1.We use the "cache" in our implementation of forward and backward propagation to pass useful values to the next layer in the forward propagation. True/False? 1 point

○ True

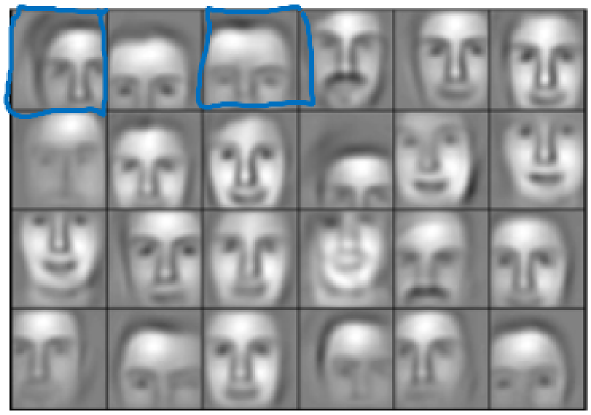
○ False

2.During the backpropagation process, we use gradient descent to change the hyperparameters. True/False? 1 point

○ True

○ False

3.Considering the intermediate results below, which layers of a deep neural network are they likely to belong to? 1 point



○ Later layers of the deep neural network.

○ Input layer of the deep neural network.

○ Early layers of the deep neural network.

○ Middle layers of the deep neural network.

4.We can not use vectorization to calculate $da^{[l]}$ in backpropagation, we must use a for loop over all the examples. True/False? 1 point

○ False

○ True

5.Assume we store the values for $n^{[l]}$ in an array called layer_dims, as follows: layer_dims = [n_x, 4,3,2,1]. So layer 1 has four hidden units, layer 2 has 3 hidden units, and so on. Which of the following for-loops will allow you to initialize the parameters for the model? 1 point

○

```
for i in range(len(layer_dims)-1):

parameter['W' + str(i+1)] = np.random.randn(layer_dims[i], layer_dims[i+1]) * 0.01

parameter['b' + str(i+1)] = np.random.randn(layer_dims[i+1], 1) * 0.01
```

○

```
for i in range(1, len(layer_dims)/2):

parameter['W' + str(i)] = np.random.randn(layer_dims[i], layer_dims[i-1]) * 0.01

parameter['b' + str(i)] = np.random.randn(layer_dims[i], 1) * 0.01
```

○

```
for i in range(len(layer_dims)-1):

parameter['W' + str(i+1)] = np.random.randn(layer_dims[i+1], layer_dims[i]) * 0.01

parameter['b' + str(i+1)] = np.random.randn(layer_dims[i+1], 1) * 0.01
```

○

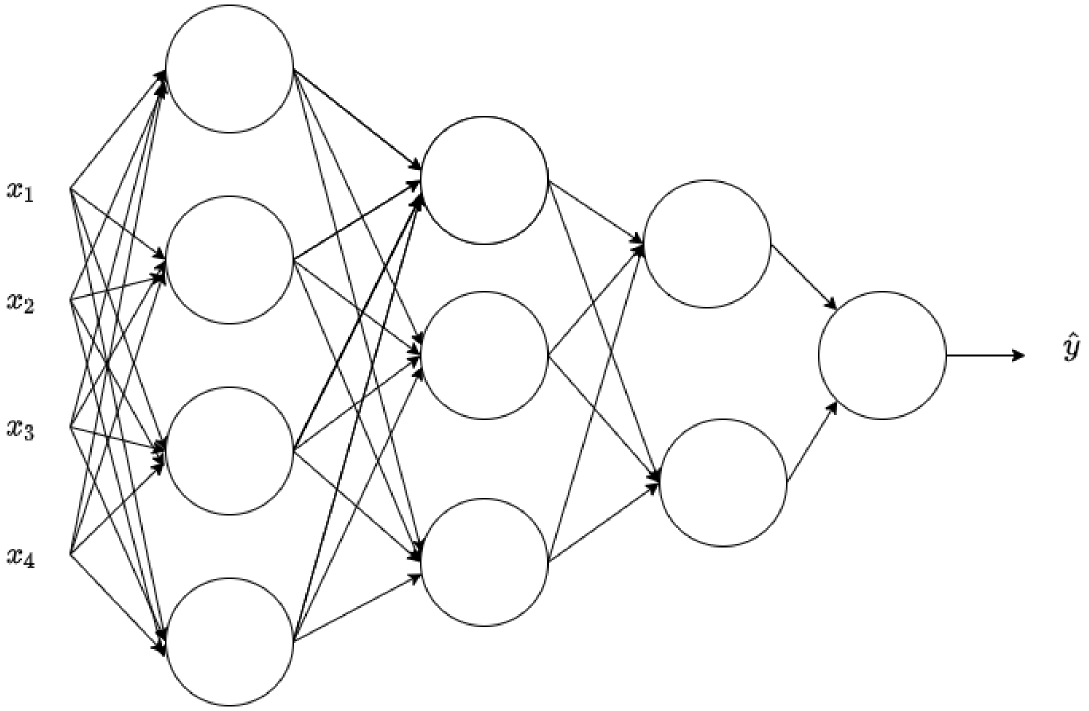
```
for i in range(len(layer_dims)):

parameter['W' + str(i+1)] = np.random.randn(layer_dims[i+1], layer_dims[i]) * 0.01

parameter['b' + str(i+1)] = np.random.randn(layer_dims[i+1], 1) * 0.01
```

