Report 2 Layer classifier

1 Analytical gradient check

I checked the analytical gradient computations by training the network on the training data with regularization turned off (lambda=0). The network overfitted the data with an accuracy on the training set of 60 % and a loss around 1.2 after training on 200 epoches and with a reasonable $\eta = 0.01$. I am sure the computation is bug free since I got reasonable result and percentage like the one in the pdf.

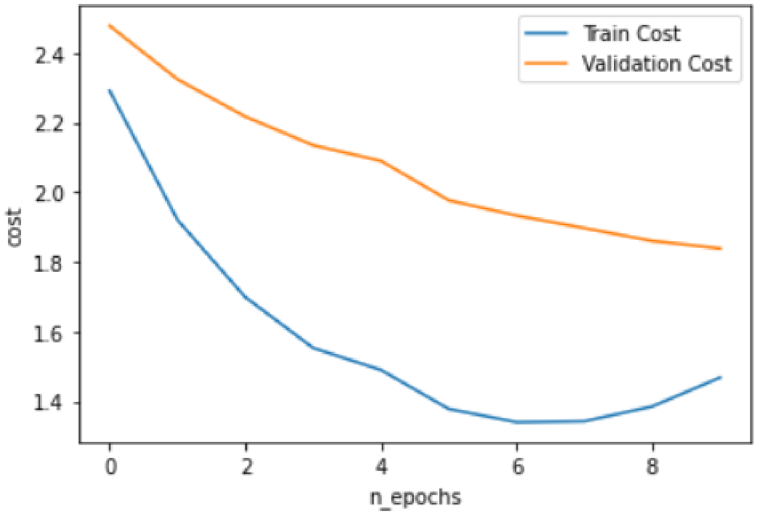

Figure 1: Loss obtained for $\lambda = 0.01$, batches = 100, $step\_size = 500$, eta_min = 1e-5 and eta_max = 1e-1

2 Curves for training and validation loss/cost

- Figure 3:

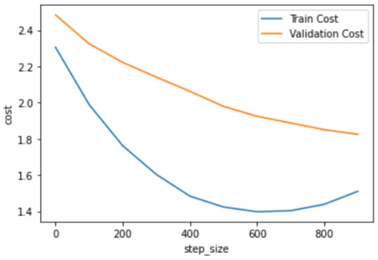

Figure 2: lambda=0.01, epochs=10, batch=100, eta min = 1e-5, eta max = 1e-1, step size = 500

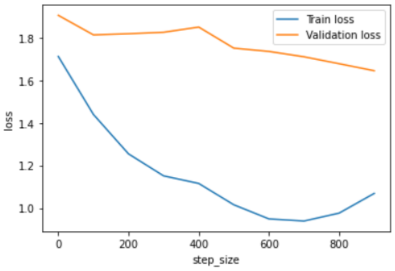

Figure 3: lambda=0.01, epochs=10, batch=100, eta min = 1e-5, eta max = 1e-1, step size = 500

- Figure 4:

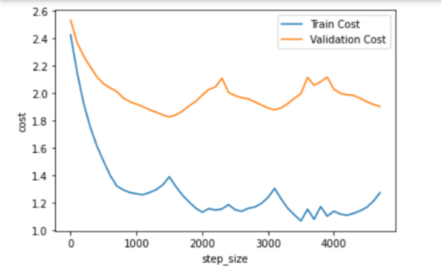

Figure 4: lambda=0.01, epochs=48, batch=100, eta min = 1e-5, eta max = 1e-1, step size = 800

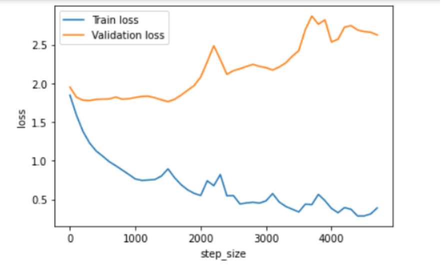

Figure 5: lambda=0.01, epochs=48, batch=100, eta min = 1e-5, eta max = 1e-1, step size = 800

We can see that increasing the step size and the number of cycles will cause the learning rate to increase and decrease more slowly, which can lead to slower convergence and a more zig-zag loss and cost plot. This is because the learning rate is changing less frequently, and so the model may take longer to find the optimal weights.

3 Range of the values for coarse search of lambda

As suggested, I started the first search between l min = -5 and l max = -1 for 8 random values and for 2 cycle each:

```
l = l_min + (l_max - l_min)*np.random.rand(1, 1)
lbd = (10 ** l) [0][0]
```

The best 3 lambdas: 8.30453e-05, 3.62624e-05 and 0.0050573 with respective accuracy of 0.5102, 0.5072 and 0.5072.

4 Range of values for the fine search of lambda

I choose the best accuracy of the coarse search and then I performed a normal search around the best accuracy found for 4 cycles.

```
lbd = np.random.normal(best_lbd, best_lbd*0.01)
```

The best 3 lambdas: 8.40041e-05, 8.35670e-05 and 8.30955e-05 with respective accuracy of 0.526, 0.5246 and 0.523.

5 Plot for best lambda

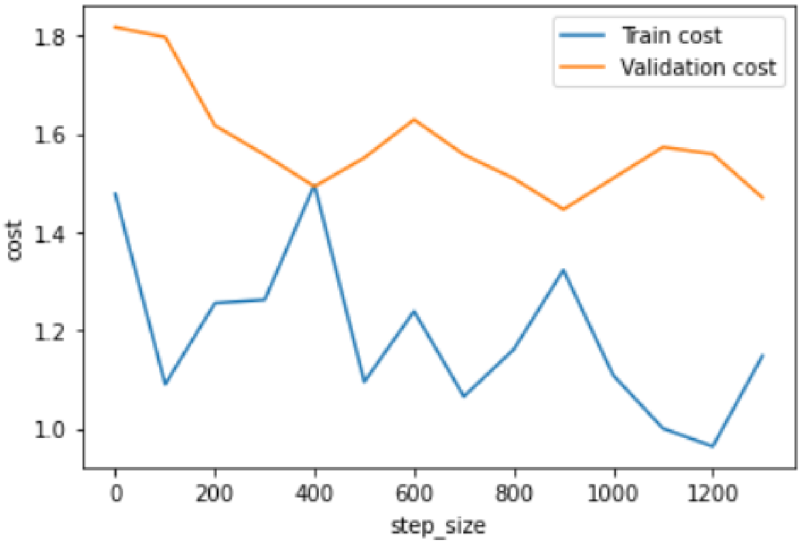

Figure 6: Cost for the best lambda

The test accuracy is 0.5217.