

**ISLAMIC UNIVERSITY OF TECHNOLOGY (IUT)**  
**ORGANISATION OF ISLAMIC COOPERATION (OIC)**

**Department of Computer Science and Engineering (CSE)**

SEMESTER FINAL EXAMINATION

WINTER SEMESTER, 2011-2012

DURATION: 3 Hours

FULL MARKS: 150

**CSE 4541: Machine Learning**

Programmable calculators are not allowed. Do not write anything on the question paper.

There are **8 (Eight)** questions. Answer any **6 (Six)** of them.

Figures in the right margin indicate marks.

1. a) Prove that "Gradient Descent algorithm can converge to local minimum even with fixed learning rate  $\alpha$ ". 10
- b) How can you make sure that your implementation of Gradient Descent algorithm is working correctly? 8
- c) Describe the effect of choosing learning rate  $\alpha$  in Gradient descent algorithm. 7
2. a) Explain the followings for logistic regression: 3+5+4
  - i. Hypothesis.
  - ii. Cost Function.
  - iii. Decision boundary.
- b) How does logistic regression draw non linear decision boundary? 8
- c) Explain logistic regression for multiclass classification. 5
3. a) What are the measuring criteria for selecting best split in decision tree induction? 5
- b) Which attribute is the best for splitting the following dataset? 20

RID	age	income	student	credit_rating	Class: buys_computer
1	youth	high	no	fair	no
2	youth	high	no	excellent	no
3	middle_aged	high	no	fair	yes
4	senior	medium	no	fair	yes
5	senior	low	yes	fair	yes
6	senior	low	yes	excellent	no
7	middle_aged	low	yes	excellent	yes
8	youth	medium	no	fair	no
9	youth	low	yes	fair	yes
10	senior	medium	yes	fair	yes
11	youth	medium	yes	excellent	yes
12	middle_aged	medium	no	excellent	yes
13	middle_aged	high	yes	fair	yes
14	senior	medium	no	excellent	no



4. a) Suppose you have already trained a Back Propagation Network (BPN) so that all connection weight are adjusted. Describe how you can predict class label for a new test data. 12
- b) Explain the method of weight adjustment of interconnections in BPN. 13
5. a) Why the Support Vector Machine (SVM) is called 'large margin classifier'? 10
- b) Explain the mechanism behind SVM with Gaussian kernels. 15

6. Suppose you have the following document term index:

d1:	1	2		
d2:	1	3	4	5
d3:	2	3	4	6
d4:	1	2	3	4
d5:	1	2	3	6

where  $D_i = i^{th}$  Document and 1,2,3,4,5,6 are terms present in documents.

d1: 1 2 means Document 1 (d1) contains term 1 and 2.

- a) Apply an agglomerative clustering procedure to the collection. Consider the Dice's coefficient measure as a measure of similarity between documents, 20

$$D(d_i, d_j) = 2 \frac{|d_i \cap d_j|}{|d_i| + |d_j|}$$

- b) Draw the resulting Dendrogram. 5

7. a) Determine the clusters in the following data set, using K means clustering algorithm: 20

Object	attribute 1 (X): weight index	attribute 2 (Y): pH
Medicine A	1	1
Medicine B	2	1
Medicine C	4	3
Medicine D	5	4

- b) What are the strength and weakness of K means clustering algorithm? 5

8. a) Draw the diagram of a neural network that solve XOR problem. 10
- b) Explain the mechanism behind Anomaly detection algorithm. 15