AI4ICPS IIT Kharagpur	**IIT KHARAGPUR AI4ICPS I HUB FOUNDATION** **H**ands-on **A**pproach to **AI, Cohort-2, July – October 2024** **Assignment 3: Linear Models for Regression and Classification**

Due date: Friday 2nd August 2024, EOD – IST.

Important Instructions for submitting solutions

1. Submit the solution to all questions in the assignment should be submitted in a **single PDF file with not more than 500 words**.
2. Any plagiarism if detected will automatically attract **zero marks** for that assignment.
3. It is preferable if the **text of PDF file can be extracted** through a PDF extractor e.g. PyPDF. For example, pictures of handwritten text are not extractable, whereas PDF generated by MS Word, Latex, etc., are.
4. Exceptionally good solutions with extractable text may receive **special appreciation** from the teachers.

Question 1

Answer the following questions briefly, in 1-2 sentences each.

a. What does linear regression try to optimize?

b. Is it possible to use linear regression to represent quadratic equations? Explain with an example.

c. Why is it crucial to detect and remove outliers?

d. What is feature scaling? When is it required?

e. State two differences between linear regression and logistic regression.

f. Why is the Mean Square Error cost function unsuitable for logistic regression?

g. What can be inferred if the cost function initially decreases but then increases or gets stuck at a high value?

h. Describe two ways to perform multi-class classification using logistic regression.

Question 2

Consider a linear regression model with two variables: h(x) = w0 + w1.x1 + w2.x2; which has been initialized with the following weights: w0 = 0; w1 = 1; w2 = 1. Consider the learning rate alpha = 0.0002. You are given the following data:

x1	60	67	71	75	78
x2	22	24	15	20	16
y	140	159	192	200	212

Write the values of the weights after performing the gradient descent algorithm for **2 iterations**. Calculate the initial mean squared error before any iterations, and the final error after updating the weights for 2 iterations. Provide the values in tables like the following:

Answer:

	w0	w1	w2
After iteration 1			
After Iteration 2			

Initial Mean squared error	
Final Mean squared error	