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

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

Expand

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

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

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

Correct

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

Correct

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

Correct

- [x] learning rate α

Which of the following statements is true?

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

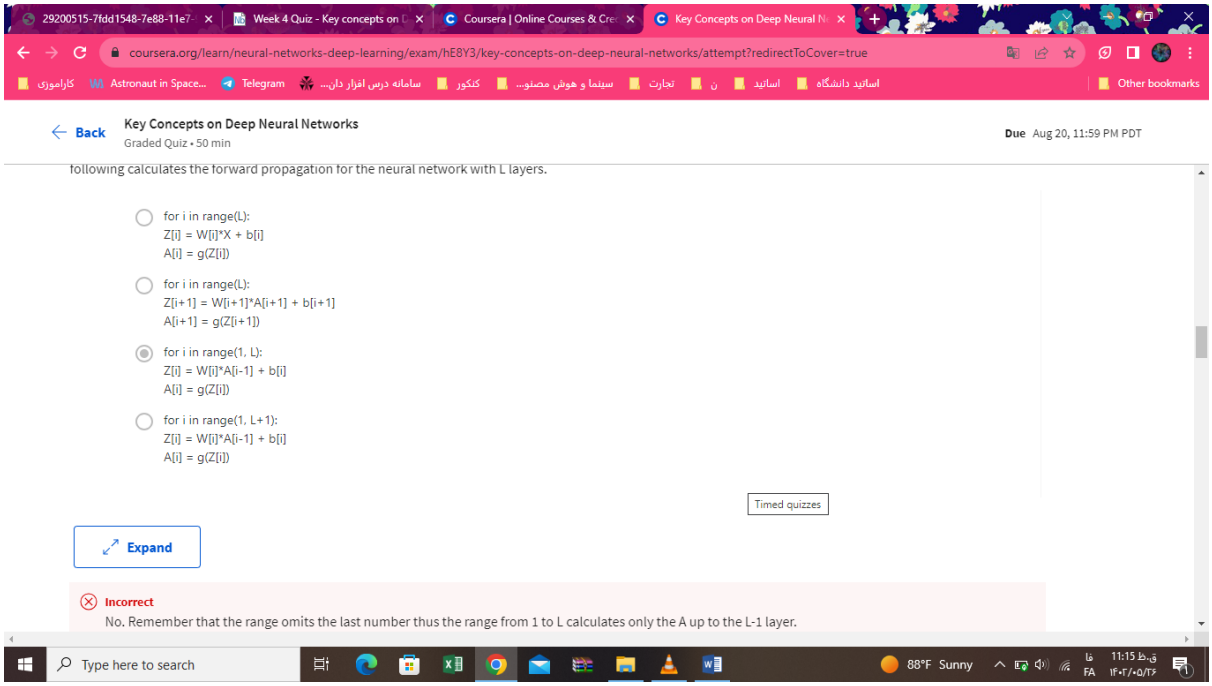
Vectorization allows you to compute forward propagation in an L-layer neural network without an explicit for-loop (or any other explicit iterative loop) over the layers l=1, 2, …,L. True/False?

- ○ True
- ◉ False

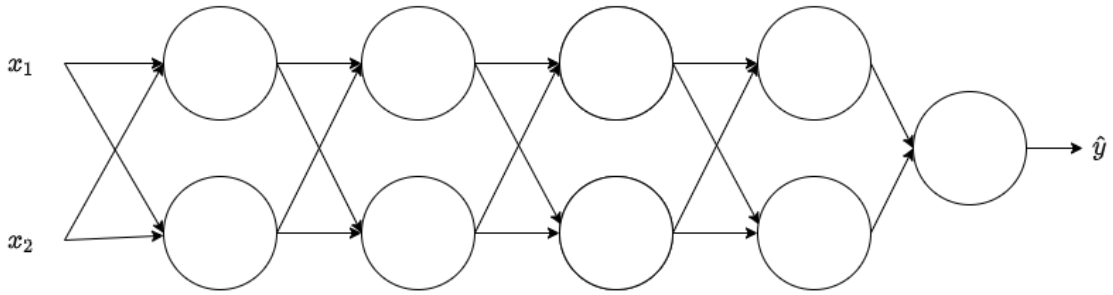
Expand

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: ($a^{[l]} = g^{[l]}(z^{[l]})$, $z^{[l]} = W^{[l]}a^{[l-1]} + b^{[l]}$, ...).



Consider the following neural network:



How many layers does this network have?

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

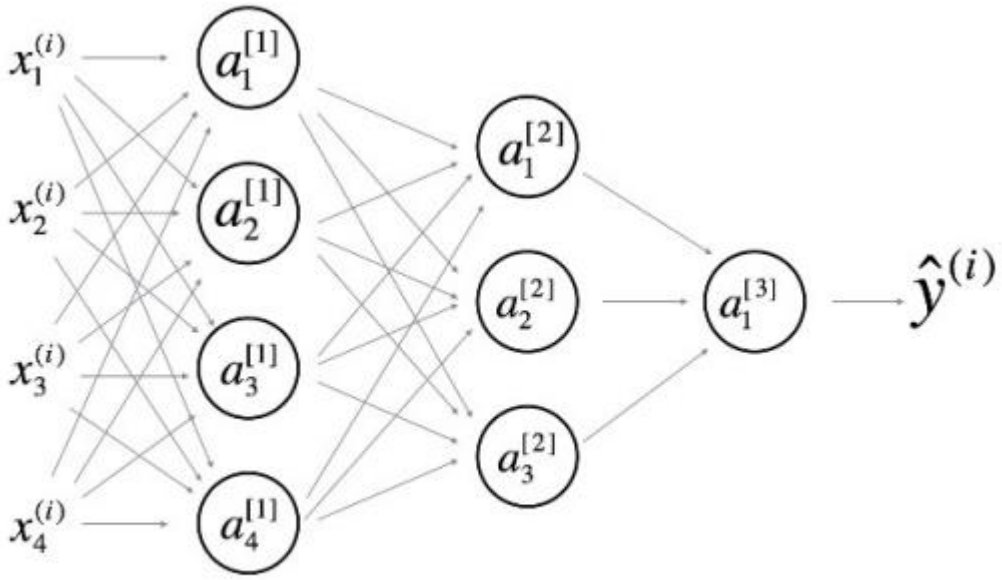
During forward propagation, in the forward function for a layer l you need to know what is the activation function in a layer (sigmoid, tanh, ReLU, etc.). During backpropagation, the corresponding backward function also needs to know what is the activation function for layer l, since the gradient depends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

- 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.



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

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

. Whereas the previous question used a specific network, in the general case what is the dimension of W^{[l]}, the weight matrix associated with layer l?

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

Expand

Correct
True