

Predict automobile price using Machine Learning

Use Machine learning to predict the price of a car based on several characteristics. The objective is to build a model to understand the factors that drive the car of the price. This will help your automobile company launch their new car in the market effectively by pricing it better.

Tasks:

- Perform EDA on the data
- Perform data cleanup as required
- Pick the best variable for making a simple linear regression model
- Perform train test split
- Build model using best variable and report the R2
- Make a multiple regression model
 - o Apply feature selection approaches discussed in the class
- Final model should be interpretable
 - o What is your understanding of the factors that drive price?

Data description:

Attribute: Attribute Range

1. symboling: -3, -2, -1, 0, 1, 2, 3

(Cars are initially assigned a risk factor symbol associated with its price. Then, if it is more risky (or less), this symbol is adjusted by moving it up (or down) the scale. Actuarians call this process "symboling". A value of +3 indicates that the auto is risky, -3 that it is probably pretty safe.)

2. make: alfa-romero, audi, bmw, chevrolet, dodge, honda, isuzu, jaguar, mazda, mercedes- benz, mercury, mitsubishi, nissan, peugot, plymouth, porsche, renault, saab, subaru, toyota, volkswagen, volvo

3. fuel-type: diesel, gas.

4. aspiration: std, turbo.

5. num-of-doors: four, two.

6. body-style: hardtop, wagon, sedan, hatchback, convertible.

7. drive-wheels: 4wd, fwd, rwd.

8. engine-location: front, rear.

9. wheel-base: continuous from 86.6 120.9.

10. length: continuous from 141.1 to 208.1. 11

width: continuous from 60.3 to 72.3.

12. height: continuous from 47.8 to 59.8.

13. curb-weight: continuous from 1488 to 4066.

14. engine-type: dohc, dohcv, l, ohc, ohcf, ohcv, rotor.

15. num-of-cylinders: eight, five, four, six, three, twelve, two.

16. engine-size: continuous from 61 to 326.

17. fuel-system: 1bbl, 2bbl, 4bbl, idi, mfi, mpfi, spdi, spfi.

18. bore: continuous from 2.54 to 3.94.

19. stroke: continuous from 2.07 to 4.17.

20. compression-ratio: continuous from 7 to 23.

21. horsepower: continuous from 48 to 288.

- 22. peak-rpm: continuous from 4150 to 6600.
- 23. city-mpg: continuous from 13 to 49.
- 24. highway-mpg: continuous from 16 to 54.
- 25. price: continuous from 5118 to 45400.