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? **1 / 1 point**

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

Expand

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

2. Which of the following are “parameters” of a neural network? (Check all that apply.) **1 / 1 point**

- [x] $b^{[l]}$ the bias vector.
- [ ] $g^{[l]}$ the activation functions.
- [x] $W^{[l]}$ the weight matrices.
- [ ] L the number of layers of the neural network.

Correct
Correct. The weight matrices and the bias vectors are the parameters of the network.

Correct
Correct. The weight matrices and the bias vectors are the parameters of the network.

Expand

Correct
Great, you got all the right answers.

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

- Early layers of the deep neural network.
- Later layers of the deep neural network.
- Input layer of the deep neural network.
- Middle 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: ($a^{[l]} = g^{[l]}(z^{[l]})$, $z^{[l]} = W^{[l]}a^{[l-1]} + b^{[l]}$, ...).

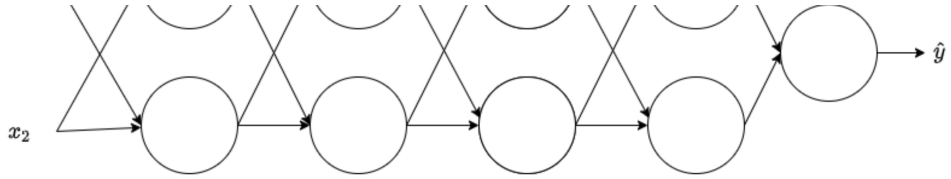
5. Suppose W[i] is the array with the weights of the i-th layer, b[i] is the vector of biases of the i-th layer, and g is the activation layer used in all layers. Which of the following calculates the forward propagation for the neural network with L layers. **1 / 1 point**

- 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])

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 5.
- ( ) The number of layers L is 4.
- ( ) The number of layers L is 6
- ( ) The number of layers L is 2.

Expand

Correct
Yes. The number of layers is the number of hidden layers + 1.

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? **0 / 1 point**

- (•) False
- ( ) True

Expand

Incorrect
No, 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. There are certain functions with the following properties: **1 / 1 point**

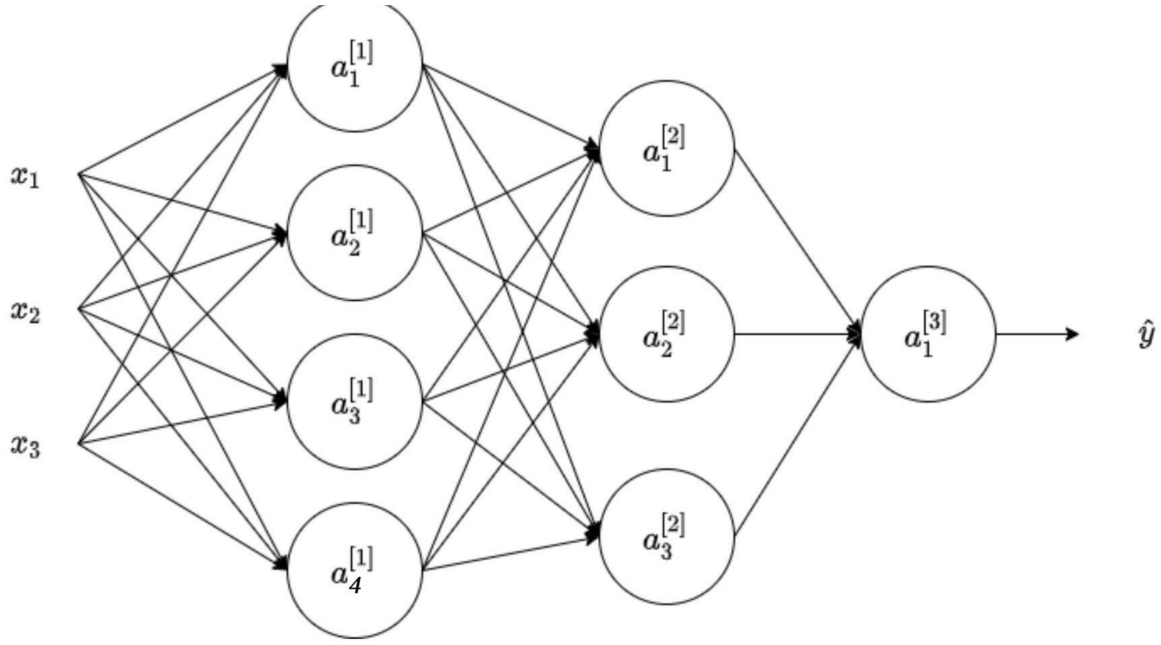
(i) To compute the function using a shallow network circuit, you will need a large network (where we measure size by the number of logic gates in the network), but (ii) To compute it using a deep network circuit, you need only an exponentially smaller network. True/False?

- ( ) False
- (•) True

Expand

Correct

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



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

- [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 (3, 1)
- [ ] $b^{[1]}$ will have shape (1, 4)
- [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 (3, 1)
- [ ] $W^{[1]}$ will have shape (3, 4)
- [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 (4, 3)
- [ ] $W^{[2]}$ will have shape (1, 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 $(1, n^{[l]})$
- (•) $b^{[l]}$ has shape $(n^{[l]}, 1)$

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