Contents

Assignment 1 Report

Rajanish B15126, Sahil Yadav B15130

September 17, 2017

Purpose

This report has been drafted as a part of **Assignment 1** under the course **CS669 : Pattern Recognition**.
There has been a sincere effort put in by the group to complete the assignment. This report is in parts some indication of the concepts that have developed along the way while learning the course. A practial application of the theoretical concepts is always a welcome experience this task being no different.

Problem Definition and Aims

We have been provided with three datasets :

1. Linearly separable artifical data
2. Non-linearly separable artifical data
3. Real world data

Here are a few features of datasets:

1. Each dataset has three classes.
2. Each data point consists of two features.

Training set : The first 75% data of each class has been taken as training set. The classifier is designed on the basis of parameters determined from this set.
Test set : The remaining 25% data of each class has been taken as test set. The classifier takes this data as input and gives the class label for this data as output.

Tasks

It is required to design classifiers considering varying constraints on covariance matrices of all classes resulting to four distinct cases. For every case we consider each dataset separately. For each point in each test dataset our classifier predicts the class it belongs to.
The results of this classification for each dataset is presented in the form of a **confusion matrix** which is further analyzed in terms of **classification accuracy**, **precision**, **recall** and **F-measure**.

In order to visually appreciate the decision surface and its performance there are three kinds of plots viz.:

1. Constant density contour plots for all classes together with the training data superposed.
2. Decision region plot for every pair of classes together with the training data superposed.
3. Decision region plot for all the classes together with the training data superposed.

Based on the assumption about the covariance matrices of all classes we are left with the following four sections :

1. Covariance matrix for all the classes is the same and is $\sigma^2\mathbf{I}$.
2. Full Covariance matrix for all the classes is the same and is Σ.
3. Covariance matric is diagonal and is different for each class.
4. Full Covariance matrix for each class is different.

Our classification criteria :

1. Classifying between two classes(C_i and C_j) at a time
   For every data point $\mathbf{x}$ in test data we calculate $g_i(\mathbf{x})$ for i and j. Then we provide classlabel for x :
   $$L(\mathbf{x}) : argmax\{i\}g_i(\mathbf{x}) \quad (1)$$
2. Classifying among three classes(C_i, C_j and C_k) at a time
   For every data point $\mathbf{x}$ in test data we calculate $g_i(\mathbf{x})$ for i, j and k. Then we provide classlabel for x :
   $$L(\mathbf{x}) : argmax\{i\}g_i(\mathbf{x}) \quad (2)$$

The following is a detailed insight into the aforementioned cases :

1 Covariance matrix for all the classes is the same and is σ^2I.

The general form of covariance matrix for a class having data which has two features is:

$$\Sigma : \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} \tag{1.1}$$

where :
σ_{11} : Covariance of feature 1 with feature 1 or variance of feature 1.
σ_{12} : Covariance of feature 1 with feature 2.
σ_{21} : Covariance of feature 2 with feature 1.
σ_{22} : Covariance of feature 2 with feature 2 or variance of feature 2.
In order to get $\Sigma = \sigma^2 I$
we need to do an averaging operation fow which we define σ_i^2 (Average covariance for each class) :

$$\sigma_i^2 = \frac{\sigma_{11} + \sigma_{12} + \sigma_{21} + \sigma_{22}}{4} \tag{1.2}$$

Furthermore we define σ^2 for n classes as :

$$\sigma^2 = \frac{\sigma_1^2 + \sigma_2^2 + ... + \sigma_n^2}{n} \tag{1.3}$$

For our particular case with three classes this reduces to :

$$\sigma^2 = \frac{\sigma_1^2 + \sigma_2^2 + \sigma_3^2}{3} \tag{1.4}$$

With this idea the training data for all three datasets is analysed to get their corresponding covariances i.e. σ_1^2, σ_1^2, and σ_1^2 using eq.(1.2). These covariance when put in eq.(1.4) give the required σ^2.

The equation for discriminating function for each class in this case is :

$$g_i(\mathbf{x}) = -\frac{1}{2\sigma^2}[\mathbf{x}^t\mathbf{x} - \mu_i\mathbf{x} + \mu_i^t\mu_i] + \ln P(C_i) \tag{1.5}$$

where :
μ_i : Mean vector of i^{th} class.
$\ln P(C_i)$: natural logarithm of prior probability of i^{th} class.

In our case we know a priori that every dataset has classes which have equal number of data points in them, thus prior probability for each class

$P(C_i)$ is same and can be negected from the dicriminating function. Therefore, the discriminanting function for i^{th} class now becomes :

$$g_i(\mathbf{x}) = -\frac{1}{2\sigma^2}[\mathbf{x}^t\mathbf{x} - \mu_i\mathbf{x} + \mu_i^t\mu_i] \tag{1.6}$$

All this is applied to every dataset in the next subsections.

1.1 Linearly Separable Data

Here are the plots for linearly separable data :

1. Decision region plot for every pair of classes together with the training data superposed.

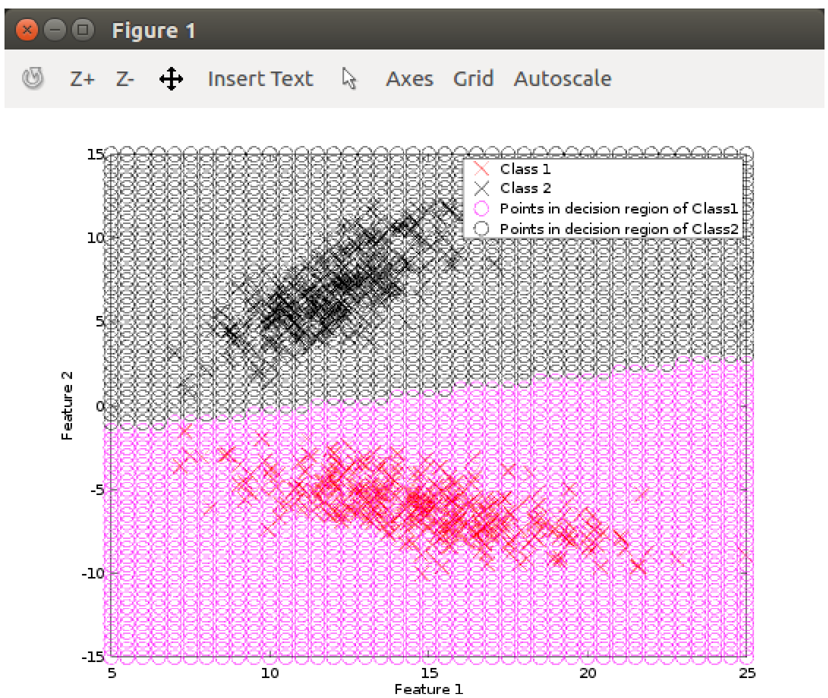
   (a) 

   Fig1 : This is the decision plot between class 1 and class 2 with their training data superposed.

(b)

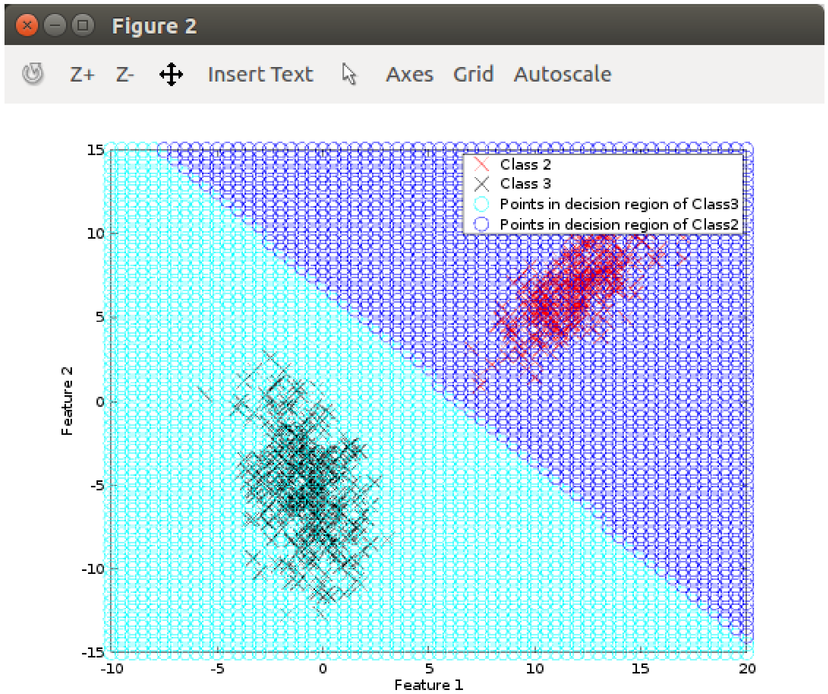


Fig2 : This is the decision plot between class 2 and class 3 with their training data superposed.

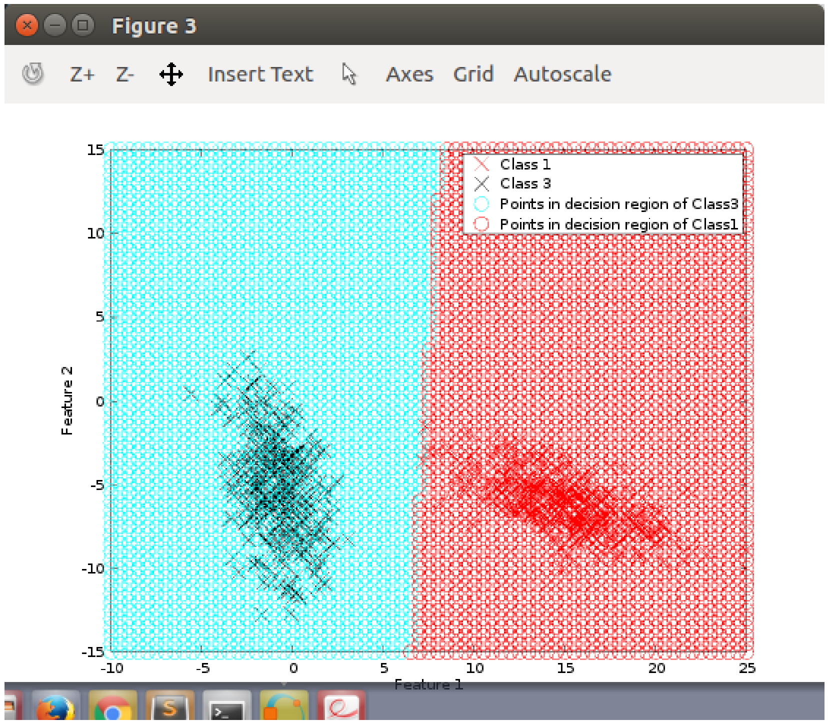
(c) 

Fig3 : This is the decision plot between class 3 and class 1 with their training data superposed.

2. Decision region plot for all the classes together with the training data superposed

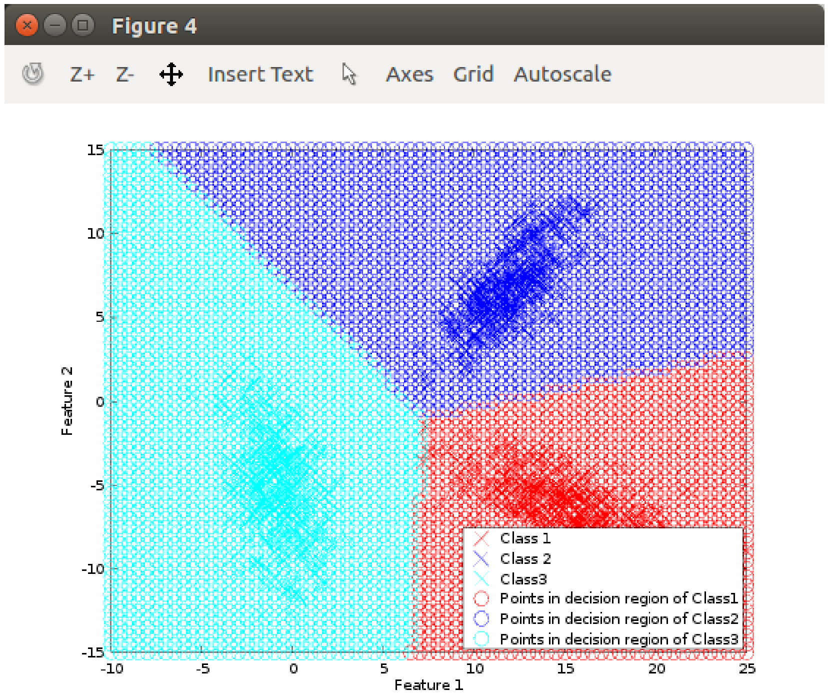
(a) 

Fig4 : This is the decision plot among class 1, class 2 and class 3 with their training data superposed.

3. Constant density contour plot for all the classes together with the training data superposed

(a)

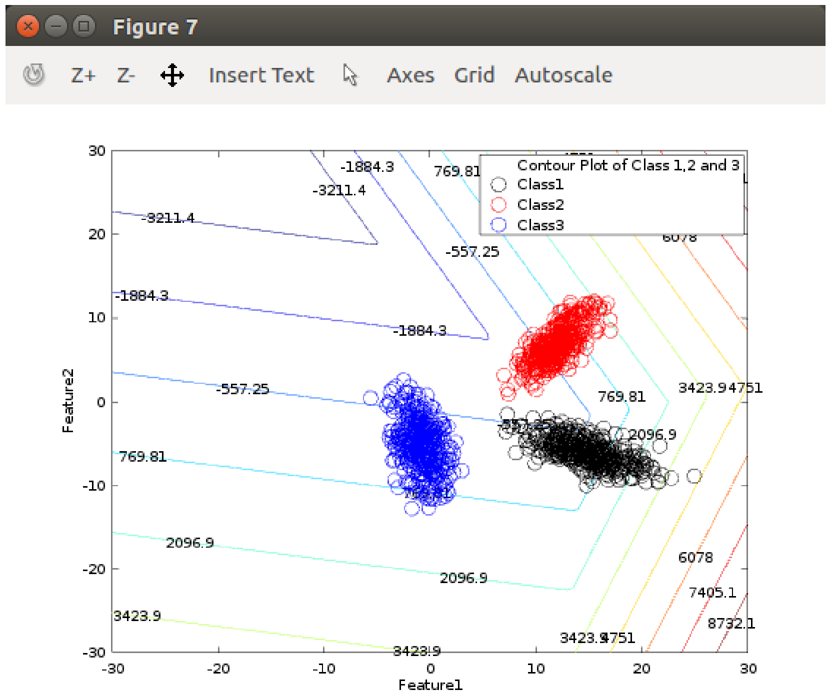


Fig5 : This is the constant density contour plot for all the classes together with the training data superposed

4. Confusion matrix

$$M = \begin{bmatrix} 125 & 0 & 0 \\ 0 & 125 & 0 \\ 0 & 0 & 125 \end{bmatrix}$$

In this matrix row(i) corresponds to the number of test data points categorized as class i.
Column(j) corresponds to the number of test data points actually in class j.

5. Calculation of performance parameters

(a) Accuracy

$$
\begin{aligned}
Accuracy &= \left(\frac{\text{Total correct classifications}}{\text{Total classifications}}\right) \times 100 \\
Accuracy &= \left(\frac{125+125+125}{125+125+125}\right) \times 100 \\
&= 100\%
\end{aligned}
$$

(b) Precision

i. Precision for Class 1

$$
\begin{aligned}
PC_1 &= \left(\frac{\text{Correct classifications to class 1}}{\text{Total classifications to class 1}}\right) \times 100 \\
PC_1 &= \left(\frac{125}{125+0+0}\right) \times 100 \\
PC_1 &= 100\%
\end{aligned}
$$

ii. Precision for Class 2

$$
\begin{aligned}
PC_2 &= \left(\frac{\text{Correct classifications to class 2}}{\text{Total classifications to class 2}}\right) \times 100 \\
PC_2 &= \left(\frac{125}{0+125+0}\right) \times 100 \\
PC_2 &= 100\%
\end{aligned}
$$

iii. Precision for Class 3

$$
\begin{aligned}
PC_3 &= \left(\frac{\text{Correct classifications to class 3}}{\text{Total classifications to class 3}}\right) \times 100 \\
PC_3 &= \left(\frac{125}{0+0+125}\right) \times 100 \\
PC_3 &= 100\%
\end{aligned}
$$

iv. Mean Precision

$$
\begin{aligned}
MPC &= \left(\frac{PC_1+PC_2+PC_3}{3}\right) \times 100 \\
MPC &= 100\%
\end{aligned}
$$

(c) Recall

i. Recall for Class 1

$$RC_1 = \left(\frac{\text{Correct classifications to class 1}}{\text{Total data points in class 1}}\right) \times 100$$
$$RC_1 = \left(\frac{125}{125}\right) \times 100$$
$$RC_1 = 100\%$$

ii. Recall for Class 2

$$RC_2 = \left(\frac{\text{Correct classifications to class 2}}{\text{Total data points in class 2}}\right) \times 100$$
$$RC_2 = \left(\frac{125}{125}\right) \times 100$$
$$RC_2 = 100\%$$

iii. Recall for Class 3

$$RC_3 = \left(\frac{\text{Correct classifications to class 3}}{\text{Total data points in class 3}}\right) \times 100$$
$$RC_3 = \left(\frac{125}{125}\right) \times 100$$
$$RC_3 = 100\%$$

iv. Mean Recall

$$MPC = \left(\frac{RC_1 + RC_2 + RC_3}{3}\right) \times 100$$
$$MPC = 100\%$$

(d) F-Measure

i. F-Measure for Class 1

$$FM_1 = \left(\frac{PC_1 \times RC_1 \times 2}{PC_1 + RC_1}\right)$$
$$FM_1 = \left(\frac{100 \times 100 \times 2}{100 + 100}\right)$$
$$FM_1 = 100$$

ii. F-Measure for Class 2

$$FM_2 = \left(\frac{PC_2 \times RC_2 \times 2}{PC_2 + RC_2}\right)$$
$$FM_2 = \left(\frac{100 \times 100 \times 2}{100 + 100}\right)$$
$$FM_2 = 100$$

iii. F-Measure for Class 3

$$
\begin{aligned}
FM_3 &= \left(\frac{PC_3 \times RC_3 \times 2}{PC_3 + RC_3}\right) \\
FM_3 &= \left(\frac{100 \times 100 \times 2}{100 + 100}\right) \\
FM_3 &= 100
\end{aligned}
$$

iv. Mean F-Measure

$$
\begin{aligned}
MFM &= \left(\frac{FM_1 + FM_1 + FM_3}{3}\right) \\
MFM &= \left(\frac{100 + 100 + 100}{3}\right) \\
MFM &= 100
\end{aligned}
$$

As all parameter for the linearly separable data points are 100% (for accuracy, mean precision and mean recall) and 100 (for mean F-measure). Thus the assumption for this case works perfectly well for the data.

