r = np.random.rand()
alpha = 0.00001 + r*0.99999

r = np.random.rand()

Graded Quiz • 50 min

Congratulations! You passed!

Grade received 80% Latest Submission Grade 80% To pass 80% or higher

Go to next item

1.	With a relatively small set of hyperparameters, it is OK to use a grid search. True/False?	0 / 1 point
	False	
	○ True	
	∠ [™] Expand	
	$igotimes$ Incorrect. When the set of hyperparameters is small like a range for $n_l=1,2,3$ grid search works fine.	
2.	In a project with limited computational resources, which three of the following hyperparameters would you choose to tune? Check all that apply.	1 / 1 point
	igspace The eta parameter of the momentum in gradient descent.	
	✓ Correct Correct. This hyperparameter can increase the speed of convergence of the training, thus is worth tuning.	
	⋈ α	
	Correct Correct. This might be the hyperparameter that most impacts the results of a model.	
	_ ε in Adam.	
	β_1, β_2 in Adam.	
	mini-batch size	
	Correct Correct. This can have a great impact on the results of the cost function, thus it is worth tuning it.	
	∠ ⁿ Expand	
	Great, you got all the right answers.	
3.	Even if enough computational power is available for hyperparameter tuning, it is always better to babysit one model ("Panda" strategy), since this will result in a more custom model. True/False?	1 / 1 point
	False	
	○ True	
	∠ ⁷ Expand	
	Correct Correct. Although it is possible to create good models using the "Panda" strategy, obtaining better results is more likely using a "caviar" strategy due to the number of tests and the nature of the deep learning process of ideas, code, and experiment.	
4.	Knowing that the hyperparameter α should be in the range of 0.00001 and 1.0 , which of the following is the	1 / 1 point

