-

There are THREE questions. Answer any TWO. Figures in the right-hand margin indicate full marks.

1. Using the Fuzzy C-Means clustering algorithm cluster the following data into two.   [12.5]
    (4,5), (4,8) and (3,7)
    Show an iteration of the algorithm.
2. Draw a hierarchy using the hierarchical clustering algorithm   [12.5]
    (1,2), (1,8), (2, 8), (1,4) and (4, 4)
3.   Using the K-Mean clustering algorithm cluster the following data into three.   [12.5]
   (1,2), (1, 5), (2, 8), (2,4) and (6, 8)
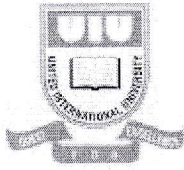
There are FIVE questions. Answer any THREE. Figures in the right-hand margin indicate full marks.

1. a) Define Machine Learning.   1.0
    b) Distinguish between unsupervised and semisupervised learning with an example.   2.0
    c) Write a short note on: a) Close test, b) Open test and c) Generalization.   3.0
    d) Mention some real life applications of Machine Learning Tools.   2.0
    e) Define overfitting. Why is overfitting not good?   2.0

2. Obtain two clusters of the following food items using the k-Mean clustering algorithm   10

Food item #	Protein content, P	Fat content, F
Food item #1	1.1	60
Food item #2	8.2	20
Food item #3	4.2	35
Food item #4	1.5	21
Food item #5	7.6	15
Food item #6	2.0	55
Food item #7	3.9	39

3. Draw a hierarchy using the hierarchical clustering algorithm based on the proximity of the   10
    following Italian Cities given in the Map
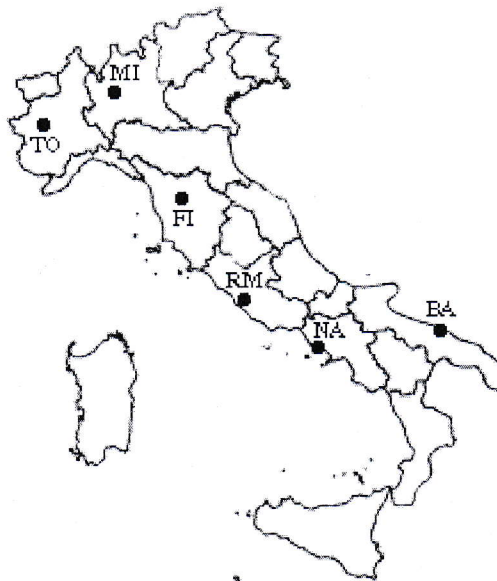
	BA	FI	MI	NA	RM	TO
BA	0	662	877	255	412	996
FI	662	0	295	468	268	400
MI	877	295	0	754	564	138
NA	255	468	754	0	219	869
RM	412	268	564	219	0	669
TO	996	400	138	869	669	0

4. a) Write down the properties of Self Organizing Map (SOM).　　　　　　　　　10

   b) Cluster the following areas of Bangladesh into two classes based on ART, DANCE, and MUSIC features using the SOM algorithm.

   Values for each feature ART, DANCE and MUSIC are:

Values	Rank
Excellent	5
Very Good	4
Good	3
Average	2
Poor	1

   Geographical areas are:

   　　　　Mymensingh: (Very Good, Good, Excellent)
   　　　　Chittagong: (Excellent, Average, Good)
   　　　　Dhaka: (Excellent, Excellent, Excellent)
   　　　　Rajshahi: (Good, Average, Poor)
   　　　　Sylhet: (Average, Good, Excellent)

   Assume the following initial weight matrix and learning rate to solve the problem.

   　　W= 　0.3　0.5　0.7
   　　　　　0.5　0.7　0.4

   　　Learning rate=0.6

   Complete one epoch of the SOM.

