

CS535 Deep Learning

Homework 2:

CIFAR 10 Image Classification using Fully Connected Neural Network

Yathartha Tuladhar
February 19th, 2018

For **Part 1** and **Part 2**, please refer to the implementation in code.

Part 3:

The learning rate, momentum constant, batch size, and number of hidden units that were giving me good results are shown in table 1.

Table 1. Hyperparameters

| Learning Rate | Momentum | Batch size | Hidden units | Epochs |
|---------------|----------|------------|--------------|--------|
| 0.001 | 0.6 | 16 | 50 | 100 |

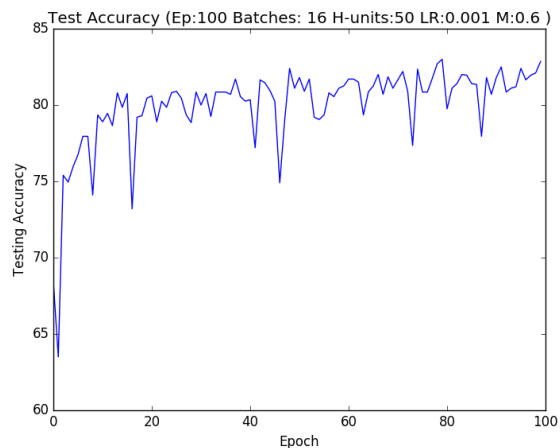
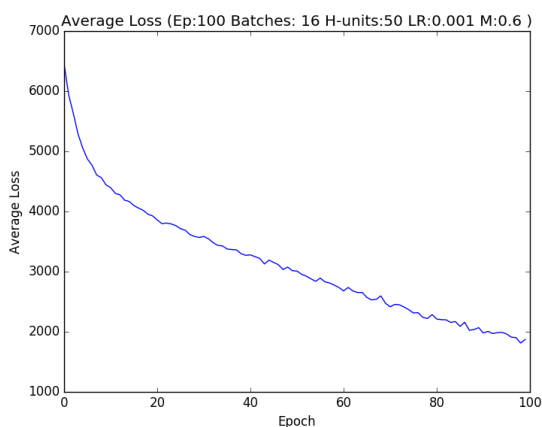
The maximum training accuracy was about 93%

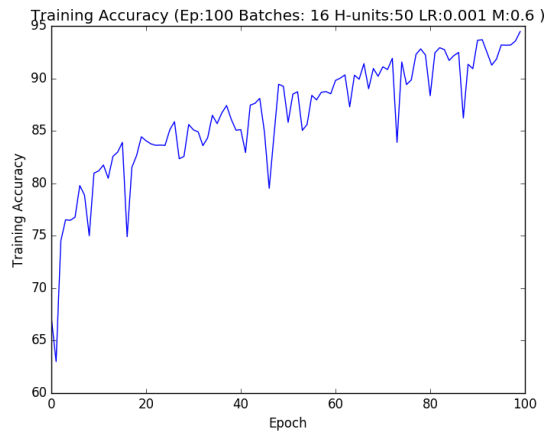
The maximum testing accuracy was about 82%

Part 4: Training Monitoring

The following graphs show the average losses, test accuracy and training accuracy for a set of Hyperparameters, and number of epochs. I tested with different Hyperparameters, as well as the number of epochs in order to get an idea for what values will create good results for the network. I kept the number of hidden units and epochs low because of computation capability of my laptop.

For the figure below, the parameters were: Epochs = 100, batch size = 16, hidden units = 50, learning rate = 0.001, momentum = 0.6

