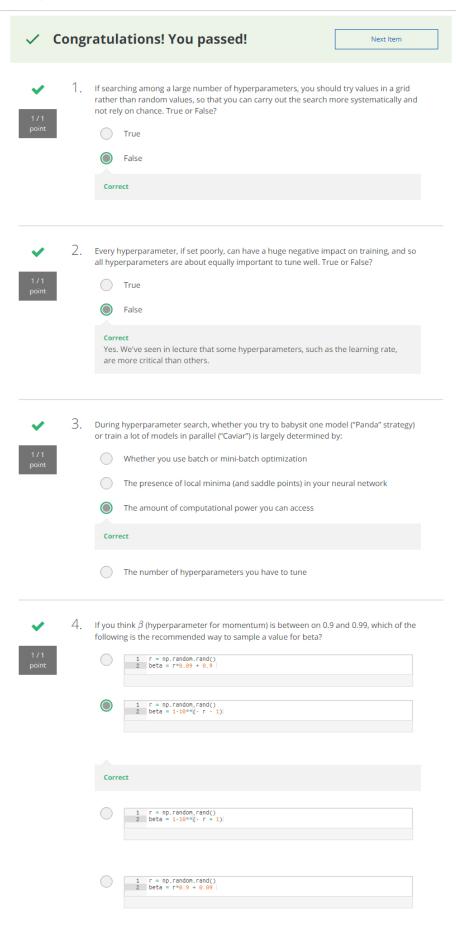
Quiz, 10 questions



1/1 point 1/1 point	 6. 	Finding good hyperparameter values is very time-consuming. So typically you should do it once at the start of the project, and try to find very good hyperparameters so that you don't ever have to revisit tuning them again. True or false? True False Correct 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? $W^{[l]}$
роте		$egin{aligned} a^{[l]} \ b^{[l]} \ \hline & z^{[l]} \ \\ \end{array}$
1/1 point	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 In case μ is too small To avoid division by zero Correct
1/1 point	8.	Which of the following statements about γ and β in Batch Norm are true? β and γ 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 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. Un-selected is correct They can be learned using Adam, Gradient descent with momentum, or RMSprop, not just with gradient descent.

