


Course No: SWE 5437
 Course Title: Artificial Intelligent.
 Date: 08/ 01/ 2015
 No. of Questions: 4
 Time: 2 H.
 Using Calculator (Yes)

University of Palestine

 Final Exam
 1st Term 2014/2015
 Total Grade: 60

Instructor Name: Eng. M. Timraz
 Student No.: _____
 Student Name: _____
 College Name: Engineering
 Dep. / Specialist: Software
 Using Dictionary (No)

First Question **No. of Branches (2)** **(15/60)**

Q1 B1: Put True or False in the specific table below. **(10/15)**

- (1) In non-exclusive clustering, points may belong to multiple clusters.
- (2) The symbol table is computed according to a frequency table and it has exactly the same dimension as the symbol table.
- (3) A good clustering method will produce high quality clusters with low intra-class similarity and high inter-class similarity.
- (4) Cluster-based categorization is more complicated than direct categorization.
- (5) A cluster is a set of points such that any point in a cluster is closer (or more similar) to every other point in the cluster than to any point not in the cluster.
- (6) To treat the ratio scaled variables, it is very good to use the interval scaled variables.
- (7) It is necessary to convert the interval data to ordinal data.
- (8) Distances and similarities play an important role in cluster analysis.
- (9) A binary attribute is an attribute that has exactly one possible value.
- (10) In an asymmetric variable, one of its values carries more importance than the other.

<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	<u>7</u>	<u>8</u>	<u>9</u>	<u>10</u>

Q1 B2 **(5/15)**

Choose the correct answer (a, b, c or d) and then put it in the specific table in the model answer note.

- (1) All of the following are requirements for good clustering except one.
 - a. Insensitive to order of input records.
 - b. Able to deal with noise and outliers.
 - c. Able to deal with high dimensionality.
 - d. must be lower than inter cluster distance.
- (2) Enumerate all possible ways of dividing the points into clusters and evaluate the goodness of each potential set of clusters by using the given
 - a. Maximum distance idea.
 - b. Cluster Types.
 - c. Objective function.
 - d. High dimensionality.
- (3) there are three cases according to how missing values can occur in data sets except one
 - a. Missing values occur in several variables.
 - b. Missing values occur in a number of records.
 - c. Missing values occur randomly in variables and records.

Course No: SWE 5437
 Course Title: Artificial Intelligent.
 Date: 08/ 01/ 2015
 No. of Questions: 4
 Time: 2 H.
 Using Calculator (Yes)

University of Palestine



Final Exam
 1st Term 2014/2015
 Total Grade: 60

Instructor Name: Eng. M. Timraz
 Student No.: _____
 Student Name: _____
 College Name: Engineering
 Dep. / Specialist: Software
 Using Dictionary (No)

d. None of the above.

(4) Jaccard coefficient method used to calculate.....

- a. the dissimilarity between objects.
- b. the dissimilarity between clusters.
- c. the distances between attributes.
- d. all of the above.

(5) Data scales are.....

- a. qualitative scales.
- b. quantitative scales.
- c. a and b.
- d. non of the above.

<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>

Second Question

No. of Branches (3)

(15/60)

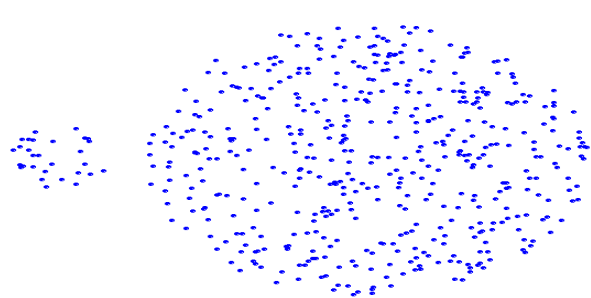
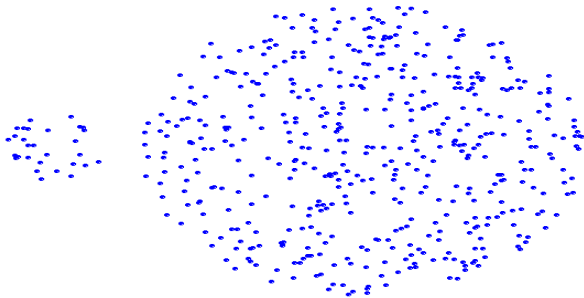
Q2 B1

(5/15)

For the following data set show the differences when you apply MIN, MAX and Group, Distance between centroids and Average techniques

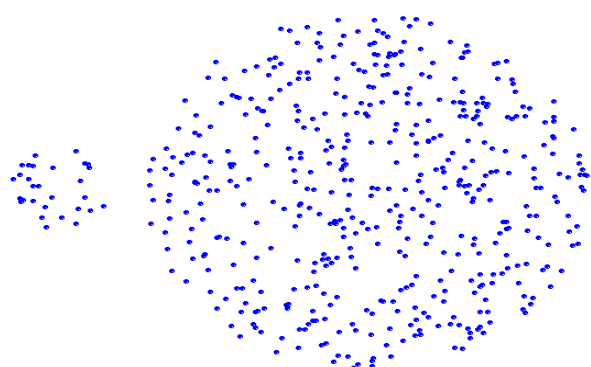
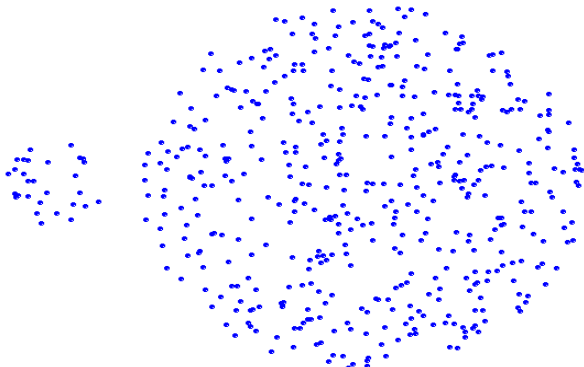
MIN:

MAX:



Group Average:

Distance between centroids





Q2 B2 (10/15)

Chose the best answer from the table "A" for the suitable of the table "B".

