Question 5 (4 points): Learning With Restarts

**Pen Data Accuracy**:

|  |  |  |
| --- | --- | --- |
| Max | Average | Standard Deviation |
| 0.9033733562035449 | 0.899199542595769 | 0.005556920212609628 |

**Car Data Accuracy:**

|  |  |  |
| --- | --- | --- |
| Max | Average | Standard Deviation |
| 0.985 | 0.977 | 0.006782329983125274 |

Question 6 (4 points): Varying The Hidden Layer

**Pen Data Accuracy**:

|  |  |  |  |
| --- | --- | --- | --- |
| # of perceptron | Average | Max | Standard Deviation |
| 0 | 0.0 | 0.0 | 0.0 |
| 5 | 0.8360205831903945 | 0.8487707261292167 | 0.009461815873226587 |
| 10 | 0.8886792452830189 | 0.89937106918239 | 0.009243357519936529 |
| 15 | 0.9068610634648371 | 0.9442538593481989 | 0.020111974833727224 |
| 20 | 0.9025157232704402 | 0.9105202973127502 | 0.007313105525589252 |
| 25 | 0.9018867924528301 | 0.907661520869068 | 0.00610306925374176 |
| 30 | 0.9049170954831333 | 0.9073756432246999 | 0.0021590860592909373 |
| 35 | 0.9053173241852488 | 0.9073756432246999 | 0.001975660668172796 |
| 40 | 0.9044596912521442 | 0.9056603773584906 | 0.0009635391392997614 |

**Car Data Accuracy**:

|  |  |  |  |
| --- | --- | --- | --- |
| # of perceptron | Average | Max | Standard Deviation |
| 0 | 0.745 | 0.745 | 0.0 |
| 5 | 0.9640000000000001 | 0.975 | 0.009165151389911688 |
| 10 | 0.978 | 0.995 | 0.012083045973594584 |
| 15 | 0.9789999999999999 | 0.985 | 0.0037416573867739447 |
| 20 | 0.9810000000000001 | 0.99 | 0.007348469228349541 |
| 25 | 0.9829999999999999 | 0.99 | 0.00509901951359279 |
| 30 | 0.9800000000000001 | 0.995 | 0.011832159566199244 |
| 35 | 0.9789999999999999 | 0.98 | 0.0020000000000000018 |
| 40 | 0.978 | 0.985 | 0.007483314773547889 |

As we can see in the chart, the neuro network accuracy for both datasets starts to increase when the number of perceptrons in the hidden layer increases. While the accuracy of the NN trained on both datasets significantly increases from 0 perceptrons to 5 perceptrons, the accuracy only slightly increases after having 5 perceptrons to more and eventually fluctuates as the number of perceptrons continues to increase. It is also worth noting that the NN trained on Pen data does not perform when there is zero hidden layer perceptron, while the NN trained on Car data does perform acceptably well when there is zero hidden layer perceptron.

Question 7 (2 points Extra Credit): Learning XOR

When I tried to run the NN with 0 hidden layers the training error and weight change basically does not change, as expected, but the final average accuracy over 5 rounds of training was 0.4 which is greater than 0. Then I ran the NN with weight change threshold = 0.0001 and max iteration = 2000 and the number of hidden layer perceptrons increasing starting from 1 until the average accuracy become 1.0, the result is quite as expected: the average accuracy increases when we add more hidden layer perceptrons.

|  |  |
| --- | --- |
| Number of hidden layer perceptrons | Average accuracy |
| 1 | 0.500000 |
| 2 | 0.650000 |
| 3 | 0.600000 |
| 4 | 0.600000 |
| 5 | 0.750000 |
| 6 | 0.750000 |
| 7 | 0.800000 |
| 8 | 0.900000 |
| 9 | 0.900000 |
| 10 | 0.850000 |
| 11 | 0.950000 |
| 12 | 1.000000 |

When I went ahead to increase the max number of perceptrons to a large number, in this case 50, I can observe the average accuracy increases rather steadily and stays to be 1.0 when the number of perceptrons is above 15 (I will not include the results after running 50 times to avoid being lengthy.)