|  |
| --- |
| Dataset 0: **Car Price Prediction Dataset**  Tác giả: **Sukhmandeep Singh Brar**  **Số lượng mẫu: 8128**  Số lượng đặc trưng: 12  Nguồn dữ liệu: Kaggle, cardekho.com  Loại kiểu dữ liệu: Văn bản, số thực, số nguyên  Tác dụng: **Car Price Prediction Dataset** bao gồm hàng loạt đặc điểm nổi bật liên quan đến ô tô, từ năm sản xuất, giá bán, số km đã lăn bánh, đến loại nhiên liệu, hình thức người bán, kiểu hộp số, số lượng chủ sở hữu trước đó, quãng đường đã đi, cùng các thông số kỹ thuật của động cơ. Những thuộc tính này mang đến cái nhìn sâu sắc và giá trị, hé lộ các yếu tố chi phối giá ô tô, đồng thời là nền tảng để xây dựng các mô hình dự đoán tinh tế, giúp ước lượng chính xác giá trị bán ra của những chiếc xe.  Định dạng: CSV  Danh sách các cột:   * Name: tên, thương hiệu xe * Year: năm sản xuất ( đơn vị: năm) * Selling\_price: giá bán (đơn vị: INR) * Km\_driven: số km đã đi * Fuel: loại nhiên liệu tiêu thụ * Seller\_type: loại người bán * Transmission: loại hộp số * Owner: số đời chủ sở hữu trước đó * Mileage: mức tiêu hao nguyên liệu của xe được đo bằng km/l hoặc km/kg * Engine: dung tích động cơ * Max\_power : công suất tối đa mà động cơ có thể tạo ra * Seats: số chỗ ngồi của xe   Biến mục tiêu: selling\_price |
| Dataset 1: **Used Car Prices**  Tác giả : **Sujay R**  Số lượng mẫu: 5847  Số lượng đặc trưng: 14  Nguồn dữ liệu: Kaggle, Cars24  Loại dữ liệu: văn bản ,số thực, số nguyên  Tác dụng: **Used Car Prices** chứa đựng nhiều chi tiết phong phú và các đặc điểm liên quan đến ô tô đã qua sử dụng. Trọng tâm của phân tích xoay quanh biến mục tiêu – giá của ô tô đã qua sử dụng, được tính bằng đơn vị lakhs, mang đến một cái nhìn sâu sắc về giá trị thị trường của những chiếc xe này.  Định dạng: CSV  Danh sách các cột   * Index : đánh thứ tự cho các hàng * Name: tên, thương hiệu của xe * Location: thành phố nơi chiếc xe được niêm yết * Year: năm sản xuất * Kilometers\_Driven: số km đã đi * Fuel\_Type: loại nhiên liệu sử dụng * Transmission: loại hộp số( tự động, thủ công) * Owner\_Type: số đời chủ sở hữu trước đó * Mileage: mức tiêu hao nguyên liệu của xe được đo bằng km/l hoặc km/kg * Engine: dung tích động cơ * Power: công suất tối đa * Seats : số chỗ ngồi * New\_Price: giá của chiếc xe mới cùng mẫu ( đơn vị Cr hoặc Lakh) * Price: giá bán hiện tại ( đơn vị: Lakh)   Biến mục tiêu: Price |
| Dataset 2: **Car Price Prediction Dataset**  Tác giả: Jackson Divakar R (Owner)  Số lượng mẫu: 2095  Số lượng đặc trưng: 14  Nguồn dữ liệu: Kaggle  Loại dữ liệu: văn bản ,số thực, số nguyên  Tác dụng : **Car Price Prediction Dataset** được xây dựng nhằm ước tính giá của xe đã qua sử dụng dựa trên nhiều đặc điểm khác nhau. Dataset này hữu ích trong các bài toán máy học và phân tích hồi quy, giúp xác định các yếu tố tác động đến giá xe trên thị trường.  Định dạng: CSV  Danh sách các cột:   * # : đánh số thứ tự * name: tên thương hiệu * year: năm sản xuất * selling\_price: giá bán (đơn vị INR) * km\_driven: số km đã đi * fuel: loại nhiên liệu sử dụng * seller\_type: loại người bán * transmission: loại hộp số( tự động hoặc thủ công) * owner: số lượng người sở hữu trước đó * seats: số chỗ ngồi * max\_power: công suất tối đa * Mileage Unit: đơn vị của Mileage * Mileage: mức tiêu hao nguyên liệu của xe * Engine: dung tích động cơ |