## Hyperparameter tuning, Batch Normalization, Programming Frameworks

10/10 points (100%)

Quiz, 10 questions

ngra	itulations! You passed!	Next Ite
<b>~</b>	1/1 points	
rather	thing among a large number of hyperparameters, you should try than random values, so that you can carry out the search more s t rely on chance. True or False?	
	True	
0	False	
Corr	ect	
	ryperparameter, if set poorly, can have a huge negative impact o yperparameters are about equally important to tune well. True of True	
0	False	
	ect We've seen in lecture that some hyperparameters, such as the le are more critical than others.	earning
<b>~</b>	1 / 1 points	
_	hyperparameter search, whether you try to babysit one model ( n a lot of models in parallel ("Caviar") is largely determined by:	"Panda" strategy)
	Whether you use batch or mini-batch optimization	
	The presence of local minima (and saddle points) in your neura	

# The amount of computational power you can access Hyperparameter tuning, Batch Normalization, Programming Frameworks<sub>Correct</sub>

10/10 points (100%)

Quiz, 10 questions

	The number of hyperparameters you have to tune		
<b>~</b>	1 / 1 points		
hyper	think $\beta$ parameter for momentum) is between on 0.9 and 0.99, which of the following is commended way to sample a value for beta?		
	1 r = np.random.rand() 2 beta = r*0.09 + 0.9		
0	1 r = np.random.rand() 2 beta = 1-10**(- r - 1)		
Corr	ect		
	1 r = np.random.rand() 2 beta = 1-10**(- r + 1)		
	1 r = np.random.rand() 2 beta = r*0.9 + 0.09		
<b>~</b>	1 / 1 points		
t once	g good hyperparameter values is very time-consuming. So typically you should do at the start of the project, and try to find very good hyperparameters so that you ver have to revisit tuning them again. True or false?		
	True		
0	False		
Correct			

### Hyperparameter tuning, Batch Normalization, Programming Frameworks

10/10 points (100%)

Quiz, 10 questions In batch normalization as presented in the videos, if you apply it on the l th layer of your neural network, what are you normalizing?

- $W^{[l]}$
- $a^{[l]}$

Correct

( ) b[!]



1/1

points

7.

In the normalization formula  $z_{norm}^{(i)} = \frac{z^{(i)} - \mu}{\sqrt{\sigma^2 + \varepsilon}}$ 

, why do we use epsilon?

- To speed up convergence
- To avoid division by zero

Correct

- To have a more accurate normalization
- In case  $\mu$  is too small



1/1 points

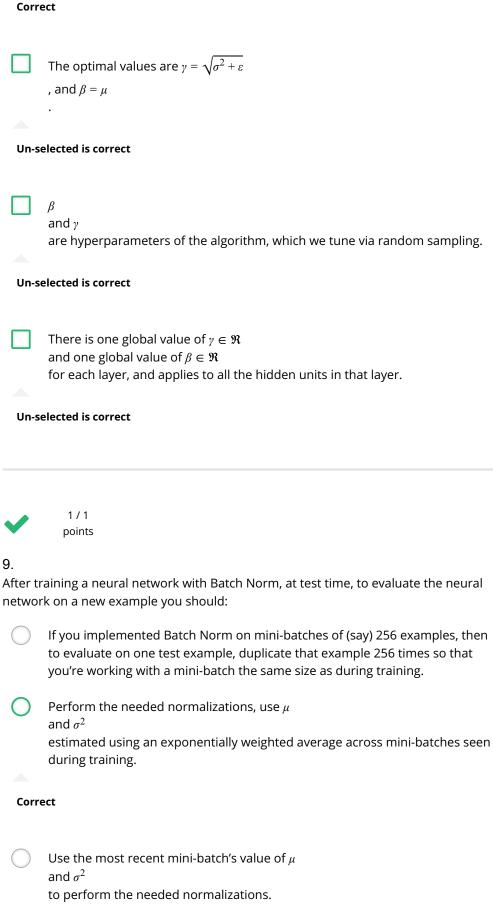
8

Which of the following statements about  $\gamma$  and  $\beta$ 

in Batch Norm are true?

- They set the mean and variance of the linear variable  $z^{[l]}$  of a given layer.
  - Correct

Quiz, 10 questions



Skip the step where you normalize using  $\mu$ 

#### [Math Processing Error] and $\sigma^2$

since a single test example cannot be normalized.

#### Hyperparameter tuning, Batch Normalization, Programming Frameworks

10/10 points (100%)

Quiz, 10 questions

