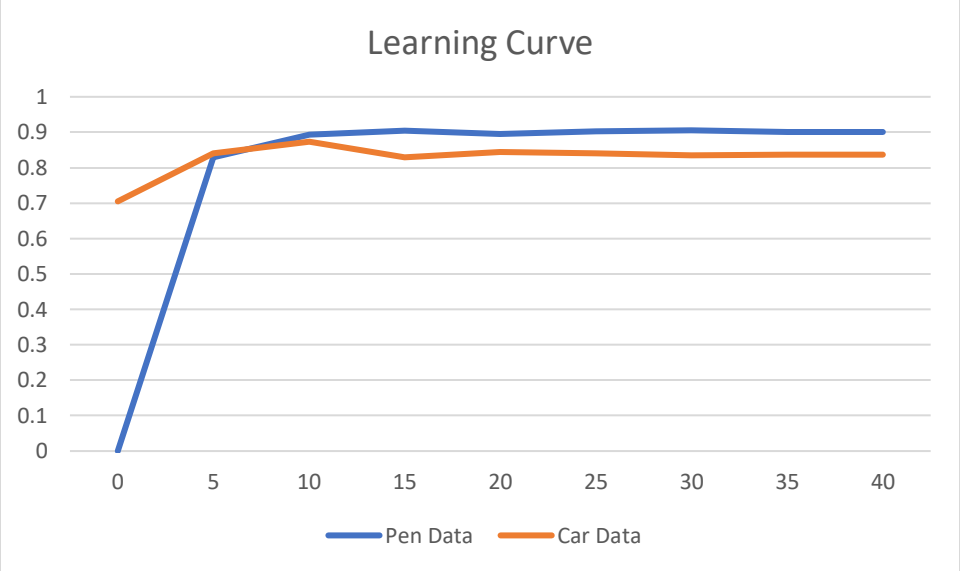
5.

| | Max | Mean | Std. Dev. |
|---|---|---|---|
| testPenData | 0.907661520869068 | 0.901658090337 | 0.00553452031576 |
| testCarData | 0.899214659686 | 0.893848167539 | 0.00395712472718 |

6.

| | Max | Mean | Std. Dev |
|---|---|---|---|
| testPenData 0 layers | 0 | 0 | 0 |
| testCarData 0 layers | 0.704842931937 | 0.704842931937 | 0 |
| testPenData 5 layers | 0.844196683819 | 0.828702115495 | 0.0123708079828 |
| testCarData 5 layers | 0.85667539267 | 0.840314136126 | 0.0168385868197 |
| testPenData 10 layers | 0.902801600915 | 0.892853058891 | 0.00819507049385 |
| testCarData 10 layers | 0.888089005236 | 0.873298429319 | 0.0106544824214 |
| testPenData 15 layers | 0.933676386507 | 0.904402515723 | 0.0147337688124 |
| testCarData 15 layers | 0.870418848168 | 0.828926701571 | 0.0284594828075 |
| testPenData 20 layers | 0.903087478559 | 0.895254431103 | 0.00617548649871 |
| testCarData 20 layers | 0.857329842932 | 0.844240837696 | 0.0119461824179 |
| testPenData 25 layers | 0.906803887936 | 0.903201829617 | 0.00232669982268 |
| testCarData 25 layers | 0.86387434555 | 0.839790575916 | 0.018251537187 |
| testPenData 30 layers | 0.907947398513 | 0.905488850772 | 0.00175855420231 |
| testCarData 30 layers | 0.857329842932 | 0.8349643979 | 0.0162081844704 |
| testPenData 35 layers | 0.904516866781 | 0.90051457976 | 0.00433554913897 |
| testCarData 35 layers | 0.846204188482 | 0.837303664921 | 0.00623069645833 |
| testPenData 40 layers | 0.903087478559 | 0.900228702115 | 0.00343529305564 |
| testCarData 40 layers | 0.852748691099 | 0.835994764398 | 0.0108930095663 |




The number of hidden layer perceptrons (x-axis) appeared to have a minimal effect on the average accuracy (y-axis) in both datasets. Both accuracies saw an increase between five and ten layers, but past that the values remained about the same. In general, I would expect the accuracies to increase with a higher number of hidden layers since more information about the data can be stored, but in this particular run of testing that did not occur.

7. Test cases for XOR are (0, 0, 0), (0, 1, 1), (1, 0, 1), and (1, 1, 0), as seen in xor.data.txt. Testing XOR with 0 hidden layers produced the following output after five runs: [0.5, 0.0, 0.5, 0.5, 0.0], which is not particularly accurate. I incremented the number of hidden layer perceptrons by 1 until the average of the five runs exceeded 75% accuracy, which only occurred with 26 perceptrons. This seems quite unusual, since the neural network can theoretically classify the data points perfectly with just two perceptrons.