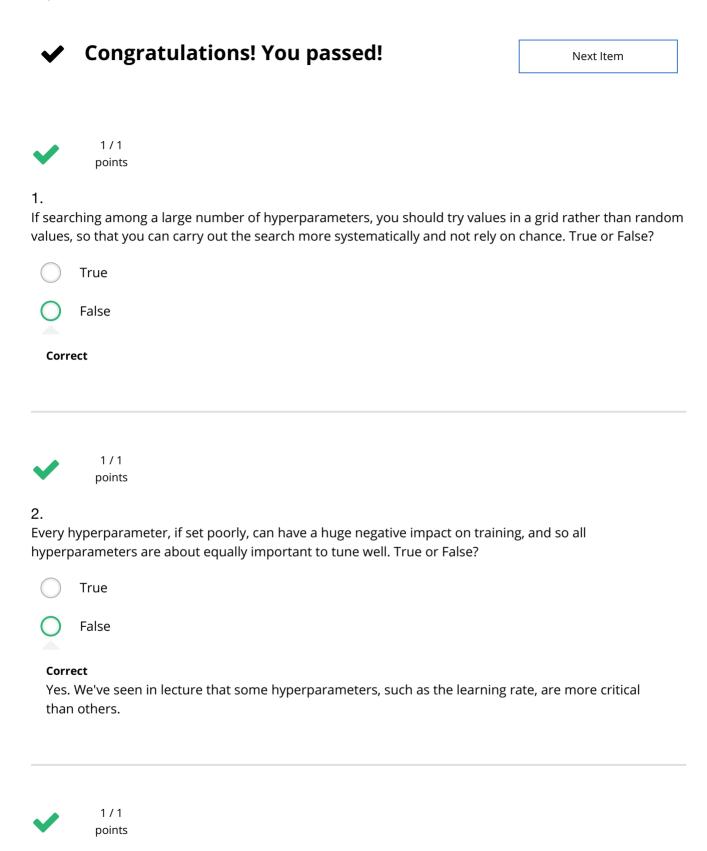
Hyperparameter tuning, Batch Normalization, Programming Frameworks

10/10 points (100.00%)

Quiz, 10 questions



models in parallel ("Caviar") is largely determined by:

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

Hyperparameter tuning, Batch Nonmalization, Programming Frameworks

10/10 points (100.00%)

Quiz, 10 questions

The amount of computational power you can access

The number of hyperparameters you have to tune

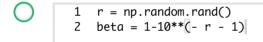


1/1 points

4

If you think β (hyperparameter for momentum) is between on 0.9 and 0.99, which of the following is the recommended way to sample a value for beta?

```
1 r = np.random.rand()
2 beta = r*0.09 + 0.9
```



Correct

```
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

5.

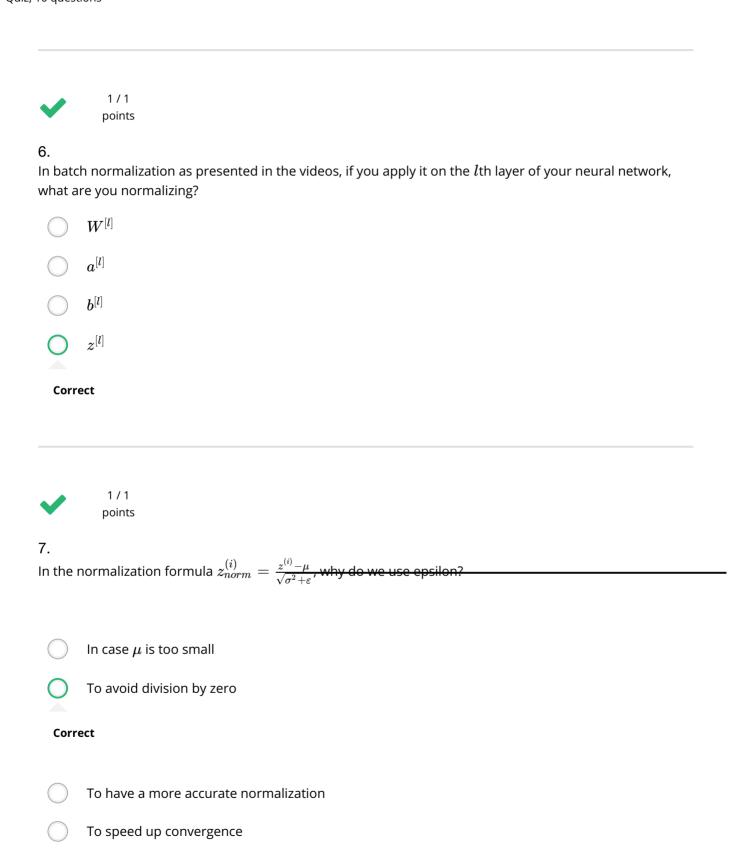
Finding good hyperparameter values is very time-consuming. So typically you should do it once at the start of the project, and try to find very good hyperparameters so that you don't ever have to revisit tuning them again. True or false?

True

Hyperparameter tuning, Batch Normalization, Programming Frameworks

10/10 points (100.00%)

Quiz, 10 questions



1/1 points

8.

Which of the following statements about γ and β in Batch Norm are true? Hyperparameter tuning, Batch Normalization, Programming 10/10 points Frameworks γ are hyperparameters of the algorithm, which we tune via random sampling. (100.00%)

Quiz, 10 questions

Un-sel	ected is	correct

Corr	They can be learned using Adam, Gradient descent with momentum, or RMSprop, not just with gradient descent. ect
Un-s	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.
Corr	They set the mean and variance of the linear variable $z^{[l]}$ of a given layer. $oldsymbol{ ext{ect}}$
	The optimal values are $\gamma=\sqrt{\sigma^2+arepsilon}$, and $eta=\mu$.
Un-s	elected is correct
~	1 / 1 points
9.	
	raining a neural network with Batch Norm, at test time, to evaluate the neural network on a new ble you should:
	sle you should: Skip the step where you normalize using μ and σ^2 since a single test example cannot be
	sle you should: Skip the step where you normalize using μ and σ^2 since a single test example cannot be normalized.

Correct

Hyperparameter tuning, Batch Normalization, Programming Frameworks

10/10 points (100.00%)

Quiz, 10 questions

