2015

CS669: Pattern Recognition

GROUP-13

IIT Mandi

13-Sep-15

Amit Kumar – B13107

Arpit Krishna – B13110

Paawan Mukker – B13218

INTRODUCTION –

In pattern recognition, machine learning and statistics, classification is the problem of identifying to which of a set of categories (sub-populations) a new observation belongs, on the basis of a training set of data containing observations (or instances) whose category membership is known. An example would be assigning a given email into "spam" or "non-spam" classes or assigning a diagnosis to a given patient as described by observed characteristics of the patient (gender, blood pressure, presence or absence of certain symptoms, etc.).

Problem Statement:

Classification of several types of data using two different types of classifiers and observe the results.

Data types provided –

- Linearly separable data.
- Non linearly separable data
  - Interlocking
  - Spiral
  - Ring
- Overlapping data
- Real World Data

- Data contains 2/3 classes and has only two attributes.
- Classifiers to implement –
  - Bayes Classifier
    - Same covariance matrix
    - Different covariance matrix
  - Naïve Bayes Classifier
    - Covariance matrix of the form: $\sigma^2 I$
    - Same covariance matrix
    - Different covariance matrix

Learning Objective:

- Observe the decision boundaries for different datasets under different classifiers and explain the reasons for them.
- Observe performance accuracy of different classifiers for different types of data sets.

CLASSIFIERS –

Bayesian decision theory is a fundamental statistical approach to the problem of pattern classification. This approach is based on quantifying the tradeoffs between various classification decisions using probability and the costs that accompany such decisions. It makes the assumption that the decision problem is posed in probabilistic terms, and that all of the relevant probability values are known.

Using Bayes' theorem, the conditional probability can be decomposed as

$$p(C_k|\mathbf{x}) = \frac{p(C_k)\ p(\mathbf{x}|C_k)}{p(\mathbf{x})}.$$

In plain English, using Bayesian probability terminology, the above equation can be written as

$$\text{posterior} = \frac{\text{prior} \times \text{likelihood}}{\text{evidence}}.$$

In practice, there is interest only in the numerator of that fraction, because the denominator does not depend on C and the values of the features F_i are given, so that the denominator is effectively constant. The numerator is equivalent to the joint probability model

$$p(C_k, x_1, \dots, x_n)$$

Now the "naive" conditional independence assumptions come into play: assume that each feature F_i is conditionally independent of every other feature F_j for $j \neq i$, given the category C. Thus, the joint model can be expressed as

$$\begin{aligned} p(C_k|x_1, \dots, x_n) &\propto p(C_k, x_1, \dots, x_n) \\ &\propto p(C_k)\ p(x_1|C_k)\ p(x_2|C_k)\ p(x_3|C_k)\ \cdots \\ &\propto p(C_k) \prod_{i=1}^{n} p(x_i|C_k)\,. \end{aligned}$$

This means that under the above independence assumptions, the conditional distribution over the class variable C is:

$$p(C_k|x_1,\dots,x_n) = \frac{1}{Z}p(C_k)\prod_{i=1}^{n} p(x_i|C_k)$$

where the evidence Z = p(x) is a scaling factor dependent only on x_1,.,.,…, x_n, that is, a constant if the values of the feature variables are known.

In machine learning, naive Bayes classifiers are a family of simple probabilistic classifiers based on applying Bayes' theorem with strong (naive) independence assumptions between the features.

Naive Bayes is a simple technique for constructing classifiers: models that assign class labels to problem instances, represented as vectors of feature values, where the class labels are drawn from some finite set. It is not a single algorithm for training such classifiers, but a family of algorithms based on a common principle: all naive Bayes classifiers assume that the value of a particular feature is independent of the value of any other feature, given the class variable. For example, a fruit may be considered to be an apple if it is red, round, and about 10 cm in diameter. A naive Bayes classifier considers each of these features to contribute independently to the probability that this fruit is an apple, regardless of any possible correlations between the color, roundness and diameter features.

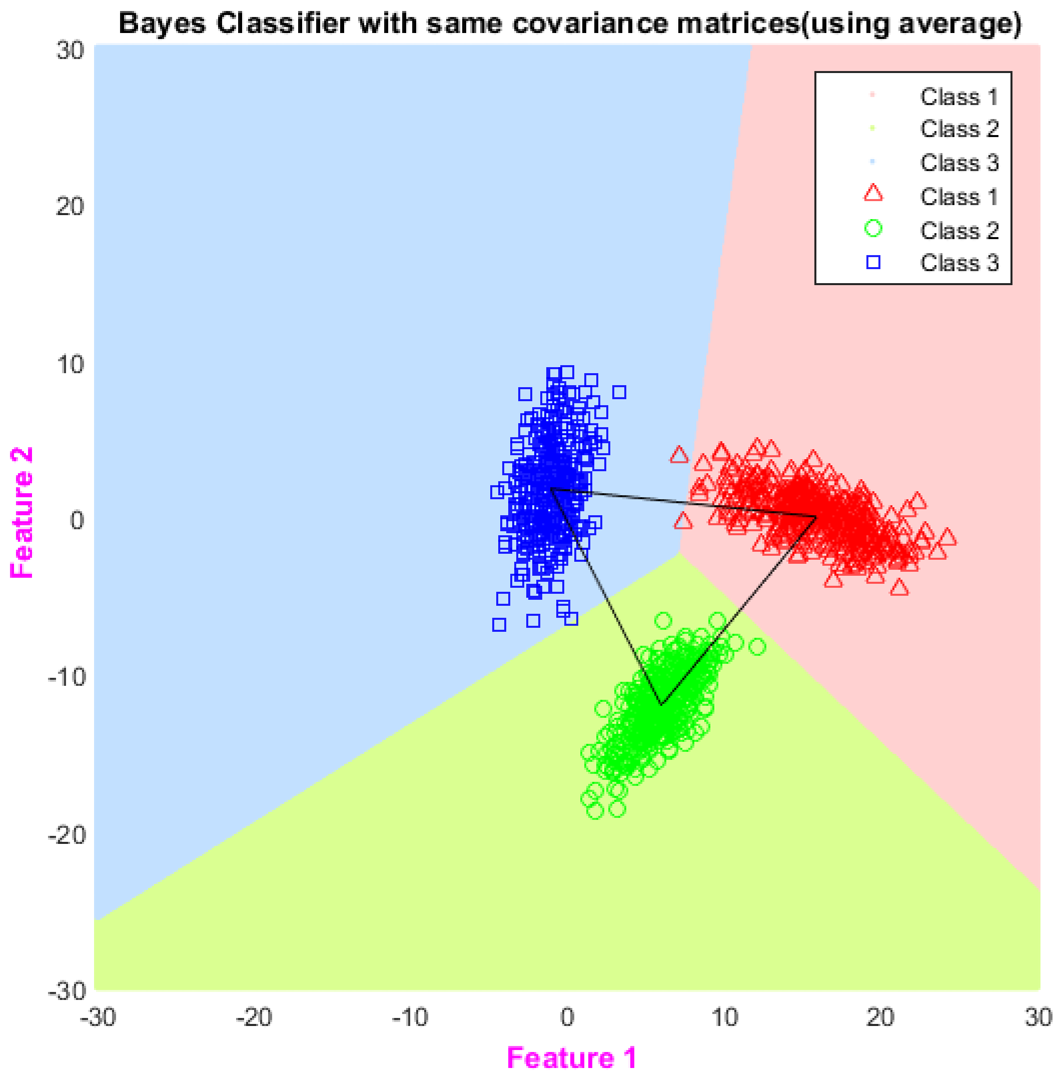
Naive Bayes classifiers are highly scalable, requiring a number of parameters linear in the number of variables (features/predictors) in a learning problem. Maximum-likelihood training can be done by evaluating a closed-form expression, which takes linear time, rather than by expensive iterative approximation as used for many other types of classifiers.

♦ **Dataset I-(a)**: 2-dimensional artificial data of 3 classes which are **linearly separable.**

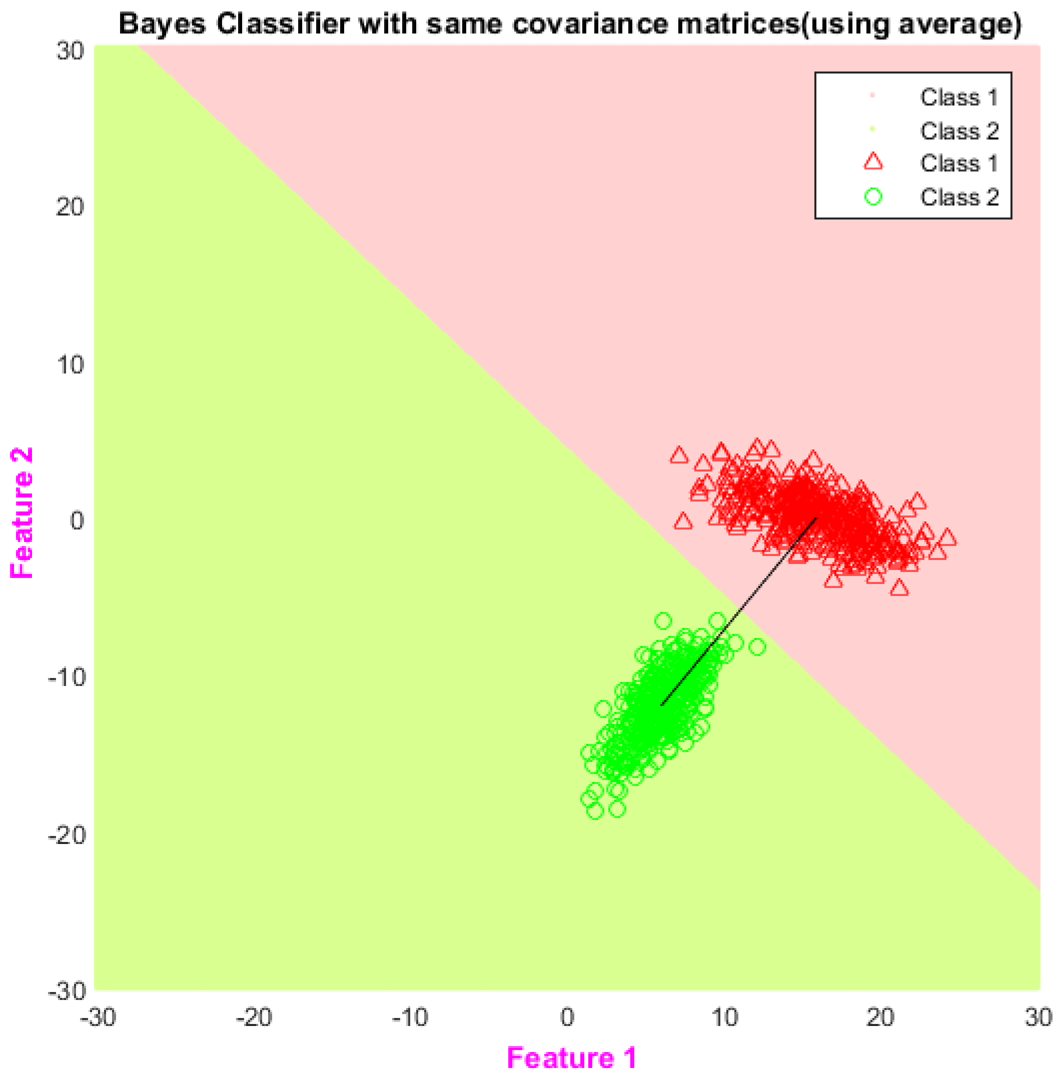
1) Bayes Classifier –

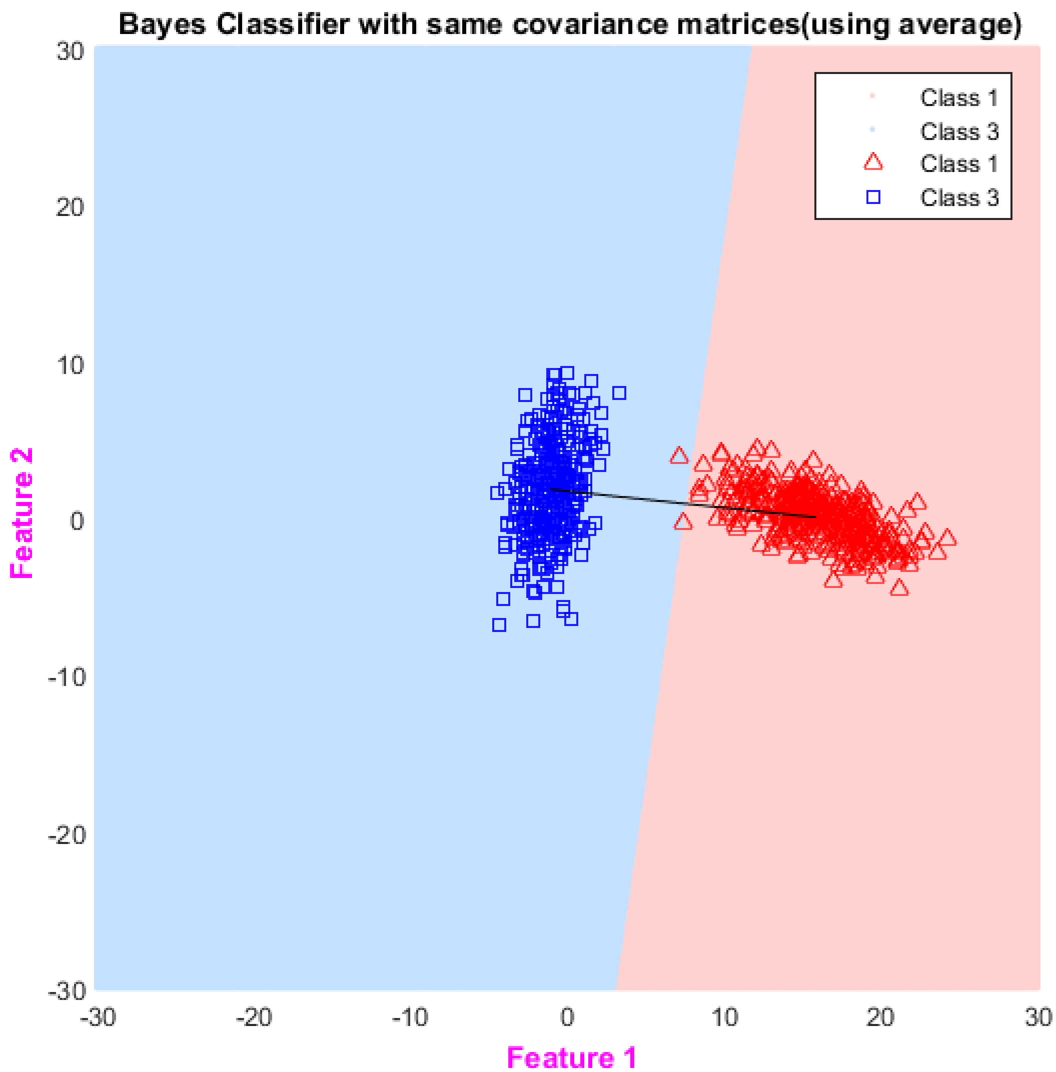
a) Same covariance matrices (by taking average)

