

1. Title: Auto-Mpg Data

2. Sources:

(a) Origin: This dataset was taken from the StatLib library which is maintained at Carnegie Mellon University. The dataset was used in the 1983 American Statistical Association Exposition.

(c) Date: July 7, 1993

3. Past Usage:

- See 2b (above)

- Quinlan, R. (1993). Combining Instance-Based and Model-Based Learning. In Proceedings on the Tenth International Conference of Machine Learning, 236-243, University of Massachusetts, Amherst. Morgan Kaufmann.

4. Relevant Information:

This dataset is a slightly modified version of the dataset provided in the StatLib library. In line with the use by Ross Quinlan (1993) in predicting the attribute "mpg", 8 of the original instances were removed because they had unknown values for the "mpg" attribute. The original dataset is available in the file "auto-mpg.data-original".

"The data concerns city-cycle fuel consumption in miles per gallon, to be predicted in terms of 3 multivalued discrete and 5 continuous attributes." (Quinlan, 1993)

5. Number of Instances: 398

6. Number of Attributes: 9 including the class attribute

7. Attribute Information:

- | | |
|------------------|-----------------------------------|
| 1. mpg: | continuous |
| 2. cylinders: | multi-valued discrete |
| 3. displacement: | continuous |
| 4. horsepower: | continuous |
| 5. weight: | continuous |
| 6. acceleration: | continuous |
| 7. model year: | multi-valued discrete |
| 8. origin: | multi-valued discrete |
| 9. car name: | string (unique for each instance) |

8. Missing Attribute Values: horsepower has 6 missing values