Shallow Neural Networks

1.

Question 1

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

0 / 1 point

Expand

Incorrect

You didn't select all the correct answers

2.

Question 2

The sigmoid function is only mentioned as an activation function for historical reasons. The tanh is always preferred without exceptions in all the layers of a Neural Network. True/False?

0 / 1 point

Expand

Incorrect

No. Although the tanh almost always works better than the sigmoid function when used in hidden layers, thus is always proffered as activation function, the exception is for the output layer in classification problems.

3.

Question 3

Which of the following is a correct vectorized implementation of forward propagation for layer 2?

1 / 1 point

Expand

Correct

Yes. The elements of layer two are represented using a superscript in brackets.

4.

Question 4

When building a binary classifier for recognizing cats (y=1) vs raccoons (y=0). Is better to use the sigmoid function as activation function for the hidden layers. True/False

0 / 1 point

Expand

Incorrect

No. Using tanh almost always works better than the sigmoid function for hidden layers.

5.

Question 5

Consider the following code:

#+begin_src python

x = np.random.rand(3, 2)

y = np.sum(x, axis=0, keepdims=True)

#+end_src

What will be y.shape?

1 / 1 point
Expand

Correct
Yes. By choosing the axis=0 the sum is computed over each column of the array, thus the resulting array is a row vector with 2 entries. Since the option keepdims=True is used the first dimension is kept, thus (1, 2).

6.
Question 6
Suppose you have built a neural network. You decide to initialize the weights and biases to be zero. Which of the following statements is true?

1 / 1 point
Expand

Correct

7.
Question 7
Logistic regression's weights should be initialized randomly rather than to all zeros, because if you initialize to all zeros, then logistic regression will fail to learn a useful decision boundary because it will fail to "break symmetry", True/False?

0 / 1 point
Expand

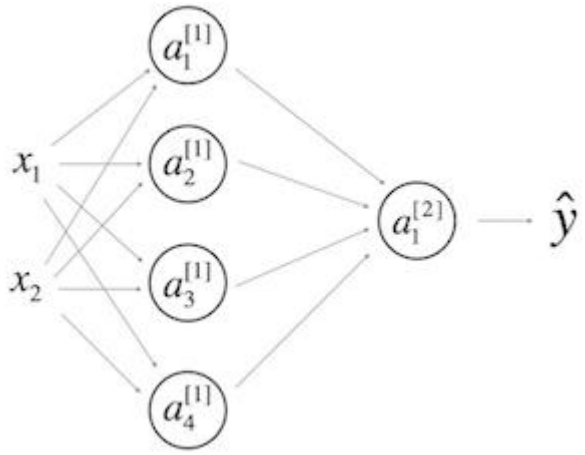
Incorrect
No, Logistic Regression doesn't have a hidden layer. If you initialize the weights to zeros, the first example x fed in the logistic regression will output zero but the derivatives of the Logistic Regression depend on the input x (because there's no hidden layer) which is not zero. So at the second iteration, the weights' values follow x's distribution and are different from each other if x is not a constant vector.

8.
Question 8
Which of the following are true about the tanh function?

0 / 1 point
Expand

Incorrect
You didn't select all the correct answers

9.
Question 9
Consider the following 1 hidden layer neural network:



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

0 / 1 point

Expand

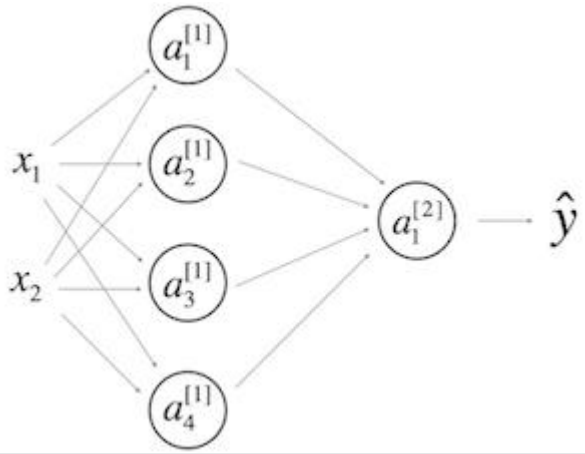
Incorrect

You didn't select all the correct answers

10.

Question 10

What are the dimensions of Z^{[1]}$Z^{[1]}$ and A^{[1]}$A^{[1]}$?



0 / 1 point

Expand

Incorrect

1.

