Danny North

2/18/16

CS 478

Dr. Martinez

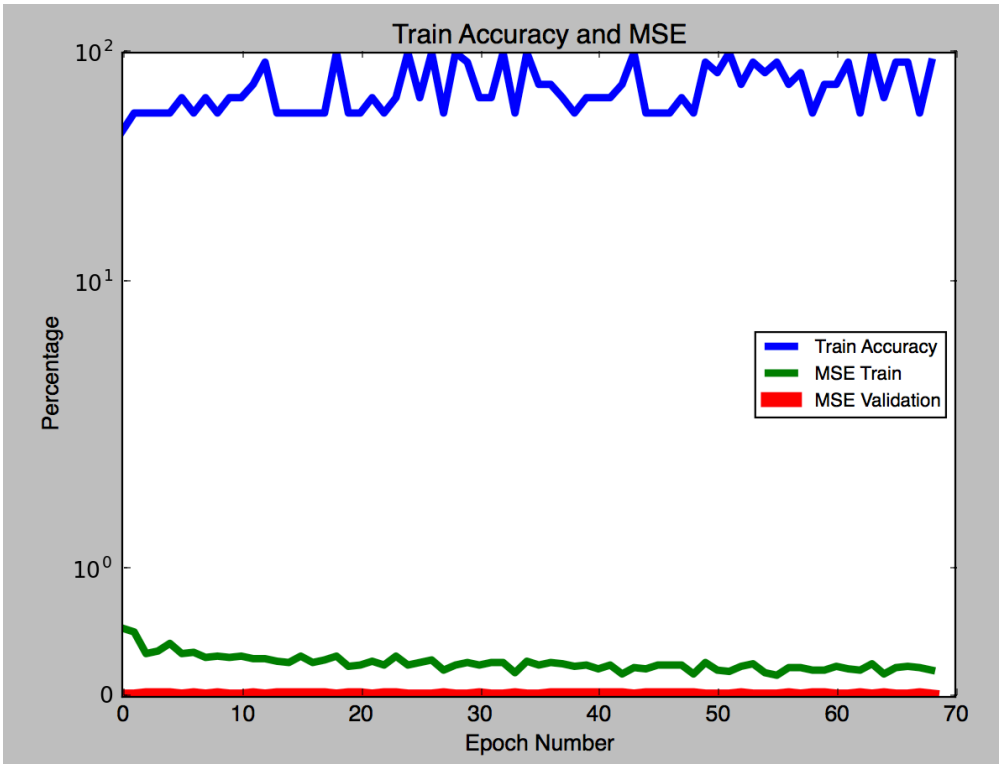
Backpropagation Lab

I. Classifying the Iris Dataset

In order to classify the iris dataset, three output nodes were needed. I used two hidden layer nodes per input, and had a learning rate of 0.1 with a momentum of 0. I also randomly split the data on a 75/25 split and further split the training data to a training and validation set of 90/10. Below is the graph with the results from one test, and the average epoch length and accuracy from five tests.

The stopping criteria that was used on this and further tests was a BSSF that didn't improve over 300 epochs, or if the BSSF reached 100 then it would continue 5 more epochs and then stop.

Average Accuracy over 5 tests: **94.8 %**
Average Number of Epochs over 5 tests: **76**

