

Congratulations! You passed!

TO PASS 80% or higher



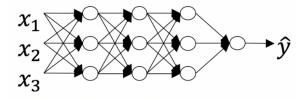
grade 100%

1. What is the "cache" used for in our implementation of forward propagation and backward propagation? □ It is used to cache the intermediate values of the cost function during training. □ We use it to pass variables computed during forward propagation to the corresponding backward propagation step. It contains useful values for box useful values for showard propagation to the corresponding backward propagation step. It contains useful values for broward propagation compute definations. □ We use it to pass variables computed during backward propagation to the corresponding forward propagation step. It contains useful values for forward propagation computed that the computed that the step. It contains useful values for forward propagation units and sends it to the backward propagation units because it is needed to compute the chain rule definations. □ Correct □ Intermity of the following, which ones are "hyperparameters"? (Check all that apply) □ It is used to keep track of the hidden layers all to the backward propagation units and sends it to the backward propagation units and sends	Key concepts on Deep Neural Networks LATEST SUBMISSION GRADE			
It is used to cache the intermediate values of the cost function during forward propagation to the corresponding backward propagation costs. Recombination and the cost of page 2015 of the company of propagation to the corresponding backward propagation costs. We used it is pass variables compaced during beloward propagation to the corresponding backward propagation costs. Propagation to the corresponding forward propagation costs.				
We use it to gass variables computed during forward propagation to the corresponding backward propagation stops; it contains useful values for backward propagation to compute demanders. (It is used to keep track of the hyperparameters that we are sarching over, to spead up computation. (It is used to keep track of the hyperparameters that we are sarching over, to spead up computation. (It is used to keep track of the hyperparameters that we are sarching over, to spead up computation. (It is used to keep track of the hyperparameters that we are sarching over, to spead up computation. (It is used to keep track of the hyperparameters that we are sarching over, to spead up computation. (It is used to keep track of the hyperparameters that we are sarching over, to spead up computation. (It is used to keep track of the hyperparameters that we are sarching over, to spead up computation. (It is used to keep track of the hyperparameters to keep the shall not dimension. (It is used to keep track of the hyperparameters to keep track of the following which ones are "hyperparameters" (Check all that apply) (It is used to the backward propagation in the forward propagation by the hyperparameters of the hyperparameters is true? (It is consistent to the following statements is true? (It is used to the following statements is true? (It is used to the following statements is true? (It is used to the following statements is true? (It is used to the following statements is true? (It is a relief in the hyperparameters of the hyperparameters of the input than the deeper layer. (It is used to the following statements is true? (It is used to the following statements is true? (It is used to the following statements is true? (It is used to the following statements is true? (It is used to the following statements is true? (It is used to the following statements is true? (It is used to the following statements is true? (It	1.	What is the "cache" used for in our implementation of forward propagation and backward propagation?	1 / 1 point	
we set to pass unable compacts during beloading prospect for the prospection of the pros				
to use to keep track of the hyperparameters that we are searching over, to speed up computation. It is used to keep track of the hyperparameters that we are searching over, to speed up computation. Correct Correct the "cache" records values from the foreweld propagation units and sends it to the backward propagation units because it is needed to compute the chain rule derivatives. 