Data set details

1. Dataset file: ‘cars_price.csv’
2. This is a regression data set.

2. The data set has 206 samples.

3. There are 25 features.

4. The target variable is the price of the car.

Attributes

- symboling
- normalized-losses
- make
- fuel-type
- aspiration
- num-of-doors
- body-style
- drive-wheels
- engine-location
- wheel-base
- length
- width
- height
- curb-weight
- engine-type
- num-of-cylinders
- engine-size
- fuel-system
- bore
- stroke
- compression-ratio
- horsepower
- peak-rpm
- city-mpg
- highway-mpg

Tasks

The following are the broad tasks to be performed:

- Load the data into your code
- Keep 20% of the data aside as test data.
- Apply the required preprocessing techniques on the data.
- Apply different machine learning techniques.
- Tune the parameters of models as required.

- Build a model named ‘final_model’.
- Report the following metrics using ‘final_model’ on the test data:
    - MSE
    - MAE
    - R2-score
- Report the importance of different features.
- **Prepare a brief report documenting the key steps involved along with the reasoning behind different decisions that you made in these steps.**

Notes

- Write textual comments in your code wherever appropriate.
- **Note:** As a data scientist, you might need to make several decisions while exploring the data. feel free to do that and mention the same in the report.
- Be as innovative and creative as possible. Visualize data wherever required.