Assignment2

Benjamin Klybor

September 20, 2018

My repository for this assignment can be found here: `https://github.com/bklybor/Assignment1.git`

Trial	epochs	batch_size	activation	normalized	learning rate	Accuracy	Loss
1	20	128	linear	no	0.01	0.098	nan
2	40	128	linear	no	0.01	0.098	nan
3	20	256	linear	no	0.01	0.098	nan
4	20	128	linear	yes	0.01	0.819	3940
5	20	128	sigmoid	no	0.01	0.427	0.0726
6	20	128	sigmoid	yes	0.01	0.573	0.0568
7	20	128	linear	no	0.20	0.098	nan
8	20	128	linear	yes	0.20	0.823	3910
9	20	128	sigmoid	yes	0.20	0.862	0.0294
10	20	128	sigmoid	yes	0.60	0.873	0.0396
11	20	128	sigmoid	yes	0.20	0.823	3910
12	20	128	softmax	yes	0.20	0.910	0.0179

Figure 1: A table of accuracies for the MNIST data set.

From Fig. 1 we can see that changing the number of epochs, batch size, or learning rate has no effect on the accuracy of our model. In fact with simple linear activation the accuracy is woefully low at about 10%. However, simply normalizng our data increases the accuracy to about 80%, far better than the random guessing of the other linear trials. Accuracy is further improved if we increase the learning rate.

Most surprising to me is that linear activation with normalization is more

accurate than sigmoid activation with normalization, but that when we raise the learning rate to the same rate for each, 0.20, the sigmoid function provides greater accuracy than the linear activation.