Car Resale Value Prediction

Splitting Data into Independent and Dependent Variables

X = car_dataset.drop(['Name','Price'],axis= 1)

Y = car_dataset['Price']

print(X)

Output:

	Unnamed: 0	Location	Year	Kilometers_Driven	Fuel_Type	Transmission \
0	0	Mumbai	2010	72000	2	0
1	1	Pune	2015	41000	1	0
2	2	Chennai	2011	46000	0	0
3	3	Chennai	2012	87000	1	0
4	4	Coimbatore	2013	40670	1	1
...	...	...	...	...	...	...
6014	6014	Delhi	2014	27365	1	0
6015	6015	Jaipur	2015	100000	1	0
6016	6016	Jaipur	2012	55000	1	0
6017	6017	Kolkata	2013	46000	0	0
6018	6018	Hyderabad	2011	47000	1	0

	Owner_Type	Mileage	Engine	Power	Seats	New_Price
0	First	26.6 km/kg	998 CC	58.16 bhp	5.0	NaN
1	First	19.67 kmpl	1582 CC	126.2 bhp	5.0	NaN
2	First	18.2 kmpl	1199 CC	88.7 bhp	5.0	8.61 Lakh
3	First	20.77 kmpl	1248 CC	88.76 bhp	7.0	NaN
4	Second	15.2 kmpl	1968 CC	140.8 bhp	5.0	NaN
...	...	...	...	...	...	...
6014	First	28.4 kmpl	1248 CC	74 bhp	5.0	7.88 Lakh
6015	First	24.4 kmpl	1120 CC	71 bhp	5.0	NaN
6016	Second	14.0 kmpl	2498 CC	112 bhp	8.0	NaN
6017	First	18.9 kmpl	998 CC	67.1 bhp	5.0	NaN
6018	First	25.44 kmpl	936 CC	57.6 bhp	5.0	NaN

[6019 rows x 12 columns]

print(Y)

0	1.75
1	12.50
2	4.50
3	6.00
4	17.74
	...

6014	4.75
6015	4.00
6016	2.90
6017	2.65
6018	2.50

Name: Price, Length: 6019, dtype: float64

Spiltting the Training data and Test Data

X_train,X_test,Y_train,Y_test = train_test_split(X,Y,test_size = 0.1,random_state = 2)

print(X_train)

	Unnamed: 0	Location	Year	Kilometers_Driven	Fuel_Type	Transmission \
2083	2083	Bangalore	2015	65000	1	1
2887	2887	Mumbai	2015	39000	0	0
1230	1230	Ahmedabad	2012	52000	1	0
307	307	Chennai	2006	58000	0	0
5555	5555	Delhi	2010	75319	0	0
...	...	...	...	...	...	...
1099	1099	Chennai	2012	50000	0	0
2514	2514	Kolkata	2010	35008	0	0
3606	3606	Delhi	2015	53000	0	0
5704	5704	Jaipur	2017	54455	1	1
2575	2575	Hyderabad	2015	120000	1	0

	Owner_Type	Mileage	Engine	Power	Seats	New_Price
2083	Second	18.56 kmpl	1995 CC	190 bhp	5.0	NaN
2887	First	18.6 kmpl	1197 CC	81.83 bhp	5.0	NaN
1230	First	21.43 kmpl	1364 CC	86.79 bhp	5.0	19.27 Lakh
307	Second	0.0 kmpl	1086 CC	null bhp	5.0	NaN
5555	Second	21.1 kmpl	1197 CC	85.80 bhp	5.0	NaN
...	...	...	...	...	...	...
1099	First	17.0 kmpl	1197 CC	80 bhp	5.0	NaN
2514	First	17.92 kmpl	1086 CC	62.1 bhp	5.0	NaN
3606	First	17.8 kmpl	1497 CC	117.3 bhp	5.0	NaN
5704	First	22.0 kmpl	1498 CC	108.495 bhp	5.0	NaN
2575	First	14.8 kmpl	2499 CC	80 bhp	10.0	NaN

[5417 rows x 12 columns]

print(X_test)

	Unnamed: 0	Location	Year	Kilometers_Driven	Fuel_Type	Transmission \
4930	4930	Coimbatore	2017	60993	1	0
463	463	Delhi	2014	31465	1	0
338	338	Pune	2012	73000	1	0

3090	3090	Bangalore	2010	72010	1	0
1826	1826	Kochi	2012	83184	0	0
...	...	...	...	...	...	...
3054	3054	Mumbai	2014	11000	0	0
887	887	Mumbai	2011	48000	0	1
418	418	Coimbatore	2018	22397	0	1
1838	1838	Hyderabad	2007	70000	0	0
1064	1064	Kolkata	2009	43002	0	0

	Owner_Type	Mileage	Engine	Power	Seats	New_Price
4930	First	24.0 kmpl	1120 CC	70 bhp	5.0	NaN
463	First	22.7 kmpl	1498 CC	89.84 bhp	5.0	NaN
338	First	21.12 kmpl	1248 CC	88.8 bhp	5.0	NaN
3090	First	16.8 kmpl	1493 CC	110 bhp	5.0	NaN
1826	First	14.28 kmpl	1798 CC	138.03 bhp	5.0	21.69 Lakh
...	...	...	...	...	...	...
3054	First	17.01 kmpl	1591 CC	121.3 bhp	5.0	NaN
887	Second	14.74 kmpl	1991 CC	184 bhp	5.0	NaN
418	First	11.5 kmpl	2996 CC	367 bhp	5.0	95.38 Lakh
1838	Second	19.81 kmpl	1086 CC	68.05 bhp	5.0	NaN
1064	First	20.36 kmpl	1197 CC	78.9 bhp	5.0	NaN

[602 rows x 12 columns]

print(Y_train)

2083	29.00
2887	5.50
1230	6.50
307	1.50
5555	2.00
	...
1099	4.00
2514	1.36
3606	6.95
5704	6.50
2575	5.50

Name: Price, Length: 5417, dtype: float64

print(Y_test)

4930	5.21

463	6.50
338	3.40
3090	4.45
1826	6.22
	...
3054	6.10
887	9.75
418	70.99
1838	2.00
1064	1.92

Name: Price, Length: 602, dtype: float64