# Unit 7 - Week 4 - Multilayer perceptron



## Assignment 04

**The due date for submitting this assignment has passed.**
**As per our records you have not submitted this assignment.**

**Due on 2019-08-28, 23:59 IST.**

Instructions:

1. **Attempt all questions.**
2. **Submission deadline**: 28th August 2019 23:59 IST
3. **Solutions to be posted**: 29th August 2019
4. Older browsers might show **unnecessary vertical bars** at the end of math equations.

1) Consider the logistic function $\varphi(v_j(n)) = \frac{1}{1+\exp(-av_j(n))}$, where $a > 0$. Choose the correct option(s) with respect to the given activation function. **2 points**

- Maximum value of $\varphi'(v_j(n)) = 0.25$for all $a > 0$.
- Maximum value of $\varphi'(v_j(n))$ is dependent on $a$.
- $\varphi'(v_j(n))$ is an even function.
- $\varphi'(v_j(n))$ is an odd function.

No, the answer is incorrect.
Score: 0
Accepted Answers:
*Maximum value of $\varphi'(v_j(n))$ is dependent on $a$.*
*$\varphi'(v_j(n))$ is an even function.*

2) (True/False): A multilayer neural network where the neurons operate in the linear region can be approximated to a single layer neural network. **1 point**

- True
- False

No, the answer is incorrect.
Score: 0
Accepted Answers:
*True*

3) As discussed in the lecture, the momentum constant $\alpha$ is normally assigned a positive value in the range $0 < \alpha \leq 1$. Suppose $\alpha$ were assigned a negative value in the range $-1 \leq \alpha < 0$, then **2 points**

- If the sign of the gradient of $E$ remains same then algorithm converges slower.
- If the sign of the gradient alternates then algorithm converges faster.
- If the sign of the gradient of $E$ remains same then algorithm converges faster.
- If the sign of the gradient alternates then algorithm converges slower.

No, the answer is incorrect.
Score: 0
Accepted Answers:
*If the sign of the gradient of $E$ remains same then algorithm converges slower.*
*If the sign of the gradient alternates then algorithm converges faster.*

4) Consider a feed-forward artificial neural network (ANN) with a single hidden layer. There are two inputs a and b, one hidden unit c, and one output unit d. This network has five weights $(w_{ca}, w_{cb}, w_{dc}, w_{da}, w_{db})$, where $w_{xy}$ represents the bias for neuron $x$ . Initialize these weights to the values $(0.1, 0.1, 0.1, 0.1, 0.1)$. For the inputs given below, assuming sigmoid activation function with slope parameter $a = 1$, compute the output $y_i^{(i)}$ of the hidden unit for $(a^{(1)}, b^{(1)}) = (1, 0)$ and $(a^{(2)}, b^{(2)}) = (0, 1)$. **1 point**

| | $a^{(i)}$ | $b^{(i)}$ | $d^{(i)}$ |
|---|---|---|---|
| $i = 1$ | 1 | 0 | 1 |
| $i = 2$ | 0 | 1 | 0 |

- $y_c^{(1)} = 0.54, y_c^{(2)} = 0.54$
- $y_c^{(1)} = 0.55, y_c^{(2)} = 0.55$
- $y_c^{(1)} = 0.55, y_c^{(2)} = 0.54$
- $y_c^{(1)} = 0.54, y_c^{(2)} = 0.55$

No, the answer is incorrect.
Score: 0
Accepted Answers:
$y_c^{(1)} = 0.55, y_c^{(2)} = 0.55$

5) In continuation with the question 4, error $E = \frac{1}{2}\left(\left(y_d^{(1)} - d^{(1)}\right)^2 + \left(y_d^{(2)} - d^{(2)}\right)^2\right)$ at the output will be **1 point**

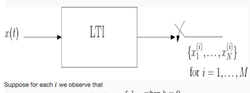
No, the answer is incorrect.
Score: 0
Accepted Answers:
*(type: Range) 0.245,0.26*

6) (True/False) Let $x_1, x_2, x_3$ be three variables taking binary values. Define $(x_1, x_2, x_3) = \overline{x_1}\ \overline{x_2}\ \overline{x_3} + x_1 x_2\ \overline{x_3} + x_1\ \overline{x_2}\ x_3 + \overline{x_1}\ x_2\ x_3$ , then the data set is linearly separable. **2 points**

- True
- False

No, the answer is incorrect.
Score: 0
Accepted Answers:
*False*

7) Consider the following linear time invariant (LTI) system shown in Figure below with input signal $x(t)$ and the output of the system is sampled randomly according to a certain distribution $f$ to get $M$ vectors $\overline{X_i} = \{x_1^{(i)}, \ldots, x_N^{(i)}\}$ for $i = 1, \ldots, M$ . **2 points**