1.2 Non-Linearly Separable Data

Here are the plots for non-linearly separable data :

1. Decision region plot for every pair of classes together with the training data superposed.

(a)

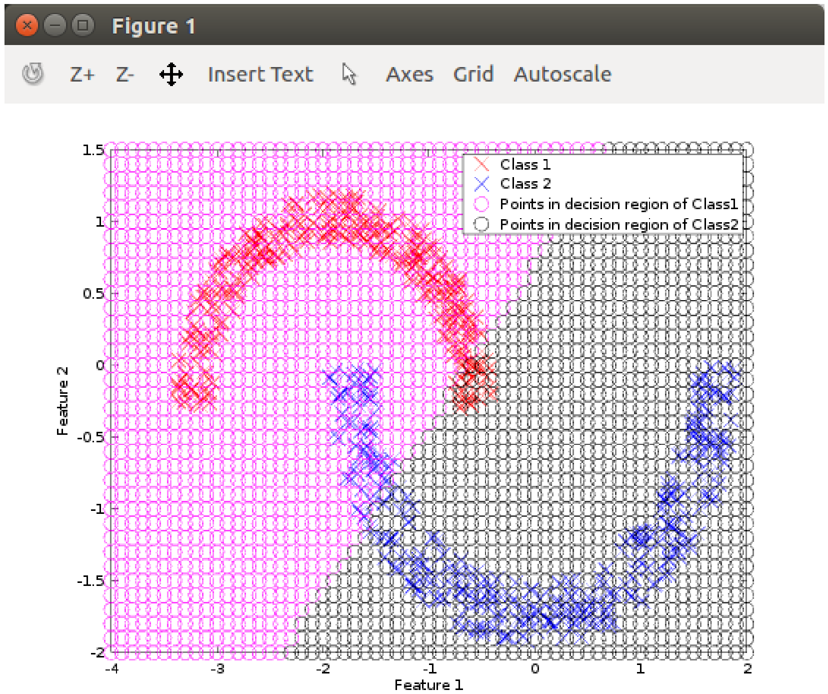


Fig1 : This is the decision plot between class 1 and class 2 with their training data superposed.

(b)

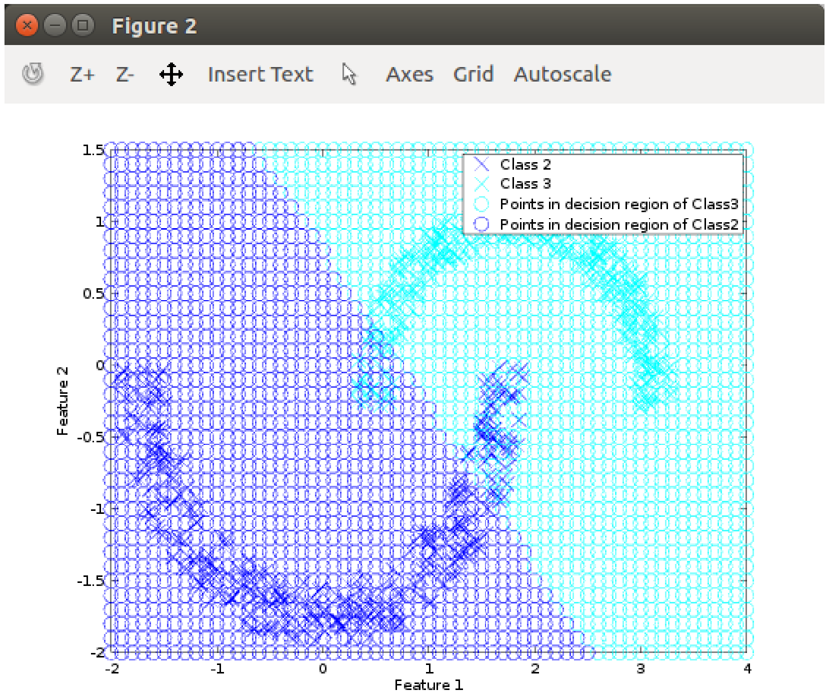


Fig2 : This is the decision plot between class 2 and class 3 with their training data superposed.

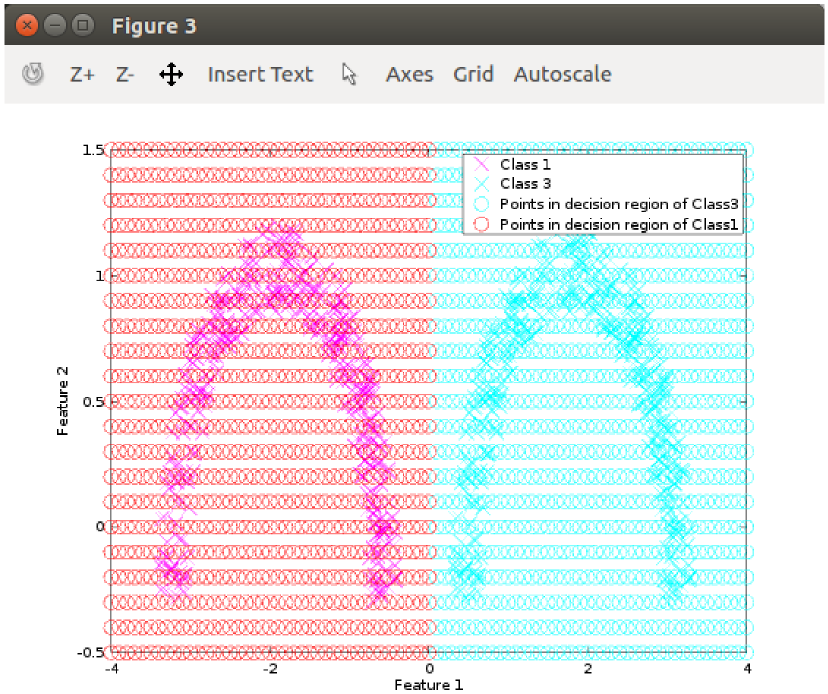
(c) 

Fig3 : This is the decision plot between class 3 and class 1 with their training data superposed.

2. Decision region plot for all the classes together with the training data superposed

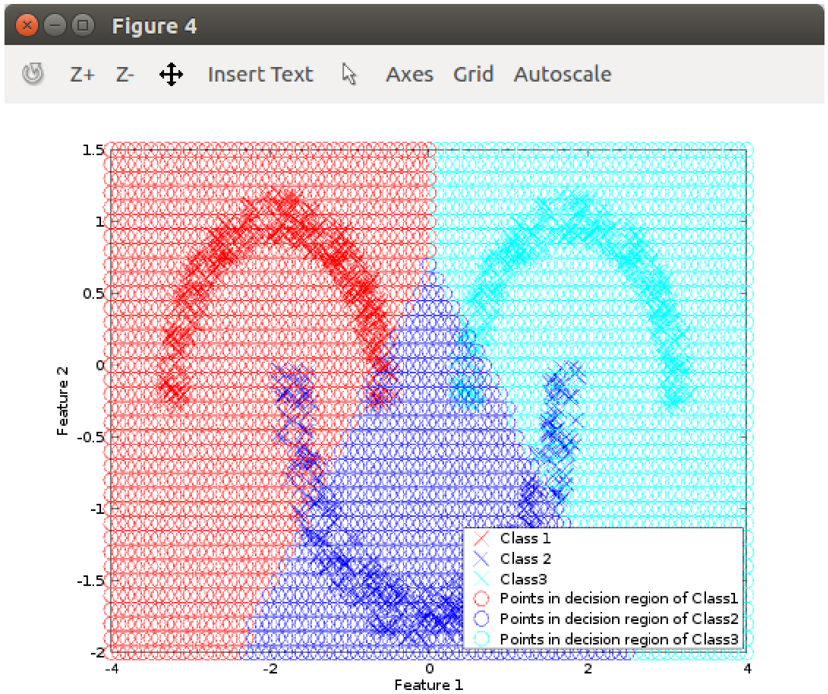
(a) 

Fig4 : This is the decision plot among class 1, class 2 and class 3 with their training data superposed.

3. Constant density contour plot for all the classes together with the training data superposed

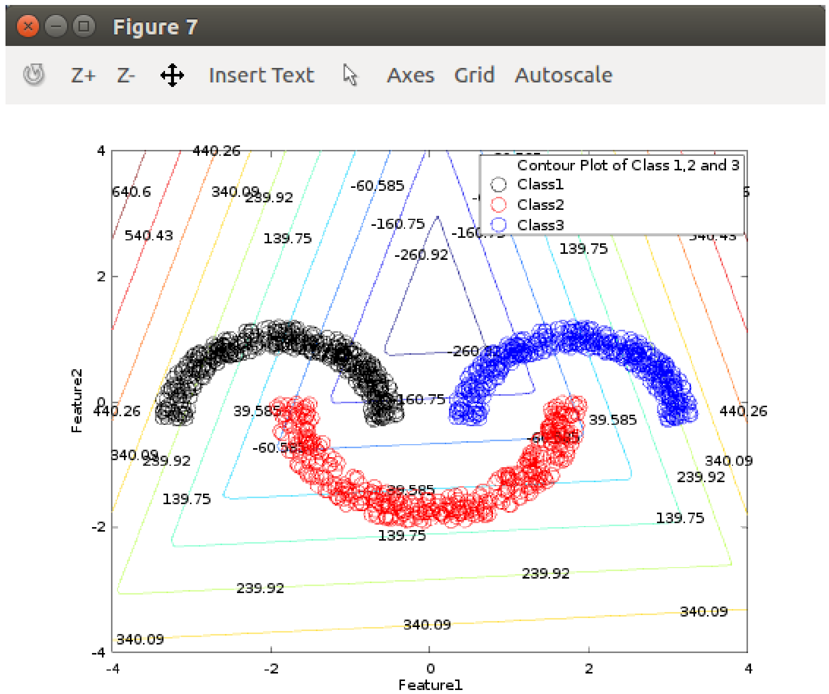
(a) 

Fig5 : This is the constant density contour plot for all the classes together with the training data superposed

4. Confusion matrix

$$M = \begin{bmatrix} 113 & 12 & 0 \\ 26 & 77 & 22 \\ 0 & 10 & 115 \end{bmatrix}$$

In this matrix row(i) corresponds to the number of test data points categorized as class i.
Column(j) corresponds to the number of test data points actually in class j.

5. Calculation of performance parameters

(a) Accuracy

$$
\begin{aligned}
Accuracy &= \left(\frac{\text{Total correct classifications}}{\text{Total classifications}}\right) \times 100 \\
Accuracy &= \left(\frac{113 + 77 + 115}{125 + 125 + 125}\right) \times 100 \\
&= 81.33\%
\end{aligned}
$$

(b) Precision

i. Precision for Class 1

$$
\begin{aligned}
PC_1 &= \left(\frac{\text{Correct classifications to class 1}}{\text{Total classifications to class 1}}\right) \times 100 \\
PC_1 &= \left(\frac{113}{113 + 26 + 0}\right) \times 100 \\
PC_1 &= 81.29\%
\end{aligned}
$$

ii. Precision for Class 2

$$
\begin{aligned}
PC_2 &= \left(\frac{\text{Correct classifications to class 2}}{\text{Total classifications to class 2}}\right) \times 100 \\
PC_2 &= \left(\frac{77}{77 + 12 + 10}\right) \times 100 \\
PC_2 &= 77.78\%
\end{aligned}
$$

iii. Precision for Class 3

$$
\begin{aligned}
PC_3 &= \left(\frac{\text{Correct classifications to class 3}}{\text{Total classifications to class 3}}\right) \times 100 \\
PC_3 &= \left(\frac{115}{115 + 22 + 0}\right) \times 100 \\
PC_3 &= 83.94\%
\end{aligned}
$$

iv. Mean Precision

$$
\begin{aligned}
MPC &= \left(\frac{PC_1 + PC_2 + PC_3}{3}\right) \times 100 \\
MPC &= 81\%
\end{aligned}
$$

(c) Recall

i. Recall for Class 1

$$\begin{aligned} RC_1 &= \left(\frac{\text{Correct classifications to class 1}}{\text{Total data points in class 1}}\right) \times 100 \\ RC_1 &= \left(\frac{113}{125}\right) \times 100 \\ RC_1 &= 90.4\% \end{aligned}$$

ii. Recall for Class 2

$$\begin{aligned} RC_2 &= \left(\frac{\text{Correct classifications to class 2}}{\text{Total data points in class 2}}\right) \times 100 \\ RC_2 &= \left(\frac{77}{125}\right) \times 100 \\ RC_2 &= 61.6\% \end{aligned}$$

iii. Recall for Class 3

$$\begin{aligned} RC_3 &= \left(\frac{\text{Correct classifications to class 3}}{\text{Total data points in class 3}}\right) \times 100 \\ RC_3 &= \left(\frac{115}{125}\right) \times 100 \\ RC_3 &= 92\% \end{aligned}$$

iv. Mean Recall

$$\begin{aligned} MPC &= \left(\frac{RC_1 + RC_2 + RC_3}{3}\right) \times 100 \\ MRC &= 81.33\% \end{aligned}$$

(d) F-Measure

i. F-Measure for Class 1

$$\begin{aligned} FM_1 &= \left(\frac{PC_1 \times RC_1 \times 2}{PC_1 + RC_1}\right) \\ FM_1 &= \left(\frac{81.29 \times 90.4 \times 2}{81.92 + 90.4}\right) \\ FM_1 &= 85.60 \end{aligned}$$

ii. F-Measure for Class 2

$$\begin{aligned} FM_2 &= \left(\frac{PC_2 \times RC_2 \times 2}{PC_2 + RC_2}\right) \\ FM_2 &= \left(\frac{77.78 \times 61.6 \times 2}{77.78 + 61.6}\right) \\ FM_2 &= 68.75 \end{aligned}$$

iii. F-Measure for Class 3

$$FM_3 = \left(\frac{PC_3 \times RC_3 \times 2}{PC_3 + RC_3}\right)$$
$$FM_3 = \left(\frac{83.94 \times 92 \times 2}{83.92 + 92}\right)$$
$$FM_3 = 87.79$$

iv. Mean F-Measure

$$MFM = \left(\frac{FM_1 + FM_1 + FM_3}{3}\right)$$
$$MFM = \left(\frac{85.60 + 68.75 + 87.79}{3}\right)$$
$$MFM = 80.71$$

We notice that the accuracy, mean precision, mean recall and F-measure around 80% but recall for class 2 is 61.6 % thus this method may lead to noticibale loss where recall for class 2 is a variable of major significance.

1.3 Real-world Data

Here are the plots for non-linearly separable data :

1. Decision region plot for every pair of classes together with the training data superposed.

(a)

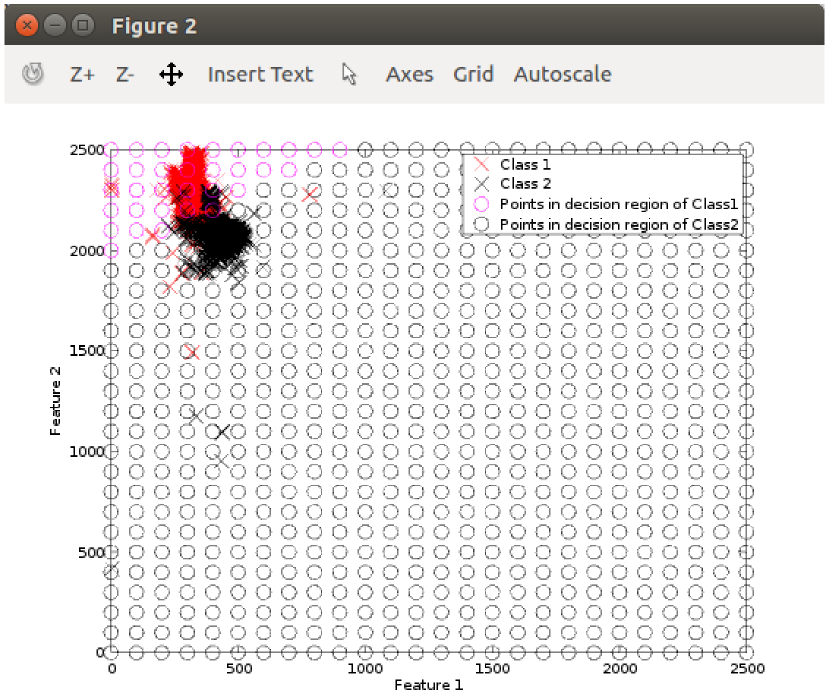


Fig1 : This is the decision plot between class 1 and class 2 with their training data superposed.

(b)

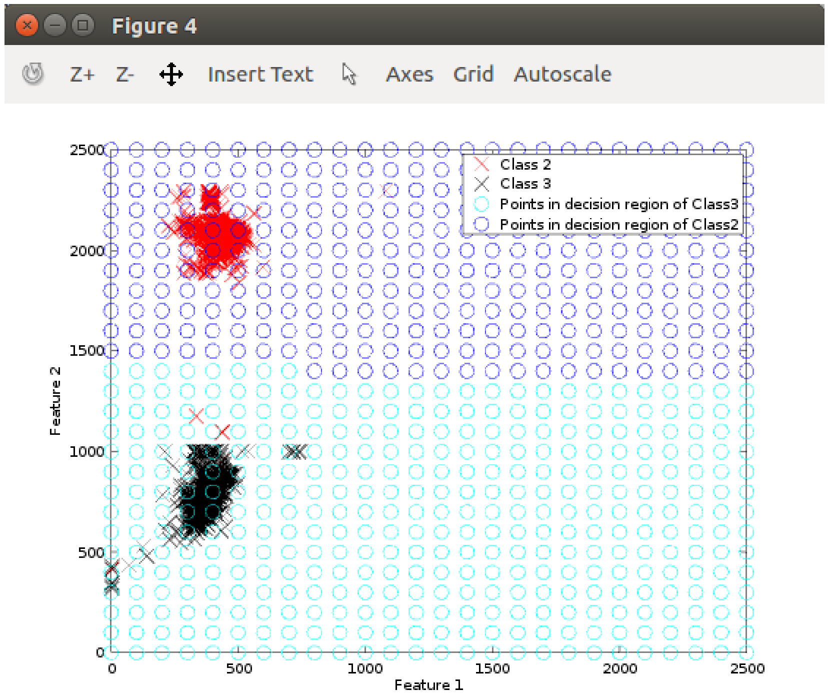


Fig2 : This is the decision plot between class 2 and class 3 with their training data superposed.

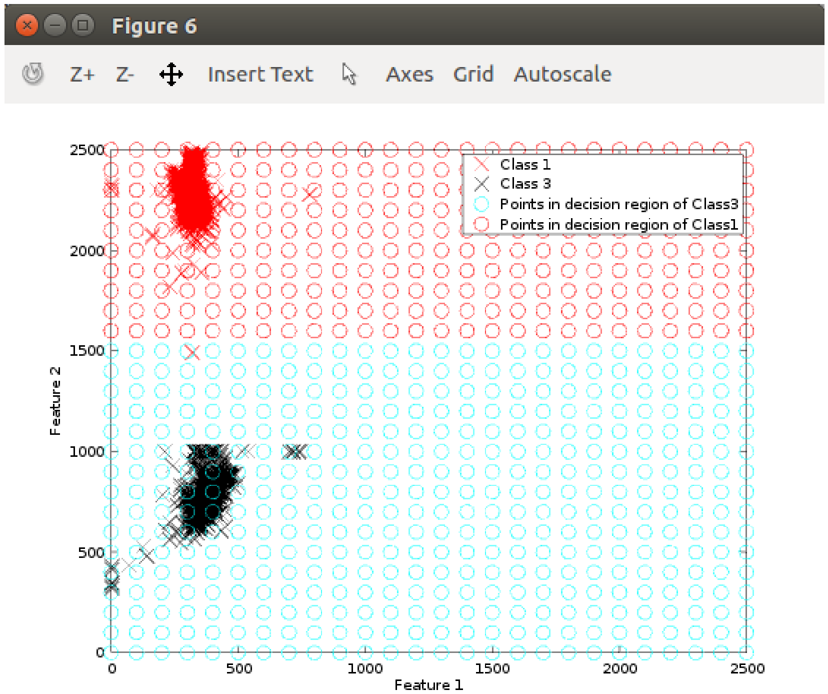
(c) 

Fig3 : This is the decision plot between class 3 and class 1 with their training data superposed.

