**Car company car price prediction**

**Due date Saturday Nov-13-2021 by 9pm**

Source: https://www.kaggle.com/goyalshalini93/car-price-prediction-linear-regression-rfe/notebook

**Problem Statement:**

An automobile company aspires to enter the US market by setting up their manufacturing unit there and producing cars locally to give competition to their US and European counterparts.

They have contracted an automobile consulting company to understand the **factors** on which the **pricing of cars depends**. Specifically, they want to understand the factors affecting the pricing of cars in the American market, since those may be very different from other markets. The company wants to know:

* Which variables are significant in predicting the price of a car
* How well those variables describe the price of a car

Based on various market surveys, the consulting firm has **gathered** a large dataset of different types of cars across the American market.

**Business Goal:**

You are required to model the price of cars with the available independent variables. It will be used by the management to understand how exactly the prices vary with the independent variables. They can accordingly manipulate the design of the cars, the business strategy etc. to meet certain price levels. Further, the model will be a good way for management to understand the pricing dynamics of a new market.

**Questions (100 points):**

1. What is T (Task), E (Experience), P (Performance) for this specific problem?
2. What is the size of the training data? How many features do we have?
3. Prepare a Jupyter python notebook to explore the given data.
4. Use the scikit-learn package to use the proper machine learning model for your analysis. Specifically:
   1. Model the relationship between car price and horsepower using linear regression.
   2. Model the relationship between car price and citymapg using linear regression.
   3. Which variable (feature) better explains the car price? How did you reach that conclusion?
   4. Prepare a brief power point presentation **which contains the following**:
      1. Title slide which has the project name, your name, your university department, and date
      2. A slide explaining the problem
      3. A slide explaining the goals or aims of the project
      4. A slide answering questions 1 and 2 above
      5. A slide highlighting the methods you used
      6. A slide with your main conclusions with metrics supporting your conclusions
      7. A future work slide and any obstacles you faced during the project such as data pre-processing, data quality, data size
      8. Convert your ppt to pdf and name it as: LastName\_FirstName\_CarPriceProject.pdf

BONUS question (20 points):

* Can you model the effect of both horsepower and citymapg using one machine learning model? What is this model called? You can use other python frameworks/libraries besides scikit-learn if you prefer.

How to submit:

* Submit your Jupyter notebook to github and eLearning
* Submit your slides in pdf format to eLearning