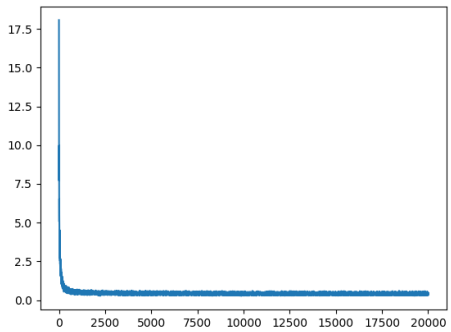
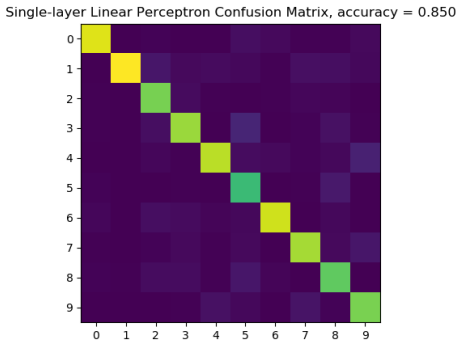
HW4 report

Dengyuan Wang

1. Single layer perceptron with linear activation function: Accuracy = 85.0%

Learning rate:0.05, Decay: 0.9 per 1K batches; Iteration: 20000

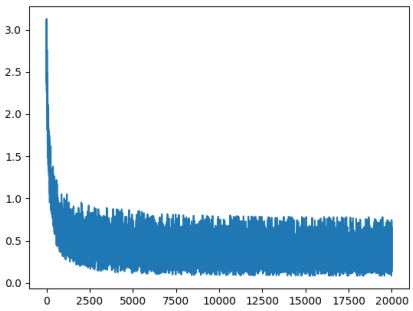


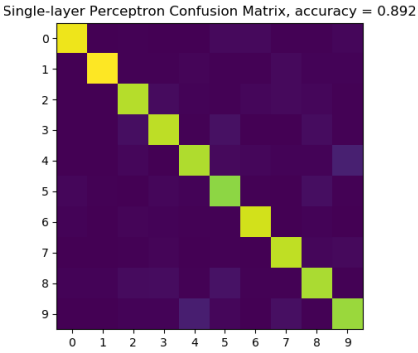


2. Single layer perceptron with RELU activation function:Accuracy = 89.2%

Learning rate:0.05, Decay: 0.9 per 1K batches; Iteration: 20000

Used Leak RELU

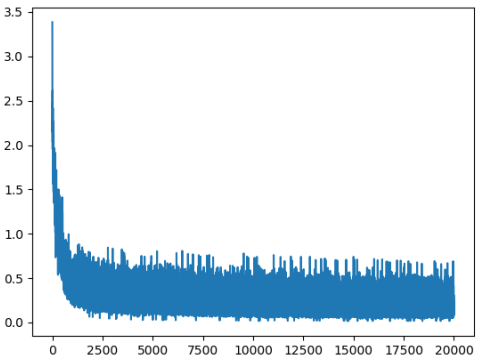


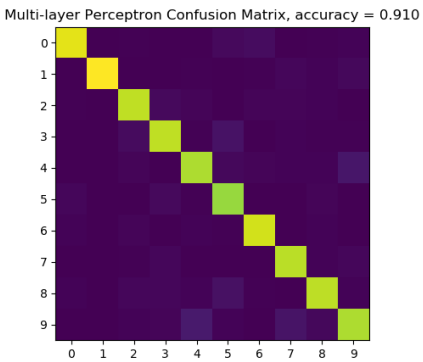


3. Two layers perceptron with RELU activation function:Accuracy = 91.0%

Learning rate:0.05, Decay: 0.9 per 1K batches; Iteration: 20000

Used Leak RELU



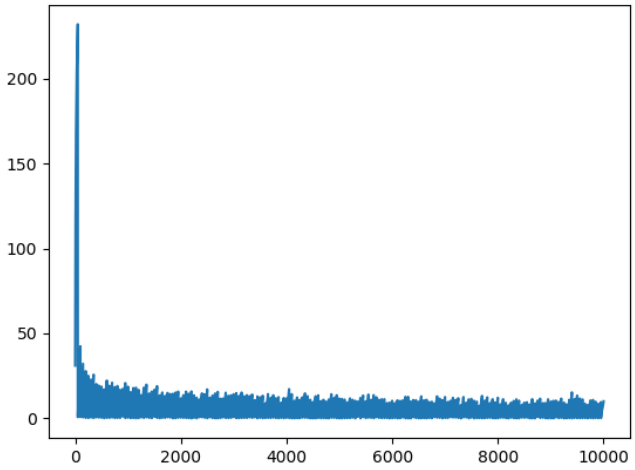


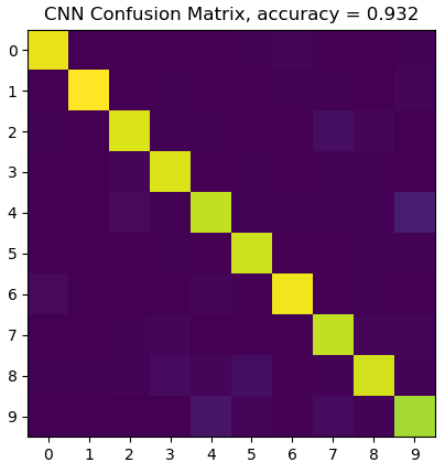
4. CNN+FC: Accuracy = 93.2%

Use Adam Optimization method: learning rate = 0.1, beta1 = 0.95, beta2=0.99

v_w = beta1*v_w + (1-beta1)*dl_dw/batchsize
s_w = beta2*s_w + (1-beta2)*(dl_dw/batchsize)**2
w -= lr* v_w/np.sqrt(s_w+1e-7)
Decay: 0.9 per 1K iteration
Iteration: 10000

Used RELU





X-axis is iteration, Y-axis is loss per batch

For Function implementation details, please refer to the comments in file cnn.py