Congratulations! You passed!

Grade received 90% **Latest Submission Grade** 90% **To pass** 80% or higher

Go to next item

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

0 / 1 point

- $A^{[l]}$
- $W^{[l]}$
- $Z^{[l]}$
- $b^{[l]}$

Expand

Incorrect
No. In the equations for backward propagation we see that $dA^{[l]}$ and finally $dW^{[l]}$ depends on $Z^{[l]}$ not on $A^{[l]}$.

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

1 / 1 point

- True
- False

Expand

Correct
Correct. During backpropagation, we use gradient descent to compute new values of $W^{[l]}$ and $b^{[l]}$. These are the parameters of the network.

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

1 / 1 point

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

Expand

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

4. Vectorization allows you to compute forward propagation in an L-layer neural network without an explicit for-loop (or any other explicit iterative loop) over the layers l=1, 2, …,L. True/False?

1 / 1 point

- True
- False

Expand

Correct

Forward propagation propagates the input through the layers, although for shallow networks we may just write all the lines ($a^{[2]} = g^{[2]}(z^{[2]}), z^{[2]} = W^{[2]}a^{[1]} + b^{[2]}, ...$) in a deeper network, we cannot avoid a for loop iterating over the layers: ($a^{[l]} = g^{[l]}(z^{[l]}), z^{[l]} = W^{[l]}a^{[l-1]} + b^{[l]}, ...$).

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(1, L+1):
  Z[i] = W[i]*A[i-1] + b[i]
  A[i] = g(Z[i])
- for i in range(L):
  Z[i] = W[i]*X + 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])

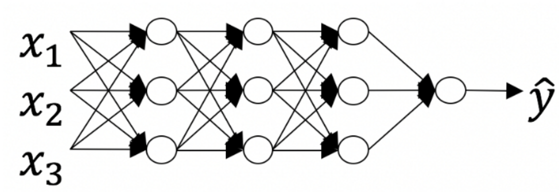
Expand

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



How many layers does this network have?

- The number of layers L is 4. The number of hidden layers is 4.
- The number of layers L is 5. The number of hidden layers is 4.
- The number of layers L is 4. The number of hidden layers is 3.
- The number of layers L is 3. The number of hidden layers is 3.

Expand

Correct

Yes. As seen in lecture, the number of layers is counted as the number of hidden layers + 1. The input and output layers are not counted as hidden layers.

7. During forward propagation, for the value of $A^{[l]}$ the value is used of $Z^{[l]}$ with the activation function $g^{[l]}$. During backward propagation we calculate $dA^{[l]}$ from $Z^{[l]}$.

1 / 1 point

- False
- True

Expand

Correct

Correct. During backward propagation we are interested in computing $dW^{[l]}$ and $db^{[l]}$. For that we use $g'^{L}, dZ^{[l]}, Z^{[l]}$, and $W^{[l]}$.

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?

1 / 1 point

- False
- True

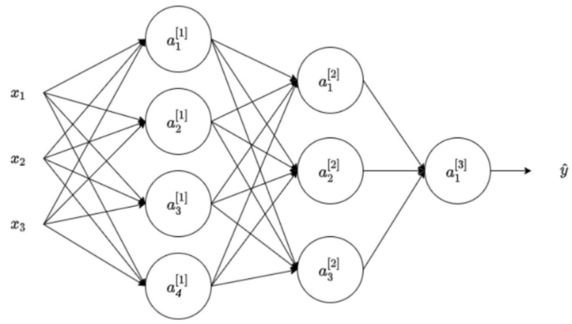
Expand

Correct

Correct. 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 is true? (Check all that apply).

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

Correct

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

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

Correct

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

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

Expand

Correct

Great, you got all the right answers.

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

Expand

Correct

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