Chapter 4

Experimental Result

We test our hypothesis with four data sets which each has its own characteristic for classification on Gaussian distribution. The first data set is for illustrating the classification in the concave area with three Gaussian distributions. The second data set is for illustrating the classification of a data set lying in the space of a zigzagged channel with three Guassian distributions as well. The third data set is for illustrating the classification of a data set lying in circular channel space. And the fourth data set is for illustrating the general classification which has two data classes distributed on a 2-dimension space with Guassian distribution. The first three data sets are generated by each of the following equation.

Set 1:

$$(y - y_0) = \frac{b}{a}(x - x_0)^2 + G(0,c), x_u \le x \le x_v \tag{4.1}$$

Set 2:

$$y = \sin(x) + G(0,c) \tag{4.2}$$

Set 3:

$$(y - y_0) = \frac{b}{a}(x - x_0)^2 + G(0,c) \tag{4.3}$$

where $\{x_0, y_0\}$, $\{x_u, y_v\}$ and $\{a, b, c\}$ are three sets of parameters that determine the locations and shapes of the samples. The Guassian distribution is determined with $G(0,c)$. The fourth data set is created by a randomization with the Guassian distribution.

Some of the results from the training sets are shown in Figures 4.1 – 4.11. Each figure consists of three sub-figures, top, middle, and bottom. The top sub-figure shows the result after all data vectors are covered by some GERBF. The redundant GERBF are, then, pruned and all the remaining GERBF are demonstrated in the middle sub-figure. The last sub-figure illustrates the new GERBF after applying the Bootstrap to adjust the size.

The correct classification rate is the rate of the testing data, computed by the following equations.

$$\sigma = \frac{c}{n} \tag{4.4}$$

where σ is the correct classification rate, n is the number of data in testing data set, and c is the number of data in testing set which correctly classified by the trained network. The correctly classification data are the data that give the output from network greater than or equal to 1- err (which err is a small constant).

We compute the percent of correct classification rate and use this rate to show the result of our algorithm. The measurement of the rate is computed by counting the number of the testing data that the trained network can classify correctly and comparing it with the number of data vectors in the testing data set for creating the ratio. Then, we compute the percent of correct classification rate by multiplying with 100. For example, On Testset1_1, there are 52 patterns in first class and 52 patterns in another class. First of all, we create a smallest size (minimum number of hidden nodes) of GERBF neural network for this testset. Then, the testing process starts the testing with a set of testing data and, then, collects the numbers of correctly classified data and computes the percent of the correct classification rate by equation 4.4.

4.1 Testset1

This test set has two parabola shapes that are flipped horizontally and shifted vertically. This type of data set will show that the elliptic node can shrink and locate in the concave area. Tables 4.1 shows the comparison of the percentage of correct classification rate before and after generalization on each distribution. The result shows that if we have enough space for generalization (enlarging our hidden nodes) then the correct classification rate will increase. Figures 4.1 to 4.3 show the result of Testset1 that uses number of hidden node on each distribution. All data vectors are in a 2-dimensional space and are shown accordingly to their coordinates in x and y axes. The data vectors in training data set are represented by the symbol ".", the other class data vectors are represented by the symbol "*" (which are gray color), and the data vectors in testing data set are represented by symbol "*" (which are black color). The solid ellipses represent the GERBF nodes and the dashed ellipses represent the generalized GERBF nodes.

Testset	Number of hidden nodes	Average epochs	Correct Classification Rate (%) before Generalization	Correct Classification Rate (%) after Generalization
1_1	2	63	90.00	95.00
1_2	2	113	93.33	96.67
1_3	2	126	80.00	100.00