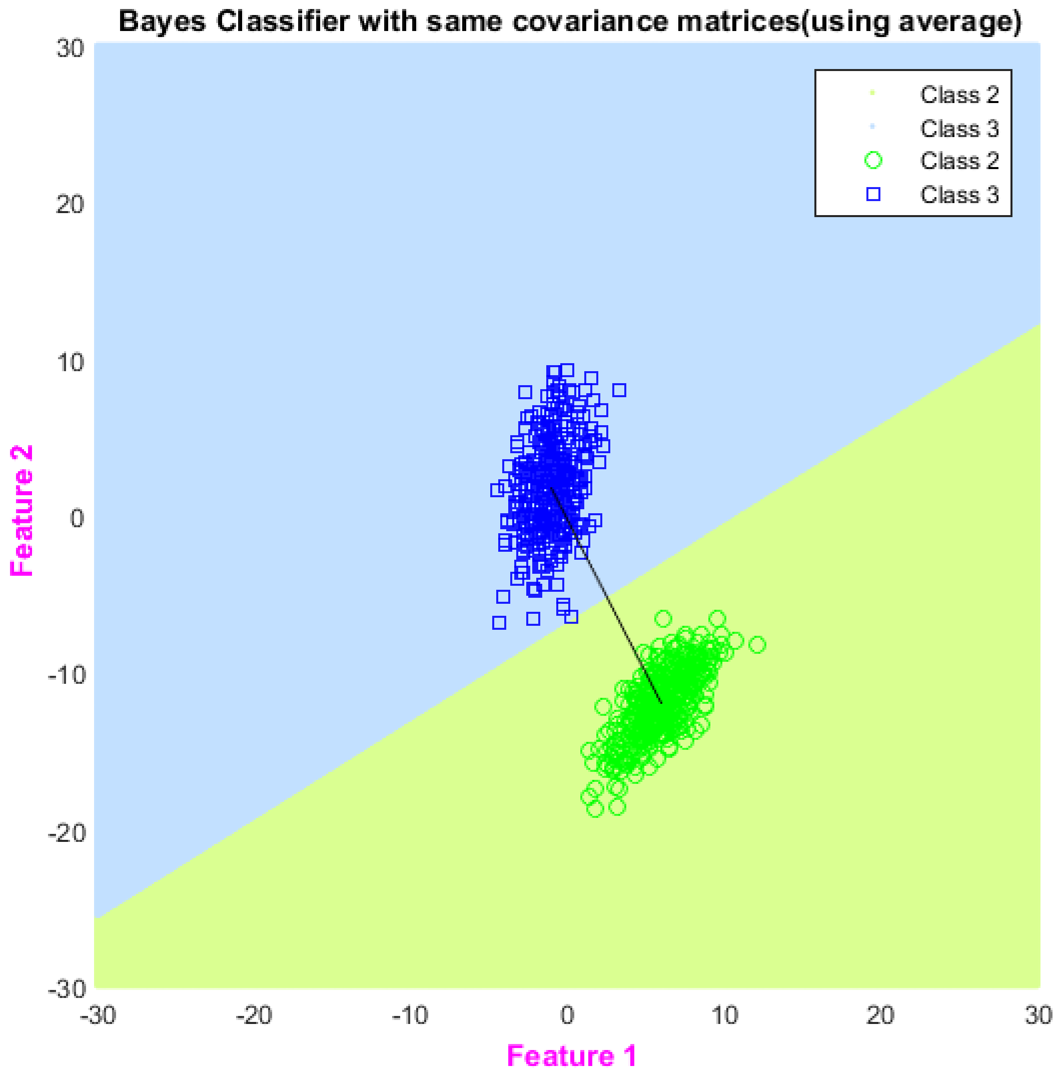
- *Decision region plot for all the classes together with the training data superposed –*



- *Decision region plot for every pair of classes -*




Bayes Classifier with same covariance matrices(using average)
Class 1
Class 3
Class 1
Class 3
Feature 2
Feature 1
30
20
10
0
-10
-20
-30
-30
-20
-10
0
10
20
30



- *Confusion Matrix based on performance for test data-*

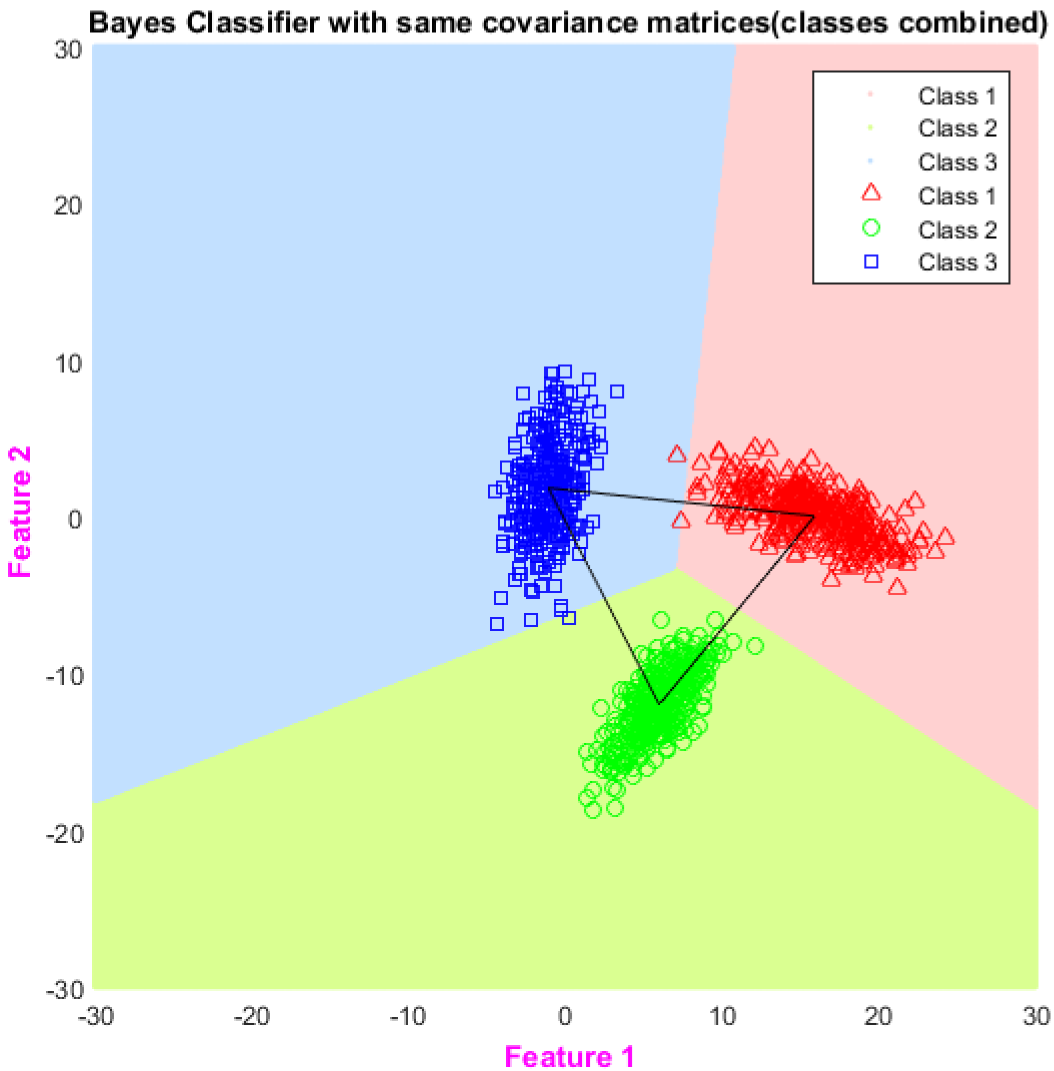
Predicted Class ⇨	CLASS 1	CLASS 2	CLASS 3
Actual Class ⇩			
Class 1	124	0	1
Class 2	0	125	0
Class 3	0	0	125

- *Classification accuracy on test data –*

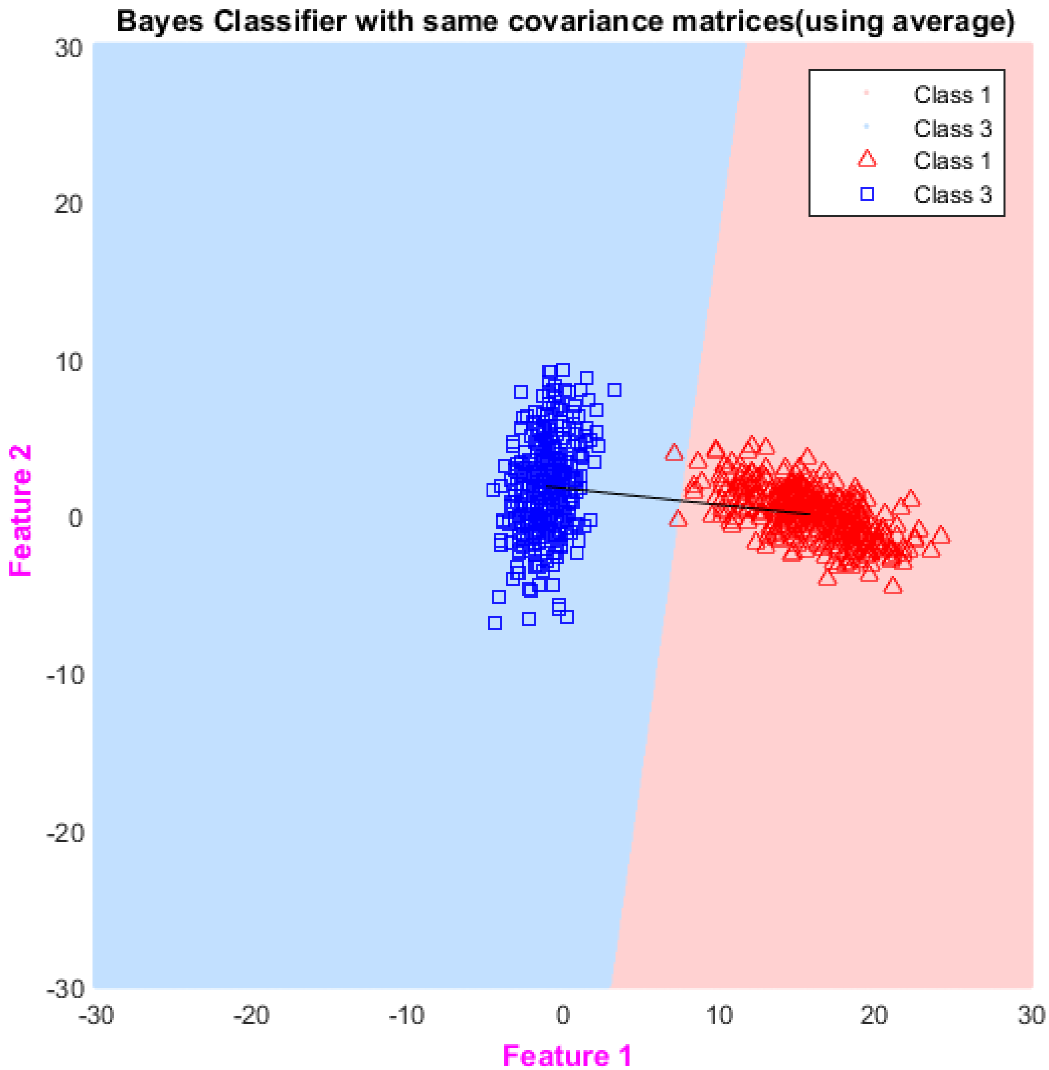
Overall Accuracy – 99.7333
Classifier Accuracy for class 1 – 99.2
Classifier Accuracy for class 2 – 100.0
Classifier Accuracy for class 3 – 100.0

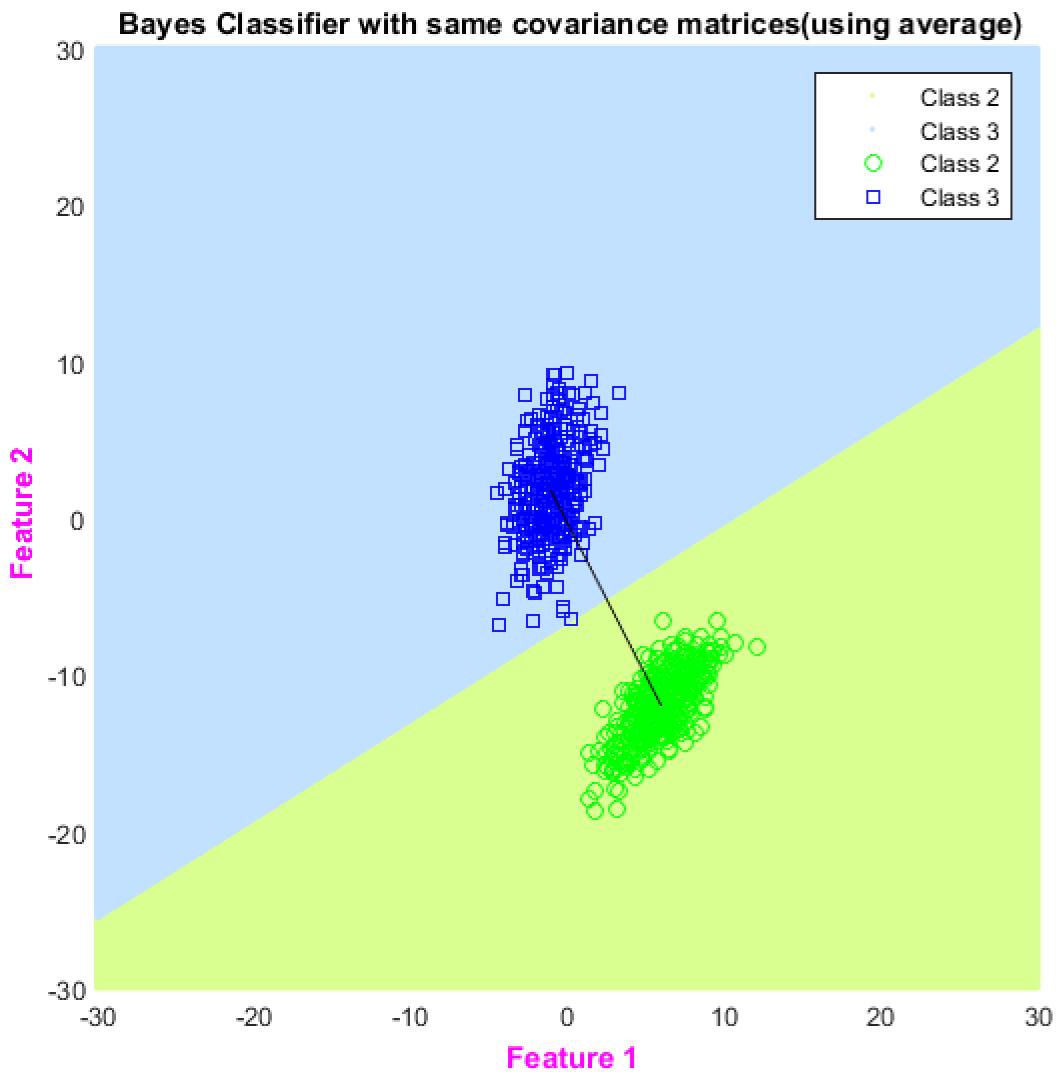
b) Same covariance matrices (from training data of all classes combined)

