National University

of Computer & Emerging Sciences-Islamabad
Chiniot-Faisalabad Campus

Dr. Hashim Yasin

Department of Computer Science.

CS4104 – Applied Machine Learning
Assignment No. 1

Instructions:

1. Submit your assignment in **soft** as well as in **hard form (Report)** within due date and time. Soft form does **not** mean the photos of the hardcopy. Late submission will result in deduction of marks.
2. Report should necessarily contain the discussion, comments and conclusion about the solution. Without report, you will not get full marks.
3. Mention your names and roll numbers clearly on your document.
4. Name your zip or other folder/file that you want to submit according to the following format: **AML_A1_RollNo_FirstName**
5. Try to solve the each task of the assignment by your own.
6. No excuse or resubmission is permissible.
7. Do your assignment in group of maximum two members.
8. There is no any restriction of the language for the programming tasks.
9. **In programming tasks, you are NOT allowed to use any built-in function of any library for specific tasks.**

_____

Question No. 1: Linear Regression & Gradient Descent

Download training dataset, which consists of the input X file and the corresponding output Y file. 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**

- Preprocess the data,
- 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: Activation Functions

Prove that the derivative of Logistic (Sigmoid) function

$$\sigma(x) = \frac{1}{1+e^{-x}} \qquad \text{is} \qquad \frac{d\sigma(x)}{dx} = \sigma(x).\,(1 - \sigma(x))$$

and of Tangent function

$$\tanh(x) = \frac{e^x + e^{-x}}{e^x - e^{-x}} \qquad \text{is} \quad \tanh'(x) = 1 - \tanh^2(x).$$

Question No. 3: 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.