Name: _____ Reg #: _____ Section: _____

National University of Computer and Emerging Sciences, Lahore Campus

Course: Bioinformatics Course Code: CS4054

Program: BS(Computer Science) Semester: Spring 2025 **Section: BCS-8A**

Duration: 30 Minutes **Total Marks: 10**
Paper Date: 24-April-2025 **Page(s): 1**

Exam: Quiz 3

Q1 Consider the following gene expression distance matrix from a microarray experiment for 5 genes. A biologist is trying to find out whether these 5 genes can be separated into clusters based on their behavior in the experimental conditions.

	$Gene_1$	$Gene_2$	$Gene_3$	$Gene_4$	$Gene_5$
$Gene_1$	1.00	0.90	0.10	0.65	0.20
$Gene_2$	0.90	1.00	0.70	0.60	0.50
$Gene_3$	0.10	0.70	1.00	0.40	0.30
$Gene_4$	0.65	0.60	0.40	1.00	0.80
$Gene_5$	0.20	0.50	0.30	0.80	1.00

Use average link to update your distance matrix. Indicate the data points belonging to each cluster. Create a dendrogram to visualize the results. If the biologist wants to group the genes into two clusters, identify the genes belonging to each cluster. [10 marks]