

## Basic Neural Network Results

### Network 1:

- Structure: [784, 50, 30, 10]
- Learning rate: 0.001

Samples: 60.000

Time taken for training: 01:06:05

Accuracy: **33.21%**

Average Loss: 0.83

Samples: 100.000

Time taken for training: 02:03:25

Accuracy: **45.65%**

Average Loss: 0.74

Samples: 200.000

Time taken for training: 4:29:16

Accuracy: **62.40%**

Average Loss: 0.56

Samples: 400.000

Time taken for training: 9:14:52

Accuracy: **76.50%**

Average Loss: 0.39

Samples: 600.000

Time taken for training: 14:02:33

Accuracy: **80.92%**

Average Loss: 0.31

Samples: 840.000

Time taken for training: 19:46:08

Accuracy: **83.78%**

Average Loss: 0.26

## Network 2:

- Structure: [784, 85, 10]
- Learning rate: 0.001

Samples: 60.000

Time taken for training: 01:59:19

Accuracy: **38.62%**

Average Loss: 0.80

Samples: 100.000

Time taken for training: 03:46:05

Accuracy: **51.05%**

Average Loss: 0.65

Samples: 200.000

Time taken for training: 08:05:32

Accuracy: **61.35%**

Average Loss: 0.52

Samples: 400.000

Time taken for training: 15:59:45

Accuracy: **72.92%**

Average Loss: 0.40

Samples: 600.000

Time taken for training: 23:53:15

Accuracy: **84.35%**

Average Loss: 0.27

Samples: 840.000

Time taken for training: 33:23:06

Accuracy: **86.27%**

Average Loss: 0.23

### Network 3:

- Structure: [784, 27, 27, 27, 10]
- Learning rate: 0.001

Samples: 60.000

Time taken for training: 00:37:49

Accuracy: **28.19%**

Average Loss: 0.86

Samples: 100.000

Time taken for training: 01:10:08

Accuracy: **33.89%**

Average Loss: 0.82

Samples: 200.000

Time taken for training: 02:34:12

Accuracy: **44.79%**

Average Loss: 0.69

Samples: 400.000

Time taken for training: 05:23:46

Accuracy: **65.38%**

Average Loss: 0.52

Samples: 600.000

Time taken for training: 08:09:35

Accuracy: **77.87%**

Average Loss: 0.38

Samples: 840.000

Time taken for training: 11:28:06

Accuracy: **81.24%**

Average Loss: 0.30

## Improved Neural Network Results

### Network 1:

- Structure: [784, 85, 10]
- Learning rate: 0.001

Samples: 10.000

Time taken for training: 00:18:33

Accuracy: **83.88%**

Average Loss: 0.607282

Samples: 20.000

Time taken for training: 00:42:22

Accuracy: **87.27%**

Average Loss: 0.471222

Samples: 50.000

Time taken for training: 01:54:22

Accuracy: **89.82%**

Average Loss: 0.360597

Samples: 100.000

Time taken for training: 03:55:30

Accuracy: **91.31%**

Average Loss: 0.301386

Samples: 200.000

Time taken for training: 07:58:24

Accuracy: **92.29%**

Average Loss: 0.262244

Samples: 300.000

Time taken for training: 12:07:41

Accuracy: **93.15%**

Average Loss: 0.236050

Note: although the fairly good hit rate above 90% had been somewhat increased, its growth is too little for the amount of time taken. However, the average loss also fell which may suggest that the network could even go farther and reach better results.

## Network 2:

- Structure: [784, 85, 10]
- Learning rate: 0.01

Samples: 1.000

Time taken for training: 00:01:56

Accuracy: **71.18%**

Average Loss: 0.878569

Samples: 5.000

Time taken for training: 00:12:03

Accuracy: **87.01%**

Average Loss: 0.451212

Samples: 10.000

Time taken for training: 00:25:10

Accuracy: **89.26%**

Average Loss: 0.377774

Samples: 50.000

Time taken for training: 02:14:25

Accuracy: **93.04%**

Average Loss: 0.256567

Note: Here, an experiment was made with a very high learning rate to see if it could reach great results with less time and training samples. Within just 2 hours it reached the mark of 93% of accuracy and with just 0.25 average loss.

