

Homework 08

1. Run the Pytorch notebook, which demonstrates how to set up and train a convolutional neural network on Fashion MNIST. Examine the Network class, which defines the architecture of the network.
 - a. Modify the architecture to eliminate the convolutional layers. The main adjustment is in the sizes of the networks and features. How much is performance affected by having only fully connected layers? How many weights are learned in the original network, and the one without convolutions?
 - b. Return to the original architecture, and eliminate the first fully connected layer. How much does this affect performance?
 - c. Perform any other modifications of interest, such as widening the fully connected layers, or deepening the network. (It is generally fairly easy to improve upon the performance of the demo network.) You can also optionally change the learning rate, batch size, and other hyper-parameters. Briefly comment on what you discover.
2. For the same Fashion MNIST demo, next examine the use of Adam and batch normalization.
 - a. Try changing the Adam learning algorithm to the more traditional stochastic gradient descent with momentum. Modify the parameter grid to vary momentum, and demonstrate a good value for momentum (between 0 and 1). It is not necessary to repeat the other hyper-parameter variations included in the demo.
 - b. Briefly explain why batch normalization may be helpful in neural networks.
 - c. (Advanced, optional:) Modify the network to apply appropriate batch normalization to each layer, using either SGD or Adam, as well as a good momentum from your experiments. Do you find batch normalization helpful for this problem?