

# Practical aspects of deep learning

10/10 points (100%)

Teste, 10 questions

✓ **Parabéns! Você foi aprovado!**

Próximo item



1/1  
pontos

1.

If you have 10,000,000 examples, how would you split the train/dev/test set?



98% train . 1% dev . 1% test



**Correto**



33% train . 33% dev . 33% test



60% train . 20% dev . 20% test



1/1  
pontos

2.

The dev and test set should:



Come from the same distribution



**Correto**





Come from different distributions

## Practical aspects of deep learning

Be identical to each other (same (x,y) pairs)

10/10 points (100%)

Teste, 10 questions



Have the same number of examples



1/1  
pontos

3.

If your Neural Network model seems to have high bias, what of the following would be promising things to try? (Check all that apply.)

☐

Get more test data



Não selecionado está correto

☐

Make the Neural Network deeper



Correto

☐

Increase the number of units in each hidden layer



Correto

☐

Add regularization



Não selecionado está correto

☐

Get more training data



Não selecionado está correto



1/1  
pontos

## Practical aspects of deep learning

Teste, 10 questions

4.

You are working on an automated check-out kiosk for a supermarket, and are building a classifier for apples, bananas and oranges. Suppose your

10/10 points (100%)

classifier obtains a training set error of 0.5%, and a dev set error of 7%.

Which of the following are promising things to try to improve your classifier? (Check all that apply.)

☐

Increase the regularization parameter  $\lambda$



Correto

☐

Decrease the regularization parameter  $\lambda$



Não selecionado está correto

☐

Get more training data



Correto

☐

Use a bigger neural network



Não selecionado está correto



1/1  
pontos

5.

What is weight decay?

☐

Gradual corruption of the weights in the neural network if it is trained on noisy data.

☐

A technique to avoid vanishing gradient by imposing a ceiling on the values of the weights.

☒

A regularization technique (such as L2 regularization) that results in gradient descent shrinking the weights on every iteration.



Correto

# Practical aspects of deep learning

☐ The process of gradually decreasing the learning rate during training

10/10 points (100%)

Teste, 10 questions



1/1  
pontos

6.

What happens when you increase the regularization hyperparameter lambda?



Weights are pushed toward becoming smaller (closer to 0)



**Correto**



Weights are pushed toward becoming bigger (further from 0)



Doubling lambda should roughly result in doubling the weights



Gradient descent taking bigger steps with each iteration (proportional to lambda)



1/1  
pontos

7.

With the inverted dropout technique, at test time:



You apply dropout (randomly eliminating units) and do not keep the  $1/\text{keep\_prob}$  factor in the calculations used in training



You do not apply dropout (do not randomly eliminate units), but keep the  $1/\text{keep\_prob}$  factor in the calculations used in training.



You apply dropout (randomly eliminating units) but keep the  $1/\text{keep\_prob}$  factor in the calculations used in training.



You do not apply dropout (do not randomly eliminate units) and do not keep the  $1/\text{keep\_prob}$  factor in the calculations used in training



**Correto**

## Practical aspects of deep learning

10/10 points (100%)

Teste, 10 questions



1/1  
pontos

8.

Increasing the parameter `keep_prob` from (say) 0.5 to 0.6 will likely cause the following: (Check the two that apply)

☐

Increasing the regularization effect



**Não selecionado está correto**

☐

Reducing the regularization effect



**Correto**

☐

Causing the neural network to end up with a higher training set error



**Não selecionado está correto**

☐

Causing the neural network to end up with a lower training set error



**Correto**



1/1  
pontos

9.

Which of these techniques are useful for reducing variance (reducing overfitting)? (Check all that apply.)

☐

Xavier initialization



**Não selecionado está correto**

# Practical aspects of deep learning

Teste, 10 questions

10/10 points (100%)



Gradient Checking

Não selecionado está correto



L2 regularization

Correto



Dropout

Correto



Vanishing gradient

Não selecionado está correto



Data augmentation

Correto



Exploding gradient

Não selecionado está correto



1/1  
pontos

10.

Why do we normalize the inputs  $x$ ?



Normalization is another word for regularization--It helps to reduce variance



It makes the parameter initialization faster



It makes it easier to visualize the data