2. Arrong the following, which ones are "hyperparameters"? (Check all that apply) bias vectors \$\textit{\textit{b}}\$ size of the holden layers \$n^{ }\$ to correct tearning rate o currect rumber of layers \$L\$ in the neural network currect weight matrices \$W^{ }\$ number of layers \$d\$ in the neural network currect activation values \$a^{ }\$ 3. Which of the following statements is true? at the depart layers of a neural network are typically computing more complex features of the input than the earlier layers. The despite layers of a neural network are typically computing more complex features of the input than the desper layers. Correct Correct The despite layers of a neural network are typically computing more complex features of the input than the desper layers. Correct Correct Correct The parameter \$\text{correct}\$ is true? The despite layers of a neural network are typically computing more complex features of the input than the desper layers. Correct Correct The parameter \$\text{correct}\$ is \$\text{correct}\$ in \$\text{correct}\$				
 ✓ Correct Correct the "cache" records values from the forward propagation units and sends it to the backward propagation units because its needed to compute the chain rule denombres. 2. Among the following, which ones are "hyperparametees" (Check all that apply) 2. Among the following, which ones are "hyperparametees" (Check all that apply) 2. Among the following which ones are "hyperparametees" (Check all that apply) 3. Size of the hidden layers n^{2/3} ✓ Correct 2. Inaming rate or ✓ Correct 3. Which of the following statements is true? 4. Which of the following statements is true? 3. Which of the following statements is true? 4. Which of the following statements is true? 5. The satisfier layers of a neural network are typically computing more complex features of the input than the earlier layers. 7. The satisfier layers of a neural network are typically computing more complex features of the input than the deeper layers. 7. Correct 4. Vectorization allows you to compute forward propagation in an L-layer neural network without an explicit for-loop for any other layers (all the lines (all all all all all all all all all al				
Correct. The Cracker Correct sets along the Correct sets of the Section of the S		It is used to keep track of the hyperparameters that we are searching over, to speed up computation.		
bias vectors b ⁽ⁱ⁾ size of the hidden layers n ⁽ⁱ⁾ √ Correct learning rate α √ Correct number of iterations √ Correct weight matrices W ⁽ⁱ⁾ √ In the neural network weight matrices W ⁽ⁱ⁾ number of layers L in the neural network √ Correct activation values a ⁽ⁱ⁾ 3. Which of the following statements is true? activation values a ⁽ⁱ⁾ 3. Which of the following statements is true? activation values a ⁽ⁱ⁾ 6. The deeper layers of a neural network are typically computing more complex features of the input than the earlier layers. The earlier layers of a neural network are typically computing more complex features of the input than the deeper layers. √ Correct		Correct, the "cache" records values from the forward propagation units and sends it to the backward		
stee of the hidden layers n ^[i] ✓ Correct learning rate α ✓ correct number of iterations ✓ correct weight matrices W ^[i] number of layers <i>L</i> in the neural network ✓ correct activation values a ^[i] attivation values a ^[i] The deeper layers of a neural network are typically computing more complex features of the input than the earlier layers. The earlier layers of a neural network are typically computing more complex features of the input than the deeper layers. ✓ Correct Correct True False Forward propagation propagations the input through the layers, although for shallow networks we may just write all the lines (a ^[i] = g ^[i] (z ^[i]), z ^[i] = W ^{[i}	2.	Among the following, which ones are "hyperparameters"? (Check all that apply.)	1/1 point	
learning rate α Correct In unifor of iterations Correct In unifor of iterations Correct In weight matrices W ^(f) In unifor of layers <i>L</i> in the neural network Correct In the deeper layers of a neural network are typically computing more complex features of the input than the earlier layers. The earlier layers of a neural network are typically computing more complex features of the input than the deeper layers. The earlier layers of a neural network are typically computing more complex features of the input than the deeper layers. Correct Correct True False False Correct Forward propagation propagates the input through the layers, although for shallow networks we may just write all the lines (a ^(f) = g ^(f) (z ^(g)), z ^(g) = W ^(g) (a ^(g) + B ^(g) , -1) in a deeper network, we cannot avoid a for loop iterating over the layers (a ^(g) = g ^(g) (z ^(g)), z ^(g) = W ^(g) (a ^(g) - 1) + B ^(g) , -1) in a deeper network, we cannot avoid a for loop iterating over the layers (a ^(g) = g ^(g) (z ^(g)), z ^(g) = W ^(g) (a ^(g) - 1) + B ^(g) , -1) in a deeper network, we cannot avoid a for loop iterating over the layers (a ^(g) = g ^(g) (z ^(g)), z ^(g) = W ^(g) (a ^(g) - 1) + B ^(g) , -1) in a deeper network, we cannot avoid a for loop iterating over the layers (a ^(g) = g ^(g) - g ^(g) = W ^(g) (a ^(g) - 1) + B ^(g) , -1) in a deeper network, we cannot avoid a for loop iterating over the layers (a ^(g) = g ^(g) (z ^(g)), z ^(g) = W ^(g) (a ^(g) - 1) + B ^(g) , -1) in a deeper network, we cannot avoid a for loop iterating over the layers (a ^(g) = g ^(g) (z ^(g)), z ^(g) = G ^(g) (z ^(g)), -1 in a cannot (a ^(g) - 1) + B ^(g) , -1 in a range(1), len(layer, diss)/21; -1 in a range(1), len(layer, diss)/		$igsquare$ bias vectors $b^{[l]}$		
learning rate α		$igstyle{igstyleigytureigstyleigytureigstyleigytureigstyleigytureigstyleigytureigstyleigytureigstyleigytureigstyleigytureigstyleigytureigytureigwyanaigatureigwyanaig$		
✓ Correct □ number of iterations ✓ correct □ weight matrices W ^(f) ☑ number of layers L in the neural network ✓ Correct □ activation values a ^(f) ③ The deeper layers of a neural network are typically computing more complex features of the input than the earlier layers. ○ The earlier layers of a neural network are typically computing more complex features of the input than the deeper layers. ✓ Correct 4. Vectorization allows you to compute forward propagation in an L-layer neural network without an explicit for-loop (or any of the layers in 1, 2,, TrueFalse? ○ True ⑥ False ✓ Correct Forward propagation propagates the input through the layers, although for shallow networks we may just write all the lines (a ^(f) = g ^(f) (z ^(f)), z ^(f) = W ^(f) (d ^(f) + B ^(f) ,) in a deeper network, we cannot avoid a for loop iterating over the layers (a ^(f) = g ^(f) (z ^(f)), z ^(f) = W ^(f) (d ^(f) + B ^(f) ,) 5. Assume we store the values for n ^(f) in an array called layer, dims, as follows: layer, dims = (n _x , 4.3.2,1). So layer 1 has four the interaction of the model? 1 for i in range(1, len(layer,dims)/2): 2 parameter(1) ^(g) + str(1) = np, random, randon (layer_dims[1], layer_dims[1-1]) + 0.01 1 for i in range(1, len(layer_dims)/2): 2 parameter(1) ^(g) + str(1) = np, random, randon (layer_dims[1], layer_dims[1-1]) + 0.01 1 for i in range(1, len(layer_dims)/2): 2 parameter(1) ^(g) + str(1) = np, random randon (layer_dims[1], layer_dims[1-1]) + 0.01 1 for i in range(1, len(layer_dims)/2): 2 parameter(1) ^(g) + str(1) = np, random randon (layer_dims[1], layer_dims[1-1]) + 0.01 1 for i in range(1, len(layer_dims)/2): 2 parameter(1) ^(g) + str(1) = np, random randon (layer_dims[1], layer_dims[1-1]) + 0.01		✓ Correct		
 ✓ correct weight matrices W⁽ⁱ⁾ ✓ number of layers <i>L</i> in the neural network ✓ correct activation values a⁽ⁱ⁾ 3. Which of the following statements is true? (activation values a⁽ⁱ⁾ (b) The deeper layers of a neural network are typically computing more complex features of the input than the earlier layers. (c) The earlier layers of a neural network are typically computing more complex features of the input than the deeper layers. (c) The earlier layers of a neural network are typically computing more complex features of the input than the deeper layers. (d) Correct (e) True (e) False (f) False (g) False (g) False (g) False (h) False		$igstyle{igstyleigwyalaig$		
 ✓ Correct weight matrices W^[I] ✓ number of layers L in the neural network ✓ Correct activation values a^[I] 3. Which of the following statements is true? The deeper layers of a neural network are typically computing more complex features of the input than the earlier layers. The earlier layers of a neural network are typically computing more complex features of the input than the deeper layers. ✓ Correct 4. Vectorization allows you to compute forward propagation in an L-layer neural network without an explicit for-loop (or any or layers). True False ✓ Correct Forward propagation propagates the input through the layers, although for shallow networks we may just write all the lines (a^[S] = g^[S](z^[S]), z^[S] = W^[I](a^[S] + b^[S],) in a deeper network, we cannot avoid a for loop iterating over the layers; (a^[S] = g^[S](z^[S]), z^[S] = W^[I](a^[S] + b^[S],), in a deeper network, we cannot avoid a for loop iterating over the layers; (a^[S] = g^[S](z^[S]), z^[S] = W^[I](a^[S] + b^[S],) in a deeper network, we cannot avoid a for loop iterating over the layers; (a^[S] = g^[S](z^[S]), z^[S] = W^[I](a^[S] + b^[S],) in a deeper network, we cannot avoid a for loop iterating over the layers; (a^[S] = g^[S](z^[S]), z^[S] = W^[I](a^[S] + b^[S],) in a deeper network, we cannot avoid a for loop iterating over the layers; (a^[S] = g^[S](z^[S]), z^[S] = W^[I](a^[S] + b^[S],) in a deeper network, we cannot avoid a for loop iterating over the layers; (a^[S] = g^[S](z^[S]), z^[S] = W^[I](a^[S] + b^[S],) in a deeper network, we cannot avoid a for loop iterating over the layers; (a^[S] = g^[S](z^[S]), z^[S] = w^[S](a^[S] + b^[S],) in a deeper network, we cannot avoid a for loop iterating over the layers; (a^[S] = g^[S](z^[S]), z^[S] = w^[S](a^[S], a^[S], a^[S], a^[S], a^[S], a^[S], a^[S], a^[S], a^[S],		✓ Correct		
