ordinal-encording

April 4, 2024

Import pandas and sklearn

[1]:
```
import pandas as pd
from sklearn.preprocessing import OrdinalEncoder
```

Read the dataset

[3]:
```
data = pd.read_csv('/content/toys.csv')
data
```

[3]:

	age	gender	fever	cough	city	has_covid
0	60	Male	103.0	Mild	Kolkata	No
1	27	Male	100.0	Mild	Delhi	Yes
2	42	Male	101.0	Mild	Delhi	No
3	31	Female	98.0	Mild	Kolkata	No
4	65	Female	101.0	Mild	Mumbai	No
..	…	…	…	…	…	…
95	12	Female	104.0	Mild	Bangalore	No
96	51	Female	101.0	Strong	Kolkata	Yes
97	20	Female	101.0	Mild	Bangalore	No
98	5	Female	98.0	Strong	Mumbai	No
99	10	Female	98.0	Strong	Kolkata	Yes

[100 rows x 6 columns]

Display the top 5 observations in the dataset

[4]:
```
data.head()
```

[4]:

	age	gender	fever	cough	city	has_covid
0	60	Male	103.0	Mild	Kolkata	No
1	27	Male	100.0	Mild	Delhi	Yes
2	42	Male	101.0	Mild	Delhi	No
3	31	Female	98.0	Mild	Kolkata	No
4	65	Female	101.0	Mild	Mumbai	No

compute the count of unique values in the column named cough

[5]:
```
data['cough'].value_counts()
```

[5]:
```
cough
Mild      62
Strong    38
Name: count, dtype: int64
```

[ ]:
```
#mild
#strong
```

[8]:
```
data_Ordinal=data[['cough']]
data_Ordinal.head()
```

[8]:
```
  cough
0  Mild
1  Mild
2  Mild
3  Mild
4  Mild
```

[9]:
```
oe = OrdinalEncoder(categories=[["Mild", "Strong"]])
```

[10]:
```
oe.fit(data_Ordinal)
```

[10]:
```
OrdinalEncoder(categories=[['Mild', 'Strong']])
```

[11]:
```
data_Transform =oe.transform(data_Ordinal)
```

[12]:
```
data_Transform
```

[12]:
```
array([[0.],
       [0.],
       [0.],
       [0.],
       [0.],
       [0.],
       [1.],
       [1.],
       [1.],
       [0.],
       [0.],
       [0.],
       [0.],
       [1.],
       [0.],
       [0.],
       [1.],
       [0.],
       [1.],
       [0.],
```

[1.],
[1.],
[0.],
[1.],
[0.],
[1.],
[0.],
[0.],
[1.],
[0.],
[1.],
[0.],
[0.],
[1.],
[0.],
[0.],
[1.],
[0.],
[0.],
[0.],
[0.],
[0.],
[0.],
[0.],
[0.],
[1.],
[0.],
[0.],
[0.],
[1.],
[0.],
[0.],
[1.],
[1.],
[1.],
[0.],
[0.],
[0.],
[1.],
[1.],
[1.],
[0.],
[1.],
[1.],
[1.],
[0.],
[0.],
[0.],

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

[15]:
```
New_data = pd.DataFrame(data_Transform, columns=['cough'])
```

[17]:
```
New_data.head()
```

[17]:

	cough
0	0.0
1	0.0
2	0.0
3	0.0
4	0.0

[18]:
```
New_data['cough'].value_counts
```

[18]:
```
<bound method IndexOpsMixin.value_counts of 0     0.0
1     0.0
2     0.0
3     0.0
4     0.0
     …
95    0.0
96    1.0
97    0.0
98    1.0
99    1.0
Name: cough, Length: 100, dtype: float64>
```