**PSL432 Theoretical Physiology**
**Department of Physiology**
**University of Toronto**

# ASSIGNMENT 2

Assigned 27 February 2024
Due 12 March 2024

This assignment is worth 40% of your final mark in the course. Please email your submission, consisting of four m-files, to douglas.tweed @utoronto.ca and ankit.roy@mail.utoronto.ca by 11:59 p.m. on 12 March. Your code will be graded partly on concision and clarity, so keep it brief and orderly. And write it in plain Matlab without using any of the Matlab toolboxes.

In what follows, INIT and SURNAME stand for your given-name initial and your surname in upper-case letters, e.g. for me, D and TWEED. And init and surname stand for your initial and surname in lower-case letters, e.g. d and tweed.

Your job is to use backpropagation with adam to train a network to recognize the handwritten digits in the MNIST data set. The network should be a 4-layer mapnet where the two middle layers are 28-by-28 maps with field-fractions of 0.25. Please construct your network so that each layer projects to the next one (as in create_net.m), but layer 2 *also* projects *directly* to layer 4 through weight matrix $W\{4\}$ — so $W\{4\}$ is a 10-by-1568 matrix with input $[y\{2\}; y\{3\}]$. And please make the connections from both middle layers to the output layer dense (that is, fully connected, *not* maplike).

In the two middle layers, use the following activation function, which we'll call *relog*:

$$y\{l\}_i = \max(\ 0, \operatorname{sgn}(v\{l\}_i) \log(\ 1 + |v\{l\}_i|\ )\ ).$$

Make the final layer of your network affine, but then apply a *softmax* function to the network's output signal:

$$y_i = \exp(y\{4\}_i) \,/\, \Sigma_j \exp(y\{4\}_j).$$

From the resulting $y$ vector and the one-hot label vector $y^*$, compute the error and loss as $e = y - y^*$ and $L = e^\mathsf{T}e/2$.

To make your mapnet, write an m-file called init_surname_create_mapnet, analogous to the posted m-file create_net. All the vectors and matrices associated with your mapnet should be initialized inside init_surname_create_mapnet. In particular, that m-file should set all biases $b$ to 0, and should initialize each neuron's weight by drawing it from a normal distribution with a mean of 0 and standard deviation of sqrt(2/$N$), where $N$ is the number of the neuron's inputs, i.e. the number of upstream neurons projecting to it directly.

For learning, write m-files init_surname_forward_relog and init_surname_backprop_relog, analogous to the posted m-files forward_relu and backprop_relu but modified to use the relog function instead of relu and to take into account the direct projection from layer 2 to layer 4. Please compute the softmax *outside* your forward_relog function, so your code calls your forward_relog and then applies softmax to its output. And compute the derivative $\partial L/\partial v\{4\}$ outside your backprop_relog function and use that derivative as the input to backprop_relog. Implement adam using the posted m-file adam.m, and set the adam $\eta$ = 0.001 and the other adam hyperparameters to their usual values.

You'll find the MNIST data set in the file MNIST.mat, along with MNIST_key.m, which explains the format of those data. Load the data set as in MNIST_key.m. Feed the images to the network in minibatches of 100 each. If you can avoid it, don't write any code, anywhere in your files, that uses a loop or list to go through the minibatch example by example; instead, for concision and speed, write your code so that all operations are done on the whole minibatch array at once, as in the posted sample code.

Run 10 epochs — that is, 10 passes through all 60,000 training examples — with 600 minibatches in each epoch. At the start of each epoch, shuffle the data, as in MNIST_key.m (this shuffling can aid learning, as it supplies a different series of 600 minibatches to the network in different epochs).

After each epoch, run your network on the entire test set (in one big “minibatch” of all 10,000 test examples), count the number of incorrect guesses, and display that number in the Command Window. If all goes well, then on most runs, the number of incorrect guesses should be 400 or better after one epoch, and should fall below 160 at least once in the first 10 epochs.

Submit an m-file called INIT_SURNAME_A2.m, along with init_surname_create_mapnet.m, init_surname_forward_relog.m, and init_surname_backprop_relog.m.