2. Decision region plot for all the classes together with the training data superposed

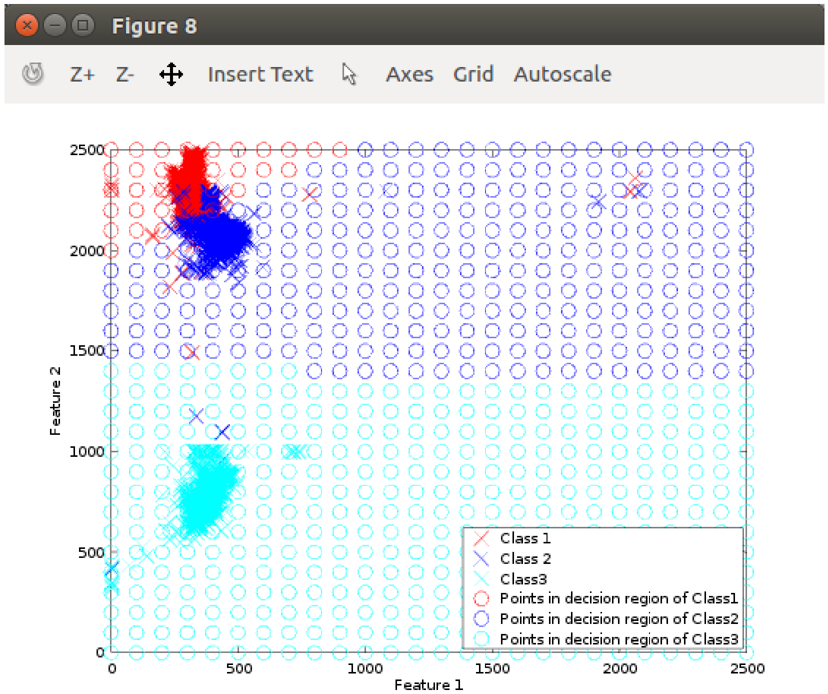
(a) 

Fig4 : This is the decision plot among class 1, class 2 and class 3 with their training data superposed.

3. Constant density contour plot for all the classes together with the training data superposed

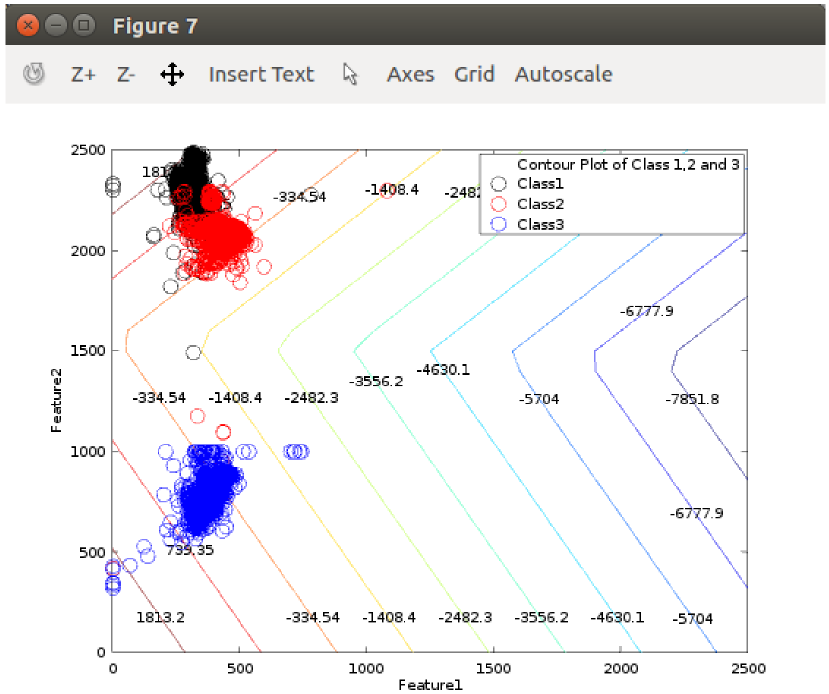
(a) 

Fig5 : This is the constant density contour plot for all the classes together with the training data superposed

4. Confusion matrix

$$M = \begin{bmatrix} 570 & 24 & 3 \\ 348 & 207 & 18 \\ 0 & 0 & 622 \end{bmatrix}$$

In this matrix row(i) corresponds to the number of test data points categorized as class i.
Column(j) corresponds to the number of test data points actually in class j.

5. Calculation of performance parameters

   (a) Accuracy

$$Accuracy \quad = \quad 78.069\%$$

   (b) Precision

i. Precision for Class 1

$$PC_1 = 62.092\%$$

ii. Precision for Class 2

$$PC_2 = 89.610\%$$

iii. Precision for Class 3

$$PC_3 = 96.734\%$$

iv. Mean Precision

$$MPC = 82.812\%$$

(c) Recall

i. Recall for Class 1

$$RC_1 = 95.477\%$$

ii. Recall for Class 2

$$RC_2 = 36.126\%$$

iii. Recall for Class 3

$$RC_3 = 100\%$$

iv. Mean Recall

$$MRC = 77.201\%$$

(d) F-Measure

i. F-Measure for Class 1

$$FM_1 = 75.248$$

ii. F-Measure for Class 2

$$FM_2 = 51.493$$

iii. F-Measure for Class 3

$$FM_3 = 98.340$$

iv. Mean F-Measure

$$MFM = 75.027$$

Here again the accuracy, mean recall, mean precision all are around 80% but some noticable observations prevail : 1. Precision for class 1 is not very stisfactory. 2. Recall for class 2 is poor. 3. Recall for class 3 is 100%.

So this method on this data may prove futile in case precision of class 1 and recall of class 2 matter.

2 Full Covariance matrix for all the classes is the same and is Σ.

Here we have :

$$\Sigma = \frac{\Sigma_1 + \Sigma_2 + \Sigma_3}{3}$$

$$g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x} - \mu_i^t)\Sigma^{-1}(\mathbf{x} - \mu_i) + \ln P(C_i)$$

2.1 Linearly Separable Data

Here are the plots for linearly separable data :

1. Decision region plot for every pair of classes together with the training data superposed.

(a)

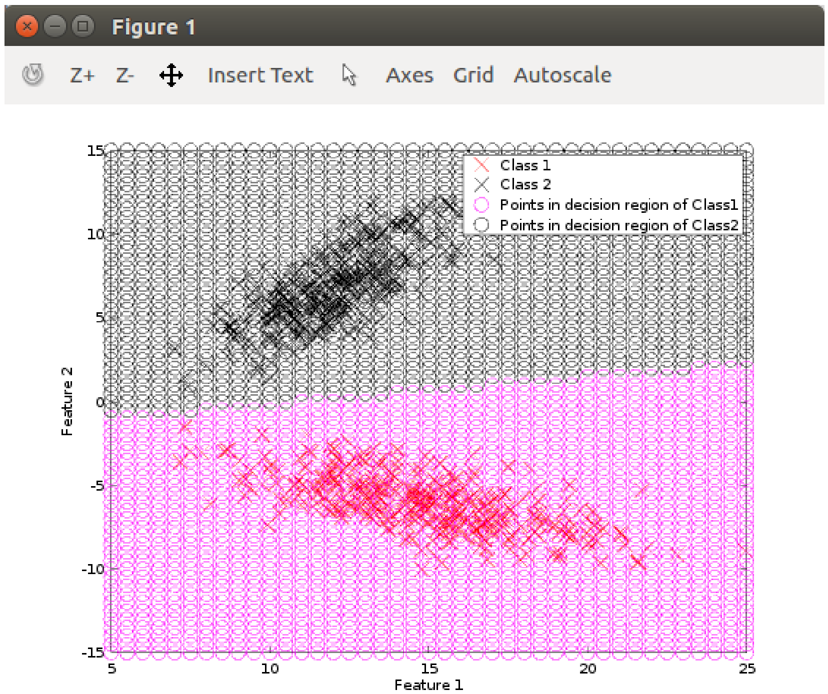


Fig1 : This is the decision plot between class 1 and class 2 with their training data superposed.

(b)

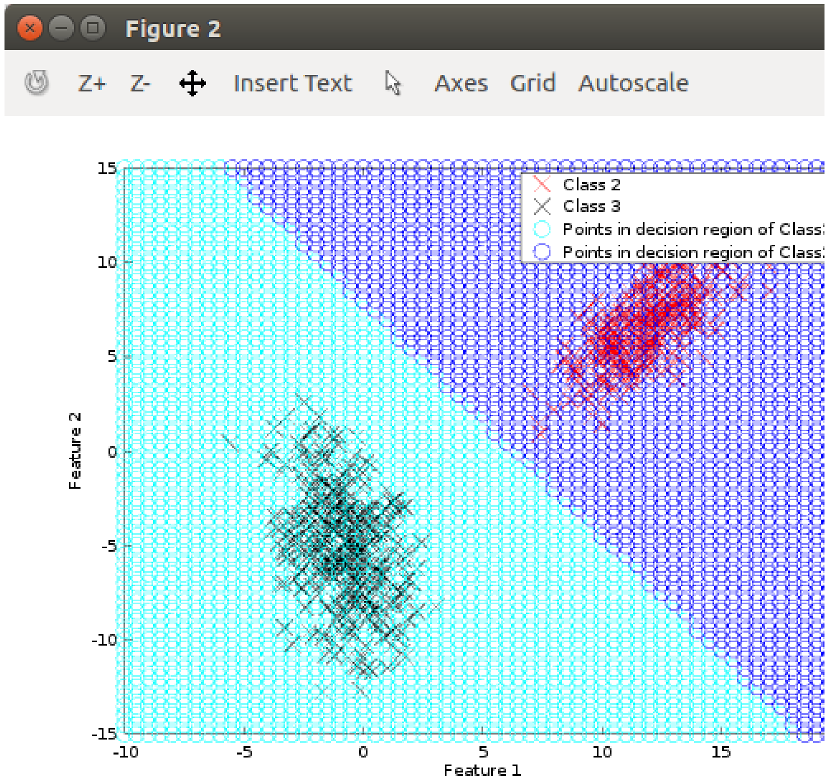


Fig2 : This is the decision plot between class 2 and class 3 with their training data superposed.

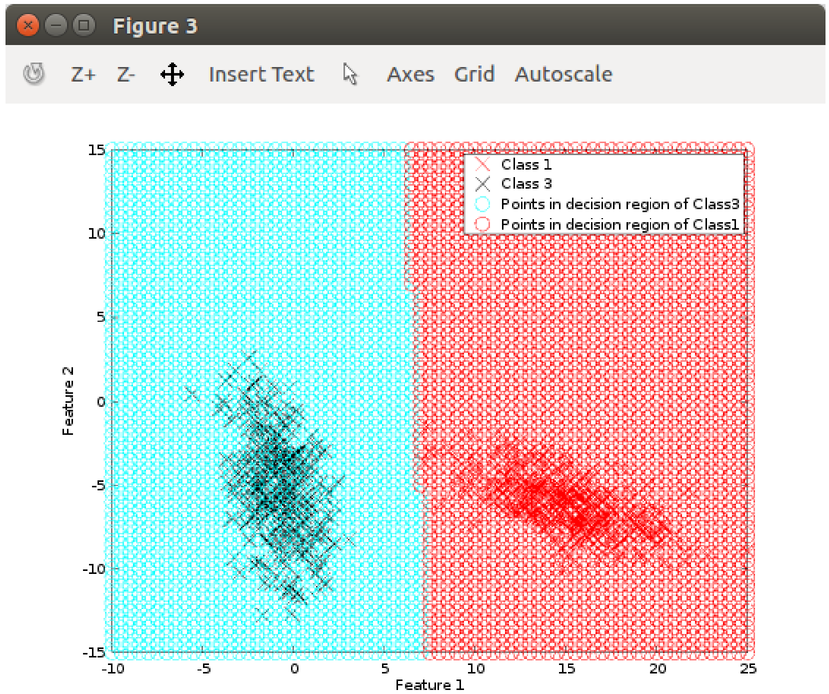
(c) 

Fig3 : This is the decision plot between class 3 and class 1 with their training data superposed.

2. Decision region plot for all the classes together with the training data superposed

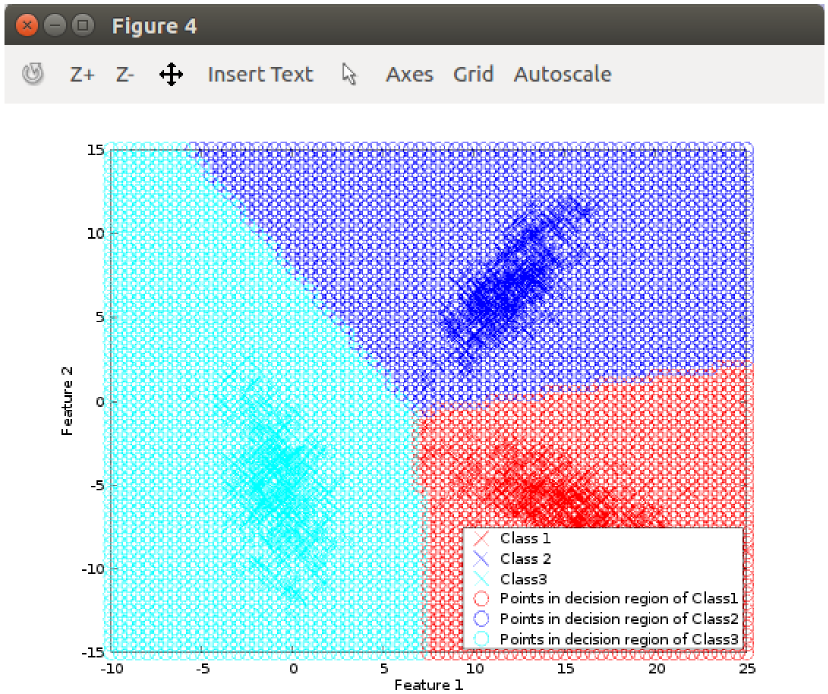
(a) 

Fig4 : This is the decision plot among class 1, class 2 and class 3 with their training data superposed.

3. Constant density contour plot for all the classes together with the training data superposed

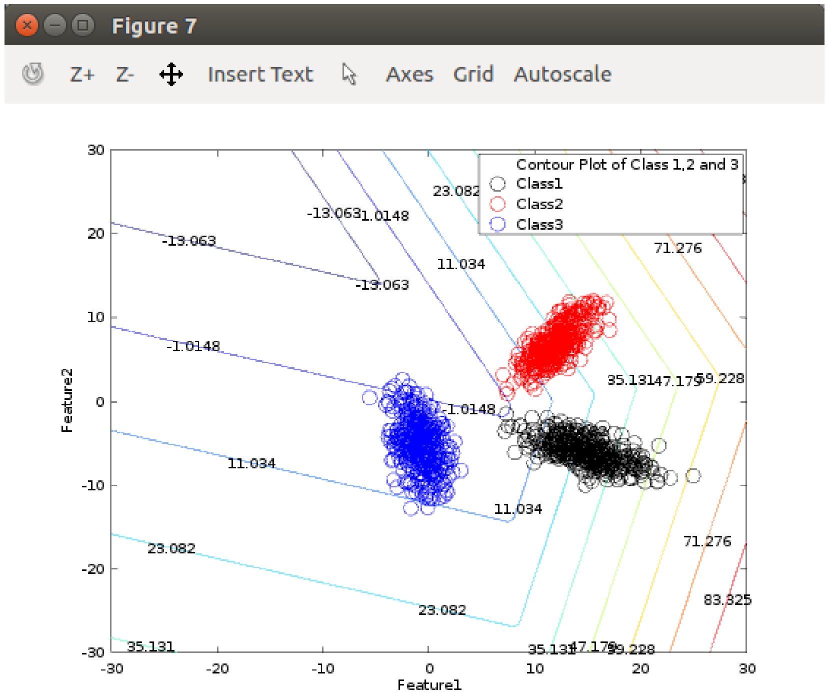
(a) 

Fig5 : This is the constant density contour plot for all the classes together with the training data superposed

4. Confusion matrix

$$M = \begin{bmatrix} 125 & 0 & 0 \\ 0 & 125 & 0 \\ 0 & 0 & 125 \end{bmatrix}$$

In this matrix row(i) corresponds to the number of test data points categorized as class i.
Column(j) corresponds to the number of test data points actually in class j.

5. Calculation of performance parameters

(a) Accuracy

$$\begin{aligned} Accuracy &= \left(\frac{\text{Total correct classifications}}{\text{Total classifications}}\right) \times 100 \\ Accuracy &= \left(\frac{125+125+125}{125+125+125}\right) \times 100 \\ &= 100\% \end{aligned}$$

(b) Precision

i. Precision for Class 1

$$\begin{aligned} PC_1 &= \left(\frac{\text{Correct classifications to class 1}}{\text{Total classifications to class 1}}\right) \times 100 \\ PC_1 &= \left(\frac{125}{125+0+0}\right) \times 100 \\ PC_1 &= 100\% \end{aligned}$$

ii. Precision for Class 2

$$\begin{aligned} PC_2 &= \left(\frac{\text{Correct classifications to class 2}}{\text{Total classifications to class 2}}\right) \times 100 \\ PC_2 &= \left(\frac{125}{0+125+0}\right) \times 100 \\ PC_2 &= 100\% \end{aligned}$$

iii. Precision for Class 3

$$\begin{aligned} PC_3 &= \left(\frac{\text{Correct classifications to class 3}}{\text{Total classifications to class 3}}\right) \times 100 \\ PC_3 &= \left(\frac{125}{0+0+125}\right) \times 100 \\ PC_3 &= 100\% \end{aligned}$$

iv. Mean Precision

$$\begin{aligned} MPC &= \left(\frac{PC_1+PC_2+PC_3}{3}\right) \times 100 \\ MPC &= 100\% \end{aligned}$$

(c) Recall

i. Recall for Class 1

$$\begin{aligned} RC_1 &= \left(\frac{\text{Correct classifications to class 1}}{\text{Total data points in class 1}}\right) \times 100 \\ RC_1 &= \left(\frac{125}{125}\right) \times 100 \\ RC_1 &= 100\% \end{aligned}$$

ii. Recall for Class 2

$$\begin{aligned} RC_2 &= \left(\frac{\text{Correct classifications to class 2}}{\text{Total data points in class 2}}\right) \times 100 \\ RC_2 &= \left(\frac{125}{125}\right) \times 100 \\ RC_2 &= 100\% \end{aligned}$$

iii. Recall for Class 3

$$\begin{aligned} RC_3 &= \left(\frac{\text{Correct classifications to class 3}}{\text{Total data points in class 3}}\right) \times 100 \\ RC_3 &= \left(\frac{125}{125}\right) \times 100 \\ RC_3 &= 100\% \end{aligned}$$

iv. Mean Recall

$$\begin{aligned} MRC &= \left(\frac{RC_1 + RC_2 + RC_3}{3}\right) \times 100 \\ MRC &= 100\% \end{aligned}$$

(d) F-Measure

i. F-Measure for Class 1

$$\begin{aligned} FM_1 &= \left(\frac{PC_1 \times RC_1 \times 2}{PC_1 + RC_1}\right) \\ FM_1 &= \left(\frac{100 \times 100 \times 2}{100 + 100}\right) \\ FM_1 &= 100 \end{aligned}$$

ii. F-Measure for Class 2

$$\begin{aligned} FM_2 &= \left(\frac{PC_2 \times RC_2 \times 2}{PC_2 + RC_2}\right) \\ FM_2 &= \left(\frac{100 \times 100 \times 2}{100 + 100}\right) \\ FM_2 &= 100 \end{aligned}$$

iii. F-Measure for Class 3

$$
\begin{aligned}
FM_3 &= \left(\frac{PC_3 \times RC_3 \times 2}{PC_3 + RC_3}\right) \\
FM_3 &= \left(\frac{100 \times 100 \times 2}{100 + 100}\right) \\
FM_3 &= 100
\end{aligned}
$$

iv. Mean F-Measure

$$
\begin{aligned}
MFM &= \left(\frac{FM_1 + FM_1 + FM_3}{3}\right) \\
MFM &= \left(\frac{100 + 100 + 100}{3}\right) \\
MFM &= 100
\end{aligned}
$$

As all parameter for the linearly separable data points are 100% (for accuracy, mean precision and mean recall) and 100 (for mean F-measure). Thus the assumption for this case works perfectly well for the data.

