

CS 410/510: Deep Learning: Computational Structures and Programming
Winter 2022
Assignment 2
Due 11:59pm, Feb 6

You may use any framework to do this assignment including Pytorch, Keras, Tensorflow, or others. As usual you submit your code (as a link in colab, as a github repository, as a gzipped file) plus a document explaining your results (see below).

The goal of this assignment is to gain some experience with CNNs. You will use the CIFAR-10 dataset for all the experiments.

1. Construct a simple CNN that follows the architecture of LeNet. However, the images have 3 channels. The first kernel will be 6 5x5x3 kernels that takes us to 28x28x6. The remainder of the network is the same. Stride is 1 for convolution and 2 for pooling layers. You need to train the network for different choices of parameters: learning rate 0.1, 0.01, 0.001; activation function = sigmoid, Tanh; loss function = MSE, cross-entropy (3x2x2 = 12 combinations). Train for a sufficient number of epochs to see a low train loss.
 - a. *Submit plots of training error vs epoch and test error vs epoch for each combination of parameters. Describe the reasons for what you observe*
 - b. *For any one trained network and parameter choice, display feature maps of 10 images at the last convolution layer. What has your network learned? Here is one helpful tutorial: <https://machinelearningmastery.com/how-to-visualize-filters-and-feature-maps-in-convolutional-neural-networks/>*
2. Select Relu, learning rate of 0.001, cross-entropy loss. Change the network to use 3x3 kernels for both convolutions rather than 5x5. Same pooling. Train for a sufficient number of epochs to see a low enough train loss.
 - a. *Plot the error (test, train) vs epoch and discuss how your results differ from the previous network*
3. Build a CNN with 5 convolution layers with 3x3 kernels and corresponding 2x2 average pooling with stride 1. Use *padding* at each convolution layer so that the size of the output matches the input after each convolution layer. Train the network for any choice of model parameters.
 - a. *How long did it take to train? Why?*
 - b. *How does the accuracy of this network compare against the previous two? Explain*

Submit a pdf of your discussion.