### Network 3:

- Structure: [784, 50, 30, 10]
- Learning rate: 0.001

Samples: 10.000

Time taken for training: 00:11:46

Accuracy: **81.64%**

Average Loss: 0.625300

Samples: 20.000

Time taken for training: 00:26:34

Accuracy: **86.05%**

Average Loss: 0.474921

Samples: 50.000

Time taken for training: 01:11:04

Accuracy: **88.95%**

Average Loss: 0.383299

Samples: 100.000

Time taken for training: 02:25:39

Accuracy: **90.80%**

Average Loss: 0.324978

Samples: 160.000

Time taken for training: 03:55:27

Accuracy: **91.39%**

Average Loss: 0.315646

Samples: 240.000

Time taken for training: 05:55:14

Accuracy: **90.30%**

Average Loss: 0.400130

Samples: 300.000

Time taken for training: 07:25:03

Accuracy: **90.31%**

Average Loss: 0.411509

Note: It seems that we reach a convergence here around 90% and that, the tested values above 160.000 samples started overfitting - since the average loss increased.

#### Network 4:

- Structure: [784, 27, 27, 27, 10]
- Learning rate: 0.001

Samples: 10.000

Time taken for training: 00:06:59

Accuracy: **76.30%**

Average Loss: 0.811425

Samples: 50.000

Time taken for training: 00:42:42

Accuracy: **87.28%**

Average Loss: 0.470756

Samples: 100.000

Time taken for training: 01:28:17

Accuracy: **89.61%**

Average Loss: 0.385539

Samples: 200.000

Time taken for training: 02:57:47

Accuracy: **90.30%**

Average Loss: 0.411521

Samples: 300.000

Time taken for training: 04:24:28

Accuracy: **89.53%**

Average Loss: 0.501933

Samples: 420.000

Time taken for training: 06:08:17

Accuracy: **88.88%**

Average Loss: 0.604547

Note: here we also can see a convergence near 90% of accuracy. It's very interesting to see that the model *clearly* dropped his performance with 300.000 and above trained samples.

.

### Network 5:

- Structure: [784, 27, 27, 27, 10]
- Learning rate: 0.0001

Samples: 50.000

Time taken for training: 0:41:29

Accuracy: **68.72%**

Average Loss: 1.110818

Samples: 100.000

Time taken for training: 01:23:57

Accuracy: **76.84%**

Average Loss: 0.740867

Samples: 160.000

Time taken for training: 02:14:51

Accuracy: **81.15%**

Average Loss: 0.585183

Samples: 240.000

Time taken for training: 03:23:10

Accuracy: **84.96%**

Average Loss: 0.500420

Samples: 300.000

Time taken for training: 04:16:08

Accuracy: **86.63%**

Average Loss: 0.461618

Samples: 350.000

Time taken for training: 04:50:06

Accuracy: **87.61%**

Average Loss: 0.441047

Samples: 400.000

Time taken for training: 05:31:30

Accuracy: **88.19%**

Average Loss: 0.423309

Samples: 480.000

Time taken for training: 06:37:07

Accuracy: **88.64%**

Average Loss: 0.412620



### Network 6:

- Structure: [784, 50, 30, 10]
- Learning rate: 0.0001

Samples: 50.000

Time taken for training: 01:00:40

Accuracy: **73.82%**

Average Loss: 1.002825

Samples: 100.000

Time taken for training: 02:19:14

Accuracy: **80.63%**

Average Loss: 0.679039

Samples: 160.000

Time taken for training: 03:49:25

Accuracy: **83.20%**

Average Loss: 0.573466

Samples: 240.000

Time taken for training: 05:58:19

Accuracy: **85.27%**

Average Loss: 0.519348

Samples: 300.000

Time taken for training: 07:26:57

Accuracy: **86.70%**

Average Loss: 0.480036

Note: Although the accuracy seemed to be growing, after 7 hours of training and the current results, I thought it wasn't worthy to keep training it. Anyways, the binary file with trained network is available in my GitHub repository.

by Felipe Colombelli

[@colombelli /MachineLearning](#)