2.2 Non-Linearly Separable Data

Here are the plots for non-linearly separable data :

1. Decision region plot for every pair of classes together with the training data superposed.

(a)

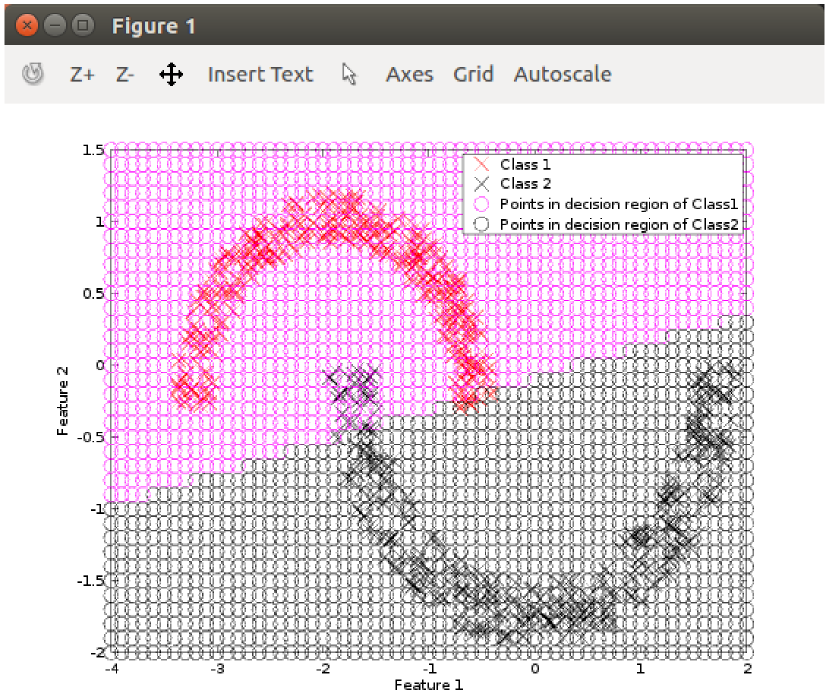


Fig1 : This is the decision plot between class 1 and class 2 with their training data superposed.

(b)

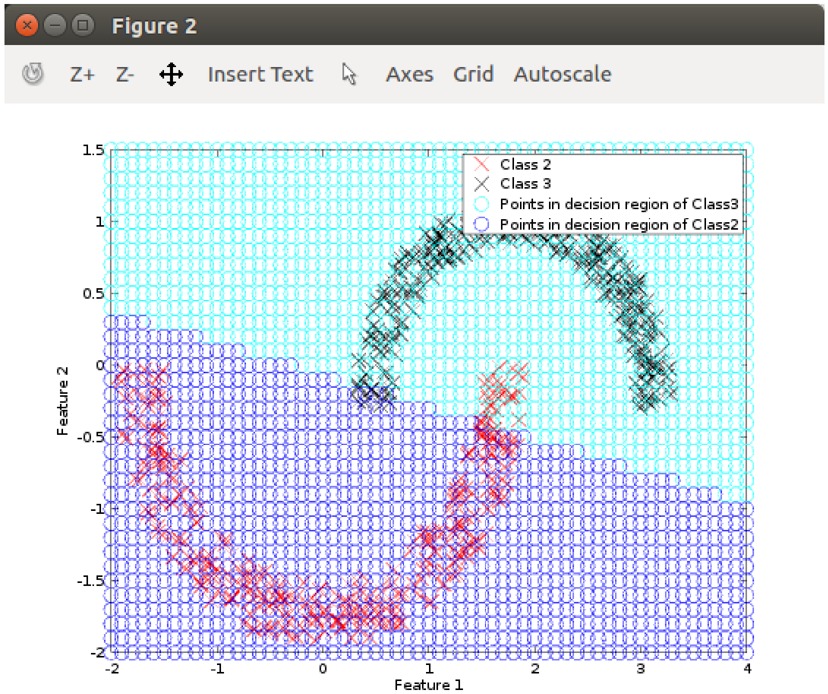


Fig2 : This is the decision plot between class 2 and class 3 with their training data superposed.

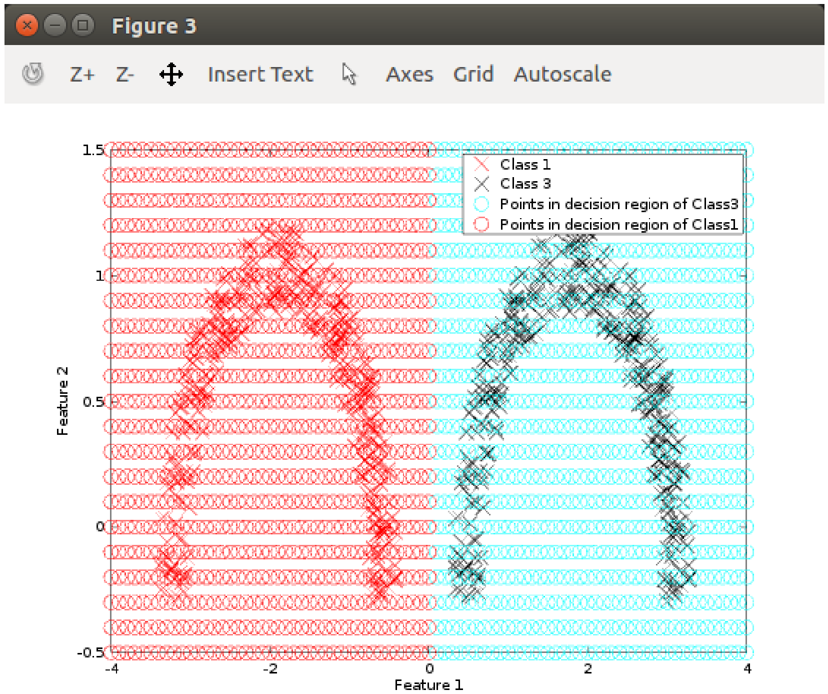
(c) 

Fig3 : This is the decision plot between class 3 and class 1 with their training data superposed.

2. Decision region plot for all the classes together with the training data superposed

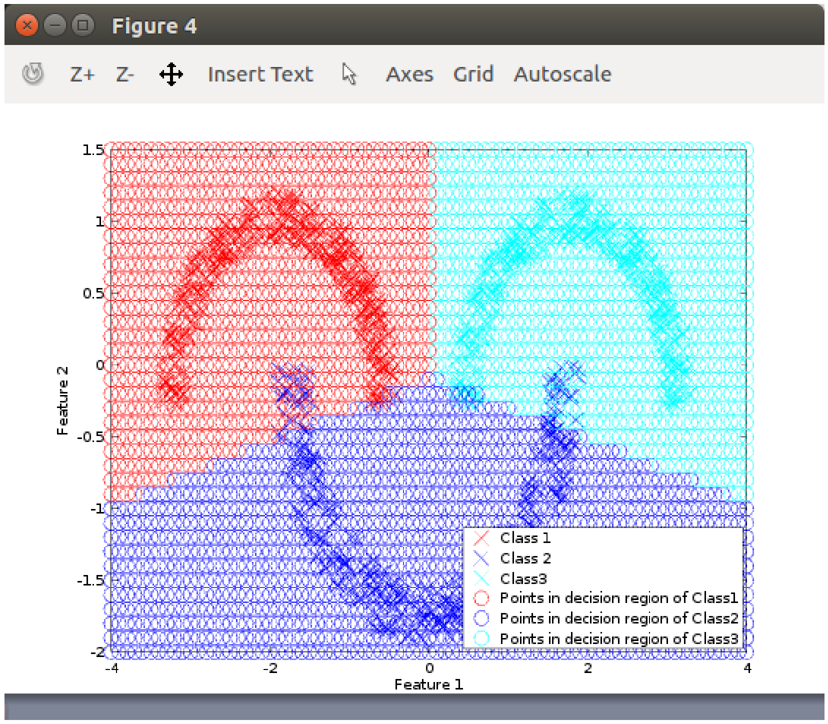
(a) 

Fig4 : This is the decision plot among class 1, class 2 and class 3 with their training data superposed.

3. Constant density contour plot for all the classes together with the training data superposed

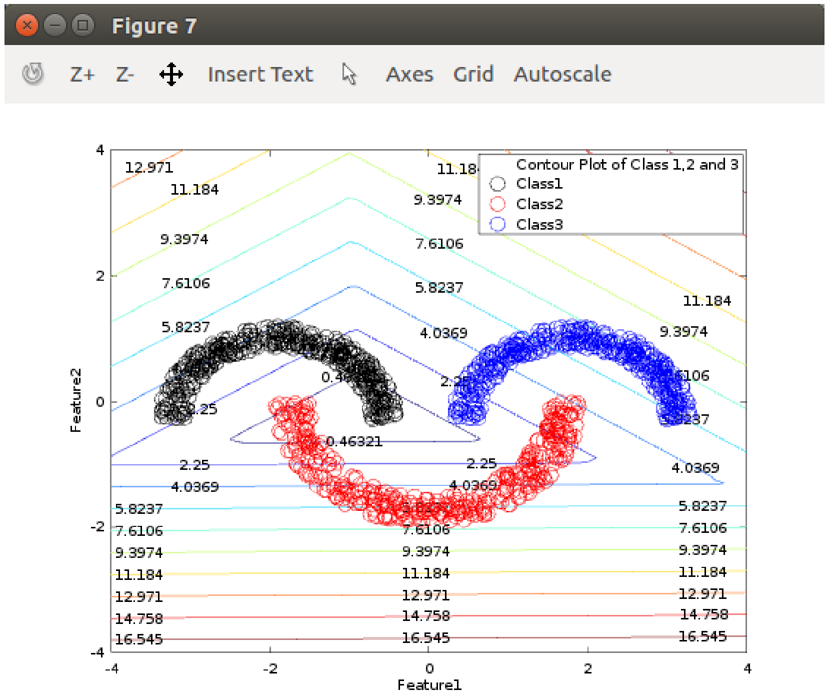
(a) 

Fig5 : This is the constant density contour plot for all the classes together with the training data superposed

4. Confusion matrix

$$M = \begin{bmatrix} 122 & 3 & 0 \\ 15 & 98 & 12 \\ 0 & 3 & 122 \end{bmatrix}$$

In this matrix row(i) corresponds to the number of test data points categorized as class i.
Column(j) corresponds to the number of test data points actually in class j.

5. Calculation of performance parameters

(a) Accuracy

$$Accuracy = 91.200\%$$

(b) Precision

i. Precision for Class 1

$$PC_1 = 89.051\%$$

ii. Precision for Class 2

$$PC_2 = 94.231\%$$

iii. Precision for Class 3

$$PC_3 = 91.045\%$$

iv. Mean Precision

$$MPC = 91.442\%$$

(c) Recall

i. Recall for Class 1

$$RC_1 = 97.600\%$$

ii. Recall for Class 2

$$RC_2 = 78.400\%$$

iii. Recall for Class 3

$$RC_3 = 97.600\%$$

iv. Mean Recall

$$MRC = 91.200\%$$

(d) F-Measure

i. F-Measure for Class 1

$$FM_1 = 93.130$$

ii. F-Measure for Class 2

$$FM_2 = 85.590$$

iii. F-Measure for Class 3

$$FM_3 = 94.208$$

iv. Mean F-Measure

$$MFM = 90.976$$

Here accuracy, precision, recall are all in or near the nineties. This case suits satisfactorily for this data set.

2.3 Real-world Data

Here are the plots for real-world data :

1. Decision region plot for every pair of classes together with the training data superposed.

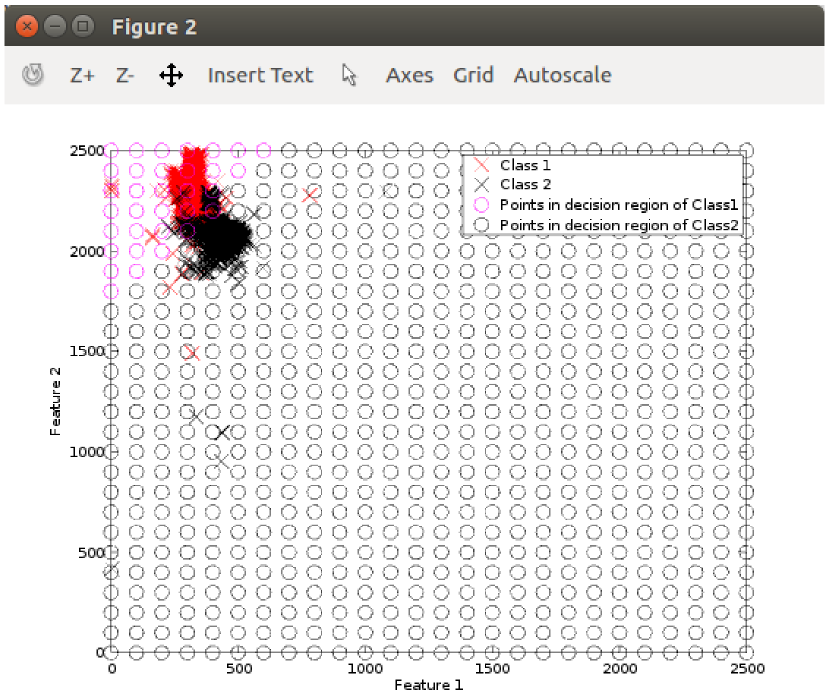
   (a) 

   Fig1 : This is the decision plot between class 1 and class 2 with their training data superposed.

(b)

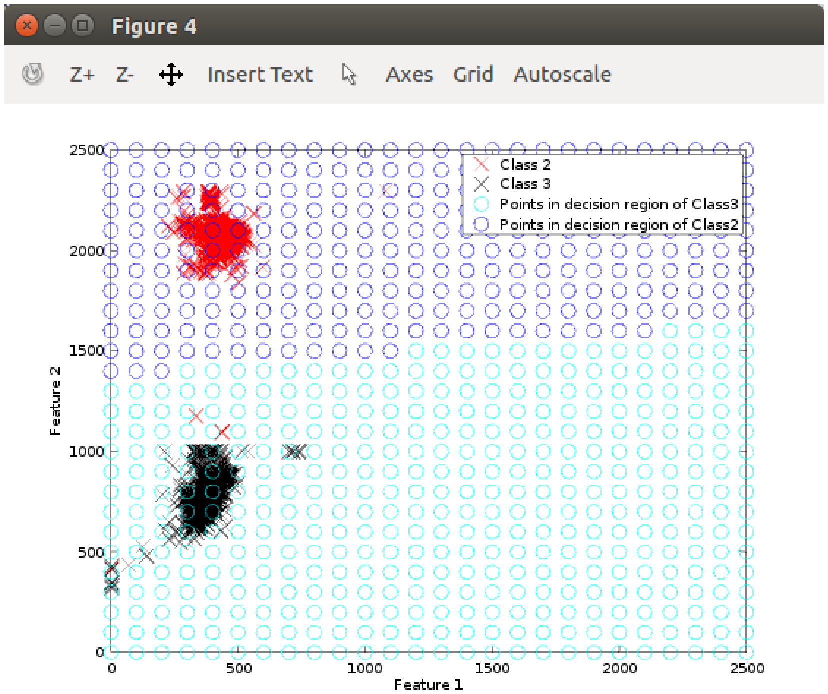


Fig2 : This is the decision plot between class 2 and class 3 with their training data superposed.

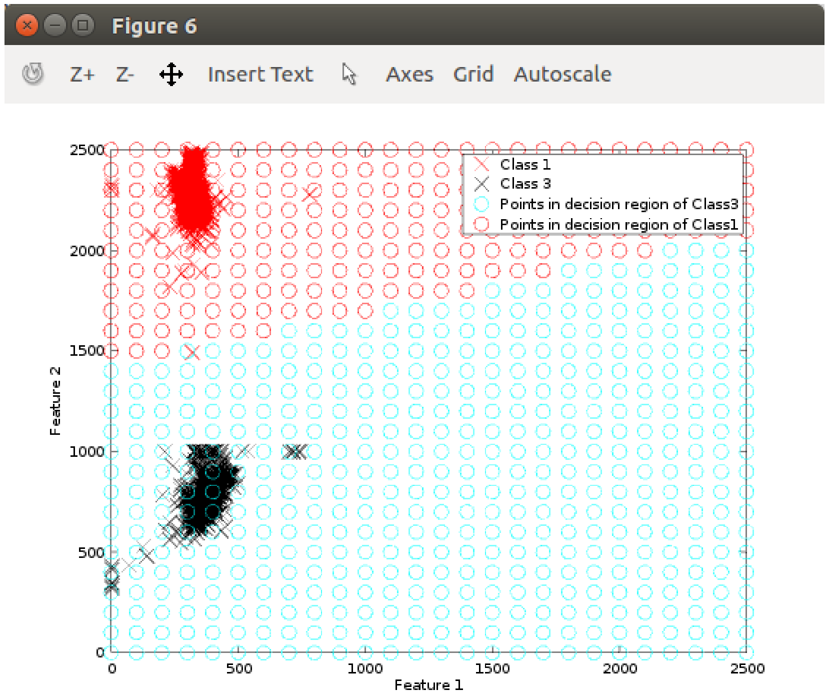
(c) 

Fig3 : This is the decision plot between class 3 and class 1 with their training data superposed.

2. Decision region plot for all the classes together with the training data superposed

(a)

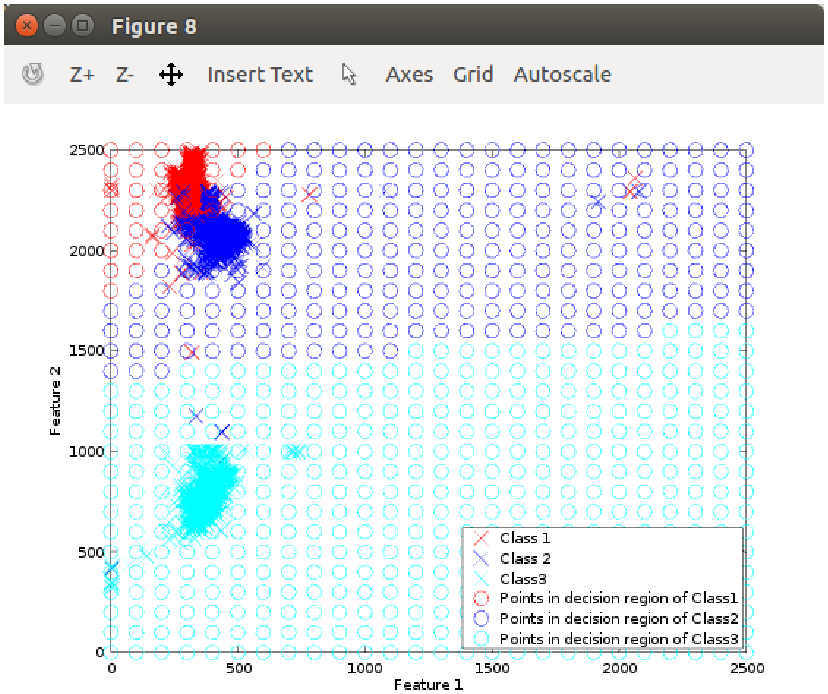


Fig4 : This is the decision plot among class 1, class 2 and class 3 with their training data superposed.

