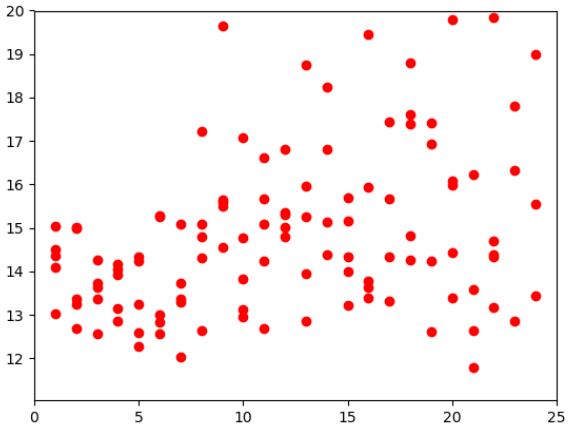
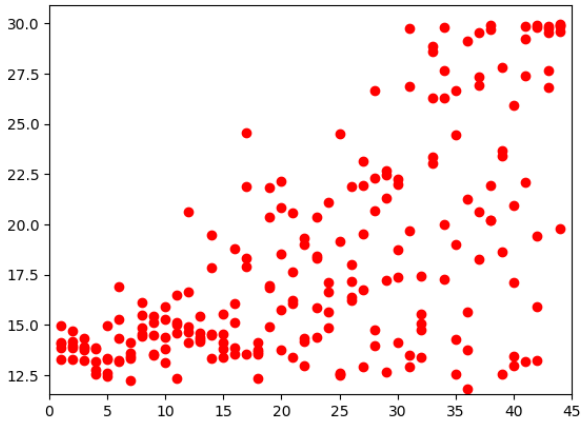
Ross Spencer

HW 06:
For my hidden layer, I decided upon 4 hidden nodes based off the following graphs.





The x-axis is how many nodes are in the hidden layer, and the y-axis is the error at each of 5 passes of the training per number of nodes (hence the 5 dots at each x value), with the learning rate fixed at .1 and the number of iterations fixed at 10. While some beyond x=4 hidden nodes had lower error at their best iteration, their higher worst cases outweighed the benefit and increased time.

The resulting confusion matrix for 4 hidden nodes with the learning rate fixed at .1 and the iterations fixed at 10 is as follows:

	0	1
0	40	15
1	20	45

Percentage Correct: 70.83333333333334
We have 40 correctly classified 0's, 45 correctly classified 1's, 15 incorrectly classified 0's, and 20 incorrectly classified 1's, with a 70% accuracy in most runs of the code.

With 4 nodes in the hidden layer and 10 iterations fixed, I tried the following learning rates and got the corresponding accuracy score from the confusion matrix method:

Learning Rate	Accuracy
0	33.33333333333333
0.001	50.0
0.01	50.0
0.05	65.83333333333333
0.1	79.16666666666666
0.2	50.0
0.3	50.0

Learning Rate	Accuracy
0.4	50.0
0.5	50.0
0.6	50.0
0.7	50.0
0.8	50.0
0.9	50.0
1.0	50.0
2.0	50.0

For changing the learning rate, alpha=0 did terribly, alpha=.1 did better than guessing, and the rest did terribly, with special consideration to alpha=0.05 for being slightly-less-terrible on this particular iteration. Similar passes of the code returned similar results. As a result, .1 appears to be the best learning rate.

For one pass with the following number of iterations, with 4 hidden nodes and a learning rate of .1, the accuracy score from the confusion matrix method looked as follows:

Number of Iterations	Accuracy
1	50
5	65
10	54.1666666
15	66.6666666
20	66.6666666
50	83.33333333
100	83.33333333
500	83.33333333
1000	100.0

Clearly at 1000 iterations the neural network is overfitted for our 120 data points. 50 seems to provide a decent balance compared to another number of iterations.

So, one time running my resulting heuristically-optimal neural net uses 4 nodes in the hidden layer, a learning rate of 0.1, and 50 iterations of back-propagation gave the following confusion matrix:

```
Confusion matrix is:
[[40.  0.]
 [20. 60.]]
Percentage Correct:  83.33333333333334
```