



NUS

National University
of Singapore

EE4305 INTRODUCTION TO FUZZY/NEURAL SYSTEM
PART II PROJECT I

Classification of Cardboard Papers Using a Multilayer Perceptron

Name: PHANG Swee King
Matric Number: U066584J
Email: king@nus.edu.sg

November 8, 2009

1 Objectives

1. Competence in setting up and training a multilayer Perceptron (MLP) using the error-backpropagation algorithm.
2. Understanding of the principles and issues of training and testing of a MLP.

2 Project Description

In this project, a MLP is to be trained to classify five types of cardboard papers based on their odors. Detecting odors in cardboard papers is an important issue in the packaging industry. It is an essential requirement that the material used for packaging does not introduce unwanted odors into the product (e.g. foodstuffs) contained in the package. The odor of cardboard paper is measured by a hybrid gas array sensor (a.k.a. an electronic nose). The odors from five different types of cardboard paper from commercial manufacturers were recorded with the electronic nose. The MLP will be trained to determine the origin of an unknown sample based on the data measured by the electronic nose.

3 Discussions

3.1 Learning Process

The assigned initial weights and learning rates play an important role in how quickly a successful network training is obtained. Using a different weight and learning rate, the number of epochs required may vary from a few hundreds to more than ten thousands. The learning process will repeats, epoch by epoch until a point where it converges to a local minima or global minimum. In general, there are a few possible way to end the training. This happens when the absolute rate of change in mean squared error per epoch is sufficiently small, synaptic weights stabilizes and mean squared error of an epoch is less than a given minimum value.

3.1.1 Initial Weight

In this program, the initial weight values are selected by the program itself. However, we are able to assigned a range of the weight value for the starting point of training process. The weight used each time we start the learning process will be randomly chosen within the specified range. Thus, the outcome of training varies in each trial. Once the correct range of weight is selected, it is possible to converge to a global minimum, which is the target of this project.

3.1.2 Learning Rates

Learning rate affects speed of gradient descent towards a minimum, both local minimum and global minimum. Small learning rates result in longer time to convergence. However, large leaning rates will result in undesired oscillatory behaviour in the process. It may lead to instability of learning process.

3.2 Limitations

Ideally, we want to find optimal weight such that the learning process will converge to a global minimum. However, error-backpropagation algorithm does not guarantee convergence to global

minimum. Hence, it is possible to get trapped in a local minimum; in which case the Total SSE field will show a value that remain approximately same value after many epochs have been completed. Gradient descent (if convergent) only guarantees reduction of network output error down to a minimum, but does not guarantee that such minimum is global.

3.3 Methods to Overcome Limitations

To avoid learning process got trapped in local minimum, a few ways can be practiced to overcome the problems. However, it does not guarantee a successful result.

1. Changing the weight value may help the learning process to leap out from the local minimum and proceed to converge to a next minimum point.
2. Continue to train until it loop out from the local minimum. However, this process takes much longer time to complete.

4 Results

4.1 10 Hidden Layers

Training Result: Successful
 Final Sweeps of Training: 2711
 Final Weight Range: 0 to 0.21
 Final Learning Rate: 0.02
 Final Squared Error Value: 0.01

Table 1: Final Training Result for 10 Hidden Layers

Results Of Training With File	TrainingData48.net
Date of Analysis	11/7/2009
Time of Analysis	12:53:23 PM
Type Of Network	IntDev
Learning Rate	0.02
Sweeps Of Training	2711
Hits	288
Misses	0
Minimum Squared Error Defining A Hit	0.01
Weight Start Settings	
Maximum	0.21
Minimum	0
Sign Option	Both
Bias Start Settings	
Maximum	0
Minimum	0
Sign Option	Both

The trained network is then tested with the following 10 data. The results of the testing is shown below.