3. Constant density contour plot for all the classes together with the training data superposed

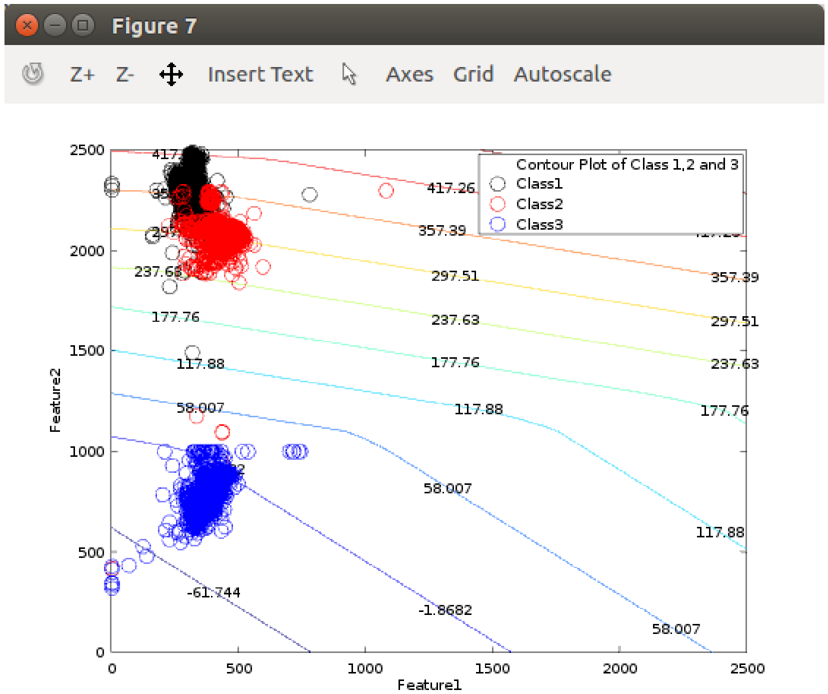
(a) 

Fig5 : This is the constant density contour plot for all the classes together with the training data superposed

4. Confusion matrix

$$M = \begin{bmatrix} 571 & 23 & 3 \\ 345 & 211 & 17 \\ 0 & 0 & 622 \end{bmatrix}$$

In this matrix row(i) corresponds to the number of test data points categorized as class i.
Column(j) corresponds to the number of test data points actually in class j.

5. Calculation of performance parameters

   (a) Accuracy

$$Accuracy \;=\; 78.348\%$$

   (b) Precision

i. Precision for Class 1

$$PC_1 = 62.336\%$$

ii. Precision for Class 2

$$PC_2 = 90.171\%$$

iii. Precision for Class 3

$$PC_3 = 96.885\%$$

iv. Mean Precision

$$MPC = 83.131\%$$

(c) Recall

i. Recall for Class 1

$$RC_1 = 95.645\%$$

ii. Recall for Class 2

$$RC_2 = 36.824\%$$

iii. Recall for Class 3

$$RC_3 = 100\%$$

iv. Mean Recall

$$MRC = 77.490\%$$

(d) F-Measure

i. F-Measure for Class 1

$$FM_1 = 75.479$$

ii. F-Measure for Class 2

$$FM_2 = 52.292$$

iii. F-Measure for Class 3

$$FM_3 = 98.418$$

iv. Mean F-Measure

$$MFM = 75.396$$

Here again the accuracy, mean recall, mean precision all are around 80% but some noticable observations prevail : 1. Precision for class 1 is not very stisfactory. 2. Recall for class 2 is poor. 3. Recall for class 3 is 100%.

3 Covariance matric is diagonal and is different for each class.

The general form of covariance matrix for a class having data which has two features is:

$$\Sigma : \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} \tag{1.7}$$

where :
σ_{11} : Covariance of feature 1 with feature 1 or variance of feature 1.
σ_{12} : Covariance of feature 1 with feature 2.
σ_{21} : Covariance of feature 2 with feature 1.
σ_{22} : Covariance of feature 2 with feature 2 or variance of feature 2.

We consider $\sigma_{12} = \sigma_{12} = 0$
So we have the covariance matrix as :

$$\Sigma : \begin{bmatrix} \sigma_{11} & 0 \\ 0 & \sigma_{22} \end{bmatrix} \tag{1.8}$$

The discriminating equation thus produced is :

$$\begin{aligned} g_i(\mathbf{x}) &= \mathbf{x}^t W_i \mathbf{x} + {w_i}^t \mathbf{x} + C_{i0} \\ W_i &= -\frac{1}{2} {\Sigma_i}^{-1} \\ w_i &= {\Sigma_i}^{-1} \mu_i \\ C_{i0} &= -\frac{1}{2} {\mu_i}^t {\Sigma_i}^{-1} \mu_i - \frac{1}{2} \ln |\Sigma_i| + \ln P(C_i) \end{aligned}$$

${}_m u_i$: Mean of i^{th} class. $\ln P(C_i)$: Prior of i^{th} class.

3.1 Linearly Separable Data

Here are the plots for linearly separable data :

1. Decision region plot for every pair of classes together with the training data superposed.

(a)

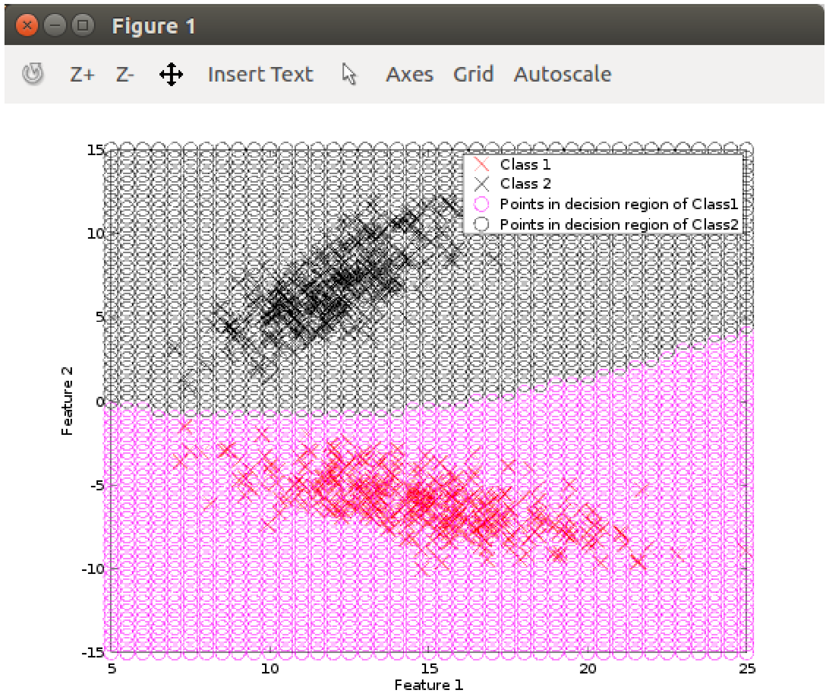


Fig1 : This is the decision plot between class 1 and class 2 with their training data superposed.

(b)

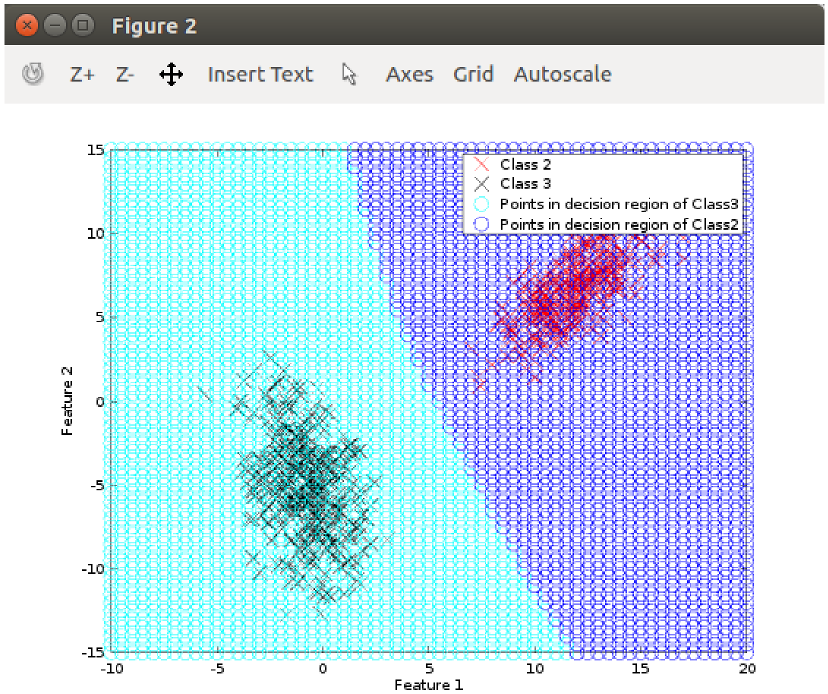


Fig2 : This is the decision plot between class 2 and class 3 with their training data superposed.

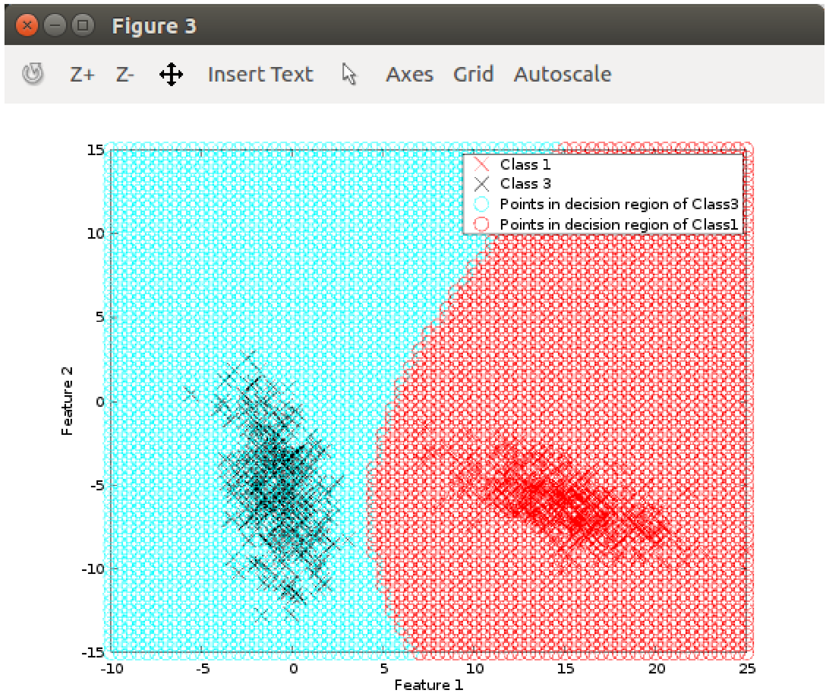
(c) 

Fig3 : This is the decision plot between class 3 and class 1 with their training data superposed.

2. Decision region plot for all the classes together with the training data superposed

(a)

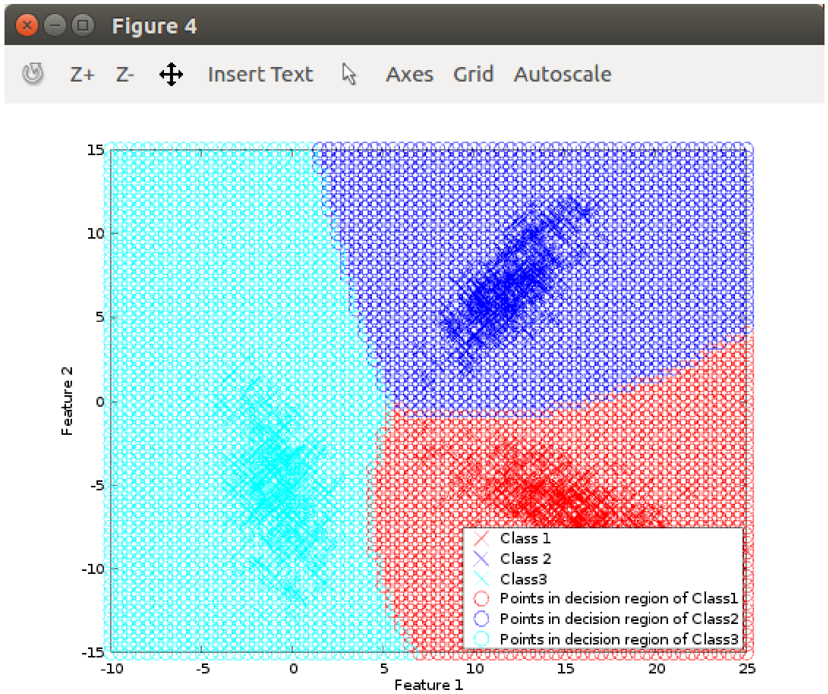


Fig4 : This is the decision plot among class 1, class 2 and class 3 with their training data superposed.

3. Constant density contour plot for all the classes together with the training data superposed

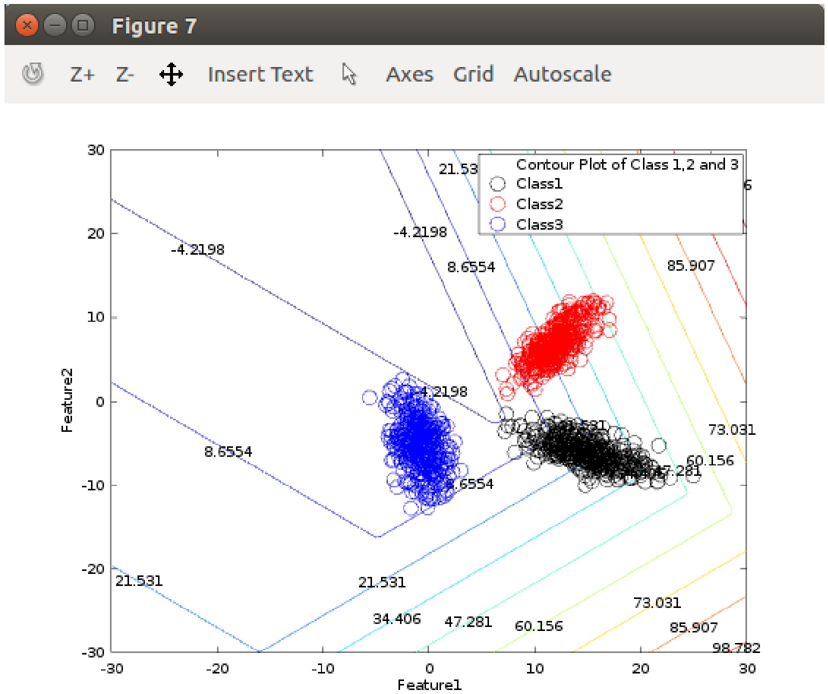
(a) 

Fig5 : This is the constant density contour plot for all the classes together with the training data superposed

4. Confusion matrix

$$M = \begin{bmatrix} 125 & 0 & 0 \\ 0 & 125 & 0 \\ 0 & 0 & 125 \end{bmatrix}$$

In this matrix row(i) corresponds to the number of test data points categorized as class i.
Column(j) corresponds to the number of test data points actually in class j.

5. Calculation of performance parameters

(a) Accuracy

$$\begin{aligned} Accuracy &= \left(\frac{\text{Total correct classifications}}{\text{Total classifications}}\right) \times 100 \\ Accuracy &= \left(\frac{125+125+125}{125+125+125}\right) \times 100 \\ &= 100\% \end{aligned}$$

(b) Precision

i. Precision for Class 1

$$\begin{aligned} PC_1 &= \left(\frac{\text{Correct classifications to class 1}}{\text{Total classifications to class 1}}\right) \times 100 \\ PC_1 &= \left(\frac{125}{125+0+0}\right) \times 100 \\ PC_1 &= 100\% \end{aligned}$$

ii. Precision for Class 2

$$\begin{aligned} PC_2 &= \left(\frac{\text{Correct classifications to class 2}}{\text{Total classifications to class 2}}\right) \times 100 \\ PC_2 &= \left(\frac{125}{0+125+0}\right) \times 100 \\ PC_2 &= 100\% \end{aligned}$$

iii. Precision for Class 3

$$\begin{aligned} PC_3 &= \left(\frac{\text{Correct classifications to class 3}}{\text{Total classifications to class 3}}\right) \times 100 \\ PC_3 &= \left(\frac{125}{0+0+125}\right) \times 100 \\ PC_3 &= 100\% \end{aligned}$$

iv. Mean Precision

$$\begin{aligned} MRC &= \left(\frac{PC_1+PC_2+PC_3}{3}\right) \times 100 \\ MRC &= 100\% \end{aligned}$$

(c) Recall

i. Recall for Class 1

$$
\begin{aligned}
RC_1 &= \left(\frac{\text{Correct classifications to class 1}}{\text{Total data points in class 1}}\right) \times 100 \\
RC_1 &= \left(\frac{125}{125}\right) \times 100 \\
RC_1 &= 100\%
\end{aligned}
$$

ii. Recall for Class 2

$$
\begin{aligned}
RC_2 &= \left(\frac{\text{Correct classifications to class 2}}{\text{Total data points in class 2}}\right) \times 100 \\
RC_2 &= \left(\frac{125}{125}\right) \times 100 \\
RC_2 &= 100\%
\end{aligned}
$$

iii. Recall for Class 3

$$
\begin{aligned}
RC_3 &= \left(\frac{\text{Correct classifications to class 3}}{\text{Total data points in class 3}}\right) \times 100 \\
RC_3 &= \left(\frac{125}{125}\right) \times 100 \\
RC_3 &= 100\%
\end{aligned}
$$

iv. Mean Recall

$$
\begin{aligned}
MPC &= \left(\frac{RC_1 + RC_2 + RC_3}{3}\right) \times 100 \\
MPC &= 100\%
\end{aligned}
$$

(d) F-Measure

i. F-Measure for Class 1

$$
\begin{aligned}
FM_1 &= \left(\frac{PC_1 \times RC_1 \times 2}{PC_1 + RC_1}\right) \\
FM_1 &= \left(\frac{100 \times 100 \times 2}{100 + 100}\right) \\
FM_1 &= 100
\end{aligned}
$$

ii. F-Measure for Class 2

$$
\begin{aligned}
FM_2 &= \left(\frac{PC_2 \times RC_2 \times 2}{PC_2 + RC_2}\right) \\
FM_2 &= \left(\frac{100 \times 100 \times 2}{100 + 100}\right) \\
FM_2 &= 100
\end{aligned}
$$

iii. F-Measure for Class 3

$$\begin{aligned} FM_3 &= \left(\frac{PC_3 \times RC_3 \times 2}{PC_3 + RC_3}\right) \\ FM_3 &= \left(\frac{100 \times 100 \times 2}{100 + 100}\right) \\ FM_3 &= 100 \end{aligned}$$

iv. Mean F-Measure

$$\begin{aligned} MFM &= \left(\frac{FM_1 + FM_1 + FM_3}{3}\right) \\ MFM &= \left(\frac{100 + 100 + 100}{3}\right) \\ MFM &= 100 \end{aligned}$$

As all parameter for the linearly separable data points are 100% (for accuracy, mean precision and mean recall) and 100 (for mean F-measure). Thus the assumption for this case works perfectly well for the data.

3.2 Non-Linearly Separable Data

Here are the plots for non-linearly separable data :

1. Decision region plot for every pair of classes together with the training data superposed.

(a)

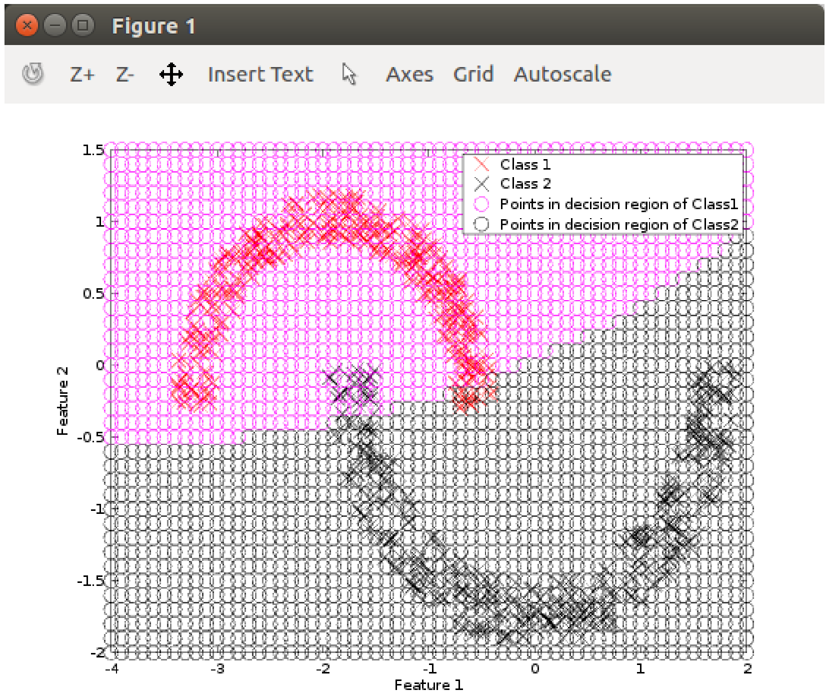


Fig1 : This is the decision plot between class 1 and class 2 with their training data superposed.

