## Congratulations! You passed!

**Grade received** 100% **Latest Submission Grade** 100% **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**

- ○ True
- ◉ False

Expand

Correct

**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**

- ◉ False
- ○ True

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**

- ○ True
- ◉ False

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.** 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? **1 / 1 point**

- ○ r = np.random.rand()
  beta = r*0.09 + 0.9
- ○ r = np.random.rand()
  beta = 1-10**(- r + 1)
- ○ r = np.random.rand()
  beta = r*0.9 + 0.09

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

Expand

Correct

**5.** Finding new values for the hyperparameters, once we have found good ones for a model, should only be done if new hardware or computational power is acquired. True/False? **1 / 1 point**

- True
- False

Expand

Correct
Correct. As the data changes for the model, it might be beneficial to tune some of the hyperparameters again.

**6.** In batch normalization as presented in the videos, if you apply it on the $l$th layer of your neural network, what are you normalizing? **1 / 1 point**

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

Expand

Correct

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

- The parameters $\beta$ and $\gamma$ of batch normalization can't be trained using Adam or RMS prop.
- One intuition behind why batch normalization works is that it helps reduce the internal covariance.
- The parameter $\epsilon$ in the batch normalization formula is used to accelerate the convergence of the model.
- There is a global value of $\gamma$ and $\beta$ that is used for all the hidden layers where batch normalization is used.

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 $\gamma$ and $\beta$ in Batch Norm are true? **1 / 1 point**

- [x] They can be learned using Adam, Gradient descent with momentum, or RMSprop, not just with gradient descent.

Correct

- [ ] The optimal values are $\gamma = \sqrt{\sigma^2 + \varepsilon}$, and $\beta = \mu$.

- [ ] $\beta$ and $\gamma$ are hyperparameters of the algorithm, which we tune via random sampling.
- [ ] There is one global value of $\gamma \in \mathfrak{R}$ and one global value of $\beta \in \mathfrak{R}$ for each layer, and these apply to all the hidden units in that layer.
- [x] They set the mean and variance of the linear variable $\tilde{z}^{[l]}$ of a given layer.

✓ Correct

Expand

**Correct**
Great, you got all the right answers.

**9.** A neural network is trained with Batch Norm. At test time, to evaluate the neural network we turn off the Batch Norm to avoid random predictions from the network. True/False? **1 / 1 point**

- (•) False
- ( ) True

Expand

**Correct**
Correct. During the test, the parameters $\mu$ and $\sigma^2$ are estimated using an exponentially weighted average across mini-batches used during training.

**10.** Which of these statements about deep learning programming frameworks are true? (Check all that apply) **1 / 1 point**

- [x] Even if a project is currently open source, good governance of the project helps ensure that it remains open even in the long term, rather than become closed or modified to benefit only one company.

✓ Correct

- [x] 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.

Expand

**Correct**
Great, you got all the right answers.