STAT 502 HW 4

Name: Kentaro Hoffman

**Parameter values**

Number of Perceptrons in Hidden Layer = 10

Slope of Transfer function = 1

Learning Rate = 0.0005

Epoch Size = 200

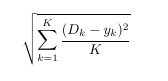
Maximum number of learning Steps = 600,000

Size of Training Set = 200

Size of Testing Set = 100

Minimum RMSE = 0.05

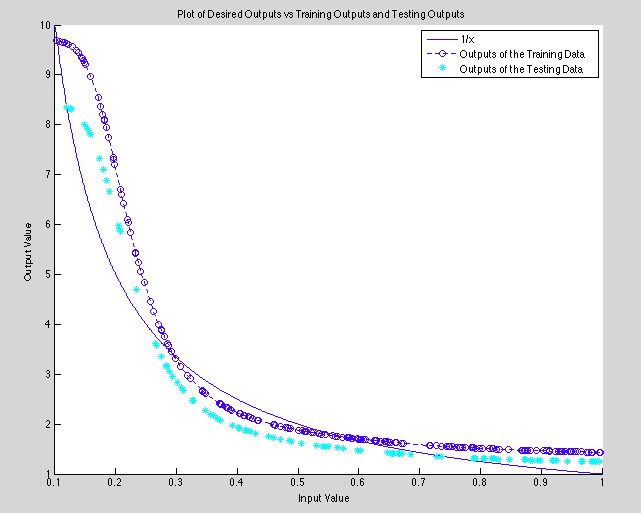
At every 30,000 iterations of the code, the RMSE of the code was calculated using the following formula:



Where K is the number of training data inputs, 200, D\_k is the desired output of the kth element of the training data. And y\_k is the output of the perceptron network for the kth input. The table and the regression line below show that as the simulation goes on, the RMSE of the training data seems to be decreasing, a sign that the neural network is correctly training on the testing data.

|  |  |  |
| --- | --- | --- |
| Iteration Number | Training RMSE | Testing RMSE |
| 30,000 | 0.2433 | 0.2922 |
| 60,000 | 0.1217 | 0.129 |
| 90,000 | 0.2441 | 0.2981 |
| 120,000 | 0.2373 | 0.1968 |
| 150,000 | 0.1913 | 0.14 |
| 180,000 | 0.2354 | 0.2678 |
| 210,000 | 0.0874 | 0.0823 |
| 240,000 | 0.1599 | 0.2989 |
| 270,000 | 0.1376 | 0.12156 |
| 300,000 | 0.0885 | 0.2857 |
| 330,000 | 0.2405 | 0.2562 |
| 360,000 | 0.0975 | 0.0857 |
| 390,000 | 0.2232 | 0.2211 |
| 420,000 | 0.0965 | 0.098 |
| 450,000 | 0.1706 | 0.2106 |
| 480,000 | 0.1575 | 0.13 |
| 510,000 | 0.2324 | 0.2324 |
| 540,000 | 0.17 | 0.179 |
| 570,000 | 0.1053 | 0.134 |
| 600,000 | 0.0935 | 0.09 |

From the plot above, we can see that the RMSE of both the training and testing are decreasing. In fact, although the testing RMSE tends to have a wider standard deviation, from the fact that they have similar linear regression lines, it seems likely that the training and testing responses should be similar, a fact that is confirmed in the next plot.

However, with good RMSE for the training data alone, it is hard to say whether the neural network is over fitting on the training data or not. But with the graph below, we can say that this does not seem to be the case. The graph of outputs for the Testing Data is very similar to the output of the Training data. Both of which are also similar to the graph of the true desired output, 1/x. Moreover, the RMSE of the training and testing data are nearly identical, indicating that the current neural network is producing outputs that easily generalize to other types of testing data.

|  |  |  |
| --- | --- | --- |
|  | RMSE | Standard Deviation |
| Training | 0.0994 | 1.8290 |
| Testing | 0.1097 | 1.3352 |