National University
o f   C o m p u t e r   &   E m e r g i n g   S c i e n c e s - I s l a m a b a d
C h i n i o t - F a i s a l a b a d   C a m p u s

Dr. Hashim Yasin

Department of Artificial Intelligence & Data Science.

AI3002 –Machine Learning
Assignment No. 1

Assignment Submission Guidelines:

1. Submit your assignment in **soft form (Code + Report)** within the due date and time. Soft form does not mean submitting photos of the hardcopy. Late submissions will result in a deduction of marks.
2. The **report** must include a discussion, comments, and a conclusion about your solution. Submitting without a report will result in a loss of full marks.
3. Name the zip or other folder/file that you submit using the following format: **ML_A1_RollNo_FirstName**.
4. Ensure that you solve each task of the assignment on your own.
5. You are allowed to do your assignment in **groups of a maximum of two members**.
6. There is no restriction on the programming language used for the tasks.
7. For programming tasks, you are NOT allowed to use any built-in functions or libraries for specific tasks.

_____

Question No. 1: Linear Regression & Gradient Descent

Download training dataset, which consists of the input X file (DataX) and the corresponding output Y file (DataY). The input X file (DataX) has three attributes: the living area, the number of bedrooms, and the number of floors, while the output Y file (DataY) represents the house prices in response to these attributes. There are $m = 50$ training examples. Perform the following tasks,

(a) Gradient Descent

Implement gradient descent algorithm with a learning rate = 0.02.

$$\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta_0, \ldots, \theta_n)$$

(b) Consider the closed-form solution to a least square fit given as under. Implement it in order to calculate the values of the parameters for the same dataset.

$$\boldsymbol{\theta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$$

(c) Visualize the data and the results of the linear regression graphically. Also, discuss the results in detail.

(d) Make comparison of both results obtained in case of (a) and (b).

Question No. 2: Logistic Regression

Implement the logistic regression in order to classify the houses into two classes, "Costly" and "Not Costly", using the same input data "DataX" and the classes in "ClassY" file provided in Question 1. The "ClassY" file contains two values: **1** for "Costly" class and **0** for "Not Costly" class.