Peter Bui ID: 973296176

Description:

This experiment takes two data sets, mnist_train.csv, and 'mnist_test.csv. They contain pixel values corresponding to images of handwritten digits. The purpose of the experiment is to train a group of 10 perceptrons to correctly identify each example's 784 inputs over three different learning rates (.001, .01, .1). Starting with randomized weights, they are adjusted over the course of 50 epochs as the perceptrons calculate them with a given input. The weights change if their calculated outputs do not match the target label of the current example. Predictions are stored by taking the largest calculated value out of the 10 perceptrons, and their accuracy is put against the total amount of predictions. A confusion matrix is also utilized to see the perceptron's choice versus the correct choice. This shows which digits are commonly misinterpreted.

Learning Rate 1 (0.001):

Plot:

Any oscillations or overfitting?

There appears to be small bumps

Confusion Matrix:

Training

Learning Rate: .001

Correct Answer

Predicted Answer

```
[2.85331e+05 5.80000e+01 2.25100e+03 1.81500e+03 8.18000e+02 3.23500e+03 2.22900e+03 1.03600e+03 3.01700e+03 1.69100e+03]
[5.50000e+01 3.21581e+05 3.16600e+03 1.58900e+03 1.43500e+03 1.19200e+03 1.11300e+03 1.08100e+03 7.31700e+03 1.39200e+03]
[ 1703.   2825. 259537.   9834.   2393.   3080.   3619.   4251.   8044.  1695.]
[  908.   1471.   7731. 259900.    905.  14483.    488.   2346.  12805.  6034.]
[  543.    383.   3145.    877. 261398.   4052.   2116.   3324.   3682.  14521.]
[ 2194.   2370.   1986.  14558.   1803. 221336.   5922.   1504.  15059.  5123.]
[2.31000e+03 5.52000e+02 4.30200e+03 1.44400e+03 2.21500e+03 5.75700e+03 2.76974e+05 2.75000e+02 2.80500e+03 2.16000e+02]
[  390.    933.   4072.   3225.   1911.   1627.    321. 282318.   2003. 17580.]
[ 2225.   6282.  10024.   8369.   4436.  12351.   2850.   1800. 230355.  6701.]
[  491.    645.   1686.   4939.  14786.   3937.    268.  15315.   7463. 242497.]
```

Test

Learning Rate: .001

Correct Answer

Predicted Answer

```
[4.7935e+04 7.0000e+00 1.9800e+02 1.4300e+02 5.0000e+01 3.5800e+02 2.8500e+02 1.6600e+02 3.3600e+02 2.5700e+02]
[0.0000e+00 5.5508e+04 2.5800e+02 5.9000e+01 2.7000e+01 1.9600e+02 1.1600e+02 2.6200e+02 5.2200e+02 1.9900e+02]
[  142.   203. 46632.  1070.   319.   369.   272.   814.   769.   162.]
[8.1000e+01 1.1600e+02 8.5800e+02 4.5477e+04 1.3000e+02 1.8630e+03 4.5000e+01 2.3200e+02 1.4500e+03 9.0700e+02]

[9.0000e+01 4.5000e+01 4.8100e+02 6.8000e+01 4.5156e+04 5.5100e+02 2.4100e+02 4.6500e+02 4.8000e+02 2.1330e+03]
[  206.   234.   326.  1382.   146. 37845.   839.   139.  2167.   772.]
[2.2800e+02 1.3000e+01 5.4100e+02 1.4400e+02 3.1600e+02 7.4800e+02 4.5628e+04 5.3000e+01 3.9100e+02 4.0000e+01]
[   70.    79.   537.   379.   190.   333.   109. 47471.   377.  1788.]
[  206.   488.  1497.  1083.   606.  1704.   285.   210. 41252.   897.]
[4.2000e+01 5.7000e+01 2.7200e+02 6.9500e+02 2.1600e+03 6.3300e+02 8.0000e+01 1.5880e+03 9.5600e+02 4.3295e+04]
```

Which are classified most accurately:

1 was guessed correctly the most

Followed by 0 and 7

Which are the most confused:

Guessed 5 when it was 9

Guessed 5 when it was 3, 8

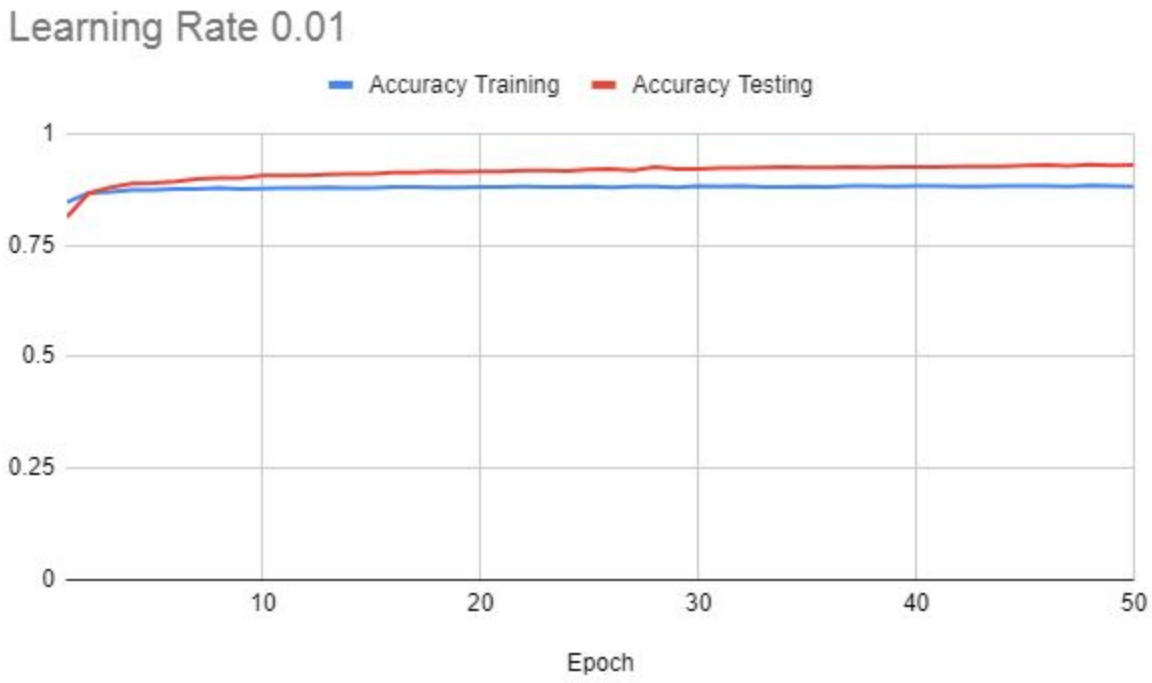
Guessed 7 when it was 9

Guessed 8 when it was 2, 5

Guessed 9 when it was 7

Learning Rate 2 (0.01):

Plot:



Any oscillations or overfitting?

There still appears to be small bumps

Confusion Matrix:

Training

Learning Rate: .01

Correct Answer

Predicted Answer

```
[2.85333e+05 6.20000e+01 2.57400e+03 1.96100e+03 6.76000e+02 3.43700e+03  2.12200e+03 9.79000e+02 2.85900e+03 1.93100e+03]
[6.50000e+01 3.20378e+05 3.35900e+03 1.81300e+03 1.48900e+03 1.25200e+03  1.13100e+03 1.26000e+03 8.03100e+03 1.51700e+03]
[  1684.   3367. 259513.   9881.   2195.   3067.   3603.   3951.   7765.  1698.]
[  1073.   1457.   7200. 259244.   1061.  14039.    655.   2572.  12579.  6366.]
[   537.    437.   3471.    846. 261520.   3837.   2265.   3858.   3873.  14618.]
[  2065.   2257.   2276.  14608.   1796. 221761.   5738.   1066.  15417.  5486.]
[2.13500e+03 3.82000e+02 4.35000e+03 1.84300e+03 2.58200e+03 5.87800e+03 2.76914e+05 3.28000e+02 2.71800e+03 2.34000e+02]
[3.89000e+02 8.51000e+02 4.13400e+03 3.07500e+03 2.09100e+03 1.45500e+03 1.94000e+02 2.81104e+05 1.98600e+03 1.77590e+04]
[  2157.   7083.   9733.   8659.   4301.  12185.   3093.   1518. 229414. 6204.]
[7.12000e+02 8.26000e+02 1.29000e+03 4.62000e+03 1.43890e+04 4.13900e+03 1.85000e+02 1.66140e+04 7.90800e+03 2.41637e+05]
```