Pattern: 10 +19.50 +16.60 +2.40 +27.30 +18.00 +10.70 +11.70 +4.80
+9.70 +20.50 +2.90 +21.00 +8.30 +7.30 +3.40

Response (A): +.07 +.05 +.00 +.90 +.04 +.00
Desired (D): +.00 +.00 +.00 +1.00 +.00 +.00
Error (D-A): -.07 -.05 +.00 +.10 -.04 +.00

=====

Pattern: 9 +11.70 +11.20 +2.40 +15.60 +10.20 +19.50 +6.30 +5.30
+6.80 +11.70 +2.90 +7.30 +3.40 +3.40 +2.90

Response (A): +.05 +.00 +.89 +.00 +.05 +.03
Desired (D): +.00 +.00 +1.00 +.00 +.00 +.00
Error (D-A): -.05 +.00 +.11 +.00 -.05 -.03

=====

Pattern: 8 +17.10 +15.60 +2.40 +26.80 +17.10 +12.20 +10.70 +2.90
+7.30 +20.00 +2.90 +22.40 +9.20 +8.80 +3.40

Response (A): +.01 +.94 +.00 +.07 +.01 +.01
Desired (D): +.00 +1.00 +.00 +.00 +.00 +.00
Error (D-A): -.01 +.06 +.00 -.07 -.01 -.01

=====

Pattern: 7 +17.50 +15.10 +2.90 +25.80 +16.10 +9.20 +11.20 +3.40
+7.30 +19.00 +2.90 +17.10 +7.30 +6.30 +3.40

Response (A): +.07 +.05 +.00 +.90 +.04 +.00
Desired (D): +1.00 +.00 +.00 +.00 +.00 +.00
Error (D-A): +.93 -.05 +.00 -.90 -.04 +.00

=====

Pattern: 6 +2.40 +1.90 +1.90 +2.90 +1.40 +12.70 +1.90 +1.90
+2.40 +1.90 +2.90 +2.90 +2.40 +2.40 +2.90

Response (A): +.01 +.02 +.03 +.00 +.01 +.97
Desired (D): +.00 +.00 +.00 +.00 +.00 +1.00
Error (D-A): -.01 -.02 -.03 +.00 -.01 +.03

=====

Pattern: 5 +18.50 +15.60 +2.90 +26.30 +12.20 +12.20 +8.80 +7.80
+9.20 +20.00 +2.90 +8.30 +4.40 +3.40 +6.30

Response (A): +.00 +.00 +.07 +.05 +.93 +.01
Desired (D): +.00 +.00 +.00 +.00 +1.00 +.00
Error (D-A): +.00 +.00 -.07 -.05 +.07 -.01

=====

```

Pattern:  4 +20.00 +18.00 +1.90 +29.30 +17.50 +15.60 +13.10 +3.90
+8.80 +22.40 +2.40 +21.00 +8.30 +7.30 +3.40
-----
Response (A): +.07 +.05 +.00 +.90 +.04 +.00
Desired (D): +.00 +.00 +.00 +1.00 +.00 +.00
Error (D-A): -.07 -.05 +.00 +.10 -.04 +.00
=====

Pattern:  3 +13.10 +11.20 +1.90 +17.50 +10.70 +10.20 +6.80 +5.30
+6.80 +13.10 +2.90 +7.80 +3.90 +2.90 +2.90
-----
Response (A): +.03 +.00 +.86 +.00 +.10 +.02
Desired (D): +.00 +.00 +1.00 +.00 +.00 +.00
Error (D-A): -.03 +.00 +.14 +.00 -.10 -.02
=====

Pattern:  2 +19.00 +17.10 +2.40 +30.70 +17.10 +15.10 +11.70 +3.90
+7.80 +22.40 +2.90 +22.40 +9.20 +8.80 +2.90
-----
Response (A): +.01 +.94 +.00 +.08 +.01 +.01
Desired (D): +.00 +1.00 +.00 +.00 +.00 +.00
Error (D-A): -.01 +.06 +.00 -.08 -.01 -.01
=====

Pattern:  1 +18.00 +16.10 +2.40 +27.30 +15.10 +16.10 +11.20 +3.90
+7.30 +20.00 +2.90 +17.50 +7.80 +6.80 +3.90
-----
Response (A): +.81 +.02 +.02 +.14 +.00 +.00
Desired (D): +1.00 +.00 +.00 +.00 +.00 +.00
Error (D-A): +.19 -.02 -.02 -.14 +.00 +.00
=====

```

As we can see, besides pattern 7, the other patterns error are considerably small within ± 0.19 range. Therefore, the performance of the trained network in classifying the samples in the test data set is considered acceptable.