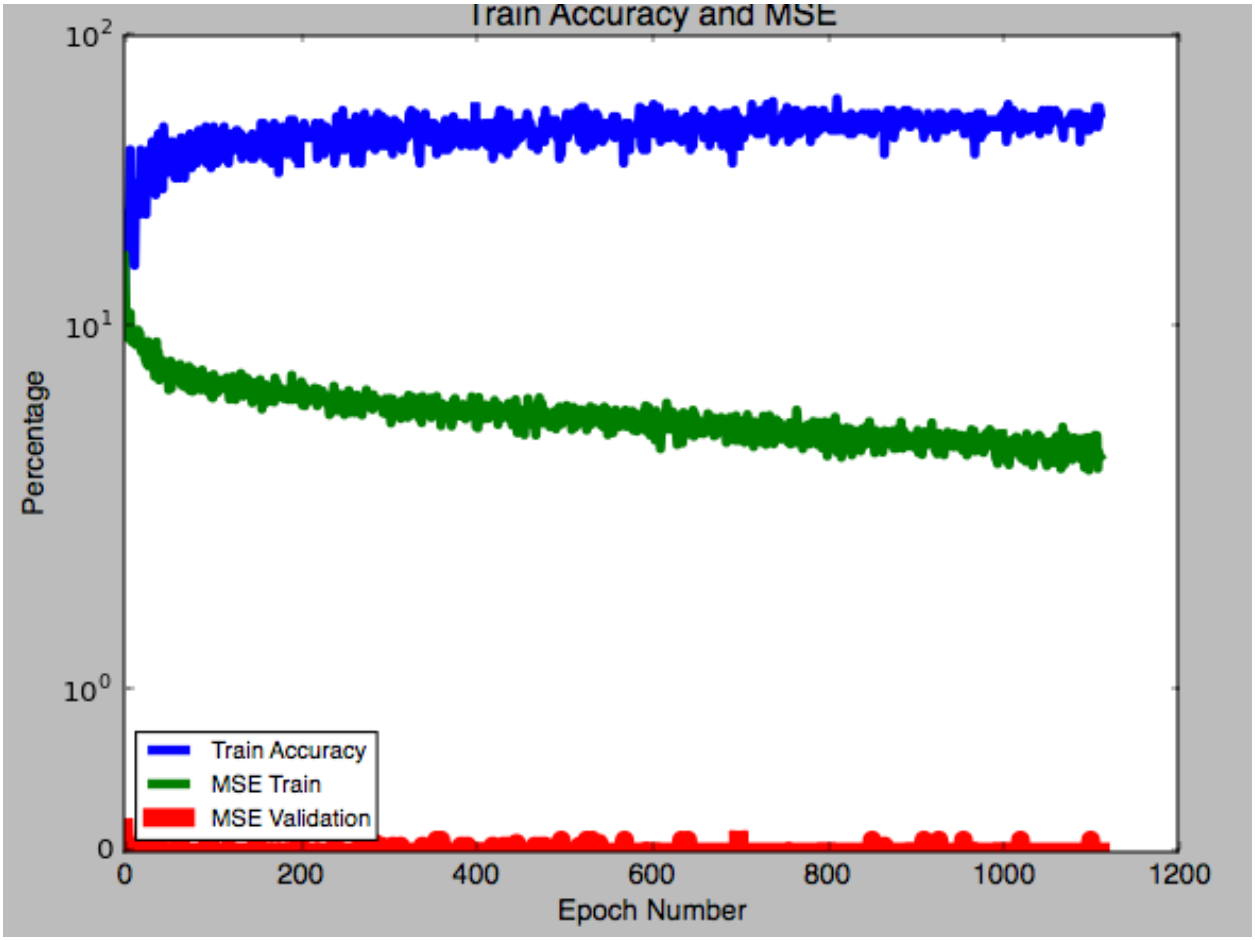


II. Classifying the Vowel Dataset

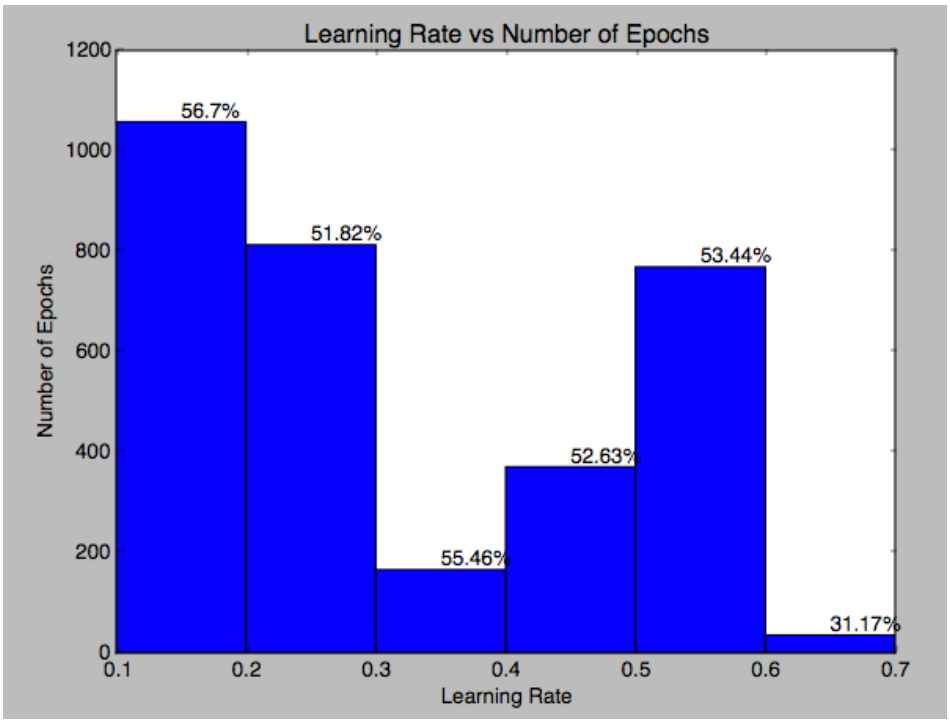
This dataset was a lot larger and more difficult to train. Using a learning rate of 0.1 and a momentum of 0.1, I was able to get results in the 50% - 60% range.