Question 1

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

0 / 1 point

Expand

Incorrect

You didn't select all the correct answers

2.

Question 2

The sigmoid function is only mentioned as an activation function for historical reasons. The tanh is always preferred without exceptions in all the layers of a Neural Network. True/False?

1 / 1 point

Expand

Correct

Yes. Although the tanh almost always works better than the sigmoid function when used in hidden layers, thus is always proffered as activation function, the exception is for the output layer in classification problems.

3.

Question 3

Which of these is a correct vectorized implementation of forward propagation for layer ll, where $1 \leq l \leq L1 \leq l \leq L$?

1 / 1 point

Expand

Correct

4.

Question 4

The use of the ReLU activation function is becoming more rare because the ReLU function has no derivative for $c=0c=0$. True/False?

1 / 1 point

Expand

Correct

Yes. Although the ReLU function has no derivative at $c=0c=0$ this rarely causes any problems in practice. Moreover it has become the default activation function in many cases, as explained in the lectures.

5.

Question 5

Consider the following code:

#+begin_src python

x = np.random.rand(4, 5)

y = np.sum(x, axis=1)

#+end_src

What will be y.shape?

0 / 1 point

Expand

Incorrect

No. By using axis=1 the sum is computed over each row of the array, thus the resulting array is a column vector with 4 entries. Since the option keepdims was not used the array doesn't keep the second dimension.

6.

Question 6

Suppose you have built a neural network with one hidden layer and tanh as activation function for the hidden layer. You decide to initialize the weights to small random numbers and the biases to zero. The first hidden layer's neurons will perform different computations from each other even in the first iteration. True/False?

1 / 1 point

Expand

Correct

7.

Question 7

A single output and single layer neural network that uses the sigmoid function as activation is equivalent to the logistic regression. True/False

1 / 1 point

Expand

Correct

Yes. The logistic regression model can be expressed by $\hat{y} = \sigma\left(W \, x + b \right)$$y^ =\sigma(Wx+b)$. This is the same as $a^{[1]} = \sigma(W^{[1]} \, X + b)$$a^{[1]}=\sigma(W^{[1]}X+b)$.

8.

Question 8

Which of the following is true about the ReLU activation functions?

1 / 1 point

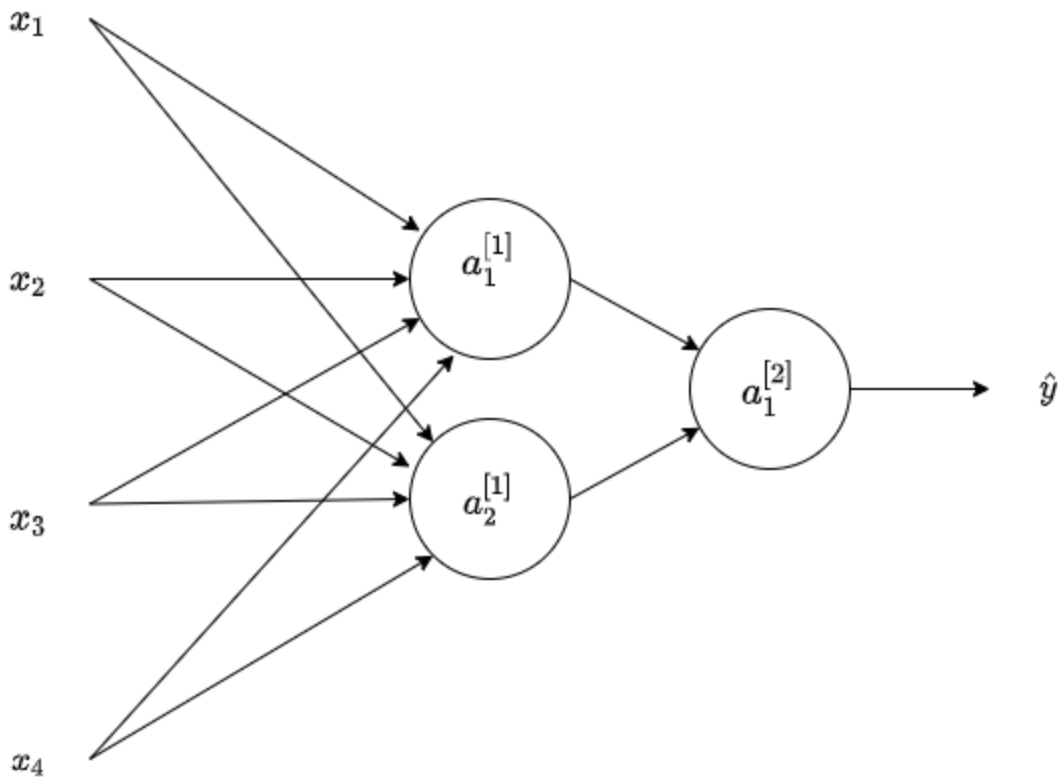
Expand

Correct

9.