weight matrices W ^[i] In uniber of layers L in the neural network Correct activation values a ^[i] Which of the following statements is true? The deeper layers of a neural network are typically computing more complex features of the input than the earlier layers. The earlier layers of a neural network are typically computing more complex features of the input than the deeper layers. The earlier layers of a neural network are typically computing more complex features of the input than the deeper layers. The earlier layers of a neural network are typically computing more complex features of the input than the deeper layers. The earlier layers of a neural network are typically computing more complex features of the input than the deeper layers. The earlier layers of a neural network are typically computing more complex features of the input than the deeper layers. The earlier layers of a neural network are typically computing more complex features of the input than the deeper layers. The earlier layers of a neural network are typically computing more complex features of the input than the deeper layers. The earlier layers of a neural network are typically computing more complex features of the input than the earlier layers. The earlier layers of a neural network are typically computing more complex features of the input than the earlier layers. The earlier layers of the input than the earlier layer neural network without an explicit for-loop (or any 1/1) point write all the lines (a ^[i] = g ^[i] (2 ^[i]), z ^[i] = W ^[i] (2 ^[i]), z ^[i] in a deeper neural network without an explicit for-loop (or any 1/1) point layer. The earlier layers of a neural network are typically computing more complex features of the input than the earlier layer. The deeper layers of the input than the earlier layer. The deeper layers of the input than the earlier layer. The deeper layers of the input than the earlier layer. The deeper layers of the input than the earlier layer. The deeper layers of the input than t		✓ number of iterations		
 ✓ correct 3. Which of the following statements is true? 1/1 point The deeper layers of a neural network are typically computing more complex features of the input than the earlier layers. The earlier layers of a neural network are typically computing more complex features of the input than the deeper layers. ✓ correct 4. Vectorization allows you to compute forward propagation in an <i>L</i>-layer neural network without an explicit for-loop (or any other explicit iterative loop) over the layers [=1, 2,, L. True/False? ☐ True ⑥ False ✓ correct Forward propagation propagates the input through the layers, although for shallow networks we may just write all the lines (a^[2] = g^[2](2^[2]), 2^[2] = W^[2](a^[1] + b^[2],) in a deeper network, we cannot avoid a for loop iterating over the layers: (a^[3] = g^[3](2^[3]), 2^[3] = W^[3](a^[3] + b^[3],) in a deeper network, we cannot avoid a for loop iterating over the layers: (a^[3] = g^[3](2^[3]), 2^[3] = W^[3](a^[3] + b^[3],) in a deeper network, we cannot avoid a for loop iterating over the layers: (a^[3] = g^[3](2^[3]), 2^[3] = W^[3](a^[3] + b^[3],) in a deeper network, we cannot avoid a for loop iterating over the layers: (a^[3] = g^[3](2^[3]), 2^[3] = W^[3](a^[3] + b^[3],) in a deeper network, we cannot avoid a for loop iterating over the layers: (a^[3] = g^[3](2^[3]), 2^[3] = W^[3](a^[3] + b^[3],) in a deeper network, we cannot avoid a for loop iterating over the layers: (a^[3] = g^[3](2^[3]), 2^[3] = g^[3](a^[3]), 2^[3] = p^[3](a^[3]), 2^[3] = p^[3](a^[3]		✓ Correct		
 ✓ Correct activation values a 		$igcup$ weight matrices $W^{[l]}$		