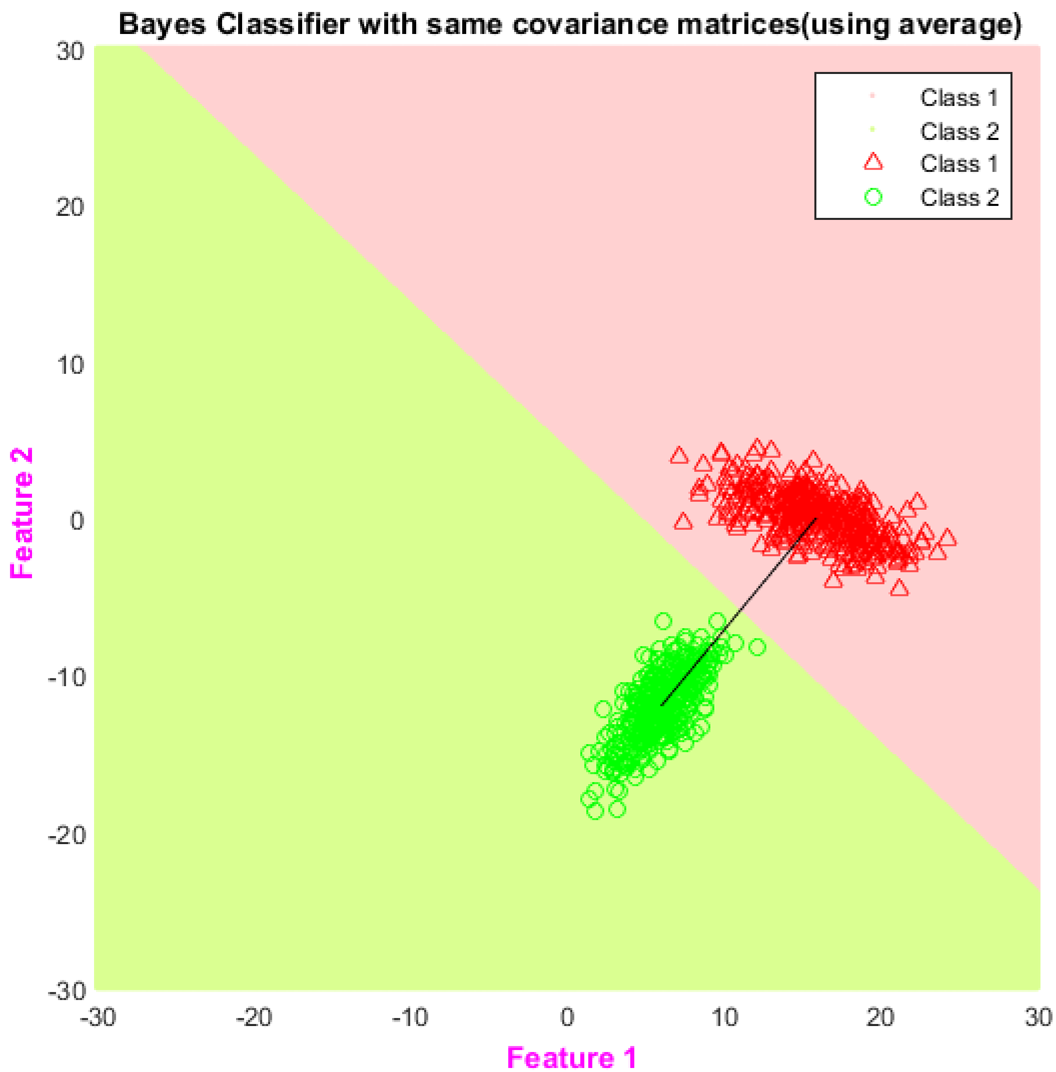
- *<u>Decision region plot for all the classes together with the training data superposed -</u>*



- *Decision region plot for every pair of classes –*




Bayes Classifier with same covariance matrices(using average)
Class 2
Class 3
Class 2
Class 3
Feature 2
Feature 1
30
20
10
0
-10
-20
-30
-30
-20
-10
0
10
20
30



- *<u>Confusion Matrix based on performance for test data-</u>*

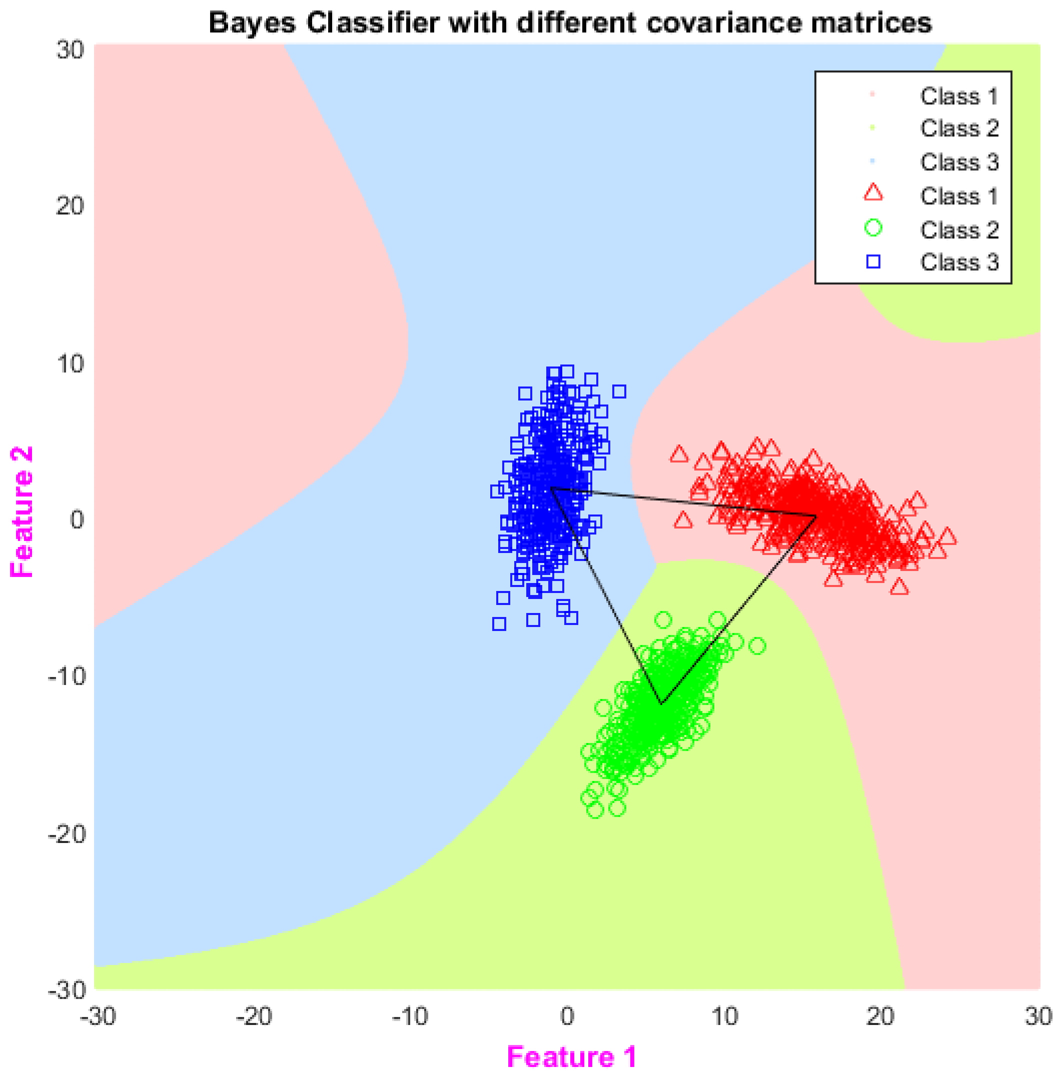
Predicted Class ⇨	CLASS 1	CLASS 2	CLASS 3
Actual Class ⇩			
Class 1	124	0	1
Class 2	0	125	0
Class 3	0	0	125

- *<u>Classification accuracy on test data –</u>*

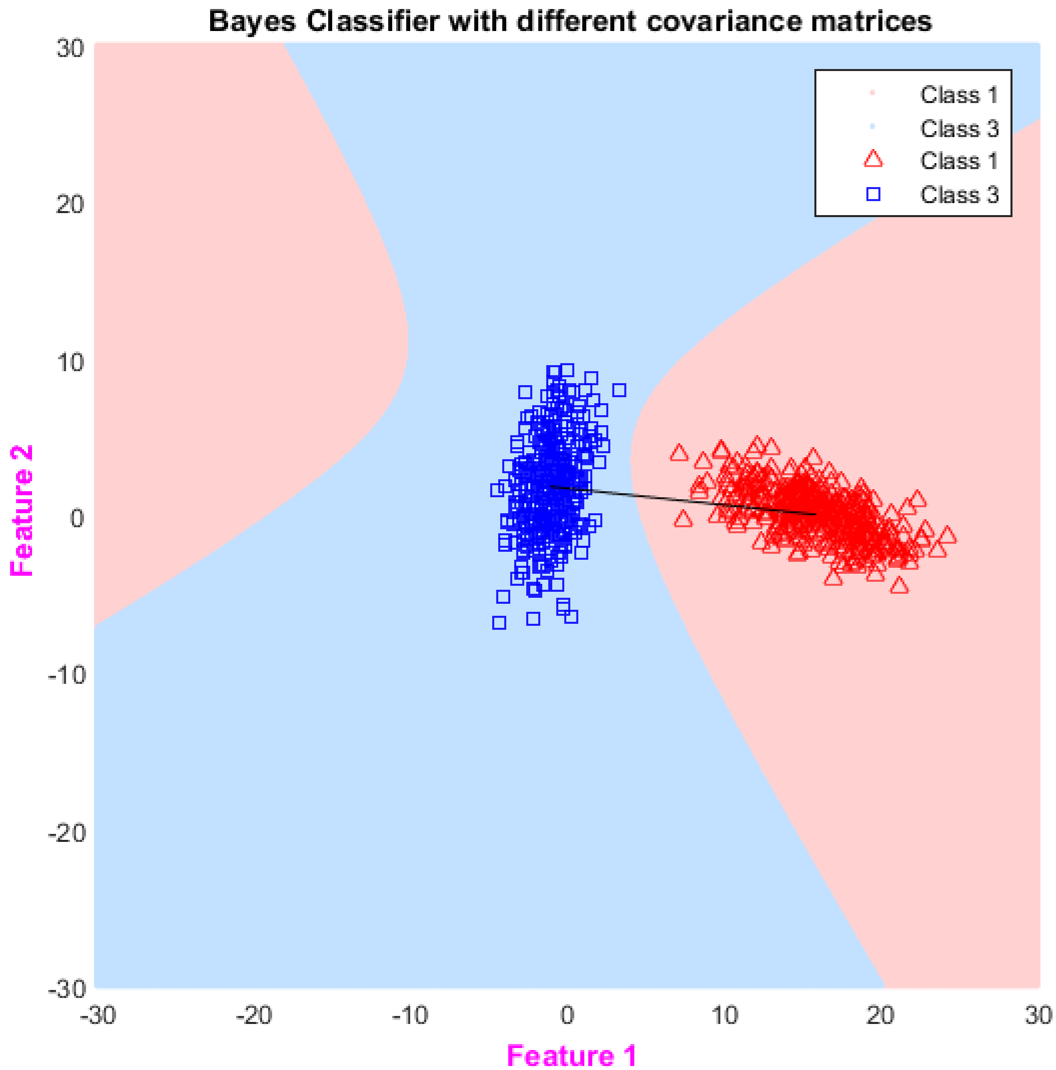
- Overall Accuracy – 99.7333
- Classifier Accuracy for class 1 – 99.2
- Classifier Accuracy for class 2 – 100.0
- Classifier Accuracy for class 3 – 100.0

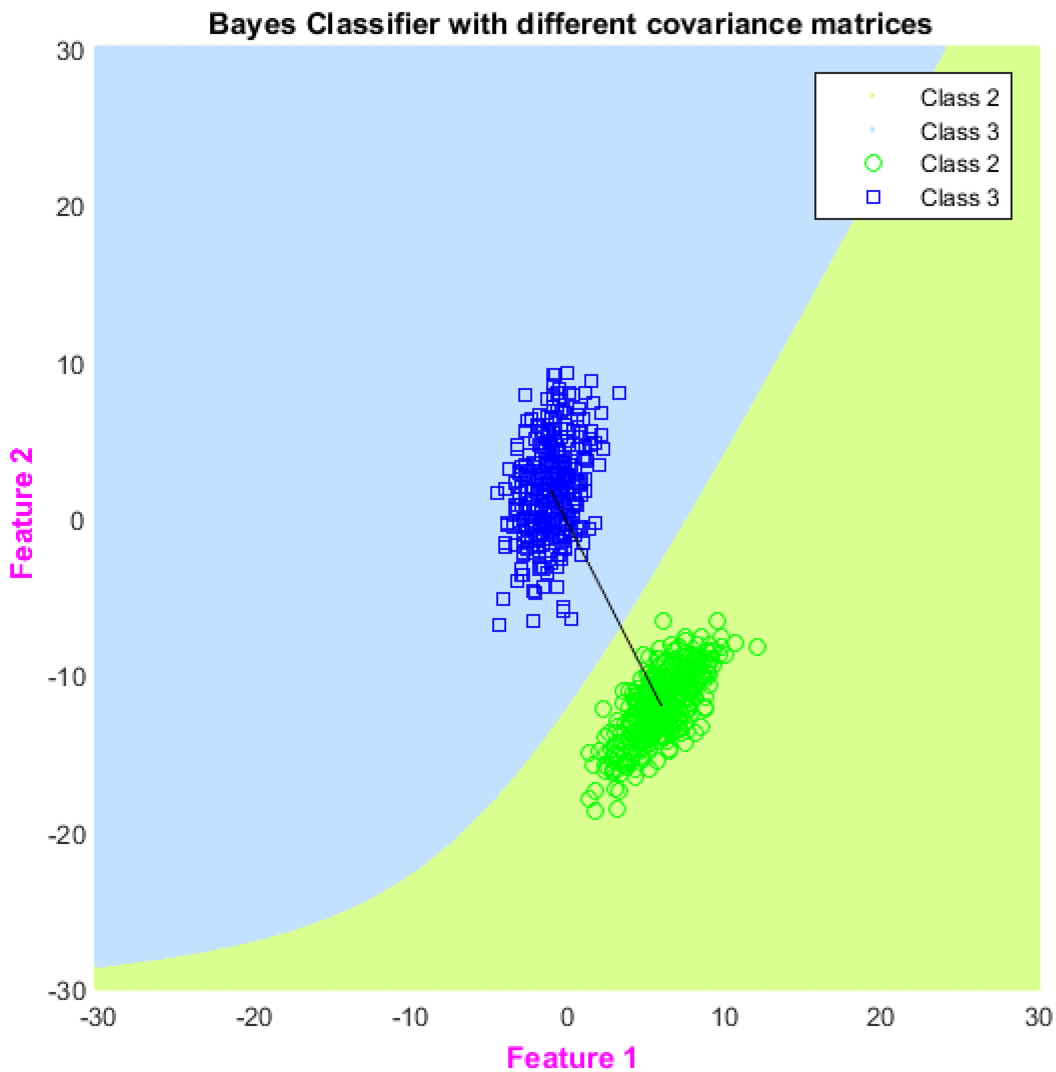
c) Covariance matrix for each class is different

