✓ **Congratulations! You passed!**

Grade received 100%   **Latest Submission Grade** 100%   **To pass** 80% or higher

Go to next item

---

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

○ 33% train. 33% dev. 33% test

○ 60% train. 20% dev. 20% test

◉ 98% train. 1% dev. 1% test

⤢ **Expand**

✓ **Correct**

---

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?   **1 / 1 point**

○ He won't be able to measure the bias of the model.

○ Not having a test set is unacceptable under any circumstance.

○ He won't be able to measure the variance of the model.

◉ He might be overfitting to the dev set.

⤢ **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. A model developed for a project is presenting high bias. One of the sponsors of the project offers some resources that might help reduce the bias. Which of the following additional resources has a better chance to help reduce the bias?   **1 / 1 point**

○ Gather more data for the project.

○ Use different sources to gather data and better test the model.

◉ Give access to more computational resources like GPUs.

⤢ **Expand**

✓ **Correct**
Yes. This can allow the developers to try bigger networks, train for more cycles, and test different architectures.

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.)

1 / 1 point

☑ Increase the regularization parameter lambda

> ✓ **Correct**

☐ Decrease the regularization parameter lambda

☑ Get more training data

> ✓ **Correct**

☐ Use a bigger neural network

⟋ **Expand**

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

5. Which of the following are regularization techniques?

1 / 1 point

☐ Increase the number of layers of the network.

☑ Dropout.

> ✓ **Correct**
> Correct. Using dropout layers is a regularization technique.

☐ Gradient Checking.

☑ Weight decay.

> ✓ **Correct**
> Correct. Weight decay is a form of regularization.

⟋ **Expand**

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

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

1 / 1 point

○ False

⦿ True

⟋ **Expand**

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

1 / 1 point

○ You apply dropout (randomly eliminating 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) 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.

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

⤢ **Expand**

✓ **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

⤢ **Expand**

9. Which of these techniques are useful for reducing variance (reducing overfitting)? (Check all that apply.)

1 / 1 point

☑ Dropout

> ✓ Correct

☐ Gradient Checking

☐ Exploding gradient

☐ Vanishing gradient

☐ Xavier initialization

☑ L2 regularization

> ✓ Correct

✓ Data augmentation

↗ **Expand**

10. Why do we normalize the inputs x?

1 / 1 point

○ It makes it easier to visualize the data

● It makes the cost function faster to optimize

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

○ It makes the parameter initialization faster

↗ **Expand**