Assigned: 28 October 2021

Homework #4 – MLP Fundamentals

EE 541: Fall 2021

Due: Saturday, 06 November 2021 at 23:59. Late penalty: 10% per 24-hours before 08 November at 23:59. Submission instructions will follow separately on canvas.

1. Train an MLP for Fashion MNIST. Use one hidden layer with 100 nodes and ReLU activations. Use a dropout layer for the hidden layer output with 30% dropout rate and use an L2 regularizer with coefficient $\lambda = 0.0001$. You may use the demo scripts from lecture and discussion.

   Evaluate this trained model on the test data and compute a "confusion matrix" – i.e., find the rate of classification rate conditioned on the true class. Element (i, j) of the confusion matrix is the rate at which the network decides class j when class i is the correct label (ground truth).

   a. Use the seaborn python package or like generate a heatmap showing the confusion matrix.

   b. Consider class m. List the class most likely confused for class m for $m \in \{0, \ldots, 9\}$.

   c. Which two classes (types of clothing) are most likely to be confused overall?

2. For this problem, you will train two models on Fashion MNIST:

   1. One hidden layer with ReLU activation, 128 nodes. No regularization, no dropout.

   2. One hidden layer with ReLU activation, 48 nodes. L2 regularization with coefficient $\lambda = 0.0001$ and dropout with rate 0.2 at the hidden layer.

   Train for 40 epochs. You may use the demo scripts from lecture and discussion.

   Produce histograms for the weights of these two networks -- a separate histogram for the two layers (the input and hidden layers) in each case. Describe the qualitative differences between these histograms.