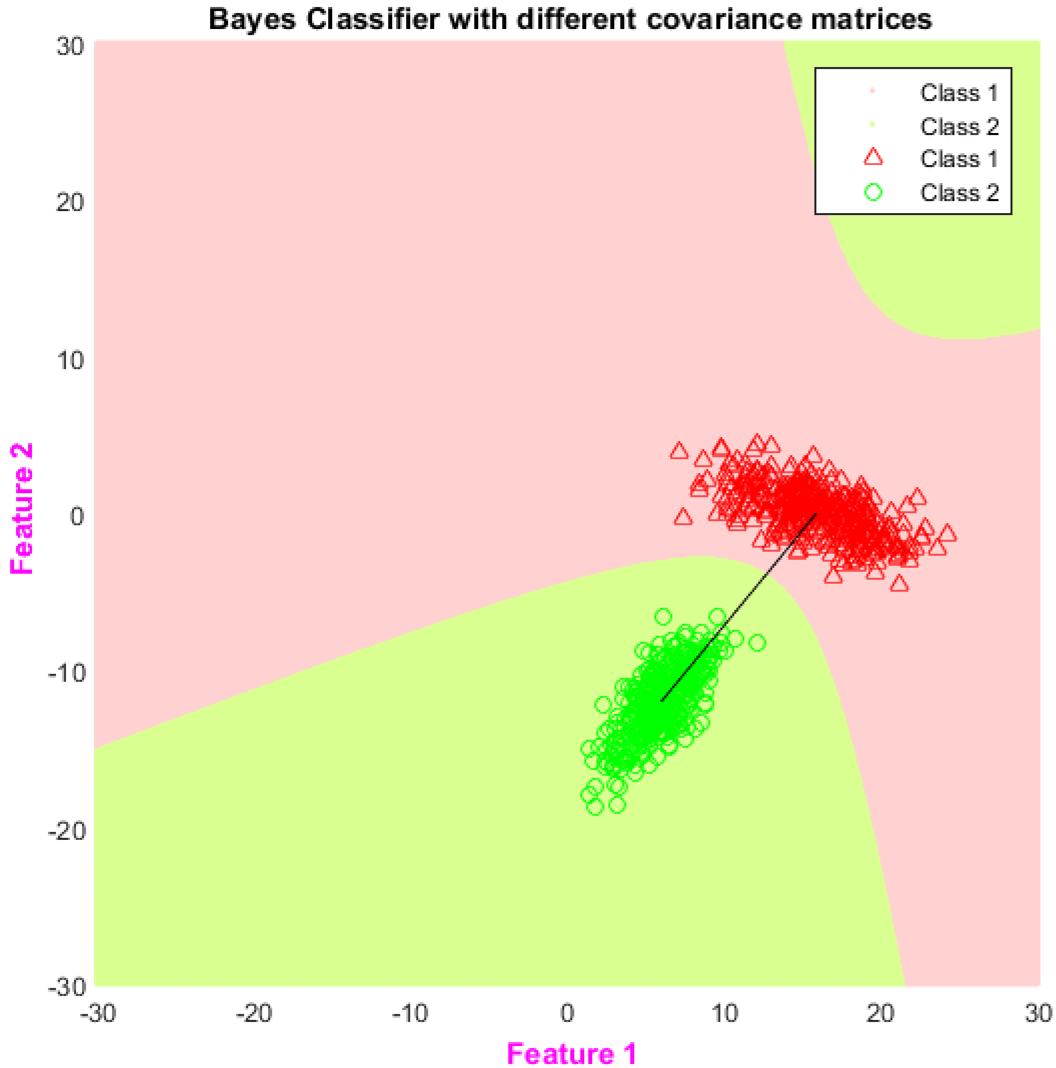
- *Decision region plot for all the classes together with the training data superposed -*



- *Decision region plot for every pair of classes -*




Bayes Classifier with different covariance matrices
Class 2
Class 3
Class 2
Class 3
Feature 2
Feature 1
30
20
10
0
-10
-20
-30
-30
-20
-10
0
10
20
30



- *Confusion Matrix based on performance for test data-*

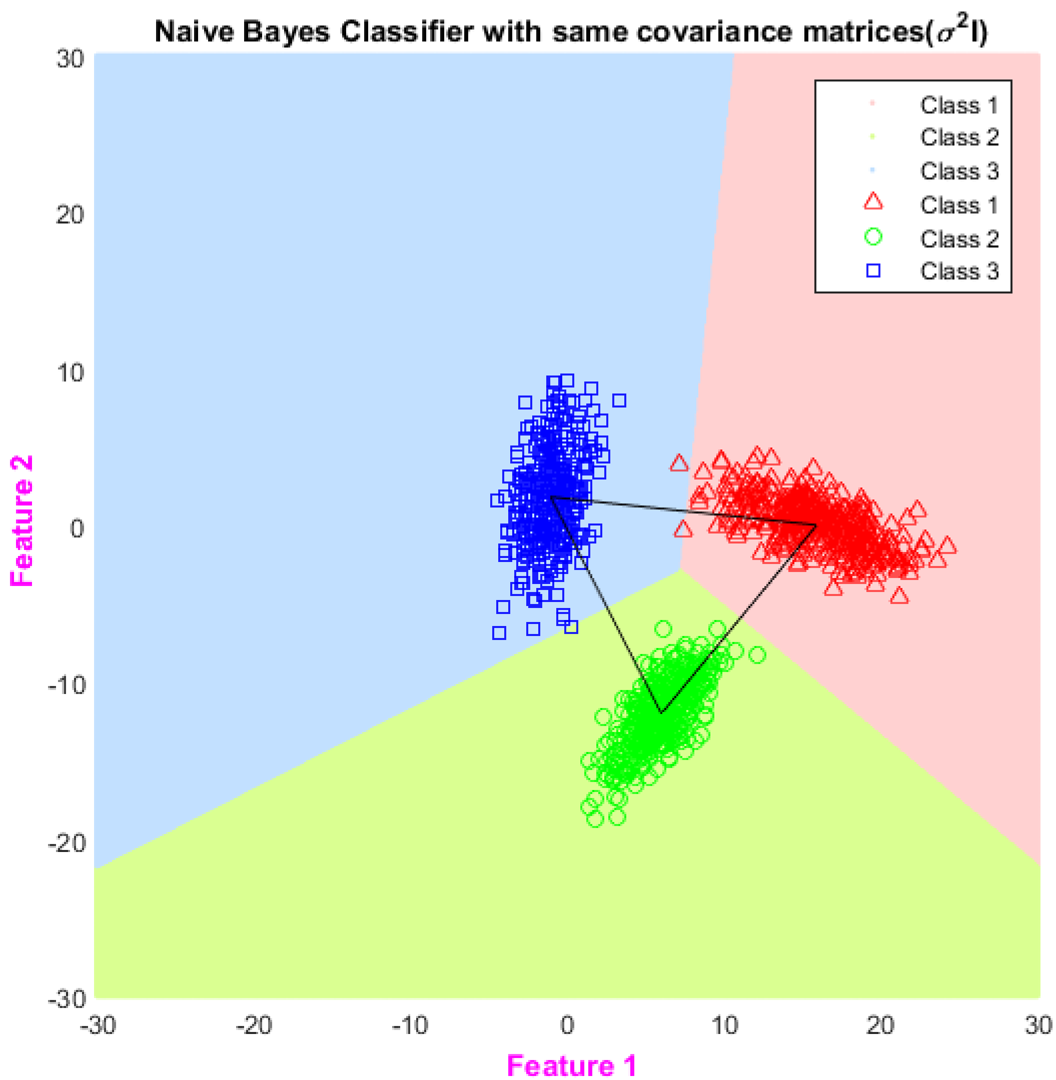
Predicted Class ⇨	CLASS 1	CLASS 2	CLASS 3
Actual Class ⇩			
Class 1	125	0	0
Class 2	0	125	0
Class 3	0	0	125

- *Classification accuracy on test data –*
  Overall Accuracy - 100.0
  Classifier Accuracy for class 1 - 100.0
  Classifier Accuracy for class 2 - 100.0
  Classifier Accuracy for class 3 - 100.0

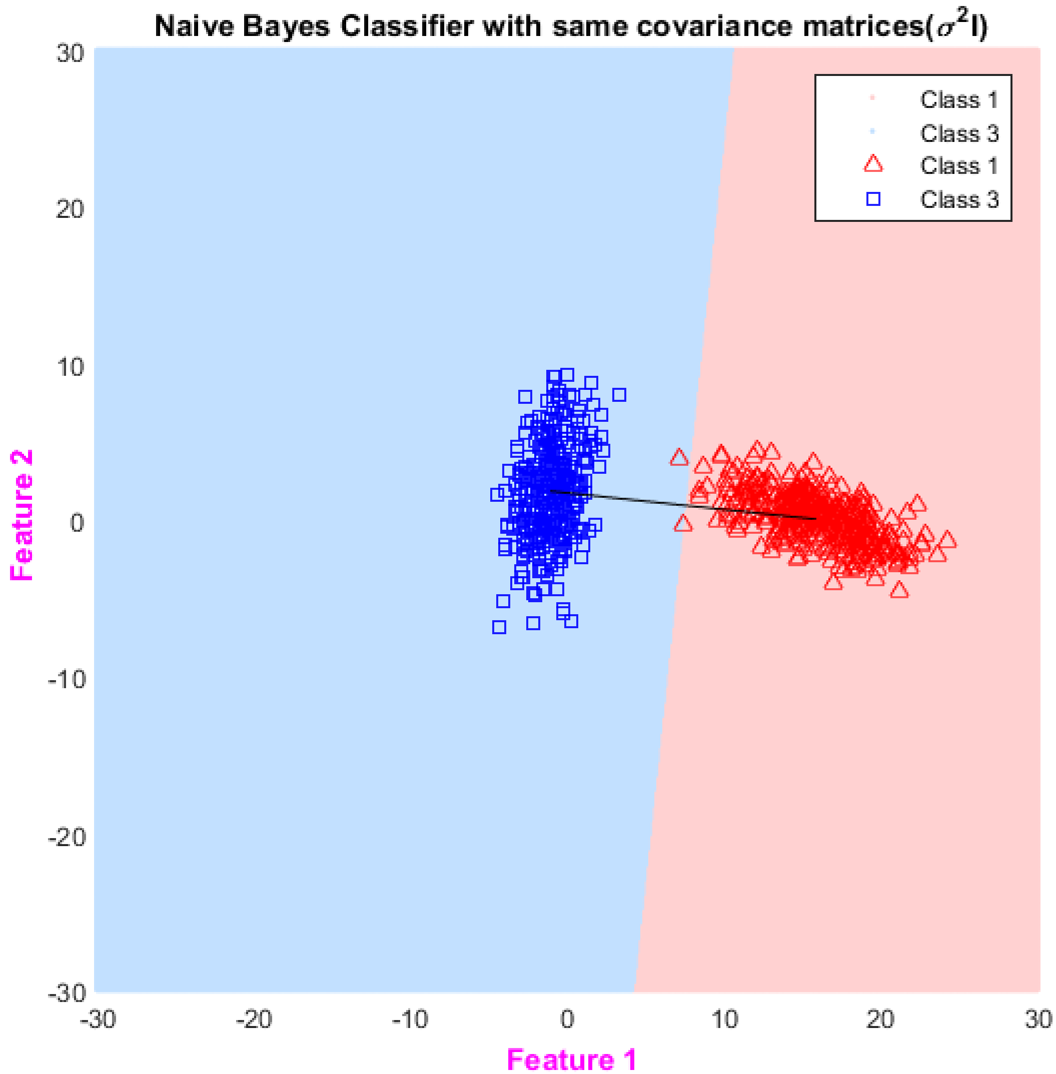
2) Naive Bayes Classifier –

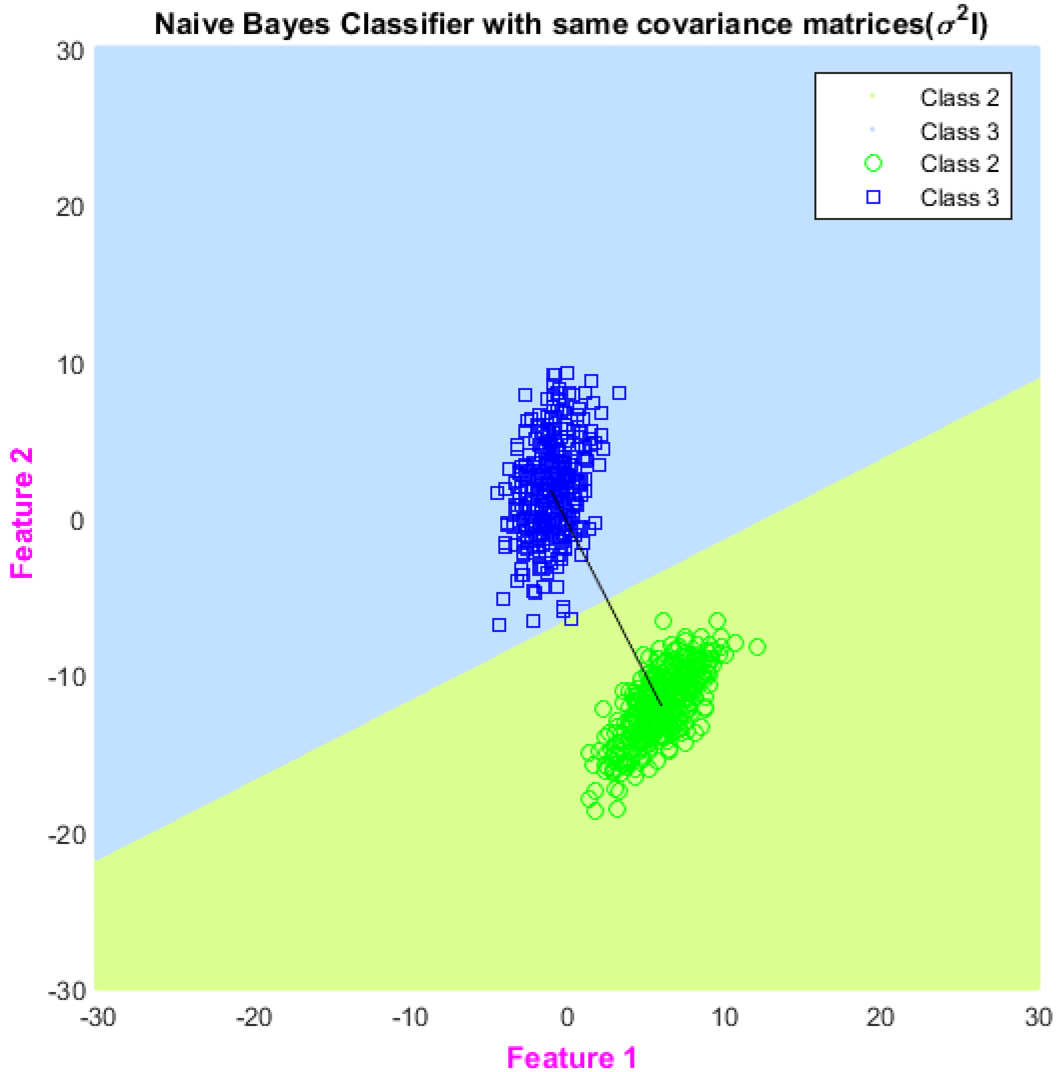
a) Same covariance matrices (σ^2 I)