Test

Learning Rate: .01

Correct Answer

Predicted Answer

```
[4.7961e+04 0.0000e+00 2.7000e+02 1.0900e+02 4.4000e+01 3.1700e+02  2.3000e+02 2.0000e+02 3.6000e+02 2.7200e+02]
[1.0000e+00 5.5389e+04 2.8200e+02 7.7000e+01 4.8000e+01 1.8800e+02  1.3700e+02 3.4900e+02 8.1400e+02 3.0200e+02]
[  151.   205. 46774.  1146.   311.   324.   247.   760.  1092.   189.]
[  106.   158.   699. 45525.    92.  1793.    87.   256.  1500.   758.]
[8.2000e+01 2.9000e+01 4.9900e+02 8.2000e+01 4.5424e+04 4.2700e+02  2.0000e+02 3.7500e+02 6.1100e+02 1.8690e+03]
[  134.   229.   235.  1331.   135. 38158.   826.   164.  1766.   807.]
[2.4500e+02 9.2000e+01 5.3900e+02 1.0300e+02 3.4000e+02 7.8300e+02  4.5773e+04 8.4000e+01 4.6200e+02 2.7000e+01]
[   59.    57.   609.   414.   220.   308.   123. 47381.   330.  1961.]
[  173.   555.  1411.  1089.   605.  1626.   266.   198. 40709.   881.]
[8.8000e+01 3.6000e+01 2.8200e+02 6.2400e+02 1.8810e+03 6.7600e+02  1.1000e+01 1.6330e+03 1.0560e+03 4.3384e+04]
```

Which are classified most accurately:

1 was guessed correctly the most

Followed by 0 and 7

Which are the most confused:

Guessed 3 when it was 5, 8

Guessed 4 when it was 9

Guessed 5 when it was 3, 8

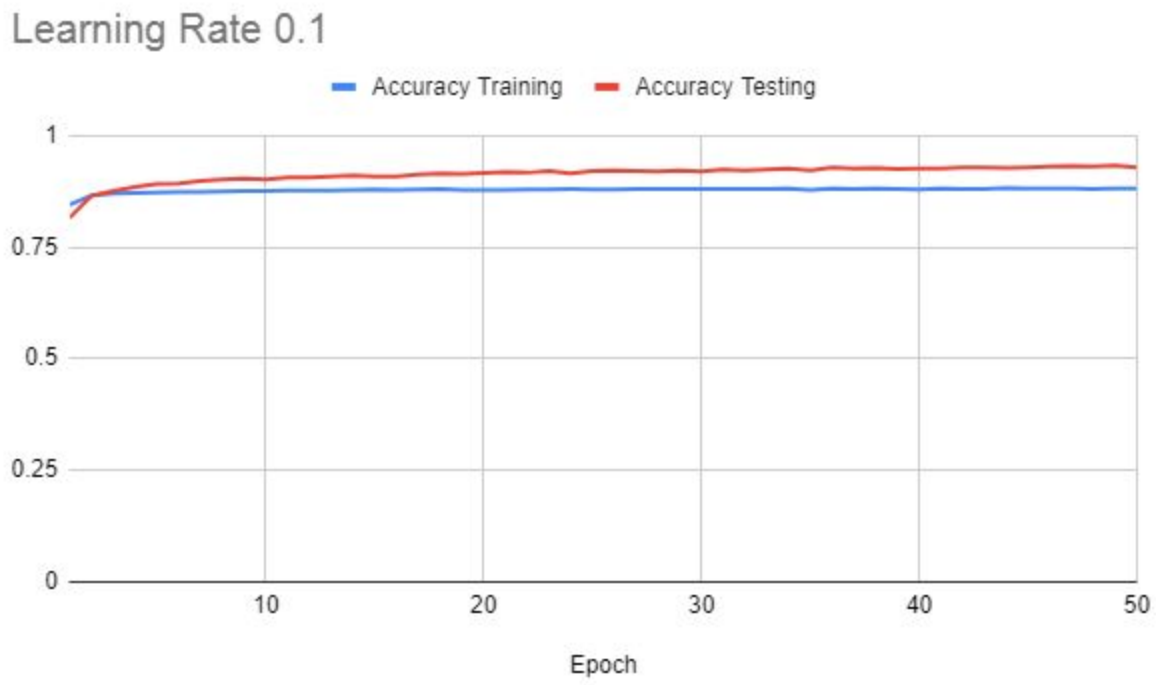
Guessed 7 when it was 9

Guessed 8 when it was 5

Guessed 9 when it was 4,7

Learning Rate 3 (0.1):

Plot:



Any oscillations or overfitting?

