

```
In [74]: import numpy as np
import pandas as pd
```

```
In [75]: df = pd.read_csv('customer.csv')
```

```
In [76]: df.sample(5)
```

```
Out[76]:
```

	age	gender	review	education	purchased
33	89	Female	Good	PG	Yes
10	98	Female	Good	UG	Yes
9	74	Male	Good	UG	Yes
48	39	Female	Good	UG	Yes
30	73	Male	Average	UG	No

```
In [77]: df = df.iloc[:,2:]
```

```
In [78]: df.head()
```

```
Out[78]:
```

	review	education	purchased
0	Average	School	No
1	Poor	UG	No
2	Good	PG	No
3	Good	PG	No
4	Average	UG	No

```
In [82]: from sklearn.preprocessing import OrdinalEncoder
```

```
In [89]: X_train
```

```
Out[89]:
```

	review	education
16	Poor	UG
44	Average	UG
13	Average	School
48	Good	UG
42	Good	PG
38	Good	School

17	Poor	UG
4	Average	UG
5	Average	School
24	Average	PG
47	Good	PG
32	Average	UG
37	Average	PG
26	Poor	PG
15	Poor	UG
41	Good	PG
30	Average	UG
21	Average	PG
9	Good	UG
20	Average	School
23	Good	School
12	Poor	School
33	Good	PG
14	Poor	PG
2	Good	PG
8	Average	UG
7	Poor	School
1	Poor	UG
36	Good	UG
46	Poor	PG
19	Poor	PG
18	Good	School
22	Poor	PG
29	Average	UG
49	Good	UG
28	Poor	School
43	Poor	PG
27	Poor	PG
10	Good	UG
11	Good	UG

```
In [90]: oe = OrdinalEncoder(categories=[['Poor', 'Average']
```

```
In [91]: oe.fit(X_train)
```

```
Out[91]: OrdinalEncoder(categories=[['Poor', 'Average', 'Good'], ['School', 'UG', 'PG']])
```

```
In [92]: X_train = oe.transform(X_train)
```

```
In [93]: X_train
```

```
Out[93]: array([[0., 1.],
 [1., 1.],
 [1., 0.],
 [2., 1.],
 [2., 2.],
 [2., 0.],
 [0., 1.],
 [1., 1.],
 [1., 0.],
 [1., 2.],
 [2., 2.],
 [1., 1.],
 [1., 2.],
 [0., 2.],
 [0., 1.],
 [2., 2.],
 [1., 1.],
 [1., 2.],
 [2., 1.],
 [1., 0.],
 [2., 0.],
 [0., 0.],
 [2., 2.],
 [0., 2.],
 [2., 2.],
 [1., 1.],
 [0., 0.],
 [0., 1.],
 [2., 1.],
 [0., 2.],
 [0., 2.],
 [2., 0.],
 [0., 2.],
 [1., 1.],
 [2., 1.],
 [0., 0.],
 [0., 2.],
 [0., 2.],
 [2., 1.],
 [2., 1.]])
```

```
In [86]: oe.categories_
```

```
Out[86]: [array([ 'Poor' , 'Average' , 'Good' ], dtype=object),  
         array(['School', 'UG', 'PG'], dtype=object)]
```

```
In [87]: X_train
```

```
Out[87]: array([[1., 2.],  
               [1., 0.],  
               [0., 1.],  
               [2., 0.],  
               [2., 1.],  
               [1., 1.],  
               [2., 1.],  
               [0., 0.],  
               [2., 1.],  
               [0., 2.],  
               [1., 1.],  
               [0., 2.],  
               [2., 0.],  
               [2., 2.],  
               [0., 0.],  
               [2., 2.],  
               [2., 1.],  
               [0., 2.],  
               [0., 1.],  
               [2., 0.],  
               [2., 2.],  
               [0., 2.],  
               [0., 0.],  
               [2., 0.],  
               [2., 0.],  
               [1., 1.],  
               [2., 2.],  
               [0., 1.],  
               [2., 2.],  
               [1., 1.],  
               [2., 2.],  
               [0., 1.],  
               [1., 0.],  
               [0., 0.],  
               [0., 2.],  
               [0., 2.],  
               [1., 2.],  
               [2., 1.],  
               [1., 0.],  
               [2., 0.]])
```

```
In [95]: from sklearn.preprocessing import LabelEncoder
```

```
In [96]: le = LabelEncoder()
```

```
In [97]: le.fit(y_train)
```

```
Out[97]: LabelEncoder()
```

```
In [98]: le.classes_
```

```
Out[98]: array(['No', 'Yes'], dtype=object)
```

```
In [99]: y_train = le.transform(y_train)
         y_test = le.transform(y_test)
```

```
In [100... y_train
```

```
Out[100... array([1, 0, 0, 1, 1, 0, 1, 0, 1, 1, 1, 1, 1, 0,
        0, 1, 0, 0, 1, 1, 0, 0,
               1, 1, 0, 0, 1, 0, 1, 0, 1, 0, 1, 1, 0, 0,
        0, 0, 1, 1])
```

```
In [ ]:
```