In [ ]:

```
import pandas as pd
```

In [ ]:

```
df = pd.read_csv("/content/weatherPrediction.csv")
df
```

Out[ ]:

	Outlook	Temperature	Humidity	Wind	Play cricket
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
5	Rain	Cool	Normal	Strong	No
6	Overcast	Cool	Normal	Strong	Yes
7	Sunny	Mild	High	Weak	No
8	Sunny	Cool	Normal	Weak	Yes
9	Rain	Mild	Normal	Weak	Yes
10	Sunny	Mild	Normal	Strong	Yes
11	Overcast	Mild	High	Strong	Yes
12	Overcast	Hot	Normal	Weak	Yes
13	Rain	Mild	High	Strong	No

In [ ]:

```
df[  ["Outlook","Temperature"] ]
```

Out[ ]:

	Outlook	Temperature
0	Sunny	Hot
1	Sunny	Hot
2	Overcast	Hot
3	Rain	Mild
4	Rain	Cool
5	Rain	Cool
6	Overcast	Cool
7	Sunny	Mild
8	Sunny	Cool
9	Rain	Mild
10	Sunny	Mild
11	Overcast	Mild
12	Overcast	Hot
13	Rain	Mild

In [ ]:

```
df["Outlook"] = df["Outlook"].map({"Sunny":0, "Overcast":1,"Rain":2})
df
```

Out[ ]:

	Outlook	Temperature	Humidity	Wind	Play cricket
0	0	Hot	High	Weak	No
1	0	Hot	High	Strong	No
2	1	Hot	High	Weak	Yes
3	2	Mild	High	Weak	Yes
4	2	Cool	Normal	Weak	Yes
5	2	Cool	Normal	Strong	No
6	1	Cool	Normal	Strong	Yes
7	0	Mild	High	Weak	No
8	0	Cool	Normal	Weak	Yes
9	2	Mild	Normal	Weak	Yes
10	0	Mild	Normal	Strong	Yes
11	1	Mild	High	Strong	Yes
12	1	Hot	Normal	Weak	Yes
13	2	Mild	High	Strong	No

In [ ]:

```
df["Temperature"] = df["Temperature"].map({"Cool":0,"Mild":1,"Hot":2})
df["Humidity"] = df["Humidity"].map({"Normal":0,"High":1})
df["Wind"] = df["Wind"].map({"Weak":0,"Strong":1})
df
```

Out[ ]:

	Outlook	Temperature	Humidity	Wind	Play cricket
0	0	2	1	0	No
1	0	2	1	1	No
2	1	2	1	0	Yes
3	2	1	1	0	Yes
4	2	0	0	0	Yes
5	2	0	0	1	No
6	1	0	0	1	Yes
7	0	1	1	0	No
8	0	0	0	0	Yes
9	2	1	0	0	Yes
10	0	1	0	1	Yes
11	1	1	1	1	Yes
12	1	2	0	0	Yes
13	2	1	1	1	No

In [ ]:

```
x_data = df[["Outlook","Temperature","Humidity","Wind"]].values
x_data
```

Out[ ]:

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

In [ ]:

```
y_data = df["Play cricket"].values
y_data
```

Out[ ]:

```
array(['No', 'No', 'Yes', 'Yes', 'Yes', 'No', 'Yes', 'No', 'Yes', 'Yes',
       'Yes', 'Yes', 'Yes', 'No'], dtype=object)
```

In [ ]:

```
from sklearn.tree import DecisionTreeClassifier
classifier = DecisionTreeClassifier(criterion="entropy")
classifier = classifier.fit(x_data,y_data)
```

In [ ]:

```
from sklearn import tree
import graphviz
dot_data = tree.export_graphviz(classifier, out_file=None,filled=True,
                                feature_names=["Outlook","Temperature","Humidity","Wind"
],
                                class_names=["No","Yes"])
graph = graphviz.Source(dot_data)
graph
```

Out[ ]:

In [ ]:

```
test_samples = [[0, 0, 0, 2]]
classifier.predict(test_samples)
```

Out[ ]:

```
array(['Yes'], dtype=object)
```

In [ ]:

```
#Using One-hot encoder
df = pd.read_csv("/content/weatherPrediction.csv")
df
```

Out[ ]:

	Outlook	Temperature	Humidity	Wind	Play cricket
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

	Outlook	Temperature	Humidity	Wind	Play cricket
5	Rain	Cool	Normal	Strong	No
6	Overcast	Cool	Normal	Strong	Yes
7	Sunny	Mild	High	Weak	No
8	Sunny	Cool	Normal	Weak	Yes
9	Rain	Mild	Normal	Weak	Yes
10	Sunny	Mild	Normal	Strong	Yes
11	Overcast	Mild	High	Strong	Yes
12	Overcast	Hot	Normal	Weak	Yes
13	Rain	Mild	High	Strong	No

In [ ]:

```
one_hot_data = pd.get_dummies(df[ ['Outlook', 'Temperature', 'Humidity', 'Wind'] ])
#print the new dummy data
one_hot_data
```

Out[ ]:

	Outlook_Overcast	Outlook_Rain	Outlook_Sunny	Temperature_Cool	Temperature_Hot	Temperature_Mild	Humidity_High	H
0	0	0	1	0	1	0	1	
1	0	0	1	0	1	0	1	
2	1	0	0	0	1	0	1	
3	0	1	0	0	0	1	1	
4	0	1	0	1	0	0	0	
5	0	1	0	1	0	0	0	
6	1	0	0	1	0	0	0	
7	0	0	1	0	0	1	1	
8	0	0	1	1	0	0	0	
9	0	1	0	0	0	1	0	
10	0	0	1	0	0	1	0	
11	1	0	0	0	0	1	1	
12	1	0	0	0	1	0	0	
13	0	1	0	0	0	1	1	

In [ ]:

```
x_data = one_hot_data.values
y_data = df["Play cricket"].values
```

In [ ]:

```
y_data
```

Out[ ]:

```
array(['No', 'No', 'Yes', 'Yes', 'Yes', 'No', 'Yes', 'No', 'Yes', 'Yes',
       'Yes', 'Yes', 'Yes', 'No'], dtype=object)
```

In [ ]:

```
from sklearn.tree import DecisionTreeClassifier
classifier = DecisionTreeClassifier(criterion="gini")
classifier = classifier.fit(x_data,y_data)
```

In [ ]:

```
from sklearn import tree
import graphviz
dot_data = tree.export_graphviz(classifier, out_file=None,filled=True,
                                feature_names=one_hot_data.columns.values,
                                class_names=["No","Yes"])
graph = graphviz.Source(dot_data)
graph
```

Out[ ]:

In [ ]:

```
```