

## Week 3: Hyperparameter Tuning, Batch Normalization, Programming framework

1. 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?

**Ans:** False

2. 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?

**Ans:** False

3. During hyperparameter search, whether you try to babysit one model or train a lot of models in parallel is largely determined by,

**Ans:** The amount of computational power

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

**Ans:**

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

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?

**Ans:** False

6. In batch normalization as presented in course, if you apply it on the  $l^{th}$  layer of your neural network, what are you normalizing?

**Ans:**  $z^{[l]}$

7. In the normalization formula, why do we use epsilon?

**Ans:** Avoiding divide by zero.

8. Which of the following statements about  $\gamma$  and  $\beta$  in BatchNorm are true?

**Ans:**

- They can be learned using Adam, Gradient descent with momentum, or RMSprop, not just with gradient descent.
- They set the mean and variance of the linear variable  $z^{[l]}$  of a given layer.

9. After training a neural network with Batch Norm, at test time, to evaluate the neural network on a new example you should,

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

10. Which one of the following statements about programming frameworks is true?

**Ans:** 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.

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.