Question 5 Learning With Restarts

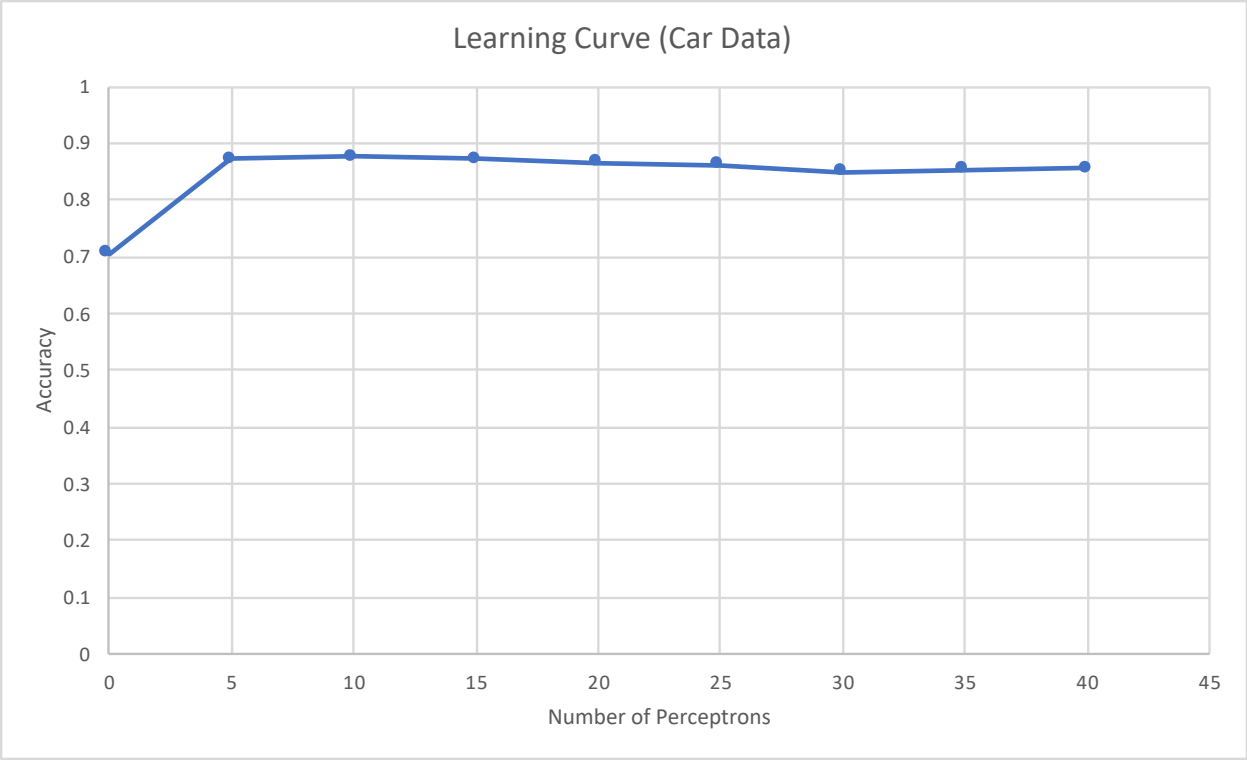
('Pen Data Average Accuracy: ', 0.904802744425386)
('Pen Data Max Accuracy: ', 0.9068038879359634)
('Pen Data St Dev Accuracy: ', 0.0021771792755471524)
[0.9068038879359634, 0.906232132647227, 0.9042309891366496, 0.900800457404231, 0.9059462550028587]
('Car Data Average Accuracy: ', 0.8407068062827225)
('Car Data Max Accuracy: ', 0.8592931937172775)
('Car Data St Dev Accuracy: ', 0.01328775801383297)
[0.8592931937172775, 0.8462041884816754, 0.818717277486911, 0.8363874345549738, 0.8429319371727748]