Question 9

Consider the following 1 hidden layer neural network:



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

1 / 1 point

Expand

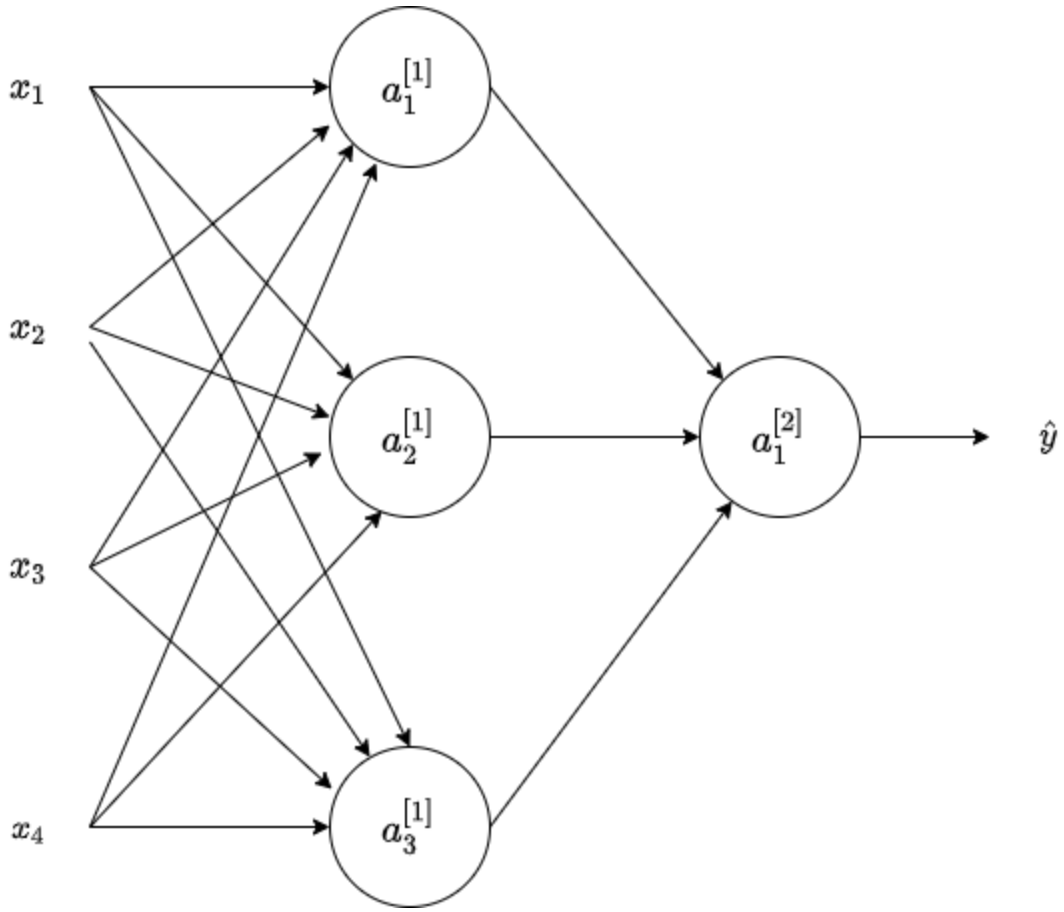
Correct

Great, you got all the right answers.

10.

Question 10

Consider the following 1 hidden layer neural network:



What are the dimensions of Z^{[1]}$Z_{[1]}$ and A^{[1]}$A_{[1]}$?

1 / 1 point

Expand

Correct

Yes. The Z^{[1]}$Z_{[1]}$ and A^{[1]}$A_{[1]}$ are calculated over a batch of training examples. The number of columns in Z^{[1]}$Z_{[1]}$ and A^{[1]}$A_{[1]}$ is equal to the number of examples in the batch, m. And the number of rows in Z^{[1]}$Z_{[1]}$ and A^{[1]}$A_{[1]}$ is equal to the number of neurons in the first layer.