Your grade: 100%

Next item →

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

---

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

- ⦿ 98% train . 1% dev . 1% test
- ◯ 33% train . 33% dev . 33% test
- ◯ 60% train . 20% dev . 20% test

✓ **Correct**

---

2. The dev and test set should: 1 / 1 point

- ◯ Be identical to each other (same (x,y) pairs)
- ◯ Come from different distributions
- ◯ Have the same number of examples
- ⦿ Come from the same distribution

✓ **Correct**

---

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.)   1 / 1 point

- ☑ Make the Neural Network deeper

✓ **Correct**

- ☐ Add regularization
- ☑ Increase the number of units in each hidden layer

✓ **Correct**

- ☐ Get more training data

---

4. Your classifier for bananas and oranges gets a training set error of 0.1% and a development set error of 11%.   1 / 1 point

Which of the following statements are true? (Check all that apply.)

- ☐ The model is overfitting the development set.
- ☐ The model has a very high bias.
- ☑ The model has a high variance.

✓ **Correct**
The large gap between training and development set errors is a hallmark of high variance.

- ☑ The model is overfitting the training set.

✓ **Correct**
This is a classic indication of overfitting, where the model performs exceptionally well on the training data but poorly on unseen data.

---

5. What is weight decay?   1 / 1 point

- ◯ Gradual corruption of the weights in the neural network if it is trained on noisy data.
- ⦿ A regularization technique (such as L2 regularization) that results in gradient descent shrinking the weights on every iteration.
- ◯ A technique to avoid vanishing gradient by imposing a ceiling on the values of the weights.
- ◯ The process of gradually decreasing the learning rate during training.

✓ **Correct**

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

1 / 1 point

○ False

● True

> ⊘ **Correct**
> Correct. By increasing the regularization parameter the magnitude of the weight parameters is reduced. This helps reduce the variance.

7. With the inverted dropout technique, at test time:

1 / 1 point

● You do not apply dropout (do not randomly eliminate units) and do not keep the 1/keep_prob factor in the calculations used in training

○ You do not apply dropout (do not randomly eliminate units), but keep the 1/keep_prob factor in the calculations used in training.

○ You apply dropout (randomly eliminating units) and do not keep the 1/keep_prob factor in the calculations used in training

○ You apply dropout (randomly eliminating units) but keep the 1/keep_prob factor in the calculations used in training.

> ⊘ **Correct**

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 the hyperparameter lambda.

☑ 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.

☐ Increase the value of keep_prob in dropout.

☐ Use Xavier initialization.

10. Suppose that a model uses, as one feature, the total number of kilometers walked by a person during a year, and another feature is the height of the person in meters. What is the most likely effect of normalization of the input data?

1 / 1 point

○ It will increase the variance of the model.

○ It will make the data easier to visualize.

● It will make the training faster.

○ It won't have any positive or negative effects.

> ⊘ **Correct**
> Correct. Since the difference between the ranges of the features is very different, this will likely cause the process of gradient descent to oscillate, making the optimization process longer.