Congratulations! You passed!

Grade received 80%
Latest Submission Grade 80%
To pass 80% or higher

Go to next item

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? **1 / 1 point**

Expand

Correct

2. In a project with limited computational resources, which three of the following hyperparameters would you choose to tune? Check all that apply. **1 / 1 point**

Expand

Correct
Great, you got all the right answers.

3. Using the "Panda" strategy, it is possible to create several models. True/False? **0 / 1 point**

↙↗ **Expand**

⊗ **Incorrect**
Incorrect. Following the "Panda" analogy, it is possible to babysit a model until a certain point and then start again to produce a different one.

4. Knowing that the hyperparameter α should be in the range of 0.001 and 1.0. Which of the following is the recommended way to sample a value for α? **1 / 1 point**

↙↗ **Expand**

✓ **Correct**
Yes. This gives a random number between $0.001 = 10^{-3}$ and 10^{0}.

5. Once good values of hyperparameters have been found, those values should be changed if new data is added or a change in computational power occurs. True/False? **1 / 1 point**

↙↗ **Expand**

✓ **Correct**
Correct. The choice of some hyperparameters such as the batch size depends on conditions such as hardware and quantity of data.

6. When using batch normalization it is OK to drop the parameter $W^{[l]}$ from the forward propagation since it will be subtracted out when we compute $\tilde{z}^{[l]} = \gamma z^{[l]}_{\text{normalize}} + \beta^{[l]}$. True/False? **1 / 1 point**

↙↗ **Expand**

✓ **Correct**
Correct. The parameter $W^{[l]}$ doesn't get subtracted during the batch normalization process, although it gets re-scaled.

7. Which of the following are true about batch normalization? **1 / 1 point**

Expand

Correct

Yes. Internal covariance is a name to express that there has been a change in the distribution of the activations. Since after each iteration of gradient descent the parameters of a layer change, we might think that the activations suffer from covariance shift.

8. Which of the following statements about γ and β in Batch Norm are true? **0 / 1 point**

↙↗ **Expand**

⊗ **Incorrect**
You chose the extra incorrect answers.

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 point**

↙↗ **Expand**

✓ **Correct**

10. Which of the following are some recommended criteria to choose a deep learning framework? **1 / 1 point**

↙↗ **Expand**

Correct

Correct. The running speed is a major factor, especially when working with large datasets.