Table 4.1 The result of Testset 1

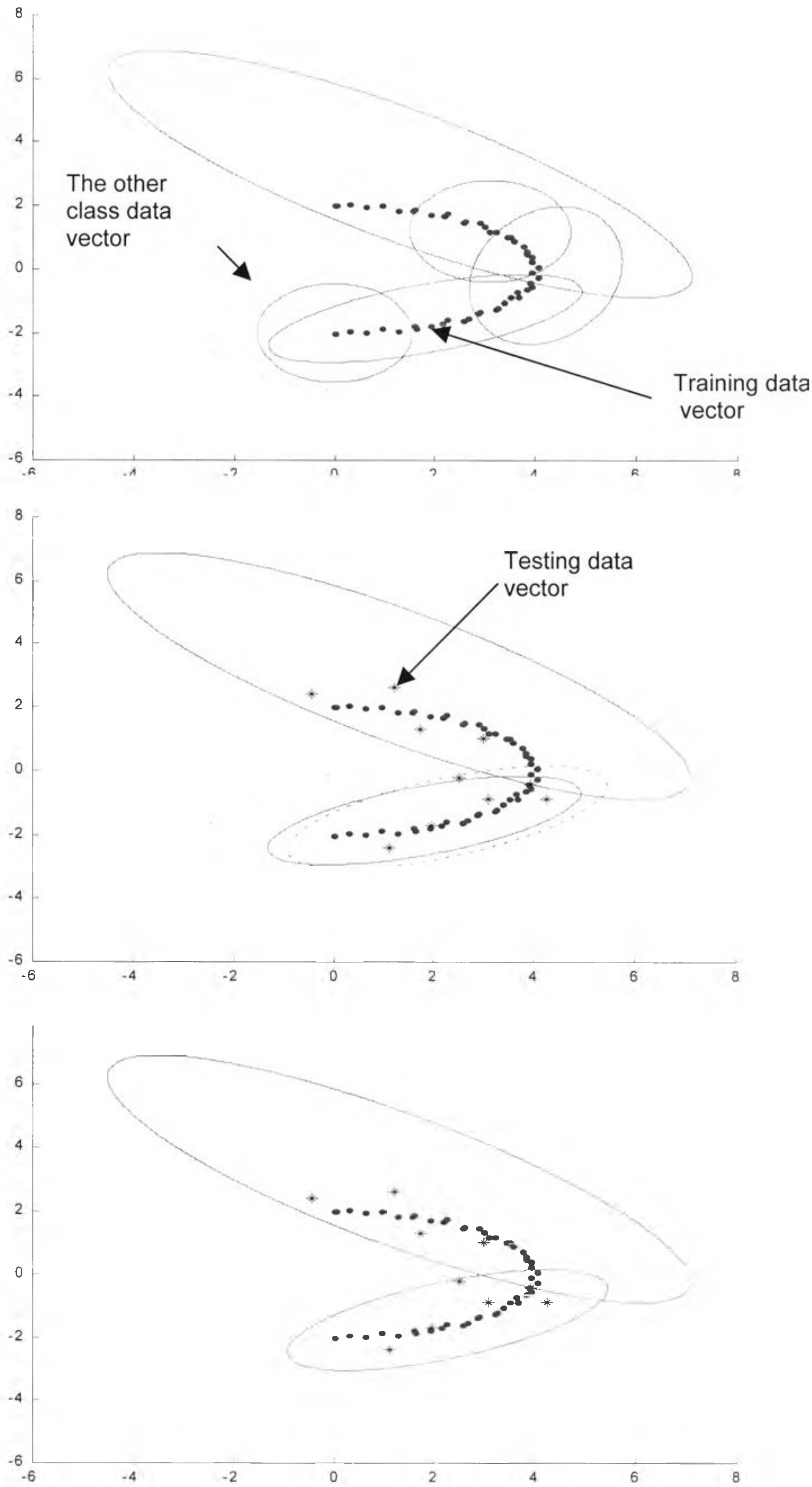

Figure 4.1 The result of Testset1_1.

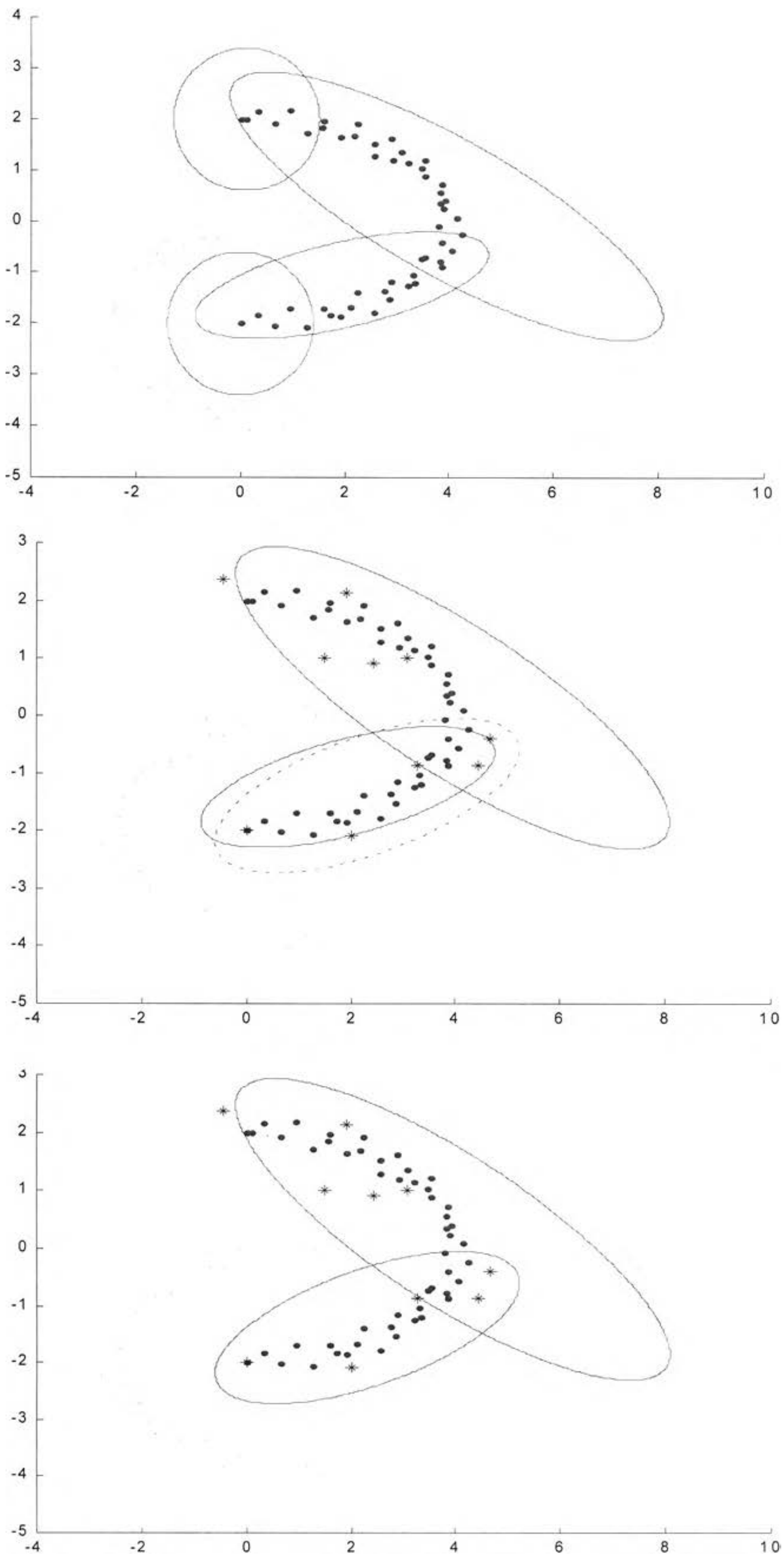

Figure 4.2 The result of Testset1_2.

