Deep Learning in Data Science
dd2424

Report to Assignment 1 (for Bonus)

Yiming Fan
yimingf@kth.se

1 Introduction

This assignment aims at training a multiple linear one-layer network using *gradient descent* method. The dataset used in this assignment is CIFAR-10.

2 Methods & Mechanisms

We tried mainly three methods on improving the training accuracy:

1. All training data from 5 batches are used for training (49000 out of 50000) and validation (1000 out of 50000).

2. We trained the network with long enough (50) epochs. This is a reasonable boundary before overfitting.

3. We found a set of (not perfect but) sound parameters and they are: $\eta = 0.02$ (and decay by 0.9 after every 10 epochs), $\lambda = 0.01$, 500 batches (each contains 98 data) and 50 epochs.

3 Results

The test accuracy is around 44% and the highest accuracy achieved is 44.07%.

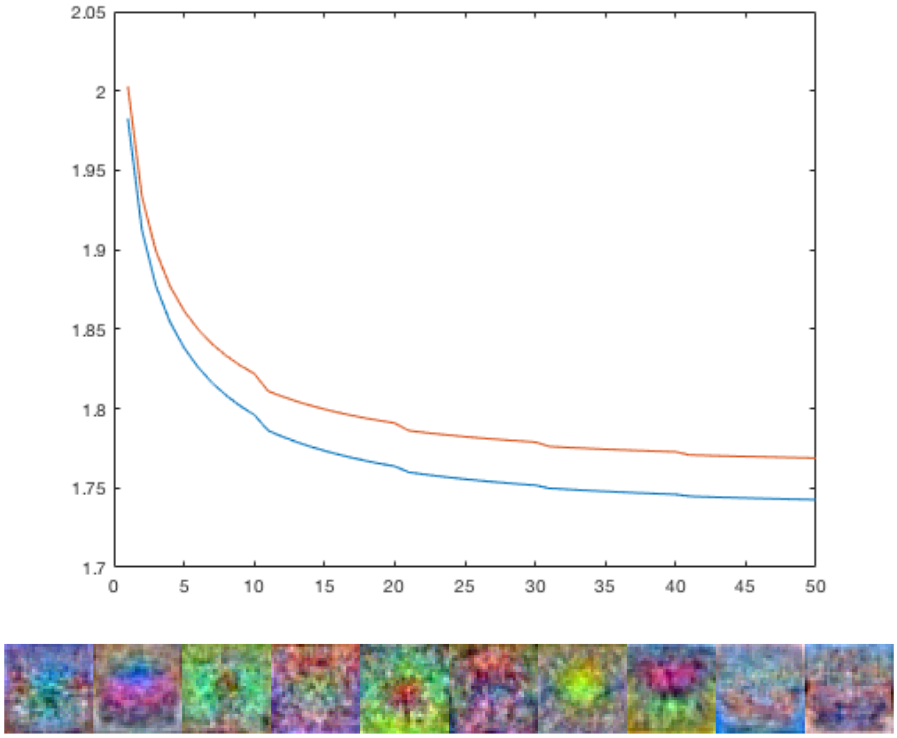

Figure 1: Result figure and the learnt matrix. The test accuracy is around 44%.