

CS 598 Homework 1

Mohan Sun*

We implement a fully connected one-hidden-layer neural network and train the model with MNIST data. Figure 1 shows training results and timing. Our model have an accuracy of 97.97% which satisfies the requirement of the assignment.

Model the model is a modified and extended version of Logisticregression.py. We keep the i/o and data processing part, keep the structure of Epoch loop, mini-batch loop, and forward as well as backward as function. We modified forward to:

```
def forward(x,y, model):  
    Z = model['W'] @ x + model['b1']  
    H = sigmoid(Z)  
    U = model['C'] @ H + model['b2']  
    p = softmax_function(U)  
    return Z, H, p
```

*Even if this document use "we" and "our", the homework is done by only Mohan "Fred" Sun without any collaboration. Netid: mohans2

and backward to:

```
def backward(x, y, Z, H, p, model, model_grads):  
    dU = np.copy(p)  
    dU[y] = dU[y] - 1  
    dC = dU @ H.T  
    delta = model['C'].T @ dU  
    db1 = delta * sigmoidp(Z)  
    dW = db1 @ x.T  
    model_grads['W']=dW  
    model_grads['b1']=db1  
    model_grads['b2']=dU  
    model_grads['C']=dC  
    return model_grads
```

This basically reflects the model used in lecture notes. The only difference here is that we use sigmoid instead of ReLU to avoid overflow issue with exponentials in softmax.

parameters For this model, we set $D_h = 100$ and $LR^{(0)} = 0.1$. LR is scaled down by 10 for every 5 epochs.

```
Test accuracy: 0.9797000000
→ CS598 python CS598_Mohan_Sun_HW1.py
No. 0 epoch accuracy: 0.9300833333
No. 1 epoch accuracy: 0.9685833333
No. 2 epoch accuracy: 0.9753666667
No. 3 epoch accuracy: 0.9798500000
No. 4 epoch accuracy: 0.9806333333
No. 5 epoch accuracy: 0.9839166667
No. 6 epoch accuracy: 0.9920000000
No. 7 epoch accuracy: 0.9948666667
No. 8 epoch accuracy: 0.9952666667
No. 9 epoch accuracy: 0.9961333333
No. 10 epoch accuracy: 0.9963500000
No. 11 epoch accuracy: 0.9968333333
No. 12 epoch accuracy: 0.9972833333
No. 13 epoch accuracy: 0.9967000000
No. 14 epoch accuracy: 0.9974500000
No. 15 epoch accuracy: 0.9972500000
No. 16 epoch accuracy: 0.9973333333
No. 17 epoch accuracy: 0.9971500000
No. 18 epoch accuracy: 0.9975166667
No. 19 epoch accuracy: 0.9973500000
Train took 527.1324031353 seconds
Test accuracy: 0.9797000000
```

Figure 1: **Neural Network Training Result and Timing.** Our model have a test accuracy of 97.97% in the last run. Training takes approximately 527 seconds.