Average Accuracy over 5 tests: **56.4 %**

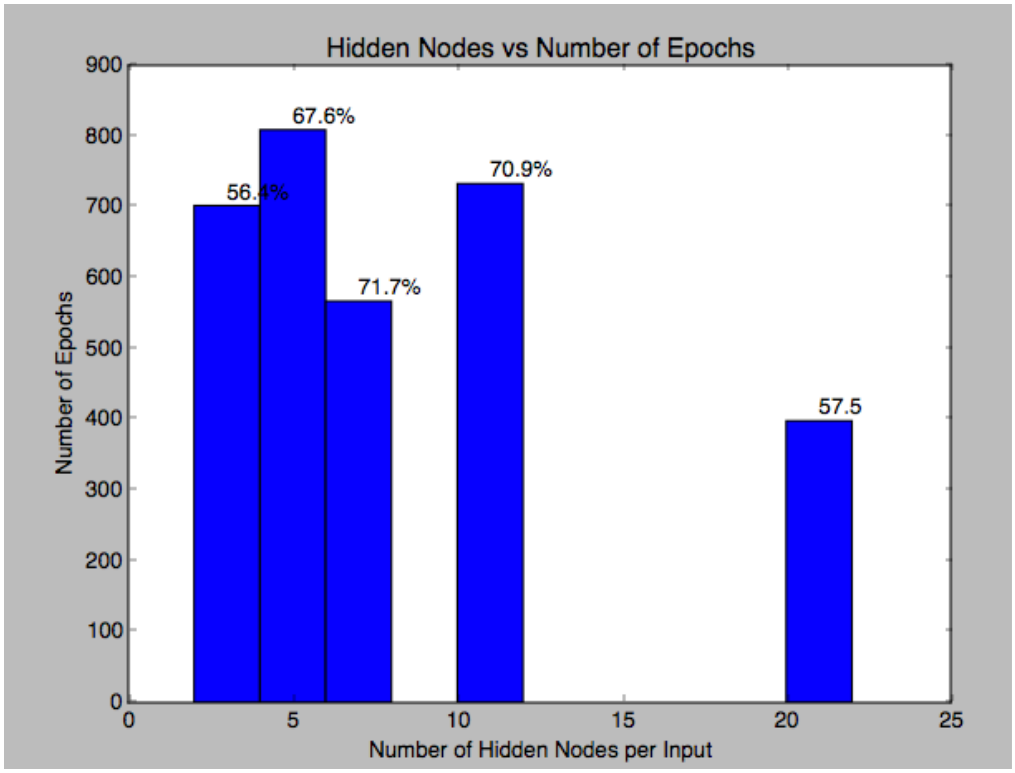
Average Number of Epochs over 5 tests: **702**



