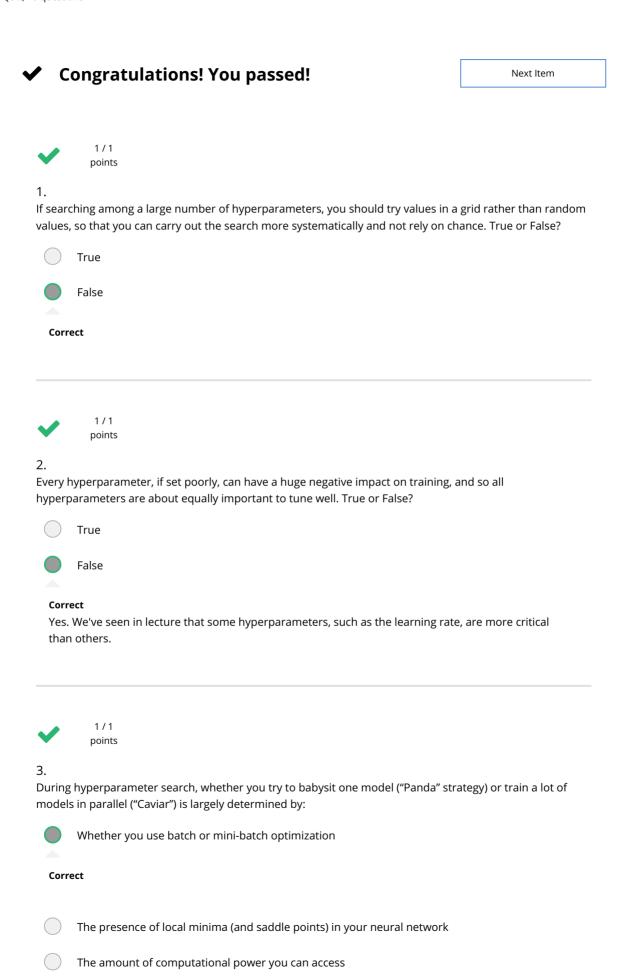


Hyperparameter tuning, Batch Normalization, Programming Frameworks

Quiz, 10 questions

10/10 points (100%)



The number of hyperparameters you have to tune. Hyperparameter tuning, Batch Normalization, Programming Frameworks

10/10 points (100%)

Quiz, 10 questions



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?





Correct



1/1 points

6.

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



 $z^[l]$

Hyperparameter tuning, Batch Normalization, Programming Frameworks

10/10 points (100%)

Quiz, 10 questions $W^{[}l^{]}$

\bigcirc a

$$b^{[l]}$$



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 have a more accurate normalization
- To avoid division by zero

Correct

- In case μ is too small
- To speed up convergence



1/1 points

8.

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



Correct

eta and γ are hyperparameters of the algorithm, which we tune via random sampling.

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

Hyperparameter । अवस्था अवस्था

Quiz, 10 quenticelected is correct

10/10 points (100%)

~	1/1 points
	aining a neural network with Batch Norm, at test time, to evaluate the neural network on a new e you should:
	Use the most recent mini-batch's value of μ and σ^2 to perform the needed normalizations.
Corre	ect
	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.
Un-se	elected is correct
	Perform the needed normalizations, use μ and σ^2 estimated using an exponentially weighted average across mini-batches seen during training.
Un-se	elected is correct
	Skip the step where you normalize using μ and σ^2 since a single test example cannot be normalized.
Un-se	elected is correct
~	1 / 1 points
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 the it remains open even in the long term, rather than become closed or modified to benefit only one company.
Corre	ect
	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	ect