r = -5*np.random.r alpha = 10**r		
	and()	
r = -4*np.random.r	rand()	
alpha = 10**r		
∠ ⁷ Expand		
Z Expand		
⊘ Correct	50	
Yes. This will gener	rate a random value between 10^{-5} and 10^{0} chosen randomly in a logarithmic scale.	
Finding good byperpara	meter values is very time-consuming. So typically you should do it once at the start of the	
	ery good hyperparameters so that you don't ever have to tune them again. True or false?	1/1 point
0		
False		
○ True		
∠ ⁷ Expand		
⊘ Correct		
in batch normalization a: you normalizing?	s presented in the videos, if you apply it on the l th layer of your neural network, what are	1/1 point
○ <i>p</i> ^[t]		
z ^[I]		
$\bigcirc a^{[l]}$		
$\bigcirc W^{[l]}$		
7		
∠ ⁷ Expand		
⊘ Correct		
When using normalizatio	on:	1 / 1 point
When using normalizatio $z_{norm}^{(i)}=rac{z^{(i)}-\mu}{\sqrt{\sigma^2+arepsilon}}$	in:	1 / 1 point
$z_{norm}^{(i)} = rac{z^{(i)} - \mu}{\sqrt{\sigma^2 + arepsilon}}$	on: e normalization of $z^{(i)}$ may fail since division by 0 may be produced due to rounding	1/1 point
$z_{norm}^{(i)} = rac{z^{(i)} - \mu}{\sqrt{\sigma^2 + arepsilon}}$		1 / 1 point
$z_{norm}^{(i)}=rac{z^{(i)}-\mu}{\sqrt{\sigma^2+arepsilon}}$ In case σ is too small, the		1/1 point
$z_{norm}^{(i)}=rac{z^{(i)}-\mu}{\sqrt{\sigma^2+\epsilon}}$ In case σ is too small, the errors. True/False?		1 / 1 point
$z_{norm}^{(i)}=rac{z^{(i)}-\mu}{\sqrt{\sigma^2+\epsilon}}$ In case σ is too small, the errors. True/False?		1/1 point
$ z_{norm}^{(i)} = \frac{z^{(i)} - \mu}{\sqrt{\sigma^2 + \varepsilon}} $ In case σ is too small, the errors. True/False? \bigcirc True		1/1 point
$ z_{norm}^{(i)} = \frac{z^{(i)} - \mu}{\sqrt{\sigma^2 + \varepsilon}} $ In case σ is too small, the errors. True/False? \bigcirc True		1/1 point
$ z_{norm}^{(i)} = \frac{z^{(i)} - \mu}{\sqrt{\sigma^2 + \varepsilon}} $ In case σ is too small, the errors. True/False?		1/1 point
$ z_{norm}^{(i)} = \frac{z^{(i)} - \mu}{\sqrt{\sigma^2 + \varepsilon}} $ In case σ is too small, the errors. True/False? \bigcirc True		1/1 point
$z_{norm}^{(i)} = \frac{z^{(i)} - \mu}{\sqrt{\sigma^2 + \varepsilon}}$ In case σ is too small, the errors. True/False? True False		1/1 point
$z_{norm}^{(i)} = \frac{z^{(i)} - \mu}{\sqrt{\sigma^2 + \varepsilon}}$ In case σ is too small, the errors. True/False? True False False Correct Correct. The norm.	e normalization of $z^{(i)}$ may fail since division by 0 may be produced due to rounding a normalization of $z^{(i)}$ may fail since division by 0 may be produced due to rounding the normalization of $z^{(i)}$ and $z^{(i)}$ and $z^{(i)}$ and $z^{(i)}$ and $z^{(i)}$ are $z^{(i)}$ and $z^{(i)}$ and $z^{(i)}$ are $z^{(i)}$ and $z^{(i)}$ are $z^{(i)}$ and $z^{(i)}$ and $z^{(i)}$ are $z^{(i)}$	1/1 point
$z_{norm}^{(i)} = \frac{z^{(i)} - \mu}{\sqrt{\sigma^2 + \varepsilon}}$ In case σ is too small, the errors. True/False? True False False Correct Correct. The norm.	e normalization of $z^{(i)}$ may fail since division by 0 may be produced due to rounding	1/1 point
$z_{norm}^{(i)} = \frac{z^{(i)} - \mu}{\sqrt{\sigma^2 + \varepsilon}}$ In case σ is too small, the errors. True/False? True False False Correct Correct. The norm.	e normalization of $z^{(i)}$ may fail since division by 0 may be produced due to rounding a normalization of $z^{(i)}$ may fail since division by 0 may be produced due to rounding the normalization of $z^{(i)}$ and $z^{(i)}$ and $z^{(i)}$ and $z^{(i)}$ and $z^{(i)}$ are $z^{(i)}$ and $z^{(i)}$ and $z^{(i)}$ are $z^{(i)}$ and $z^{(i)}$ are $z^{(i)}$ and $z^{(i)}$ and $z^{(i)}$ are $z^{(i)}$	1/1 point
$z_{norm}^{(i)} = \frac{z^{(i)} - \mu}{\sqrt{\sigma^2 + \varepsilon}}$ In case σ is too small, the errors. True/False? True False Correct Correct. The norm parameter prevent	e normalization of $z^{(i)}$ may fail since division by 0 may be produced due to rounding a lization formula uses a smoothing parameter ϵ so in $z^{(i)}_{\rm norm}=\frac{z^{(i)}-\mu}{\sqrt{\sigma^2+\epsilon}}$ use of the ϵ is that the denominator be 0.	
$z_{norm}^{(i)} = \frac{z^{(i)} - \mu}{\sqrt{\sigma^2 + \varepsilon}}$ In case σ is too small, the errors. True/False? True False Correct Correct. The norm parameter prevent	e normalization of $z^{(i)}$ may fail since division by 0 may be produced due to rounding a normalization of $z^{(i)}$ may fail since division by 0 may be produced due to rounding the normalization of $z^{(i)}$ and $z^{(i)}$ and $z^{(i)}$ and $z^{(i)}$ and $z^{(i)}$ are $z^{(i)}$ and $z^{(i)}$ and $z^{(i)}$ are $z^{(i)}$ and $z^{(i)}$ are $z^{(i)}$ and $z^{(i)}$ and $z^{(i)}$ are $z^{(i)}$	1/1 point 0/1 point
$z_{norm}^{(i)} = \frac{z^{(i)} - \mu}{\sqrt{\sigma^2 + \varepsilon}}$ In case σ is too small, the errors. True/False? True False False Correct Correct. The norm parameter prevent Which of the following an wind patch	e normalization of $z^{(i)}$ may fail since division by 0 may be produced due to rounding alization formula uses a smoothing parameter ϵ so in $z_{\mathrm{norm}}^{(i)}=\frac{z^{(i)}-\mu}{\sqrt{\sigma^2+\epsilon}}$ use of the ϵ is that the denominator be 0.	