activation values $a^{[i]}$ 3. Which of the following statements is true? ① The deeper layers of a neural network are typically computing more complex features of the input than the earlier layers. ① The earlier layers of a neural network are typically computing more complex features of the input than the deeper layers. ② Correct 4. Vectorization allows you to compute forward propagation in an <i>L</i> -layer neural network without an explicit for-loop (or any other explicit iterative loop) over the layers = 1, 2,, L. True/False? ② True ③ False ② Correct Forward propagation propagates the input through the layers, although for shallow networks we may just write all the lines ($a^{[i]} = g^{[i]}(z^{[i]})$, $z^{[i]} = W^{[i]}a^{[i]} + b^{[i]}$,) in deeper network, we cannot avoid a for loop iterating over the layers: ($a^{[i]} = g^{[i]}(z^{[i]})$, $z^{[i]} = W^{[i]}a^{[i-1]} + b^{[i]}$,). 5. Assume we store the values for $n^{[i]}$ in an array called layer, dims, as follows: layer_dims = {n_i, 4, 3,2,1}. So layer 1 has four hidden units, layer 2 has 3 hidden units and so on. Which of the following for-loops will allow you to initialize the parameters for the mode? ② 1 for i in range(1, len(layer_dims)/2): 2		lacksquare number of layers L in the neural network		
 Which of the following statements is true? The deeper layers of a neural network are typically computing more complex features of the input than the earlier layers. The earlier layers of a neural network are typically computing more complex features of the input than the deeper layers. ✓ Correct Vectorization allows you to compute forward propagation in an <i>L</i>-layer neural network without an explicit for-loop (or any other explicit iterative loop) over the layers =1, 2,, L. True/False? True False ✓ 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](z^[1] + b^[2],) in a deeper network, we cannot avoid a for loop iterating over the layers: (a^[1] = g^[1](z^[2]), z^[2] = W^[1](a^[1] + b^[1],). 5. Assume we store the values for n^[2] in an array called layer, dims, as follows: layer_dims = [n₂, 4,3,2,1]. So layer 1 has four hidden units, layer 2 has 3 hidden units and so on. Which of the following for-loops will allow you to initialize the parameters for the model? 1 for i in range(1, len(layer_dims)/2): 2 parameter 1		✓ Correct		
 The deeper layers of a neural network are typically computing more complex features of the input than the earlier layers. The earlier layers of a neural network are typically computing more complex features of the input than the deeper layers. Correct Vectorization allows you to compute forward propagation in an <i>L</i>-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 Correct Forward propagation propagates the input through the layers, although for shallow networks we may just write all the lines (a^[2] = g^[3](e^[2]), e^[2] = W^[3] a^[1] + b^[3],) in a deeper network, we cannot avoid a for loop iterating over the layers: (a^[1] = g^[3](e^[2]), e^[3] = W^[3] a^[1] + b^[3],). 5. Assume we store the values for n^[4] in an array called layer_dims, as follows: layer_dims = [n₂, 4,3,2,1]. So layer 1 has four hidden units, layer 2 has 3 hidden units and so on. Which of the following for-loops will allow you to initialize the parameters for the model? 1 for 1 in range(1, len(layer_dims)/2): 2		\square activation values $a^{[l]}$		
layers.	3.	Which of the following statements is true?	1 / 1 point	
The earlier layers of a neural network are typically computing more complex features of the input than the deeper layers. ✓ Correct 4. Vectorization allows you to compute forward propagation in an <i>L</i> -layer neural network without an explicit for-loop (or any 1/1 point other explicit iterative loop) over the layers !=1, 2,,L. True/False? ○ True ⑥ False ✓ Correct Forward propagation propagates the input through the layers, although for shallow networks we may just write all the lines (a!² = g²²(z²²), z²² = W²² a!¹ + b²³,, in a deeper network, we cannot avoid a for loop iterating over the layers: (a!² = g¹²(z²²), z²² = W²¹ a²¹ + b²³,), in a deeper network, we cannot avoid a for loop iterating over the layers: (a!² = g¹²(z²²), z²² = W²¹ a²¹ + b²³,), 5. Assume we store the values for n²¹ in an array called layer_dims, as follows: layer_dims = [n₂, 4,3,2,1]. So layer 1 has four 1/1 point hidden units, layer 2 has 3 hidden units and so on. Which of the following for-loops will allow you to initialize the parameters for the model? ○ 1 for i in range(1, len(layer_dims)/2): 2				
