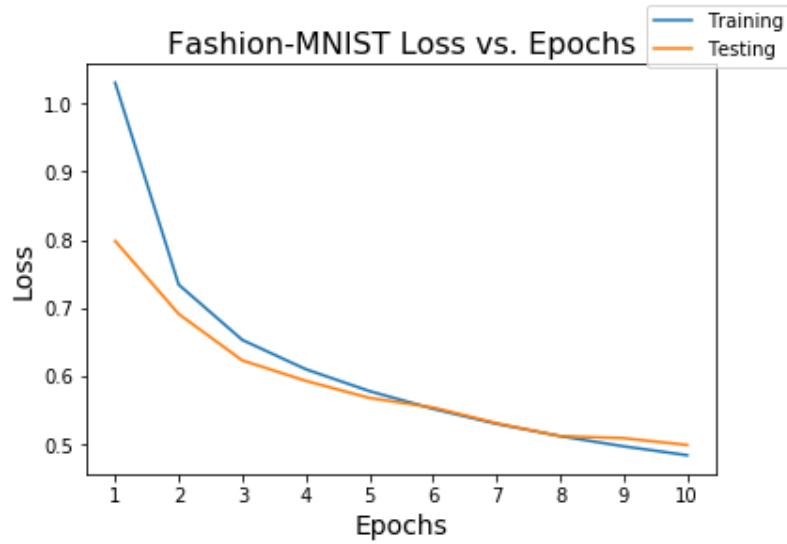


Stuart Harley

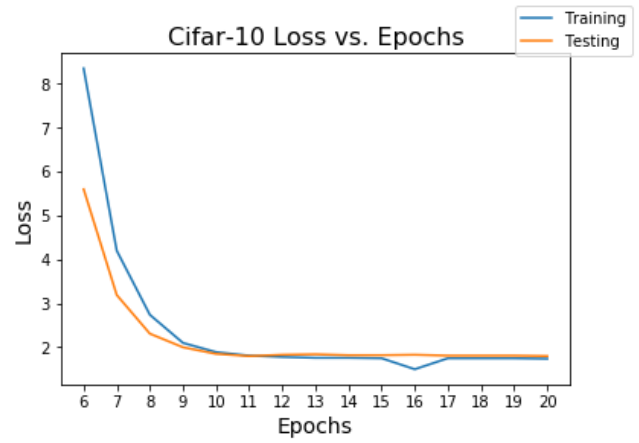
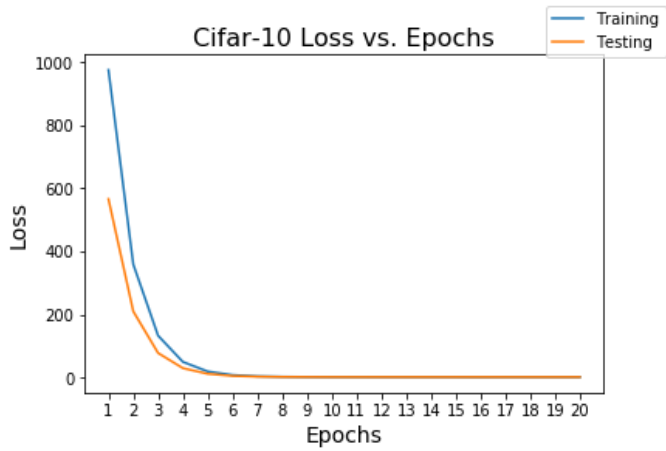
Deep Learning Lab 7: Completing the From-Scratch Library Report

Fashion-MNSIT dataset:



Epoch	Training Accuracy	Testing Accuracy	Training Loss	Testing Loss
1	65.15%	68.82%	1.03	.798
2	73.31%	73.69%	.734	.691
3	76.54%	77.99%	.653	.623
4	78.69%	79.58%	.610	.593
5	79.95%	80.76%	.578	.568
6	80.85%	81.39%	.552	.554
7	81.72%	81.79%	.530	.531
8	82.34%	82.24%	.512	.512
9	82.89%	82.64%	.497	.509
10	83.30%	82.89%	.484	.499

CIFAR-10 dataset:



Epoch	Training Accuracy	Testing Accuracy	Training Loss	Testing Loss
1	16.64%	20.23%	974.50	566.63
2	22.51%	23.70%	358.87	209.66
3	26.25%	27.70%	133.22	78.32
4	30.22%	32.86%	50.19	29.96
5	34.23%	36.76%	19.61	12.15
6	36.90%	38.76%	8.35	5.60
7	38.36%	39.22%	4.20	3.19
8	39.64%	38.94%	2.67	2.31
9	40.65%	38.89%	2.10	2.00
10	41.28%	40.39%	1.89	1.85
11	41.77%	40.91%	1.81	1.80
12	42.14%	39.40%	1.78	1.83
13	42.53%	39.57%	1.76	1.84
14	42.80%	39.87%	1.76	1.82
15	42.96%	40.59%	1.75	1.82
16	43.23%	40.47%	1.75	1.83
17	43.46%	41.02%	1.75	1.81
18	43.45%	41.06%	1.75	1.81
19	43.69%	41.11%	1.75	1.81
20	43.74%	41.49%	1.74	1.80

Conclusion: This lab was more difficult than past labs. Deriving the backpropagation equations was straightforward enough since we worked on doing that in class a lot. Writing the unit tests were also not bad. You just had to do calculate a simple example by hand and make sure it was all working. Since we did all of the testing of the layers before implementing a network, we were able to be confident that they were working correctly.

The Fashion-MNIST dataset was not too difficult to create a good network for. Also, I found the model trained better without the regularization layers, so I just used 2 linear layers with a ReLU in between and a Softmax + Cross Entropy at the end. I chose a learning rate of .01 which seemed to work well. After only 10 epochs the accuracy was above 80% and the training curves were still going down. Therefore, if I had trained it for longer, I could have most likely gotten better accuracies.

The CIFAR-10 dataset was much harder to create a good model for. I used a learning rate of .001 because when I tried .01 it was too big. And I chose a middle layer of 100 nodes based on the recommendation from the professor. However, it still did not improve accuracy quickly. I ran it for 20 epochs and ended with an accuracy of only ~42%. This took significantly longer to run than the Fashion-MNIST model since there were more weights to learn. I also had to implement the stable version of softmax & cross entropy for this network. I was getting inf and nan errors before doing that. It was easy to test that the edited softmax & cross entropy layer still worked after implementing it because we had already made unit tests for that which was nice. The accuracy curves were still going down (slowly), and if I were to continue training and/or lower the learning rate after some # of epochs, and/or use a different network architecture, I likely would have been able to achieve a better accuracy.

One thing that was annoying about this lab was how long it took to train the networks. The Fashion-MNIST was not bad, taking less than 5 min for 10 epochs. But the CIFAR-10 took a couple hours. Therefore, figuring out good hyperparameters for the networks was difficult because you had to wait for results to start to come in before you could try another value(s).