4.2 15 Hidden Layers

Training Result: Successful
Final Sweeps of Training: 19103
Final Weight Range: 0 to 0.42
Final Learning Rate: 0.02
Final Squared Error Value: 0.01

Table 2: Final Training Result for 15 Hidden Layers

Results Of Training With File	TrainingData48.net
Date of Analysis	11/7/2009
Time of Analysis	1:34:25 PM
Type Of Network	IntDev
Learning Rate	0.02
Sweeps Of Training	19103
Hits	60
Misses	0
Minimum Squared Error Defining A Hit	0.01
Weight Start Settings	
Maximum	0.42
Minimum	0
Sign Option	Both
Bias Start Settings	
Maximum	0
Minimum	0
Sign Option	Both

The trained network is then tested with the following 10 data. The results of the testing is shown below.

Pattern: 10 +19.50 +16.60 +2.40 +27.30 +18.00 +10.70 +11.70 +4.80 +9.70 +20.50 +2.90 +21.00 +8.30 +7.30 +3.40 ----- Response (A): +.05 +.02 +.02 +.94 +.00 +.01 Desired (D): +.00 +.00 +.00 +1.00 +.00 +.00 Error (D-A): -.05 -.02 -.02 +.06 +.00 -.01 =====
Pattern: 9 +11.70 +11.20 +2.40 +15.60 +10.20 +19.50 +6.30 +5.30 +6.80 +11.70 +2.90 +7.30 +3.40 +3.40 +2.90 ----- Response (A): +.01 +.00 +.97 +.03 +.02 +.01 Desired (D): +.00 +.00 +1.00 +.00 +.00 +.00 Error (D-A): -.01 +.00 +.03 -.03 -.02 -.01 =====

Pattern: 8 +17.10 +15.60 +2.40 +26.80 +17.10 +12.20 +10.70 +2.90
+7.30 +20.00 +2.90 +22.40 +9.20 +8.80 +3.40

Response (A): +.03 +.96 +.00 +.01 +.01 +.01
Desired (D): +.00 +1.00 +.00 +.00 +.00 +.00
Error (D-A): -.03 +.04 +.00 -.01 -.01 -.01
=====

Pattern: 7 +17.50 +15.10 +2.90 +25.80 +16.10 +9.20 +11.20 +3.40
+7.30 +19.00 +2.90 +17.10 +7.30 +6.30 +3.40

Response (A): +.90 +.04 +.00 +.06 +.03 +.00
Desired (D): +1.00 +.00 +.00 +.00 +.00 +.00
Error (D-A): +.10 -.04 +.00 -.06 -.03 +.00
=====

Pattern: 6 +2.40 +1.90 +1.90 +2.90 +1.40 +12.70 +1.90 +1.90
+2.40 +1.90 +2.90 +2.90 +2.40 +2.40 +2.90

Response (A): +.00 +.02 +.01 +.01 +.01 +.98
Desired (D): +.00 +.00 +.00 +.00 +.00 +1.00
Error (D-A): +.00 -.02 -.01 -.01 -.01 +.02
=====

Pattern: 5 +18.50 +15.60 +2.90 +26.30 +12.20 +12.20 +8.80 +7.80
+9.20 +20.00 +2.90 +8.30 +4.40 +3.40 +6.30

Response (A): +.05 +.02 +.02 +.00 +.96 +.00
Desired (D): +.00 +.00 +.00 +.00 +1.00 +.00
Error (D-A): -.05 -.02 -.02 +.00 +.04 +.00
=====

Pattern: 4 +20.00 +18.00 +1.90 +29.30 +17.50 +15.60 +13.10 +3.90
+8.80 +22.40 +2.40 +21.00 +8.30 +7.30 +3.40

Response (A): +.05 +.02 +.02 +.94 +.00 +.01
Desired (D): +.00 +.00 +.00 +1.00 +.00 +.00
Error (D-A): -.05 -.02 -.02 +.06 +.00 -.01
=====

Pattern: 3 +13.10 +11.20 +1.90 +17.50 +10.70 +10.20 +6.80 +5.30
+6.80 +13.10 +2.90 +7.80 +3.90 +2.90 +2.90

