

## Congratulations! You passed!

recommended way to sample a value for lpha?

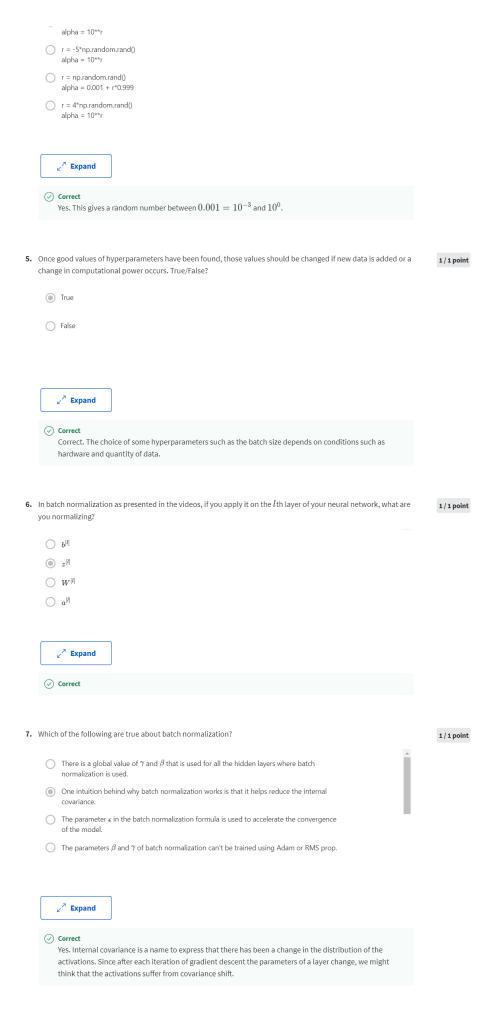
r = -3\*np.random.rand()

Grade Latest Submission received 100% Grade 100%

To pass 80% or higher

Go to next item

1.	If searching among a large number of hyperparameters, you should try values in a grid rather than random values, so that you can carry out the search more systematically and not rely on chance. True or False?	1/1 point
	○ True	
	False	
	∠ <sup>2</sup> Expand	
	<b>⊘</b> Correct	
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.	
	$\beta_1, \beta_2$ in Adam.	
	\$\$\epsilon\$\$ in Adam.	
	✓ The \$\$\beta\$\$ parameter of the momentum in gradient descent.	
	<ul> <li>Correct         Correct. This hyperparameter can increase the speed of convergence of the training, thus is worth tuning.     </li> </ul>	
	∠ <sup>n</sup> Expand	
3.	Using the "Panda" strategy, it is possible to create several models. True/False?	1/1 point
	True	
	○ False	
	∠ <sup>™</sup> Expand	
	Correct Correct. Following the "Panda" analogy, it is possible to babysit a model until a certain point and then start again to produce a different one.	
4.	Knowing that the hyperparameter $lpha$ should be in the range of $0.001$ and $1.0$ . Which of the following is the	1/1 point



$\bigcirc$ The optimal values to use for $\gamma$ and $\beta$ are $\gamma=\sqrt{\sigma^2+\epsilon}$ and $\beta=\mu$ .	
The parameters	
$\gamma^{[I]}$	
$^{\omega_{norm}}-\sqrt{\sigma^2}$ .	_
$igcup$ The parameters $\gamma^{[l]}$ and $eta^{[l]}$ can be learned only using plain gradient descent.	
$ u^{\gamma} $ Expand $ u^{\gamma} $ Correct Correct. When applying the linear transformation $ u^{(l)} = \beta^{[l]} z_{norm}^{(l)} + \gamma^{[l]} $ we set the variance and mean of $ u^{[l]} $ .	1
After training a neural network with Batch Norm, at test time, to evaluate the neural network on a new example you should:	1/1 point
<ul> <li>Perform the needed normalizations, use \( \mu \) and \( \sigma^2 \) estimated using an exponentially weighted average across mini-batches seen during training.</li> </ul>	
If you implemented Batch Norm on mini-batches of (say) 256 examples, then to evaluate on one test example, duplicate that example 256 times so that you're working with a pactifulate salter as 'rouring' trianing:	
Ouse the most recent mini-batch's value of $\mu$ and $\sigma^2$ to perform the needed normalizations.	
Skip the step where you normalize using $\mu$ and $\sigma^2$ since a single test example cannot be normalized.	
∠ <sup>n</sup> Expand	
○ Correct	
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?  True  False	1/1 point
∠ <sup>?</sup> Expand	
Correct Correct. To ensure that a project will remain open source in the long run it must have a good governance body too.	