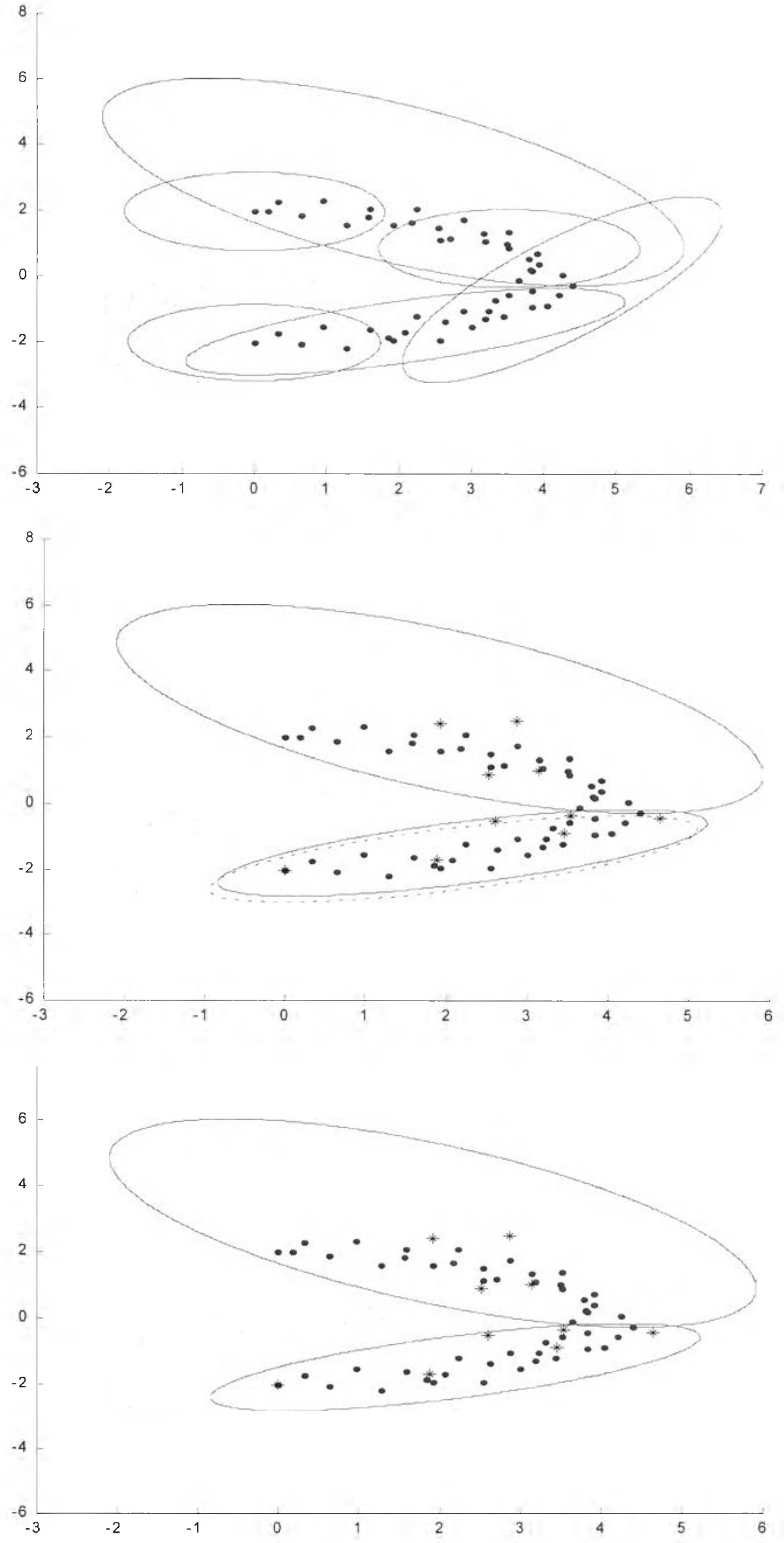

Figure 4.3 The result of Testset1_3.

4.2 Testset2

The second data set has three sine curves. These curves are used to show how our GERBFs shrink to lie in space channel of different classes of data. Tables 4.2 summarizes the number of epochs used for each training with different distribution and how better the classification is after we increase the generalization to network. We found that a correct classification rate from a smaller distribution data set is higher than from the larger distribution set because the small distribution provides more empty space for enlarging the GERBF area than the large distribution. Notice that there are no improvements of the correct classification rates for test sets 2_2 and 2_3. It is because the size of the space channel in each test case is smaller than the estimated size obtained by Bootstrap.

Testset	Number of hidden nodes	Average epochs	Correct Classification Rate (%) before Generalization	Correct Classification Rate (%) after Generalization
2_1	3	34	97.88	98.18
2_2	3	46	97.33	97.33
2_3	3	36	84.50	84.50

Table 4.2 The result of Testset 2.

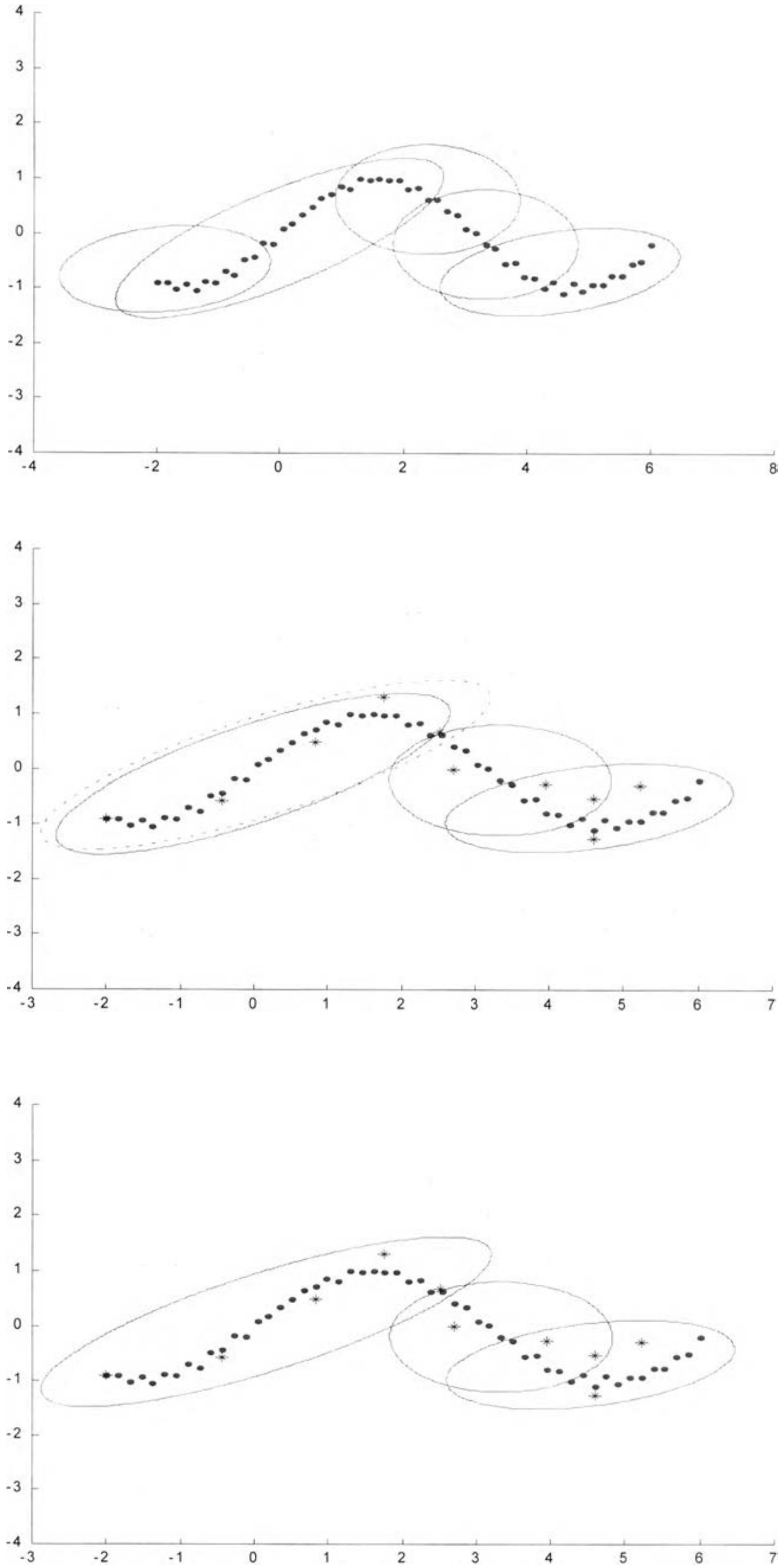

Figure 4.4 The result of Testset2_1.