There continues to be small bumps

There doesn't seem to be any overfitting on any of these

Confusion Matrix:

Training

Learning Rate: .1

Correct Answer

Predicted Answer

```
[2.84605e+05 6.20000e+01 2.61800e+03 1.63400e+03 9.28000e+02 3.26100e+03 2.34800e+03 1.12700e+03 3.24800e+03 1.53900e+03]
[6.90000e+01 3.20701e+05 3.39300e+03 1.72200e+03 1.35700e+03 1.47800e+03 9.57000e+02 1.04700e+03 8.56600e+03 1.46700e+03]
[ 1799.   3174. 259299.  10665.   2335.   2817.   3455.   3996.   7648. 1498.]
[  966.   1595.   7844. 259125.   1068.  14578.    617.   2509.  12296. 6725.]
[  474.    440.   3224.    848. 259878.   4154.   2130.   3576.   3346. 15304.]
[ 2625.   1959.   2222.  14377.   1617. 220907.   5803.   1178.  14860. 5595.]
[2.30300e+03 6.27000e+02 4.26400e+03 1.82700e+03 2.72000e+03 6.03200e+03 2.77222e+05 2.29000e+02 2.67500e+03 1.58000e+02]
[  422.    976.   3763.   3078.   2284.   1404.    391. 281386.   2048. 18119.]
[ 2293.   6777.   9789.   8688.   4715.  12028.   2761.   1347. 229943.  6698.]
[5.94000e+02 7.89000e+02 1.48400e+03 4.58600e+03 1.51980e+04 4.39100e+03 2.16000e+02 1.68550e+04 7.92000e+03 2.40347e+05]
```

Test

Learning Rate: .1

Correct Answer

Predicted Answer

```
[4.8125e+04 0.0000e+00 2.9200e+02 9.2000e+01 4.1000e+01 2.5500e+02  2.0600e+02 1.6900e+02 3.9800e+02 2.5700e+02]
[0.0000e+00 5.5218e+04 3.1300e+02 9.6000e+01 4.3000e+01 2.3500e+02  7.6000e+01 2.9500e+02 8.2800e+02 2.6100e+02]
[  125.   258. 46796.  1037.   309.   353.   308.   805.   796.   194.]
[   72.   130.   619. 45616.   140.  1685.    63.   211.  1281.   816.]
[6.9000e+01 2.2000e+01 5.3700e+02 5.1000e+01 4.5142e+04 5.5700e+02 2.1700e+02 3.5000e+02 5.5200e+02 1.8650e+03]
[  136.   245.   221.  1578.   186. 38269.   822.   195.  1995.   808.]
[1.9300e+02 6.5000e+01 4.8400e+02 7.5000e+01 3.3500e+02 6.3000e+02 4.5719e+04 6.3000e+01 2.9400e+02 2.7000e+01]
[   67.    89.   583.   377.   220.   285.    65. 47383.   394.  1931.]
[  177.   684.  1384.   895.   630.  1676.   383.   147. 41140.   930.]
[3.6000e+01 3.9000e+01 3.7100e+02 6.8300e+02 2.0540e+03 6.5500e+02 4.1000e+01 1.7820e+03 1.0220e+03 4.3361e+04]
```

Which are classified most accurately:

- 1 was guessed correctly the most
- Followed by 0 and 7

Which are the most confused:

- Guessed 2 when it was 3
- Guessed 3 when it was 5, 8
- Guessed 4 when it was 9
- Guessed 5 when it was 3, 8
- Guessed 7 when it was 9
- Guessed 8 when it was 5
- Guessed 9 when it was 4,7

Compare learning rates:

Are there any differences between them:

One difference appears to be that having a higher learning rate causes the perceptrons to get more accurate more quickly, as noticed in the testing dataset's lines. Smaller datasets probably has this effect since the jump is more visible. However, looking at the confusion matrix, it appears that having a faster learning rate causes there to be more misinterpretations. I guess making more mistakes helps it learn faster

Conclusions:

Numpy made a lot of calculations faster, the code could be a bit more sleek with getting target values and calculating accuracies. (Used several if/else blocks) A lot of problems occurred because data was not properly pre-processes (shuffling data before getting labels) There was a resolved issue where a perceptron was outputting -1 instead of 0 for values less than 0. For some reason, python displayed some of the confusion matrix cells in scientific notation, but the values were still reasonable.