Kathlene Ngo

Melanie Zhang

Adam Kagel

Yunxuan He

MAT 128B Project#2

REPORT SUMMARY

**i. Download *MNIST\_all.mat* from Greenbaum and Chartier textbook website. Read pages 179-180.**

**ii. Plotting digits**

File: Project2Part2Ex1.m

%% Here we are implementing a program that reads digits from the data base given by the textbook’s website (with MNIST\_all.mat)

digit = train0(1,:);

digitImage = reshape(digit, 28, 28);

image(rot90(flipud(digitImage),-1));

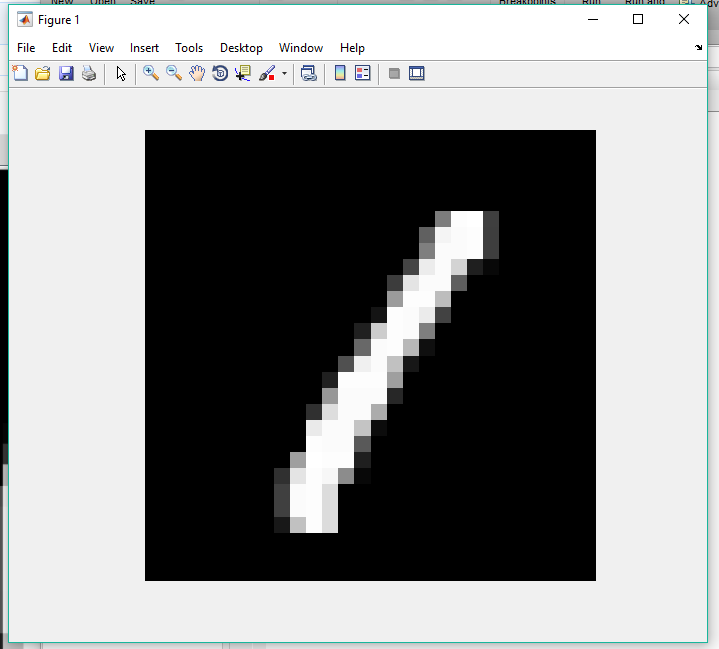
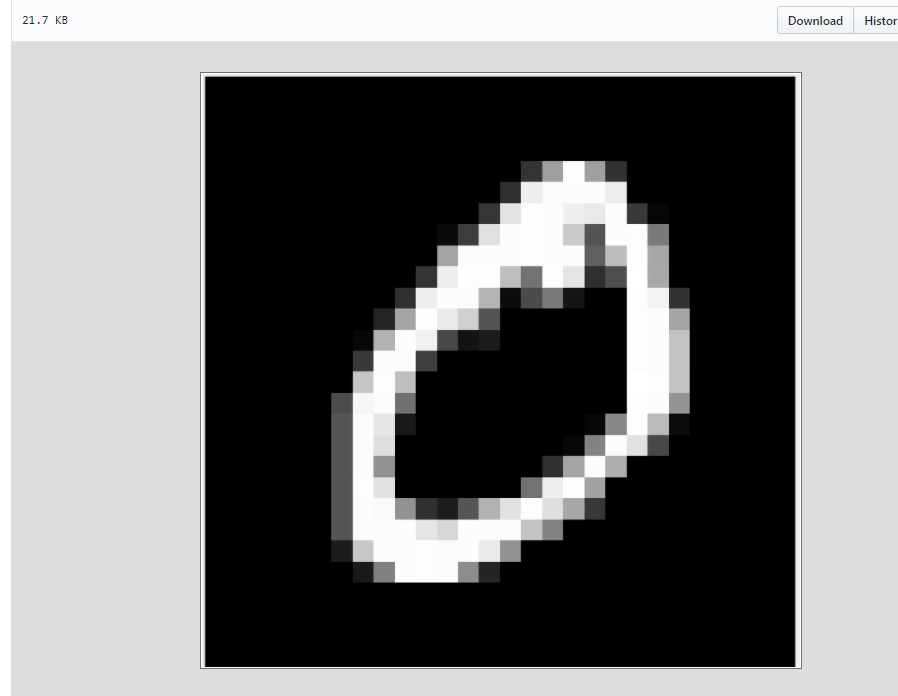
colormap(gray(256)), axis square tight off;

digit = train1(1,:);

digitImage = reshape(digit, 28, 28);

image(rot90(flipud(digitImage),-1));

colormap(gray(256)), axis square tight off;