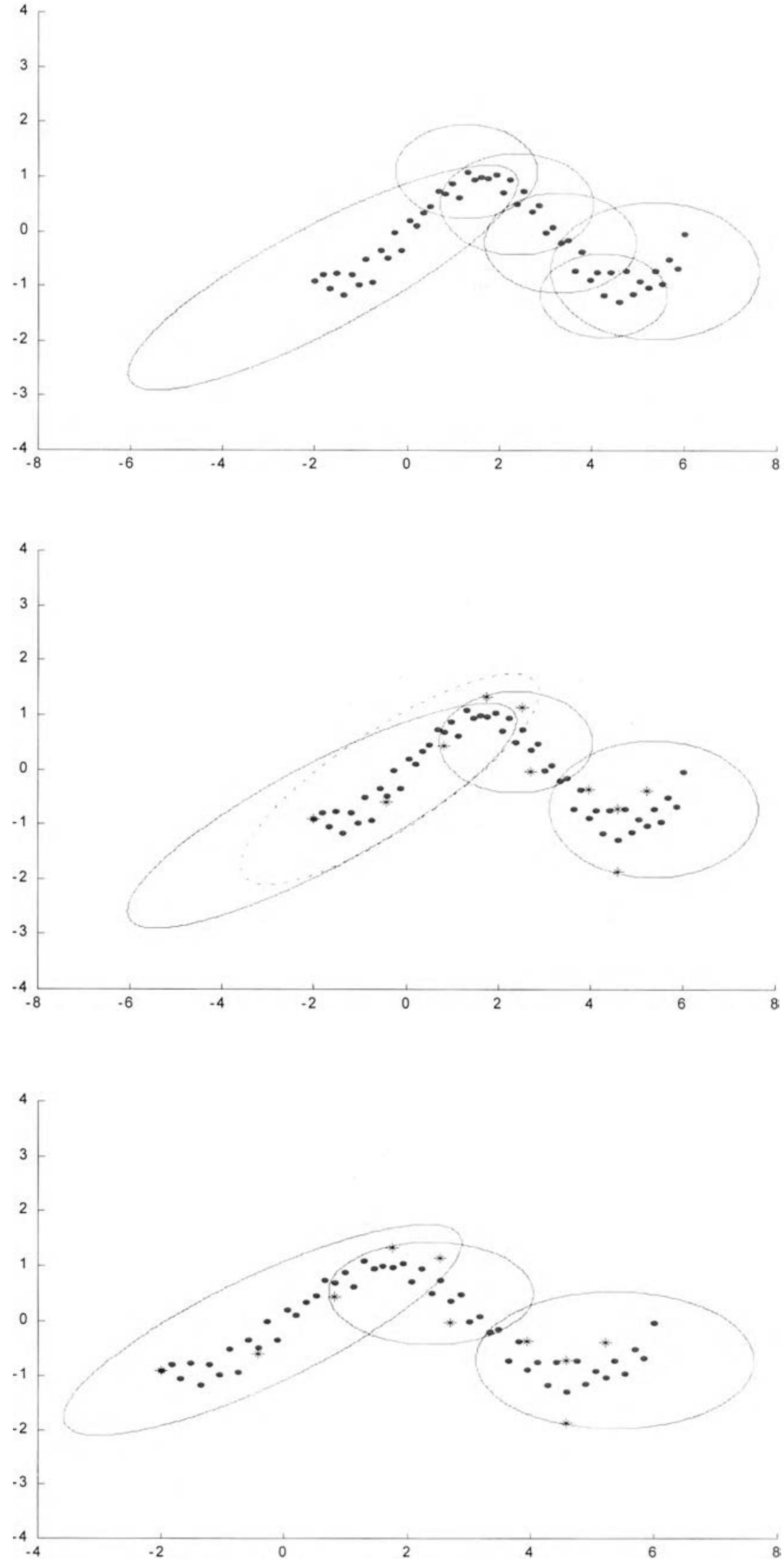

Figure 4.5 The result of Testset2_2.

