

# Homework 3 Report

Sean Kane

November 15, 2018

## 1 Problem 1

MNIST Neural Network

### 1.1 System Description

Description of all choices made - num of hidden neurons, learning rate, momentum, output thresholds, rules for choosing initial weights, criterion for deciding when to stop training, etc

This neural network is built with 784 input layers - each corresponding to one pixel of a 28 x 28 image - 150 hidden neurons, and ten output neurons. Each of the ten output neurons corresponded to an identified number. For example, if the output of the zeroth neuron was a one, this would predict a zero, a one at the first neuron meant a prediction of one. A learning rate of 0.01 and an alpha value of 0.5 was used for the implementation of momentum.

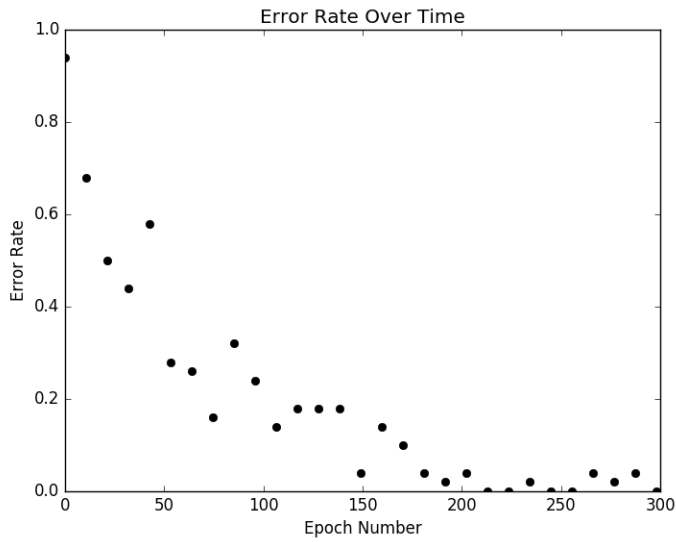
During training the output threshold was determined by which neuron had the highest output value of over 0.75. If there was not a neuron over 0.75, the output was null and the system was trained as if it was incorrect. During testing, the predicted neuron was the neuron with the highest value, regardless of whether it was over 0.75.

Training of the neural network occurred until one of two conditions was met. The first condition was a maximum number of epochs set at 800. This number was chosen because I wanted the machine to see each test point an average of ten times, and with a mini batch size of 50 and a test sample size of 4000, the epoch on average would be 800. The second condition, which occasionally would cause training to end before the max epochs, was if the percent change in loss dropped below .01 percent. Percent change was measured as:

$$PercentChange = \frac{Loss_{N-1} - Loss_N}{Loss_{N-1}}$$

### 1.2 Results

Performance of final network on training and test set using confusion matrix. Plot time series of error during training saved every tenth epoch Testing accuracy: 0.8461 after 297 epochs.



### 1.3 Confusion Matrix

0	1	2	3	4	5	6	7	8	9
979	0	5	3	0	36	9	0	0	0
13	2,116	88	170	59	100	53	137	17	40
24	2	960	80	0	12	21	3	67	1
14	0	4	850	0	84	0	0	2	2
8	18	28	18	1,779	55	43	70	71	361
12	0	2	21	7	963	3	0	27	4
22	6	25	8	21	16	1,388	6	27	6
43	30	141	75	62	52	33	1,800	34	162
0	0	0	0	0	0	0	0	391	0
4	7	16	9	44	26	16	15	69	944

### 1.4 Error Rate Graph

### 1.5 Analysis of Results

Describe, discuss, and interpret results you got, and why you think they as as they are.

## 2 Problem 2

Autoencoder

### 2.1 System Description

Description of learning rate, momentum, initial weights, when to stop training.

## 2.2 Results

Performance of final network on the training set and test set. Two values which should be plotted as two bars side by side. Plot same error in the same way, separated by each digit - two bars for 0, two for 1, etc. Plot time series of the overall training error during training using the data saved at every tenth epoch.

## 2.3 Features

Plot images for a large number of features, just like data shown in first100.jpg

## 2.4 Analysis of Results

Describe, discuss, and interpret results you got and why you think they are. Comment on features you found, and what they suggest. Comment on which digits turned out to be easiest to reconstruct and which ones more difficult, and why you think that happened.