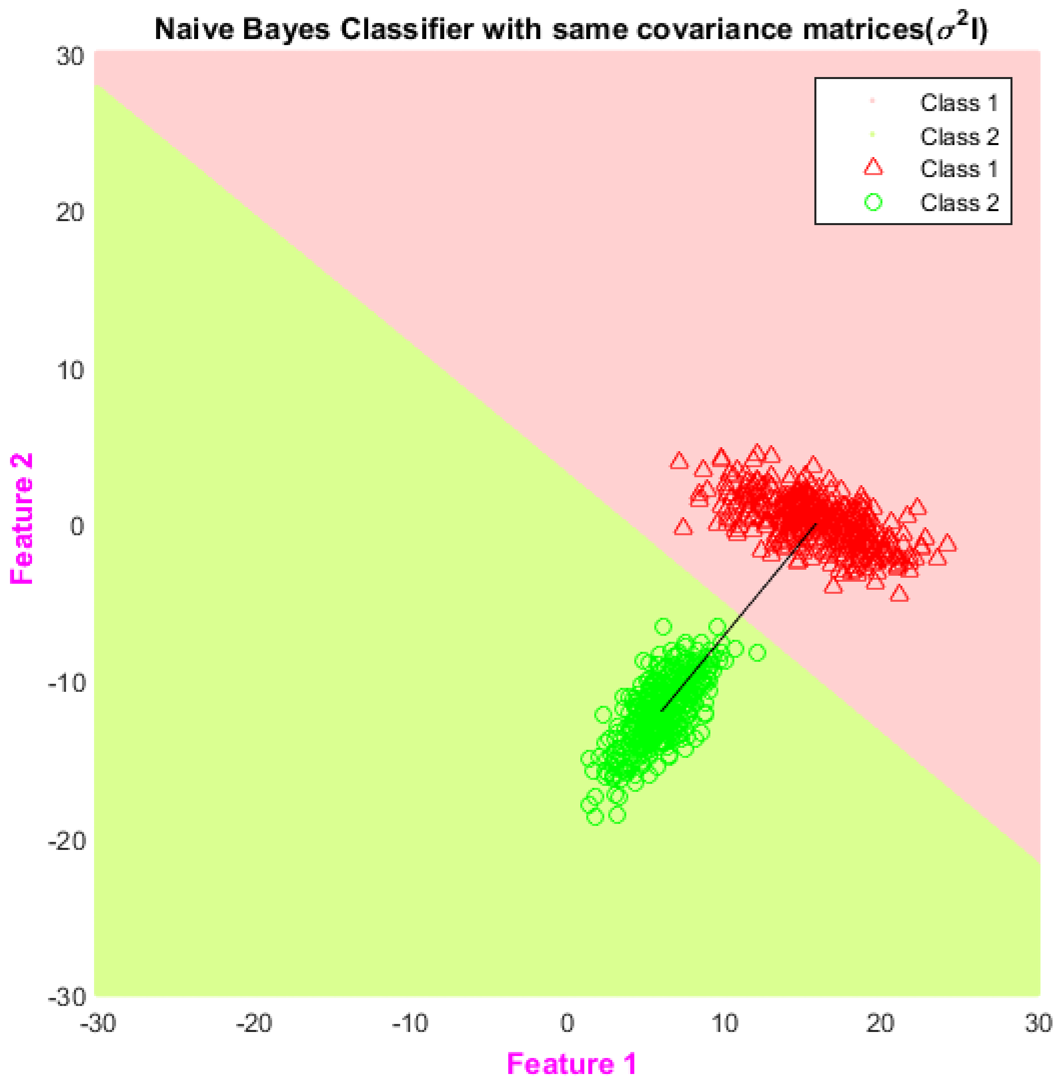
- *Decision region plot for all the classes together with the training data superposed -*



- *Decision region plot for every pair of classes –*




Naive Bayes Classifier with same covariance matrices(σ^2I)
Class 2
Class 3
Class 2
Class 3
Feature 2
Feature 1
30
20
10
0
-10
-20
-30
-30
-20
-10
0
10
20
30



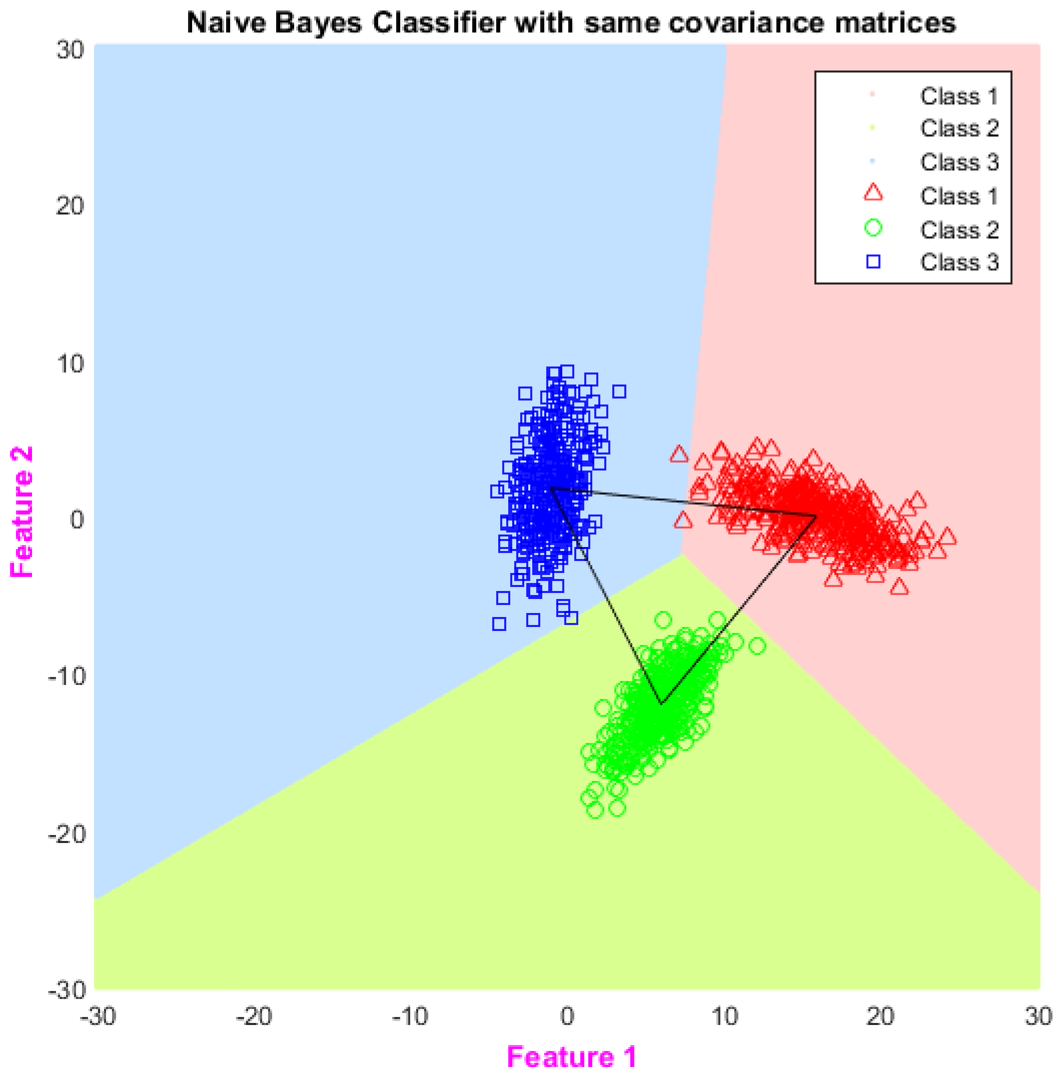
- *Confusion Matrix based on performance for test data-*

Predicted Class ⇨	CLASS 1	CLASS 2	CLASS 3
Actual Class ⇩			
Class 1	124	0	1
Class 2	0	125	0
Class 3	0	0	125

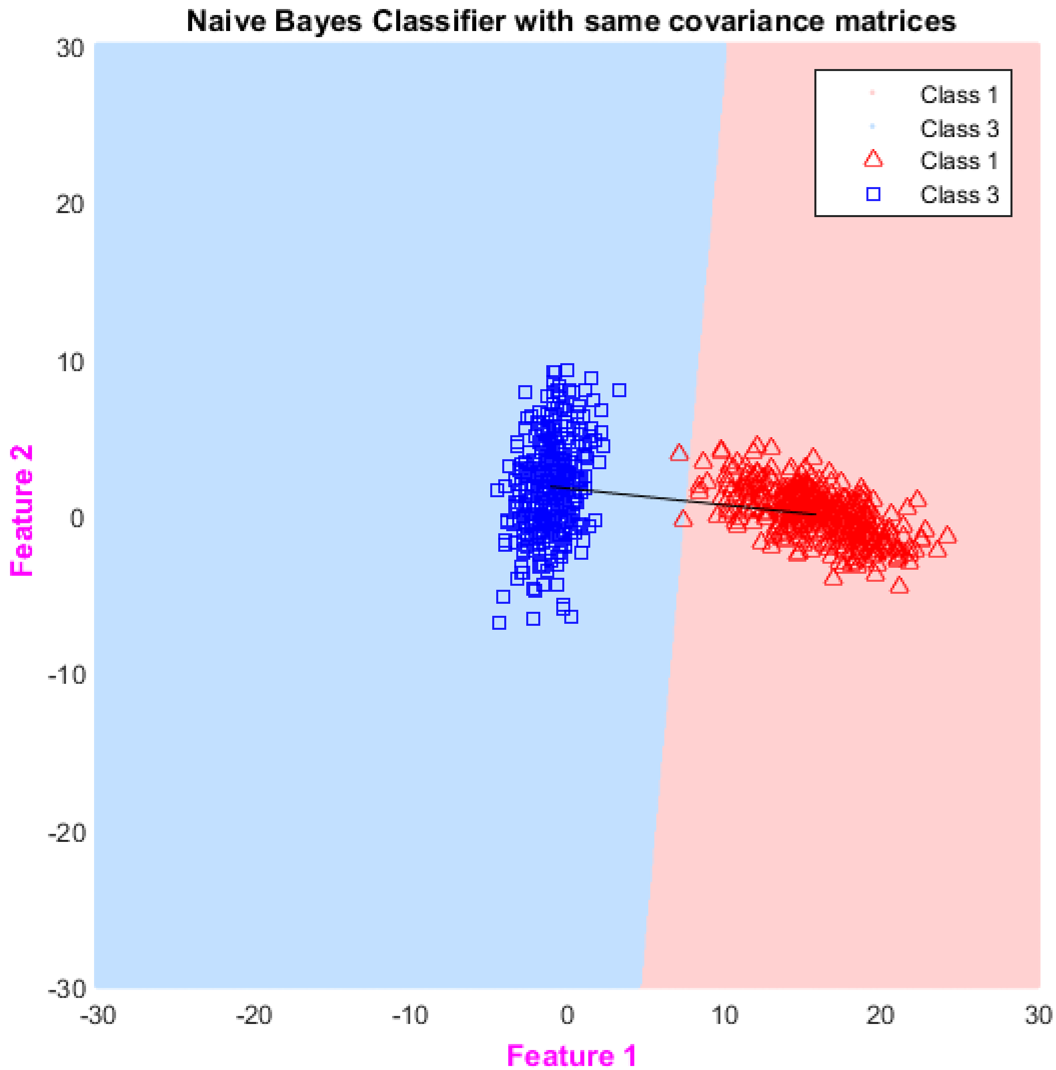
- *Classification accuracy on test data –*
  Overall Accuracy – 99.7333
  Classifier Accuracy for class 1 – 99.2
  Classifier Accuracy for class 2 – 100.0
  Classifier Accuracy for class 3 – 100.0

