Dakota Crowder

CSCEA415 Machine Learning

Project 1 Artificial Neural Network

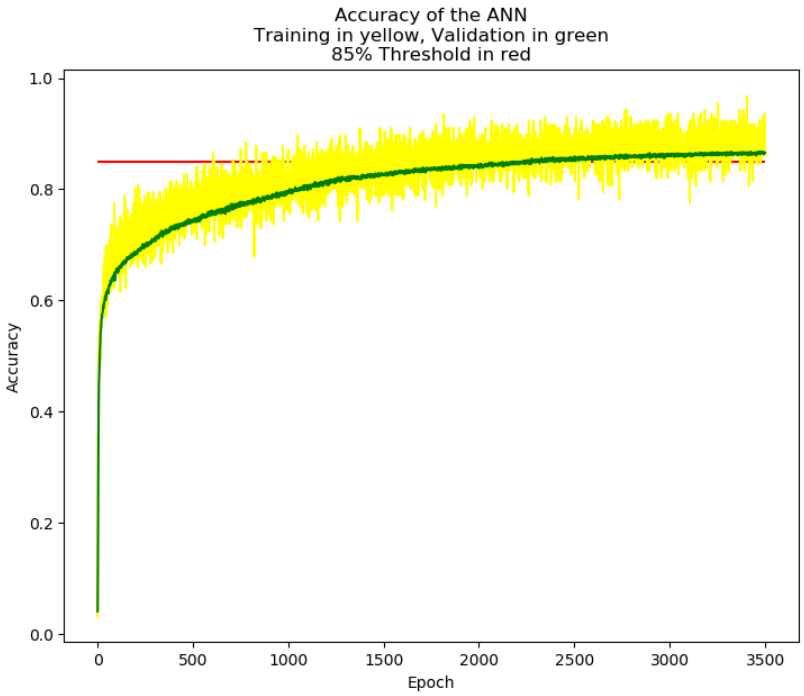
Write Up

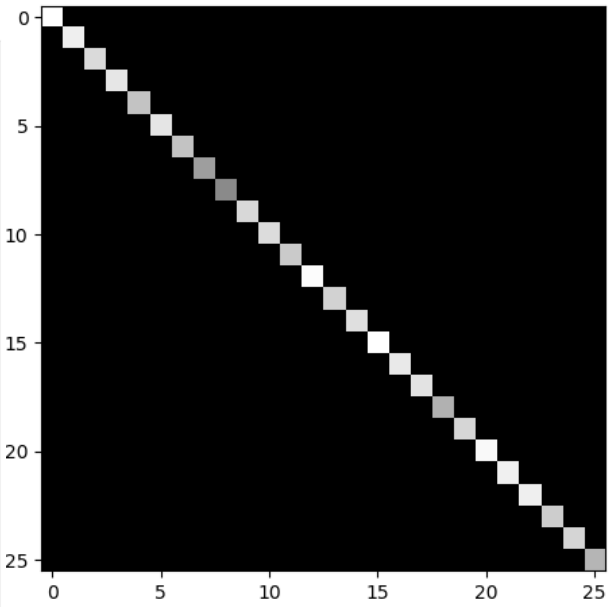
Constants for Each run:

Batch size is 160, hidden neuron count is 100, epochs set to each run to make it get to the 85% or better. An epoch is considered a run of 16000 letters from the training set, which is the entire size of the training set, validation is done on the entire validation set. Batch size is set to 160 just to have the easy calculation that 100 batches is 1 epoch.

Results for different Training Rate:

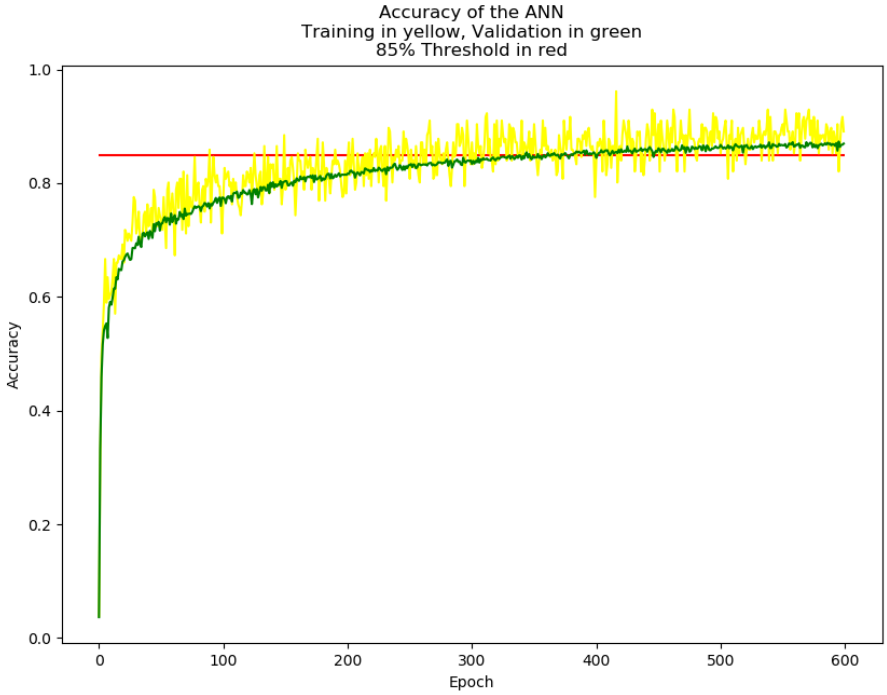
μ = 0.001, 3500 epochs

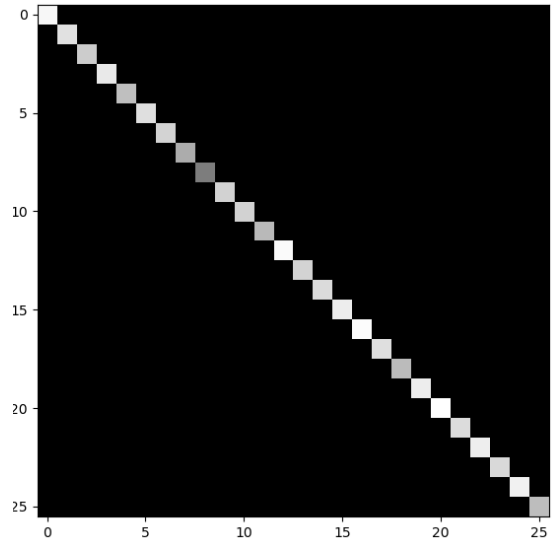




85% threshold Reached at epoch 2180, which took about 10x longer than 0.01 learning rate. Learning was very slow, it took so long to get anywhere close to 85%

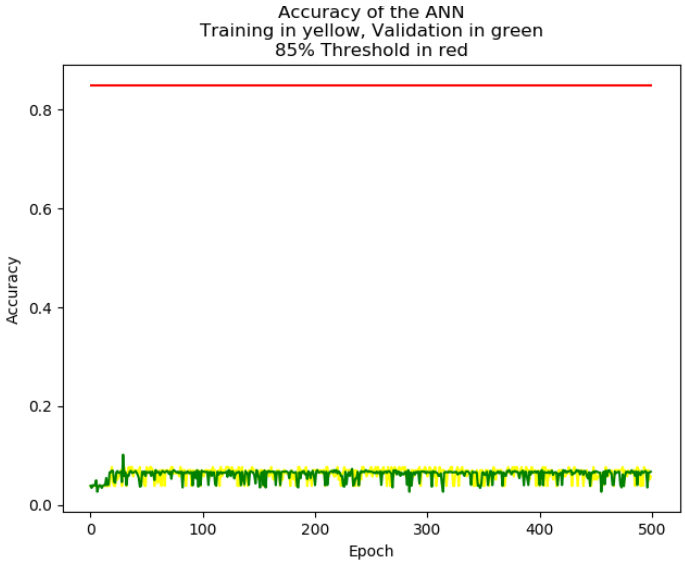
μ = 0.01, 600 epochs

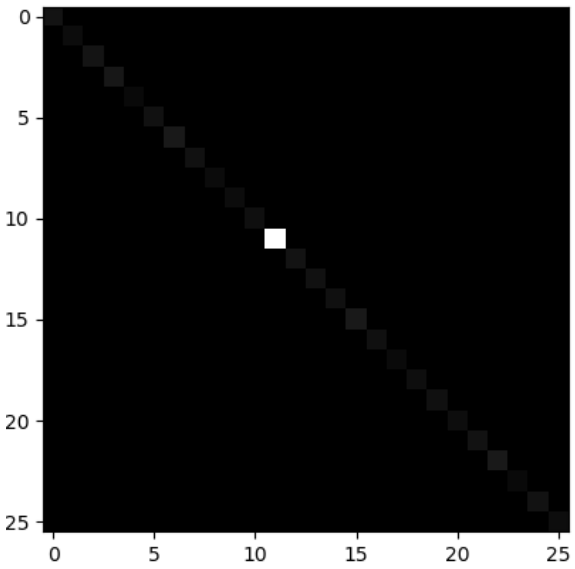




The heat map shows that the diagonal line is the most common, which is expected, and the first epoch that is at the 85% threshold is 344, the rest is just getting better. The learning get good very quickly, then slowly makes it to 85% to 87%.