 ✓ Correct 4. Vectorization allows you to compute forward propagation in an <i>L</i>-layer neural network without an explicit for-loop (or any other explicit iterative loop) over the layers !=1, 2,,L. True/False? ○ True ⑥ False ✓ 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^[3] = W^[3](z^[3]) in a deeper network, we cannot avoid a for loop iterating over the layers: (a^[4] = g^[4](z^[4]), z^[4] = W^[4] a^[4-1] + b^[4],). 5. Assume we store the values for n^[4] in an array called layer_dims, as follows: layer_dims = [n_x, 4,3,2,1]. So layer 1 has four 1/1 point hidden units, layer 2 has 3 hidden units and so on. Which of the following for-loops will allow you to initialize the parameters for the model? ○ 1 for i in range(1, len(layer_dims)/2): 2 parameter ['W' + str(1)] = np. random. randn(layer_dims[1], layer_dims[1-1]) * 0.01 ○ 1 for i in range(1, len(layer_dims)/2): 2 parameter ['W' + str(1)] = np. random. randn(layer_dims[1], layer_dims[1-1]) * 0.01 		The earlier layers of a neural network are typically computing more complex features of the input than the deeper		
 4. Vectorization allows you to compute forward propagation in an <i>L</i>-layer neural network without an explicit for-loop (or any other explicit iterative loop) over the layers !=1, 2,,L. True/False? ○ True ⑥ False ✓ 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^[3] = W^[3] g^[3] + b^[3],) in a deeper network, we cannot avoid a for loop iterating over the layers: (a^[4] = g^[4](z^[4]), z^[4] = W^[4] g^[4] = W^[4] g^[4] + b^[4],). 5. Assume we store the values for n^[4] in an array called layer_dims, as follows: layer_dims = [n_x, 4,3,2,1]. So layer 1 has four 1/1 point hidden units, layer 2 has 3 hidden units and so on. Which of the following for-loops will allow you to initialize the parameters for the model? ○ 1 for i in range(1, len(layer_dims)/2): 2		•		
other explicit iterative loop) over the layers I=1, 2,,L. True/False? True False ✓ 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] + β ^[2] ,) in a deeper network, we cannot avoid a for loop iterating over the layers: (a ^[2] = g ^[2] (z ^[2]), z ^[2] = W ^[2] a ^[1] + β ^[2] ,). 5. Assume we store the values for n ^[2] in an array called layer_dims, as follows: layer_dims = [n _x , 4,3,2,1]. So layer 1 has four 1/1 point hidden units, layer 2 has 3 hidden units and so on. Which of the following for-loops will allow you to initialize the parameters for the model? 1 for i in range(1, len(layer_dims)/2): 2 parameter['W' + str(1)] = np. random. randn(layer_dims[i], layer_dims[i-1]) * 0.01 1 for i in range(1, len(layer_dims)/2): 2 parameter['W' + str(i)] = np. random. randn(layer_dims[i], layer_dims[i-1]) * 0.01		Correct		
True	4.		1 / 1 point	
Correct Forward propagation propagates the input through the layers, although for shallow networks we may just write all the lines (a ^[2] = g ^[3] (e ^[2]), e ^[2] = W ^[3] a ^[1] + b ^[2] ,) in a deeper network, we cannot avoid a for loop iterating over the layers: (a ^[1] = g ^[1] (e ^[2]), e ^[1] = W ^[1] a ^[1-1] + b ^[1] ,). 5. Assume we store the values for n ^[3] in an array called layer_dims, as follows: layer_dims = [n ₂ , 4,3,2,1]. So layer 1 has four hidden units, layer 2 has 3 hidden units and so on. Which of the following for-loops will allow you to initialize the parameters for the model? 1 for i in range(1, len(layer_dims)/2): 2 parameter['W' + str(i)] = np.random.randn(layer_dims[i], layer_dims[i-1]) * 0.01 1 for i in range(1, len(layer_dims)/2): 2 parameter['W' + str(i)] = np.random.randn(layer_dims[i], layer_dims[i-1]) * 0.01				
Forward propagation propagates the input through the layers, although for shallow networks we may just write all the lines $(a^{[2]} = g^{[3]}(e^{[2]}), e^{[2]} = W^{[3]}a^{[1]} + b^{[2]},)$ in a deeper network, we cannot avoid a for loop iterating over the layers: $(a^{[1]} = g^{[1]}(e^{[2]}), e^{[1]} = W^{[1]}a^{[1]} + b^{[1]},)$. 5. Assume we store the values for $n^{[1]}$ in an array called layer_dims, as follows: layer_dims = $[n_x, 4.3, 2, 1]$. So layer 1 has four widden units, layer 2 has 3 hidden units and so on. Which of the following for-loops will allow you to initialize the parameters for the model? 1		False		
hidden units, layer 2 has 3 hidden units and so on. Which of the following for-loops will allow you to initialize the parameters for the model? 1		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		
<pre>2</pre>	5.	hidden units, layer 2 has 3 hidden units and so on. Which of the following for-loops will allow you to initialize the	1/1 point	
<pre>parameter['W' + str(i)] = np.random.randn(layer_dims[i], layer_dims[i-1]) * 0.01</pre>		<pre>parameter['W' + str(i)] = np.random.randn(layer_dims[i], layer_dims[i-1]) * 0.01</pre>		
		<pre>1 for i in range(1, len(layer_dims)/2): 2 parameter('w' + str(i)] = np.random.randn(layer_dims[i], layer_dims[i-1]) * 0.01</pre>		



