

▼ Knowledge Discovery & Data Mining Lab-07

Name: Gurvinder Kaur Matharu


PRN: 20190802077

▼ AIM:

To implement Decision Tree using pyrhon.

```
import pandas as pd
from sklearn.tree import DecisionTreeClassifier
from sklearn.preprocessing import LabelEncoder
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score
from sklearn import tree
```

```
data = pd.read_csv('weather.csv')
data.head()
```

	Outlook	Temperature	Humidity	Wind	Played tennis	
0	Sunny	Hot	High	Weak	No	
1	Sunny	Hot	High	Strong	No	
2	Overcast	Hot	High	Weak	Yes	
3	Rain	Mild	High	Weak	Yes	
4	Rain	Cool	Normal	Weak	Yes	

```
data.shape
```

```
(14, 5)
```

```
data.isnull().sum()
```

```
Outlook      0
Temperature  0
Humidity      0
Wind          0
Played tennis 0
dtype: int64
```

```
X = data.iloc[:,0:4].values
y = data.iloc[:,4].values
```

```
labelencoder= LabelEncoder()
X[:,0] = labelencoder.fit_transform(X[:,0])
X[:,1] = labelencoder.fit_transform(X[:,1])
X[:,2] = labelencoder.fit_transform(X[:,2])
X[:,3] = labelencoder.fit_transform(X[:,3])
y = labelencoder.fit_transform(y)
```

```
X
```

```
array([[2, 1, 0, 1],
       [2, 1, 0, 0],
       [0, 1, 0, 1],
       [1, 2, 0, 1],
       [1, 0, 1, 1],
```

```
[1, 0, 1, 0],
[0, 0, 1, 0],
[2, 2, 0, 1],
[2, 0, 1, 1],
[1, 2, 1, 1],
[2, 2, 1, 0],
[0, 2, 0, 0],
[0, 1, 1, 1],
[1, 2, 0, 0]], dtype=object)
```

y

```
array([0, 0, 1, 1, 1, 0, 1, 0, 1, 1, 1, 1, 1, 0])
```

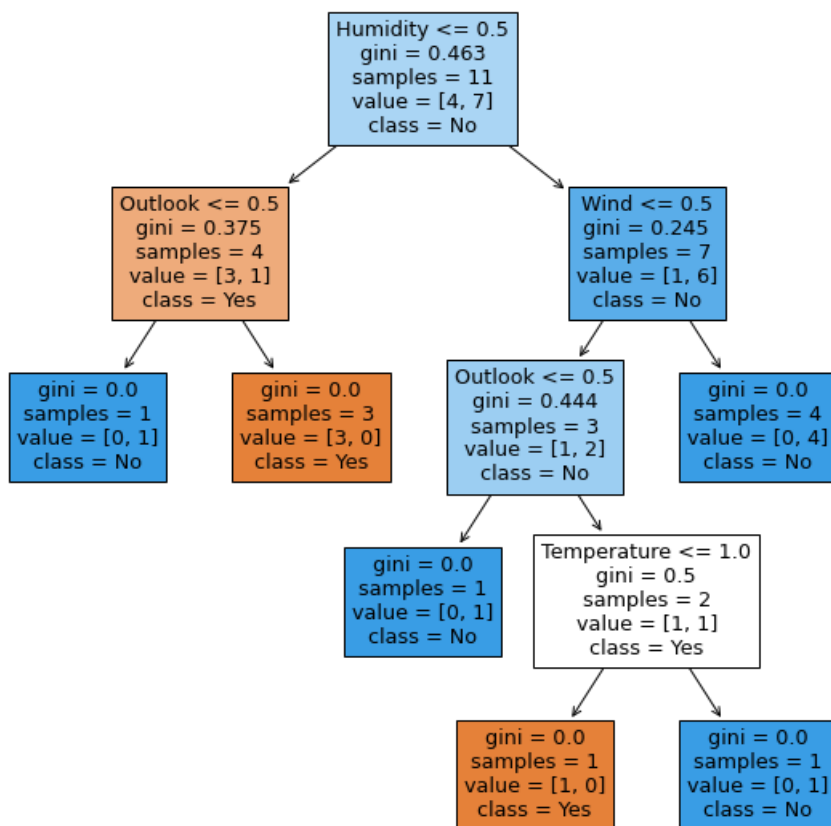
```
# splitting the dataset in training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2)
```

```
clf = DecisionTreeClassifier()
clf = clf.fit(X_train, y_train)
y_pred = clf.predict(X_test)
```

```
print("Accuracy:", accuracy_score(y_test, y_pred))
```

```
Accuracy: 0.6666666666666666
```

```
from matplotlib import pyplot as plt
fig = plt.figure(figsize=(10,10))
tree.plot_tree(clf,
                feature_names=['Outlook', 'Temperature', 'Humidity', 'Wind'],
                class_names=['Yes', 'No'],
                filled=True);
```



✓ 1s completed at 9:15 PM

