

Car Price Prediction using RANDOM FOREST REGRESSION

1 Problem Predicting Car price

2 Dataset

X-parameter | Y-price
Car price based on other car numerical parameters

3 Summarize Dataset

(10, 2)
dataset.shape
dataset.head(5)

4 Segregate Dataset into X & Y

Input - Taking only Numerical Values
dataset.select_dtypes(exclude=['object']).columns
Output
Y = dataset['price']

5 Scaling our Independent Variables(X)

1 Scale

Imagine you have a number, 50, and you want to rescale it from a range of 0-100 to a range of 0-1.

Rescaling Formula:

- Original Range: 0 to 100
- Target Range: 0 to 1

The formula to rescale is:

$$\text{Rescaled Value} = \frac{\text{Original Value} - \text{Original Min}}{\text{Original Max} - \text{Original Min}} \times (\text{Target Max} - \text{Target Min}) + \text{Target Min}$$

For the number 50:

- Original Min: 0
- Original Max: 100
- Target Min: 0
- Target Max: 1

So:

$$\text{Rescaled Value} = \frac{50 - 0}{100 - 0} \times (1 - 0) + 0 = 0.5$$

Summary

Scaling in this context means adjusting a value from one range to another range. For example, scaling 50 from a 0-100 range to a 0-1 range gives 0.5.

2 Rescaling

6 Splitting dataset to train & test for Validation

train_test_split

7 Algorithm

RANDOM FOREST

8 Train using Random Forest

9 Prediction & Validation

r2_score

