Week 01: Quiz

Congratulations! You passed!

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

Go to next item

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.
- (•) 60% train. 20% dev. 20% test.
- 33% train. 33% dev. 33% 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? **1 / 1 point**

- He won't be able to measure the variance of the model.
- (•) 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.

3. If your Neural Network model seems to have high variance, what of the following would be promising things to try? **1 / 1 point**

- [ ] Make the Neural Network deeper
- [x] Add regularization

  ✓ Correct

- [ ] Get more test data
- [x] Get more training data

  ✓ Correct

- [ ] Increase the number of units in each hidden layer

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 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**

- [x] Increase the regularization parameter lambda

  ✓ Correct

- [ ] Decrease the regularization parameter lambda
- [x] Get more training data

  ✓ Correct

- [ ] Use a bigger neural network

5. Which of the following are regularization techniques? **1 / 1 point**

- [x] Dropout.

  Correct
  Correct. Using dropout layers is a regularization technique.

- [ ] Gradient Checking.
- [ ] Increase the number of layers of the network.
- [x] Weight decay.

  Correct
  Correct. Weight decay is a form of regularization.

Expand

Correct
Great, you got all the right answers.

6. The regularization hyperparameter must be set to zero during testing to avoid getting random results. True/False? **1 / 1 point**

- (•) False
- ( ) True

Expand

Correct
Correct. The regularization parameter affects how the weights change during training, this means during backpropagation. It has no effect during the forward propagation that is when predictions for the test are made.

7. Which of the following are true about dropout? 1 / 1 point

- [ ] 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.

- [x] 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.

- [ ] It helps to reduce the bias of a model.

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

- (•) Increasing the regularization effect.
- ( ) Reducing the regularization effect.
- ( ) Causing the neural network to have a higher variance.

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 these techniques are useful for reducing variance (reducing overfitting)? (Check all that apply.) **1 / 1 point**

- [x] Data augmentation

  ✓ **Correct**

- [ ] Xavier initialization
- [ ] Exploding gradient
- [ ] Vanishing gradient
- [x] L2 regularization

  ✓ **Correct**

- [ ] Gradient Checking
- [x] Dropout

  ✓ **Correct**

Expand

Correct
Great, you got all the right answers.

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 make the data easier to visualize.
- (•) It will make the training faster.
- ( ) It won't have any positive or negative effects.
- ( ) It will increase the variance of the model.

Expand

Correct
Correct. Since the difference between the ranges of the features is very different, this will likely cause the