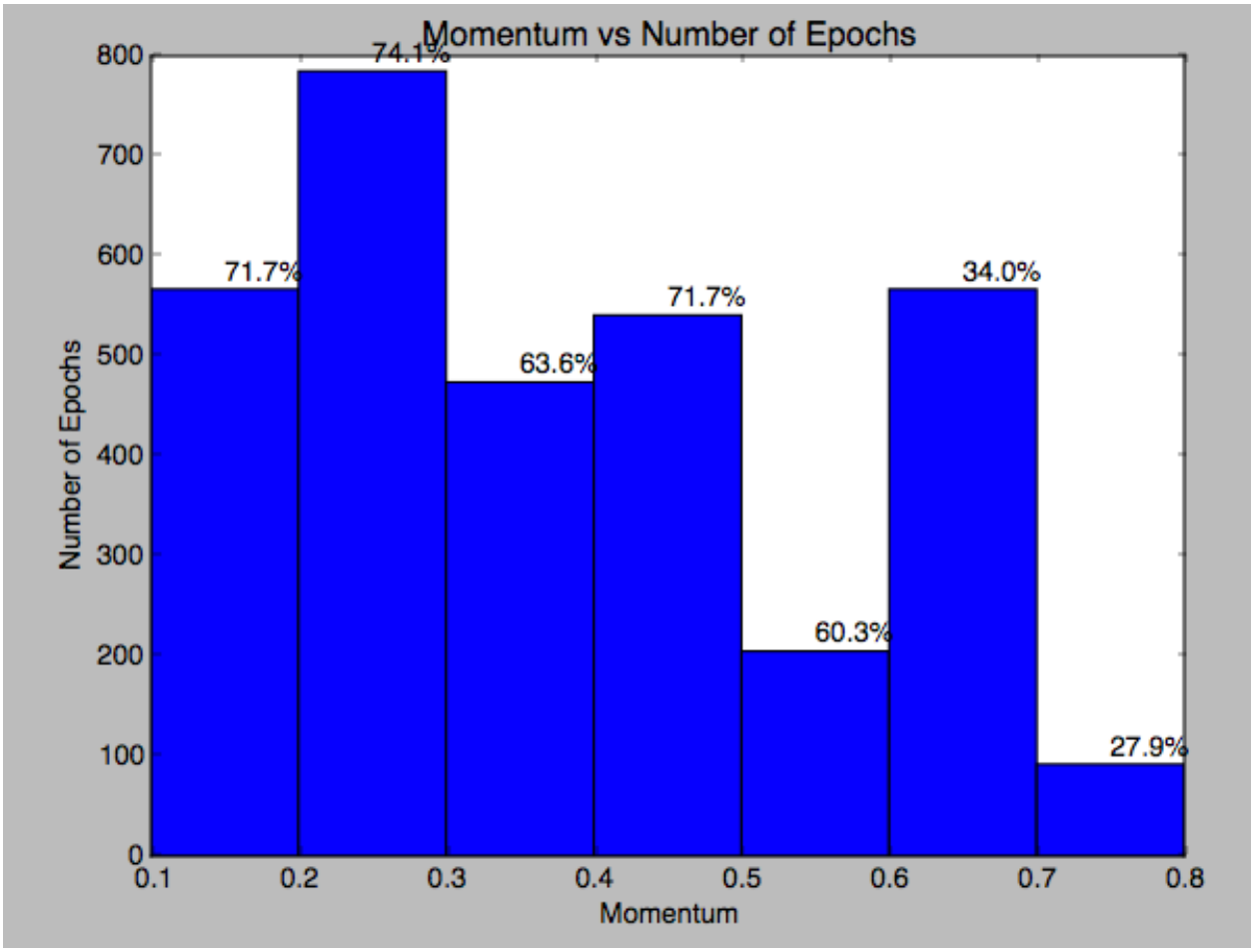
The graph below shows the learning rate vs the number of epochs it took to reach the BSSF. Above the bars is a percentage. This was the percentage that the BSSF was settled at, and I believe it's an important and interesting number to look at. You will notice that although the learning rate of 0.6 settles VERY quickly, the accuracy is very poor. This is why I stopped at this epoch, because I would prefer a higher accuracy with more epochs as long as there can be a trade-off. I believe a learning rate of 0.3 is the winner in that aspect. (It is important to note that these tests were only run once and are not an average, meaning this graph could be drastically different if taken the time to run all of the tests.)



The next graph shows the number of hidden nodes per input node vs the number of epochs. Again, the accuracies are shown in these graphs and we can see that at first, the more nodes hidden nodes we have per input, the higher our accuracy. In fact, we went from about 52% accuracy to 71% accuracy in less epochs; that's a big jump! However, it falls off after the test of 6 hidden nodes per input.

