← Back **Key Concepts on Deep Neural Networks**
Graded Quiz • 50 min

Due Aug 28, 11:59 PM +03

Try again once you are ready

Grade received 70%

Latest Submission Grade 70%

To pass 80% or higher

Try again

1. What is stored in the 'cache' during forward propagation for latter use in backward propagation?

1 / 1 point

● $Z^{[l]}$

○ $W^{[l]}$

○ $A^{[l]}$

○ $b^{[l]}$

Expand

Correct
Yes. This value is useful in the calculation of $dW^{[l]}$ in the backward propagation.

2. Among the following, which ones are "hyperparameters"? (Check all that apply.)

1 / 1 point

☑ size of the hidden layers $n^{[l]}$

✓ **Correct**

☑ number of layers L in the neural network

✓ **Correct**

□ activation values $a^{[l]}$

□ bias vectors $b^{[l]}$

☑ number of iterations

✓ **Correct**

□ weight matrices $W^{[l]}$

☑ learning rate α

✓ **Correct**

Expand

✓ **Correct**

Great, you got all the right answers.

3. Which of the following statements is true?

1 / 1 point

● The deeper layers of a neural network are typically computing more complex features of the input than the earlier layers.

○ The earlier layers of a neural network are typically computing more complex features of the input than the deeper layers.

Expand

✓ **Correct**

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

0 / 1 point

● True

○ False

Expand

⊗ **Incorrect**

Incorrect. We can use vectorization in backpropagation to calculate $dA^{[l]}$ for each layer. This computation is done over all the training examples.

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?

0 / 1 point

●
```
for i in range(1, len(layer_dims)):
parameter['W' + str(i)] = np.random.randn(layer_dims[i-1], layer_dims[i]) * 0.01
parameter['b' + str(i)] = np.random.randn(layer_dims[i], 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(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], 1) * 0.01

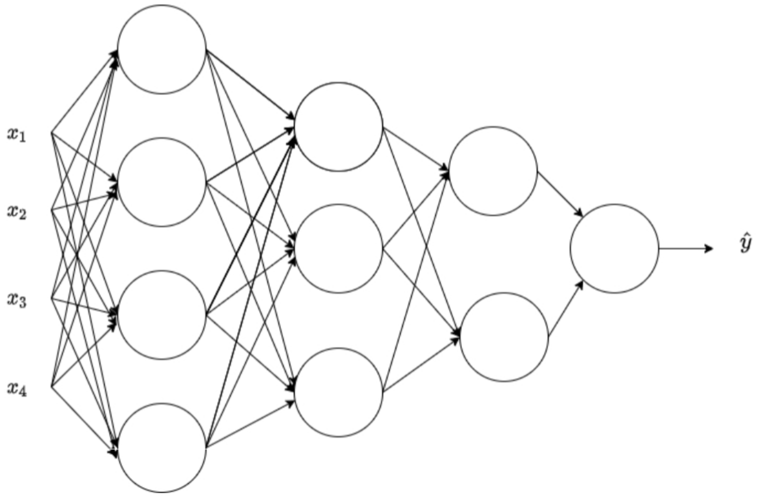
○ for i in range(1, len(layer_dims)):
parameter['W' + str(i)] = np.random.randn(layer_dims[i], layer_dims[i-1]) * 0.01
parameter['b' + str(i)] = np.random.randn(layer_dims[i], 1) * 0.01

Expand

ⓧ **Incorrect**

6. Consider the following neural network:

1 / 1 point



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

● 4, 4, 3, 2, 1

○ 4, 3, 2, 1

○ 4, 3, 2

○ 4, 4, 3, 2

Expand

✓ **Correct**
Yes. The $n^{[l]}$ are the number of units in each layer, notice that $n^{[0]} = n_x$.

7. During forward propagation, in the forward function for a layer l you need to know what is the activation function

1 / 1 point

in a layer (sigmoid, tanh, ReLU, etc.). During backpropagation, the corresponding backward function also needs to know what is the activation function for layer l, since the gradient depends on it. True/False?

○ False

● True

Expand

Correct
Yes, as you've seen in week 3 each activation has a different derivative. Thus, during backpropagation you need to know which activation was used in the forward propagation to be able to compute the correct derivative.

8. A shallow neural network with a single hidden layer and 6 hidden units can compute any function that a neural network with 2 hidden layers and 6 hidden units can compute. True/False?

0 / 1 point

○ False

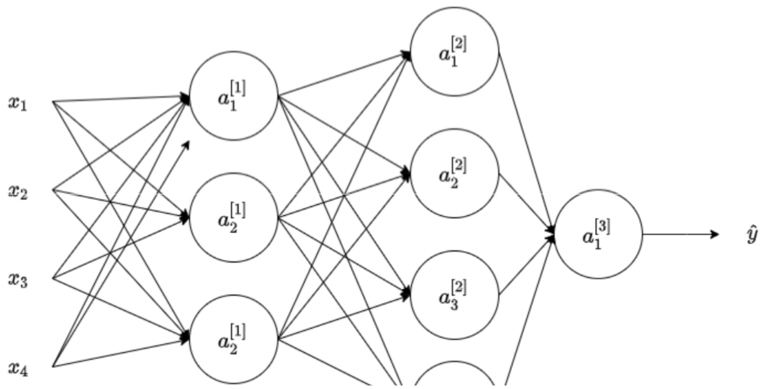
● True

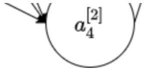
Expand

Incorrect
Incorrect. As seen during the lectures there are functions you can compute with a "small" L-layer deep neural network that shallower networks require exponentially more hidden units to compute.

9. Consider the following 2 hidden layers neural network:

1 / 1 point





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

☐ $W^{[2]}$ will have shape (3, 4)

☐ $b^{[1]}$ will have shape (4, 1)

☐ $W^{[2]}$ will have shape (1, 3)

☑ $W^{[1]}$ will have shape (3, 4)

✓ **Correct**
Yes. More generally, the shape of $W^{[l]}$ is $(n^{[l]}, n^{[l-1]})$.

☐ $b^{[1]}$ will have shape (1, 3)

☑ $W^{[2]}$ will have shape (4, 3)

✓ **Correct**
Yes. More generally, the shape of $W^{[l]}$ is $(n^{[l]}, n^{[l-1]})$.

☐ $W^{[2]}$ will have shape (3, 1)

☑ $b^{[1]}$ will have shape (3, 1)

✓ **Correct**
Yes. More generally, the shape of $b^{[l]}$ is $(n^{[l]}, 1)$.

☐ $W^{[1]}$ will have shape (4, 3)

Expand

Correct
Great, you got all the right answers.

10. In the general case if we are training with m examples what is the shape of $A^{[l]}$?

1 / 1 point

○ $(m, n^{[l]})$

○ $(n^{[l+1]}, m)$

● $(n^{[l]}, m)$

○ $(m, n^{[l+1]})$

Expand

✓ **Correct**

Yes. The number of rows in $A^{[1]}$ corresponds to the number of units in the l-th layer.