(b)

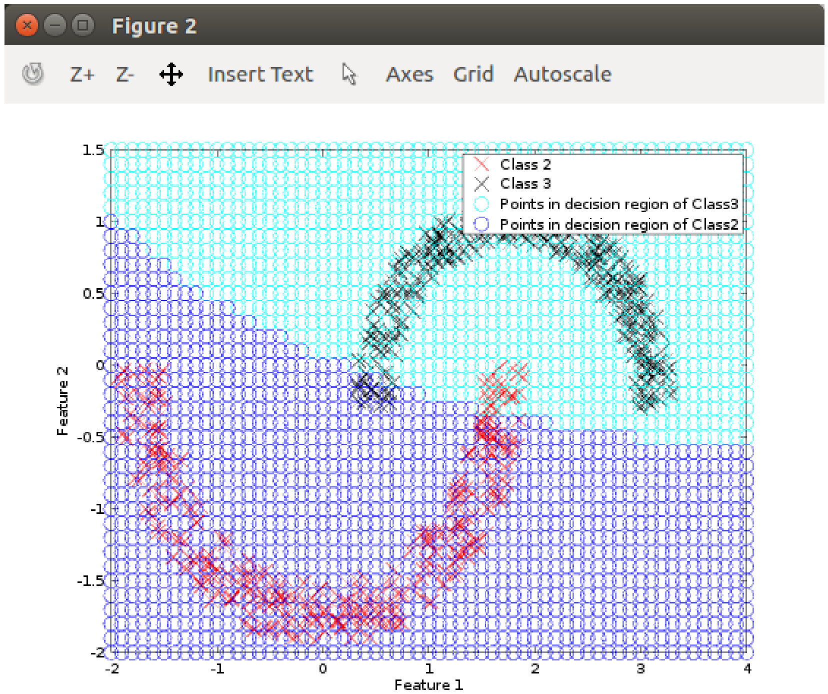


Fig2 : This is the decision plot between class 2 and class 3 with their training data superposed.

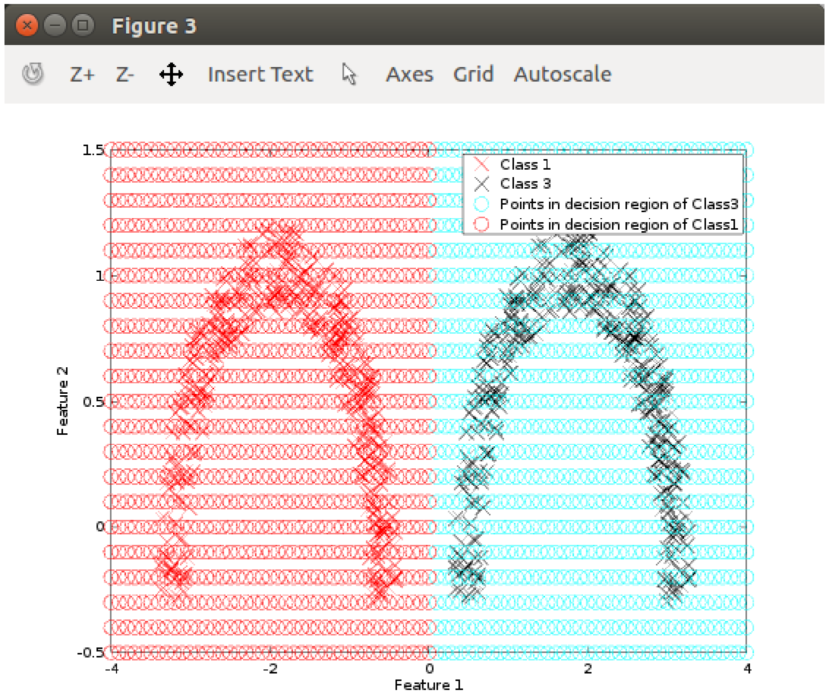
(c) 

Fig3 : This is the decision plot between class 3 and class 1 with their training data superposed.

2. Decision region plot for all the classes together with the training data superposed

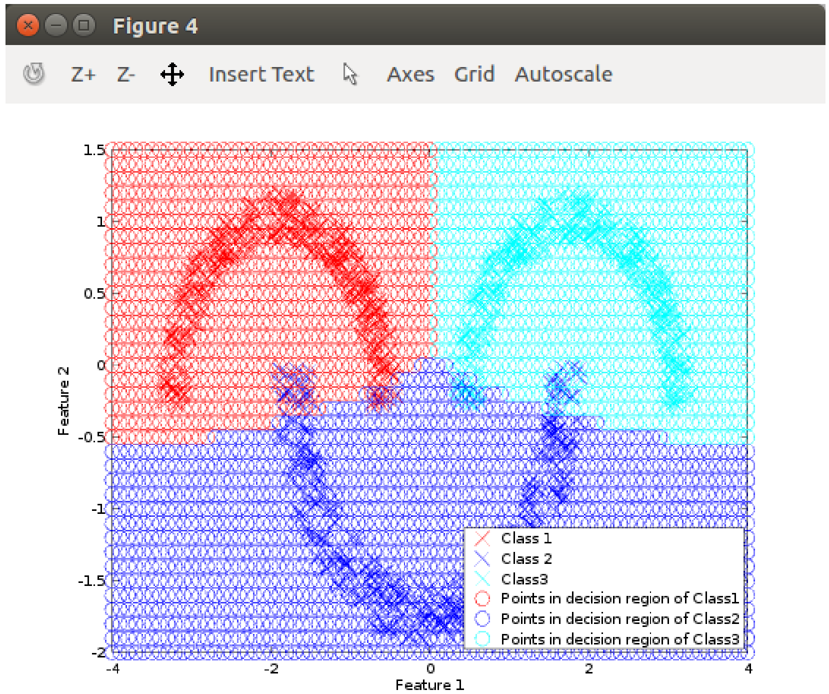
(a) 

Fig4 : This is the decision plot among class 1, class 2 and class 3 with their training data superposed.

3. Constant density contour plot for all the classes together with the training data superposed

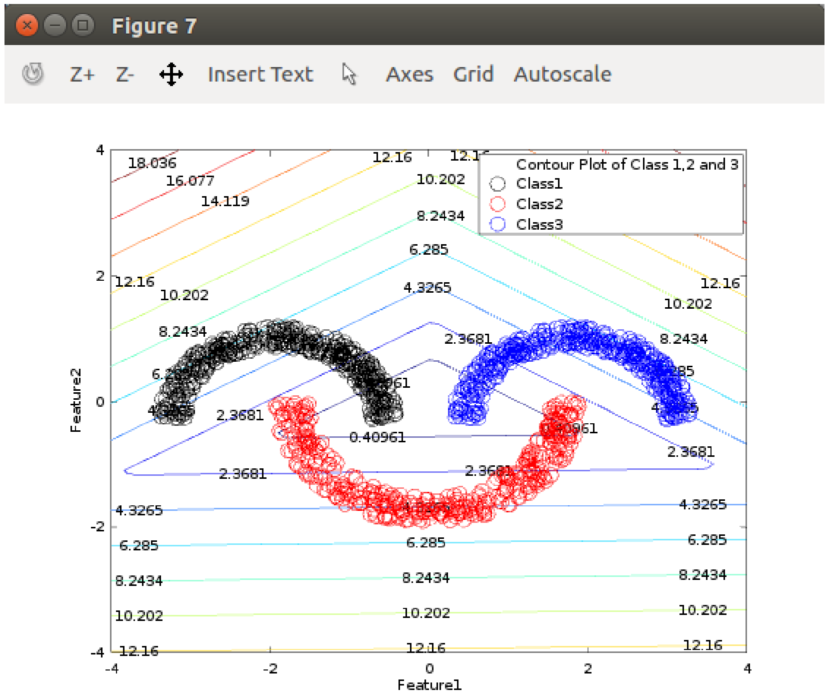
(a) 

Fig5 : This is the constant density contour plot for all the classes together with the training data superposed

4. Confusion matrix

$$M = \begin{bmatrix} 117 & 8 & 0 \\ 12 & 106 & 7 \\ 0 & 3 & 122 \end{bmatrix}$$

In this matrix row(i) corresponds to the number of test data points categorized as class i.
Column(j) corresponds to the number of test data points actually in class j.

5. Calculation of performance parameters

   (a) Accuracy

$$Accuracy \quad = \quad 92\%$$

   (b) Precision

i. Precision for Class 1

$$PC_1 = 90.698\%$$

ii. Precision for Class 2

$$PC_2 = 90.598\%$$

iii. Precision for Class 3

$$PC_3 = 94.574\%$$

iv. Mean Precision

$$MPC = 91.957\%$$

(c) Recall

i. Recall for Class 1

$$RC_1 = 93.600\%$$

ii. Recall for Class 2

$$RC_2 = 84.800\%$$

iii. Recall for Class 3

$$RC_3 = 97.600\%$$

iv. Mean Recall

$$MRC = 92\%$$

(d) F-Measure

i. F-Measure for Class 1

$$FM_1 = 92.126$$

ii. F-Measure for Class 2

$$FM_2 = 87.603$$

iii. F-Measure for Class 3

$$FM_3 = 96.063$$

iv. Mean F-Measure

$$MFM = 91.931$$

Here accuracy, precision, recall are all in or near the nineties. This case suits satisfactorily for this data set.

3.3 Real-world Data

Here are the plots for real world separable data :

1. Decision region plot for every pair of classes together with the training data superposed.

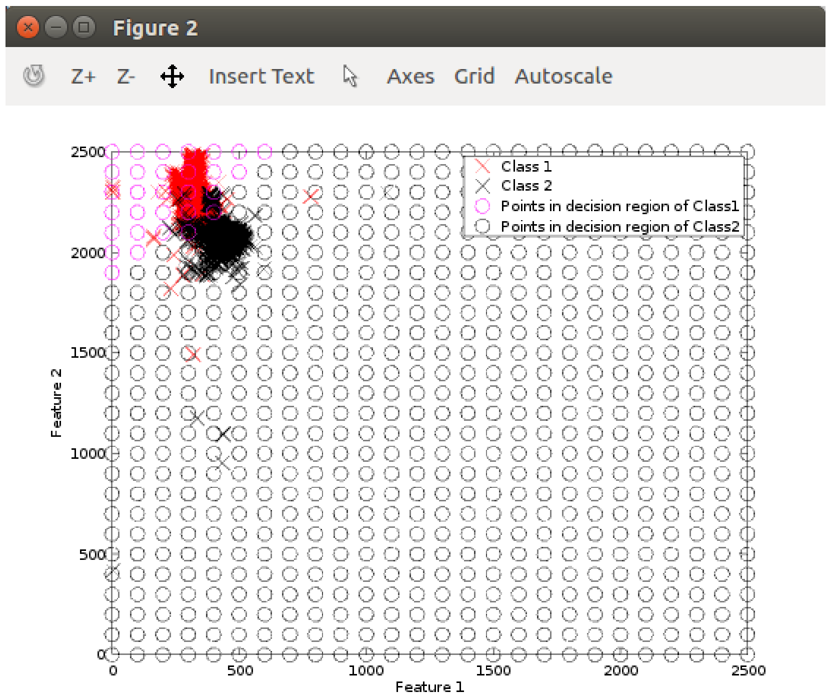
   (a) 

   Fig1 : This is the decision plot between class 1 and class 2 with their training data superposed.

(b)

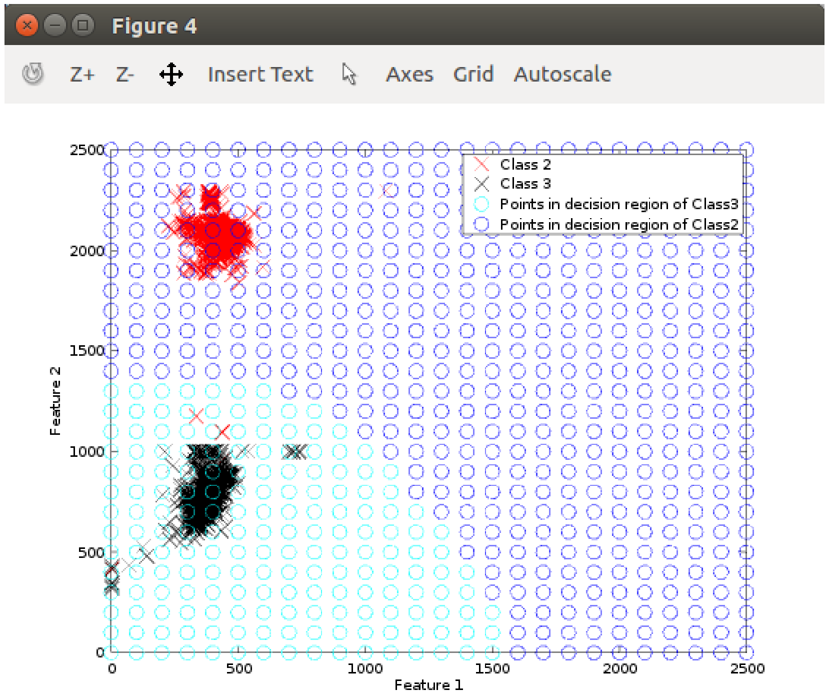


Fig2 : This is the decision plot between class 2 and class 3 with their training data superposed.

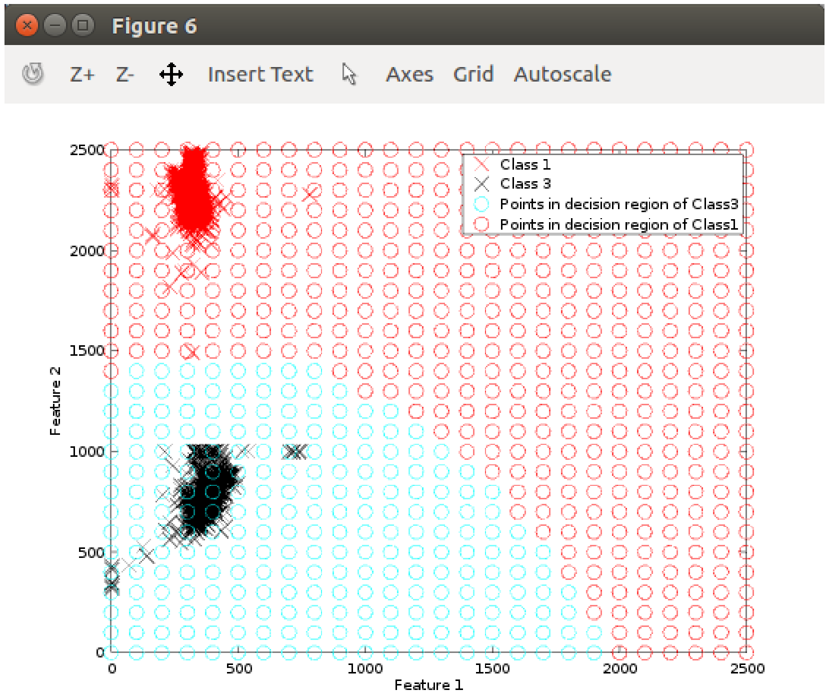
(c) 

Fig3 : This is the decision plot between class 3 and class 1 with their training data superposed.

2. Decision region plot for all the classes together with the training data superposed

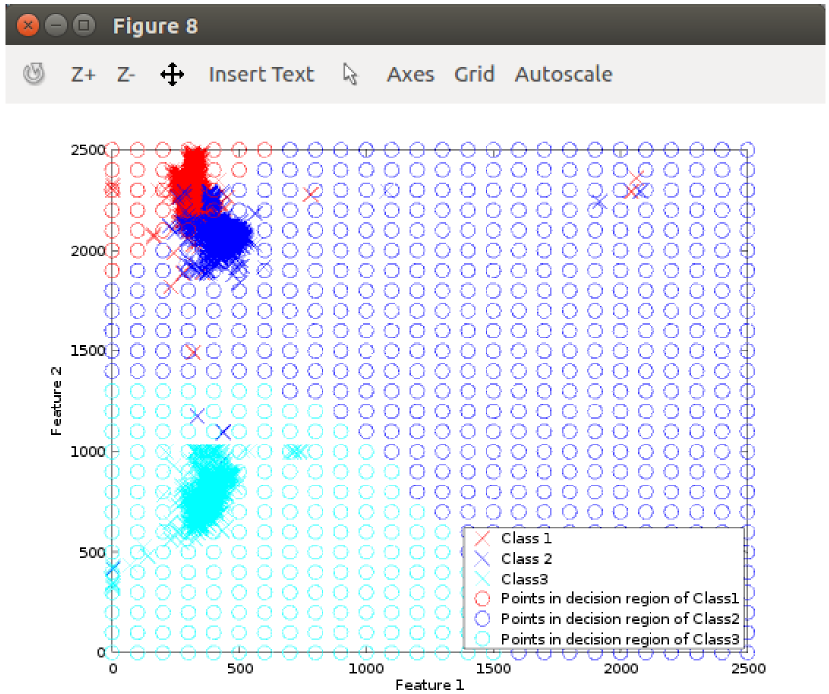
(a) 

Fig4 : This is the decision plot among class 1, class 2 and class 3 with their training data superposed.

3. Constant density contour plot for all the classes together with the training data superposed

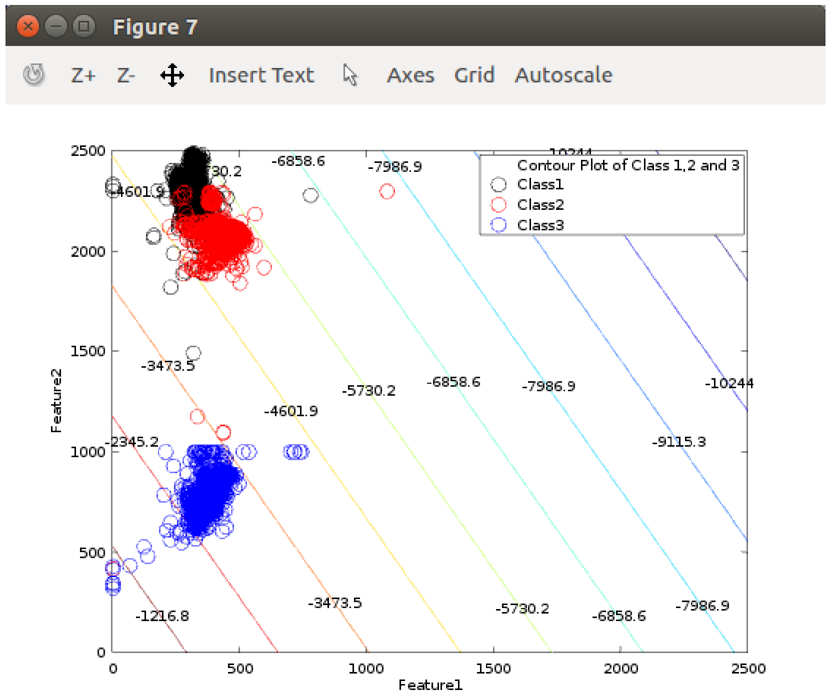
(a) 

Fig5 : This is the constant density contour plot for all the classes together with the training data superposed

4. Confusion matrix

$$M = \begin{bmatrix} 571 & 23 & 3 \\ 346 & 210 & 17 \\ 0 & 0 & 622 \end{bmatrix}$$

In this matrix row(i) corresponds to the number of test data points categorized as class i.
Column(j) corresponds to the number of test data points actually in class j.