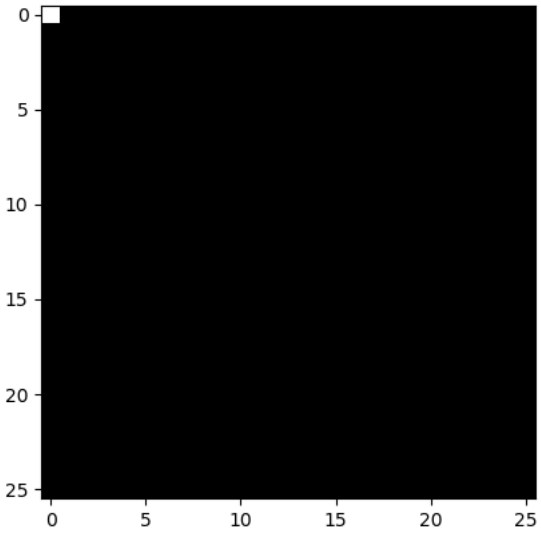
μ = 0.1, 500 epochs

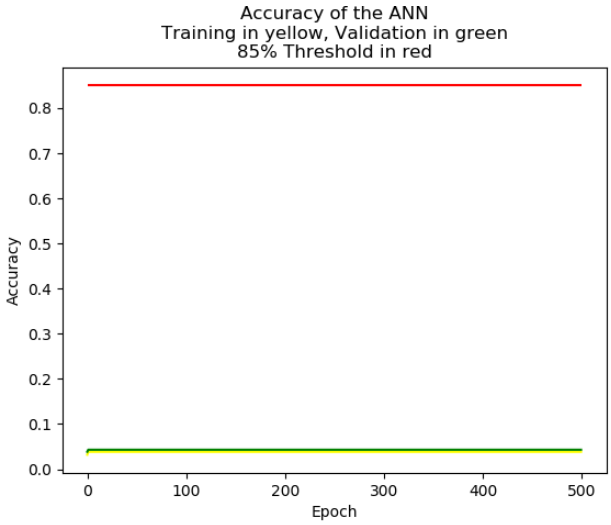




The main issue with this learning rate, is that the math quickly goes to the point of overflow, yet still runs, so I assume that the machine just doesn't do back propagation. It is just basically guessing the same letters repeatedly.

