Congratulations! You passed!

Grade received 100%

Latest Submission Grade 100%

To pass 80% or higher

Retake the assignment in **7h 40m**

Go to next item

1. Which of the following are true? (Check all that apply.) **1 / 1 point**

- [ ] X is a matrix in which each row is one training example.
- [ ] $a^{[2](12)}$ denotes activation vector of the 12^{th} layer on the 2^{nd} training example.
- [ ] $a_4^{[2]}$ is the activation output of the 2^{nd} layer for the 4^{th} training example
- [x] $a_4^{[2]}$ is the activation output by the 4^{th} neuron of the 2^{nd} layer

  ✓ **Correct**

- [x] $a^{[2](12)}$ denotes the activation vector of the 2^{nd} layer for the 12^{th} training example.

  ✓ **Correct**

- [x] X is a matrix in which each column is one training example.

- [x] $a^{[2](12)}$ denotes the activation vector of the 2^{nd} layer for the 12^{th} training example.

  ✓ **Correct**

- [x] X is a matrix in which each column is one training example.

  ✓ **Correct**

- [x] $a^{[2]}$ denotes the activation vector of the 2^{nd} layer.

  ✓ **Correct**

Expand

Correct
Great, you got all the right answers.

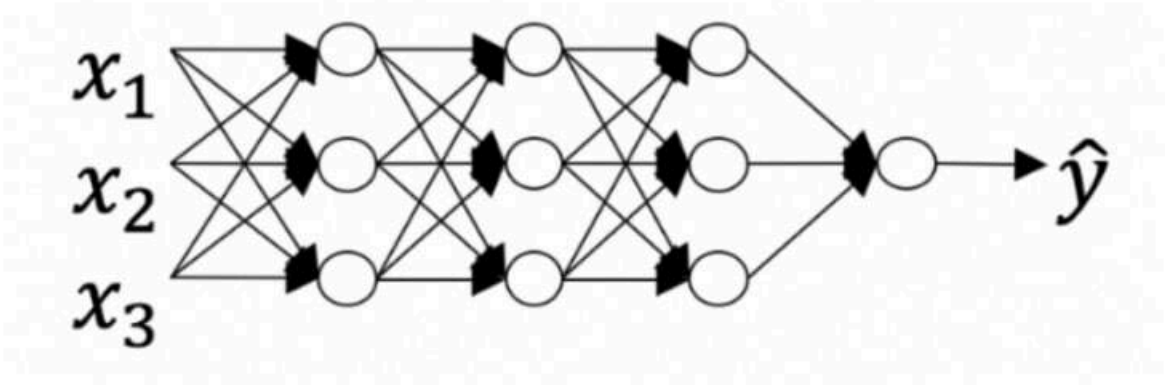
2. The sigmoid function is only mentioned as an activation function for historical reasons. The tanh is always preferred without exceptions in all the layers of a Neural Network. True/False? 1 / 1 point

- ○ True
- ◉ False

Expand

Correct
Yes. Although the tanh almost always works better than the sigmoid function when used in hidden layers, thus is always proffered as activation function, the exception is for the output layer in classification problems.

3. Which of the following represents the activation output of the second neuron of the third layer applied to the fourth example? 1 / 1 point



- $a_2^{[3](4)}$
- $a_2^{[4](3)}$
- $a_4^{[3](2)}$
- $a_3^{[4]2}$

- $a_2^{[3](4)}$
- $a_2^{[4](3)}$
- $a_4^{[3](2)}$
- $a_3^{[4]2}$

Expand

Correct

Yes. The superscript in brackets indicates the layer number, the superscript in parenthesis represents the number of examples, and the subscript the number of the neuron.

4. You are building a binary classifier for recognizing cucumbers (y=1) vs. watermelons (y=0). Which one of these activation functions would you recommend using for the output layer? 1 / 1 point

- (●) sigmoid
- ( ) tanh
- ( ) Leaky ReLU
- ( ) ReLU

Expand

Correct
Yes. Sigmoid outputs a value between 0 and 1 which makes it a very good choice for binary classification. You can classify as 0 if the output is less than 0.5 and classify as 1 if the output is more than 0.5. It can be done with tanh as well but it is less convenient as the output is between -1 and 1.

5. Consider the following code: 1 / 1 point

A = np.random.randn(4,3)

B = np.sum(A, axis = 1, keepdims = True)

What will be B.shape? (If you're not sure, feel free to run this in python to find out).

- (4, 1)
- (4,)
- (1, 3)
- (3,)

Expand

Correct
Yes, we use (keepdims = True) to make sure that A.shape is (4,1) and not (4,). It makes our code more robust.

6. Suppose you have built a neural network with one hidden layer and tanh as activation function for the hidden layer. You decide to initialize the weights to small random numbers and the biases to zero. The first hidden layer's neurons will perform different computations from each other even in the first iteration. True/False? 1 / 1 point

- False No. Since the weights are most likely different, each neuron will do a different computation.
- True Yes. Since the weights are most likely different, each neuron will do a different computation.

