LAHORE UNIVERSITY OF MANAGEMENT SCIENCES
Department of Electrical Engineering

EE 514 (CS 535) Machine Learning – Spring 2025
Quiz 4

Name: ______________________

Campus ID: ______________________

Total Marks: 10
Time Duration: 15 minutes

Question 1 (8 marks)

We consider a dataset with two features: study hours (x_1) and sleep hours (x_2), with the corresponding exam scores (y).

Study Hours (x_1)	Sleep Hours (x_2)	Exam Score (y)
5.0	10	75
6.0	12	78
6.5	13	85
7.0	14	88

(a) [1 mark] Write a linear model for predicting y using both features.

Solution: The linear model for predicting the exam score y using both study hours (x_1) and sleep hours (x_2) is:

$$y = \theta_0 + \theta_1 x_1 + \theta_2 x_2.$$

Here θ_0 is the intercept (bias term), θ_1 is the coefficient for study hours (x_1), and θ_2 is the coefficient for sleep hours (x_2).

(b) [2 marks] Formulate the matrix equation $Y = X\theta$, clearly defining Y, X, and θ.

Solution: The matrix equation $Y = X\theta$ is formulated as follows:

$$Y = \begin{bmatrix} 75 \\ 78 \\ 85 \\ 88 \end{bmatrix}, \quad X = \begin{bmatrix} 1 & 5.0 & 10 \\ 1 & 6.0 & 12 \\ 1 & 6.5 & 13 \\ 1 & 7.0 & 14 \end{bmatrix}, \quad \theta = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \end{bmatrix}$$

Here Y is the vector of exam scores, X is the design matrix with a column of 1's (for the intercept) and the features x_1 and x_2, and θ is the vector of parameters to be estimated.

(c) [1 mark] Write the least-squares objective function for estimating θ.

Solution: The least-squares objective function for estimating θ is:

$$L(\theta) = \frac{1}{2} \sum_{i=1}^{n} \left(y^{(i)} - (\theta_0 + \theta_1 x_1^{(i)} + \theta_2 x_2^{(i)}) \right)^2$$

In matrix form, this can be written as:

$$L(\theta) = \frac{1}{2}(Y - X\theta)^T (Y - X\theta)$$

The goal is to minimize $L(\theta)$ with respect to θ.

(d) [4 marks] Now consider using only one feature (study hours, x_1). Rewrite the model and find the two parameters of the model.

Solution: If we use only the study hours (x_1) as the feature, the linear model becomes:

$$y = \theta_0 + \theta_1 x_1$$

For this simplified model, the matrices are:

$$Y = \begin{bmatrix} 75 \\ 78 \\ 85 \\ 88 \end{bmatrix}, \quad X = \begin{bmatrix} 1 & 5.0 \\ 1 & 6.0 \\ 1 & 6.5 \\ 1 & 7.0 \end{bmatrix}, \quad \theta = \begin{bmatrix} \theta_0 \\ \theta_1 \end{bmatrix}$$

The least-squares solution is given by:

$$\theta = (X^T X)^{-1} X^T Y$$

$$X^T X == \begin{bmatrix} 4 & 24.5 \\ 24.5 & 152.25 \end{bmatrix}$$

$$X^T Y == \begin{bmatrix} 326 \\ 2011.5 \end{bmatrix}$$

$$\theta = (X^T X)^{-1} X^T Y == \begin{bmatrix} 40.2 \\ 6.67 \end{bmatrix}$$

Thus, the linear model using only study hours is:

$$y = 40.2 + 6.67 x_1$$

Question 2 (2 marks)

1. How does high bias affect the performance of a machine-learning model?
   a) It leads to poor performance on the training data but an excellent generalization.
   b) It results in the model failing to capture the patterns in the data, causing poor performance on both training and test sets.
   c) It increases the variance of the model, making it sensitive to small changes in data.
   d) It reduces the number of parameters in the model, improving flexibility.
2. Which approach can help reduce high variance in a model?
   a) Increasing the model complexity
   b) Decreasing the amount of training data
   c) Applying regularization techniques like L1 or L2
   d) Reducing the number of features used for training