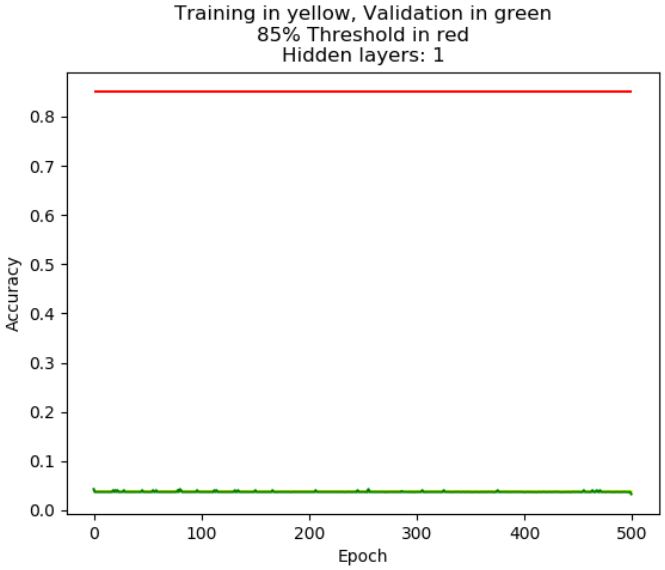
μ = 0.5, 500 epochs

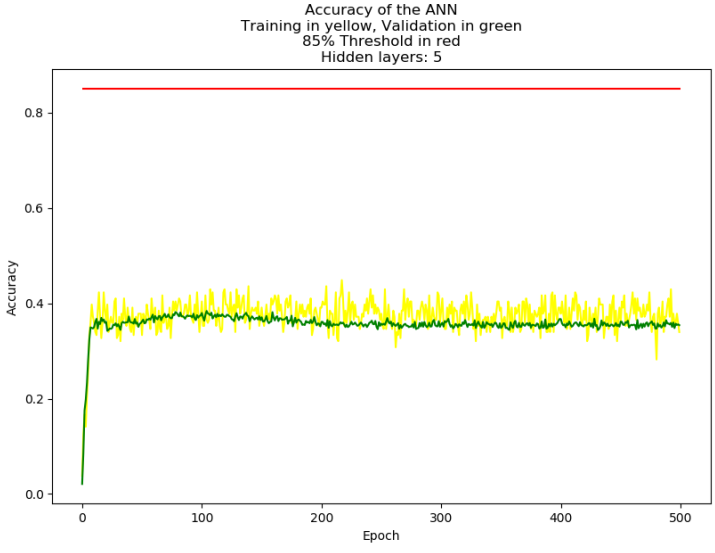


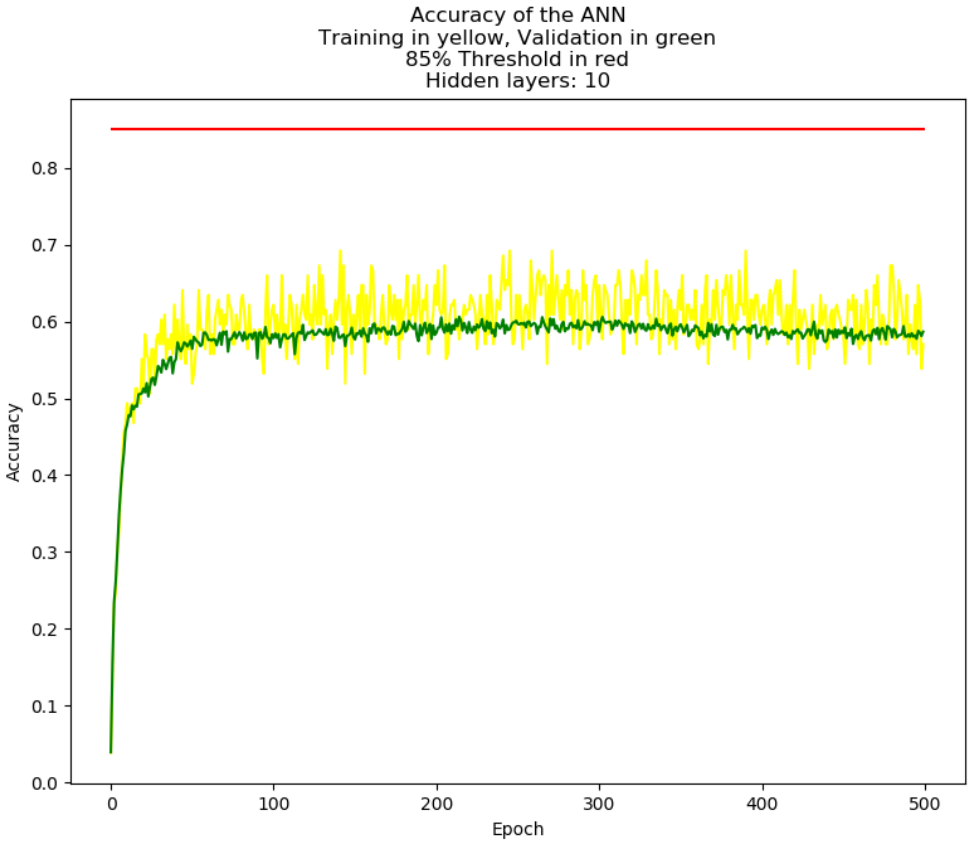


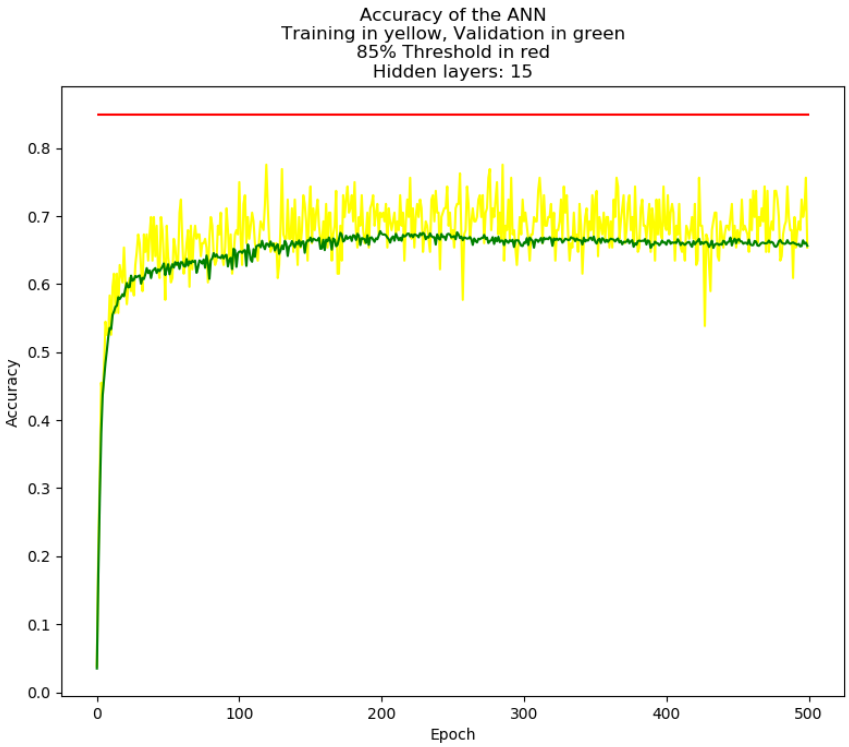
The same issue that was with the learning rate 0.1 is present here, the math gets too large, so there is overflow and no learning is done.

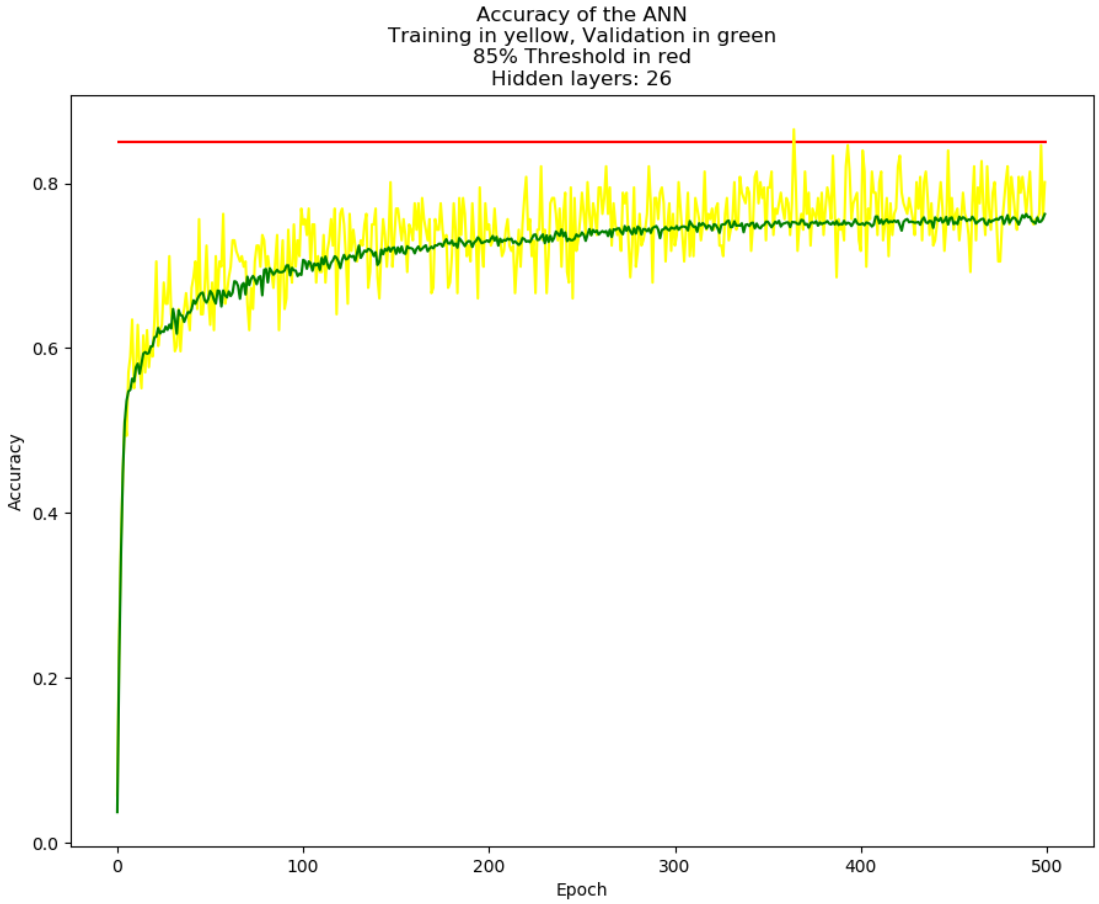
I wanted to do some testing with different sizes for the hidden layer. I will use 0.01 since that has the best and quickest learning. It says hidden layers, it isn't multiple layers, it is the number of neurons in the layer, that is a typo. The more the neurons in the hidden layer, the better it is at learning. Though there is a trade off as the more the neurons the longer the computation takes Those tests are going to be place here:

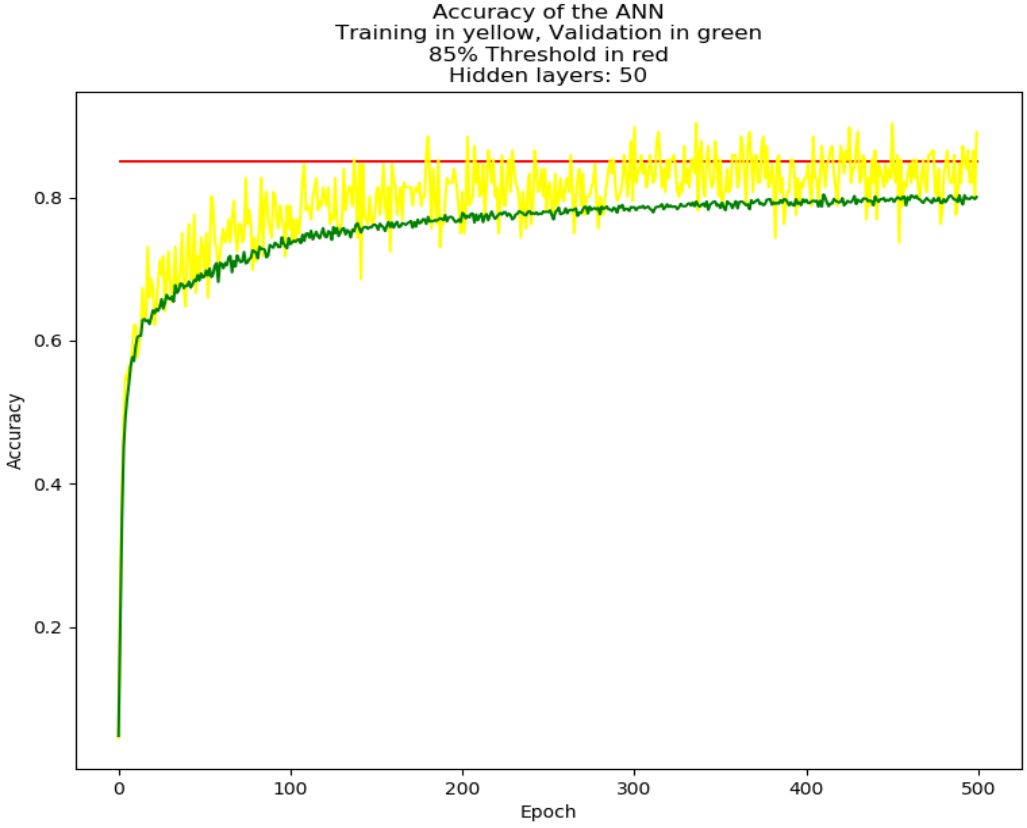


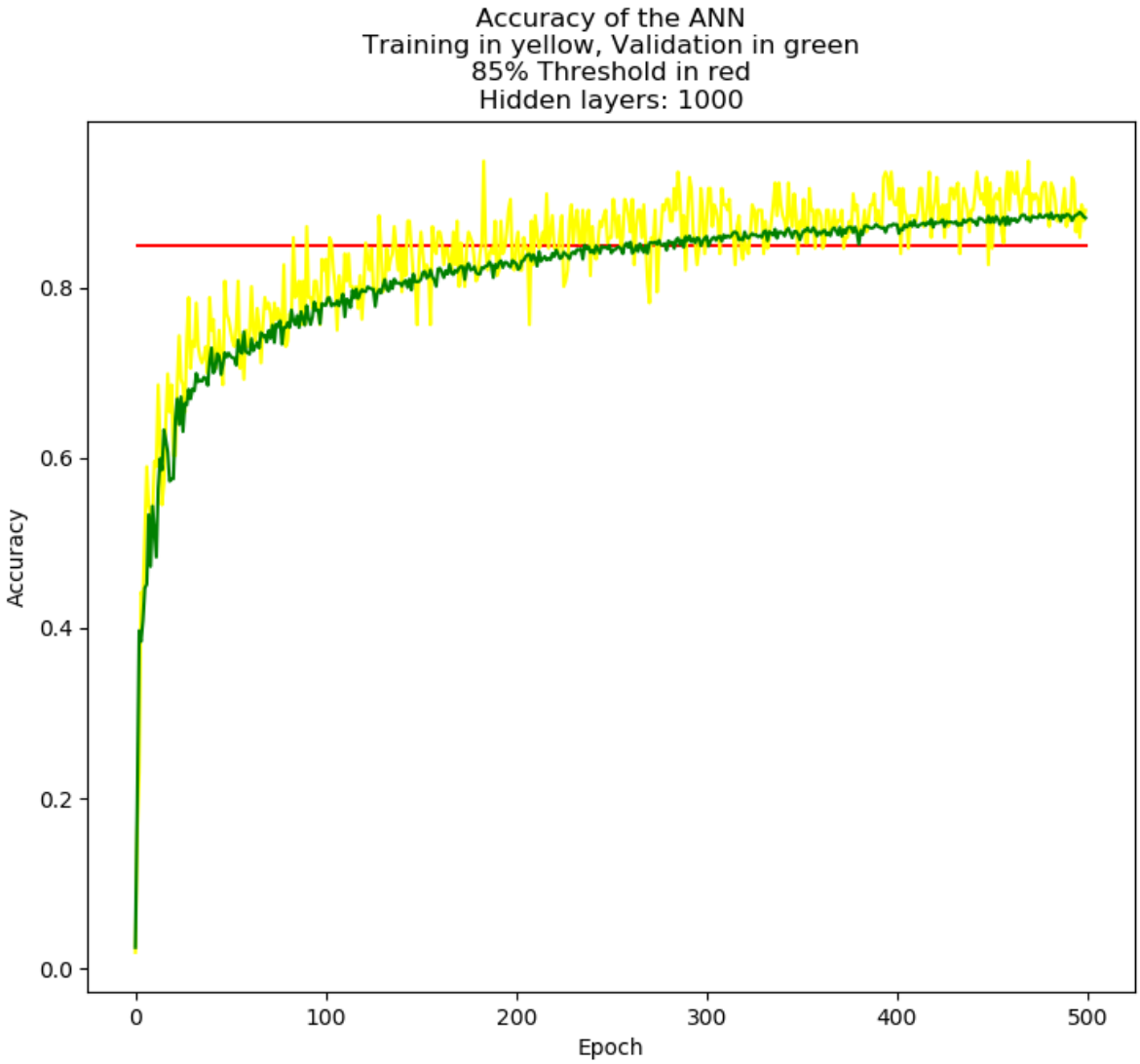



Accuracy of the ANN
Training in yellow, Validation in green
85% Threshold in red
Hidden layers: 10
Accuracy
Epoch
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0
100
200
300
400
500


Accuracy of the ANN
Training in yellow, Validation in green
85% Threshold in red
Hidden layers: 15
Accuracy
Epoch
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0
100
200
300
400
500


Accuracy of the ANN
Training in yellow, Validation in green
85% Threshold in red
Hidden layers: 26
Accuracy
Epoch
0.0
0.2
0.4
0.6
0.8
0
100
200
300
400
500


Accuracy of the ANN
Training in yellow, Validation in green
85% Threshold in red
Hidden layers: 50
Accuracy
Epoch
0.2
0.4
0.6
0.8
0
100
200
300
400
500


Accuracy of the ANN
Training in yellow, Validation in green
85% Threshold in red
Hidden layers: 1000
Accuracy
0.0
0.2
0.4
0.6
0.8
0
100
200
300
400
500
Epoch