Response (A): +.01 +.00 +.97 +.03 +.02 +.01
Desired (D): +.00 +.00 +1.00 +.00 +.00 +.00
Error (D-A): -.01 +.00 +.03 -.03 -.02 -.01
=====

```

Pattern:  2 +19.00 +17.10 +2.40 +30.70 +17.10 +15.10 +11.70 +3.90
+7.80 +22.40 +2.90 +22.40 +9.20 +8.80 +2.90
-----
Response (A): +.06 +.96 +.00 +.00 +.02 +.01
Desired (D): +.00 +1.00 +.00 +.00 +.00 +.00
Error (D-A): -.06 +.04 +.00 +.00 -.02 -.01
=====

Pattern:  1 +18.00 +16.10 +2.40 +27.30 +15.10 +16.10 +11.20 +3.90
+7.30 +20.00 +2.90 +17.50 +7.80 +6.80 +3.90
-----
Response (A): +.98 +.01 +.02 +.00 +.00 +.00
Desired (D): +1.00 +.00 +.00 +.00 +.00 +.00
Error (D-A): +.02 -.01 -.02 +.00 +.00 +.00
=====

```

In this testing result, the biggest error that occur was just 0.10. Therefore, the performance of the trained network in classifying the samples in the test data set is considered very good. This result is predictable as compare this network with the trained network with lesser hidden layers (10 layers), in theory, would perform better.

4.3 5 Hidden Layers

Training Result: Successful
Final Sweeps of Training: 42000
Final Weight Range: 0.75 to 1
Final Learning Rate: 10
Final Squared Error Value: 0.02

Table 3: Final Training Result for 5 Hidden Layers

Results Of Training With File	TrainingData48.net
Date of Analysis	11/7/2009
Time of Analysis	3:32:09 PM
Type Of Network	IntDev
Learning Rate	10
Sweeps Of Training	42000
Hits	288
Misses	0
Minimum Squared Error Defining A Hit	0.02
Weight Start Settings	
Maximum	1
Minimum	0.75
Sign Option	Both
Bias Start Settings	
Maximum	0
Minimum	0
Sign Option	Both

The trained network is then tested with the following 10 data. The results of the testing is shown below.

```

Pattern:  10 +19.50 +16.60 +2.40 +27.30 +18.00 +10.70 +11.70 +4.80
+9.70 +20.50 +2.90 +21.00 +8.30 +7.30 +3.40
-----
Response (A): +.04 +.03 +.00 +.90 +.04 +.00
Desired (D): +.00 +.00 +.00 +1.00 +.00 +.00
Error (D-A): -.04 -.03 +.00 +.10 -.04 +.00
=====

Pattern:  9 +11.70 +11.20 +2.40 +15.60 +10.20 +19.50 +6.30 +5.30
+6.80 +11.70 +2.90 +7.30 +3.40 +3.40 +2.90
-----
Response (A): +.05 +.00 +.94 +.00 +.04 +.01
Desired (D): +.00 +.00 +1.00 +.00 +.00 +.00
Error (D-A): -.05 +.00 +.06 +.00 -.04 -.01
=====

```

Pattern: 8 +17.10 +15.60 +2.40 +26.80 +17.10 +12.20 +10.70 +2.90
+7.30 +20.00 +2.90 +22.40 +9.20 +8.80 +3.40

Response (A): +.00 +.95 +.00 +.06 +.02 +.04
Desired (D): +.00 +1.00 +.00 +.00 +.00 +.00
Error (D-A): +.00 +.05 +.00 -.06 -.02 -.04
=====

Pattern: 7 +17.50 +15.10 +2.90 +25.80 +16.10 +9.20 +11.20 +3.40
+7.30 +19.00 +2.90 +17.10 +7.30 +6.30 +3.40

Response (A): +.68 +.00 +.01 +.24 +.00 +.01
Desired (D): +1.00 +.00 +.00 +.00 +.00 +.00
Error (D-A): +.32 +.00 -.01 -.24 +.00 -.01
=====

Pattern: 6 +2.40 +1.90 +1.90 +2.90 +1.40 +12.70 +1.90 +1.90
+2.40 +1.90 +2.90 +2.90 +2.40 +2.40 +2.90

Response (A): +.12 +.05 +.01 +.00 +.00 +.96
Desired (D): +.00 +.00 +.00 +.00 +.00 +1.00
Error (D-A): -.12 -.05 -.01 +.00 +.00 +.04
=====