5. a) What are the merits of Fuzzy C-Means algorithm over the k-Mean algorithm?　　　10

   b) In Image Processing Red(R), Green(G) and Blue(B) are defined as follows:

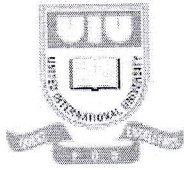   R=(255, 0, 0)
   G=(0, 255, 0)
   B=(0, 0, 255)

   Segment the following pixels into probabilistic value of the above principal components of colors based on the Fuzzy C-Means Clustering algorithm.

   　　　　Pixel-I: (145, 200, 255)
   　　　　Pixel-II: (200, 255, 145)
   　　　　Pixel-III: (255, 145, 200)

   Find one iteration of the algorithm.

United International University (UIU)

Dept. of Computer Science & Engineering (CSE)
Mid term Year: 2012 Trimester: Fall Program: MSCSE
Course: CSE 6023 Machine Learning, Marks: 30, Time: 2 hours

There are SEVEN questions. Answer any FOUR. Figures in the right-hand margin indicate full marks.

1.	a)	Define Machine Learning.	7.5
	b)	Distinguish between supervised and unsupervised learning with an example.	
	c)	Write a short note on: a) Close test, b) Open test and c) Generalization.	

| 2. | Cluster the following geographical areas into two classes based on ART, DANCE, and MUSIC features using the k-Mean clustering algorithm. | 7.5 |

Values for each feature ART, DANCE and MUSIC are:

Values	Rank
Excellent	5
Very Good	4
Good	3
Average	2
Poor	1

Geographical areas are:

Mymensingh: (Very Good, Good, Excellent)
Chittagong: (Excellent, Average, Excellent)
Dhaka: (Excellent, Excellent, Excellent)
Rajshahi: (Average, Average, Excellent)
Sylhet: (Average, Good, Excellent)

| 3. | Solve the QUES. 2 using the modified k-Mean algorithm. | 7.5 |

| 4. | Draw a hierarchy using the hierarchical clustering algorithm for the data given in QUES. 2. | 7.5 |

| 5. | a) | Write down the properties of Self Organizing Map (SOM). | 7.5 |
| | b) | Solve the QUES. 2 using the SOM algorithm assuming the following initial weight matrix and learning rate. | |

W= 0.3 0.4 0.7
    0.5 0.7 0.3

Learning rate=0.6
Complete one epoch of the SOM.

| 6. | a) | What are the merits of Fuzzy C-Means algorithm over the k-Mean algorithm? | 7.5 |
| | b) | In human body of a man some cells are affected by CANCER, some are affected by TUMOR and some are affected by both. Using the Fuzzy C-Means algorithm find the two class probabilities for the following three cells. | |

Feature Vector: (CANCER, TUMOR)
Values for each feature CANCER and TUMOR are:

Values	Rank
Highly Effected	5
Average Level	3
Normal Level	1
Not Effected	0

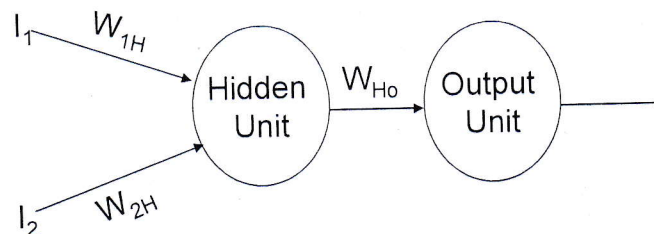Cell-I: (Average Level, Highly Effected)
Cell-II: (Highly Effected, Normal Level)
Cell-III: (Average Level, Normal Level)
Assume, the value of $\epsilon = 0.4$

---

7. Find the output value for the following test pattern using the following Multilayer Neural Network (MLN) architecture.   **7.5**

Input (I_1)	Input(I_2)	Output
0.8	0.9	1.0
0.4	0.6	unknown

Initial Weight:

$W_{1H} = -0.30$

$W_{2H} = +0.30$

$W_{HO} = +0.25$

Initial ΔWeight:

$\Delta W_{1H} = +0.35$

$\Delta W_{2H} = -0.32$

$\Delta W_{HO} = +0.28$

Momentum: $\mu = 0.55$ Learning rate: $\eta = 0.01$