

Nick Nelson
Neural Networks

Note: Resubmitting because I fixed the bug that was preventing more than one hidden layer to work. I also went and did some work with Weka.

Please submit a link to your GitHub repository for your class.

<https://github.com/nicknlsn/CS450/blob/master/week05/neural/neural.py>

Describe your overall approach to implementing the algorithm in code. How are your classes/data structures organized? How do you keep track of the necessary pieces for back-propagation.

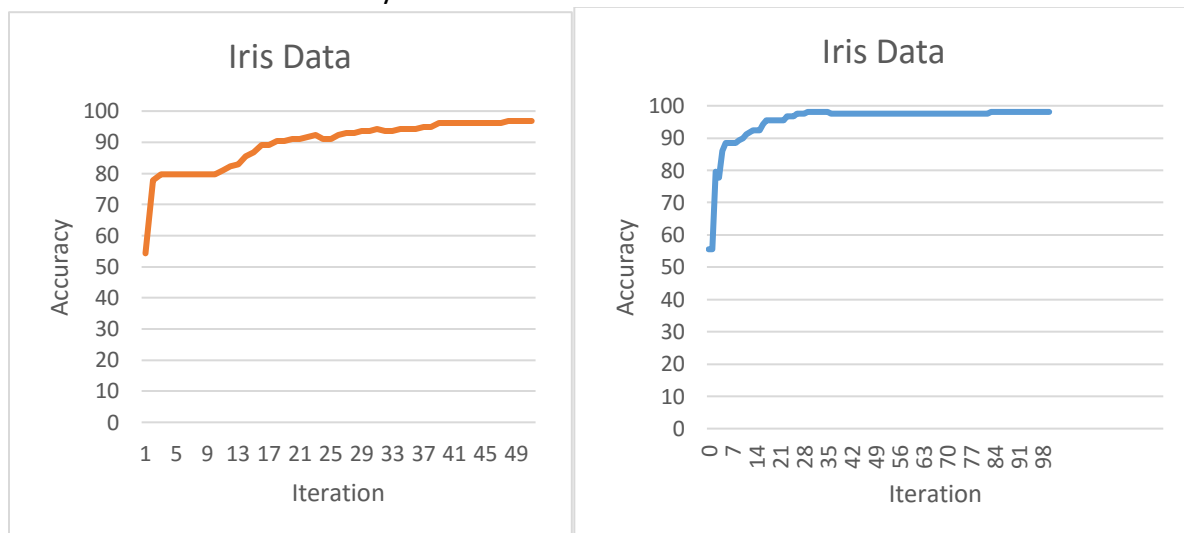
I use a NeuralLayer class to store the weights, activations, and errors of a node. This class knows how many nodes there are based on the shape of the weights array. Then I have a NeuralNetwork class that will create an array of NeuralLayers for however many layers are put into the network. This class keeps an output layer of type NeuralLayer separate from the array of hidden layers.

Describe the part of the assignment that gave you the most trouble, and how you overcame it.

The hardest part was dynamically handling different numbers of hidden layers. I got it to work for one hidden layer, but there is still at least one bug for multiple layers, which has me stumped so it isn't fixed. A network with one hidden layer seems to work for some datasets but not all, and I am not sure why.

Produce at least one graph to show the training progress for the Iris dataset:

The first graph shows the network accuracy over 51 iterations. The network has one hidden layer with 4 nodes, and the validation set came out with 98.5% accuracy. The second graph shows the data with the same network structure, but with 101 iterations. The validation set came out with 100% accuracy.

