicolor', 'versicolor', 'versicolor',  
                'versicolor', 'setosa', 'versicolor', 'versicolor', 'setosa',  
                'setosa', 'virginica', 'versicolor', 'setosa', 'setosa',  
                'virginica', 'setosa', 'setosa', 'versicolor', 'versicolor',  
                'setosa', 'virginica', 'versicolor', 'setosa', 'virginica',  
                'virginica', 'versicolor', 'setosa', 'virginica', 'versicolor',  
                'versicolor', 'virginica', 'setosa', 'virginica', 'setosa',  
                'setosa'], dtype=object)
```

```
In [18]: #Checking Score  
  
from sklearn.metrics import accuracy_score  
score = accuracy_score(y_test,predicted_values)  
score
```

```
Out[18]: 0.9777777777777777
```

here i can check the accuracy of the model 90/10

```
In [19]: from sklearn.model_selection import train_test_split  
X_train,X_test,y_train,y_test = train_test_split(x,y, test_size=0.1, random_state=0) #here i can s
```

```
In [20]: model.fit(X_train,y_train)  
predicted_values = model.predict(X_test)  
predicted_values
```

```
Out[20]: array(['virginica', 'versicolor', 'setosa', 'virginica', 'setosa',  
        'virginica', 'setosa', 'versicolor', 'versicolor', 'versicolor',  
        'virginica', 'versicolor', 'versicolor', 'versicolor',  
        'versicolor'], dtype=object)
```

```
In [21]: #Checking Score  
  
from sklearn.metrics import accuracy_score  
score = accuracy_score(y_test,predicted_values)  
score
```

```
Out[21]: 1.0
```

here i can get the prediction values using this trained model

```
In [22]: from sklearn.tree import DecisionTreeClassifier  
  
        # Create and Fit our model  
  
        model = DecisionTreeClassifier().fit(x,y)  
        model
```

```
Out[22]: DecisionTreeClassifier()
```

```
In [23]: model.predict([[5.1,3.5,1.4,0.2]])
```

```
Out[23]: array(['setosa'], dtype=object)
```

```
In [24]: model.predict([[3.6,3.1,2.4,0.3]])
```

```
Out[24]: array(['setosa'], dtype=object)
```

```
In [ ]:
```