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

MID SEMESTER EXAMINATION **WINTER SEMESTER, 2019-2020**

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

CSE 6257: Advanced Pattern Recognition

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

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

Figures in the right margin indicate marks.

---

1. a) Why is the issue of generalization important while learning the classification model? Explain with necessary figures. 5

   b) In the multi-category case, a set of samples is said to be linearly separable if there exists a linear machine that can classify them all correctly. If any samples labelled ω_i can be separated from all others by a single hyperplane, we shall say the samples are totally linearly separable. Show that totally linearly separable samples must be linearly separable, but that the converse need not be true. 10
   Note: Use example classes in the 2D feature space.

   c) Consider the three-category linear machine with discriminant functions, where i represents class: $g_i(x) = w_i^t x + b_i, i = 1,2,3.$ 5+5

   i. For the special case where x is two-dimensional vector and the bias weights b_i are zero, sketch the weight vectors with their tails at the origin, the three lines joining their heads, and the decision boundaries.

   ii. Do the decision boundaries change when a constant vector c is added to each of the three weight vectors? Justify your answer.

2. a) Suppose the training data in a 2D feature space in a two-class problem is given as shown in Table 1: 20

Table 1

Class	$[x_1\ x_2]^T$
ω_1	[1 1] [2 0] [-1 1]
ω_2	[1 0] [0 0] [-1 0]

Derive the equation of the decision boundary to linearly separate them. Show all necessary calculations along with the choice of your cost function.

   b) In a neural network, why do you need a nonlinear activation function? 5

3. a) Define the log-loss function L and cost function J for binary classification problem with m examples. Derive the derivatives of the cost function J with respect to z (total input) and w (weight vector) for a logistic regression unit. 8
   Note: Use vectorization notation. Assume standard notations and symbols.

   b) In case of a multi-layer neural network, what happens if all weights are initialized to zero? 5

   c) For a deep neural network, show all necessary calculations to find the final derivative of the loss function L with respect to a (output) for the $(l\text{-}1)^{th}$ layer. 12