$z_{norm}^{(i)} = \frac{z^{(i)} - \mu}{\sqrt{\sigma^2 + \epsilon}}$ In case σ is too small, the errors. True/False? True False Correct Correct. The norm-parameter prevent	e normalization of $z^{(i)}$ may fail since division by 0 may be produced due to rounding alization formula uses a smoothing parameter ϵ so in $z_{\mathrm{norm}}^{(i)}=\frac{z^{(i)}-\mu}{\sqrt{\sigma^2+\epsilon}}$ use of the ϵ is that the denominator be 0.	
z(i) z(i) z(i)	e normalization of $z^{(i)}$ may fail since division by 0 may be produced due to rounding alization formula uses a smoothing parameter ϵ so in $z_{\rm norm}^{(i)}=\frac{z^{(i)}-\mu}{\sqrt{\sigma^2+\epsilon}}$ use of the ϵ is that the denominator be 0.	
z(i) z(i) µ z(norm = z(i) _ µ √σ²+e In case σ is too small, the errors. True/False? ☐ True ☐ False ✓ Expand ✓ Correct Correct. The norm parameter prevent Which of the following at ✓ When using batch "learned" or trained ✓ Correct Correct. Batch norm parameter.	e normalization of $z^{(i)}$ may fail since division by 0 may be produced due to rounding alization formula uses a smoothing parameter ϵ so in $z^{(i)}_{\rm norm}=\frac{z^{(i)}-\mu}{\sqrt{\sigma^2+\epsilon}}$ use of the ϵ is that the denominator be 0. The true about batch normalization? In a normalization we introduce two new parameters $\gamma^{[l]}, \beta^{[l]}$ that must be d.	
$z_{norm}^{(i)} = \frac{z^{(i)} - \mu}{\sqrt{\sigma^2 + \varepsilon}}$ In case σ is too small, the errors. True/False? True False False Correct. The norm parameter prevent Which of the following an which is the following and the following are the following and the following are the following and the following are the following and the following and the following are th	e normalization of $z^{(i)}$ may fail since division by 0 may be produced due to rounding alization formula uses a smoothing parameter ϵ so in $z_{\rm norm}^{(i)}=\frac{z^{(i)}-\mu}{\sqrt{\sigma^2+\epsilon}}$ use of the ϵ is that the denominator be 0. The true about batch normalization? In the comparameters $\gamma^{[i]}$, $\beta^{[i]}$ that must be d.	
$z_{norm}^{(i)} = \frac{z^{(i)} - \mu}{\sqrt{\sigma^2 + \varepsilon}}$ In case σ is too small, the errors. True/False? True False False Correct. The norm parameter prevent Which of the following an which is the following and the following are the following and the following are the following and the following are the following and the following and the following are th	e normalization of $z^{(i)}$ may fail since division by 0 may be produced due to rounding alization formula uses a smoothing parameter ϵ so in $z^{(i)}_{\rm norm}=\frac{z^{(i)}-\mu}{\sqrt{\sigma^2+\epsilon}}$ use of the ϵ is that the denominator be 0. The true about batch normalization? In a normalization we introduce two new parameters $\gamma^{[l]}, \beta^{[l]}$ that must be d.	
$z_{norm}^{(i)} = \frac{z^{(i)} - \mu}{\sqrt{\sigma^2 + \epsilon}}$ In case σ is too small, the errors. True/False? True False False Correct. Correct. The norm parameter prevent Which of the following an Which of the following an Which of the following an True or trained: Correct. Batch in $z^{(i)} = \beta z_{norm}^{(i)} + \beta z_{norm}^{(i)}$ $\beta^{[i]}$ and $\gamma^{[i]}$ are hypocale.	e normalization of $z^{(i)}$ may fail since division by 0 may be produced due to rounding alization formula uses a smoothing parameter ϵ so in $z_{\mathrm{norm}}^{(i)} = \frac{z^{(i)} - \mu}{\sqrt{\sigma^2 + \epsilon}}$ use of the ϵ is that the denominator be 0. The true about batch normalization? In a normalization we introduce two new parameters $\gamma^{[l]}$, $\beta^{[l]}$ that must be d. Ormalization uses two parameters β and γ to compute γ . Perparameters that must be tuned by random sampling in a logarithmic	
$z_{norm}^{(i)} = \frac{z^{(i)} - \mu}{\sqrt{\sigma^2 + \varepsilon}}$ In case σ is too small, the errors. True/False? True False False Correct Correct. The norm- parameter prevent Which of the following at When using batch "learned" or trained Correct. Batch in $z^{(i)} = \beta z_{norm}^{(i)} + \gamma z_{norm}^{(i)}$ $\beta^{[i]}$ and $\gamma^{[i]}$ are byt scale.	e normalization of $z^{(i)}$ may fail since division by 0 may be produced due to rounding alization formula uses a smoothing parameter ϵ so in $z_{\mathrm{norm}}^{(i)} = \frac{z^{(i)} - \mu}{\sqrt{\sigma^2 + \epsilon}}$ use of the ϵ is that the denominator be 0. The true about batch normalization? In a normalization we introduce two new parameters $\gamma^{[l]}$, $\beta^{[l]}$ that must be d. Ormalization uses two parameters β and γ to compute γ . Perparameters that must be tuned by random sampling in a logarithmic	