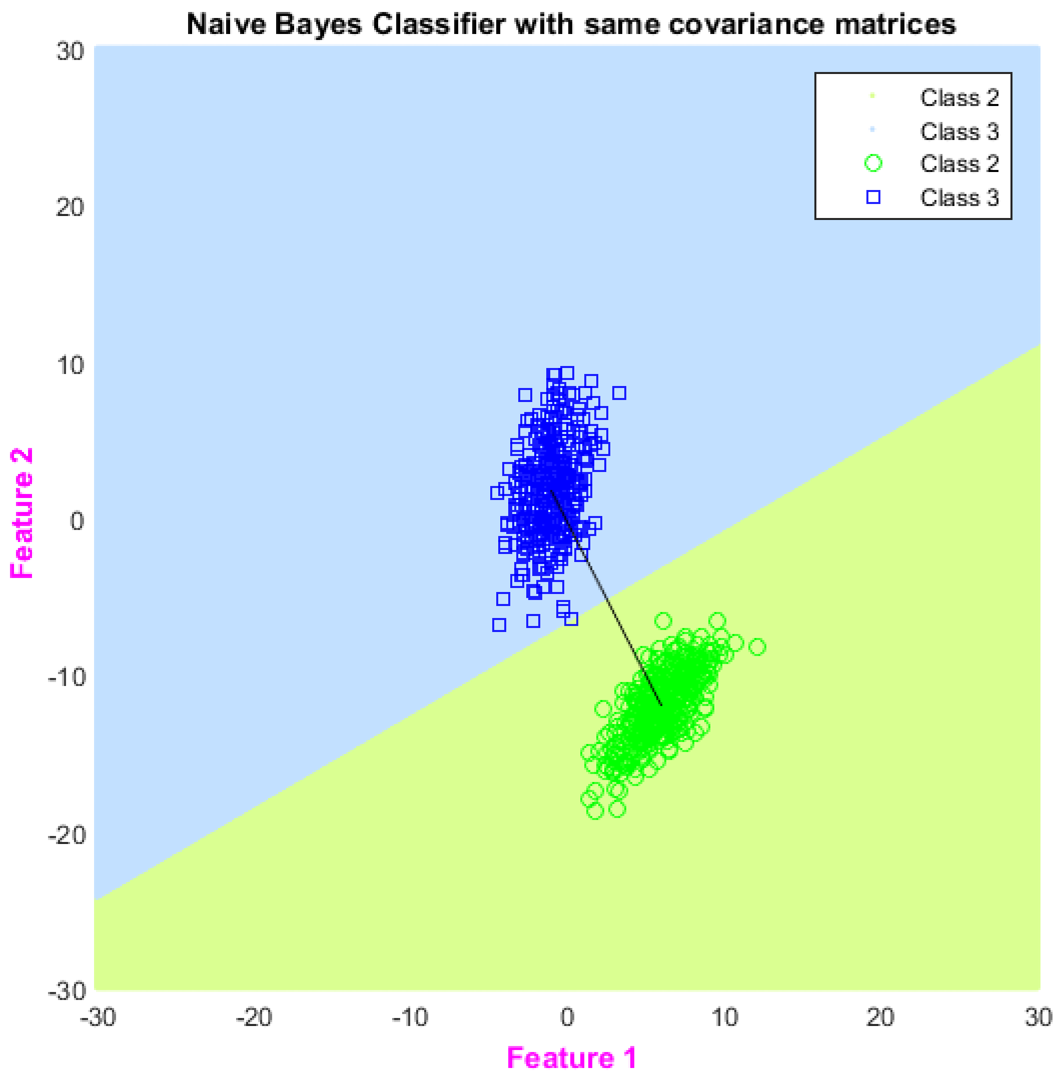
b) Same covariance matrices and is C

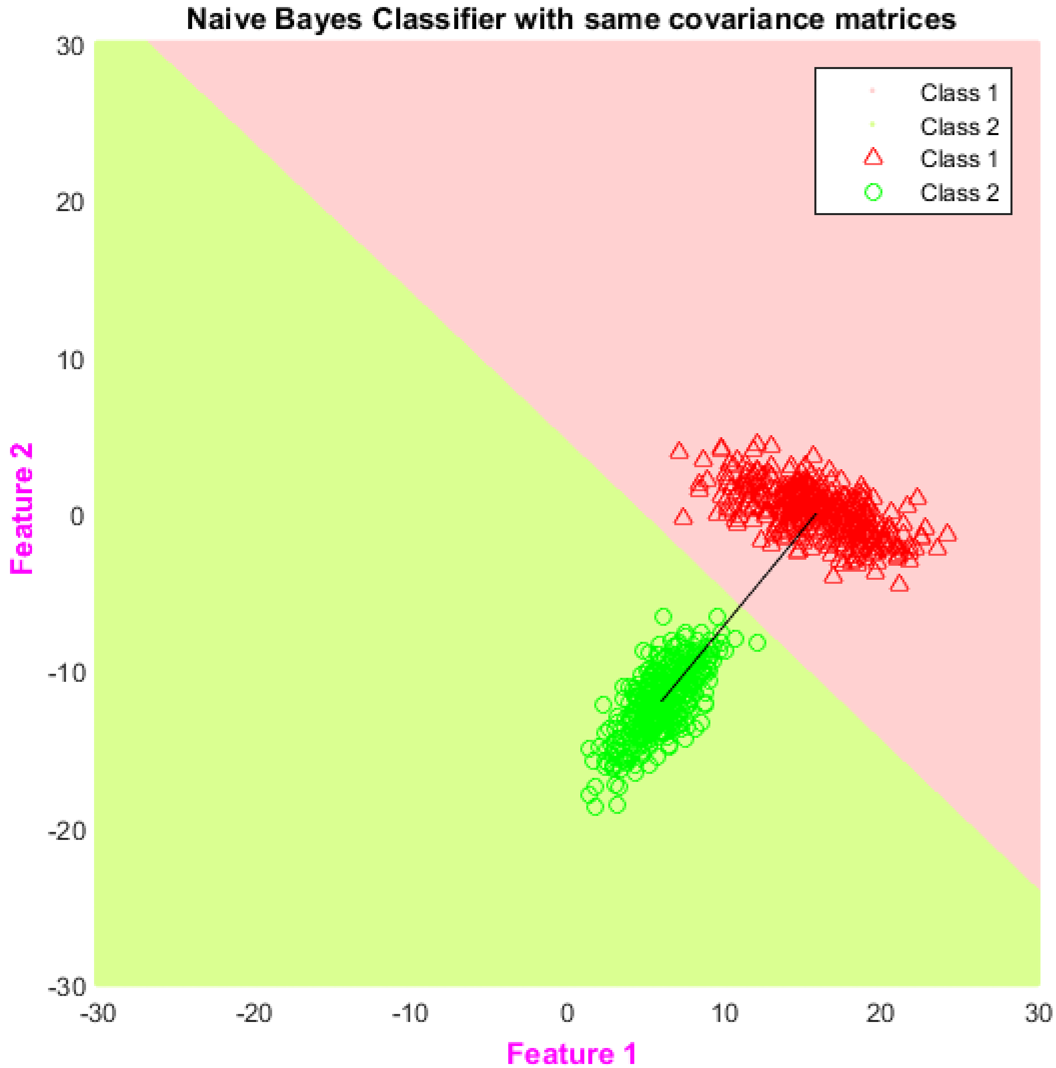
- *Decision region plot for all the classes together with the training data superposed -*



- *Decision region plot for every pair of classes -*




Naive Bayes Classifier with same covariance matrices
Class 2
Class 3
Class 2
Class 3
Feature 2
Feature 1
30
20
10
0
-10
-20
-30
-30
-20
-10
0
10
20
30



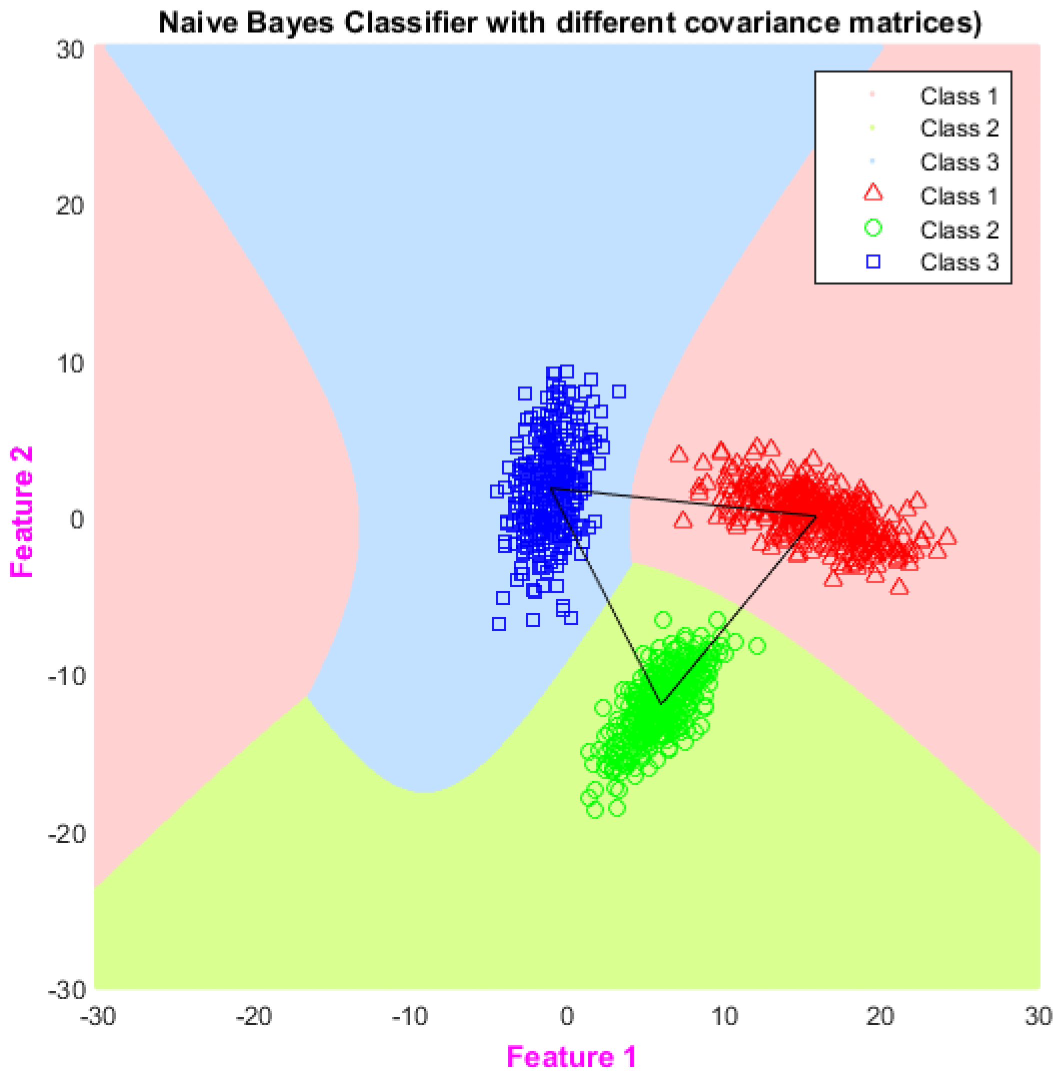
- *Confusion Matrix based on performance for test data-*

Predicted Class ⇨	CLASS 1	CLASS 2	CLASS 3
Actual Class ⇩			
Class 1	124	0	1
Class 2	0	125	0
Class 3	0	0	125

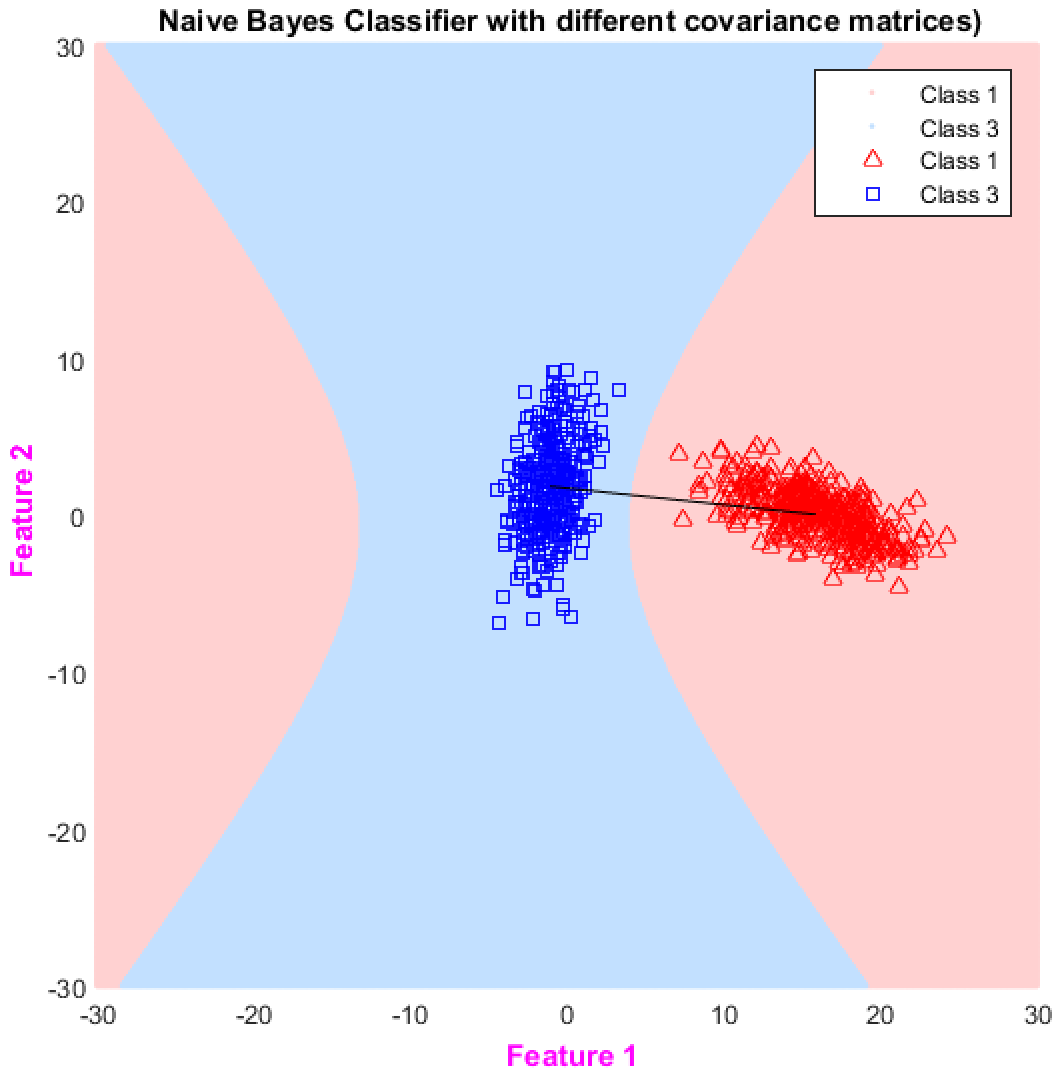
- *Classification accuracy on test data –*
  Overall Accuracy - 99.7333
  Classifier Accuracy for class 1 - 99.2
  Classifier Accuracy for class 2 - 100.0
  Classifier Accuracy for class 3 - 100.0

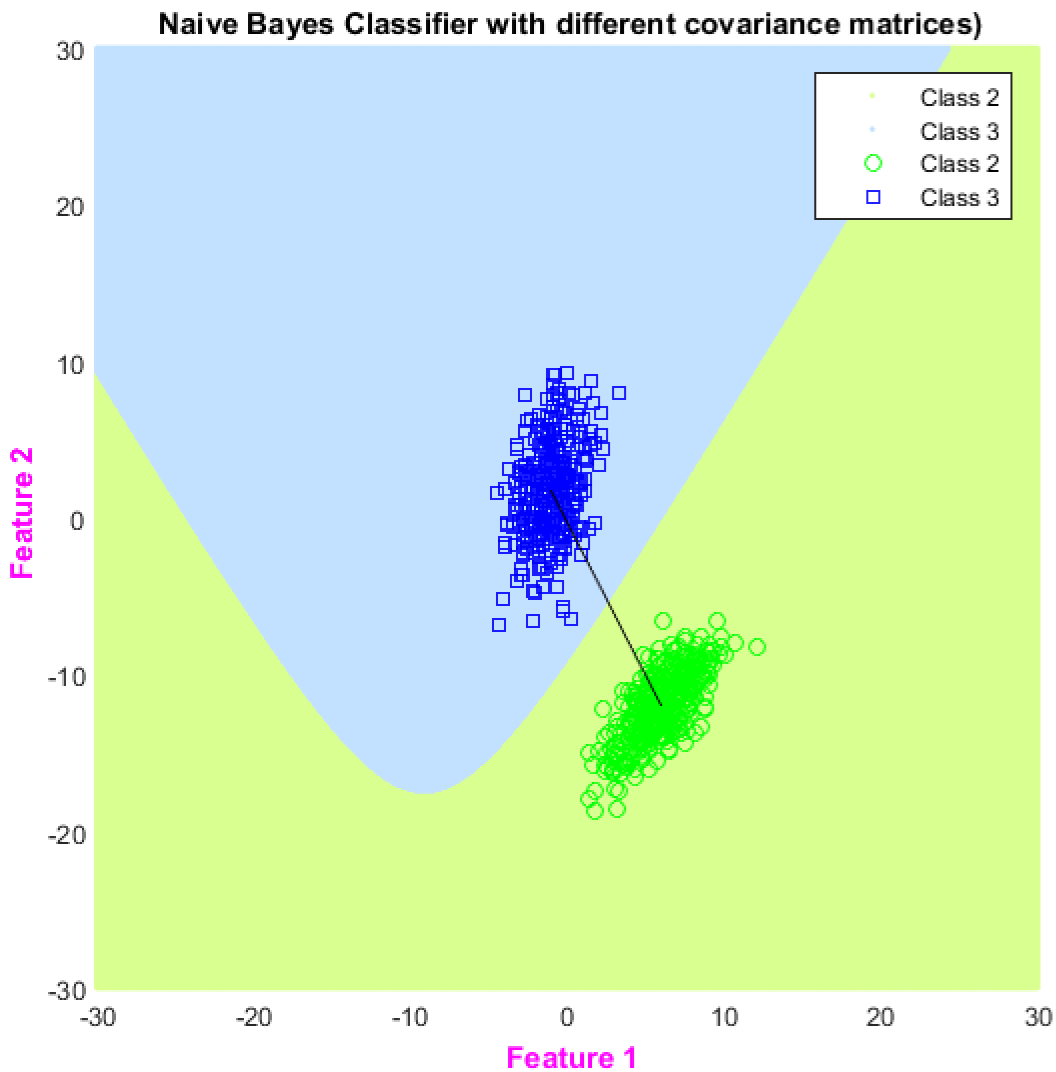
c) Covariance matrix for each class is different

