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

☐ $w_3^{[4]}$ is the column vector of parameters of the third layer and fourth neuron.

☐ $W^{[1]}$ is a matrix with rows equal to the transpose of the parameter vectors of the first layer.

☐ $w_3^{[4]}$ is the column vector of parameters of the fourth layer and third neuron.

☐ $W^{[1]}$ is a matrix with rows equal to the parameter vectors of the first layer.

☐ $w_3^{[4]}$ is the row vector of parameters of the fourth layer and third neuron.

☐ W_1 is a matrix with rows equal to the parameter vectors of the first layer.

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 point

○ True

○ False

3.Which of the following represents the activation output of the second neuron of the third layer applied to the fourth example? 1 point
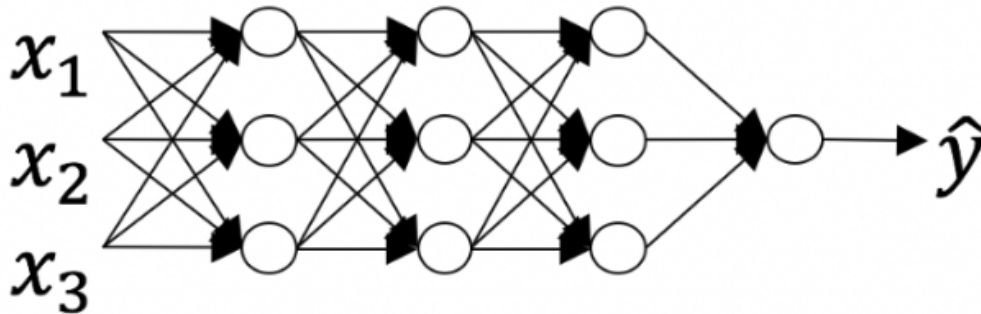
○ $a_2^{[3](4)}$

○ $a_2^{[4](3)}$

○ $a_4^{[3](2)}$

○ $a_3^{[4]2}$

4.The use of the ReLU activation function is becoming more rare because the ReLU function has no derivative for $c = 0$. True/False? 1 point

○ False

○ True

5.Consider the following code: 1 point

```python
#+begin_src python

x = np.random.rand(3, 2)

y = np.sum(x, axis=0, keepdims=True)

#+end_src
```

What will be y.shape?

○ (1, 2)

○ (3, 1)

○ (2,)

○ (3,)

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 point

○ True

  Yes. Since the weights are most likely different, each neuron will do a different computation.

○ False

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

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 point

○ True

○ False

8. Which of the following is true about the ReLU activation functions?

1 point

○ They are only used in the case of regression problems, such as predicting house prices.
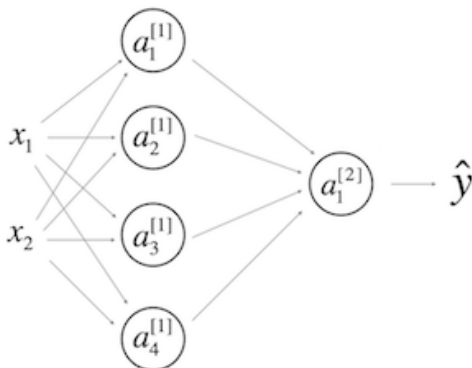
○ They cause several problems in practice because they have no derivative at 0. That is why Leaky ReLU was invented.

○ They are the go to option when you don't know what activation function to choose for hidden layers.

○ They are increasingly being replaced by the tanh in most cases.

9. Consider the following 1 hidden layer neural network:

1 point

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

☐ $W^{[2]}$ will have shape (4, 1)

☐ $b^{[2]}$ will have shape (1, 1)

☐ $b^{[1]}$ will have shape (4, 1)

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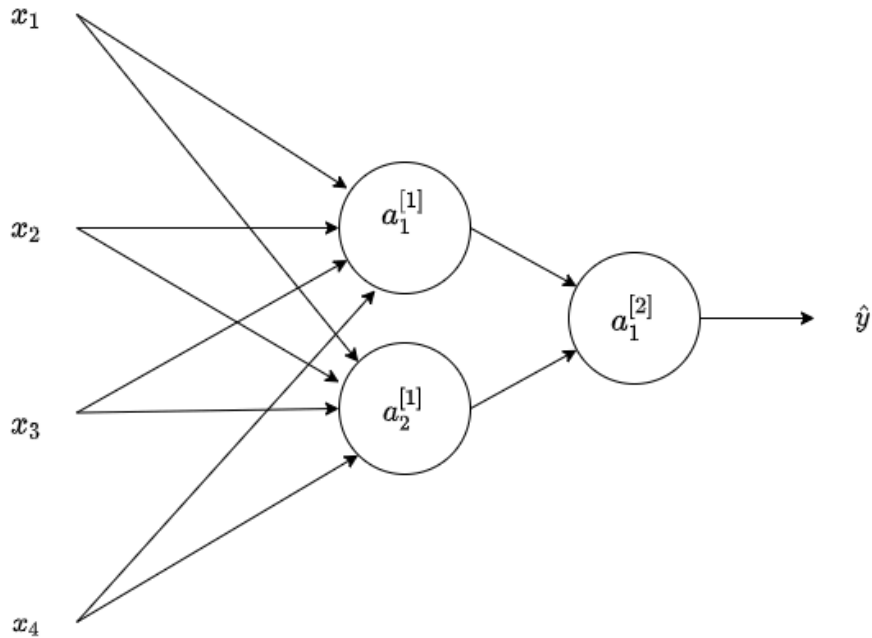
☐ $b^{[2]}$ will have shape (4, 1)

☐ $W^{[1]}$ will have shape (2, 4)

☐ $W^{[2]}$ will have shape (1, 4)

☐ $b^{[1]}$ will have shape (2, 1)

10.Consider the following 1 hidden layer neural network: 1 point

What are the dimensions of $Z^{[1]}$ and $A^{[1]}$?

○ $Z^{[1]}$ and $A^{[1]}$ are (2, m)

○ $Z^{[1]}$ and $A^{[1]}$ are (4, m)

○ $Z^{[1]}$ and $A^{[1]}$ are (2, 1)

○ $Z^{[1]}$ and $A^{[1]}$ are (4, 1)