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.
- ( ) 90% train. 5% dev. 5% test.
- ( ) 60% train. 20% dev. 20% test.

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 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.
- ( ) This will reduce the bias of the model and help improve it.

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. 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

- (●) Give access to more computational resources like GPUs.
- ( ) Use different sources to gather data and better test the model.
- ( ) Gather more data for the project.

Correct
Yes. This can allow the developers to try bigger networks, train for more cycles, and test different architectures.

4. Your classifier for bananas and oranges gets a training set error of 0.1% and a development set error of 11%. 1 / 1 point

Which of the following statements are true? (Check all that apply.)

- [ ] The model is overfitting the development set.
- [x] The model has a high variance.

Correct
The large gap between training and development set errors is a hallmark of high variance.

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

Correct
This is a classic indication of overfitting, where the model performs exceptionally well on the training data but poorly on unseen data.

5. What is weight decay? 1 / 1 point

- ○ A technique to avoid vanishing gradient by imposing a ceiling on the values of the weights.
- ○ 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 regularization technique (such as L2 regularization) that results in gradient descent shrinking the weights on every iteration.

✓ **Correct**

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

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

✓ **Correct**

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

- ☑ It helps to reduce overfitting.

✓ **Correct**
Correct. The dropout is a regularization technique and thus helps to reduce the overfit.

- [x] In practice, it eliminates units of each layer with a probability of 1- keep_prob.

Correct
Correct. The probability that dropout doesn't eliminate a neuron is keep_prob.

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

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

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

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

- [ ] Decrease the value of the hyperparameter lambda.

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 increase the variance of the model.
- ◉ It will make the training faster.
- ○ It won't have any positive or negative effects.

Correct
Correct. Since the difference between the ranges of the features is very different, this will likely cause the process of gradient descent to oscillate, making the optimization process longer.