

Assignment 1

Luis Dale Gascon
Computer Science
Towson University
lgascon1@students.towson.edu

I. INTRODUCTION

II. DATASET

The dataset is a modified version from Carnegie Mellon University's StatLib datasets. Each row contains information about a vehicle's horsepower, weight, acceleration, with an endgoal of predicting a vehicle's mpg

III. EXPLORATORY DATA ANALYSIS

IV. DATA PREPROCESSING

There are missing values for horsepower

A. Data Binning

We bin the data with clustering-based binning by using the k-means clustering algorithm to group the data based on similarities.

V. REGRESSION

VI. CONCLUSION

REFERENCES