EECS260: Homework 2 Write Up

Audrey Der (861221280)

Due: May 16, 2020

1 Introduction

This assignment asks us to implement a linear classifier and train a one layer NN using two different loss functions and SGD, normalizing the data as appropriate. We concatenate 1s to act as a bias term and initialize $\boldsymbol{W}$ and $\boldsymbol{V}$ to be identically and independently drawn from the Gaussian distribution with mean 0 and variances $\frac{1}{k}$ and $\frac{1}{100}$, respectively.

2 Linear Classifier

Trained on minibatch SGD, batch size 100, $\eta = 10^{-12}$, run for 1000 iterations. As a baseline, it achieves approximately a 23.5% misclassification rate, or a 75.5% accuracy rate. See Figure 1.

3 Training the Neural Net

A neural net was implemented using two different final transformations applied to $f(\boldsymbol{x}) = \boldsymbol{v}^T \text{ReLU}(\boldsymbol{W}\boldsymbol{x})$. This is under $\eta = 10^{-4.175}$. See Figure 2.

- Quadratic Loss, $\sigma(x) = x$
- Logistic Loss, $\sigma(x) = \frac{1}{1+e^{-x}}$

4 Results and Conclusions

Based on the graphs we can see that for some fixed learning rate, a wider network will produce better results. Logistic loss seemed to result in better performance, resulting in smoother curves, particularly for the wider networks. The wider NNs performed better than the linear classifier under the specified conditions.

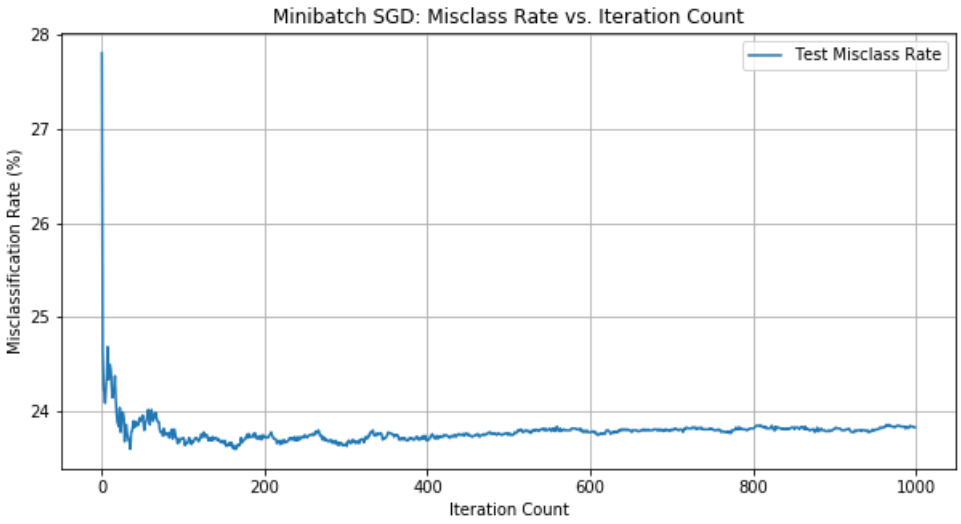

Figure 1: Linear Classifier

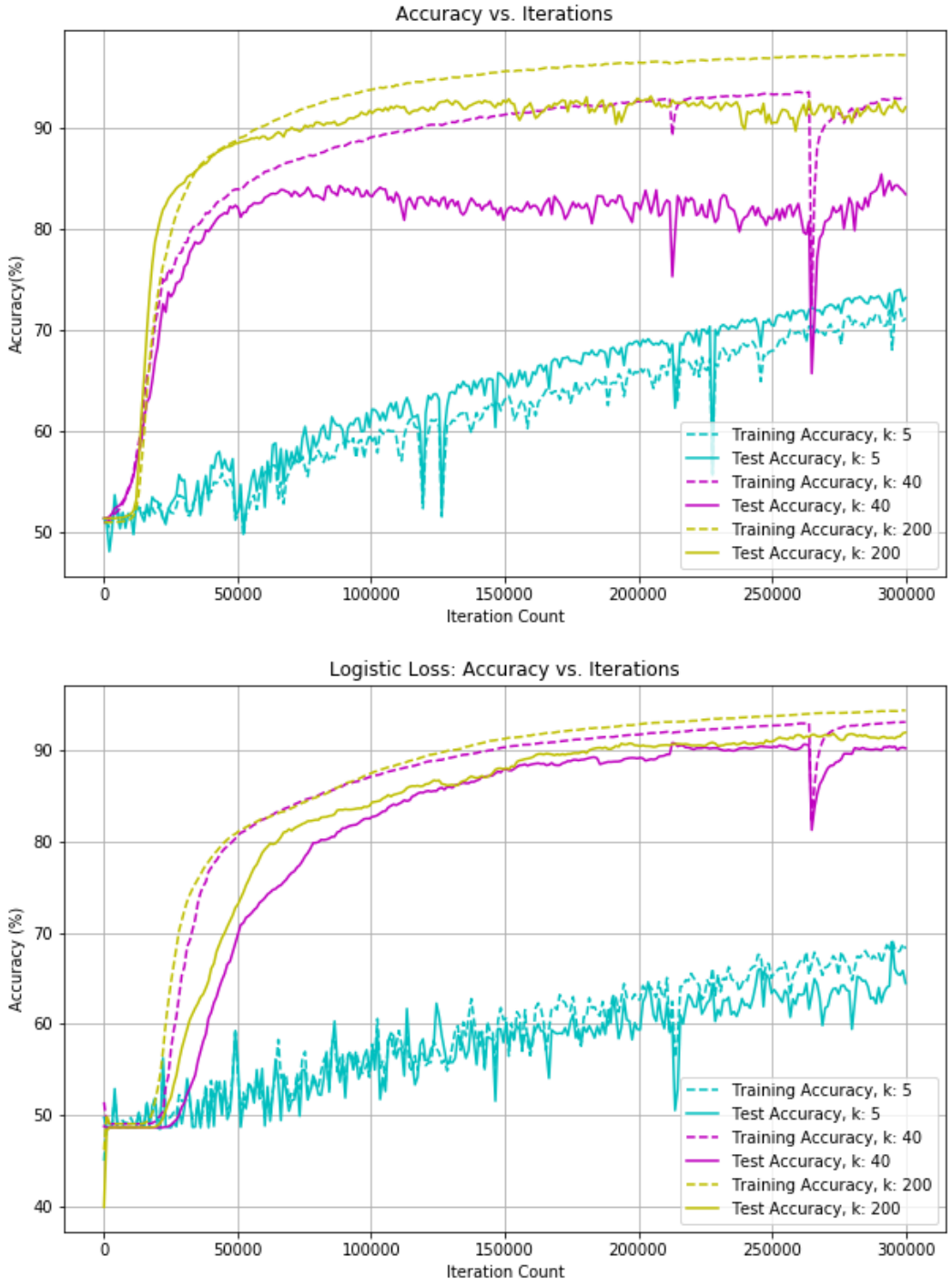

Figure 2: Accuracy vs. Iteration Count