ML 10 - Data Encoding (Ordinal and Label Encoding) By Virat Tiwari

December 1, 2023

1 ORDINAL AND LABEL ENCODER

CASE 1 -LABEL ENCODER

[1]:
```
import pandas as pd
```

[2]:
```
from sklearn.preprocessing import LabelEncoder
```

[3]:
```
# Here we create a dataframe with feature "color"

df=pd.DataFrame({
    "color":["red","blue","green","green","red","blue"]
})
```

[4]:
```
df
```

[4]:
```
   color
0    red
1   blue
2  green
3  green
4    red
5   blue
```

[6]:
```
# Create an instance of label encoder

encoder=LabelEncoder()
```

[8]:
```
# fit_transform(df["color"]) function is used for assingninh=g the labels of␣
 ↪color like 2 for red , 0 for blue , 1 for green

encoder.fit_transform(df["color"])
```

[8]:
```
array([2, 0, 1, 1, 2, 0])
```

CASE 2 - For Ranking We have to use ORDINAL ENCODING -

[9]:
```
import pandas as pd
from sklearn.preprocessing import OrdinalEncoder
```

[11]:
```
# Here we create a dataframe with feature "size"

df=pd.DataFrame({
    "size":["small","medium","large","medium","small","large"]
})
```

[12]:
```
df
```

[12]:

	size
0	small
1	medium
2	large
3	medium
4	small
5	large

[20]:
```
# OrdinalEncoder ( ) is used for assigning the catrgorical data into the␣
 ↪numerical data
# categories=[ [ ] ] is used for assigning the rank of variable

encoder=OrdinalEncoder(categories=[["small","medium","large"]])
```

[21]:
```
# fit_transform(df[[ ]]) is used for tranforming the data

encoder.fit_transform(df[["size"]])
```

[21]:
```
array([[0.],
       [1.],
       [2.],
       [1.],
       [0.],
       [2.]])
```

THANK YOU SO MUCH !!

YOURS VIRAT TIWARI :)