

Homework 6

GitHub: https://github.com/pballou/ECGR_4105/tree/master/Homework/homework_6

1. Fully connected networks for all ten classes of CIFAR-10, all training conducted over 300 epochs.

- a. One hidden layer of size 512.

Parameters:

Activation function: Tanh

Learning rate: .002

Optimizer: SGD

Batch size: 64

Results:

Final training loss: 1.147

Validation accuracy: .49

Run time: ~60 minutes

Comments: Not very good overall. Only using one hidden layer definitely limits the learning capability of the network. I could have improved it more, but because training took so long, I was unable to try all the configurations I wanted to.

- b. Three hidden layers with sizes: 1028, 256, 128.

Parameters:

Activation function: LeakyReLU

Learning rate: .002

Optimizer: SGD

Batch size: 64

Results:

Final training loss: .001599

Validation accuracy: .54

Run time: ~60 minutes

Comments: Much better than the model in part a – training loss significantly improved. Unfortunately, there was a bit of overfitting as the validation set only got .54 accuracy. As with part a, I could have improved this model a lot more, but time greatly limited what parameter combinations I was able to test.

2. Convolutional neural networks for all ten classes of CIFAR-10, all training conducted over 300 epochs.

- a. Two hidden layers.

Parameters:

Activation function: Tanh

Learning rate: .008
Optimizer: SGD
Batch size: 128

Results:

Final training loss: .628
Validation accuracy: .66
Run time: ~75 minutes

Comments: This model had the best validation accuracy, which is interesting since its training loss was still .62 – not nearly as good as 1b. This means that this model did not overfit as much, but .66 is still not great accuracy. I think, I would need more time to experiment with the parameters and train for more epochs.

b. Three hidden layers.

Parameters:

Activation function: LeakyReLU, Tanh
Learning rate: .02
Optimizer: SGD
Batch size: 256

Results:

Final training loss: .566
Validation accuracy: .60
Run time: ~75 minutes

Comments: This model had less training loss than the simpler model in part a, but had a worse validation accuracy, which likely means there was overfitting. I think the main lesson learned is you really need more time or a better GPU in order to be able to train for more epochs and test a variety of parameters. Overall, the CNNs performed than the fully connected neural nets.