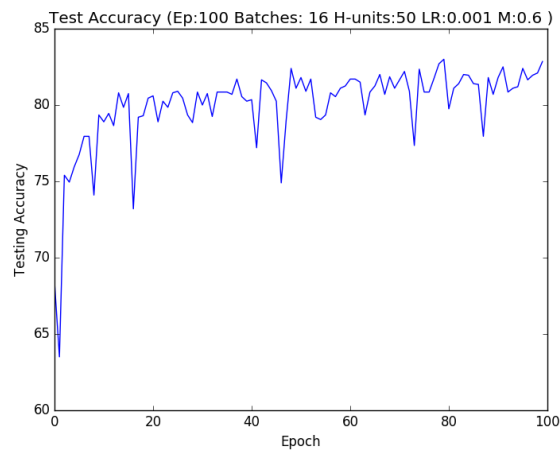




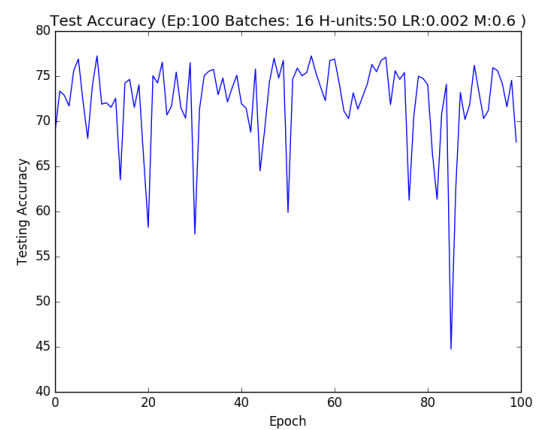
Part 5:

Test accuracy with different learning rates:

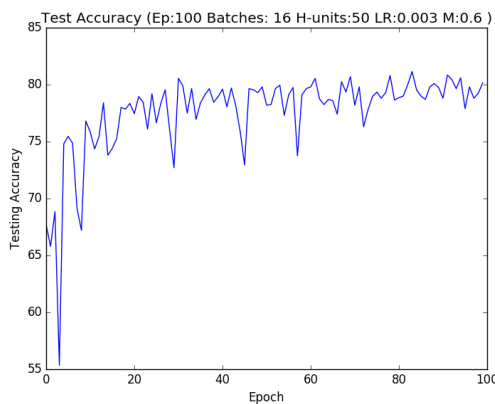
For the figure below, the parameters were: Epochs = 100, batch size = 16, hidden units = 50, and momentum = 0.6



LR = 0.001, accuracy ~82%



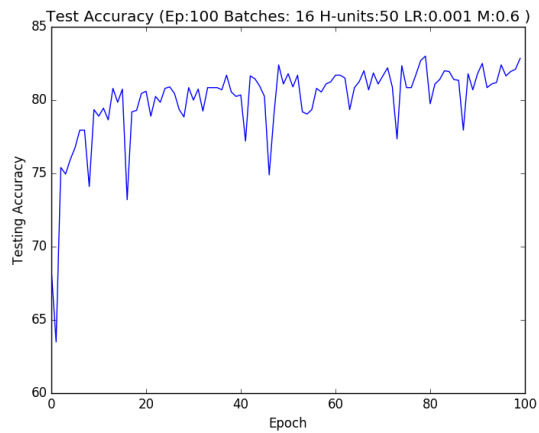
LR= 0.002, accuracy ~73%



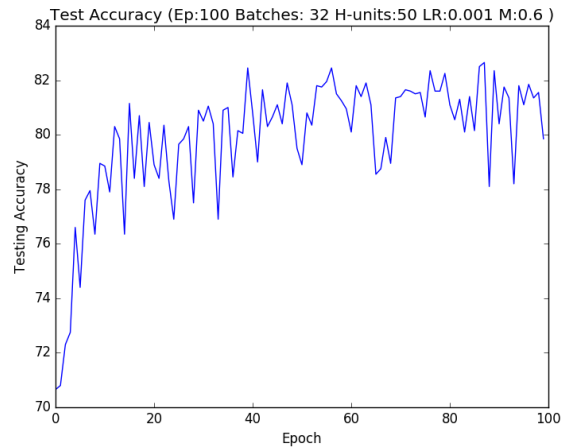
LR = 0.003, accuracy ~ 79%

Test accuracy with different batch sizes:

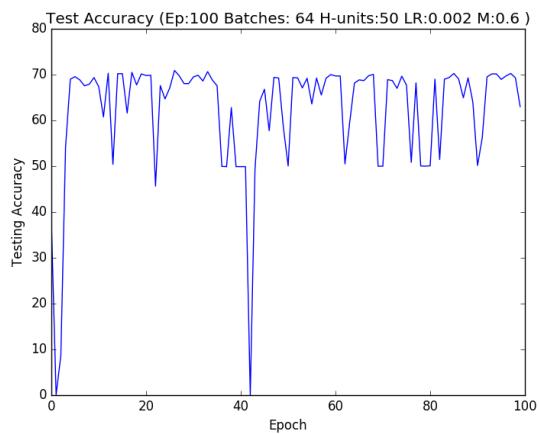
For the figure below, the parameters were: Epochs = 100, hidden units = 50, learning rate = 0.001 and momentum = 0.6



Batch size = 16, accuracy ~82%



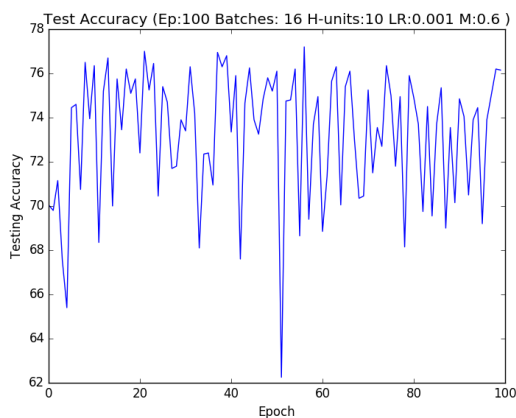
Batch size = 32, accuracy ~80%



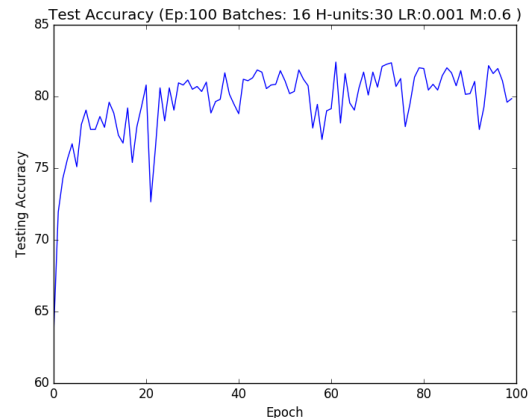
Batch size = 64, accuracy ~60%

Test accuracy with different number of hidden units:

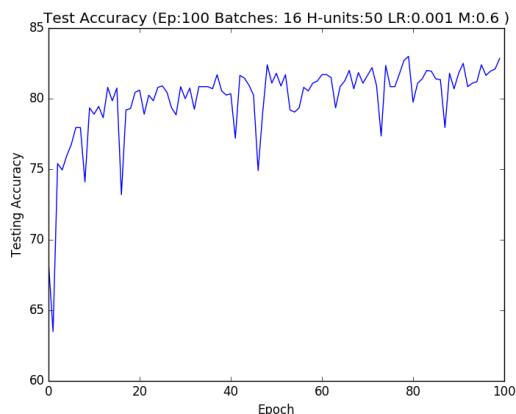
For the figure below, the parameters were: Epochs = 100, batch size = 16, learning rate = 0.001 and momentum = 0.6



Hidden units = 10, accuracy ~73%



Hidden units = 30, accuracy ~80%



Hidden units = 50, accuracy ~82%

Part 6: Discussion

The maximum accuracy of the neural network in this assignment was about 82% on the test set, which is very good for a single hidden layer network, with a small training set of 10000. I think that without using convolutional neural network, the accuracy would peak at around 85% for using only one hidden layer neural network.

Learning rate:

The learning rate seems to be a very sensitive parameter, since increasing it beyond a certain limit will cause the network to not learn anything. At this point it is bouncing/skipping past the local optimum.

Hidden units:

Very less number of hidden units (e.g. 10), will cause the test and train accuracy to be low. This is because the network is not expressive enough. However, increasing the number of hidden units beyond a certain point will not further increase the test/train accuracies. Moreover, the variance was larger in using lower number of hidden units.

Increasing the number of hidden units from 10 to 30 had more effect on the testing accuracy than compared to increasing the number of hidden units from 30 to 50.

Batch size:

Batch size of 16 worked best in my neural network implementation. Batch size of 64 (maximum) performed the worst. It seems that there is less variance in the test accuracy when using a smaller batch size.