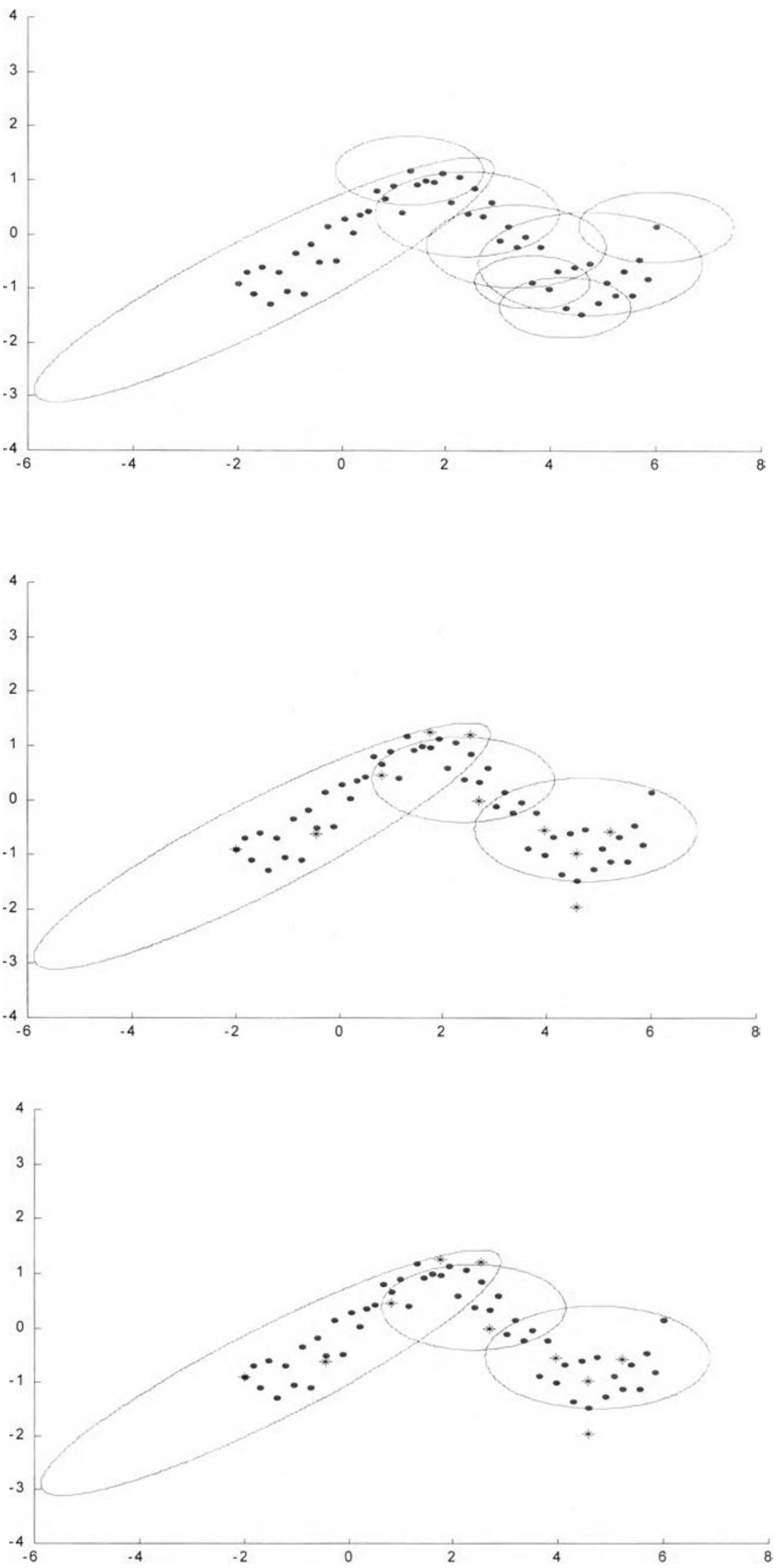

Figure 4.6 The result of Testset2_3.

4.3 Testset3

The three ring shapes show that our GERBFs work correctly with the rotation in any dimension in the limited space. The generalization does not work here due to the same reasons as those of testsets 2_2 and 2_3.

Testset	Number of hidden nodes	Average epochs	Correct Classification Rate (%) before Generalization	Correct Classification Rate (%) after Generalization
3_1	6	296	100.00	100.00
3_2	7	150	100.00	100.00
3_3	8	118	100.00	100.00

Table 4.3 The result of Testset 3.

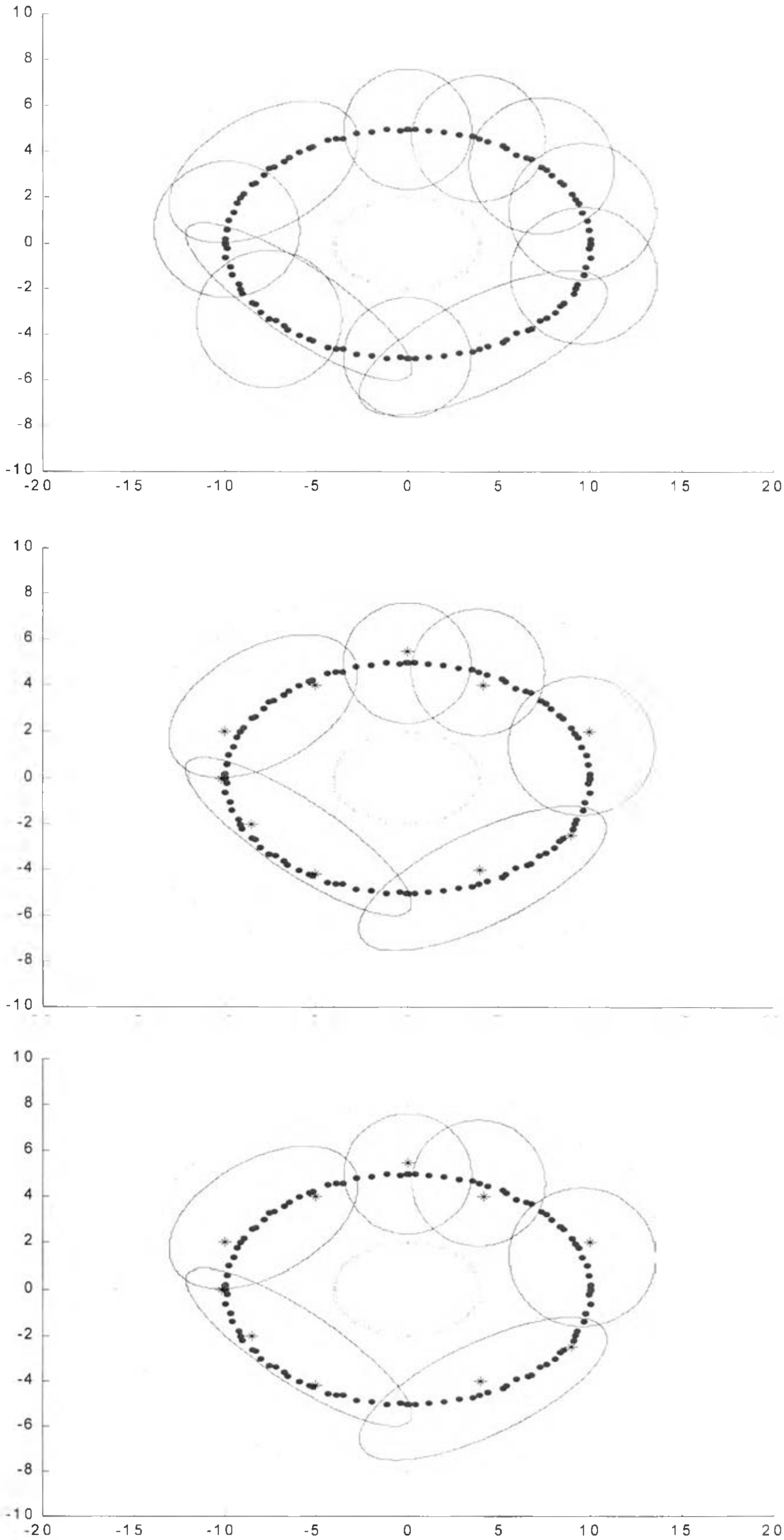

Figure 4.7 The result of Testset3_1.

