Congratulations! You passed!

Grade received 80% | Latest Submission Grade 80% | 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 might be overfitting to the dev set.
- 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.

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? **1 / 1 point**

- [ ] Increase the number of units in each hidden layer
- [ ] Get more test data
- [x] Get more training data

✓ Correct

- [ ] Make the Neural Network deeper
- [x] 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%) **0 / 1 point**

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

Expand

Incorrect
No. First, the high bias problem must be addressed.

5. What is weight decay? **1 / 1 point**

- ○ The process of gradually decreasing the learning rate during training.
- ○ A technique to avoid vanishing gradient by imposing a ceiling on the values of the weights.
- ○ 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.

Expand

Correct

6. What happens when you increase the regularization hyperparameter lambda? **1 / 1 point**

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

Expand

Correct

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 1- keep_prob.

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 keep_prob.
- [x] It helps to reduce the variance of a model.

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

Expand

Correct
Great, you got all the right answers.

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
- [x] Reducing the regularization effect

Correct

- [ ] Causing the neural network to end up with a higher training set error
- [x] Causing the neural network to end up with a lower training set error

Correct

Expand

Correct
Great, you got all the right answers.

9. Which of the following actions increase the regularization of a model? (Check all that apply) **0 / 1 point**

- [x] Increase the value of the hyperparameter lambda.

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

- [ ] Increase the value of keep_prob in dropout.
- [ ] Decrease the value of the hyperparameter lambda.
- [ ] Make use of data augmentation.
- [x] Normalizing the data.

This should not be selected
Incorrect. Data normalization doesn't affect the variance of the model.

Expand

Incorrect
You didn't select all the correct answers

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

Correct