# Assignment 4 and 5

**Problem Set 4 & 5: Neural Networks**

In this overloaded problem set, we will study neural network training from the perspective of maximum likelihood and maximum a posteriori parameter learning, and Bayesian parameter learning (optional). We will also study various neural network architectures such as neural network with L2 regularization, autoencoder, generative adversarial network, recurrent neural network. We will use MNIST data set (our old friend) and Tensorflow, but the skills learned from this problem set will be easily transferable to other images, non-images, and neural network packages.

1. (5 points) We will train a neural network to identify the digit on a image in the MNIST data set from a training data set. This neural network has 10 softmax output nodes generating where m=0,1,...,9. Let be the images arranged into a vector, be the label of the image , be the synaptic weights of the neural network, and be the index of a pattern in the training data set. Demonstrate that a neural network to maximize the log likelihood of observing the training data is one that has softmax output nodes and minimizes the criterion function of the negative log probability of training data set: . Demonstrate that a neural network to maximize the a posterior likelihood of observing the training data given a Gaussian prior of the weight distribution is one that minimizes the criterion function with L2 regularization .

2 (a). (5 points) Build a neural network with 1 hidden layer of 30 sigmoid nodes, and an output layer10 softmax nodes from 1000 training images (100 images per digit). Train the network for 30 complete epochs, using mini-batches of 10 training examples at a time, a learning rate $\eta=0.1$. Plot the training error, testing error, criterion function on training data set, criterion function on testing data set of a separate 1000 testing images (100 images per digit), and the learning speed of the hidden layer (the average absolute changes of weights divided by the values of the weights).

2 (b). (5 points) Repeat 2 (a) with 2 hidden layers of 30 sigmoid nodes each, 3 hidden layers of 30 sigmoid nodes each, and with and without L2 regularization and . (You will repeat 2(a) for 5 times: 1 for 2 hidden layer network; 1 for 3 hidden layer network; and 1 times each for 1, 2, 3 hidden layers with regularization.)

2 (c). (5 points) Construct and train convolutional neural network for MNIST classification. Regularize the training of the neural network through dropout. Regularize the training of neural network through augment your selection of 1000 images by rotating them for 1-3 degrees clockwise and counter clockwise, and shifting them for 3 pixels in 8 different directions. You can find many tutorials on those techniques, and our emphasize is that we understand those techniques.

3. (Optional) Train GAN to generate the images for the 10 digits from random noise. Train autoencoder network with linear and sigmoid activation functions for principle component analysis. Train recurrent neural network to accept the 28 rows and output the digit of the image.

4. (Optional) Train Bayesian neural network with variational and sampling based method using Edward and Tensorflow. We will cover Bayesian neural network in the lecture.