

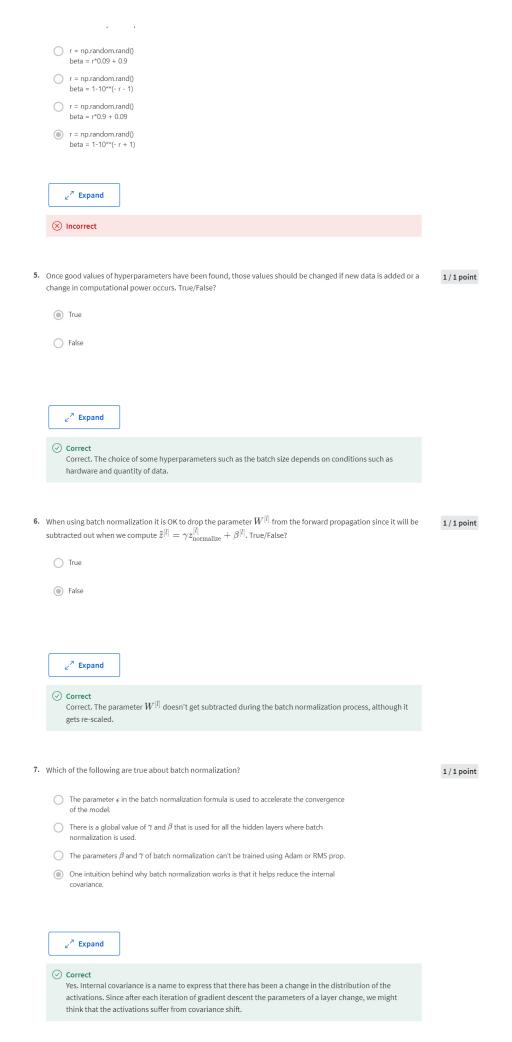
## Congratulations! You passed!

Grade Latest Submission received 80% Grade 80%

To pass 80% or higher

Go to next item

1.	Which of the following are true about hyperparameter search?	1/1 point
	Choosing random values for the hyperparameters is convenient since we might not know in advance which hyperparameters are more important for the problem at hand.	
	<ul> <li>Choosing values in a grid for the hyperparameters is better when the number of hyperparameters to tune is high since it provides a more ordered way to search.</li> </ul>	
	When sampling from a grid, the number of values for each hyperparameter is larger than when using random values.	
	When using random values for the hyperparameters they must be always uniformly distributed.	
	∠ <sup>n</sup> Expand	
	○ Correct     Correct. Different problems might be more sensitive to different hyperparameters.	
2.	If it is only possible to tune two parameters from the following due to limited computational resources. Which two would you choose?	1 / 1 point
	α	
	Correct Correct. This might be the hyperparameter that most impacts the results of a model.	
	$\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ $	
	igspace The $eta$ parameter of the momentum in gradient descent.	
	<ul> <li>Correct</li> <li>Correct. This hyperparameter can increase the speed of convergence of the training, thus is worth tuning.</li> </ul>	
	$\beta_1, \beta_2$ in Adam.	
	∠ <sup>∞</sup> Expand	
	⊙ Correct     Great, you got all the right answers.	
3.	Using the "Panda" strategy, it is possible to create several models. True/False?	0 / 1 point
	False	
	○ True	
	∠ <sup>n</sup> Expand	
	Name of the "Panda" analogy, it is possible to babysit a model until a certain point and then start again to produce a different one.	



0	14/la: -la -f	4la a fa 11 a			la a 4 I			+:	
٥.	Which of	the lollo	wing are	true a	Dout	oaten i	normai	ization:	

1/1 point

$\gamma^{[l]}$ and $\gamma^{[l]}$ are hyperparameters that must be tuned by random sampling in a logarithmic
cale.

The parameters  $\gamma^{[l]}$  and  $\beta^{[l]}$  set the variance and mean of  $\tilde{z}^{[l]}$ .

### ✓ Correct

Correct. When applying the linear transformation  $\tilde{z}^{(l)} = \beta^{[l]} z_{norm}^{(l)} + \gamma^{[l]}$  we set the variance and mean of  $\tilde{z}^{[l]}$ .

When using batch normalization we introduce two new parameters  $\gamma^{[l]}, \beta^{[l]}$  that must be "learned" or trained.

#### ✓ Correct

Correct. Batch normalization uses two parameters  $\beta$  and  $\gamma$  to compute  $\tilde{z}^{(i)}=\beta z_{norm}^{(i)}+\gamma$ .

 $z_{norm}^{(i)} = \frac{z^{(i)} - \mu}{\sqrt{\sigma^2}}$ 

## ∠ Expand

**⊘** Correct

Great, you got all the right answers.

9. A neural network is trained with Batch Norm. At test time, to evaluate the neural network we turn off the Batch Norm to avoid random predictions from the network. True/False?

1/1 point

○ True

### ∠<sup>7</sup> Expand

Correct. During the test, the parameters  $\mu$  and  $\sigma^2$  are estimated using an exponentially weighted average across mini-batches used during training.

10. If a project is open-source, it is a guarantee that it will remain open source in the long run and will never be modified to benefit only one company. True/False?

1/1 point

# O True

False

### ∠ Expand

✓ Correct

Correct. To ensure that a project will remain open source in the long run it must have a good governance body too.