1. We use the "cache" in our implementation of forward and backward propagation to pass useful values to the next layer in the forward propagation. True/False?

0 / 1 point

- (selected) True
- False

Expand

**Incorrect**

Incorrect. The "cache" is used in our implementation to store values computed during forward propagation to be used in backward propagation.

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

1 / 1 point

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

✓ Correct

- [x] learning rate $\alpha$

✓ Correct

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

✓ Correct

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

✓ Correct

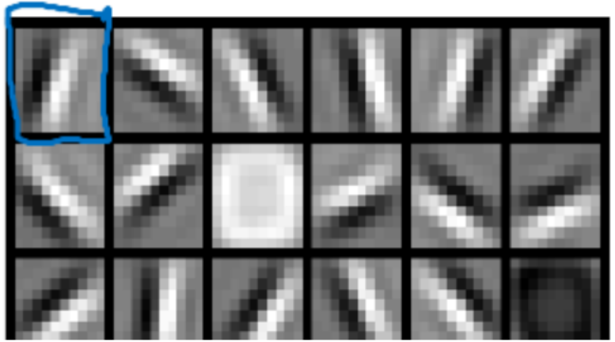
Expand

**Correct**

Great, you got all the right answers.

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. 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**

- ○ True
- ◉ False

Expand

**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.

**0 / 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+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])

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

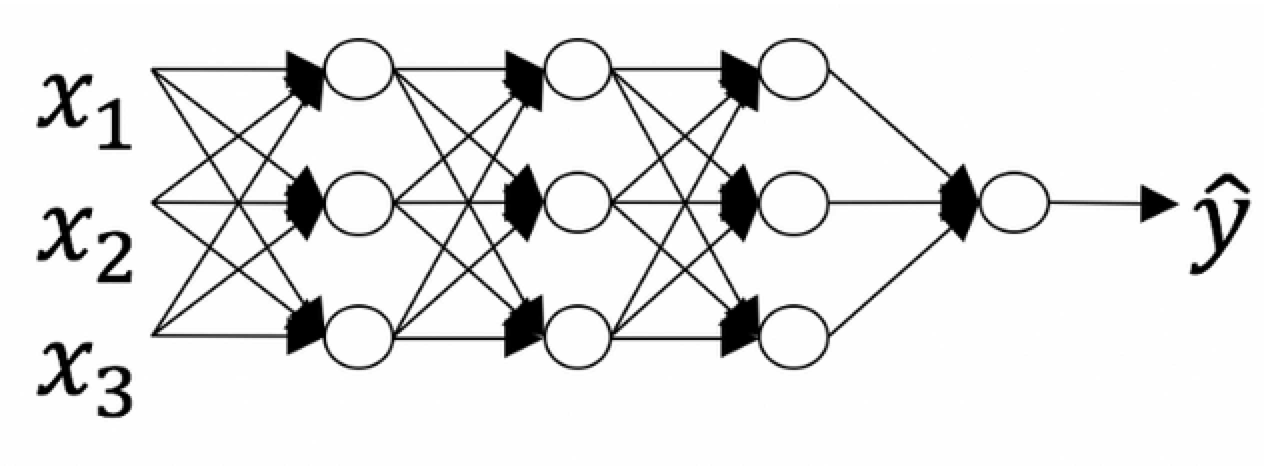
Expand

**Incorrect**

No. Remember that the range omits the last number thus the range from 1 to L calculates only the A up to the L-1 layer.

6. Consider the following neural network.

1 / 1 point




How many layers does this network have?

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

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

1 / 1 point

- [x] True
  Yes. The gradient of the output layer depends on the difference between the value computed during the forward propagation process and the target values.
- [ ] False
  No. The gradient of the output layer depends on the difference between the value computed during the forward propagation process and the target values.

Expand

Correct

**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

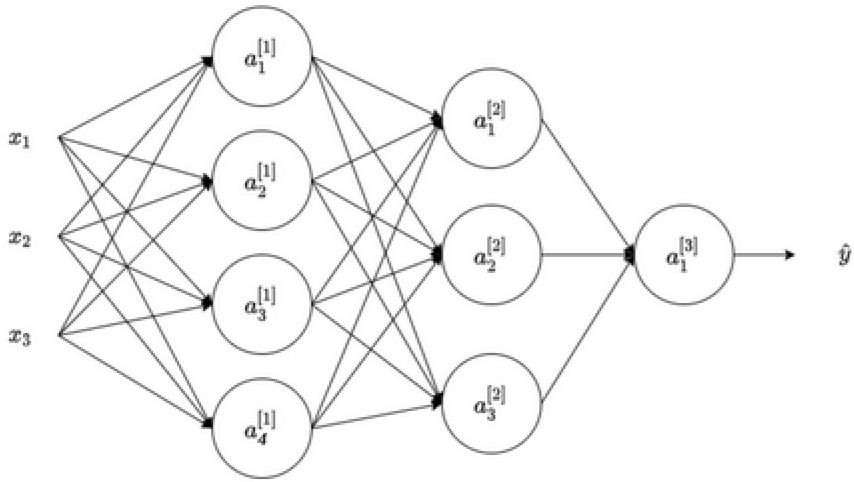
Expand

**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).

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

Correct

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

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

Correct

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

**10.** In the general case if we are training with $m$ examples what is the shape of $A^{[l]}$?

1 / 1 point

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

Expand

**Correct**

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