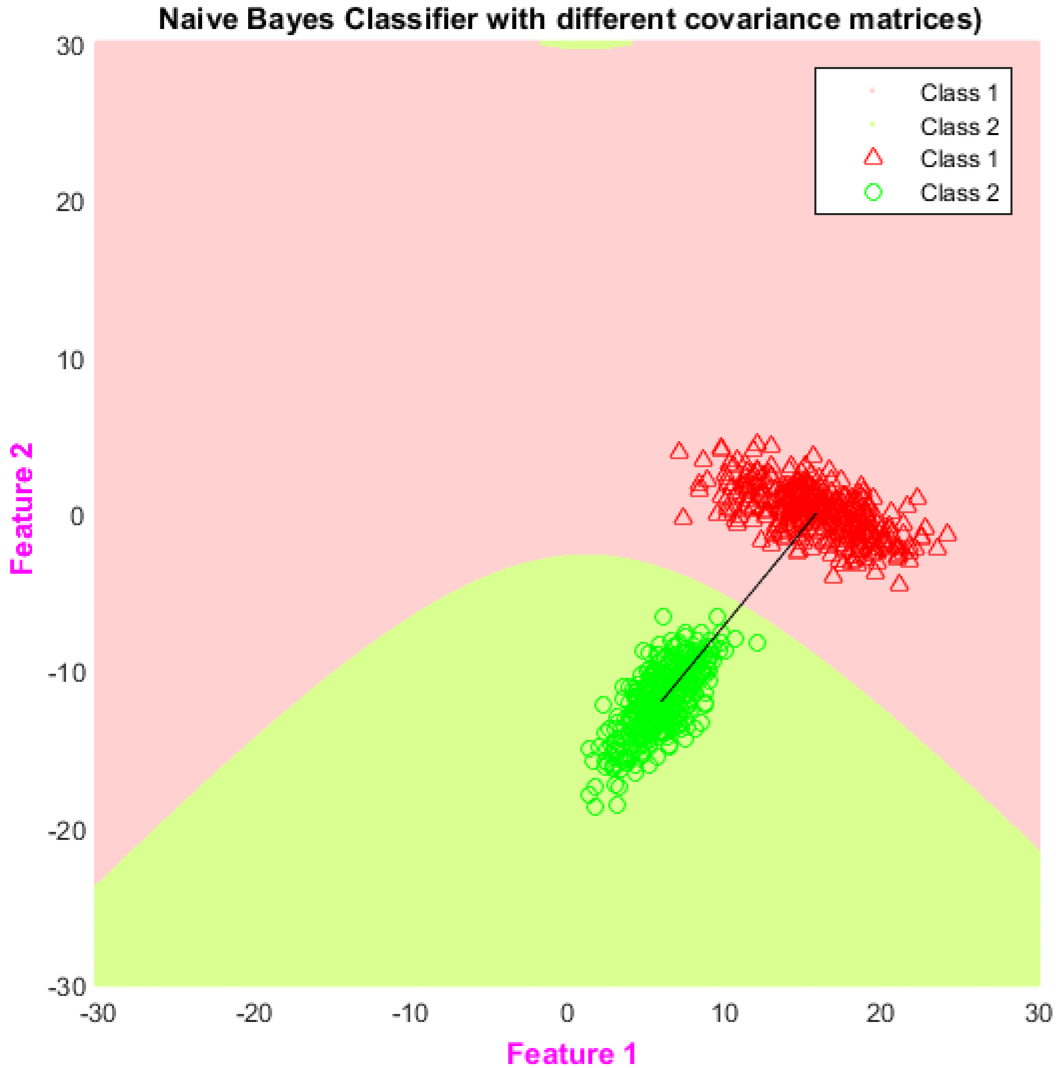
- *Decision region plot for all the classes together with the training data superposed -*



- *Decision region plot for every pair of classes -*




Naive Bayes Classifier with different covariance matrices)
Class 2
Class 3
Class 2
Class 3
Feature 2
Feature 1
30
20
10
0
-10
-20
-30
-30
-20
-10
0
10
20
30



- *<u>Confusion Matrix based on performance for test data-</u>*

Predicted Class ⇨	CLASS 1	CLASS 2	CLASS 3
Actual Class ⇩			
Class 1	125	0	0
Class 2	0	125	0
Class 3	0	0	125

- *<u>Classification accuracy on test data –</u>*
  Overall Accuracy - 100.0
  Classifier Accuracy for class 1 - 100.0
  Classifier Accuracy for class 2 - 100.0
  Classifier Accuracy for class 3 - 100.0

♦ Observations –

1) **Bayes Classifier –**

   a) **Same covariance matrices (by taking average)**
   The decision boundaries are all of **linear** nature. They pass through the **midpoint** of line joining the means of two classes. The boundaries are **not necessarily** perpendicular to this line.

   b) **Same covariance matrices (from training data of all classes combined)**
   The decision boundaries are all of **linear** nature. They pass through the **midpoint** of line joining the means of two classes. The boundaries are **not necessarily** perpendicular to this line.

   c) **Covariance matrix for each class is different**
   The decision boundaries are all of **Quadratic** nature. For this dataset they turn out to be either **Parabola** or **Hyperbola**. They do **not necessarily** pass through the **midpoint** of line joining the means of two classes.

2) **Naïve Bayes Classifier –**

   a) **Same covariance matrices (σ^2 I)**
   The decision boundaries are all of **linear** nature. They pass through the **midpoint** of line joining the means of two classes. The boundaries are perpendicular to this line.

   b) **Same covariance matrices and is C**
   The decision boundaries are all of **linear** nature. They pass through the **midpoint** of line joining the means of two classes. The boundaries are **not necessarily** perpendicular to this line.

   c) **Covariance matrix for each class is different**

   The decision boundaries are all of **Quadratic** nature. For this dataset they turn out to be either **Parabola** or **Hyperbola**. They do **not necessarily** pass through the **midpoint** of line joining the means of two classes. The classification results are more or less similar to that of Bayes.

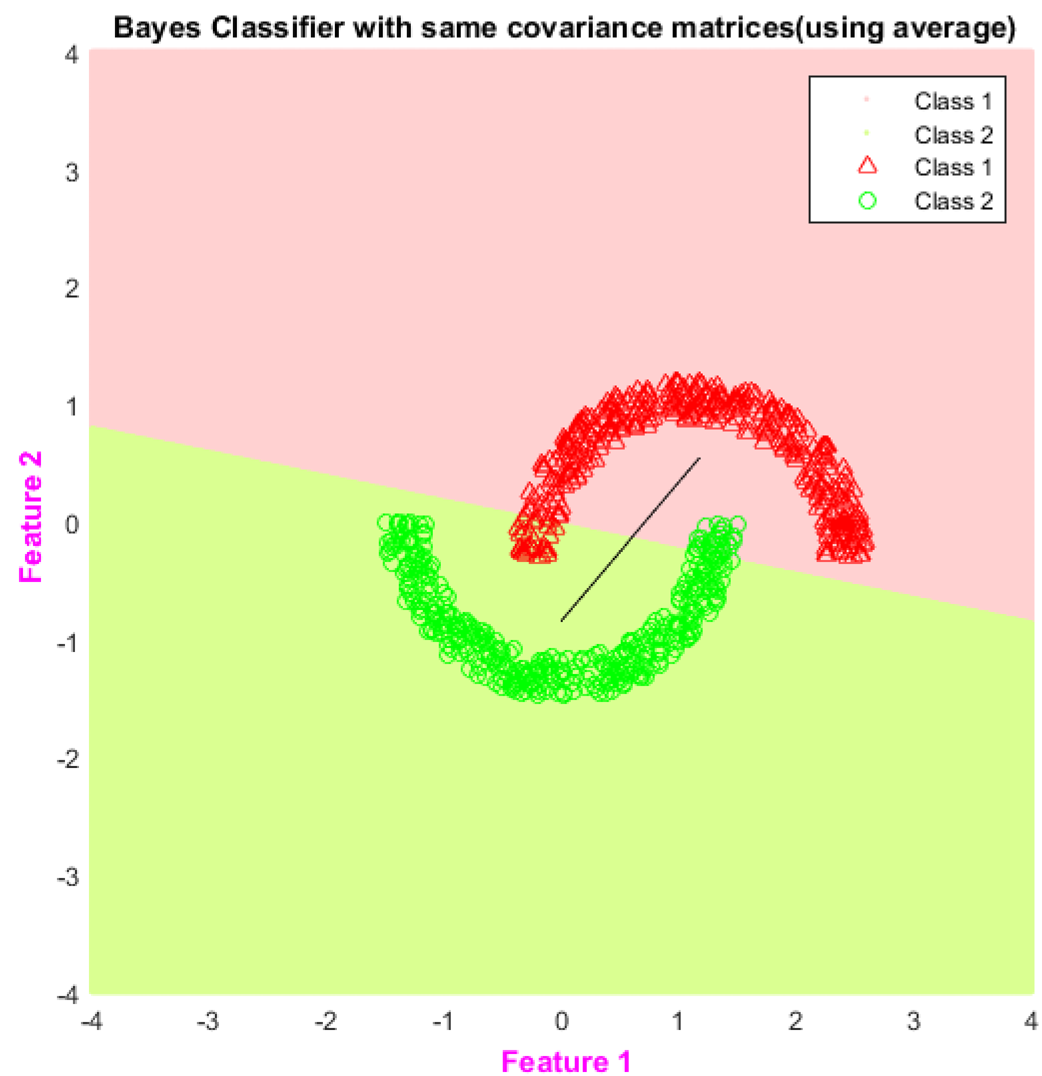
♦ **Dataset I-(b)**: 2-dimensional artificial data of 2 classes that are **nonlinearly separable**

<u>INTERLOCKING CLASSES –</u>

1) Bayes Classifier –

a) Same covariance matrices (by taking average)

- *<u>Decision region plot for all the classes together with the training data superposed -</u>*



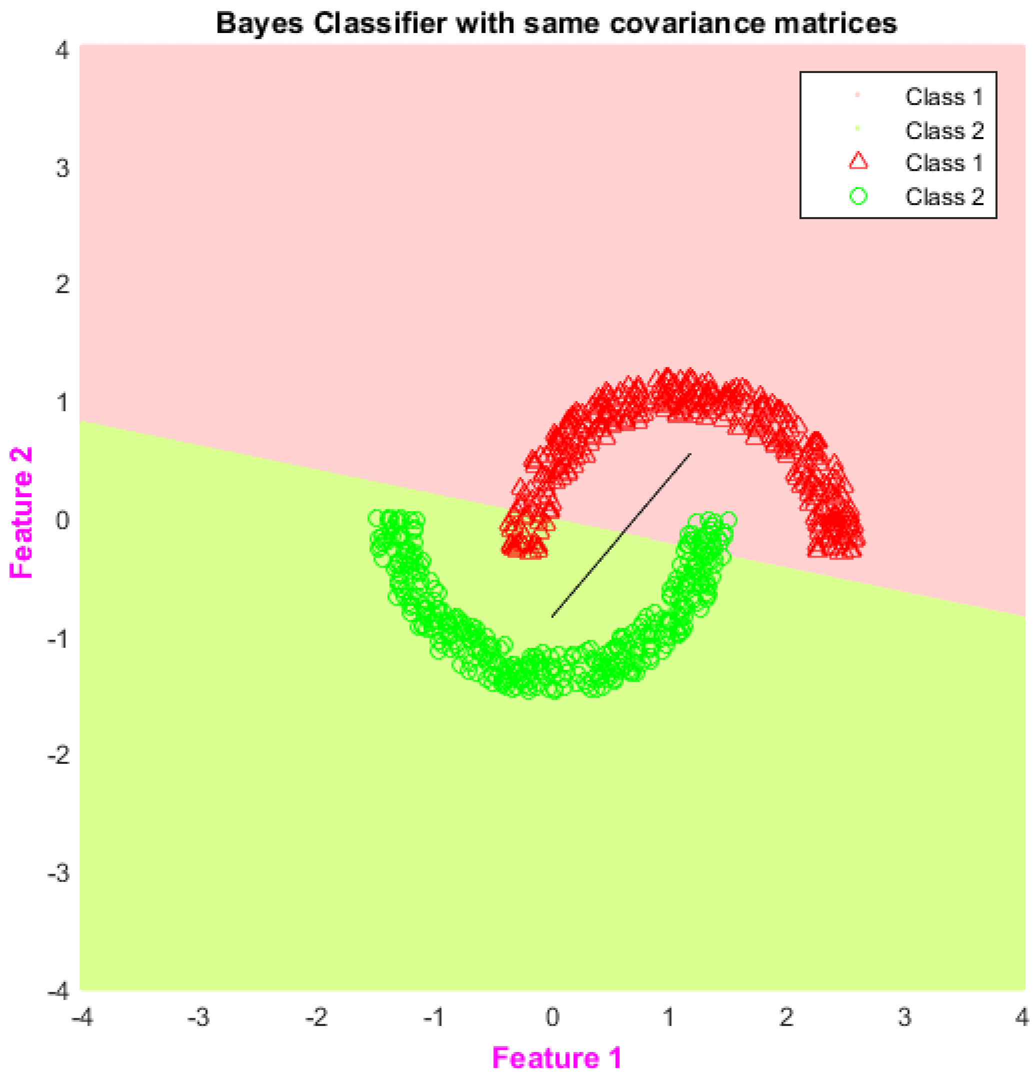
- *<u>Confusion Matrix based on performance for test data-</u>*

Predicted Class ⇨	CLASS 1	CLASS 2
Actual Class ⇩		
Class 1	118	7
Class 2	4	121

- *<u>Classification accuracy on test data –</u>*
  Overall Accuracy – 95.6
  Classifier Accuracy for class 1 – 94.4
  Classifier Accuracy for class 2 – 96.8

b) Same covariance matrices (from training data of all classes combined)

- *Decision region plot for all the classes together with the training data superposed -*