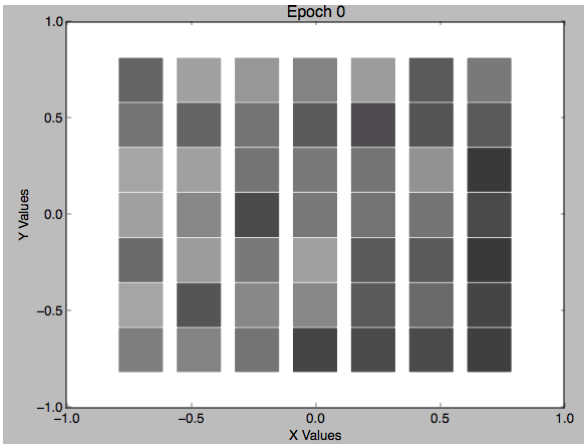


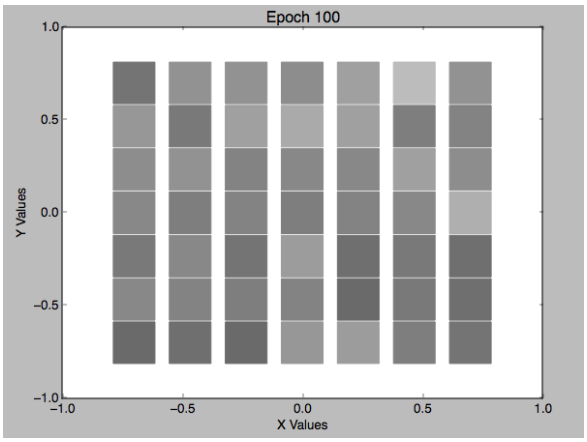
The final test was testing the different momentums. The graph below is visualized the same way as the other two, and you can see that the accuracy for the 0.2 momentum was the best, however it also took the most epochs to complete. And although 0.7 had the least amount of momentums, it was also the worst accuracy.

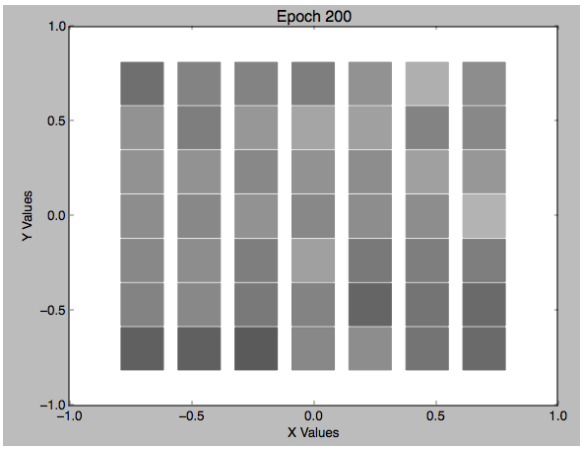


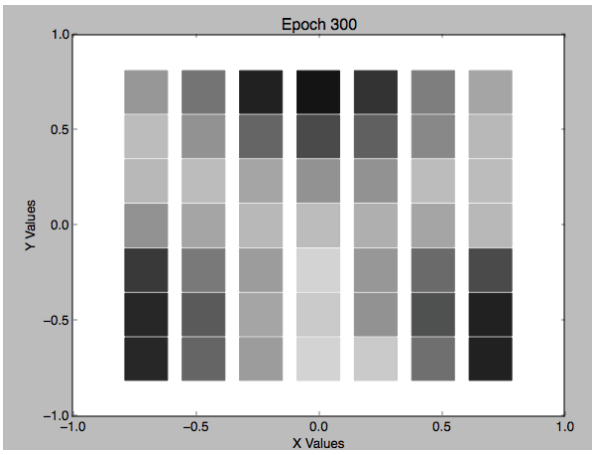
III. Interesting Experiment

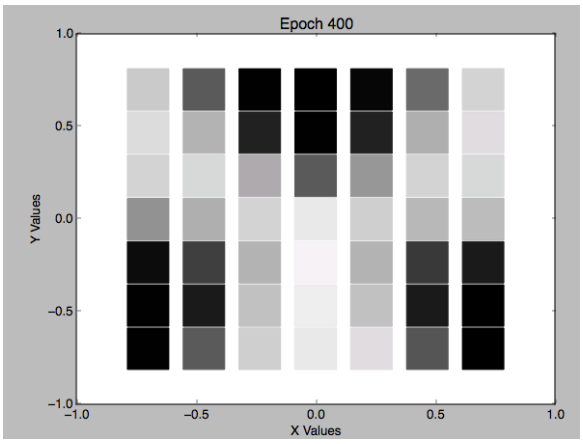
For the experiment portion of this lab, I decided to test the visualizations of running the Back propagation on the dataset that forms a Y on convergence. This was an interesting experiment because I was able to visually see the difference the backpropagation algorithm had on the data. Because the first couple hundred epochs produce something that is non-recognizable, it is interesting to note that the last thousand epochs between 1000 and 2000 has a very small and fine-tuned change. Once the change is propagated to be partly correct, then we can already assume it's mostly correct and that's something interesting to consider for future labs and problems solved. The visualizations for this test are below, and the grayscale prediction range from 0 being white and 1 being black, so the outputs and accuracies are shown in the image, since it is meant to take a Y shape.

