CDS 503: Machine Learning
Academic Session: Semester 2, 2022/2023

School of Computer Sciences, USM, Penang

LAB EXERCISE (LAB 7)

Lab 7 Exercise

The lab exercise uses auto mpg datasets. The data is a technical spec of cars.

Attributes Information:

- mpg (miles per gallon): continuous
- cylinders: multi-valued discrete
- displacement: continuous
- horsepower: continuous
- weight: continuous
- acceleration: continuous
- model year: multi-valued discrete
- origin: multi-valued discrete
- car name: string (unique for each instance)

We would like to categorize the cars into several groups based on how far a car can travel if you put just one gallon of petrol (mpg) in its tank.

1. Load the dataset
2. Normalize the dataset
3. Create a subset of the dataset that consists of mpg and displacement
4. Cluster the dataset (subset) using Hierarchical Agglomerative Clustering

---

Post your solution on Lab 07 Submission at **elearn@usm**. Make sure to include your name and lab# on the submission post.

Format: in .ipynb

The due date is **5th January 2023 5.00 pm**