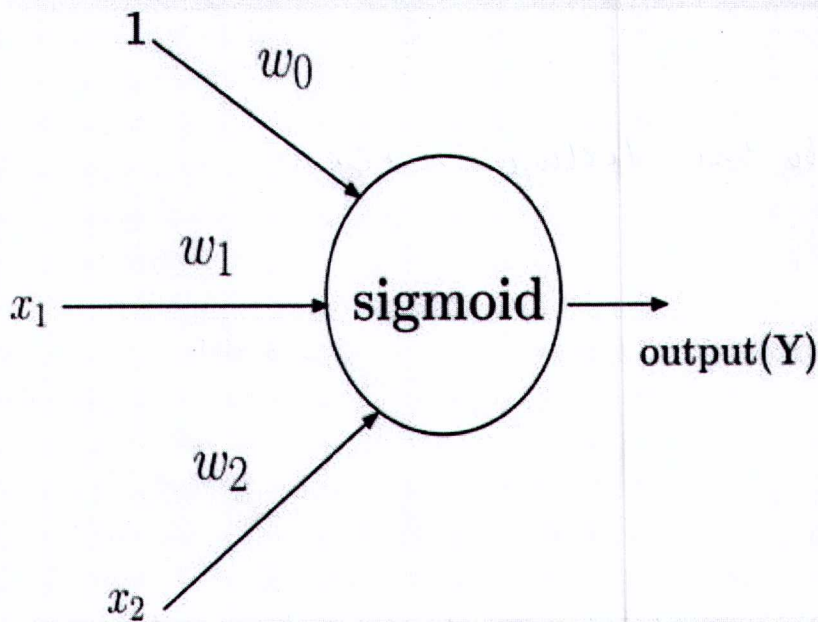


Week 6 Assignment:

Topic: Neural network

1. What would be a correct option for weight $W [w_0, w_1, w_2]$, so that the following sigmoid unit will function as an AND gate?



- A) [-5,10,10]
- B) [-10,15,15]
- C) [-10,15,5]
- D) [-10,9,9]

Input to the sigmoid unit = $w_0 + x_1 w_1 + x_2 w_2$.

For only option D, the unit will function as AND gate.

2. The back-propagation learning algorithm applied to a two layer neural network
- A) always finds the globally optimal solution.
 - B) finds a locally optimal solution which may be globally optimal.**
 - C) never finds the globally optimal solution.
 - D) finds a locally optimal solution which is never globally optimal

The back-propagation learning algorithm finds a locally optimal solution which may be globally optimal.

3. Which of the following is true?

- A) In batch gradient descent we update the weights and biases of the neural network after forward pass over each training example.
- B) In batch gradient descent we update the weights and biases of our neural network after forward pass over all the training examples.**
- C) Each step of stochastic gradient descent takes more time than each step of batch gradient descent.
- D) None of these three options is correct

Refer to the lecture notes.

4. In a neural network, which one of the following techniques is **NOT** useful to reduce overfitting?

- A) Dropout
- B) Regularization
- C) Batch normalization
- D) Adding more layers**

Adding more layers does not reduce overfitting.

5. For an image recognition problem (such as recognizing a cat in a photo), which architecture of neural network has been found to be better suited for the tasks?

- A) Multi layer perceptron
- B) Recurrent neural network
- C) Convolutional neural network**
- D) Perceptron

Refer lecture notes.

6. In training a batch neural network, after running the first few epochs, you notice that the loss does not decrease. The reasons for this could be

1. The learning rate is low.
2. The neural net is stuck in local minima
3. The neural net has too many units in the hidden layer

A) 1 or 2

B) 1 or 3

C) 2 or 3

D) 1 only

Refer lecture notes.

7. What is the sequence of steps followed in training a perceptron?

1. For a sample input, compute an output
2. Initialize weights of perceptron randomly
3. Go to the next batch of dataset
4. If the prediction does not match the output, change the weights

A) 2,1,4,3

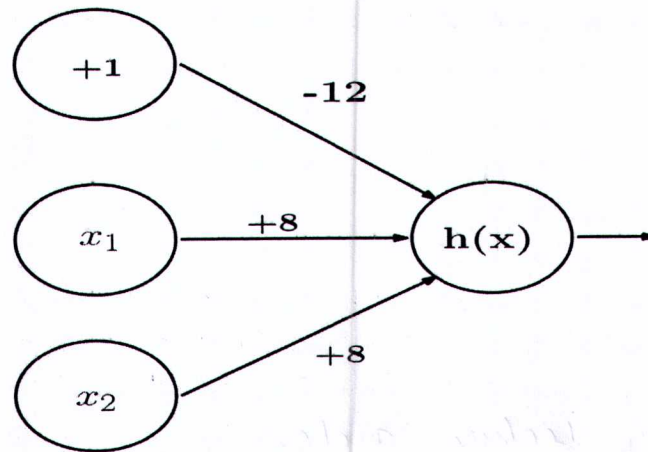
B) 1,4,3,2

C) 1,2,3,4

D) 2,3,4,1

Refer lecture notes.

8. The neural network given below takes two binary valued inputs $x_1, x_2 \in \{0,1\}$ and the activation function is the binary threshold function ($h(x)=1$ if $x>0$; 0 otherwise). Which of the following logical functions does it compute?



- A) OR
- B) AND
- C) NAND
- D) None of these

Input to the output node = $8x_1 + 8x_2 - 12$

x_1	x_2	$8x_1 + 8x_2 - 12$	$h(x)$
0	0	-12	0
0	1	-4	0
1	0	-4	0
1	1	4	1

} AND gate

9. What are the steps for using a gradient descent algorithm?

1. Calculate error between the actual value and the predicted value
2. Repeat until you find the best weights of network
3. Pass an input through the network and get values from output layer
4. Initialize random values for weight and bias
5. Go to each neurons which contributes to the error and change its respective values to reduce the error

- A) 4,3,1,5,2
- B) 1,2,3,4,5
- C) 3,4,5,2,1
- D) 2,3,4,5,1

Refer lecture notes.

10. A 4-input neuron has bias of 0 and weights 1, 2, 3 and 4. The transfer function is given by $f(v) = \max(0, v)$. The inputs are 4, 10, 5 and 20 respectively. The output will be

- A) 238
- B) 119**
- C) 75
- D) 121

$$\begin{aligned} & \max(0, 1 \times 4 + 2 \times 10 + 3 \times 5 + 4 \times 20) \\ &= \max(0, 119) \\ &= 119. \end{aligned}$$