Kevin Tran
01 December 2018
CS 460 – Machine Learning
Alfred Nehme
Assignment 2

Artificial Neural Network for Grayscale Images of Arabic Numerals

Kevin Tran

Parameters

The parameters that had to be implemented in the ANN was determined through manual trial and error after getting the ANN to function correctly. All stress testing of various units was done by taking 1000 random samples of the 10000 instances in the random sample mnist_test.csv. Initially testing was around 10 hidden units with 10 epochs and a learning rate of 0.01. The weights initialized in this ANN was +/- 0.05. Program also was modified in the JVM to add additional RAM to run tests. Tests results are considered as an average (the test itself in a case is ran three times).

In terms of the actual testing, I am assuming that all results will be trained in order to get the highest accuracy. Anything that is sampled is randomly sampled.

Hidden Units

The number of hidden units were first set to 1 unit to check and ensure ANN function. Afterwards, ANN units were then set to 10. From there, I have eyeballed and checked the average outputs of incrementing this value and felt that I should initiate full checks starting at 31. Tests were done all the way until 35. From there, I counted the number of max occurrences that each hidden unit test has held per category of test sample accuracy (refer figure 1). From there I have reached a conclusion that 33 had more best

	31 average	32 average	33 average	34 average	35 average
1	10.7	10.3	10.23333333	10.633333	9.566666667
100	13.66666667	9.833333333	12.93333333	10.766667	10
500	13.76666667	11.4	13.53333333	19.233333	10
1000	21.93333333	26.13333333	28.43333333	22.533333	27.23333333
1500	33.96666667	39.96666667	32.73333333	35.266667	38.86666667
2000	46.96666667	49.33333333	46.96666667	50.4	44.4
3000	67.1	68.43333333	68.23333333	69.3	68.33333333
4000	77.86666667	77.3	77.36666667	78.566667	80.8
5000	85.13333333	84.13333333	86.1	84.533333	84.96666667
6000	85.96666667	88.16666667	87.5	87.066667	88.1
7000	88	88.4	89.33333333	88.1	88.53333333
8000	89.3	89.3	89.9	89.9	90.16666667
9000	89.43333333	89.73333333	90.73333333	90.166667	89.2
10000	90.06666667	89.56666667	89.1	90.466667	90.06666667
20000	92.36666667	92.1	91.56666667	92.5	92.76666667
30000	92.9	92.8	93.4	92.466667	92.16666667
40000	93.26666667	93.9	93.43333333	93.133333	94.8
50000	93.13333333	94.1	94.26666667	93.133333	93.83333333
60000	94.53333333	94.56666667	94.33333333	93.833333	94.76666667

Figure 1

Hidden Unit Best Averages by Node Count vs Test Size

averages in my ANN testing scheme and so I used that value in my ANN.

Learning Rate

The learning rate was experimented next by changing it in 0.001 increments. Adding or subtracting that value did not create a better accuracy so I kept it as it for 0.01.

Epochs

Per configuration run with hidden units set at 33, I also eyeballed the change between numbers from 10 and decided to keep track of testing at 12 epochs to 19 epochs. Attached is a graph showing the differences in accuracy vs. test sizes between different epoch settings (figure 2). Since it seemed that epochs over 16 will not have much significant differences, I decided to terminate testing early and use 16 epochs as my default setting.

Kevin Tran
01 December 2018
CS 460 – Machine Learning
Alfred Nehme
Assignment 2

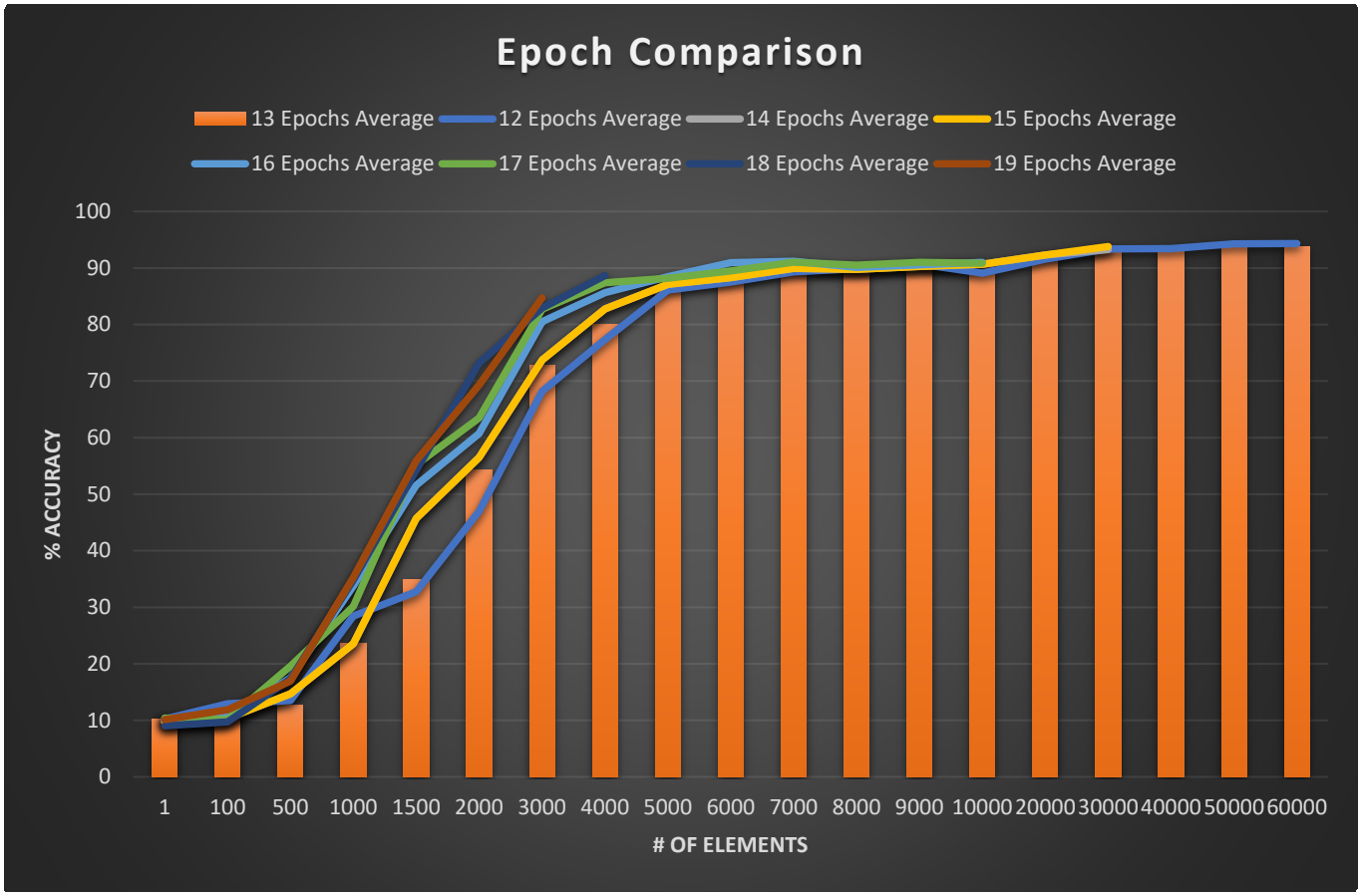

Figure 2