ANN Report for Homework 3

Palani Johnson

September 25, 2021

Training

Training ANNs with 1 Hidden Layer on MNIST:

Table 1: mini batch = 10, num epochs=10

ETA/HLS	10	25	50
$\eta = 0.5$	90.02%	92.37%	93.36%
$\eta = 0.25$	87.27%	90.93%	83.53%
$\eta = 0.125$	82.13%	87.32%	73.66%

Training ANNs with 2 Hidden Layers on MNIST:

Table 2: $h_1 = 10$, mini batch = 10, num epochs=10

ETA/h_2	10	25	50
$\eta = 0.5$	89.32%	90.26%	90.76%
$\eta = 0.25$	88.18%	88.72%	90.02%
$\eta = 0.125$	82.00%	86.73%	79.86%

Table 3: $h_1 = 25$, mini batch = 10, num epochs=10

ETA/h_2	10	25	50
$\eta = 0.5$	92.16%	92.14%	93.04%
$\eta = 0.25$	89.62%	90.93%	90.59%
$\eta = 0.125$	87.64%	88.95%	80.54%

Table 4: $h_1 = 50$, mini batch = 10, num epochs=10

ETA/h_2	10	25	50
$\eta = 0.5$	92.48%	93.21%	93.59%
$\eta = 0.25$	90.10%	91.61%	92.38%
$\eta = 0.125$	88.56%	89.35%	90.41%

Observations

I noticed that the value of η had the largest impact on the final result NN after 10 epochs. Its is to be expected, however I also expected deeper neural nets to perform better. This was not the case and the nets tended to perform just as well with one layer as with two. It seems the number of nodes in the first hidden layer generally determines how well the the network performs. While deeper nets can marginally improve performance, the training time disproportionately increased.