dummie variables

May 11, 2023

[3]:
```
import pandas as pd
df=pd.read_csv(r"C:\Users\Rakesh\Downloads\carprices.csv")
df
```

[3]:

	Car Model	Mileage	Sell Price($)	Age(yrs)
0	BMW X5	69000	18000	6
1	BMW X5	35000	34000	3
2	BMW X5	57000	26100	5
3	BMW X5	22500	40000	2
4	BMW X5	46000	31500	4
5	Audi A5	59000	29400	5
6	Audi A5	52000	32000	5
7	Audi A5	72000	19300	6
8	Audi A5	91000	12000	8
9	Mercedez Benz C class	67000	22000	6
10	Mercedez Benz C class	83000	20000	7
11	Mercedez Benz C class	79000	21000	7
12	Mercedez Benz C class	59000	33000	5

[6]:
```
df1=pd.get_dummies(df['Car Model'])
df1
```

[6]:

	Audi A5	BMW X5	Mercedez Benz C class
0	0	1	0
1	0	1	0
2	0	1	0
3	0	1	0
4	0	1	0
5	1	0	0
6	1	0	0
7	1	0	0
8	1	0	0
9	0	0	1
10	0	0	1
11	0	0	1
12	0	0	1

[10]:
```
final=pd.concat([df,df1],axis='columns')
final
```

[10]:

	Car Model	Mileage	Sell Price($)	Age(yrs)	Audi A5	BMW X5	\
0	BMW X5	69000	18000	6	0	1	
1	BMW X5	35000	34000	3	0	1	
2	BMW X5	57000	26100	5	0	1	
3	BMW X5	22500	40000	2	0	1	
4	BMW X5	46000	31500	4	0	1	
5	Audi A5	59000	29400	5	1	0	
6	Audi A5	52000	32000	5	1	0	
7	Audi A5	72000	19300	6	1	0	
8	Audi A5	91000	12000	8	1	0	
9	Mercedez Benz C class	67000	22000	6	0	0	
10	Mercedez Benz C class	83000	20000	7	0	0	
11	Mercedez Benz C class	79000	21000	7	0	0	
12	Mercedez Benz C class	59000	33000	5	0	0	

	Mercedez Benz C class
0	0
1	0
2	0
3	0
4	0
5	0
6	0
7	0
8	0
9	1
10	1
11	1
12	1

[15]:
```
final1=final.drop(['Car Model','Audi A5'],axis='columns')
final1
```

[15]:

	Mileage	Sell Price($)	Age(yrs)	BMW X5	Mercedez Benz C class
0	69000	18000	6	1	0
1	35000	34000	3	1	0
2	57000	26100	5	1	0
3	22500	40000	2	1	0
4	46000	31500	4	1	0
5	59000	29400	5	0	0
6	52000	32000	5	0	0
7	72000	19300	6	0	0
8	91000	12000	8	0	0
9	67000	22000	6	0	1

10	83000	20000	7	0	1
11	79000	21000	7	0	1
12	59000	33000	5	0	1

[17]:
```
from sklearn.linear_model import LinearRegression
model=LinearRegression()
```

[22]:
```
x=final1.drop('Sell Price($)',axis='columns')
y=final['Sell Price($)']
```

[23]:
```
model.fit(x,y)
```

[23]:
```
LinearRegression()
```

[33]:
```
model.predict([[4500,4,0,1]])
```

```
C:\Users\Rakesh\anaconda3\lib\site-packages\sklearn\base.py:450: UserWarning: X
does not have valid feature names, but LinearRegression was fitted with feature
names
  warnings.warn(
```

[33]:
```
array([51981.26203334])
```

[32]:
```
model.predict([[86000,7,1,0]])
```

```
C:\Users\Rakesh\anaconda3\lib\site-packages\sklearn\base.py:450: UserWarning: X
does not have valid feature names, but LinearRegression was fitted with feature
names
  warnings.warn(
```

[32]:
```
array([11080.74313219])
```

[31]:
```
model.score(x,y)
```

[31]:
```
0.9417050937281082
```

[ ]:
```

```