

I implemented the reverse-mode autodiff and a fully connected neural network. I experimented with different architectures and below is the results. The training happened for at most 100 epochs. I stopped the training for some of them earlier when no further improvement was observed for validation loss.

The first architecture (one hidden layer with one neuron) performed worst among all and had the highest amount of underfitting. As the complexity of the models increased, the models performed better. The last three models performed best, such as the network with two hidden layers and 20 neurons per layer. Furthermore, I continued training of this model for 100 epochs, however it did not seem to overfit. In other words, the training and validation loss reached a certain amount and continued to stay approximately the same after 100 epochs.

Network Architecture	Train Acc.	Validation Acc.	Test Acc.
One hidden layer with 1 neuron	10.40	10.00	10.40
One hidden layer with 10 neurons	78.20	66.00	62.40
One hidden layer with 15 neurons	87.20	80.80	78.40
One hidden layer with 20 neurons	96.60	80.00	79.20
Two hidden layers with 10 neurons in each layer	98.60	79.60	73.60
Two hidden layers with 20 neurons in each layer	99.60	78.80	77.60

The confusion matrix and the table below are obtained after training the best-performing model. We can observe here that “class 1” has the lowest amount of incorrect classifications (2%) and after that “class 0” has the second-lowest amount (7%). Also, we can observe that “class 2” and “class 3” have the highest amount of incorrect classifications (37%). Furthermore, we can see that “class 3” got confused with “class 5” four times. Also, “class 4” got confused with “class 9” and “class 7” got confused with “class 9” four times.

	0	1	2	3	4	5	6	7	8	9
0	<24>	1	.	.	1	.
1	.	<36>	.	.	.	1
2	2	.	<12>	1	1	.	1	1	1	.
3	1	.	.	<12>	.	4	.	1	1	.
4	.	1	1	.	<19>	4
5	2	.	.	1	1	<14>	1	.	.	1
6	2	<22>	.	.
7	1	.	2	.	1	.	.	<23>	1	4
8	.	1	.	3	.	1	1	.	<13>	.
9	1	.	.	2	.	<26>

(row = actual; col = predicted)

Class	Number of Incorrect Classifications	Percentage of Incorrect Classifications
0	2	7%
1	1	2%
2	7	37%
3	7	37%
4	6	24%
5	5	26%
6	2	8%
7	9	28%
8	6	32%
9	3	10%