

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 examples, how would you split the train/dev/test set? Choose the best option.

1 / 1 point

- ☐ 98% train. 1% dev. 1% test.
- ☐ 33% train. 33% dev. 33% test.
- ☒ 60% train. 20% dev. 20% test.

 Expand



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. In a personal experiment, an M.L. student decides to not use a test set, only train-dev sets. In this case which of the following is true?

- ☐ He won't be able to measure the variance of the model.
- ☐ He won't be able to measure the bias of the model.
- ☒ He might be overfitting to the dev set.
- ☐ Not having a test set is unacceptable under any circumstance.

 Expand



Correct

Yes. Although not recommended, if a more accurate measure of the performance is not necessary it is ok to not use a test set. However, this might cause an overfit to the dev set.

3. If your Neural Network model seems to have high variance, what of the following would be promising things to try?

☐ Make the Neural Network deeper

☒ Get more training data

✓ **Correct**

☐ Increase the number of units in each hidden layer

☐ Get more test data

☒ Add regularization

✓ **Correct**

↗ **Expand**

✓ **Correct**

Great, you got all the right answers.

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 19% and a dev set error of 21%. Which of the following are promising things to try to improve your classifier? (Check all that apply, suppose the human error is approximately 0%)

- ☐ Get more training data.
- ☐ Increase the regularization parameter λ .
- ☒ Use a bigger network.

 Expand



Yes. This can be helpful to reduce the bias of the model, and then we can start trying to reduce the high variance if this happens.

5. What is weight decay?

1 / 1 point

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

 Expand

 **Correct**

6. What happens when you increase the regularization hyperparameter lambda?

1 / 1 point

- ☒ Weights are pushed toward becoming smaller (closer to 0)
- ☐ Gradient descent taking bigger steps with each iteration (proportional to lambda)
- ☐ Doubling lambda should roughly result in doubling the weights
- ☐ Weights are pushed toward becoming bigger (further from 0)

 Expand

 **Correct**

7. Which of the following are true about dropout?

1 / 1 point

☒ It helps to reduce the variance of a model.

✓ **Correct**

Correct. The dropout is a regularization technique and thus helps to reduce the variance.

☐ It helps to reduce the bias of a model.

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

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

✓ **Correct**

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

↗ **Expand**

✓ **Correct**

Great, you got all the right answers.

8. Decreasing the parameter `keep_prob` from (say) 0.6 to 0.4 will likely cause the following:

1 / 1 point

- ☐ Reducing the regularization effect.
- ☐ Causing the neural network to have a higher variance.
- ☒ Increasing the regularization effect.

 Expand

 **Correct**

Correct. This will make the dropout have a higher probability of eliminating a node in the neural network, increasing the regularization effect.

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 L₂ penalization.

☐ Decrease the value of the hyperparameter lambda.

☐ Increase the value of keep_prob in dropout.

☐ Use Xavier initialization.

☒ 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.

 Expand



Correct

Great, you got all the right answers.

10. Which of the following is the correct expression to normalize the input \mathbf{x} ?

1 / 1 point

☐ $x = \frac{1}{m} \sum_{i=1}^m x^{(i)}$

☒ $x = \frac{x - \mu}{\sigma}$

☐ $x = \frac{x}{\sigma}$

☐ $r = \frac{1}{\sqrt{\frac{1}{m} \sum_{i=1}^m (r^{(i)})^2}}$

 Expand



Correct

Correct. This shifts the mean of the input to the origin and makes the variance one in each coordinate of the input examples.