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

English Due Apr 7, 2:59 PM CST

Your grade: 100%

Your latest: **100%** • Your highest: **100%** • To pass you need at least 80%. We keep your highest score.

Next item →

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

1 / 1 point

- (●) $Z^{[l]}$
- ( ) $W^{[l]}$
- ( ) $b^{[l]}$

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

- [ ] weight matrices $W^{[l]}$
- [x] number of layers L in the neural network

Correct

- [x] learning rate α

Correct

- [ ] activation values $a^{[l]}$
- [ ] bias vectors $b^{[l]}$
- [x] number of iterations

Correct

- [x] size of the hidden layers $n^{[l]}$

Correct

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

1 / 1 point

- ( ) Middle layers of the deep neural network.
- ( ) Input layer of the deep neural network.
- (●) Later layers of the deep neural network.
- ( ) Early layers of the deep neural network.

Correct
Correct. The deep layers of a neural network are typically computing more complex features such as the ones shown in the figure.

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 / 1 point

◉ False

○ True

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

5. Suppose W[i] is the array with the weights of the i-th layer, b[i] is the vector of biases of the i-th layer, and g is the activation function used in all layers. Which of the following calculates the forward propagation for the neural network with L layers. **1 / 1 point**

○

for i in range(L):

Z[i] = W[i]*X + b[i]

A[i] = g(Z[i])

◉

for i in range(1, L+1):

Z[i] = W[i]*A[i-1] + b[i]

A[i] = g(Z[i])

○

for i in range(L):

Z[i+1] = W[i+1]*A[i+1] + b[i+1]

A[i+1] = g(Z[i+1])

○

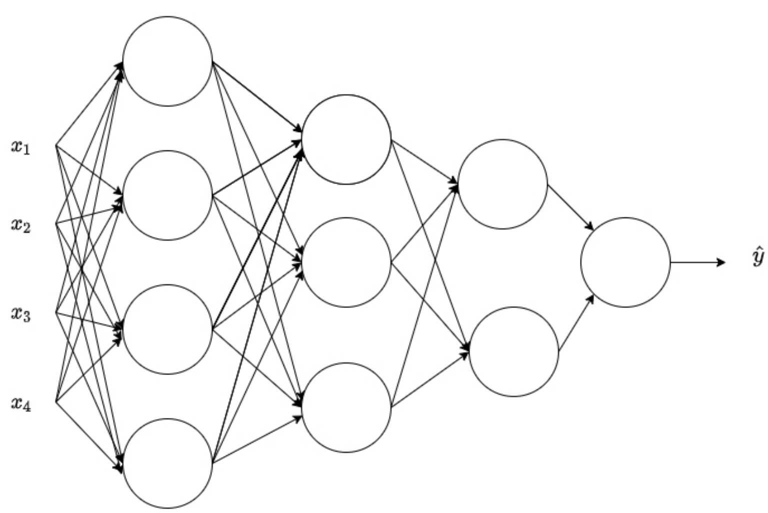
for i in range(1, L):

Z[i] = W[i]*A[i-1] + b[i]

A[i] = g(Z[i])

Correct
Yes. Remember that the range omits the last number thus the range from 1 to L+1 gives the L necessary values.

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, 3, 2, 1

◉ 4, 4, 3, 2, 1

○ 4, 4, 3, 2

○ 4, 3, 2

◯ 4, 3, 2

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 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?

1 / 1 point

- (◉) True
- ( ) False

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. For any mathematical function you can compute with an L-layered deep neural network with N hidden units there is a shallow neural network that requires only $\log N$ units, but it is very difficult to train.

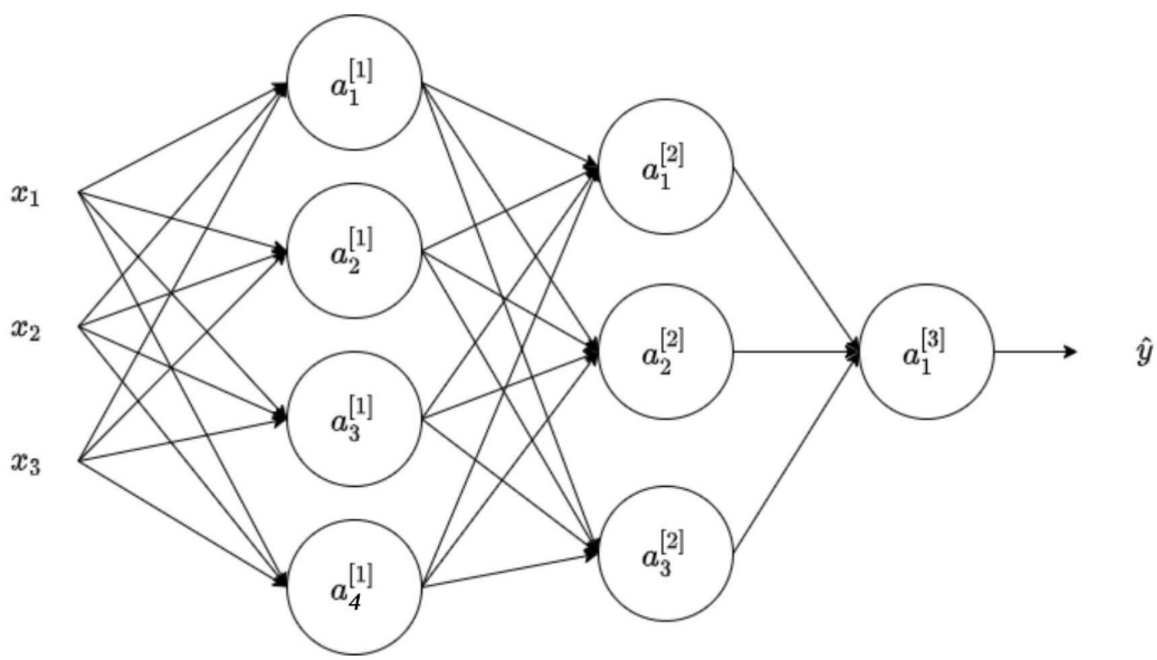
1 / 1 point

- ( ) True
- (◉) False

Correct
Correct. On the contrary, some mathematical functions can be computed using an L-layered neural network and a given number of hidden units; but using a shallow neural network the number of necessary hidden units grows exponentially.

9. Consider the following 2 hidden layers neural network:

1 / 1 point



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

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

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

- [ ] $b^{[1]}$ will have shape (3, 1)
- [ ] $W^{[2]}$ will have shape (4, 3)
- [x] $W^{[2]}$ will have shape (3, 4)

Correct

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

- [ ] $W^{[2]}$ will have shape (3, 1)
- [x] $W^{[1]}$ 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 (1, 3)
- [ ] $W^{[1]}$ will have shape (3, 4)

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 / 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 $(1, n^{[l]})$

Correct

True. $b^{[l]}$ is a column vector with the same number of rows as units in the respective layer.