Congratulations! You passed!

Grade received 90% **Latest Submission Grade** 90% **To pass** 80% or higher

Go to next item

1. With a relatively small set of hyperparameters, it is OK to use a grid search. True/False? **1 / 1 point**

- (•) True
- ( ) False

Expand

Correct
Correct. When the set of hyperparameters is small like a range for $n_l = 1, 2, 3$ grid search works fine.

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? **1 / 1 point**

- ( ) True
- (•) False

Expand

Correct
Yes. We've seen in the lecture that some hyperparameters, such as the learning rate, are more critical than others.

3. Even if enough computational power is available for hyperparameter tuning, it is always better to babysit one model ("Panda" strategy), since this will result in a more custom model. True/False? **1 / 1 point**

- (•) False
- ( ) True

Expand

Correct
Correct. Although it is possible to create good models using the "Panda" strategy, obtaining better results is more likely using a "caviar" strategy due to the number of tests and the nature of the deep learning process of ideas, code, and experiment.

4. Knowing that the hyperparameter α should be in the range of 0.00001 and 1.0, which of the following is the recommended way to sample a value for α? **1 / 1 point**

- (•)
```
r = -5*np.random.rand()
alpha = 10**r
```

- r = np.random.rand()
  alpha = 0.00001 + r*0.99999
- r = -4*np.random.rand()
  alpha = 10**r
- r = np.random.rand()
  alpha = 10**r

Expand

Correct
Yes. This will generate a random value between 10^{-5} and 10^{0} chosen randomly in a logarithmic scale.

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**

- (•) True
- ( ) False

Expand

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

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? **1 / 1 point**

- ( ) $b^{[l]}$
- ( ) $a^{[l]}$
- (•) $z^{[l]}$
- ( ) $W^{[l]}$

Expand

Correct

7. In the normalization formula $z^{(i)}_{norm} = \frac{z^{(i)}-\mu}{\sqrt{\sigma^2+\varepsilon}}$, why do we use epsilon? **1 / 1 point**

- (•) To avoid division by zero
- ( ) To speed up convergence
- ( ) To have a more accurate normalization
- ( ) In case μ is too small

Expand

Correct

8. Which of the following are true about batch normalization? **0 / 1 point**

- [ ] $\beta^{[l]}$ and $\gamma^{[l]}$ are hyperparameters that must be tuned by random sampling in a logarithmic scale.
- [ ] The parameters $\gamma^{[l]}$ and $\beta^{[l]}$ set the variance and mean of $\tilde{z}^{[l]}$.
- [x] When using batch normalization we introduce two new parameters $\gamma^{[l]}$, $\beta^{[l]}$ that must be "learned" or trained.

Correct

Correct. Batch normalization uses two parameters β and γ to compute $\tilde{z}^{(i)} = \beta z_{norm}^{(i)} + \gamma$.

- [x] $z_{norm}^{(i)} = \frac{z^{(i)} - \mu}{\sqrt{\sigma^2}}$.

This should not be selected

No. We use a small parameter ϵ in the denominator to avoid division by z.

Expand

Incorrect

You didn't select all the correct answers

9. A neural network is trained with Batch Norm. At test time, to evaluate the neural network on a new example you should perform the normalization using μ and σ^2 estimated using an exponentially weighted average across mini-batches seen during training. True/false? **1 / 1 point**

- ( ) False
- (•) True

Expand

Correct

Correct. This is a good practice to estimate the μ and σ^2 to use since at test time we might not be predicting over a batch of the same size, or it might even be a single example, thus using the μ and σ^2 of a single sample doesn't make sense.

10. If a project is open-source, it is a guarantee that it will remain open source in the long run and will never be modified to benefit only one company. True/False? **1 / 1 point**

- (•) False
- ( ) True

Expand

Correct

Correct. To ensure that a project will remain open source in the long run it must have a good governance body too.