

Congratulations! You passed!

[Go to next item](#)

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

1. If you have 20,000,000 examples, how would you split the train/dev/test set? Choose the best option.

1 / 1 point

- 60% train. 20% dev. 20% test.
- 99% train. 0.5% dev. 0.5% test.
- 90% train. 5% dev. 5% test.

[Expand](#)**Correct**

Yes. Given the size of the dataset, 0.5% of the samples are enough to get a good estimate of how well the model is doing.

2. The dev and test set should:

1 / 1 point

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

[Expand](#)**Correct**

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.
- Give access to more computational resources like GPUs.
- Use different sources to gather data and better test the model.

[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

1 / 1 point

- 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%)

4 / 4 points

- Use a bigger network.
- Increase the regularization parameter lambda.
- Get more training data.

 Expand

 Correct

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

- 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.
- A regularization technique (such as L2 regularization) that results in gradient descent shrinking the weights on every iteration.

 Expand

 Correct

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

1 / 1 point

- True
- False

 Expand

 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 apply dropout (randomly eliminating 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 do not apply dropout (do not randomly eliminate units), but 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

 Expand

 Correct

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

1 / 1 point

- Causing the neural network to have a higher variance.
- Increasing the regularization effect.
- Reducing 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 `keep_prob` in dropout.
- Use Xavier initialization.
- 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 the hyperparameter `lambda`.

 Correct

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

 Expand

 Correct

Great, you got all the right answers.

10. Why do we normalize the inputs \mathbf{x} ?

1 / 1 point

- Normalization is another word for regularization--It helps to reduce variance
- It makes it easier to visualize the data
- It makes the parameter initialization faster
- It makes the cost function faster to optimize

 Expand

 Correct

