Hajee Mohammad Danesh Science and Technology University, Dinajpur
Department of Computer Science and Engineering (CSE)
B.Sc. in CSE
Semester Final Examination 2016 (Jul-Dec)
Level 4 Semester 2, Course Code: CSE455, Credit: 3.0
Course Title: Pattern Recognition

Time: 3 hours | Total Marks: 90

NB: Figures in the right margin indicate full marks. Parts of the same question should be answered together and in the same sequence.

Section-A

Answer any Three

			marks
1.	(a)	What are the primary differences between *pattern recognition*, *machine learning*, and *data mining*? Which tasks are important in each of these areas?	2+2
	(b)	A common pattern recognition problem involves the recognition of 2-D shapes. For simplicity, consider a world of 4 classes: squares, triangles (each angle= 60°), hexagon (each angle is 120°) and circles. Propose a feature vector that can differentiate between any of the previous 4 shapes. Justify the appropriateness of it for the classification task.	5
	(c)	What is meant by *dimensionality reduction*? Why is it important? Describe the strategy of attribute subset selection used for dimensionality reduction.	1+1+2
	(d)	Given two objects represented by the tuples (22, 1, 42, 10) and (20, 0, 36, 8): i) Compute the Manhattan distance between the two objects. ii) Compute the Minkowski distance between the two objects, using q=4.	1+1
2.	(a)	For the vector x=(0,-1,0,1), and y=(1,0,-1,0), calculate the cosine and jaccard similarity.	2
	(b)	Suppose a hospital tested the age and body fat data for 18 randomly selected adults with the following result:	2*4

age	23	23	27	27	39	41	47	49	50
%fat	9.5	26.5	7.8	17.8	31.4	25.9	27.4	27.2	31.2

i) Calculate the mean, median, and standard deviation of %fat.
ii) Draw the boxplot for age.
iii) Normalize the two variables based on z-score normalization.
iv) Calculate the correlation coefficient. Are these two variables positively or negatively correlated?

	(c)	How can sampling be used as a data reduction technique? What are the differences among simple random sample, cluster sample , and stratified sample?	1+3
	(d)	What do you mean by spatial data?	1
3.	(a)	Describe the basic steps that must be followed in order to develop a pattern recognition task.	3
	(b)	What motivates the attempts to "clean" the training set?	3
	(c)	In real-world data, tuples with missing values for some attributes are a common occurrence. Describe various methods for handling this problem.	3
	(d)	Discuss with example performance of k-nn classification when (i) k is very small (ii) k is large.	1.5+1.5
	(e)	Find the centroid and radius, and diameter for the following set of patterns: (1, 1), (1, 3), (1, 4), (2, 2), (2, 3)	3
4.	(a)	What are outliers? How might you determine outliers in the data?	3
	(b)	Both Nearest Neigbour Algorithm and Square Error Clustering algorithm are partitional algorithm. Describe the main procedure of each of the algorithms.	3

(c) Suppose you have the following training examples, described by three attributes, $x1$; $x2$; $x3$, and labeled by classes $c1$ and $c2$. 8
Using the *3-NN* algorithm, decide whether new pattern p=[4,3,3] should belong to $c1$ or $c2$.

$x1$	$x2$	$x3$	Class
2	0	3	$c1$
3	1	2	$c1$
2	1	3	$c1$
0	5	0	$c2$
1	4	0	$c2$

(d) What is understood by the *curse of dimensionality*? 1

Section B
Answer any Three

marks

1. (a) What is the linear discriminate function? Explain how it can be used to find a linear classifier to discriminate between two classes? 1+4
(b) When do we need non-linear classifiers? How can SVM be used to handle this classifier? 1+3
(c) Following classifier is generated for the same training set, which has 100 instances. It has the following confusion matrices .Calculate the (i) Precision , (ii) Accuracy , and (iii) error rate 3*1.5= 4.5

		Predicted class	
		+	-
Actual class	+	50	10
	-	10	30

(d) What are inductive hypotheses and deductive hypotheses? 1.5

2. (a) How is non-separability of data handled in SVM learning? 5
(b) A major problem with the single link algorithm is that clusters consisting of long chains may be created. Describe and illustrate this concept. 4
(c) What is the gradient descent method? 2
(d) Suppose $f(x, y) = 3x^2 + 2y^3 - 2xy$. Find the minimum stationary points of this function. 4

3. (a) How Lagrange multiplier is used to solve constrained optimization problem? What is the role of it for SVM? 2+2
(b) Write short note on : 5*2 =10
   i) ROC Curves
   ii) Cross validation
   iii) Chi-square Test
   iv) PCA
   v) Overfitting
(c) What are the basic differences between supervised and unsupervised learning? 1

4. (a) Show that the weight vector w is orthogonal to the decision boundary in d dimensional space, where $d \geq 2$. 3
(b) In Support Vector Machines, what is a support vector? What is meant by the *margin* to be maximized? 2+2
(c) Write down the mathematical expression that defines a polynomial classifier 2
(d) Use a single-link clustering to obtain 3 cluster the following set of points: A = (1, 1); B = (2, 2); C = (2, 1); D = (3, 1); E = (4, 4); What is the dendrogram of this? 6