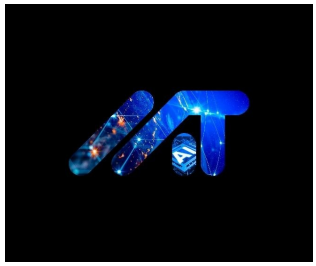

In [1]:
```
import numpy as np
import pandas as pd
```

In [2]:
```
df=pd.read_csv('customer.csv')
```

In [3]:
```
df.sample(5)
```

Out[3]:

	age	gender	review	education	purchased
1	68	Female	Poor	UG	No
24	16	Female	Average	PG	Yes
15	75	Male	Poor	UG	No
11	74	Male	Good	UG	Yes
47	38	Female	Good	PG	Yes

In [4]:
```
df=df.iloc[:,2:]
```

In [5]:
```
df.head()
```

Out[5]:

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

Train Test Split

In [18]:
```
from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(df.iloc[:,0:2],df.iloc[:,-1],test_size=0.2)
```

In [19]:
```
x_train
```

Out[19]:

	review	education
2	Good	PG
17	Poor	UG
22	Poor	PG
49	Good	UG
21	Average	PG
12	Poor	School
0	Average	School
35	Poor	School
24	Average	PG
13	Average	School
25	Good	School
48	Good	UG

46	Poor	PG
14	Poor	PG
18	Good	School
30	Average	UG
32	Average	UG
37	Average	PG
38	Good	School
5	Average	School
10	Good	UG
28	Poor	School
41	Good	PG
39	Poor	PG
3	Good	PG
33	Good	PG
42	Good	PG
45	Poor	PG
27	Poor	PG
23	Good	School
29	Average	UG
9	Good	UG
31	Poor	School
20	Average	School
11	Good	UG
7	Poor	School
4	Average	UG
47	Good	PG
36	Good	UG
15	Poor	UG

Ordinal Encoding

In ordinal encoding, each unique category value is assigned an integer value.

For example, “red” is 1, “green” is 2, and “blue” is 3.

This is called an ordinal encoding or an integer encoding and is easily reversible. Often, integer values starting at zero are used.

For some variables, an ordinal encoding may be enough. The integer values have a natural ordered relationship between each other and machine learning algorithms may be able to understand and harness this relationship.

In [20]:
```
from sklearn.preprocessing import OrdinalEncoder
```

In [21]:
```
oe = OrdinalEncoder(categories=[['Poor','Average','Good'],['School','UG','PG']])
```

In [22]:
```
oe.fit(x_train)
```

Out[22]:
```
OrdinalEncoder(categories=[['Poor', 'Average', 'Good'], ['School', 'UG', 'PG']])
```

In [23]:
```
x_train=oe.transform(x_train)
```

In [24]:
```
x_train
```

Out[24]:
```
array([[2., 2.],
       [0., 1.],
       [0., 2.],
       [2., 1.],
```

```
       [1., 2.],
       [0., 0.],
       [1., 0.],
       [0., 0.],
       [1., 2.],
       [1., 0.],
       [2., 0.],
       [2., 1.],
       [0., 2.],
       [0., 2.],
       [2., 0.],
       [1., 1.],
       [1., 1.],
       [1., 2.],
       [2., 0.],
       [1., 0.],
       [2., 1.],
       [0., 0.],
       [2., 2.],
       [0., 2.],
       [2., 2.],
       [2., 2.],
       [2., 2.],
       [0., 2.],
       [0., 2.],
       [2., 0.],
       [1., 1.],
       [2., 1.],
       [0., 0.],
       [1., 0.],
       [2., 1.],
       [0., 0.],
       [1., 1.],
       [2., 2.],
       [2., 1.],
       [0., 1.]])
```

In [25]:
```
x_test=oe.transform(x_test)
```

In [26]:
```
x_test
```

Out[26]:
```
array([[0., 2.],
       [2., 0.],
       [0., 1.],
       [2., 0.],
       [0., 2.],
       [1., 1.],
       [1., 0.],
       [1., 1.],
       [0., 2.],
       [0., 1.]])
```

In [27]:
```
oe.categories_
```

Out[27]:
```
[array(['Poor', 'Average', 'Good'], dtype=object),
 array(['School', 'UG', 'PG'], dtype=object)]
```

Label Encoding

Label Encoding refers to converting the labels into a numeric form so as to convert them into the machine-readable form. Machine learning algorithms can then decide in a better way how those labels must be operated. It is an important pre-processing step for the structured dataset in supervised learning.

In [28]:
```
from sklearn.preprocessing import LabelEncoder
```

In [29]:
```
le=LabelEncoder()
```

In [30]:
```
le.fit(y_train)
```

Out[30]:
```
LabelEncoder()
```

In [31]:
```
le.classes_
```

Out[31]:
```
array(['No', 'Yes'], dtype=object)
```

In [32]:
```
y_train=le.transform(y_train)
y_test=le.transform(y_test)
```

In [33]:
```
y_train
```

Out[33]:
```
array([0, 1, 1, 0, 0, 0, 0, 1, 1, 0, 0, 1, 0, 1, 0, 0, 1, 1, 0, 1, 1, 0,
       1, 0, 0, 1, 1, 1, 0, 0, 1, 1, 1, 1, 1, 1, 0, 1, 1, 0])
```

Thank you

Author

Muhammad Zaman Ali