Suppose for each $i$ we observe that

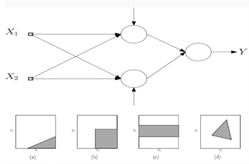
$$\mathrm{E}\left(x_{n+k}^{(i)} x_n^{(i)}\right) = \begin{cases} 1, & \text{when } k = 0 \\ 0, & \text{otherwise} \end{cases}$$

where $\mathrm{E}(\cdot)$ is an expectation operator over the distribution $f$ and we now use multi-layer perceptron (MLP) to predict $x_N$ given $x_1^{(i)}, \ldots, x_{N-1}^{(i)}$ as the input. Upon completion of training, suppose we input $\hat{x}_1, \ldots, \hat{x}_{N-1}$ to the MLP then

- Output of the MLP will be $\hat{x}_N$ with probability one.
- MLP cannot predict $\hat{x}_N$ given the input.
- Output of MLP will be 0 with probability one.
- Output of MLP will be $\hat{x}_N = \frac{\hat{x}_1 + \cdots + \hat{x}_{N-1}}{N-1}$.

No, the answer is incorrect.
Score: 0
Accepted Answers:
*MLP cannot predict $\hat{x}_N$ given the input.*

8) Consider a simple two layered neural network as in the figure and the different classification regions marked as below. **2 points**

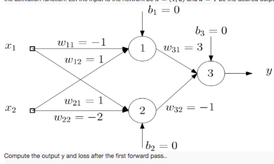



Choose the classification region(s) that can be realized using this two layered network.

- (a)
- (b)
- (c)
- (d)

No, the answer is incorrect.
Score: 0
Accepted Answers:
*(a)*
*(b)*
*(c)*

9) Consider a two layered network as in figure given below with input $x = (x_1, x_2)$. Each hidden neuron has the rectified linear unit $h(z) = \max(0, z)$ as the activation function. Let the input to the network be $x = (1, 2)$ and $d = 1$ be the desired output. The output unit is linear with loss function $J = \frac{1}{2}(y - d)^2$. **2 points**




Compute the output y and loss after the first forward pass..

- $y = 0, J = 2$
- $y = 2, J = 5$
- $y = 3, J = 2$
- $y = 4, J = 3$

No, the answer is incorrect.
Score: 0
Accepted Answers:
$y = 3, J = 2$

10) In continuation with question 9, assume the derivative of ReLU activation function, $h(z) = \max(0, z)$ is $h'(z) = \begin{cases} 1 & \text{if } z \geq 0 \\ 0 & \text{otherwise} \end{cases}$ **2 points**

Compute the local gradient at the neurons 1 and 2.

- $\frac{\partial J}{\partial v_1} = 6, \frac{\partial J}{\partial v_2} = 2$
- $\frac{\partial J}{\partial v_1} = 2, \frac{\partial J}{\partial v_2} = 0$
- $\frac{\partial J}{\partial v_1} = 3, \frac{\partial J}{\partial v_2} = -2$
- $\frac{\partial J}{\partial v_1} = 6, \frac{\partial J}{\partial v_2} = 0$

No, the answer is incorrect.
Score: 0
Accepted Answers:
$\frac{\partial J}{\partial v_1} = 6, \frac{\partial J}{\partial v_2} = 0$

11) In continuation with question 9, compute $\frac{\partial J}{\partial w_{11}}$ and $\frac{\partial J}{\partial w_{12}}$. **1 point**

- $\frac{\partial J}{\partial w_{11}} = 2, \frac{\partial J}{\partial w_{12}} = 0$
- $\frac{\partial J}{\partial w_{11}} = 1, \frac{\partial J}{\partial w_{12}} = 2$
- $\frac{\partial J}{\partial w_{11}} = -2, \frac{\partial J}{\partial w_{12}} = 2$
- $\frac{\partial J}{\partial w_{11}} = 2, \frac{\partial J}{\partial w_{12}} = 2$

No, the answer is incorrect.
Score: 0
Accepted Answers:
$\frac{\partial J}{\partial w_{11}} = 2, \frac{\partial J}{\partial w_{12}} = 0$

12) In continuation with question 9, $\frac{\partial J}{\partial w_{31}} =$

No, the answer is incorrect.
Score: 0
Accepted Answers:
*(type: Numeric) 6*

**1 point**

13) In continuation with question 9, $\frac{\partial J}{\partial w_{32}} =$

No, the answer is incorrect.
Score: 0
Accepted Answers:
*(type: Numeric) 0*

**1 point**