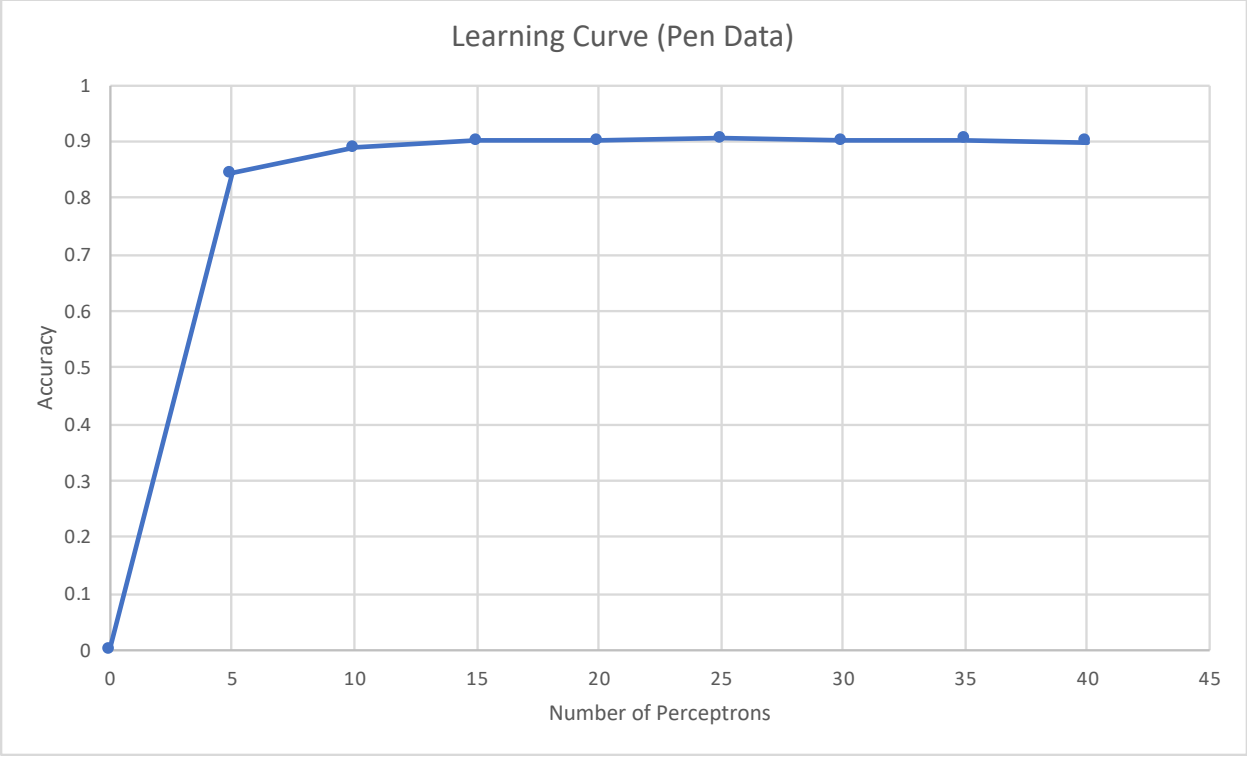
(this data can also be found in q5_output.txt)

Question 6 Varying the Hidden Layer

Car Data			
# Perceptrons	Max	Average	Standard Deviation
0	0.70484293	0.70484293	0
5	0.89070681	0.87159686	0.015215933
10	0.89528796	0.87670157	0.012801823
15	0.89136126	0.87251309	0.010804575
20	0.87434555	0.86531414	0.00892935
25	0.87696335	0.86269634	0.01165877
30	0.85994764	0.84842932	0.00980016
35	0.85863874	0.85340314	0.004571793
40	0.8645288	0.85589005	0.005418911

Pen Data			
# Perceptrons	Max	Average	Standard Deviation
0	0	0	0
5	0.85620354	0.84425386	0.009244772
10	0.9002287	0.88782161	0.007057415
15	0.90623213	0.9002287	0.004616725
20	0.9053745	0.90154374	0.004840036
25	0.90880503	0.90548885	0.001927078
30	0.90766152	0.90137221	0.006223997
35	0.90423099	0.90240137	0.001486564
40	0.90251572	0.89908519	0.003848215





The trends I noticed in these graphs is that as you increase the increase the number of hidden layers the accuracy quickly converges. In both instances the accuracy, more or less, converged at around 5 hidden layers. This shows that, after around 5 hidden layers, there is a diminishing

return in terms of accuracy gained versus time taken to run the algorithm - as adding more and more hidden layers means the model takes that much longer to train and test.

Question 7

Learning XOR

(data is present in q7_analysis.txt)

XOR is a non-linear function so it is quite difficult for a plain neural network to learn this function. I knew that increasing the number of neurons per hidden layers would increase the accuracy across the training and testing set. This was evident in the fact that after adding 43 neurons per hidden layer my neural network completely converged.

These results are what I expected, in that adding more neurons per hidden layer would cause the model to converge.

When training the data on a model with no hidden layers, I am getting an accuracy at around 25% with a max accuracy of around 50%. This is a little higher than what I would expect, but it still makes sense since training on this data with no hidden layers would result in around 1/4th of testing data result in correct predictions.