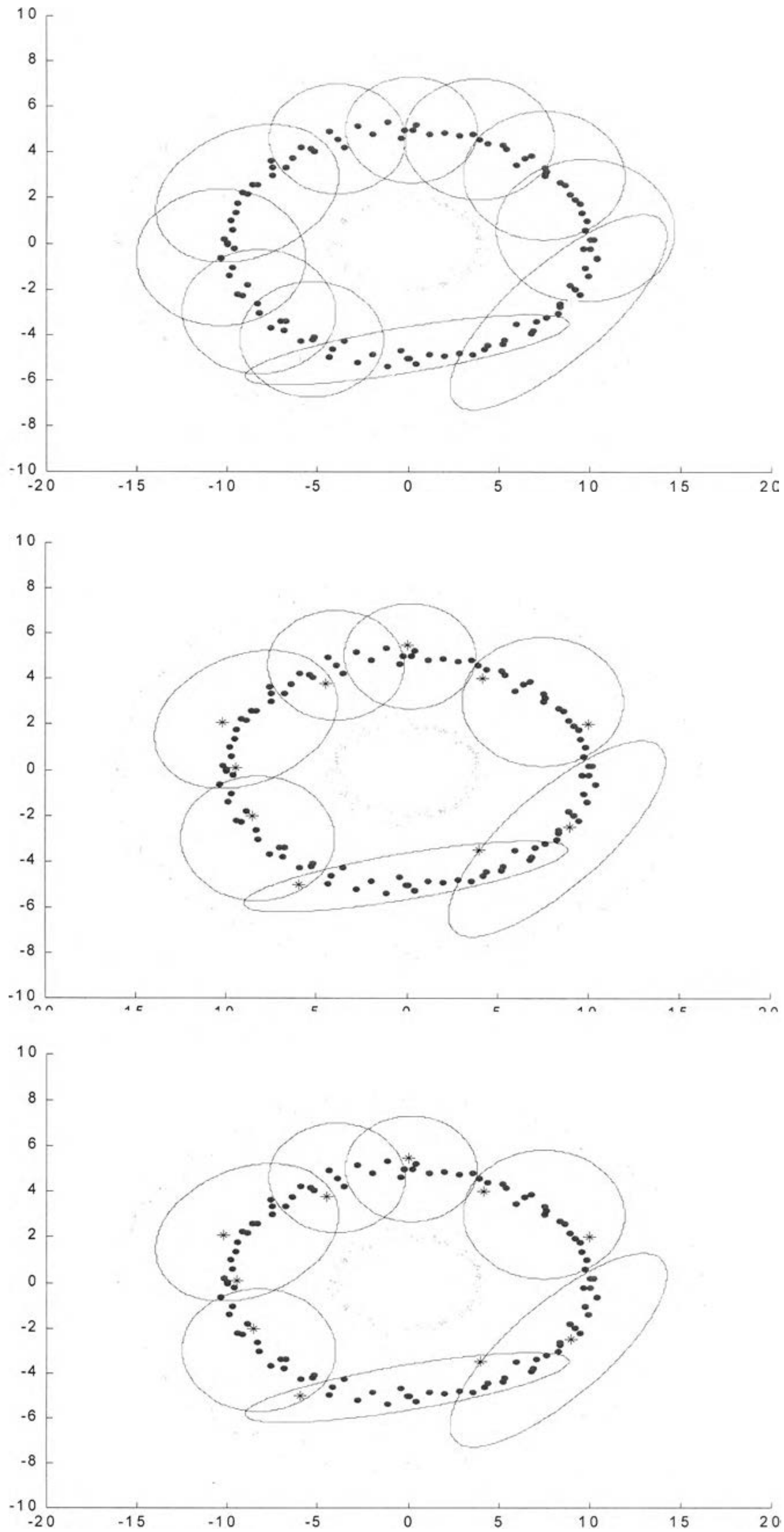

Figure4.8 The result of Testset3_2.

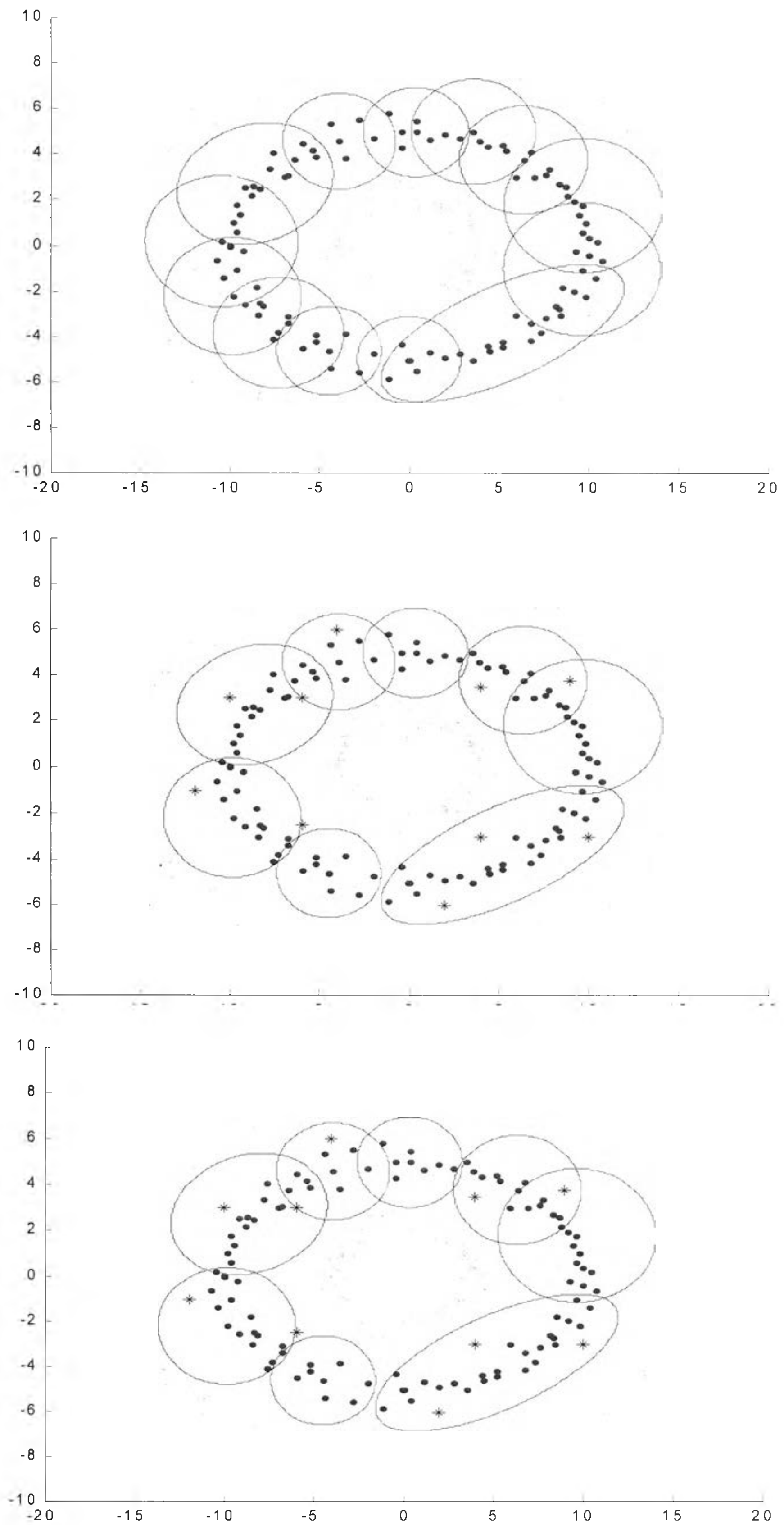

Figure4.9 The result of Testset3_3.