Expand

Correct

7. A single output and single layer neural network that uses the sigmoid function as activation is equivalent to the logistic regression. True/False

1 / 1 point

- False
- True

Expand

Correct

Yes. The logistic regression model can be expressed by $\hat{y} = \sigma\left(W\,x + b\right)$. This is the same as $a^{[1]} = \sigma(W^{[1]}\,X + b)$.

8. Which of the following is true about the tanh function? 1 / 1 point

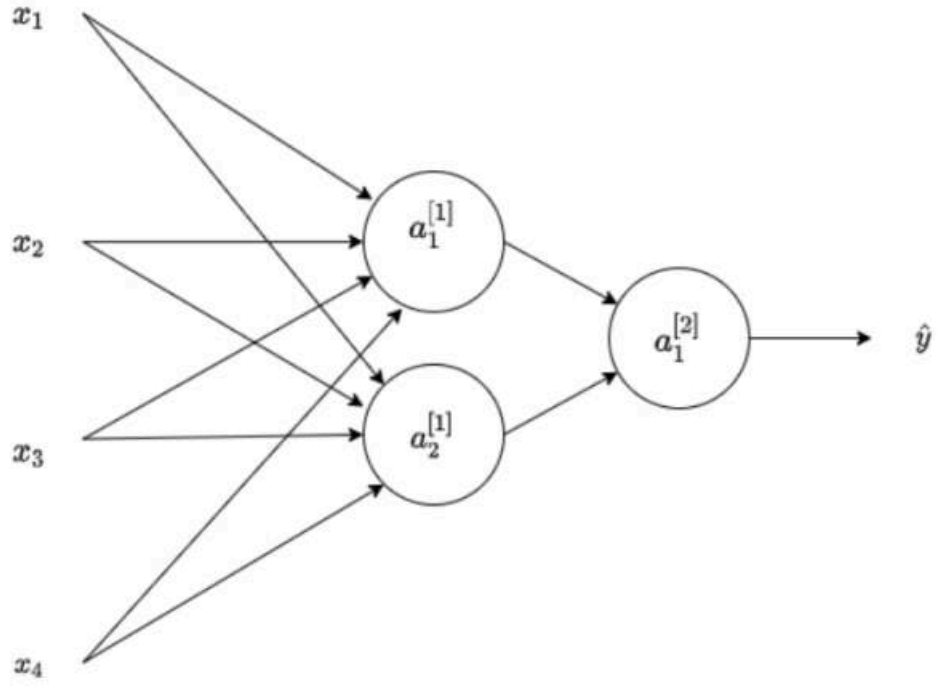
- The slope is zero for negative values.
- The derivative at $c = 0$ is not well defined.
- For large values the slope is close to zero.
- For large values the slope is larger.

Expand

Correct
Yes. We can see in the graph of the $y = tanh(c)$ how as the values of c increase the curve becomes flatter.

9. Consider the following 1 hidden layer neural network:

1 / 1 point



Which of the following statements are True? (Check all that apply).

Which of the following statements are True? (Check all that apply).

- [x] $W^{[2]}$ will have shape (1, 2)

  ✓ **Correct**

  Yes. The number of rows in $W^{[k]}$ is the number of neurons in the k-th layer and the number of columns is the number of inputs of the layer.

- [x] $W^{[1]}$ will have shape (2, 4).

  ✓ **Correct**

  Yes. The number of rows in $W^{[k]}$ is the number of neurons in the k-th layer and the number of columns is the number of inputs of the layer.

- [ ] $b^{[1]}$ will have shape (4, 2)
- [ ] $W^{[2]}$ will have shape (2, 1)
- [x] $b^{[1]}$ will have shape (2, 1).

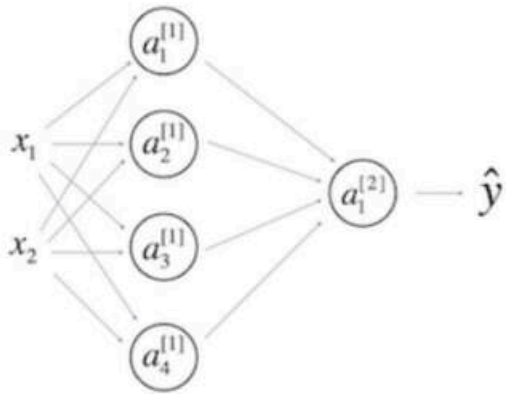
  ✓ **Correct**

  Yes. $b^{[k]}$ is a column vector and has the same number of rows as neurons in the k-th layer.

- [ ] $W^{[1]}$ will have shape (4, 2).

Correct
Great, you got all the right answers.

10. What are the dimensions of $Z^{[1]}$ and $A^{[1]}$? **1 / 1 point**



- (●) $Z^{[1]}$ and $A^{[1]}$ are (4,m)
- ( ) $Z^{[1]}$ and $A^{[1]}$ are (4,1)
- ( ) $Z^{[1]}$ and $A^{[1]}$ are (1,4)
- ( ) $Z^{[1]}$ and $A^{[1]}$ are (4,2)