WEEK-4

KEY CONCEPT ON DEEP NEURAL NETWORK

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

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

2. Which of the following are "parameters" of a neural network? (Check all that apply.)

- ☐ $b^{[l]}$ the bias vector.
- ☑ $W^{[l]}$ the weight matrices.

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

- ☐ $g^{[l]}$ the activation functions.
- ☐ L the number of layers of the neural network.

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

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

> ✓ **Correct**
> Correct. The deep layers of a neural network are typically computing more complex features such as the

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?

- (●) False
- ( ) True

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. Assume we store the values for $n^{[l]}$ in an array called layer_dims, as follows: layer_dims = [n_x, 4,3,2,1]. So lay four hidden units, layer 2 has 3 hidden units and so on. Which of the following for-loops will allow you to init the parameters for the model?

- ( )

```
for i in range(1, len(layer_dims)/2):
parameter['W' + str(i)] = np.random.randn(layer_dims[i], layer_dims[i-1]) * 0.01
parameter['b' + str(i)] = np.random.randn(layer_dims[i-1], 1) * 0.01
```

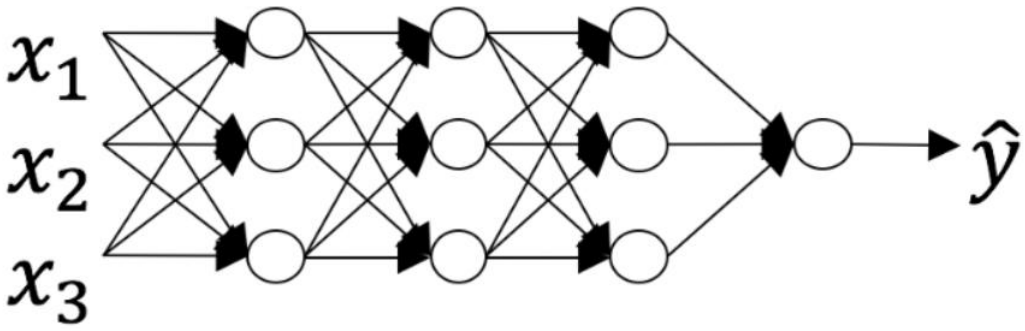
- ( )

```
for i in range(1, len(layer_dims)/2):
parameter['W' + str(i)] = np.random.randn(layer_dims[i], layer_dims[i-1]) * 0.01
parameter['b' + str(i)] = np.random.randn(layer_dims[i], 1) * 0.01
```

- (●)

```
for i in range(1, len(layer_dims)):
parameter['W' + str(i)] = np.random.randn(layer_dims[i], layer_dims[i-1]) * 0.01
parameter['b' + str(i)] = np.random.randn(layer_dims[i], 1) * 0.01
```

6. Consider the following neural network.



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 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.

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

- ○ 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.

- ◉ 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.

✓ **Correct**

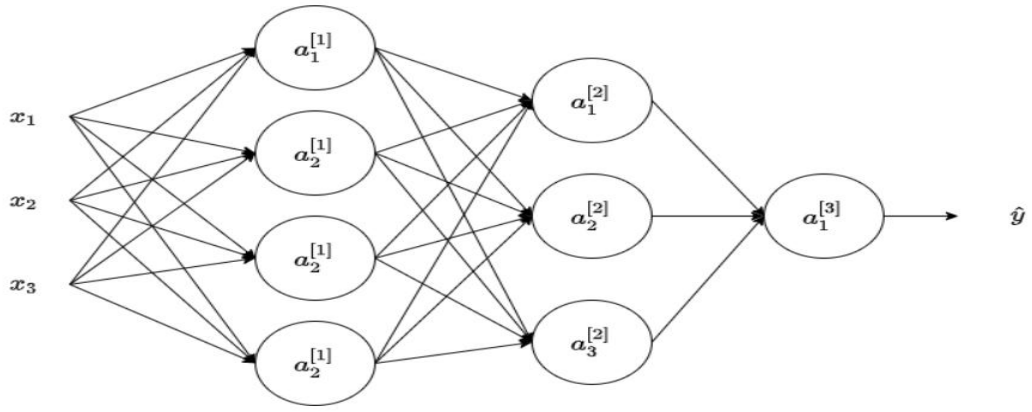
8. There are certain functions with the following properties:

(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?

- ◉ True
- ○ False

✓ **Correct**

9. Consider the following 2 hidden layers neural network:



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

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

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

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

This should not be selected
No. More generally, the shape of $W^{[l]}$ is $(n^{[l]}, n^{[l-1]})$.

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

This should not be selected
No. More generally, the shape of $W^{[l]}$ is $(n^{[l]}, n^{[l-1]})$.

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

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

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

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