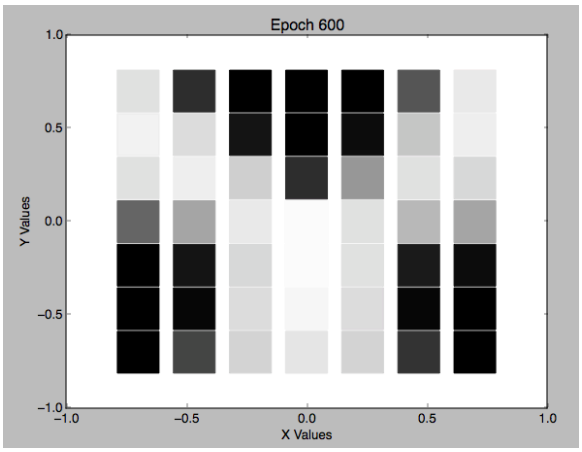

Epoch 0
Y Values
X Values
1.0
0.5
0.0
−0.5
−1.0
−1.0
−0.5
0.0
0.5
1.0

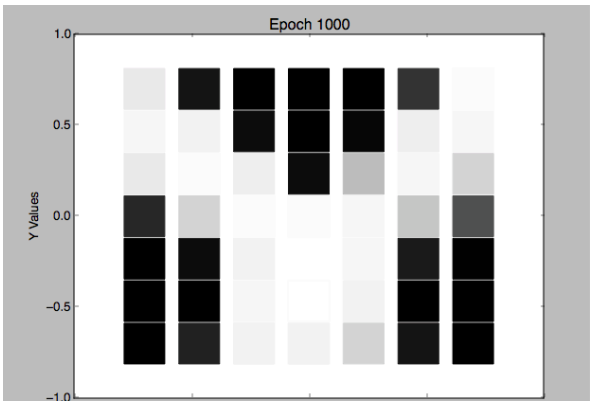

Epoch 100
Y Values
X Values
1.0
0.5
0.0
−0.5
−1.0
−1.0
−0.5
0.0
0.5
1.0

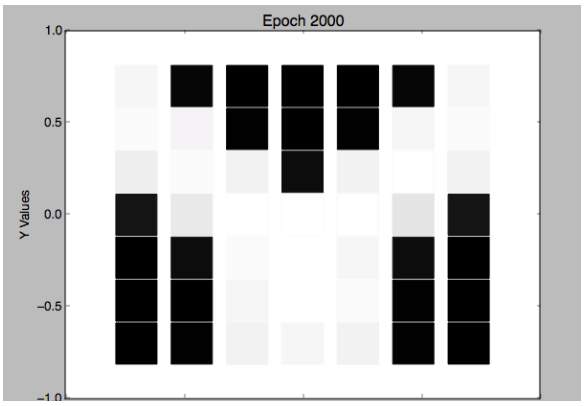

Epoch 200
Y Values
X Values
1.0
0.5
0.0
−0.5
−1.0
−1.0
−0.5
0.0
0.5
1.0


Epoch 300
Y Values
X Values
1.0
0.5
0.0
−0.5
−1.0
−1.0
−0.5
0.0
0.5
1.0


Epoch 400
Y Values
X Values
1.0
0.5
0.0
−0.5
−1.0
−1.0
−0.5
0.0
0.5
1.0


Epoch 600
Y Values
X Values
1.0
0.5
0.0
−0.5
−1.0
−1.0
−0.5
0.0
0.5
1.0


Epoch 1000
Y Values
1.0
0.5
0.0
−0.5
−1.0


Epoch 2000
Y Values
1.0
0.5
0.0
−0.5
−1.0