B.Sc. Engg. /HD CSE 5th Semester (16) 22 February 2012 (Morning)

ISLAMIC UNIVERSITY OF TECHNOLOGY (IUT)
ORGANISATION OF ISLAMIC COOPERATION (OIC)
Department of Computer Science and Engineering (CSE)

MID SEMESTER EXAMINATION **WINTER SEMESTER, 2011-2012**

DURATION: 1 Hour 30 Minutes **FULL MARKS: 75**

CSE 4541: Machine Learning

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

There are **4 (four)** questions. Answer any **3 (three)** of them.

Figures in the right margin indicate marks.

---

1. a) Explain the following with examples 4
   i. Supervised learning.
   ii. Unsupervised learning.

   b) Describe the method of parameter optimization with vectorized implementation for linear regression with single features. 10

   c) "Cost function plays an important role in error minimization for training dataset" Explain this statement in case of regression. 6

   d) When the feature normalization is used in linear regression with multiple variables? 5

2. a) Explain the effect of regularization in case of logistic regression. 7

   b) Compare and contrast between Gradient Descent and Normalized Equation for linear regression. 6

   c) Define the term "over fitting" and "under fitting". What are ways to prevent over fitting problem? 4+8

3. a) Using Entropy, find the best splitting criterion for the following dataset (figure 1). 20

age	income	student	credit_rating	Class: buys_computer
youth	high	no	fair	no
youth	high	no	excellent	no
middle_aged	high	no	fair	yes
senior	medium	no	fair	yes
senior	low	yes	fair	yes
senior	low	yes	excellent	no
middle_aged	low	yes	excellent	yes
youth	medium	no	fair	no
youth	low	yes	fair	yes
senior	medium	yes	fair	yes
youth	medium	yes	excellent	yes
middle_aged	medium	no	excellent	yes
middle_aged	high	yes	fair	yes
senior	medium	no	excellent	no

Figure 1

   b) What are the measures for selecting the best split for decision tree induction? 5

4. Explain the following of Back propagation Neural Network for multiclass classification. 4+6+6+9

a) Network representation
b) Cost function
c) Weight adjustment
d) Back propagation algorithm