

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

MID SEMESTER EXAMINATION

SUMMER SEMESTER, 2019-2020

DURATION: 1 Hour 30 Minutes

FULL MARKS: 75

CSE 4621: Machine Learning

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

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

Figures in the right margin indicate marks.

Instructions:

- i. **Closed Book:** No books (hardcopy or softcopy), notes, slides, cheat-sheets are allowed. No internet browsing.
- ii. Use standard loose pages, e.g., Legal or A4 (recommended).
- iii. Write your Student ID, Course Code, Course Name, Exam Type, Date, & Semester at the top of the first page.
- iv. All other pages should include student ID & page number at the top.
- v. Scan all pages into a single PDF file names as **StudentID_CSE4621**.
- vi. Make sure images are clear in contrast and intelligible.
- vii. Upload your script into the Google classroom as Assignment submission.
- viii. You will be given extra 15 minutes to upload at the end of the exam.
- ix. In case of difficulties during submission, contact the invigilator, and if approval is given submit through email at **hasanul@iut-dhaka.edu**
- x. For other technical difficulties contact the invigilator of the room or the course teacher.
- xi. No break is allowed during exam time.

1. a) Why are machine learning algorithms important for solving problems? Differentiate between the major types of learning approaches. 3+5
- b) When a function is called convex? Show that for a convex function the local optimum is also global. 1+4
- c) Suppose your linear regression model is suffering from high bias. What can you incorporate to reduce bias? 5
- d) Consider the following truth table for a Boolean function, f . 15+2

Table 1: Truth Table

| x_1 | x_2 | $y=f(x_1, x_2)$ |
|-------|-------|-----------------|
| 0 | 0 | 0 |
| 0 | 1 | 1 |
| 1 | 0 | 0 |
| 1 | 1 | 0 |

Find the weight parameters a logistic regression model representing the aforesaid Boolean function. Define your cost function and update-equations for the stochastic gradient descent algorithm. Select your initial weight parameter values randomly within the range $[-2, 2]$, and meticulously set your learning rate. Show all required calculations until convergence. Compute the mean square error (MSE) of your model.

2. a) The XOR function (exclusive or) returns true (1) only when one of the arguments is true (1) and another is false (0). Otherwise, it returns always false (0). Do you think it is possible to implement this function with a single logistic regression unit (i.e., neuron of an artificial neural network)? Justify your answer. 6

- b) In a multi-layer neural network, why do you need nonlinear activation functions in hidden layers? List two different activation function along with their mathematical definitions. 6+2
- c) Suppose the following figure represents the forward calculation involved in a single logistic regression unit with a logistic function, σ . 6

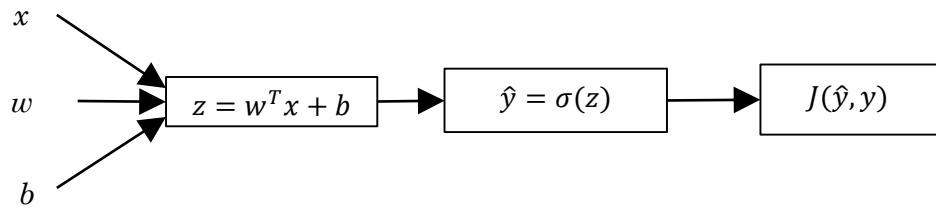


Figure 1.

Here, $J(\hat{y}, y)$ is the log-loss cost function. Solve $\frac{dJ}{dz} = \hat{y} - y$.

[Hints: Chain rule in calculus.]

3. a) “Early stopping (i.e., before reaching the global optimum) an iterative optimization method such Gradient Descent is a form of Regularization” – Do you agree or disagree with this statement. Justify your choice. 7
- b) Why do we need a Regularization term in the cost function? Compare between L1 and L2 Regularization. 2+6
- c) Given, the ridge regression cost function $J(\theta)$ as below: 5

$$J(\theta) = \frac{1}{2} \sum_{i=1}^m \left(\theta^T x^{(i)} - y^{(i)} \right)^2 + \frac{\lambda}{2} \|\theta\|^2 \quad \text{where } \lambda > 0.$$

Derive the vectorized closed form expression for the weight vector θ which minimizes the ridge regression cost function $J(\theta)$.