Pattern: 5 +18.50 +15.60 +2.90 +26.30 +12.20 +12.20 +8.80 +7.80
+9.20 +20.00 +2.90 +8.30 +4.40 +3.40 +6.30

Response (A): +.00 +.01 +.04 +.06 +.94 +.00
Desired (D): +.00 +.00 +.00 +.00 +1.00 +.00
Error (D-A): +.00 -.01 -.04 -.06 +.06 +.00
=====

Pattern: 4 +20.00 +18.00 +1.90 +29.30 +17.50 +15.60 +13.10 +3.90
+8.80 +22.40 +2.40 +21.00 +8.30 +7.30 +3.40

Response (A): +.04 +.03 +.00 +.90 +.04 +.00
Desired (D): +.00 +.00 +.00 +1.00 +.00 +.00
Error (D-A): -.04 -.03 +.00 +.10 -.04 +.00
=====

Pattern: 3 +13.10 +11.20 +1.90 +17.50 +10.70 +10.20 +6.80 +5.30
+6.80 +13.10 +2.90 +7.80 +3.90 +2.90 +2.90

Response (A): +.04 +.00 +.94 +.00 +.05 +.01
Desired (D): +.00 +.00 +1.00 +.00 +.00 +.00
Error (D-A): -.04 +.00 +.06 +.00 -.05 -.01
=====

```

Pattern:  2 +19.00 +17.10 +2.40 +30.70 +17.10 +15.10 +11.70 +3.90
+7.80 +22.40 +2.90 +22.40 +9.20 +8.80 +2.90
-----
Response (A): +.00 +.91 +.00 +.10 +.02 +.02
Desired (D): +.00 +1.00 +.00 +.00 +.00 +.00
Error (D-A): +.00 +.09 +.00 -.10 -.02 -.02
=====

Pattern:  1 +18.00 +16.10 +2.40 +27.30 +15.10 +16.10 +11.20 +3.90
+7.30 +20.00 +2.90 +17.50 +7.80 +6.80 +3.90
-----
Response (A): +.65 +.00 +.00 +.27 +.00 +.00
Desired (D): +1.00 +.00 +.00 +.00 +.00 +.00
Error (D-A): +.35 +.00 +.00 -.27 +.00 +.00
=====

```

In this testing result, the biggest error that occur was 0.32. However the results are consistent throughout the test data. Therefore, the performance of the trained network in classifying the samples in the test data set is considered acceptable. The process of training takes higher epoch to complete due to the limited number of hidden layers. Besides, the learning rate need to be assigned a higher value to higher up the training efficiency. However the high learning rate will cause undesired oscillatory behaviour of the system during training.

4.4 2 Hidden Layers

Training Result: Successful
Final Sweeps of Training: 1000
Final Weight Range: 0.75 to 0.9
Final Learning Rate: 200
Final Squared Error Value: 0.01

Table 4: Final Training Result for 2 Hidden Layers

Results Of Training With File	TrainingData48.net
Date of Analysis	11/7/2009
Time of Analysis	4:18:39 PM
Type Of Network	IntDev
Learning Rate	200
Sweeps Of Training	1000
Hits	240
Misses	48
Minimum Squared Error Defining A Hit	0.01
Weight Start Settings	
Maximum	0.9
Minimum	0.75
Sign Option	Both
Bias Start Settings	
Maximum	0
Minimum	0
Sign Option	Both

The trained network is then tested with the following 10 data. The results of the testing is shown below.

Pattern: 10 +19.50 +16.60 +2.40 +27.30 +18.00 +10.70 +11.70 +4.80 +9.70 +20.50 +2.90 +21.00 +8.30 +7.30 +3.40 ----- Response (A): +.20 +.20 +.20 +.20 +.20 +.01 Desired (D): +.00 +.00 +.00 +1.00 +.00 +.00 Error (D-A): -.20 -.20 -.20 +.80 -.20 -.01 =====
Pattern: 9 +11.70 +11.20 +2.40 +15.60 +10.20 +19.50 +6.30 +5.30 +6.80 +11.70 +2.90 +7.30 +3.40 +3.40 +2.90 ----- Response (A): +.18 +.18 +.18 +.18 +.18 +.01 Desired (D): +.00 +.00 +1.00 +.00 +.00 +.00 Error (D-A): -.18 -.18 +.82 -.18 -.18 -.01 =====