4.4 Testset4

This testset shows the working of GERBF with the general classification data set. In this case, there are two classes of data, which are distributed on x-axis and y-axis. Testset 4_1 and Testset 4_2 do not depend on each other. So the correct classification rate before generalization are different too. Table 4.4 shows the improvement of percent of the correct classification as the result of enlarging the GERBFs lying in some spaces between different classes of data. Figures 4.10 and 4.11 show the result in a 2-dimensional graph.

Testset	Number of hidden nodes	Average epochs	Correct Classification Rate(%) before Generalization	Correct Classification Rate(%) after Generalization
4_1	5	18	88.06	91.39
4_2	6	43	91.39	92.50

Table 4.4 The result of Testset 4.

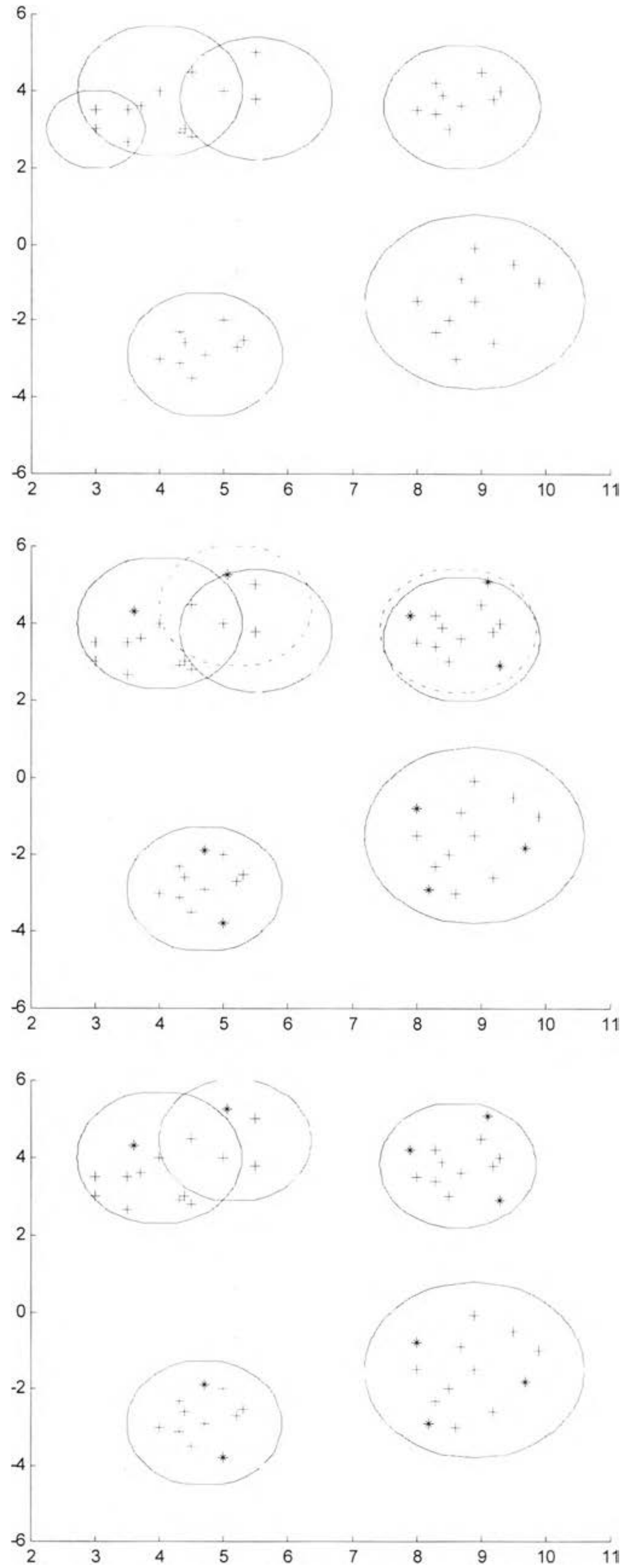

Figure 4.10 The result of Testset4_1.

