Benjamin Ye

CS/CNS/EE 156a: Learning Systems (Fall 2023)

November 17, 2023

**Bonus Exercise**

**1 Classifying digits with NNs**

**1.3 Number of Parameters**

What happens to the learning curve when you vary the number of hidden units? Specify the number of hidden units you use in each layer, and report learning curves for each different architecture you try. What trends do you notice? What is the smallest number of model parameters (not hidden units; view this via the summary function) for which you can achieve over 95% validation accuracy? An estimate is fine.

**1.4 Regularization**

Explore the usefulness of weight decay regularization (minimizing instead of just ) on the large neural net implementation give in the support code. You can do so changing the REGULARIZATION parameter in the IPython notebook, or by specifying the -r command line option in train.py in the command shell. Report learning curves for different choices of (just a few). Can you use regularization to get a similar result to one of the smaller neural nets you implemented in the previous part?

**1.5 Activations**