Congratulations! You passed!

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

Go to next item

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

- ○ 90% train. 5% dev. 5% test.
- ○ 60% train. 20% dev. 20% test.
- ◉ 99% train. 0.5% dev. 0.5% 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. When designing a neural network to detect if a house cat is present in the picture, 500,000 pictures of cats were taken by their owners. **These are used to make the training, dev and test sets.** It is decided that to increase the size of the test set, 10,000 new images of cats taken from security cameras are going to be used in the test set. Which of the following is true? 1 / 1 point

- ○ This will reduce the bias of the model and help improve it.
- ◉ This will be harmful to the project since now dev and test sets have different distributions.
- ○ This will increase the bias of the model so the new images shouldn't be used.

Expand

Correct
Yes. The quality and type of images are quite different thus we can't consider that the dev and the test sets came from the same distribution.

3. If your Neural Network model seems to have high bias, what of the following would be promising things to try? (Check all that apply.) 1 / 1 point

- ☑ Make the Neural Network deeper

✓ Correct

- ☐ Get more training data
- ☑ Increase the number of units in each hidden layer

✓ Correct

- ☐ Add regularization

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 is overfitting the train set.

  ✓ **Correct**
  Yes. This is precisely what happens when overfitting.

- [x] The model has a high variance.

  ✓ **Correct**
  No. This model has a low bias and high variance.

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

Expand

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

5. In every case it is a good practice to use dropout when training a deep neural network because it can help to prevent overfitting. True/False? **1 / 1 point**

- ( ) True
- (•) False

Expand

✓ **Correct**
Correct. In most cases, it is recommended to not use dropout if there is no overfit. Although in computer vision, due to the nature of the data, it is the default practice.

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**

- [ ] It helps to reduce the bias of a model.
- [ ] In practice, it eliminates units of each layer with a probability of keep_prob.
- [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.

- [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. 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? **1 / 1 point**

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

Expand

✓ **Correct**
Correct. A careful initialization can help reduce the vanishing gradient problem.

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

- [ ] Vanishing gradient
- [x] Data augmentation

✓ **Correct**

- [x] L2 regularization

✓ **Correct**

- [ ] Gradient Checking
- [ ] Xavier initialization
- [x] Dropout

✓ **Correct**

- [ ] Exploding gradient

Expand

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

10. Which of the following is the correct expression to normalize the input $\mathbf{X}$? **1 / 1 point**

$$\cdots \quad i=1$$

- $x = \frac{1}{m}\sum_{i=1}^{m} x^{(i)}$
- $x = \frac{x}{\sigma}$
- $x = \frac{x - \mu}{\sigma}$

Expand

Correct

Correct. This shifts the mean of the input to the origin and makes the variance one in each coordinate of the input examples.