	$igspace{}{}$ The parameters $\gamma^{[l]}$ and $eta^{[l]}$ set the variance and mean of $ar{z}^{[l]}$.				
	\checkmark Correct Correct. When applying the linear transformation $\bar{z}^{(l)}=\beta^{[l]}z_{norm}^{(l)}+\gamma^{[l]}$ we set the variance and mean of $\bar{z}^{[l]}$.				
	$egin{align*} z_{ ext{norm}}^{(i)} = rac{z^{(i)} - \mu}{\sqrt{\sigma^2}} \ . \end{aligned}$				
	∠ [™] Expand				
	Noorrect You chose the extra incorrect answers.				
9.	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			
	O Skip the step where you normalize using μ and σ^2 since a single test example cannot be normalized.				
	Ouse the most recent mini-batch's value of μ and σ^2 to perform the needed normalizations.				
	Perform the needed normalizations, use \(\mu \) and \(\sigma^2 \) estimated using an exponentially weighted average across mini-batches seen during training.				
	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.				
	∠ ³ Expand				
	© Correct				
10. Which of these statements about deep learning programming frameworks are true? (Check all that apply)					
	Even if a project is currently open source, good governance of the project helps ensure that it remains open even in the long term, rather than become closed or modified to benefit only one company.				
	✓ Correct				
	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.				
	✓ Correct				
	Deep learning programming frameworks require cloud-based machines to run.				
	∠ [⊼] Expand				