5. Calculation of performance parameters

(a) Accuracy

$$Accuracy \quad = \quad 78.292\%$$

(b) Precision

i. Precision for Class 1

$$PC_1 = 62.268\%$$

ii. Precision for Class 2

$$PC_2 = 90.129\%$$

iii. Precision for Class 3

$$PC_3 = 96.885\%$$

iv. Mean Precision

$$MPC = 83.094\%$$

(c) Recall

i. Recall for Class 1

$$RC_1 = 95.645\%$$

ii. Recall for Class 2

$$RC_2 = 36.649\%$$

iii. Recall for Class 3

$$RC_3 = 100\%$$

iv. Mean Recall

$$MRC = 77.431\%$$

(d) F-Measure

i. F-Measure for Class 1

$$FM_1 = 75.429$$

ii. F-Measure for Class 2

$$FM_2 = 52.109$$

iii. F-Measure for Class 3

$$FM_3 = 98.418$$

iv. Mean F-Measure

$$MFM = 75.319$$

Here again the accuracy, mean recall, mean precision all are around 80% but some noticable observations prevail : 1. Precision for class 1 is not very stisfactory. 2. Recall for class 2 is poor. 3. Recall for class 3 is 100%.

4 Full Covariance matrix for each class is different.

The general form of covariance matrix for a class having data which has two features is:

$$\Sigma : \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} \tag{1.9}$$

where :
σ_{11} : Covariance of feature 1 with feature 1 or variance of feature 1.
σ_{12} : Covariance of feature 1 with feature 2.
σ_{21} : Covariance of feature 2 with feature 1.
σ_{22} : Covariance of feature 2 with feature 2 or variance of feature 2.

The discriminating equation thus produced is :

$$\begin{aligned} g_i(\mathbf{x}) &= \mathbf{x}^t W_i \mathbf{x} + {w_i}^t \mathbf{x} + C_{i0} \\ W_i &= -\frac{1}{2}{\Sigma_i}^{-1} \\ w_i &= {\Sigma_i}^{-1}\mu_i \\ C_{i0} &= -\frac{1}{2}{\mu_i}^t {\Sigma_i}^{-1}\mu_i - \frac{1}{2}\ln|\Sigma_i| + \ln P(C_i) \end{aligned}$$

${}_m u_i$: Mean of i^{th} class. $\ln P(C_i)$: Prior of i^{th} class.

4.1 Linearly Separable Data

Here are the plots for linearly separable data :

1. Decision region plot for every pair of classes together with the training data superposed.

(a)

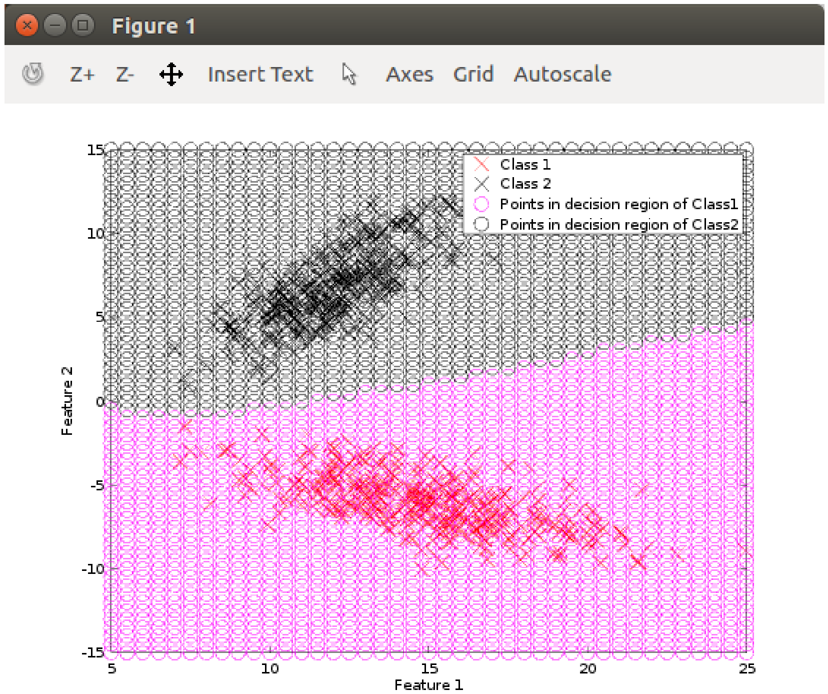


Fig1 : This is the decision plot between class 1 and class 2 with their training data superposed.

(b)

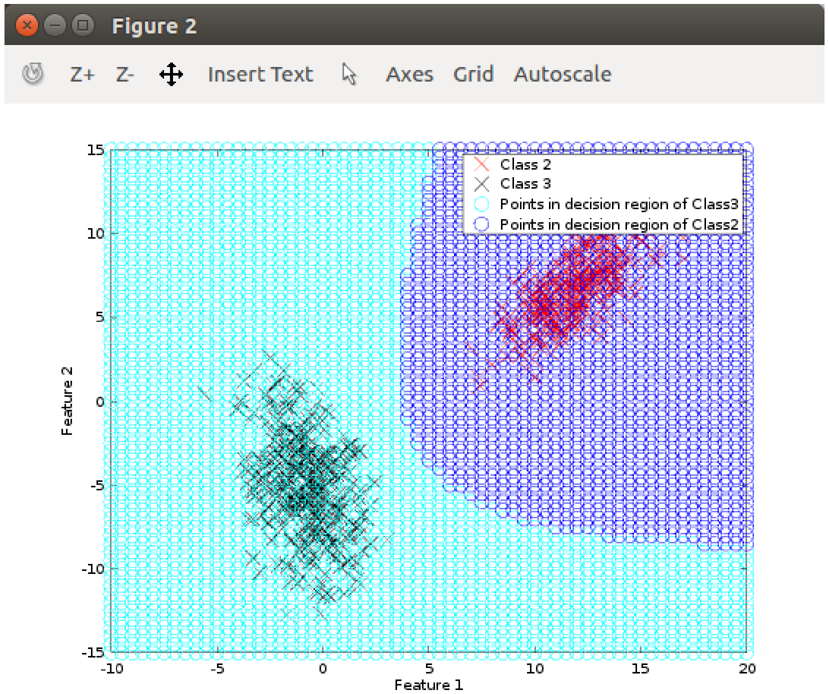


Fig2 : This is the decision plot between class 2 and class 3 with their training data superposed.

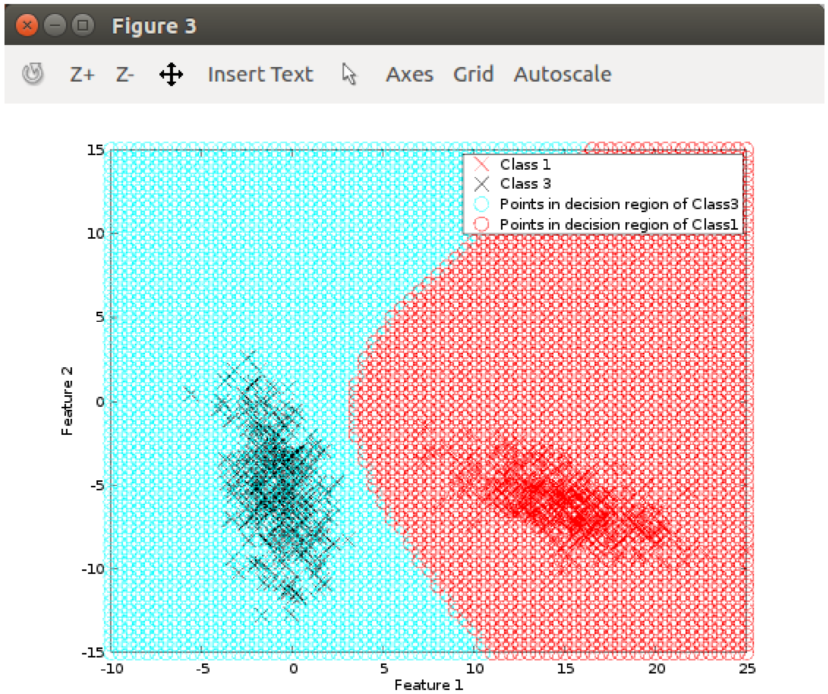
(c) 

Fig3 : This is the decision plot between class 3 and class 1 with their training data superposed.

2. Decision region plot for all the classes together with the training data superposed

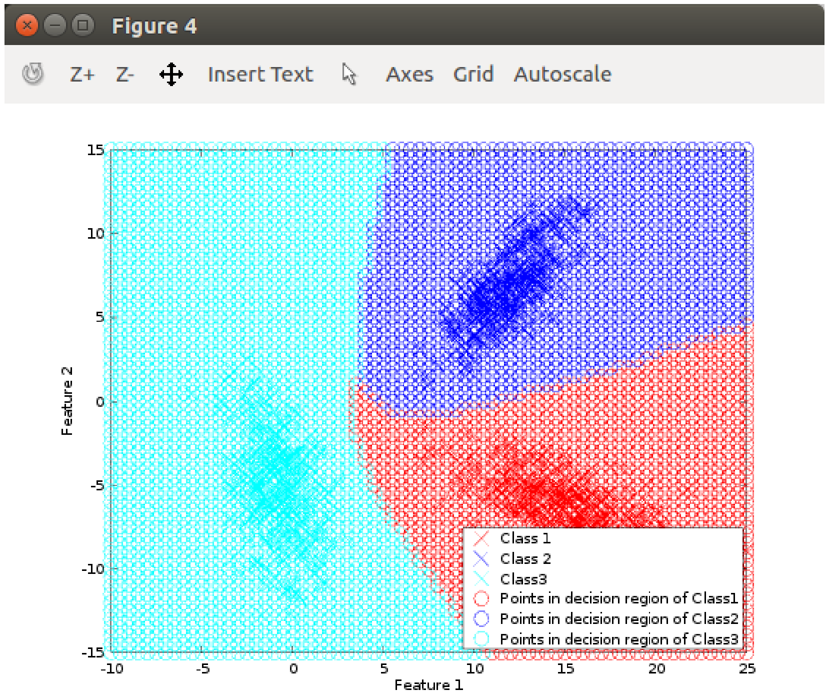
(a) 

Fig4 : This is the decision plot among class 1, class 2 and class 3 with their training data superposed.

3. Constant density contour plot for all the classes together with the training data superposed

(a)

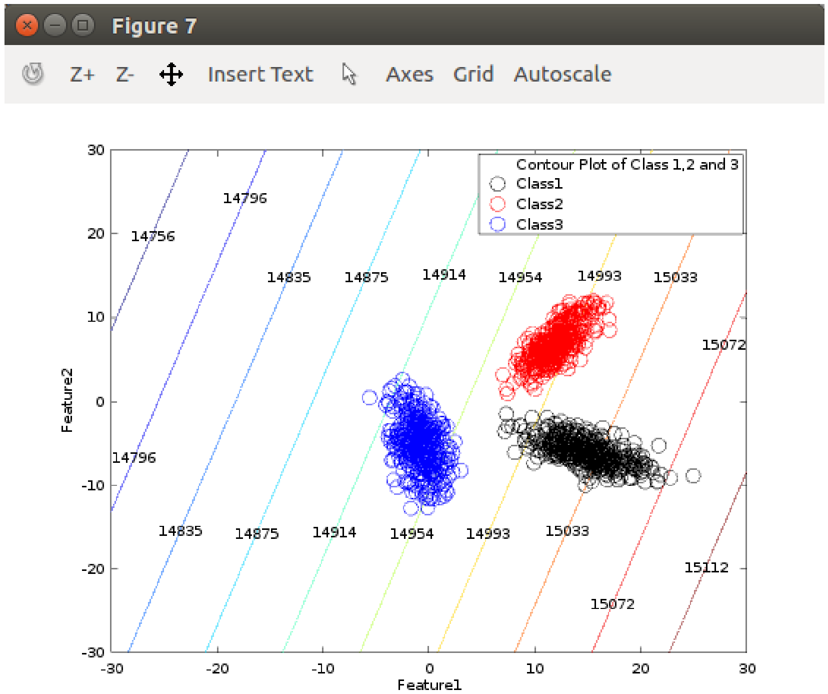


Fig5 : This is the constant density contour plot for all the classes together with the training data superposed

4. Confusion matrix

$$M = \begin{bmatrix} 125 & 0 & 0 \\ 0 & 125 & 0 \\ 0 & 0 & 125 \end{bmatrix}$$

In this matrix row(i) corresponds to the number of test data points categorized as class i.
Column(j) corresponds to the number of test data points actually in class j.

5. Calculation of performance parameters

(a) Accuracy

$$
\begin{aligned}
Accuracy &= \left(\frac{\text{Total correct classifications}}{\text{Total classifications}}\right) \times 100 \\
Accuracy &= \left(\frac{125+125+125}{125+125+125}\right) \times 100 \\
&= 100\%
\end{aligned}
$$

(b) Precision

i. Precision for Class 1

$$
\begin{aligned}
PC_1 &= \left(\frac{\text{Correct classifications to class 1}}{\text{Total classifications to class 1}}\right) \times 100 \\
PC_1 &= \left(\frac{125}{125+0+0}\right) \times 100 \\
PC_1 &= 100\%
\end{aligned}
$$

ii. Precision for Class 2

$$
\begin{aligned}
PC_2 &= \left(\frac{\text{Correct classifications to class 2}}{\text{Total classifications to class 2}}\right) \times 100 \\
PC_2 &= \left(\frac{125}{0+125+0}\right) \times 100 \\
PC_2 &= 100\%
\end{aligned}
$$

iii. Precision for Class 3

$$
\begin{aligned}
PC_3 &= \left(\frac{\text{Correct classifications to class 3}}{\text{Total classifications to class 3}}\right) \times 100 \\
PC_3 &= \left(\frac{125}{0+0+125}\right) \times 100 \\
PC_3 &= 100\%
\end{aligned}
$$

iv. Mean Precision

$$
\begin{aligned}
MRC &= \left(\frac{PC_1+PC_2+PC_3}{3}\right) \times 100 \\
MRC &= 100\%
\end{aligned}
$$

(c) Recall

i. Recall for Class 1

$$
\begin{aligned}
RC_1 &= \left(\frac{\text{Correct classifications to class 1}}{\text{Total data points in class 1}}\right) \times 100 \\
RC_1 &= \left(\frac{125}{125}\right) \times 100 \\
RC_1 &= 100\%
\end{aligned}
$$

ii. Recall for Class 2

$$
\begin{aligned}
RC_2 &= \left(\frac{\text{Correct classifications to class 2}}{\text{Total data points in class 2}}\right) \times 100 \\
RC_2 &= \left(\frac{125}{125}\right) \times 100 \\
RC_2 &= 100\%
\end{aligned}
$$

iii. Recall for Class 3

$$
\begin{aligned}
RC_3 &= \left(\frac{\text{Correct classifications to class 3}}{\text{Total data points in class 3}}\right) \times 100 \\
RC_3 &= \left(\frac{125}{125}\right) \times 100 \\
RC_3 &= 100\%
\end{aligned}
$$

iv. Mean Recall

$$
\begin{aligned}
MPC &= \left(\frac{RC_1 + RC_2 + RC_3}{3}\right) \times 100 \\
MPC &= 100\%
\end{aligned}
$$

(d) F-Measure

i. F-Measure for Class 1

$$
\begin{aligned}
FM_1 &= \left(\frac{PC_1 \times RC_1 \times 2}{PC_1 + RC_1}\right) \\
FM_1 &= \left(\frac{100 \times 100 \times 2}{100 + 100}\right) \\
FM_1 &= 100
\end{aligned}
$$

ii. F-Measure for Class 2

$$
\begin{aligned}
FM_2 &= \left(\frac{PC_2 \times RC_2 \times 2}{PC_2 + RC_2}\right) \\
FM_2 &= \left(\frac{100 \times 100 \times 2}{100 + 100}\right) \\
FM_2 &= 100
\end{aligned}
$$

iii. F-Measure for Class 3

$$\begin{aligned} FM_3 &= \left(\frac{PC_3 \times RC_3 \times 2}{PC_3 + RC_3}\right) \\ FM_3 &= \left(\frac{100 \times 100 \times 2}{100 + 100}\right) \\ FM_3 &= 100 \end{aligned}$$

iv. Mean F-Measure

$$\begin{aligned} MFM &= \left(\frac{FM_1 + FM_1 + FM_3}{3}\right) \\ MFM &= \left(\frac{100 + 100 + 100}{3}\right) \\ MFM &= 100 \end{aligned}$$

As all parameter for the linearly separable data points are 100% (for accuracy, mean precision and mean recall) and 100 (for mean F-measure). Thus the assumption for this case works perfectly well for the data.

4.2 Non-Linearly Separable Data

Here are the plots for non-linearly separable data :

1. Decision region plot for every pair of classes together with the training data superposed.

(a)

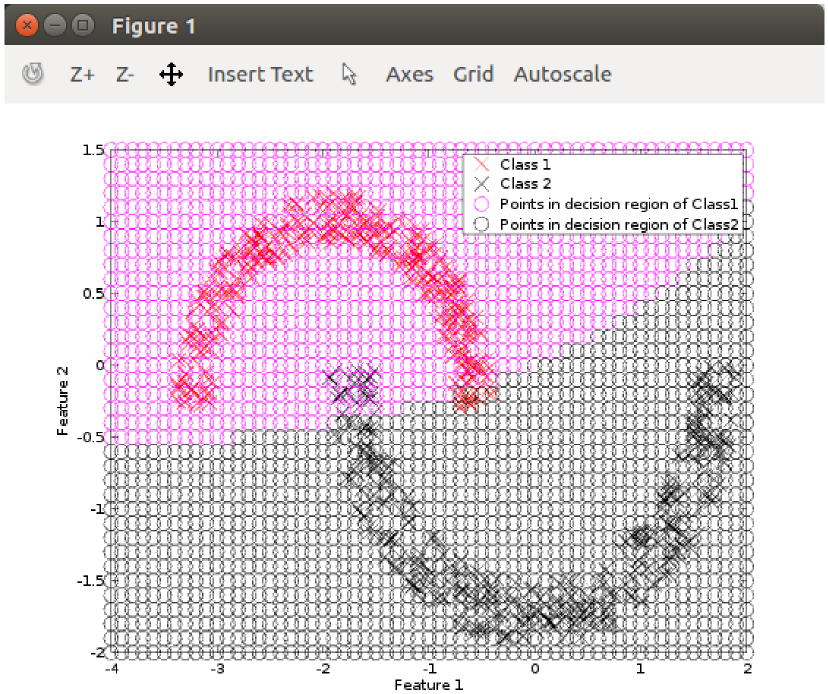


Fig1 : This is the decision plot between class 1 and class 2 with their training data superposed.

(b)

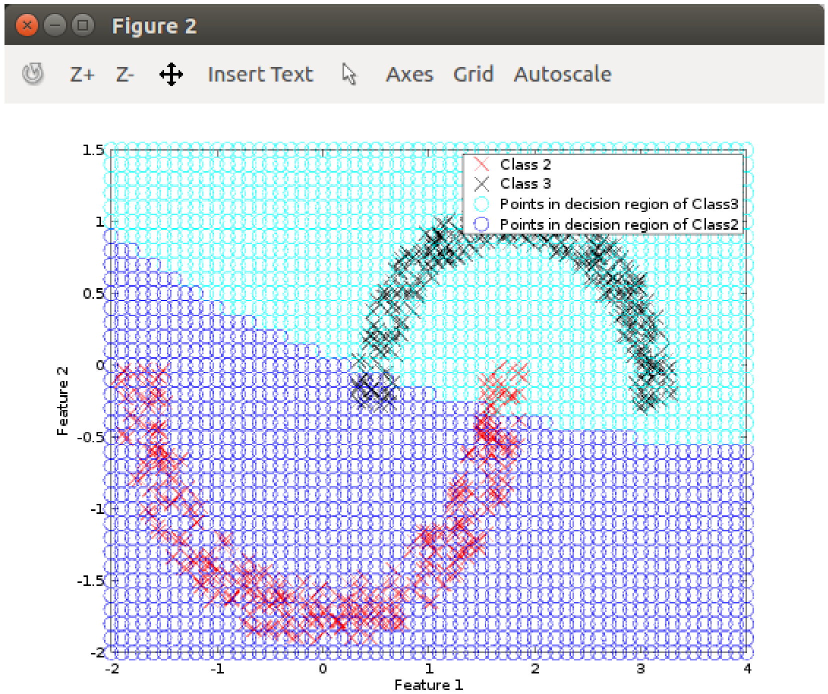


Fig2 : This is the decision plot between class 2 and class 3 with their training data superposed.

