

1. Tên bộ dữ liệu: Auto-Mpg Data

2. Nguồn dữ liệu:

Nguồn gốc: Tập dữ liệu được thu thập từ thư viện StatLib tại Đại học Carnegie Mellon. Bộ dữ liệu được sử dụng vào năm 1983 tại Triển lãm Hiệp hội Thống kê Hoa Kỳ.

Thời gian ra đời: 7/7/1993

Đường dẫn: <https://archive.ics.uci.edu/ml/datasets/Auto+MPG>

3. Thông tin về bộ dữ liệu:

Tập dữ liệu là phiên bản sửa đổi của bộ dữ liệu gốc được cung cấp từ thư viện StatLib.

Ross Quinlan (1993) sử dụng để dự đoán thuộc tính “mpg”.

Bộ dữ liệu gốc có sẵn trong tập tin "auto-mpg.data-original".

Chức năng: Bộ dữ liệu được dùng để mô tả chu kỳ tiêu thụ nhiên liệu của thành phố cho số dặm đi được cho mỗi gallon (miles per gallon) được dự đoán bởi 3 giá trị rời rạc (cylinder, model year, origin) và 5 thuộc tính liên tục (các biến còn lại).

Thông tin: Bộ dữ liệu cung cấp thông số kỹ thuật của các loại xe ô tô.

4. Số lượng các mẫu: 398.

5. Số lượng các thuộc tính: 9

6. Mô tả các thuộc tính

STT	Tên thuộc tính	Ý nghĩa thuộc tính	Kiểu dữ liệu	Miền giá trị
1	Mpg (<i>miles per gallon</i>)	Số dặm đi được cho một gallon	Số thực dương	(1,100)
2	Cylinders	Số xi lanh (inches)	Số nguyên dương	(0,10)
3	Displacement	Là số đo thể tích xi lanh được quét bởi tất cả các piston của động cơ piston, không bao gồm các buồng đốt (inches)	Số thực dương	(0,500)
4	Horsepower	Mã lực	Số thực dương	(0,300)
5	Weight	Khối lượng xe (lbs)	Số thực dương	(1000,5000)

6	Acceleration	Thời gian để tăng tốc từ 0 lên 60 dặm/ giờ	Số thực dương	(0,30)
7	Model year	Năm mẫu xe ra đời	Số dương	[70,82]
8	Origin	Nguồn gốc, xuất xứ	Số dương, (1: American, 2: European, 3: Japanese)	[1,3]
9	Car name	Tên xe	Chuỗi	(0,50)

7. Mất mát dữ liệu: Thuộc tính horsepower bị mất 6 giá trị.