Try again once you are ready

Grade received 70%

Latest Submission Grade 70%

To pass 80% or higher

Try again

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

- (•) 99% train. 0.5% dev. 0.5% test.
- ( ) 60% train. 20% dev. 20% 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. 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? **0 / 1 point**

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

Expand

Incorrect
You did not choose an option.

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

- [ ] Get more test data
- [x] Add regularization

Correct

- [ ] Make the Neural Network deeper
- [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. Working on a model to classify bananas and oranges your classifier gets a training set error of 0.1% and a dev set error of 11%. Which of the following two are true? **1 / 1 point**

- [x] The model has a high variance.

✓ **Correct**

No. This model has a low bias and high variance.

- [ ] The model has a very high bias.
- [ ] The model is overfitting the dev set.
- [x] The model is overfitting the train set.

✓ **Correct**

Yes. This is precisely what happens when overfitting.

Expand

Correct

Great, you got all the right answers.

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

- [x] Weight decay.

✓ **Correct**

Correct. Weight decay is a form of regularization.

- [x] Dropout.

✓ **Correct**

Correct. Using dropout layers is a regularization technique.

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

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

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. During training a deep neural network that uses the tanh activation function, the value of the gradients is practically zero. Which of the following is most likely to help the vanishing gradient problem? **0 / 1 point**

- Use a larger regularization parameter.
- Use Xavier initialization.
- Increase the number of layers of the network.
- Increase the number of cycles during the training.

Expand

Incorrect
Incorrect. A high number of layers in a neural network can lead to vanishing or exploding gradients.

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

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

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

Make use of data augmentation.

Correct

Correct. Data augmentation has a way to generate "new" data at a relatively low cost. Thus making use of data augmentation can reduce the variance.

Increase the value of the hyperparameter lambda.

Correct

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

Expand

Incorrect

You chose the extra incorrect answers.

10. Why do we normalize the inputs x? **1 / 1 point**

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

Expand

Correct