

# Homework 5 (*due: Feb 24*)

## MACHINE LEARNING - COSC 4360

Department of Computer Science and Electrical Engineering

Spring 2025

### Exercises

Create a **New Project** for every exercise. Take a screenshot of the source code along with its output and place the **source code** and the **screenshot** in a **zipped folder** named **LastNameFirstName\_HW5**

#### Exercise 1

Given the following dataset: *vehicles.csv* as well as the supporting file *vehicles.names*, find the *five* most important **weighted coefficients** and print their names (drop the **make** column from the *data frame*). The *target* variable is **mpg**.

**Note:** You may use any *built-in* functions you wish.

#### Exercise 2

Based on the *five* most important features from Ex. 1, create a **6D plot** with **mpg** being assigned to **marker color**.

**Note:** Tutorial: [Multi-dimensional plots](#). The source code can be downloaded from [here](#).

#### Exercise 3

Predict the **mpg** given the following unscaled data point: 6, 163, 111, 3.9, 2.77, 16.45, 0, 1, 4, 4.

**Note:** You may use any *built-in* functions you wish.

#### Exercise 4 (*Optional*)

Given the following dataset: *materials.csv*, create a **4D plot**.

**Note:** Tutorial: [Multi-dimensional plots](#). The source code can be downloaded from [here](#).

#### Exercise 5 (*Optional*)

Given the following dataset: *materials.csv*, implement your **own** version of the **Principal Component Analysis** (PCA) algorithm. In addition, define your own function for **standardizing** the data. Finally, use **PCA=2** and create a **scatter plot** with the **two Principal Components**.

**Note:** Submit through **Canvas**