Deep Learning
Midterm Exam

October 16, 2020

Instructor: Jia Xu

Student name: ————————————————————————

Student ID: ————————————————————————

Student email address: ————————————————————————

- **Read these instructions carefully**

- Do *not* turn this page until you are instructed to do so

- Fill-in your personal info as indicated above

- You have 90 minutes

- There are 10 questions.

- Each question worths the same (10 points)

- This is an open-book test.

good luck!

1 Question

1. A neural network with multiple hidden layers and sigmoid nodes can form non-linear decision boundaries.

   TRUE   FALSE

2. It is possible to represent a XOR function with a neural network without a hidden layer.

   TRUE   FALSE

3. As model complexity increases, bias will decrease while variance will increase.

   TRUE   FALSE

4. Choose one of following statements that describes the process of task execution (classifying input signal) in a feed-forward, back-propagation neural network:

   Choose one of following statements that describes the process of learning:

   a. Activation levels are propagated from the inputs through the hidden layers to the outputs.

   b. Activation levels are propagated from the outputs through the hidden layers to the inputs.

   c. Weights on the links are modified based on messages propagated from input to output.

   d. Weights on the links are modified based on messages propagated from output to input.

   e. Connections in the network are modified, gradually shortening the path from input to output.

   f. Weights at the input level are compared to the weights at the output level, and modified to reduce the discrepancy.

2 Question

What is the risk with tuning hyperparameters using a test dataset?

3 Question

Explain what effect will the following operations generally have on the bias and variance of your model. Fill in one of increases, decreases or no change in each of the cells:

	Bias	Variance
Regularizing the weights		
Increasing the size of the layers (more hidden units per layer)		
Using dropout[1] to train a deep neural network		
Getting more training data (from the same distribution as before)		

4 Question

What problem(s) will result from using a learning rate thats too high? How would you detect these problems? What problem(s) will result from using a learning rate thats too low? How would you detect these problems?

5 Question

What is a recurrent net? How do we apply backprop to a recurrent net?

6 Question

You are given a dataset with 1 million samples, each sample has 10,000 variables, some of which are counts (e.g. how many times people have clicked on a particular link, visited a particular page, or bought a particular product), others are measurements whose distributions seem close to Gaussian. Describe the steps to prepare the data.

7 Question

You are given X-ray images of patients where malignant tumors have been labeled at the pixel level (for each X-ray, you have another image where each pixel is labeled 1 for tumor and -1 for non tumor). Describe how you could use deep learning to solve the problem of detecting malignant tumors. specify which type of model you would use, and how you would go about setting up the dataset to train it.

8 Question

Assume the artificial neural network below, with mean square error loss and gold output of 3. Compute the values of all weights w_i after performing a SGD update with learning rate 0.1.
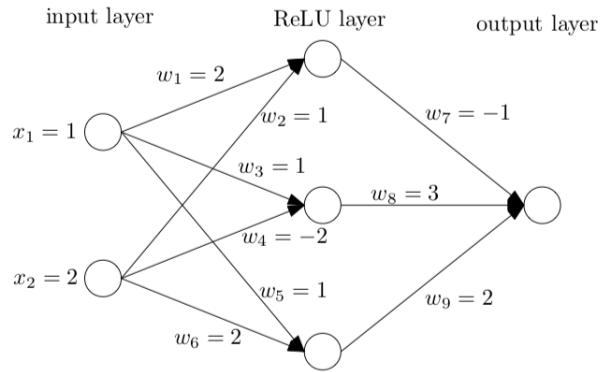
9 Question (10 points)

Suppose that you are training a Speech Recognition System on 100 million sentences. Your goal is to improve the accuracy of the translation system. Your small team of experts and engineers has 1 day only to improve the system. which of the following task/tasks would you?

1. Preprocessing

2. Change of the training data

3. Implement a new training algorithm

4. Add new features

5. Change pruning threshold in search

   Justify your answer:

10 Question (10 points)

1. What is the loss function of the feed-forward neural networks, explain what each symbol in it, represents:

2. Describe the basic ideas underlying Gradient Descent Learning Algorithms, and how the above equation would be used for such training.

3. Outline how, in general, you could estimate the generalization ability of a trained neural network.

4. How could you modify the above equation to improve the generalization ability of your trained neural network?