

Architecture	Accuracy	Etas	Mini Batch Size	Epochs
784, 30, 60, 120, 10	95.01	50	20	5
784, 23, 23, 23, 23, 10	94.43	25	15	5
784, 30, 10	94.66	50	15	5
784, 23, 23, 10	94.56	30	20	5
784, 100, 10	86.34	50	15	5
Ensemble	96.39			

Figure 1: Table of performance of ANNs

1 Description

This table shows that the ensemble does a bit better than all of the individual trained models. The etas seem to be semi random in terms of their accuracy vs performance alterations. You would be a little justified in saying the larger batch size of 20 performed better. The primary control for performance appears to be the neural networks structure. Larger layers require more training. A single larger layer can perform better then multiple layers. Overall, ensembles seem to work. The concept of the ensemble should have more error as we have repeated chances to get something wrong. In actuality it is not a repeated chance of failure, but a weighted average of proximity to the appropriate answer. It in essence says that a bunch of neural networks are in this area, it is more likely to be in this area.