Pattern: 8 +17.10 +15.60 +2.40 +26.80 +17.10 +12.20 +10.70 +2.90
+7.30 +20.00 +2.90 +22.40 +9.20 +8.80 +3.40

Response (A): +.20 +.20 +.20 +.20 +.20 +.01
Desired (D): +.00 +1.00 +.00 +.00 +.00 +.00
Error (D-A): -.20 +.80 -.20 -.20 -.20 -.01
=====

Pattern: 7 +17.50 +15.10 +2.90 +25.80 +16.10 +9.20 +11.20 +3.40
+7.30 +19.00 +2.90 +17.10 +7.30 +6.30 +3.40

Response (A): +.20 +.20 +.20 +.20 +.20 +.01
Desired (D): +1.00 +.00 +.00 +.00 +.00 +.00
Error (D-A): +.80 -.20 -.20 -.20 -.20 -.01
=====

Pattern: 6 +2.40 +1.90 +1.90 +2.90 +1.40 +12.70 +1.90 +1.90
+2.40 +1.90 +2.90 +2.90 +2.40 +2.40 +2.90

Response (A): +.02 +.01 +.02 +.01 +.01 +.98
Desired (D): +.00 +.00 +.00 +.00 +.00 +1.00
Error (D-A): -.02 -.01 -.02 -.01 -.01 +0.02
=====

Pattern: 5 +18.50 +15.60 +2.90 +26.30 +12.20 +12.20 +8.80 +7.80
+9.20 +20.00 +2.90 +8.30 +4.40 +3.40 +6.30

Response (A): +.20 +.20 +.20 +.20 +.20 +.01
Desired (D): +.00 +.00 +.00 +.00 +1.00 +.00
Error (D-A): -.20 -.20 -.20 -.20 +.80 -.01
=====

Pattern: 4 +20.00 +18.00 +1.90 +29.30 +17.50 +15.60 +13.10 +3.90
+8.80 +22.40 +2.40 +21.00 +8.30 +7.30 +3.40

Response (A): +.20 +.20 +.20 +.20 +.20 +.01
Desired (D): +.00 +.00 +.00 +1.00 +.00 +.00
Error (D-A): -.20 -.20 -.20 +.80 -.20 -.01
=====

Pattern: 3 +13.10 +11.20 +1.90 +17.50 +10.70 +10.20 +6.80 +5.30
+6.80 +13.10 +2.90 +7.80 +3.90 +2.90 +2.90

Response (A): +.20 +.20 +.20 +.20 +.20 +.01
Desired (D): +.00 +.00 +1.00 +.00 +.00 +.00
Error (D-A): -.20 -.20 +.80 -.20 -.20 -.01
=====

```

Pattern:  2 +19.00 +17.10 +2.40 +30.70 +17.10 +15.10 +11.70 +3.90
+7.80 +22.40 +2.90 +22.40 +9.20 +8.80 +2.90
-----
Response (A): +.20 +.20 +.20 +.20 +.20 +.01
Desired (D): +.00 +1.00 +.00 +.00 +.00 +.00
Error (D-A): -.20 +.80 -.20 -.20 -.20 -.01
=====

Pattern:  1 +18.00 +16.10 +2.40 +27.30 +15.10 +16.10 +11.20 +3.90
+7.30 +20.00 +2.90 +17.50 +7.80 +6.80 +3.90
-----
Response (A): +.20 +.20 +.20 +.20 +.20 +.01
Desired (D): +1.00 +.00 +.00 +.00 +.00 +.00
Error (D-A): +.80 -.20 -.20 -.20 -.20 -.01
=====

```

In this testing result, clearly the trained network did not work well on the test data provided. The trained network is limited by the small number of hidden layers available to make adjustment. Hence this is not a good training result.

5 Conclusions

1. Network is essentially a black-box. It does not provide intuitive explanation for computed result because what can be learned are operational parameters, not general abstract knowledge of a domain.
2. There are two factors affecting the learning process, namely, the initial weight and the learning rates.
3. The number of hidden layers could affect the learning outcome of the network.
4. Error-back propagation algorithm does not guarantee convergence to global minimum.