Cogs 109
Yushan Wang
A14459085

Fuel Consumption Analysis and Prediction

17th November 2020

Intro to Dataset

The dataset is collected from Kaggle.com Below is the link of the dataset:

https://www.kaggle.com/anderas/car-consume

I generally took a review of the present notebooks of this dataset: It has regression for prediction. It also has basic data visualization and analysis. Below is the link for the notebooks of this dataset:

https://www.kaggle.com/anderas/car-consume/notebooks

Data Exploration and Visualization

I would use python to explore and visualize the data. As far as I am concerned, the amount of data is not tremendous. Python should provide proper visualization. If better visualization is required, I am planning to use tableau as a complement.

Research Question

This study aims at a possible solution to predict the fuel consumption based on the data we had. The main objective is to examine whether an approach of machine learning could be a viable option to predict fuel consumption.

Planned Analysis

I proposed to design 2 models:

Model 1: using all parameters we have for fuel consumption prediction

Model 2: Using only driver-based variables (speed and road-distance)

The Machine Learning Models I will choose to use in this model would be Linear Regression, Support Vector Regression and Neural Network(Optional).

The Evaluation of the models:

I will use different train/split ratios to train the two types of the model, then I would compare the Balanced Accuracy and Absolute Errors of the two models with different Machine Learning Model Applications.

Planned Report

I will report the performance of my choices of models, and also report how well my optimized model would perform in solving the fuel consumption prediction question.