



**Cairo University**  
**Faculty of Computers and Artificial Intelligence**



**Midterm Exam**

**Department: Bioinformatics**

**Course Name: NN and Learning Machine**

**Course Code: BIO451**

**Instructor(s): Dr. Hanaa Bayomi**

**Name:.....**

**Date: 19/11/2019**

**Duration: 1 hour**

**Total Marks: 20**

**ID:.....**

**تعليمات هامة**

- حيازة التليفون المحمول مفتوحا داخل لجنة الامتحان يعتبر حالة غش تستوجب العقاب وإذا كان ضروري الدخول بالمحمول فيوضع مغلق في الحقيبة.
- لا يسمح بدخول سماعة الأذن أو البلوتوث.
- لايسمح بدخول أي كتب أو ملازم أو أوراق داخل اللجنة والمخالفة تعتبر حالة غش.

**Question 1**

**[12 marks]**

**- Answer the following Questions:**

a. [2 marks] Compare between Supervised and unsupervised Learning techniques?

- Supervised learning
  - Learning from labelled data
  - Classification, Regression, Prediction, Function Approximation
- Unsupervised learning
  - Learning from unlabelled data
  - Clustering, Visualization, Dimensionality Reduction

b. [1 marks] In the gradient descent algorithm, how can we choose the learning rate value?

By try and error , example : try 1, 0.3, 0.1, 0.03, 0.01 etc and plot the learning curve to understand whether the value should be increased or decreased.

c. [1 mark] What is an epoch?

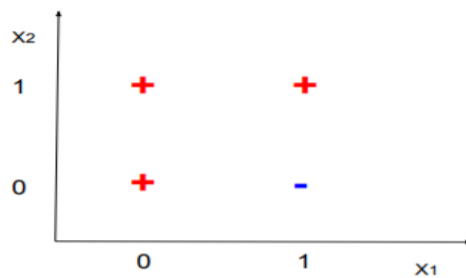
**Answer: An epoch is when all of the data in the training set is presented to the neural network once.**

- d. [2 marks] Can you represent the following Boolean function with a single unit from a neural network? (yes,No), explain why in 1-2 sentences.

**Solution: Yes, you can represent this function with a single unit from a neural network, since it is linearly separable**

| A | B | f(A,B) |
|---|---|--------|
| 1 | 1 | 0      |
| 0 | 0 | 0      |
| 1 | 0 | 1      |
| 0 | 1 | 0      |

- e. In the following Figure, we show three positive samples (“+” for  $Y = 1$ ) and one negative samples (“-” for  $Y = 0$ ). Complete the following questions.



1. [1.5 marks] Which model is better: Logistic Regression or Linear Regression? Explain why.

**Solution: Logistic Regression. Because Logistic Regression predicts values between 0 and 1, which is consistent with the target space  $Y$ , but Linear Regression predicts any values.**

2. [1.5 marks] Is there any logistic regression classifier using  $X_1$  and  $X_2$  that can perfectly classify the examples in Figure? How about if we change label of point (0, 1) from “+” to “-”?

**Solution: Logistic regression forms linear decision surface. Because data points in Figure is linear separable, they can be perfectly classified. But if we make the change, then data points are not linearly separable so no logistic regression classifier can perfectly classify the examples.**

- f. [2 marks] Suppose we have trained a linear regression model  $y = ax + b$  where  $a = 0.5$  and  $b = 1.0$ , on a set of training data points  $D = \{(1.0, 1.6), (1.5, 1.5), (3.0, 2.4)\}$ . Calculate the mean squared errors of this model on  $D$ .

**Solution:**  $MSE = \frac{0.1^2 + 0.25^2 + 0.1^2}{6} = 0.01375$

- g. [1 mark] Consider  $Y$  (output) Binary,  $X_i$  (features) continuous,  $X = \langle X_1, X_2, \dots, X_n \rangle$ , the number of estimated parameters in logistic regression.

**N+1**

**Question 2 True Or False**

**[5 marks]**

|                                                                                                                                                     |          |
|-----------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| A neural network with multiple hidden layers and sigmoid nodes can form non-linear decision boundaries.                                             | <b>T</b> |
| For logistic regression, with parameters optimized using a gradient method, setting parameters to 0 is an acceptable initialization.                | <b>T</b> |
| It is possible to represent a XOR function with a neural network without a hidden layer.                                                            | <b>F</b> |
| Let's say, a "Linear regression" model perfectly fits the training data, this means that you always have test error zero.                           | <b>F</b> |
| We can get multiple local optimum solutions if we solve a linear regression problem by minimizing the sum of squared errors using gradient descent. | <b>F</b> |

**Question 3**

**[3 marks]**

**Which type of machine learning algorithms is appropriate for the following problems?**

- a) Take a collection of different genes and find a way to automatically these genes into groups that are somehow similar or related by expression level or location.

**Clustering – unsupervised**

- b) Identification of faces in images.

**Classification**

- c) Calculation of the expected price of a stock in the stock market.

**Prediction –Regression**

*Good Luck*

*Dr. Hanaa Bayomi*