Note: In here I am adding True or False, it can be Right or Wrong both first Check If option is right or wrong under that question summary. And Rest question which I am giving multiple answer question or single choice question both are only right answers, and there is lot's of changing questions are there so please check all the question first and then only select the correct option. Also In Multiple Question I have selected all the option so it will show you all he correct option so based on that select the right answers.

Final Attempt:

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.

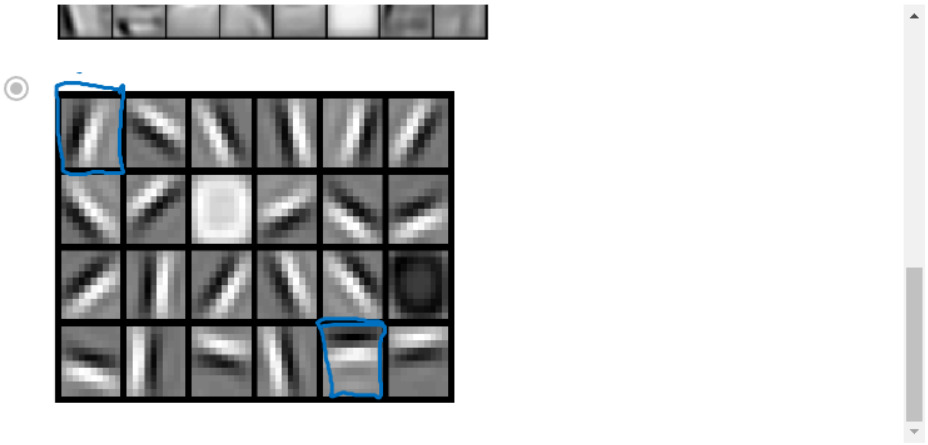
2. During the backpropagation process, we use gradient descent to change the hyperparameters. True/False? **1 / 1 point**

- ○ 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.

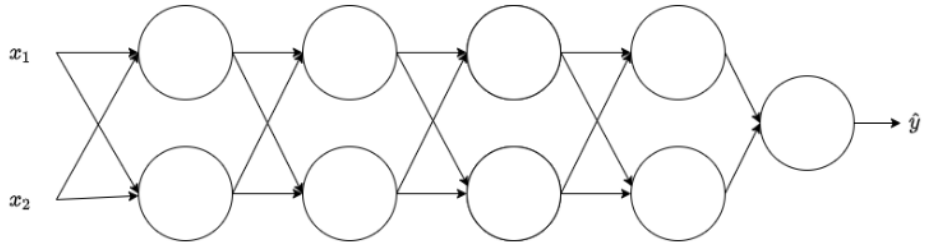
4. We can not use vectorization to calculate $da^{[l]}$ in backpropagation, we must use a for loop over all the examples. True/False? 1 / 1 point

- False
- True

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.

6. Consider the following neural network: **1 / 1 point**



How many layers does this network have?

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

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? **1 / 1 point**

- 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.

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**

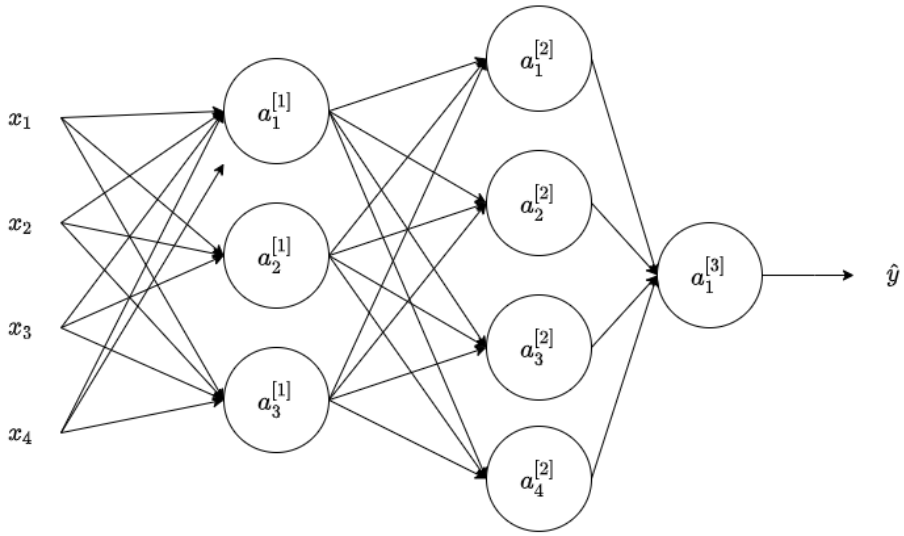
- ○ True
- ◉ False

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: **0 / 1 point**



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

- ☑ $W^{[2]}$ will have shape (1, 3)

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

- ☑ $b^{[1]}$ will have shape (3, 1)

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

- ☑ $b^{[1]}$ will have shape (4, 1)

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

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

  ✓ **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 (4, 3)

  ! **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 (1, 3)

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

- [x] $W^{[2]}$ 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, 4)

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

Expand

⊗ **Incorrect**
You chose the extra incorrect answers.

10. 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?

1 / 1 point

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

Expand

Correct
True