Problem Set

December 2021

1. Implement a three-layer FC MLP for FashionMNIST. Use Adam and an initial learning rate of 1e-2, batchsize of 256, and 20 epochs. The layer output sizes should be 100, 50, 10. The activation should be ReLU.

Create a validation set. For regularization, we will compare weight decay and a basic version of the WEISSI regularizer. Set the weight decay value to $\lambda = 1e-5$. For WEISSI, use $\lambda_{we} = 1e-6$ and $\lambda_{wc} = 1e-5$ to start. You may need to modify the loss function in order to add regularization.

Vary the values of λs for both models (try to change the scale of λs). Graph the accuracy with respect to the choices of λs for each of the models. What observations do you have about the sensitivity of the models to the choices of λs?

We will now look at how the number of neurons in each layer affects the output. Vary the output sizes of the first and second layers so that the ratio of the first output and the second output is $1, 2, 4, 8, 16$ and graph the resulting accuracy for each model. Is there a significant difference in the models?