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

   ○ 60% train. 20% dev. 20% test.

   ⦿ 99% train. 0.5% dev. 0.5% test.

   ○ 90% train. 5% dev. 5% 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. 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

   ☐ Get more training data

   ☐ Add regularization

   ☑ Make the Neural Network deeper

   > ✓ **Correct**

   ☑ Increase the number of units in each hidden layer

   > ✓ **Correct**

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?

- [ ] The model has a very high bias.

- [x] *D: The model has a high variance.

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

- [x] The model is overfitting the train set.

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

- [ ] The model is overfitting the dev set.

5. What is weight decay?

- (●) A regularization technique (such as L2 regularization) that results in gradient descent shrinking the weights on every iteration.

- ( ) 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 technique to avoid vanishing gradient by imposing a ceiling on the values of the weights.

> ✓ **Correct**

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

- (●) False

- ( ) True

> ✓ **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 1- keep_prob.

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

☑ It helps to reduce the variance of a model.

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

☐ It helps to reduce the bias of a model.

☐ In practice, it eliminates units of each layer with a probability of keep_prob.

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.

○ Causing the neural network to have a higher variance.

○ Reducing the regularization effect.

> ✓ **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 the following actions increase the regularization of a model? (Check all that apply)

1 / 1 point

☐ Increase the value of keep_prob in dropout.

☐ Decrease the value of the hyperparameter lambda.

☑ Increase the value of the hyperparameter lambda.

> ✓ **Correct**
> Correct. When increasing the hyperparameter lambda, we increase the effect of the L_2 penalization.

☐ Use Xavier initialization.

☑ Decrease the value of keep_prob in dropout.

> ✓ **Correct**
> Correct. When decreasing the keep_prob value, the probability that a node gets discarded during training is higher, thus reducing the regularization effect.

○ Increasing the regularization effect.

○ Causing the neural network to have a higher variance.

○ Reducing the regularization effect.

> ✓ **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 the following actions increase the regularization of a model? (Check all that apply)

1 / 1 point

☐ Increase the value of keep_prob in dropout.

☐ Decrease the value of the hyperparameter lambda.

☑ Increase the value of the hyperparameter lambda.

> ✓ **Correct**
> Correct. When increasing the hyperparameter lambda, we increase the effect of the L_2 penalization.

☐ Use Xavier initialization.

☑ Decrease the value of keep_prob in dropout.

> ✓ **Correct**
> Correct. When decreasing the keep_prob value, the probability that a node gets discarded during training is higher, thus reducing the regularization effect.

10. Which of the following is the correct expression to normalize the input x?

1 / 1 point

◉ $x = \frac{x - \mu}{\sigma}$

○ $x = \frac{x}{\sigma}$

○ $x = \frac{1}{m} \sum_{i=1}^{m} x^{(i)}$

○ $x = \frac{1}{m} \sum_{i=1}^{m} (x^{(i)})^2$

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