6.Consider the following neural network:

1 point



What are all the values of $n^{[0]}$, $n^{[1]}$, $n^{[2]}$, $n^{[3]}$ and $n^{[4]}$?

○ 4, 4, 3, 2

○ 4, 3, 2, 1

○ 4, 4, 3, 2, 1

○ 4, 3, 2

7.If L is the number of layers of a neural network then $dZ^{[L]} = A^{[L]} - Y$. True/False? 1 point

○ False

○ True

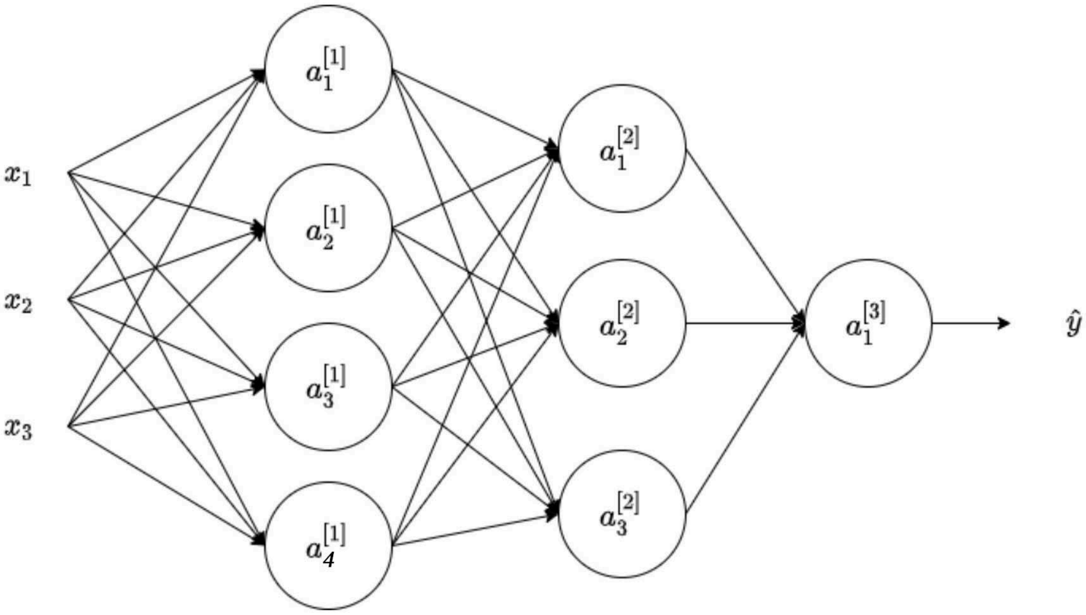
8.There are certain functions with the following properties: 1 point

(i) To compute the function using a shallow network circuit, you will need a large network (where we measure size by the number of logic gates in the network), but (ii) To compute it using a deep network circuit, you need only an exponentially smaller network. True/False?

○ True

○ False

9.Consider the following 2 hidden layers neural network: 1 point



Which of the following statements is true? (Check all that apply).

□ $W^{[2]}$ will have shape (3, 4)

□ $W^{[1]}$ will have shape (4, 3)

□ $b^{[1]}$ will have shape (1, 4)

□ $W^{[2]}$ will have shape (3, 1)

□ $W^{[1]}$ will have shape (3, 4)

□ $W^{[2]}$ will have shape (1, 3)

□ $W^{[2]}$ will have shape (4, 3)

□ $b^{[1]}$ will have shape (3, 1)

□ $b^{[1]}$ will have shape (4, 1)

10.Whereas the previous question used a specific network, in the general case what is the dimension of $b^{[l]}$, the bias vector associated with layer l?

1 point

○ $b^{[l]}$ has shape $(n^{[l]}, 1)$

○ $b^{[l]}$ has shape $(n^{[l+1]}, 1)$

○ $b^{[l]}$ has shape $(1, n^{[l-1]})$

○ $b^{[l]}$ has shape $(1, n^{[l]})$