(c)

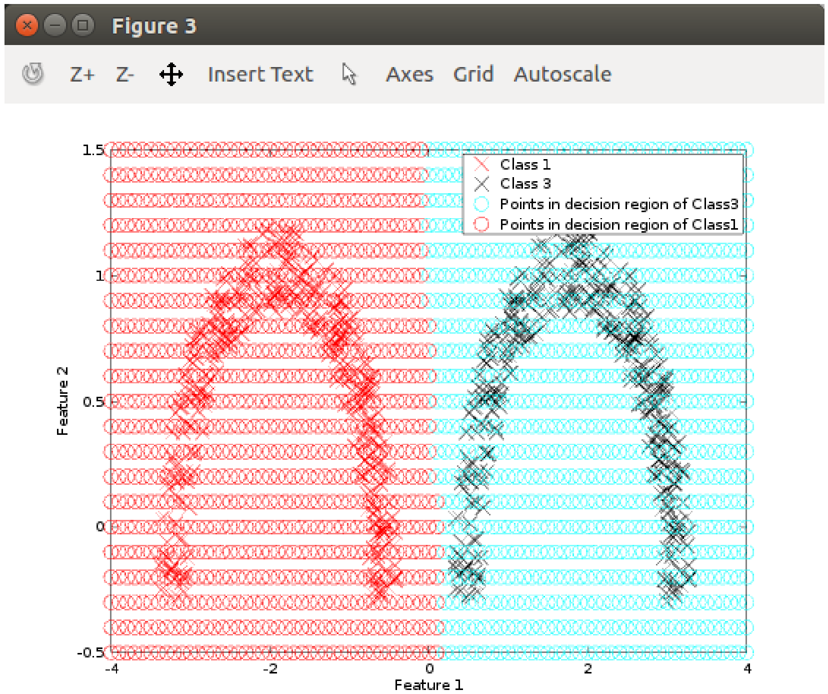


Fig3 : This is the decision plot between class 3 and class 1 with their training data superposed.

2. Decision region plot for all the classes together with the training data superposed

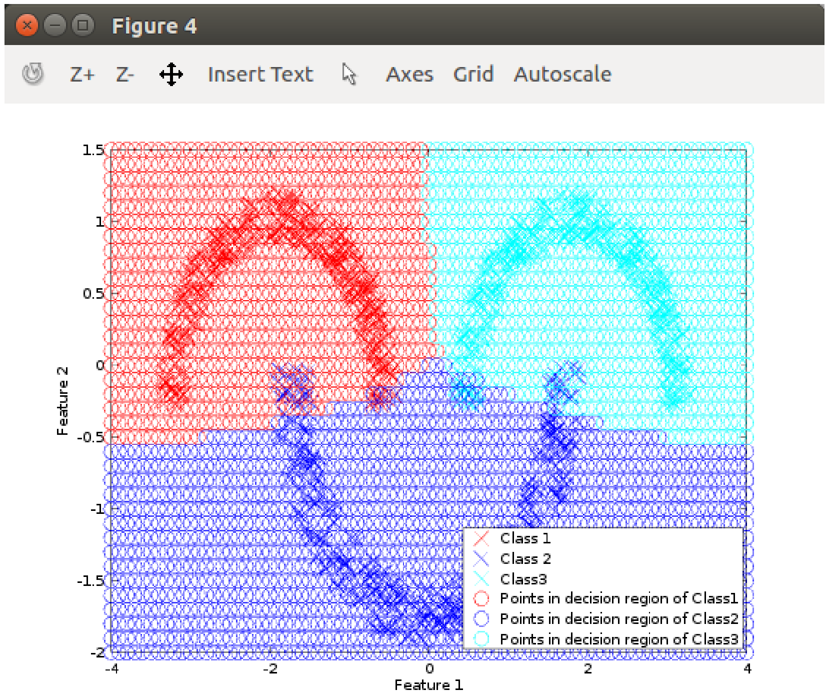
(a) 

Fig4 : This is the decision plot among class 1, class 2 and class 3 with their training data superposed.

3. Constant density contour plot for all the classes together with the training data superposed

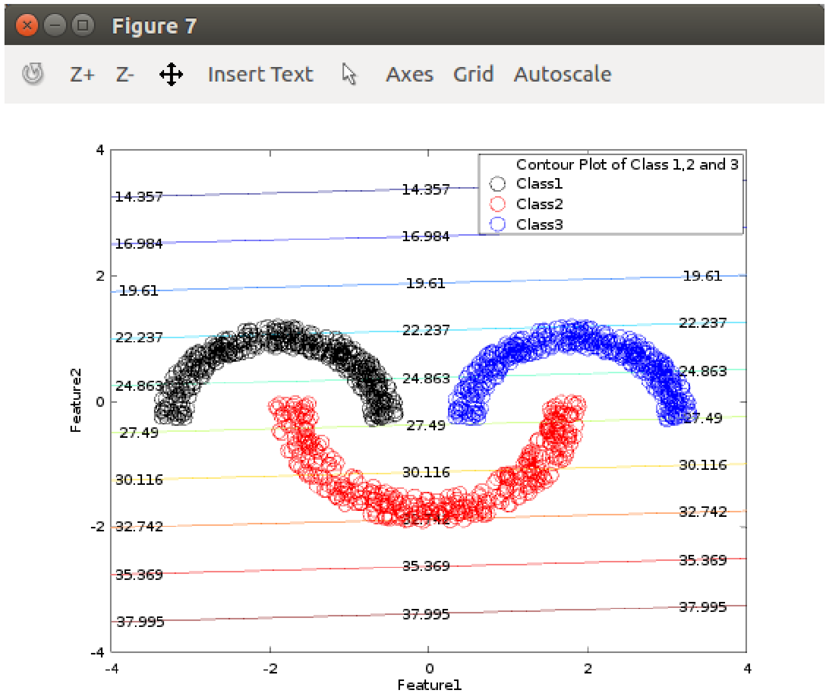
(a) 

Fig5 : This is the constant density contour plot for all the classes together with the training data superposed

4. Confusion matrix

$$M = \begin{bmatrix} 117 & 8 & 0 \\ 12 & 107 & 6 \\ 0 & 4 & 121 \end{bmatrix}$$

In this matrix row(i) corresponds to the number of test data points categorized as class i.
Column(j) corresponds to the number of test data points actually in class j.

5. Calculation of performance parameters

(a) Accuracy

$$Accuracy \quad = \quad 92\%$$

(b) Precision

i. Precision for Class 1

$$PC_1 = 90.698\%$$

ii. Precision for Class 2

$$PC_2 = 89.916\%$$

iii. Precision for Class 3

$$PC_3 = 95.276$$

iv. Mean Precision

$$MPC = 91.963\%$$

(c) Recall

i. Recall for Class 1

$$RC_1 = 93.600\%$$

ii. Recall for Class 2

$$RC_2 = 85.600\%$$

iii. Recall for Class 3

$$RC_3 = 96.800\%$$

iv. Mean Recall

$$MRC = 92\%$$

(d) F-Measure

i. F-Measure for Class 1

$$FM_1 = 92.126$$

ii. F-Measure for Class 2

$$FM_2 = 87.705$$

iii. F-Measure for Class 3

$$FM_3 = 96.032$$

iv. Mean F-Measure

$$MFM = 91.954$$

Here accuracy, precision, recall are all in or near the nineties. This case suits satisfactorily for this data set.

4.3 Real-world Data

Here are the plots for real world separable data :

1. Decision region plot for every pair of classes together with the training data superposed.

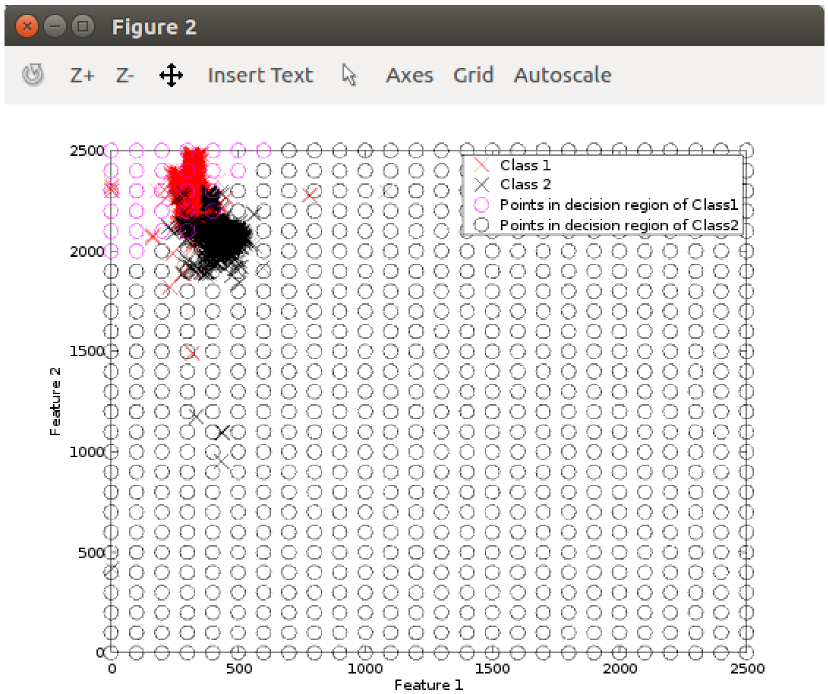
   (a) 

   Fig1 : This is the decision plot between class 1 and class 2 with their training data superposed.

(b)

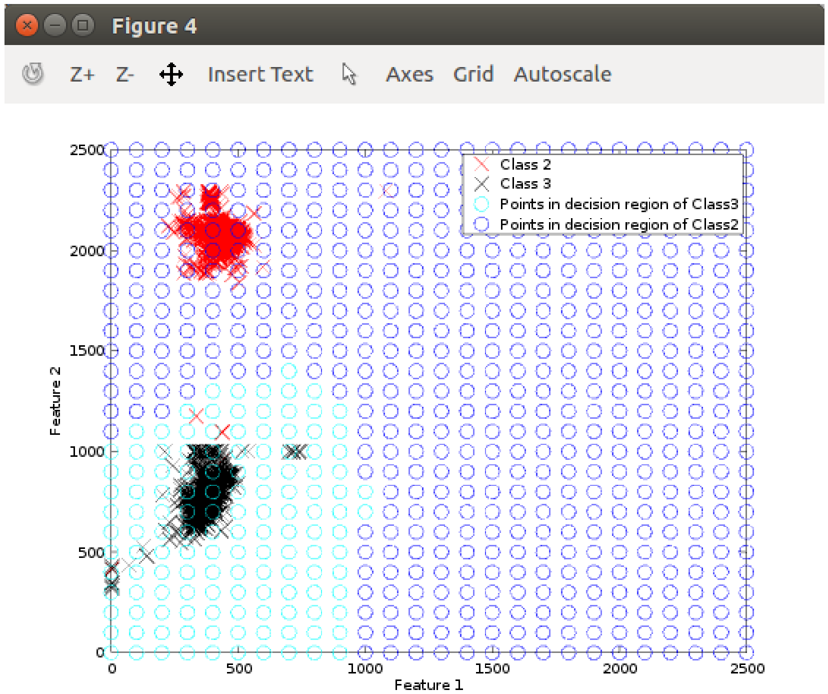


Fig2 : This is the decision plot between class 2 and class 3 with their training data superposed.

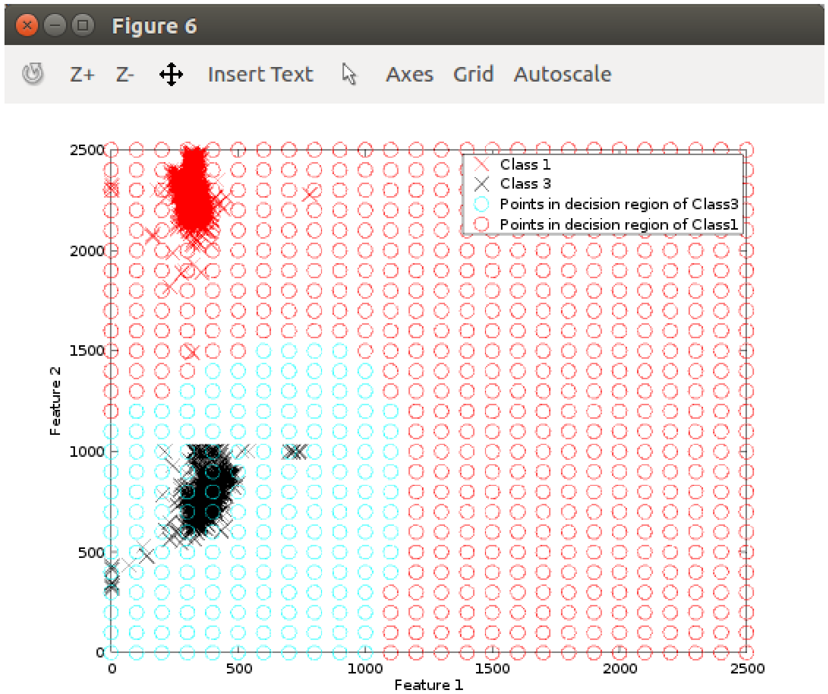
(c) 

Fig3 : This is the decision plot between class 3 and class 1 with their training data superposed.

2. Decision region plot for all the classes together with the training data superposed

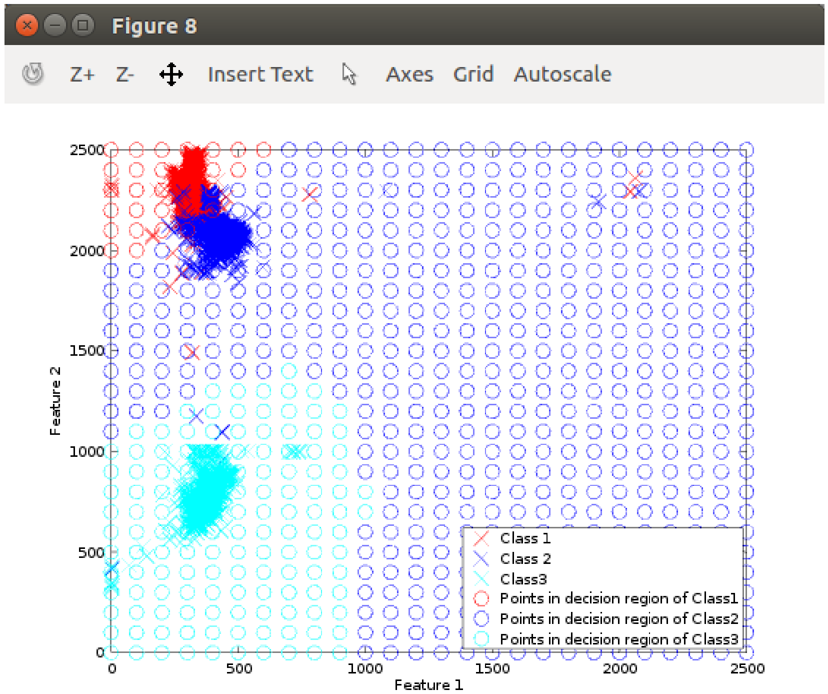
(a) 

Fig4 : This is the decision plot among class 1, class 2 and class 3 with their training data superposed.

3. Constant density contour plot for all the classes together with the training data superposed

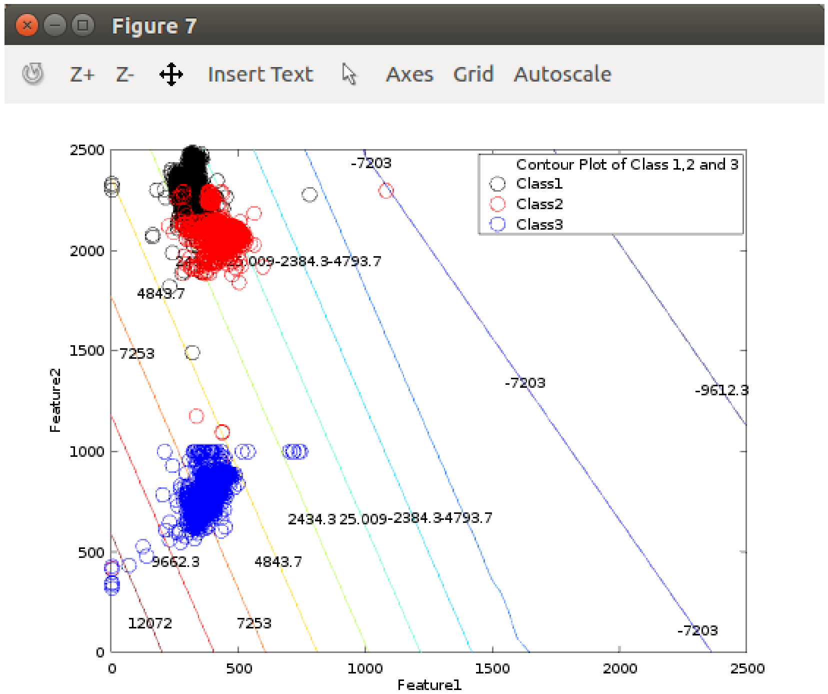
(a) 

Fig5 : This is the constant density contour plot for all the classes together with the training data superposed

4. Confusion matrix

$$M = \begin{bmatrix} 571 & 23 & 3 \\ 347 & 211 & 15 \\ 0 & 1 & 621 \end{bmatrix}$$

In this matrix row(i) corresponds to the number of test data points categorized as class i.
Column(j) corresponds to the number of test data points actually in class j.

5. Calculation of performance parameters

   (a) Accuracy

$$Accuracy \quad = \quad 78.292\%$$

   (b) Precision

i. Precision for Class 1

$$PC_1 = 62.200\%$$

ii. Precision for Class 2

$$PC_2 = 89.787\%$$

iii. Precision for Class 3

$$PC_3 = 97.183\%$$

iv. Mean Precision

$$MPC = 83.057\%$$

(c) Recall

i. Recall for Class 1

$$RC_1 = 95.645\%$$

ii. Recall for Class 2

$$RC_2 = 36.824\%$$

iii. Recall for Class 3

$$RC_3 = 99.839$$

iv. Mean Recall

$$MRC = 77.436\%$$

(d) F-Measure

i. F-Measure for Class 1

$$FM_1 = 75.380$$

ii. F-Measure for Class 2

$$FM_2 = 52.228$$

iii. F-Measure for Class 3

$$FM_3 = 98.493$$

iv. Mean F-Measure

$$MFM = 75.367$$

Here again the accuracy, mean recall, mean precision all are around 80% but some noticable observations prevail : 1. Precision for class 1 is not very stisfactory. 2. Recall for class 2 is poor. 3. Recall for class 3 is 100%.

Overall we conclude that :

1. For linearly separated dataset we should choose the case I assumptions as it provides similar performance to other cases but is less costly to calculate.

2. For non-linearly separated data any of the cases 2, 3, or 4 may be choosen. Although 3 and 4 provide better performance.

3. For real world data there's no favourite method as all perform similarly.