6. Consider the following neural network.

1/1 point



How many layers does this network have?

- igotimes The number of layers L is 4. The number of hidden layers is 3.
- \bigcirc The number of layers L is 3. The number of hidden layers is 3.
- \bigcirc The number of layers L is 4. The number of hidden layers is 4.
- \bigcirc The number of layers L is 5. The number of hidden layers is 4.



Yes. As seen in lecture, the number of layers is counted as the number of hidden layers \pm 1. The input and output layers are not counted as hidden layers.

7. During forward propagation, in the forward function for a layer $\it l$ you need to know what is the activation function in a 1/1 point layer (Sigmoid, tanh, ReLU, etc.). During backpropagation, the corresponding backward function also needs to know what is the activation function for layer \emph{l} , since the gradient depends on it. True/False?

True

O False



Yes, as you've seen in the 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. There are certain functions with the following properties:

1 / 1 point

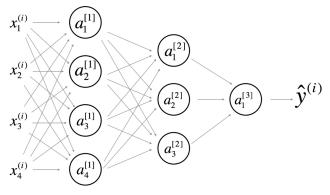
(i) To compute the function using a shallow network circuit, you will need a large network (where we measure size by the number of logic gates in the network), but (ii) To compute it using a deep network circuit, you need only an exponentially smaller network. True/False?

True

O False

9. Consider the following 2 hidden layer neural network:

1 / 1 point



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

 $lacksquare W^{[1]}$ will have shape (4, 4)



