# CAR EVALUATION

***FIRASATH ALI***



# Table of Contents

[Table of Contents 1](#_TOC_250005)

1. [Introduction 2](#_TOC_250004)
2. [Overview 2](#_TOC_250003)
3. [Data Overview 2](#_TOC_250002)
4. [Variables 2](#_TOC_250001)
5. EDA 3
6. [References 3](#_TOC_250000)



## Introduction:

This In our increasingly data-driven world, the power of technology and machine learning can be harnessed to simplify complex decision-making processes. This project delves into the realm of automobile assessment, where we aim to create an automatic system for car evaluation. Specifically, we will focus on aiding prospective car buyers in making informed choices by considering the physical attributes of vehicles. The foundation of our analysis lies in a dataset sourced from ics.uci.edu, which comprises 1,727 rows of data, each with seven distinct attributes.

## Overview:

Cars are essentially part of our regular day-to-day life. There are various kinds of cars produced by different manufacturers subsequently the buyers have a decision to make. When an individual considers buying a car, there are numerous aspects that could influence his/her choice of which kind of car he/she is keen on. The choice buyer or drivers have generally relied upon the price, safety, and how luxurious and how spacious the car is.

A car evaluation database is significant structure information that everyone should take a look at for the car features and is useful in decision making. This dataset is labeled according to the specification of PRICE, COMFORT, and SAFETY.

The objective of this report is especially to determine the decision making, identifying the car variables like car price with other various variables to decide between good acceptable cars from the unaccepted values from the target value.

1. **Dataset Overview:**

The dataset we've obtained contains crucial information about various cars, each described by six attributes. These attributes serve as the key determinants in categorizing each car into one of four classes: unacceptable, acceptable, good, or very good. The importance of this project is evident in its potential to empower individuals to make well-informed decisions when it comes to purchasing an automobile, a significant financial investment.

**4. Variables:**

* **Buying Price** : v-high, high, med, low  
  **Maintenance Cost** : v-high, high, med, low  
  **Number of doors**: 2, 3, 4, 5-more  
  **Number of persons**: 2, 4, more  
  **lug\_boot**: small, med, big  
  **Safety**: low, med, high  
  **Decision**: unacceptable, acceptable, good or very good

**5. EDA:**

In the course of this project, we have employed various data visualization techniques, particularly count plots, to gain a comprehensive understanding of the car evaluation dataset. Count plots have allowed us to visualize the distribution of data across different variables, unveiling the frequency and uniqueness of values within each attribute. One crucial observation that emerges from our exploration is the imbalance in the 'class' variable. We notice a higher prevalence of 'unacc' values compared to the other classes. This observation signifies a pronounced class imbalance, making the task at hand a challenging unbalanced multiclass classification problem.

Unbalanced class distribution, as observed in this dataset, presents a unique set of challenges for machine learning and classification models. In our case, 'unacc' cars are overrepresented, while other categories may have fewer instances. This imbalance can lead to model bias and may impact the ability to accurately classify cars into the various evaluation categories. Therefore, addressing class imbalance becomes a critical aspect of this project. Techniques such as resampling, using different evaluation metrics, or exploring more advanced machine learning models tailored for imbalanced datasets may be necessary to ensure that the classification results are reliable and robust. Acknowledging and managing this class imbalance is vital in developing a trustworthy car evaluation system.

**6. Reference:**

* <https://archive.ics.uci.edu/dataset/19/car+evaluation>
* <https://towardsdatascience.com/learn-rabbitmq-for-event-driven-architecture-eda-e1e7377db2b>
* <https://www.scalablepath.com/data-science/exploratory-data-analysis>
* <https://www.geeksforgeeks.org/what-is-exploratory-data-analysis/>
* <https://www.digitalocean.com/community/tutorials/exploratory-data-analysis-python>