Week 1: Practical aspects of deep learning

1. If you have 10,000,000 examples, how would you split the train/dev/test set?

   Ans: $98 : 1 : 1$

2. The dev and test set should:

   Ans: Should come from the same distribution

3. If your Neural Network model seems to have high variance, which of the following would be promising things to try?

   Ans: Add regularization, Get more training data

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.)

   Ans: Increase the regularization parameter λ and get more training data. Note: The network has over-fitting/high variance seen in train error, and under-fitting/high bias seen via dev set.

5. What is weight decay?

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

6. What happens when you increase the regularization hyperparameter λ

   Ans: Weights are pushed toward becoming smaller.

7. With the inverted dropout technique, at test time,

   Ans: You do not apply dropout (do not randomly eliminate units) and do not retain the $\frac{1}{prob}$ in the calculations used in training

8. Increasing the parameter `keep-prob` from say 0.5 to 0.6 will likely cause the following: (Check the two that apply)

   Ans:

   - Reducing the regularization effect
   - Causing the neural network to end up with a lower training set error

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

   Ans:

   - Dropout
   - L2 regularization
   - Data augmentation

10. Why do we normalize the inputs?

    Ans: It makes the cost function faster to optimize