No.	"A"	"B"	No.
1	Well separated clusters.		
2	Euclidean distance		
3	Density based clusters		
4	Scatter Matrix		
5	Mahalanobis Distance	$d(x, y) = \left(\sum_{j=1}^d (x_j - y_j)^2 \right)^{\frac{1}{2}}$	
6	Manhattan Distance	$c_{rs} = \frac{1}{n} \sum_{i=1}^n (x_{ir} - \bar{x}_r)(x_{is} - \bar{x}_s).$	
7	Contiguity based clusters.	$d(x, y) = \sum_{k=1}^d x_k - y_k $	
8	Maximum Distance	$d(x, y) = \max_{1 \leq k \leq d} x_k - y_k .$	
9	Covariance Matrix	$d(x, y) = \sqrt{(x - y)\Sigma^{-1}(x - y)^T},$	
10	Centre based clusters.	$M_w(C) = \sum_{i=1}^k \sum_{x \in C_i} (x - z_i)^T (x - z_i),$	



Q3 B1

(15/60)

Rearrange the persuade code for the following Algorithm (conventional k-means algorithm). **(10/15)**

No.	Item	No.
1	Require:	
2	$ni = \arg \min_{1 \leq j \leq k} di j ;$	
3	Number of Clusters k ,	
4	until no further changes of cluster membership occur in a complete iteration	
5	Recompute the cluster means of any changed clusters above;	
6	Assign case i to cluster ni ;	
7	{ Ci is the i th cluster }	
8	{ Iteration Phase }	
9	Data set D ,	
10	repeat	
11	$(C1, C2, \dots, Ck) =$ Initial partition of D .	
12	$di j =$ distance between case i and cluster j ;	
13	Output results.	
14	Dimensions d :	
15	{ 1. Initialization Phase }	

Course No: SWE 5437
Course Title: Artificial Intelligent.
Date: 08/ 01/ 2015
No. of Questions: 4
Time: 2 H.
Using Calculator (Yes)

University of Palestine



Final Exam
1st Term 2014/2015
Total Grade: 60

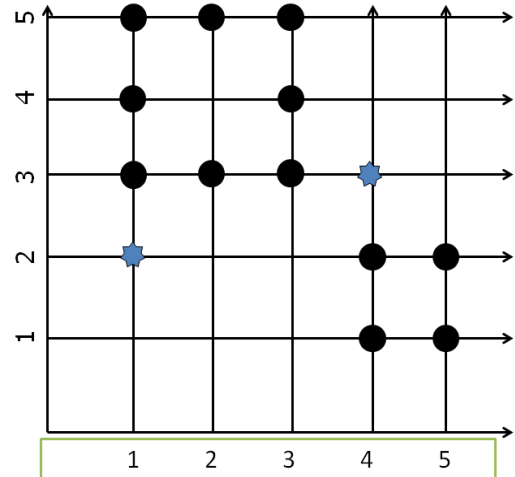
Instructor Name: Eng. M. Timraz
Student No.: _____
Student Name: _____
College Name: Engineering
Dep. / Specialist: Software
Using Dictionary (No)

Q3 B2

(5/15)

Apply K-means algorithm by drawing only for each step to know how much iterations will the star prototypes converging?

Please draw each step.



Course No: SWE 5437
Course Title: Artificial Intelligent.
Date: 08/ 01/ 2015
No. of Questions: 4
Time: 2 H.
Using Calculator (Yes)

University of Palestine



Final Exam
1st Term 2014/2015
Total Grade: 60

Instructor Name: Eng. M. Timraz
Student No.: _____
Student Name: _____
College Name: Engineering
Dep. / Specialist: Software
Using Dictionary (No)

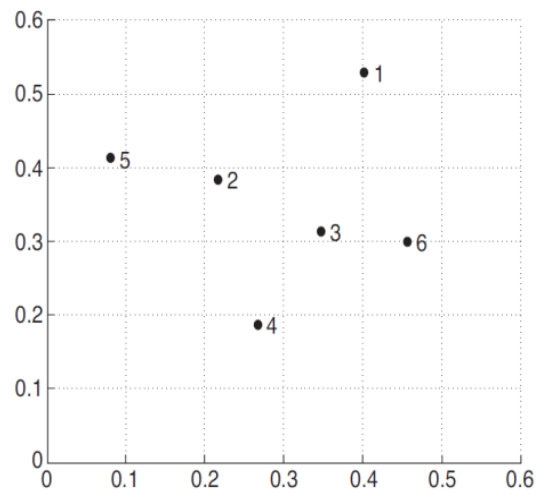
Q4 B1

(15/60)

Two dimensional data set with 7 points as in the showing figure.
Represent the Hierarchical Dendrogram using the following methods?

1. The Single link method.
2. The Complete link method.

Note: first of all you have to calculate the Euclidean distance matrix



Course No: SWE 5437
Course Title: Artificial Intelligent.
Date: 08/ 01/ 2015
No. of Questions: 4
Time: 2 H.
Using Calculator (Yes)

University of Palestine



Final Exam
1st Term 2014/2015
Total Grade: 60

Instructor Name: Eng. M. Timraz
Student No.: _____
Student Name: _____
College Name: Engineering
Dep. / Specialist: Software
Using Dictionary (No)

Good Luck