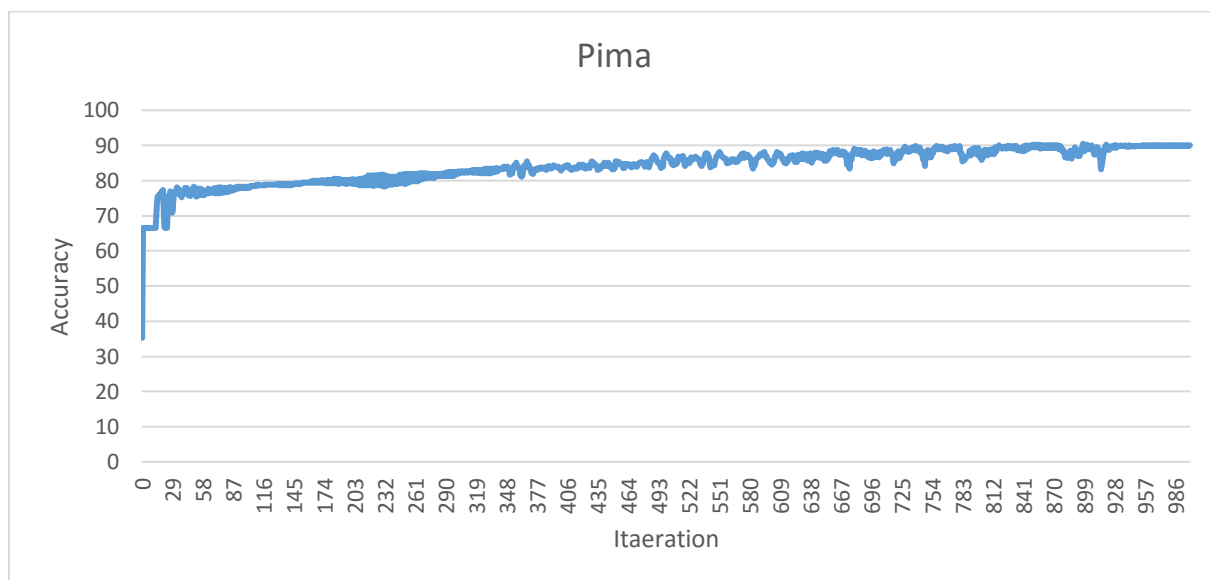
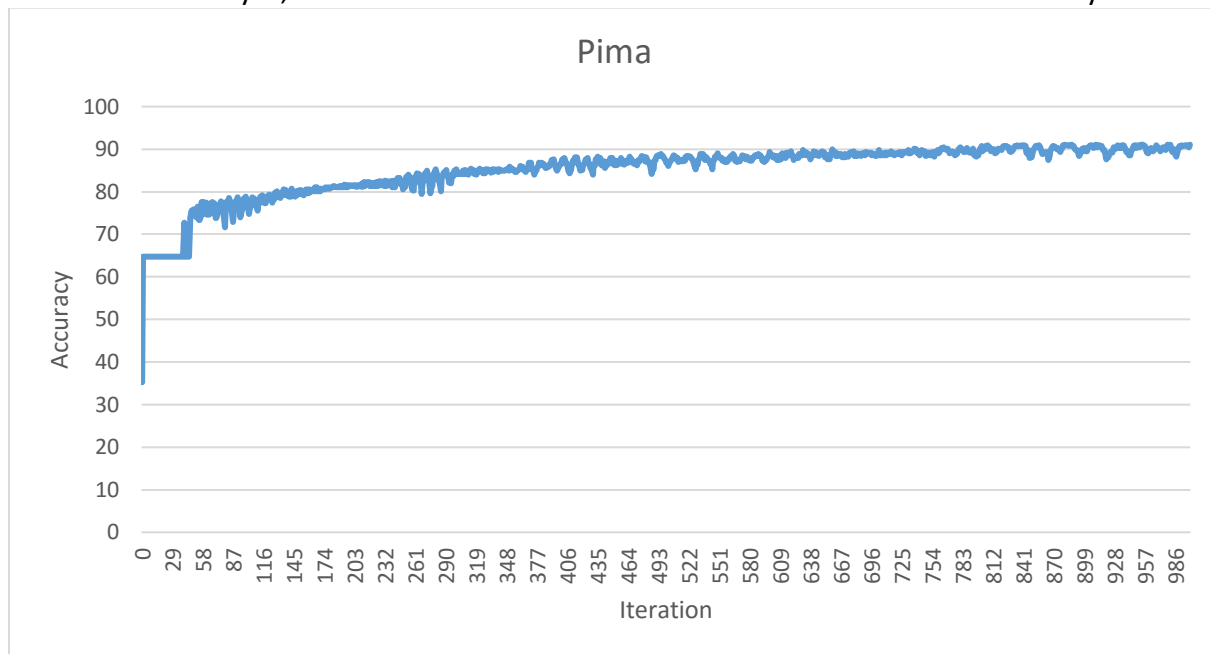


Compare your results on the Iris dataset to those of an existing implementation.

Could not get an existing implementation to work.

Produce at least one graph to show the training progress for the Diabetes dataset.

The data was preprocessed with the StandardScaler, and the suggestions from the book were followed to 'bucket' the ages, and make the max number of pregnancies an eight. More preprocessing would probably increase the accuracy. The first graph shows 1001 iterations, and two hidden layers with 8 nodes in the first and 6 in the second, and the validation set returned with 81% accuracy. The second graph shows the same number of iterations, but with 8 nodes in the first hidden layer, and 3 in the second. The validation set returned 78% accuracy.



Compare your results on the Diabetes dataset to those of an existing implementation.

Could not get one to work.

Describe any efforts you made to go above and beyond.

I went to try and get the adults data set to work. But I could not figure out how to preprocess the data. I think one of the problems was that there were too many rows, and if I had more time I probably would get it to work.

Please state which category you feel best describes your assignment and give a 1-2 sentence justification for your choice: A) Some attempt was made, B) Developing, but significantly deficient, C) Slightly deficient, but still mostly adequate, D) Meets requirements, E) Shows creativity and excels above and beyond requirements.

E. Because I experimented with additional data sets. I tried using the adults, poker, and bank data sets. I was able to preprocess the bank data set the best, also it was still giving me a few errors. And after experimenting with it I discovered it was pretty hard to get high accuracy with that data set.