Major Project: Predictive Modeling of Car Prices for Market Entry Strategy in the US Auto Industry

Problem Statement

A Chinese automobile company Geely Auto 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 the Chinese market. 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 data set of different types of cars across the America market.

What We Are Trying to Solve

We 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.

Dataset Information

The dataset contains information on various cars, with several attributes (or columns) describing different aspects of each car. Here is an explanation of each column present in the data:

Column Information

- ➢ car_ID: A unique identifier for each car.
- ➢ symboling: An insurance risk rating, ranging from -3 to 3, where higher values indicate a higher risk.
- ➢ CarName: The name of the car, which includes both the brand and model.
- ➢ fueltype: The type of fuel used by the car (e.g., gas, diesel).
- ➢ aspiration: Indicates whether the car has a standard (std) or turbocharged (turbo) engine.
- ➢ doornumber: The number of doors on the car (e.g., two, four).
- ➢ carbody: The body style of the car (e.g., convertible, hatchback, sedan, wagon, hardtop).
- ➢ drivewheel: The type of drive system the car uses (e.g., fwd for front-

wheel drive, rwd for rear-wheel drive, 4wd for four-wheel drive).
- ➤ enginelocation: The location of the engine in the car (e.g., front, rear).
- ➤ wheelbase: The distance between the front and rear axles (measured in inches).
- ➤ carlength: The overall length of the car (measured in inches).
- ➤ carwidth: The overall width of the car (measured in inches).
- ➤ carheight: The overall height of the car (measured in inches).
- ➤ curbweight: The weight of the car without passengers or cargo (measured in pounds).
- ➤ enginetype: The type of engine (e.g., dohc, ohcv, ohc, rotor, dohcv, l, ohcf).
- ➤ cylindernumber: The number of cylinders in the engine (e.g., two, three, four, five, six, eight, twelve).
- ➤ enginesize: The size of the engine (measured in cubic inches).
- ➤ fuelsystem: The type of fuel system (e.g., mpfi, 2bbl, 1bbl, spdi, 4bbl, idi, spfi).
- ➤ boreratio: The ratio of the bore (diameter of the cylinder) to the stroke (length of the cylinder).
- ➤ stroke: The length of the piston stroke within the cylinder.
- ➤ compressionratio: The ratio of the maximum to minimum volume in the cylinder.
- ➤ horsepower: The power output of the engine (measured in horsepower).
- ➤ peakrpm: The maximum revolutions per minute (RPM) of the engine.
- ➤ citympg: The fuel efficiency of the car in city driving conditions (measured in miles per gallon).
- ➤ highwaympg: The fuel efficiency of the car in highway driving conditions (measured in miles per gallon).
- ➤ price: The price of the car (measured in USD), which is the target variable for prediction.

Model Comparison Report
Create a report stating the performance of multiple models on this data and suggest the best model for production.

Report on Challenges faced
Create a report which should include challenges you faced on data and what technique used with proper reason.
Note:-All above tasks have been created on a single jupyter notebook and share the same while final submission of project to the following email_ID: krutanic@gmail.com