

Joseph Cornelius  
John DiTaranto

### **Final Project: Part 1**

3. For this project we chose to use the “Automobile” dataset from the UCI Machine Learning Repository.

<https://archive.ics.uci.edu/ml/datasets/Automobile>

4. For this specific dataset we are going to be addressing the dataset as a regression algorithm, specifically comparing the effects of horsepower, number of cylinders, & curb-weight on the expected MPG’s (city & highway) of the vehicles. Since we will be giving the machine the desired output values of the MPG, this will be a supervised model. Once this is done, we will compare its accuracy to that of if we simply gave the model the entire datasheet, rather than what we deemed were the more important variables.

In addition to comparing this to the nearest neighbors approach, we will also be using feature selection techniques to test our own model, as described in the previous paragraph.