COMSATS University Islamabad, Attock Campus

Department of Computer Science

Program: BS(AI)

Spring 2023: Assignment 3		**Course: Machine Learning Fundamentals AIC354**
Due Date: 01/06/2023	**Dated: 29/05/2023**	**Marks: 15**
Name:	**Reg. No.:**	
Note:- Don't write anything on Question Paper except your Name & Reg. No.		

The student will submit the hard copy to CR before due timing. The CR is responsible to submit the class assignments till due date in office.

Question#1 [CLO-3(SO-2,4)] [5 Marks]

Make clusters of the given below data points by applying the K Means clustering technique. Apply the Elbow method and find out the best number of clusters in k=2,3,4,5,6.

Data Points → Features ↓	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
X	1	2	3	4	5	1	2	3	4	5
Y	3	4	1	2	1	4	1	2	5	4

For the initialization of mean values, the values are given for different number of clusters:

k=2	m1=(1,1)	m2=(1,4)	-	-	-	-
k=3	m1=(1,1)	m2=(1,4)	m3=(2,5)	-	-	-
k=4	m1=(1,1)	m2=(1,4)	m3=(2,5)	m4=(4,1)	-	-
k=5	m1=(1,1)	m2=(1,4)	m3=(2,5)	m4=(4,1)	m5=(5,5)	-
k=6	m1=(1,1)	m2=(1,4)	m3=(2,5)	m4=(4,1)	m5=(5,5)	m6=(3,3)

Question#2 [CLO-3(SO-2,4)] [5 Marks]

Make dendrogram of clusters by applying the Hierarchical Agglomerative Single Linkage Clustering technique.

How many clusters are there in the dendrogram for threshold (distance) = 3

The dataset is given by:

Data Points → Features ↓	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
X	1	2	3	4	5	1	2	3	4	5
Y	3	4	1	2	1	4	1	2	5	4

Question#3 [CLO-3(SO-2,4)] [5 Marks]

Convert the following data to PCA domain:

//Note that the data has the same range, there is no need to standardize the data.

	X1 (feature 1)	X2 (feature 2)	PCA1	PCA2
Sample 1	**1**	**2**	**?**	**?**
Sample 2	**2**	**4**	**?**	**?**
Sample 3	**1**	**1**	**?**	**?**
Sample 4	**2**	**2**	**?**	**?**