

Sprint 0: Fuel Efficient and Budget Friendly Car Recommendation System

Institute: BrainStation

Author: Sundus Yawar

Capstone Supervisor: Amir Rahimizadeh

Problem Area

I am interested in solving the problem of finding budget and environmentally friendly car. How might we use machine learning to find a car that is environmentally friendly and fuel efficient within our budget? This is an important problem because the cost of living have increased, last summer gas prices went till 216cents/litre. And this summer we have seen increase in gas prices from 145cents/litre to 169cents/litre. We also have a lot of cars on the road contributing to GHG emissions into the environment and not many people can afford to buy an electric car as they are costly and there is inflation and salaries are still the same as before.

The User

The users are millennials, zillennials and generation Z as these generations care a lot about our environmental impact and are mainly looking for either our first car or another car that is cost effective and environmentally friendly due to rising cost of living.

The Big Idea

I found someone on Kaggle using the same Fuel Consumption dataset from the open data by government of Canada website but they seem to be using it to determine factors that influence CO2 emissions. While what I am trying to do is build a recommendation system combining different datasets to provide users suggestion on buying a car within their budget that's environmentally friendly. This is the project I found that is using similar dataset but for different purpose: [CO2 Emission by Vehicles | Kaggle](#)

The Impact

As mentioned in the problem area, it will help people be environmentally responsible while staying within their budget because not everyone can afford an electric vehicle and the cost of living has increased. So, it will solve 2 problems, help users choose fuel efficient cars with low CO2 emissions killing 2 birds with one stone.

The Data

I will need to join the fuel consumption rating dataset with cars prices datasets.

[Fuel consumption ratings - Open Government Portal \(canada.ca\)](#)

The website has data for each year and will need to be aggregated, and it contains the following columns and subcolumns.

Model

Year

Make

Model

Vehicle Class

Engine Size

Cylinders

Transmission

Fuel
Fuel Consumption
City (L/100 km)
Hwy (L/100 km)
Comb (L/100 km)
Comb (mpg)
CO2 Emissions
CO2
Smog

[Canada Car Sale | Kaggle](#)

Name
Production Year
Price
Color
Type

[Used Car Price Data | Kaggle](#) dataset is from 2 years ago and the price is in Indian rupees so I'll need to convert to Canadian dollar using exchange rate from 2 years ago.

car_data.csv

1. Model: Model name of the car that is sold.
2. Selling Price: The selling price of the car. This is your target feature.
3. Kilometers Driven: Number of kilometers that the car has already driven.
4. Year: Year of purchase of the car.
5. Owner: Information on a number of the previous owners.
6. Fuel Type: Fuel type of car.
7. Transmission: Transmission type of car.
8. Insurance: Insurance information of the car.
9. Car Condition: Current car condition. A rating out of 5.

model_data.csv

1. Model: Model name of the car that is sold.
2. Current Price: Current price of the car.

[Car Prices Dataset | Kaggle](#)

Train.csv - 19237 rows x 18 columns (Includes Price Columns as Target)

- Attributes
- ID

- Price: price of the care(Target Column)
- Levy
- Manufacturer
- Model
- Prod. year
- Category
- Leather interior
- Fuel type
- Engine volume
- Mileage
- Cylinders
- Gear box type
- Drive wheels
- Doors
- Wheel
- Color
- Airbags

The Alternative

Alternatively, I'd like to help users buy cars of more reliable make i.e. ones that are less likely to be recalled. I found dataset for vehicles with recalls by make, model year etc. The user can input the car they'd like to purchase and then check its likelihood for being recalled. Or get recommendation on vehicles that are less likely to have recall based on their budget. The impact it'll have is that for people who are on a budget, and with rising cost of living can be paired with a reliable car that won't cause to many issues for them financially.

[Vehicle Recalls Database - Open Government Portal \(canada.ca\)](https://open.canada.ca/data/en/dataset/5966-273x-0008-0001-1000-000000000000)