[2]:
```
import pandas as pd
import numpy as np
import os
import sys
```

[4]:
```
df = pd.read_csv("User_Data.csv")
```

[6]:
```
df
```

[6]:

	User ID	Gender	Age	EstimatedSalary	Purchased
0	15624510	Male	19	19000	0
1	15810944	Male	35	20000	0
2	15668575	Female	26	43000	0
3	15603246	Female	27	57000	0
4	15804002	Male	19	76000	0
...	...	...	...	...	...
395	15691863	Female	46	41000	1
396	15706071	Male	51	23000	1
397	15654296	Female	50	20000	1
398	15755018	Male	36	33000	0
399	15594041	Female	49	36000	1

400 rows × 5 columns

[8]:
```
x = df.iloc[:, [2,3]].values
y = df.iloc[:, 4].values
```

[10]:
```
from sklearn.model_selection import train_test_split
x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.25, random_state=0)
```

[11]:
```
#feature scaling
from sklearn.preprocessing import StandardScaler
sc = StandardScaler()
x_train = sc.fit_transform(x_train)
x_test = sc.transform(x_test)
```

[14]:
```
#fitting the classifier to the training set
from sklearn.tree import DecisionTreeClassifier
classifier = DecisionTreeClassifier(criterion='entropy', random_state=0)
```

[16]:
```
classifier.fit(x_train, y_train)
```

[16]:
```
DecisionTreeClassifier(criterion='entropy', random_state=0)
```

[18]:
```
y_pred = classifier.predict(x_test)
```

[20]:
```
from sklearn.metrics import confusion_matrix
cm = confusion_matrix(y_test, y_pred)
print(cm)
```

```
[[62  6]
 [ 3 29]]
```