IE 534 - Homework 3

Jacob Heglund - jheglun2

September 28, 2018

1 Instructions

HW3: Train a deep convolution network on a GPU with PyTorch for the CIFAR10 dataset. The convolution network should use (A) dropout, (B) trained with RMSprop or ADAM, and (C) data augmentation. For 10 percent extra credit, compare dropout test accuracy (i) using the heuristic prediction rule and (ii) Monte Carlo simulation. For full credit, the model should achieve 80-90 percent Test Accuracy. Submit via Compass (1) the code and (2) a paragraph (in a PDF document) which reports the results and briefly describes the model architecture. Due September 28 at 5:00 PM.

2 Implementation

The network uses the architecture suggested in the lecture notes, with 8 convolutional layers, and 3 fully connected layers. It doesn't seem to want to train effectively though, and the network as implemented achieves around 60 percent accuracy after 50 epochs. In order to achieve higher accuracy, I would mess around with a learning rate schedule and fine-tune other hyperparameters.

I spent 2 days building a data-loader from scratch. I then realized that the built-in Pytorch data-loader was way better, and could do shuffling and data-augmentation trivially. For my data augmentation, I used the Pytorch transform module to apply random horizontal and vertical flipping, as well as a random crop that pads the image. It also normalizes the images from some values I found here: https://github.com/kuangliu/pytorch-cifar/issues/19.

Getting this thing to actually train was a wonder in itself, and I spent two days combing through every line of code to see why it wouldn't train. Accuracy stayed at 10 percent until a miracle occurred; I realized that the Pytorch implementation of ADAM requires a learning rate to be specified, and if unspecified, will be set to 0.