축하합니다! 통과하셨습니다!

받은 학점 80% 최신 제출물 학점 80% 통과 점수: 80% 이상

다음 항목으로 이동

1. Which of the following are true? (Check all that apply.) 0 / 1점

- [x] $w_3^{[4]}$ is the column vector of parameters of the fourth layer and third neuron.
  - ✓ **Correct**
    Yes. The vector $w_j^{[i]}$ is the column vector of parameters of the i-th layer and j-th neuron of that layer.
- [ ] $w_3^{[4]}$ is the column vector of parameters of the third layer and fourth neuron.
- [x] W_1 is a matrix with rows equal to the parameter vectors of the first layer.
  - ! **This should not be selected**
    No. The notation convention is that the superscript number in brackets indicates the number of layers.
- [ ] $w_3^{[4]}$ is the row vector of parameters of the fourth layer and third neuron.
- [ ] $W^{[1]}$ is a matrix with rows equal to the parameter vectors of the first layer.
- [x] $W^{[1]}$ is a matrix with rows equal to the transpose of the parameter vectors of the first layer.
  - ✓ **Correct**
    Yes. We construct $W^{[1]}$ stacking the parameter vectors $w_j^{[1]}$ of all the neurons of the first layer.

더 보기

⊗ 틀립니다
You chose the extra incorrect answers.

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 / 1점

- ○ True
- ◉ False

더 보기

✓ 맞습니다
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. Which of these is a correct vectorized implementation of forward propagation for layer l, where $1 \leq l \leq L$? 1 / 1점

- ○ $Z^{[l]} = W^{[l]}A^{[l]} + b^{[l]}$
  $A^{[l+1]} = g^{[l]}(Z^{[l]})$
- ◉ $Z^{[l]} = W^{[l]}A^{[l-1]} + b^{[l]}$
  $Z^{[l]} = W^{[l]}A^{[l-1]} + b^{[l]}$
  $A^{[l]} = g^{[l]}(Z^{[l]})$
  $A^{[l]} = g^{[l]}(Z^{[l]})$
- ○ $$Z^{[l]} = W^{[l-1]} A^{[l]}+ b^{[l-1]}$$
  $$A^{[l]} = g^{[l]}(Z^{[l]})$$

더 보기

✓ 맞습니다

4. The use of the ReLU activation function is becoming more rare because the ReLU function has no derivative for $c = 0$. True/False? 1 / 1점

- ◉ False
- ○ True

더 보기

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

5. Consider the following code: 1 / 1점

#+begin_src python

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

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

#+end_src

What will be y.shape?

- (5,)
- (1, 5)
- (4, 1)
- (4,) ✓

더 보기

맞습니다
Yes. 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. 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 / 1점

- True Yes. Since the weights are most likely different, each neuron will do a different computation. ✓
- False No. Since the weights are most likely different, each neuron will do a different computation.

더 보기

맞습니다

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 / 1점

- True ✓
- False

더 보기

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

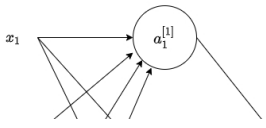
8. Which of the following is true about the ReLU activation functions? 0 / 1점

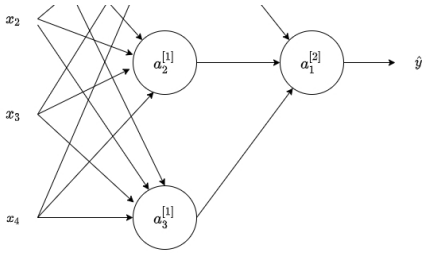
- They are the go to option when you don't know what activation function to choose for hidden layers.
- They are only used in the case of regression problems, such as predicting house prices.
- They are increasingly being replaced by the tanh in most cases.
- They cause several problems in practice because they have no derivative at 0. That is why Leaky ReLU was invented. ✓

더 보기

틀립니다

9. Consider the following 1 hidden layer neural network: 1 / 1점





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

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

  ✓ **Correct**
  Yes. $b^{[k]}$ is a column vector and has the same number of rows as neurons in the k-th layer.

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

  ✓ **Correct**
  Yes. The number of rows in $W^{[k]}$ is the number of neurons in the k-th layer and the number of columns is the number of inputs of the layer.

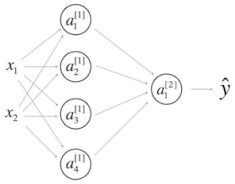
- [ ] $b^{[2]}$ will have shape (3, 1)
- [ ] $$W^{[1]}$$ will have shape (4, 3).
- [x] $$b^{[2]}$$ will have shape (1,1)

  ✓ **Correct**
  Yes. $$b^{[k]}$$ is a column vector and has the same number of rows as neurons in the k-th layer.

더 보기

맞습니다
Great, you got all the right answers.

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



- ( ) $Z^{[1]}$ and $A^{[1]}$ are (1,4)
- ( ) $Z^{[1]}$ and $A^{[1]}$ are (4,2)
- ( ) $Z^{[1]}$ and $A^{[1]}$ are (4,1)
- (•) $Z^{[1]}$ and $A^{[1]}$ are (4,m)

더 보기

맞습니다