**TECH3200 Assessment 2 Dataset**

**Theme:** Australian Vehicle Prices

**Description:**

This dataset contains the latest information on car prices in Australia for the year 2023. It covers various brands, models, types, and features of cars sold in the Australian market. It provides useful insights into the trends and factors influencing the car prices in Australia. The dataset includes information such as brand, year, model, car/suv, title, used/new, transmission, engine, drive type, fuel type, fuel consumption, kilometres, colour (exterior/interior), location, cylinders in engine, body type, doors, seats, and price. The dataset has over 16,000 records of car listings from various online platforms in Australia.

**Key Features:**

* Brand: Name of the car manufacturer
* Year: Year of manufacture or release
* Model: Name or code of the car model
* Car/Suv: Type of the car (car or suv)
* Title: Title or description of the car
* UsedOrNew: Condition of the car (used or new)
* Transmission: Type of transmission (manual or automatic)
* Engine: Engine capacity or power (in litres or kilowatts)
* DriveType: Type of drive (front-wheel, rear-wheel, or all-wheel)
* FuelType: Type of fuel (petrol, diesel, hybrid, or electric)
* FuelConsumption: Fuel consumption rate (in litres per 100 km)
* Kilometres: Distance travelled by the car (in kilometres)
* ColourExtInt: Colour of the car (exterior and interior)
* Location: Location of the car (city and state)
* CylindersinEngine: Number of cylinders in the engine
* BodyType: Shape or style of the car body (sedan, hatchback, coupe, etc.)
* Doors: Number of doors in the car
* Seats: Number of seats in the car
* Price: Price of the car (in Australian dollars)

In the previous workshops, you have learned and seen some commonly used Python libraries,

such as Pandas, Matplotlib, NumPy, SciKit-Learn (SK-Learn) etc. Use the given dataset “**Australian Vehicle Prices.csv**”, and you need to choose two applications that you learnt to develop predictive models (e.g., to predict the car price based on the features and location. You need to explain the process in the context of machine learning algorithms as a presentation format. The assessment must be completed in class within a time limit.

For detailed instructions, please go back to the assessment 2 instruction PDF file on MyKBS

* **Part A: Part A: Simulation of Machine Learning (60 minutes)**
* **Part B: Part B: Evaluation and Presentation (30 minutes)**

**Please note this is a group assessment and you are required to complete Part A and B within the 3-hour workshop time of week 9 and submit by the end of the week 9 workshop.**