File: MNISTData.m

ImageExample#2.png

ImageFor#2.png

%% Then, we compute the average digit using command T(1,:) = mean(train0), T(2,:) = mean(train1);etc and plot them, we get the image shown on the book.

%%17a

T(1,:) = mean(train0);

T(2,:) = mean(train1);

T(3,:) = mean(train2);

T(4,:) = mean(train3);

T(5,:) = mean(train4);

T(6,:) = mean(train5);

T(7,:) = mean(train6);

T(8,:) = mean(train7);

T(9,:) = mean(train8);

T(10,:) = mean(train9);

for i = 1:10,

subplot(2,5,i);

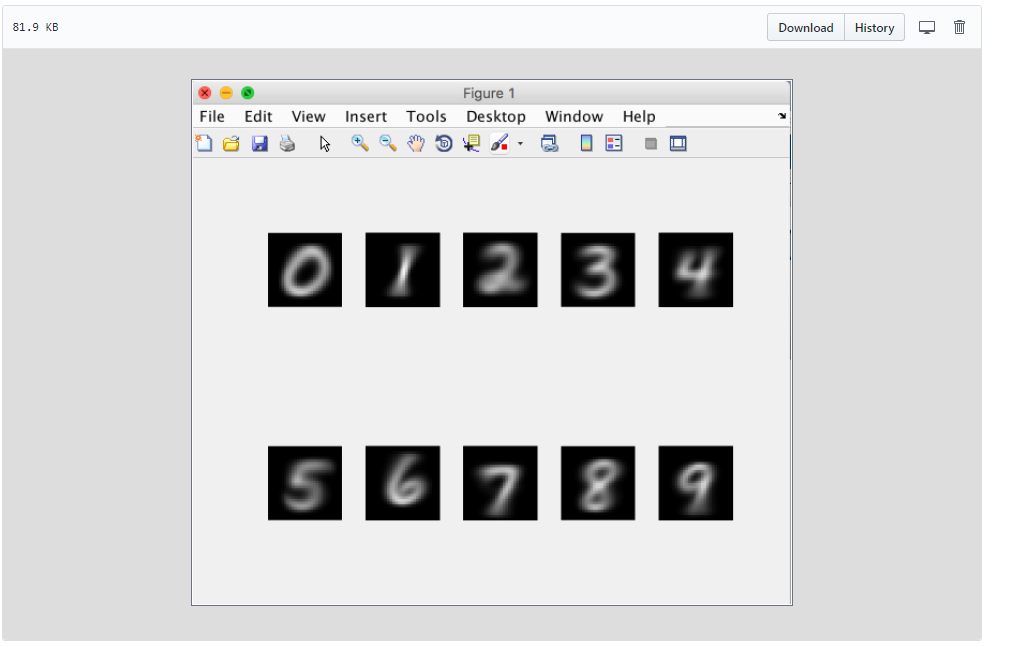
j = T(i,:);

digitimage1 = reshape(j, 28, 28);

image(rot90(flipud(digitimage1),-1));

colormap(gray(256)), axis square tight off

end;



**iii- A neuron. Implement a neuron where F is given as the logistic function. Take the derivative and analyze the function.**

File: Neuron.m

Also see attached paper.

function [ OUT ] = Neuron( InputList, InputWeight )

%UNTITLED2 Summary of this function goes here

% Detailed explanation goes here

% Format for InputList [x1;x2;x3;etc]

% Format for InputWeight [y1 y2 y3 etc];

NET = InputWeight \* InputList

OUT = 1/(1+exp(-NET));

