## Hyperparameter tuning, Batch Normalization, Programming Frameworks

9/10 points (90.00%)

Quiz, 10 questions

✓ Congratulations! You passed!	Next Item
1/1 points	
1. If searching among a large number of hyperparameters, you should try values in a ខ្ values, so that you can carry out the search more systematically and not rely on cha	
True False	
Correct	
1/1 points	
<ol><li>Every hyperparameter, if set poorly, can have a huge negative impact on training, ar about equally important to tune well. True or False?</li></ol>	nd so all hyperparameters are
True	
False	
<b>Correct</b> Yes. We've seen in lecture that some hyperparameters, such as the learning rate, others.	are more critical than
1/1 points	
3.  During hyperparameter search, whether you try to babysit one model ("Panda" strategy.	tegy) or train a lot of models in

parallel ("Caviar") is largely determined by:

Whether you use batch or mini-batch optimization

10 que <b>Corre</b>	Darameter tuning, Batch Normalization, Programming  Work®punt of computational power you can access  stions ct	9/10 points (90.00%)
	The number of hyperparameters you have to tune	
	1/1 points	
	nink $eta$ (hyperparameter for momentum) is between on 0.9 and 0.99, which of the fol nended way to sample a value for beta?	lowing is the
	1 r = np.random.rand() 2 beta = r*0.09 + 0.9	
	1 r = np.random.rand() 2 beta = 1-10**(- r - 1)	
Corre	t 1 r = np.random.rand()	
	2 beta = 1-10**(- r + 1)	
	1 r = np.random.rand() 2 beta = r*0.9 + 0.09	

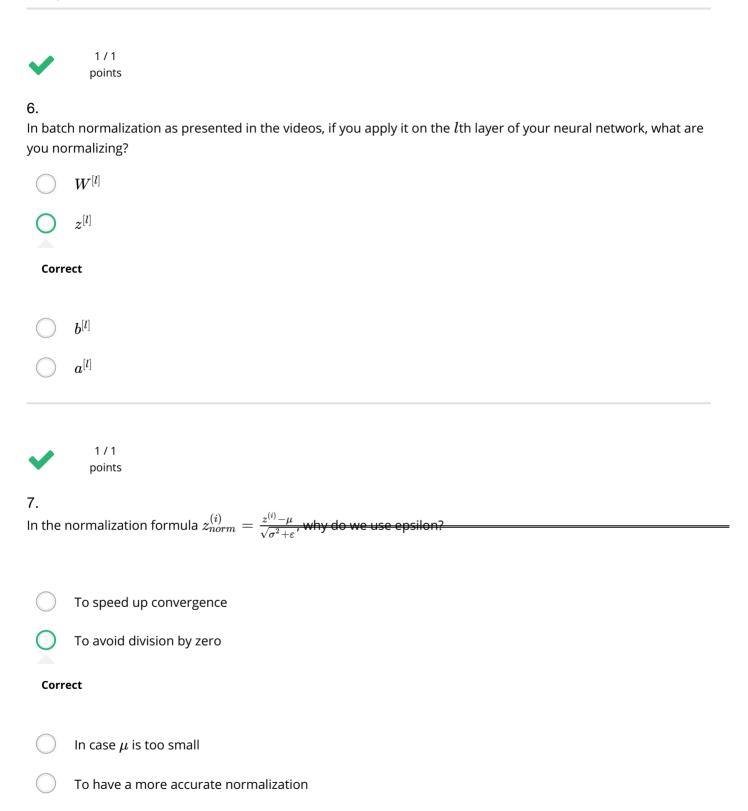
True



## Hyperparameter tuning, Batch Normalization, Programming Frameworks

9/10 points (90.00%)

Quiz, 10 questions



X

0/1 points

8.

Which of the following statements about  $\gamma$  and  $\beta$  in Batch Norm are true?

## Hyperparameteratuning Batich Mormalization Programming not just 9/10 products Frameworks.

Ouiz, 10 questions Correct There is one global value of  $\gamma \in \Re$  and one global value of  $\beta \in \Re$  for each layer, and applies to all the hidden units in that layer. This should not be selected  $\beta$  and  $\gamma$  are hyperparameters of the algorithm, which we tune via random sampling. **Un-selected** is correct The optimal values are  $\gamma = \sqrt{\sigma^2 + \varepsilon}$ , and  $\beta = \mu$ . **Un-selected** is correct They set the mean and variance of the linear variable  $z^{[l]}$  of a given layer. Correct 1/1 points After training a neural network with Batch Norm, at test time, to evaluate the neural network on a new example you should: 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 mini-batch the same size as during training. Perform the needed normalizations, use  $\mu$  and  $\sigma^2$  estimated using an exponentially weighted average across mini-batches seen during training. Correct Use 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.

## Hyperparameter tuning, Batch Normalization, Programming Frameworks

9/10 points (90.00%)

Qι

	points ,
uiz, 10 que <b>10.</b>	estions
Which c	of these statements about deep learning programming frameworks are true? (Check all that apply)
	A programming framework allows you to code up deep learning algorithms with typically fewer lines of code than a lower-level language such as Python.
Corre	ct
	Deep learning programming frameworks require cloud-based machines to run.
Un-se	elected is correct
	Even if a project is currently open source, good governance of the project helps ensure that the it remains open even in the long term, rather than become closed or modified to benefit only one company.
Corre	ct
	□

