Congratulations! You passed!

Grade received 100% **Latest Submission Grade** 100% **To pass** 80% or higher

Go to next item

1. What is the "cache" used for in our implementation of forward propagation and backward propagation?

1 / 1 point

- ○ It is used to cache the intermediate values of the cost function during training.
- ○ We use it to pass variables computed during backward propagation to the corresponding forward propagation step. It contains useful values for forward propagation to compute activations.
- ◉ We use it to pass Z computed during forward propagation to the corresponding backward propagation step. It contains useful values for backward propagation to compute derivatives.
- ○ It is used to keep track of the hyperparameters that we are searching over, to speed up computation.

Expand

Correct
Correct, the "cache" records values from the forward propagation units and are used in backward propagation units because it is needed to compute the chain rule derivatives.

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

1 / 1 point

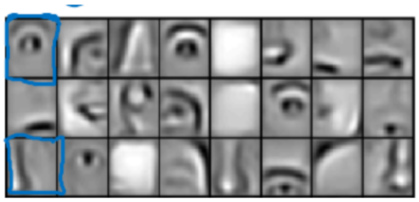
- ○ True
- ◉ False

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. Which of the following is more likely related to the early layers of a deep neural network?

1 / 1 point

○ 

Expand

Correct
Yes. The early layer of a neural network usually computes simple features such as edges and lines.

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 (selected)

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

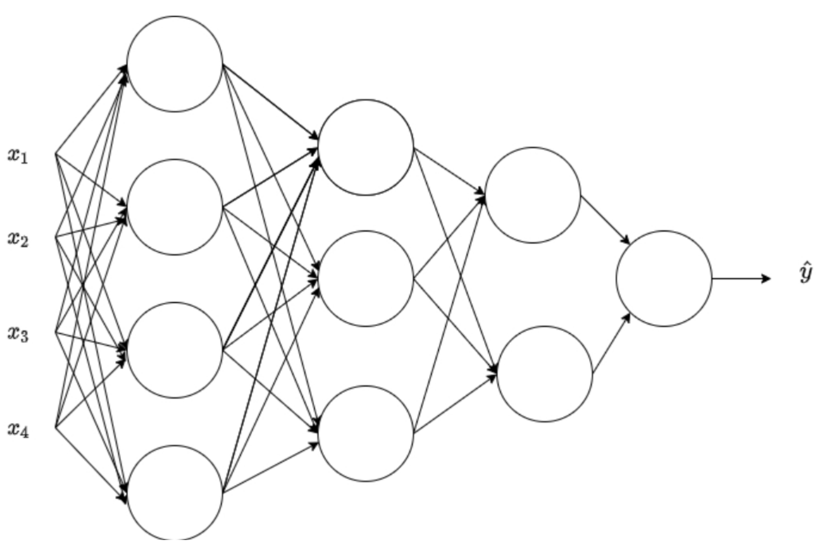
- ```
  for i in range(L):
  Z[i+1] = W[i+1]*A[i+1] + b[i+1]
  A[i+1] = g(Z[i+1])
  ```
- ```
  for i in range(L):
  Z[i] = W[i]*X + b[i]
  A[i] = g(Z[i])
  ```
- (selected)
  ```
  for i in range(1, L+1):
  Z[i] = W[i]*A[i-1] + b[i]
  A[i] = g(Z[i])
  ```
- ```
  for i in range(1, L):
  Z[i] = W[i]*A[i-1]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


  ```

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

- (selected) 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. If L is the number of layers of a neural network then $dZ^{[L]} = A^{[L]} - Y$. True/False?

1 / 1 point

- False
- (selected) True

Expand

Correct
Correct. The gradient of the output layer depends on the difference between the value computed during the forward propagation process and the target values.

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

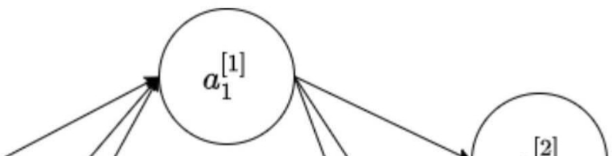
- (selected) False
- True

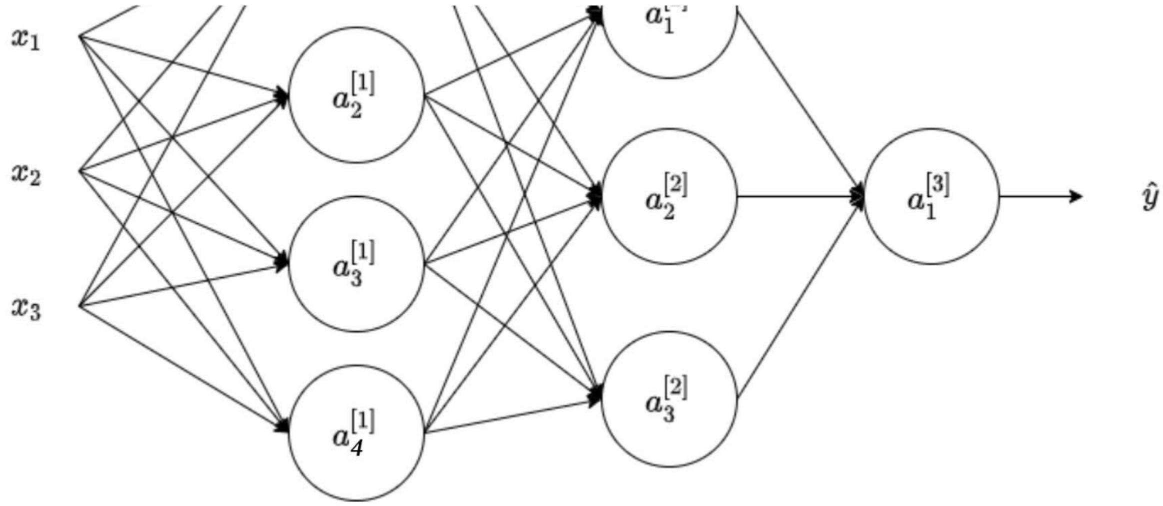
Expand

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).

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

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

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

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

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

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

- (•) $(n^{[l]}, m)$
- ( ) $(m, n^{[l+1]})$
- ( ) $(n^{[l+1]}, m)$
- ( ) $(m, n^{[l]})$

Expand

Correct

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