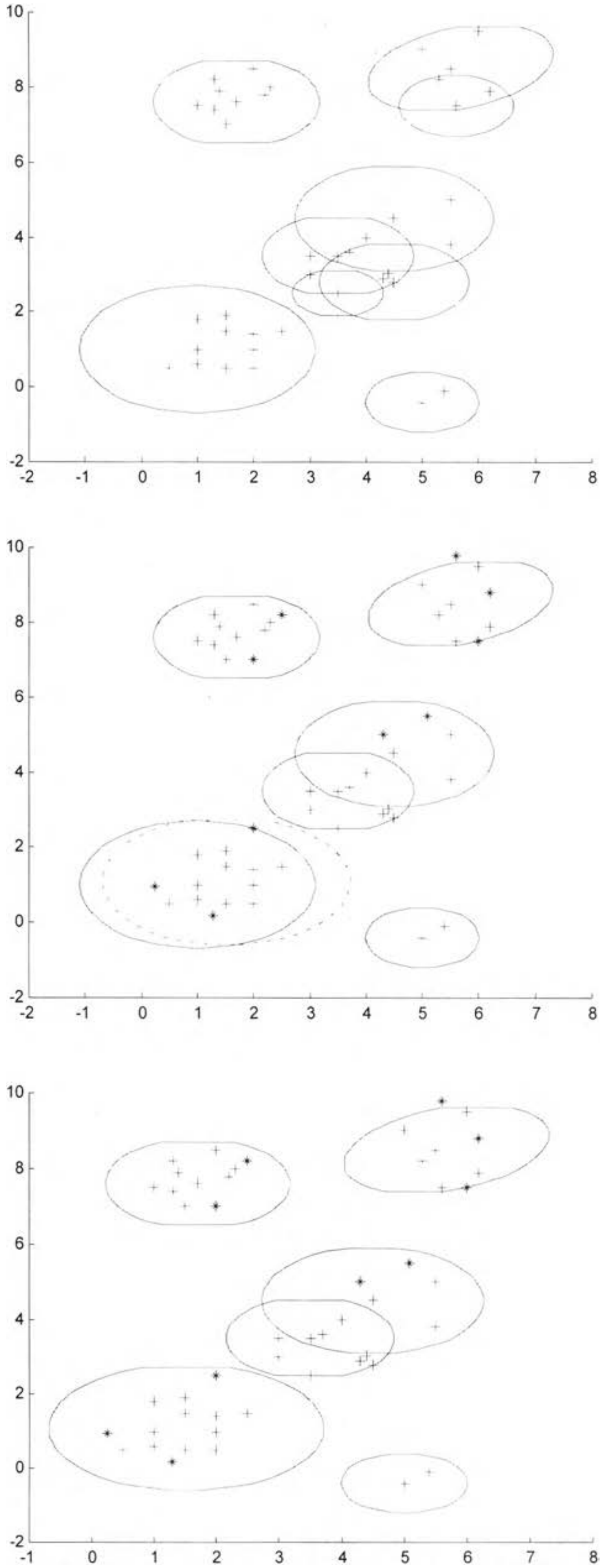

Figure 4.11 The result of Testset4_2.

Table 4.5 shows the comparison of our results and the results in [5]. Testset 2_1 to 2_3 and Testset 4_1 to 4_2 are not compared because those test sets are not the sets in [5]. And Figure 4.12 shows the greater results of the percent of the correct classification rate that come from generalization. The value on x-axis is name of our tested data and the value on y-axis is the percent of the correct classification rate.

Testset	Number of hidden nodes	Number of hidden nodes from [5]	Average epochs	Average epochs from [5]	Correct classification rate (%)	Correct classification rate (%) from [5]
1_1	2	3	48	61	95.00	94.20
1_2	2	3	36	73	95.00	94.68
1_3	3	4	34	123	99.00	94.97
2_1	3	-	45	-	95.08	-
2_2	6	-	66	-	95.48	-
2_3	7	-	28	-	89.00	-
3_1	6	7	256	230	100.00	94.17
3_2	7	7	41	384	94.00	95.00
3_3	8	9	24	579	95.00	94.03
4_1	5	-	86.94	-	91.39	-
4_2	6	-	91.39	-	92.5	-

Table 4.5 The result from all training data sets compared with those in [5]

Figure 4.12 shows the graph of comparison between average correct classification rate before (a continuous line) and after (a dashed line) generalization.

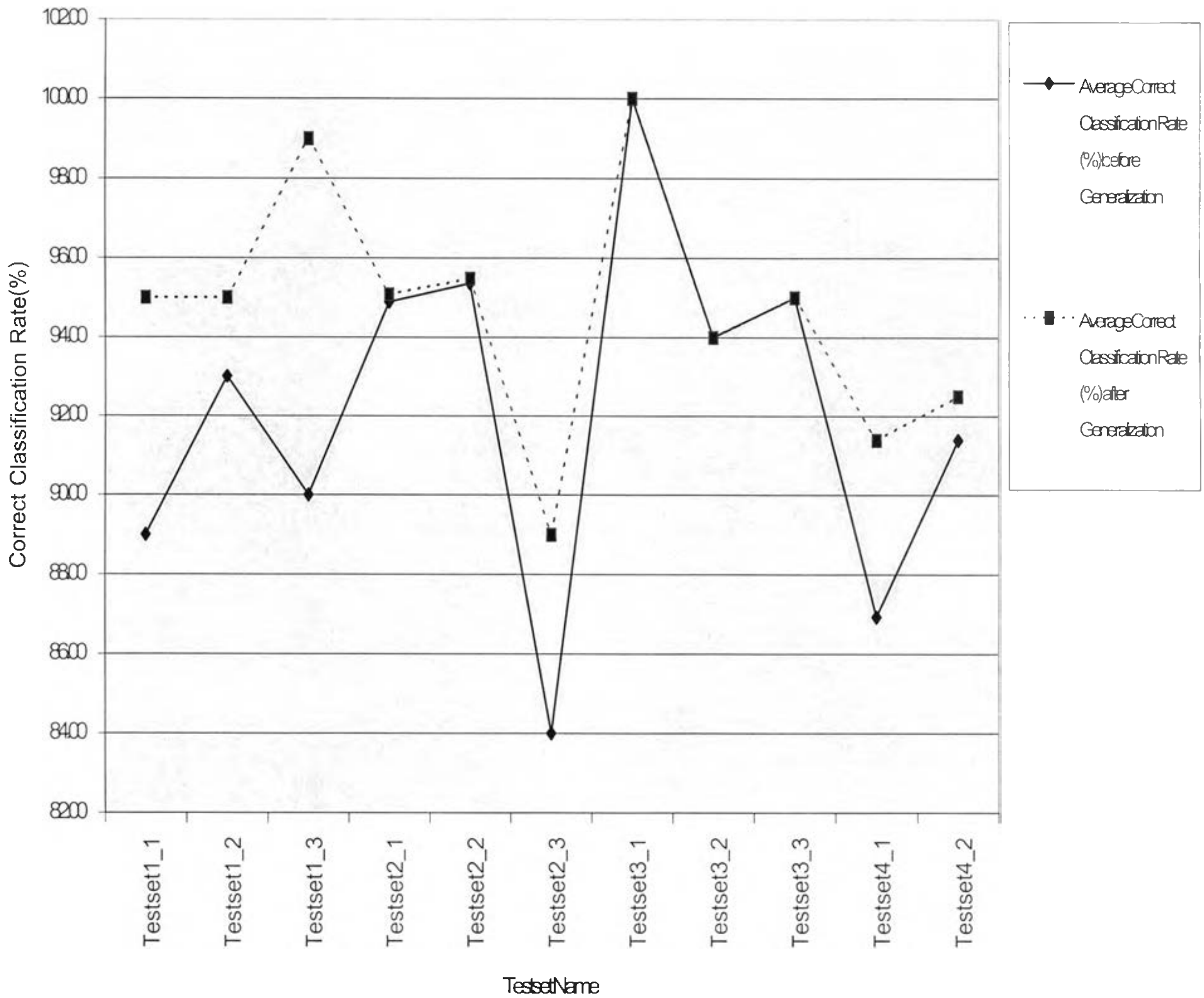

Figure 4.12 The comparison between the percentage of the correct classification rate before and after the generalization.