End

File: Project2Part3.m

% Give several pairs of InputList and InputWeight values

InputList = [1; 2; 3];

InputWeight = [1 2 3];

Neuron(InputList,InputWeight)

% This gives the case where NET = 14 OUT = 9.999991684719722e-01

I = [.1;.1;.1];

O = [.2 .2 .2];

Neuron(I, O)

% This gives the case where NET = 6.000000000000001e-02 and the

% corresponding OUT = 5.149955016194100e-01

J = [.01;.01;.01];

P = [.02 .02 .02];

Neuron(J, P)

% This gives NET = 6.000000000000001e-04 and the OUT = 5.001499999954999e-01

% From the above three cases we observe that for F = logistics function, as

% NET gets smaller, the value of OUT is getting smaller too. This implies

% that smaller value of NET gives small value of OUT and vice versa.

We notice that the derivative of the sigmoidal (logistic) function from Figure 2 has a nice expression in terms of the OUT value.

F(NET) = OUT = 1/(1+exp(-NET))

F’(NET) = OUT(1-OUT) = (1/(1+exp(-NET)))(1-(1/(1+exp(-NET)))

Using test values, we notice that the smaller value of the NET, then the smaller value of the OUT.

We can also use the exponential function using b (constant as the base) such that F(NET)=OUT=1/(1+b^(-NET))

**iv. Multilayer Network. Implement a network with a variable number of hidden networks.**

File: **MultiLayerNetwork.m**

function [ O ] = MultiLayerNetwork( InputList, InputWeights)

% PART 4!!!

% Format for InputList [x1;x2;x3;etc]

% Format for InputWeights cell array [W1;W2;W3;etc] with each Wi as a

% weight matrix

O = InputList;

for i=1:length(InputWeights) % last layer is indexed at length+1

NET = InputWeights{i} \* O; % "NET = Xw", where X is the ith matrix

% contained in InputWeights and w is the

% input vector

O = 1/(1+exp(-NET)); % output is O = F(NET)

O = O.';

end

The last layer is NOT part of the network. It contains the values of the training set and comparing to output.

INPUT LAYER does NOT do the calculations; Neurons in the HIDDEN and the OUTPUT layer contain NET and OUT.

However many weight matrices = hidden layers

The last layer is at length+1 and does not have a weight matrix.

**File : Project2Part5.m**

% PART 5!

% Initialize the network by assigning a random (small) number to each

% weight

n = 3; % number of layers

WeightMatrices = cell(1,n);

for i=1:n

WeightMatrices{i} = rand([4 4]); % weights for each layer, initialized

% to be random numbers between 0 and 1

end

input = [1; 2; 3; 4];

MultiLayerNetwork(input, WeightMatrices)

In this file, we are simply initializing network taking a random (small) number to each weight.

vi.- Training the network

A network will learn by iteratively adapting the values of wi,j

. Each input is associated to an output. These are called training pairs

Here is our algorithm (from PDF):

* Select next training pair (INPUT, OUTPUT) and apply the INPUT to the network (“forward pass”)
* Calculate output using the network (“forward pass”)
* Calculate error between the network’s output and the desired output (“reverse pass”)
  + ERROR (=|TARGET−OUT|)
* Adjust weights that minimizes the error (“reverse pass”)
* Repeat steps for each training pair

Calculations are performed by layers, in the hidden layer(s) before any calculation is performed in the output layer.

Forward: Weights between neurons can be represented by matrix W and NET= XW (X being the input vector) (Output vector is input vector for next iteratin)

Reverse: ERROR signal when comparing OUTPUT with the TARGET value. (neuron p, hidden layer j to neuron q, output layer k)

vii.- Dependence on parameters

The learning of the network (i.e. the minimization

Of the error) will depend on the number of layers, the number of neurons per layer, and for fixed values of these two parameters the network will also depend on the size of the training set. Set up a study in which you change the values of these parameters and report the error(s) you obtain (you will obtain an error for the training set and another for the test test–which should be very similar to each other provided the test and training set are similar enough).