Quiz, 10 questions

# **Congratulations! You passed!** Next Item If searching among a large number of hyperparameters, you should try values in a grid rather than random values, so that you can carry out the search more systematically and not rely on chance. True or False? 1/1 points True False Correct Every hyperparameter, if set poorly, can have a huge negative impact on training, and so all hyperparameters are about equally important to tune well. True or False? 1/1 True points False Correct Yes. We've seen in lecture that some hyperparameters, such as the learning rate, are more critical than others. During hyperparameter search, whether you try to babysit one model ("Panda" strategy) or train a lot of models in parallel ("Caviar") is largely determined by: 0/1 Whether you use batch or mini-batch optimization points

Quiz, 10 questions



The amount of computational power you can access

8/10 points (80%)

# This should not be selected

The number of hyperparameters you have to tune



If you think  $\beta$  (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?

1/1 points

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



```
1 r = np.random.rand()
2 beta = 1-10**(-r-1)
```

#### Correct





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



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

1/1 points









Correct





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

1/1 points

To have a more accurate normalization



To avoid division by zero

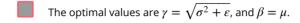
Correct

- In case  $\mu$  is too small
- To speed up convergence

Frameworks
Qui2,10 questions
points

They can be learned using Adam, Gradient descent with momentum, or RM \$80%, not just with gradient descent.

Correct



This should not be selected

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** 

eta and  $\gamma$  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  $\mathbf{z}^{[l]}$  of a given layer.

Correct



**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 points

Use the most recent mini-batch's value of  $\mu$  and  $\sigma^2$  to perform the needed normalizations.

Correct

Hyperparameter tuning, Batch Norm on mini-batches of (say) 256 examples, then to tuning, Batch Normal ization, Programming so that you's 10 points working with a mini-batch the same size as during training. (80%)

# Un-selected is correct

Perform the needed normalizations, use  $\mu$  and  $\sigma^2$  estimated using an exponentially weighted average across mini-batches seen during training.

### **Un-selected is correct**

Skip the step where you normalize using  $\mu$  and  $\sigma^2$  since a single test example cannot be normalized.

# **Un-selected is correct**



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

1/1 points

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.

# 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.

# **Un-selected is correct**

# Hyperparameter tuning, Batch Normalization, Programming Frameworks Quiz, 10 questions

8/10 points (80%)

