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 [ ]:

```
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 [ ]:

```
from sklearn.model_selection import train_test_split
x_data_train, x_data_test, y_data_train, y_data_test = train_test_split(x_data, y_data,t
est_size=0.2,random_state=1 )
```

In [ ]:

```
print(len(x_data_train), len(x_data_test))
```

```
11 3
```

In [ ]:

```
from sklearn.tree import DecisionTreeClassifier
classifier = DecisionTreeClassifier(criterion="entropy",min_samples_split=6,max_depth=3)
classifier = classifier.fit(x_data_train,y_data_train)
```

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 [ ]:

```
classifier.score(x_data_test,y_data_test)
```

Out[ ]:

```
0.6666666666666666
```