Your grade: 100%

Your latest: **100%** • Your highest: **100%** • To pass you need at least 80%. We keep your highest score.

[Next item →]

---

1. If you have 10,000 examples, how would you split the train/dev/test set? Choose the best option.

      ⦿ 60% train. 20% dev. 20% test.

      ○ 33% train. 33% dev. 33% test.

      ○ 98% train. 1% dev. 1% test.

   1 / 1 point

   > ✓ **Correct**
   > Yes. This might be considered a small data set, not in the range of big data. Thus a more classical (old) best practice should be used.

2. When designing a neural network to detect if a house cat is present in the picture, 500,000 pictures of cats were taken by their owners. **These are used to make the training, dev and test sets.** It is decided that to increase the size of the test set, 10,000 new images of cats taken from security cameras are going to be used in the test set. Which of the following is true?

      ○ This will increase the bias of the model so the new images shouldn't be used.

      ⦿ This will be harmful to the project since now dev and test sets have different distributions.

      ○ This will reduce the bias of the model and help improve it.

   1 / 1 point

   > ✓ **Correct**
   > Yes. The quality and type of images are quite different thus we can't consider that the dev and the test sets came from the same distribution.

3. If your Neural Network model seems to have high bias, what of the following would be promising things to try? (Check all that apply.)

      ☐ Get more training data

      ☐ Add regularization

      ☑ Increase the number of units in each hidden layer

   1 / 1 point

   > ✓ **Correct**

      ☑ Make the Neural Network deeper

   > ✓ **Correct**

4. You are working on an automated check-out kiosk for a supermarket, and are building a classifier for apples, bananas and oranges. Suppose your classifier obtains a training set error of 0.5%, and a dev set error of 7%. Which of the following are promising things to try to improve your classifier? (Check all that apply.)

      ☑ Increase the regularization parameter lambda

   1 / 1 point

   > ✓ **Correct**

      ☐ Decrease the regularization parameter lambda

      ☑ Get more training data

   > ✓ **Correct**

      ☐ Use a bigger neural network

5. In every case it is a good practice to use dropout when training a deep neural network because it can help to prevent overfitting. True/False?

      ○ True

      ⦿ False

   1 / 1 point

   > ✓ **Correct**
   > Correct. In most cases, it is recommended to not use dropout if there is no overfit. Although in computer vision, due to the nature of the data, it is the default practice.

6. To reduce high variance, the regularization hyperparameter lambda must be increased. True/False?

      ⦿ True

   1 / 1 point

○ False

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

☐ It helps to reduce the bias of a model.

☐ In practice, it eliminates units of each layer with a probability of keep_prob.

☑ It helps to reduce overfitting.

⊘ **Correct**
Correct. The dropout is a regularization technique and thus helps to reduce the overfit.

☑ In practice, it eliminates units of each layer with a probability of 1- keep_prob.

⊘ **Correct**
Correct. The probability that dropout doesn't eliminate a neuron is keep_prob.

8. Increasing the parameter keep_prob from (say) 0.5 to 0.6 will likely cause the following: (Check the two that apply)　　　**1 / 1 point**

☐ Increasing the regularization effect

☑ Reducing the regularization effect

⊘ **Correct**

☐ Causing the neural network to end up with a higher training set error

☑ Causing the neural network to end up with a lower training set error

⊘ **Correct**

9. Which of the following actions increase the regularization of a model? (Check all that apply)　　　**1 / 1 point**

☑ Increase the value of the hyperparameter lambda.

⊘ **Correct**
Correct. When increasing the hyperparameter lambda, we increase the effect of the L_2 penalization.

☑ Decrease the value of keep_prob in dropout.

⊘ **Correct**
Correct. When decreasing the keep_prob value, the probability that a node gets discarded during training is higher, thus reducing the regularization effect.

☐ Decrease the value of the hyperparameter lambda.

☐ Use Xavier initialization.

☐ Increase the value of keep_prob in dropout.

10. Why do we normalize the inputs x?　　　　　　　　　　　　　　　　**1 / 1 point**

○ Normalization is another word for regularization--It helps to reduce variance

⦿ It makes the cost function faster to optimize

○ It makes it easier to visualize the data

○ It makes the parameter initialization faster

⊘ **Correct**