Hyperparameter tuning, Batch Normalization, Programming Frameworks

10/10 points (100%)

Quiz, 10 questions

Congratulations! You passed!	Next Item
1/1 points	
1. If searching among a large number of hyperparameters, you should try valu than random values, so that you can carry out the search more systematical chance. True or False?	
True	
False	
Correct	
 1/1 points 2. Every hyperparameter, if set poorly, can have a huge negative impact on tra 	ining, and so all
hyperparameters are about equally important to tune well. True or False?	
True	
○ False	
Correct Yes. We've seen in lecture that some hyperparameters, such as the learning more critical than others.	ng rate, are
1/1 points	
3.	

train a lot of models in parallel ("Caviar") is largely determined by:

During hyperparameter search, whether you try to babysit one model ("Panda" strategy) or

Whether you use batch or mini-batch optimization

	Whether you use succin or mini succin optimization	
Hyperparai Framework		10/10 points (100%)
Quiz, 10 questions	The amount of computational power you can access	
Co	prrect	
	The number of hyperparameters you have to tune	
~	1 / 1 points	
4.		
If yo	u think eta (hyperparameter for momentum) is between on 0.9 and 0.99, which of the wing is the recommended way to sample a value for beta?	
	1 r = np.random.rand() 2 beta = r*0.09 + 0.9	
C	1 r = np.random.rand() 2 beta = 1-10**(- r - 1)	
Co	prrect	
	1 r = np.random.rand() 2 beta = 1-10**(- r + 1)	
	1 r = np.random.rand() 2 beta = r*0.9 + 0.09	
✓	1 / 1 points	
at th	ing good hyperparameter values is very time-consuming. So typically you should do se start of the project, and try to find very good hyperparameters so that you don't ex to revisit tuning them again. True or false?	
) True	

False

Hyperparameter tuning, Batch Normalization, Programming Frameworks

10/10 points (100%)

Quiz, 10 questions



1/1 points

6.

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



 $z^{|\cdot|}$

Correct

_		
)	$W^{[l]}$
		VV 1-1

$$\bigcirc b^{[l]}$$



1/1 points

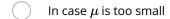
7.

In the normalization formula $z_{norm}^{(i)}=rac{z^{(i)}-\mu}{\sqrt{\sigma^2+arepsilon}}$, why do we use epsilon?



To avoid division by zero

Correct



To speed up convergence

To have a more accurate normalization



points

8.

Which of the following statements about γ and β in Batch Norm are true?

517		Oddiscra Offine Godiscs From Top Officers tides. John for Free	
Hyperna	L rame	There is one global value of $\gamma\in\mathfrak{R}$ and one global value of $\beta\in\mathfrak{R}$ for each layer, applies to all the hidden units in that layer. eter tuning, Batch Normalization, Programming	
Framewo	orks.	elected is correct	10/10 points (100%)
Quiz, 10 questi			
		They can be learned using Adam, Gradient descent with momentum, or RMSpropnot just with gradient descent.),
	Corre	ect	
		The optimal values are $\gamma = \sqrt{\sigma^2 + arepsilon}$, and $\beta = \mu.$	
	Un-se	elected is correct	
		They set the mean and variance of the linear variable $z^{\left[l ight]}$ of a given layer.	
	Corre	ect	
		eta and γ are hyperparameters of the algorithm, which we tune via random sampl	ing.
	Un-se	elected is correct	
	0.1.5		
	~	1 / 1 points	
	9.		
		raining a neural network with Batch Norm, at test time, to evaluate the neural netwew example you should:	<i>r</i> ork
		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.	
		Skip the step where you normalize using μ and σ^2 since a single test example can be normalized.	nnot
		Use the most recent mini-batch's value of μ and σ^2 to perform the needed normalizations.	
	0	Perform the needed normalizations, use μ and σ^2 estimated using an exponential weighted average across mini-batches seen during training.	ally
	<u> </u>		
	Corre	ect	

Hyperparameter tuning, Batch Normalization, Programming Frameworks

10/10 points (100%)

Quiz, 10 questi Which of these statements about deep learning programming frameworks are true? (Check all that apply)

	F-97
	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	ert .
COII	
	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.
Corr	ect
Un-s	Deep learning programming frameworks require cloud-based machines to run.





