

```
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
import matplotlib.pyplot as plt

data_train = pd.read_csv('play_tennis_train.csv')
data_test = pd.read_csv('play_tennis_test.csv')

from sklearn import preprocessing
le = preprocessing.LabelEncoder()

data_train['day'] = le.fit_transform(data_train['day'])
data_train['humidity'] = le.fit_transform(data_train['humidity'])
data_train['outlook'] = le.fit_transform(data_train['outlook'])
data_train['temp'] = le.fit_transform(data_train['temp'])
data_train['wind'] = le.fit_transform(data_train['wind'])
data_train['play'] = le.fit_transform(data_train['play'])

data_train.head()
```

```
↳
```

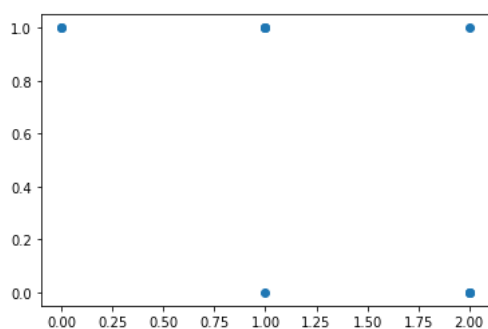
	day	outlook	temp	humidity	wind	play
0	0	2	1	0	1	0
1	2	2	1	0	0	0
2	3	0	1	0	1	1
3	4	1	2	0	1	1
4	5	1	0	1	1	1

```
data_test = data_test.apply(le.fit_transform)
data_test.head()
```

	day	outlook	temp	humidity	wind	play
0	0	2	1	1	0	1
1	1	0	1	0	0	1
2	2	0	0	1	1	1
3	3	1	1	0	0	0

```
x_train = data_train.drop('play',axis = 1)
y_train = data_train.play
```

```
plt.scatter(x_train.outlook,y_train)
plt.show()
```



```
plt.hist(x_train.outlook)
```

```
(array([2., 0., 0., 0., 0., 4., 0., 0., 0., 4.]),
 array([0., 0.2, 0.4, 0.6, 0.8, 1., 1.2, 1.4, 1.6, 1.8, 2. ]),
 <a list of 10 Patch objects>)
```



```
from sklearn import tree
clf = tree.DecisionTreeClassifier(criterion = 'entropy')
clf
```

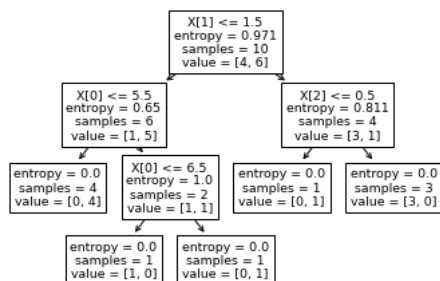
```
DecisionTreeClassifier(criterion='entropy')
~ |  |  |
```

```
clf = clf.fit(x_train, y_train)
clf
```

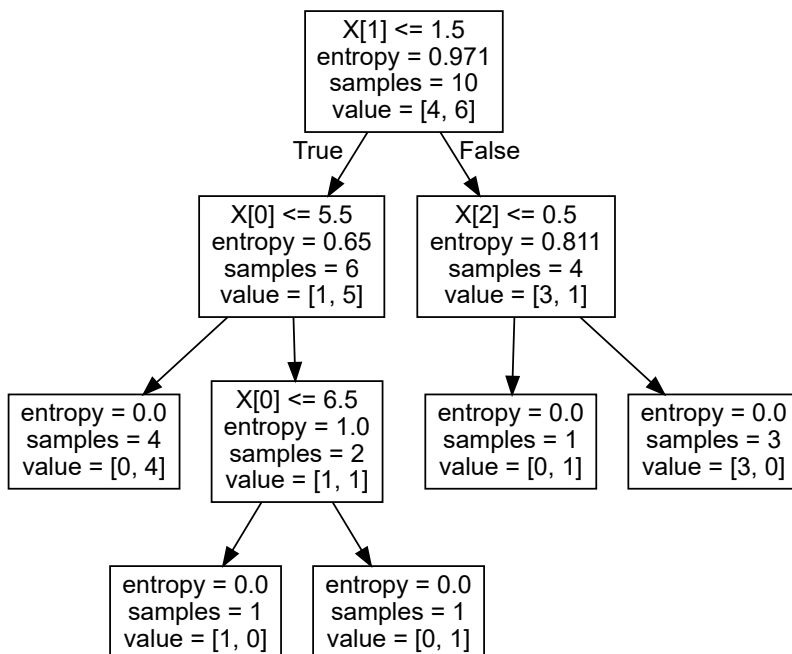
```
DecisionTreeClassifier(criterion='entropy')
```

```
tree.plot_tree(clf)
```

```
[Text(0.5, 0.875, 'X[1] <= 1.5\nentropy = 0.971\nsamples = 10\nvalue = [4, 6]'),
 Text(0.25, 0.625, 'X[0] <= 5.5\nentropy = 0.65\nsamples = 6\nvalue = [1, 5]'),
 Text(0.125, 0.375, 'entropy = 0.0\nsamples = 4\nvalue = [0, 4]'),
 Text(0.375, 0.375, 'X[0] <= 6.5\nentropy = 1.0\nsamples = 2\nvalue = [1, 1]'),
 Text(0.25, 0.125, 'entropy = 0.0\nsamples = 1\nvalue = [1, 0]'),
 Text(0.5, 0.125, 'entropy = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.75, 0.625, 'X[2] <= 0.5\nentropy = 0.811\nsamples = 4\nvalue = [3, 1]'),
 Text(0.625, 0.375, 'entropy = 0.0\nsamples = 1\nvalue = [0, 1]'),
 Text(0.875, 0.375, 'entropy = 0.0\nsamples = 3\nvalue = [3, 0]')]
```



```
import graphviz
dot_data = tree.export_graphviz(clf, out_file = None)
graph = graphviz.Source(dot_data)
graph
```



```
x_test = data_test.drop('play',axis=1)
y_test = data_test['play']

y_pred = clf.predict(x_test)
```

```
y_pred == y_test
```

```
0    False  
1     True  
2     True  
3    False  
Name: play, dtype: bool
```

[Colab paid products](#) - [Cancel contracts here](#)