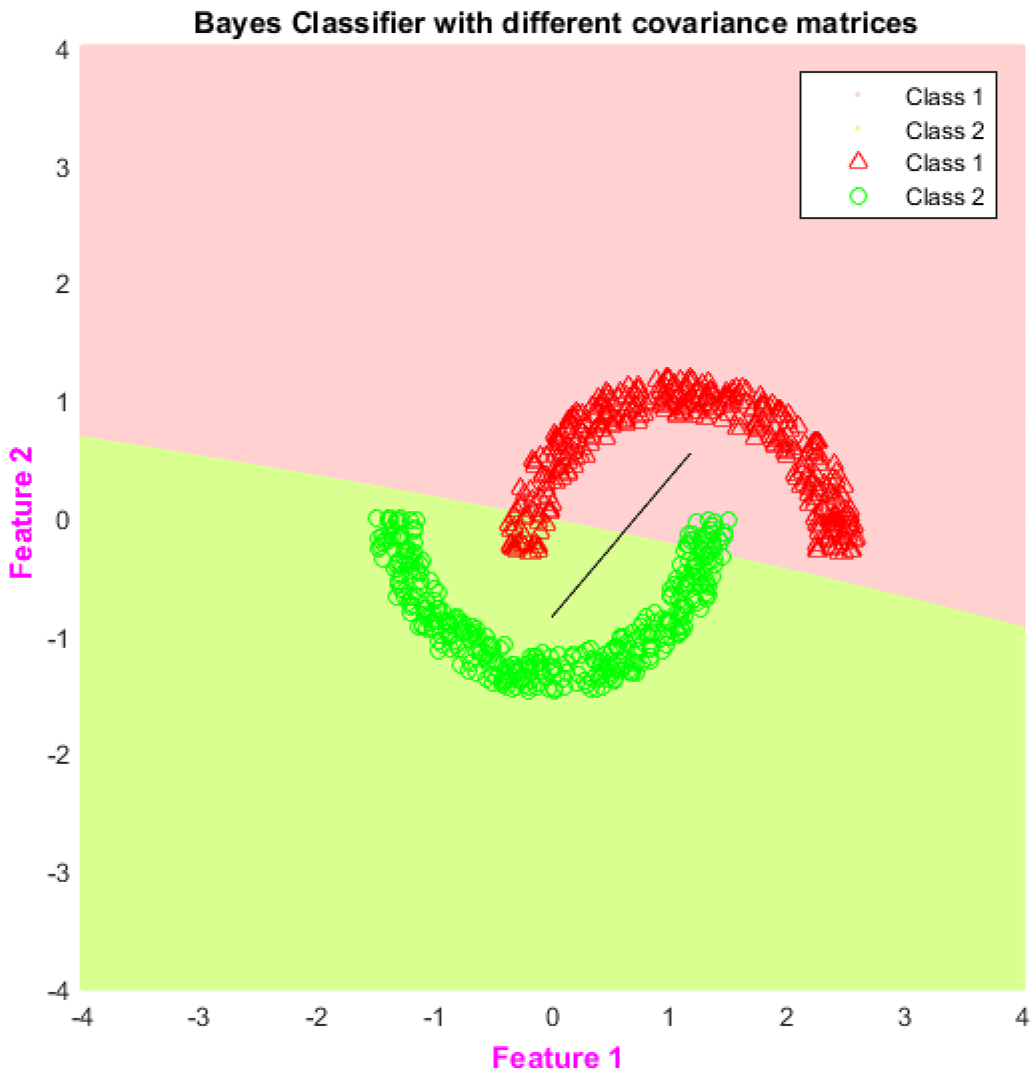
- *Confusion Matrix based on performance for test data-*

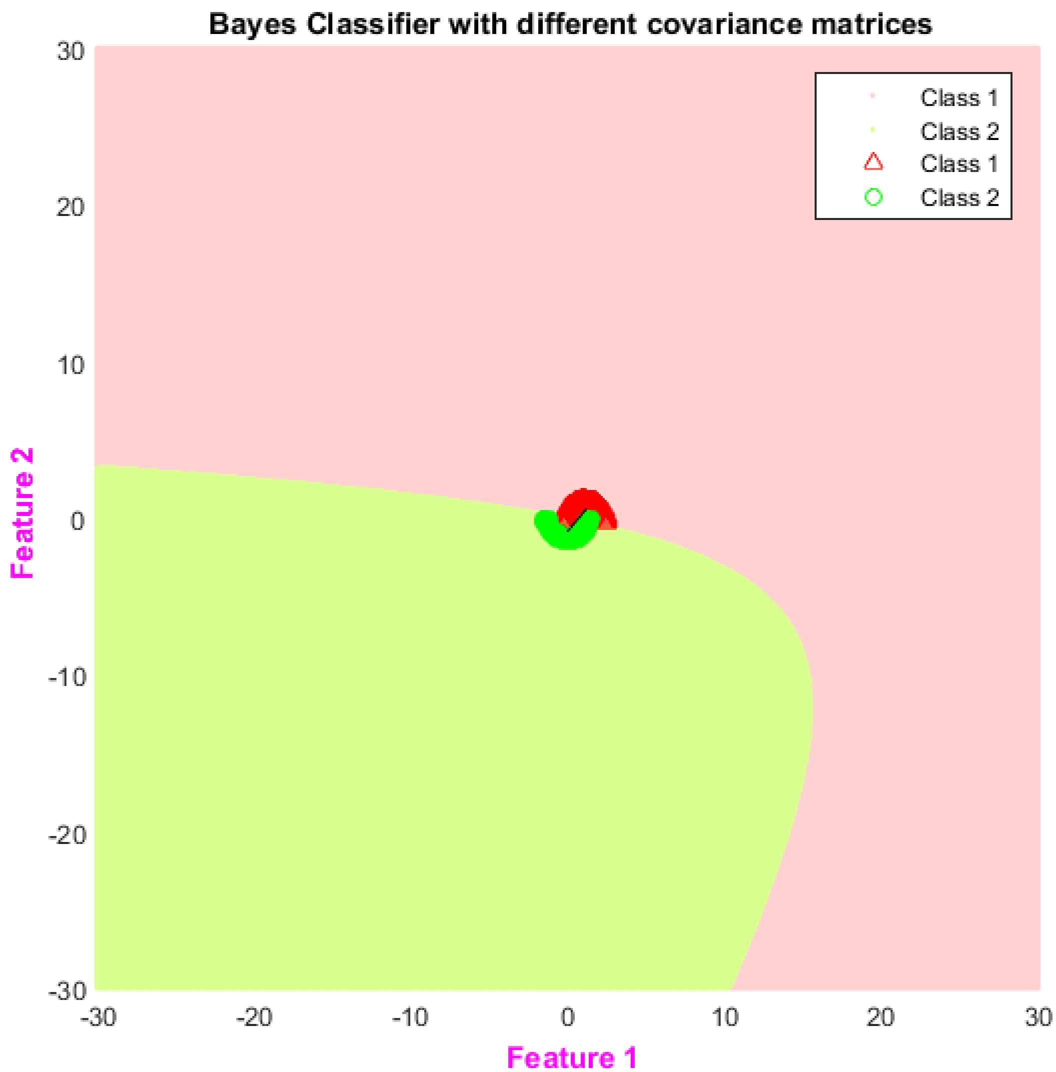
Predicted Class ⇨	CLASS 1	CLASS 2
Actual Class ⇩		
Class 1	118	7
Class 2	4	121

- *Classification accuracy on test data –*
  Overall Accuracy – 95.6
  Classifier Accuracy for class 1 - 94.4
  Classifier Accuracy for class 2 – 96.8

b) Covariance matrix for each class is different

- *<u>Decision region plot for all the classes together with the training data superposed -</u>*





Bayes Classifier with different covariance matrices
Class 1
Class 2
Class 1
Class 2
Feature 2
Feature 1
30
20
10
0
-10
-20
-30
-30
-20
-10
0
10
20
30

- *Confusion Matrix based on performance for test data-*

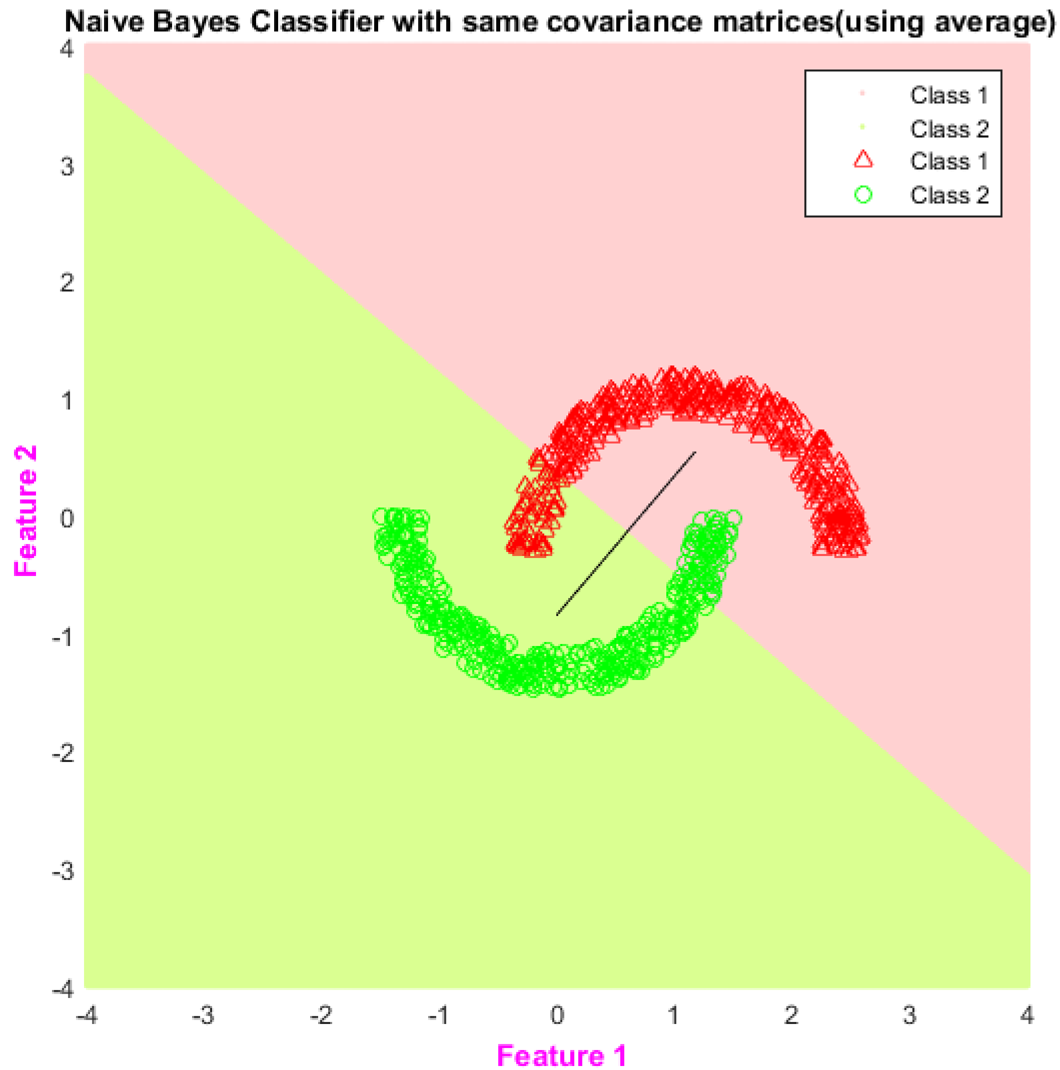
Predicted Class ⇨	CLASS 1	CLASS 2
Actual Class ⇩		
Class 1	118	7
Class 2	3	122

- *Classification accuracy on test data –*
  Overall Accuracy – 96.0
  Classifier Accuracy for class 1 – 94.4
  Classifier Accuracy for class 2 – 97.6

2) Naive Bayes Classifier –

a) Same covariance matrices (σ^2 I)

- *Decision region plot for all the classes together with the training data superposed -*



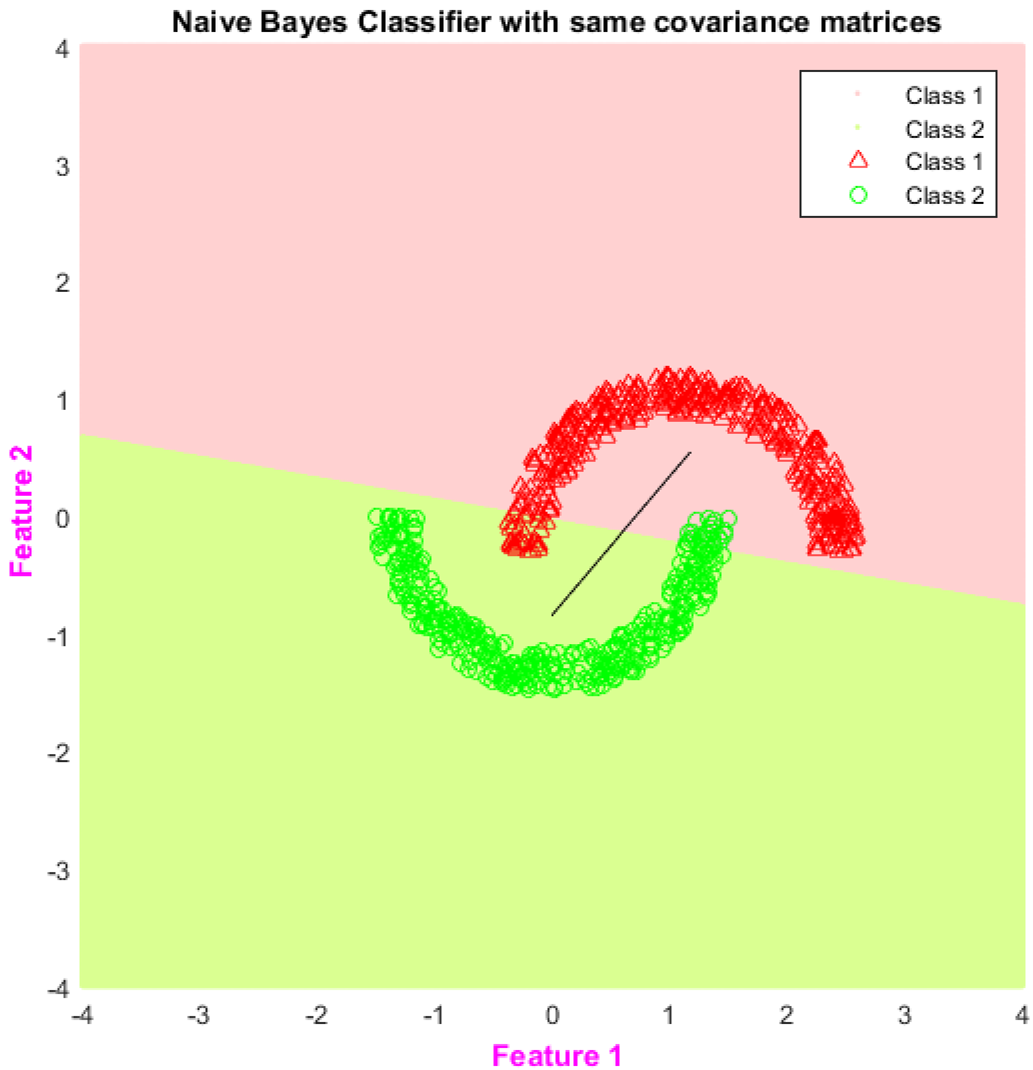
- *Confusion Matrix based on performance for test data-*

Predicted Class ⇨	CLASS 1	CLASS 2
Actual Class ⇩		
Class 1	106	19
Class 2	12	113

- *Classification accuracy on test data –*
  Overall Accuracy – 87.6
  Classifier Accuracy for class 1 – 84.8
  Classifier Accuracy for class 2 – 90.4

b) Same covariance matrices and is C

- *<u>Decision region plot for all the classes together with the training data superposed -</u>*



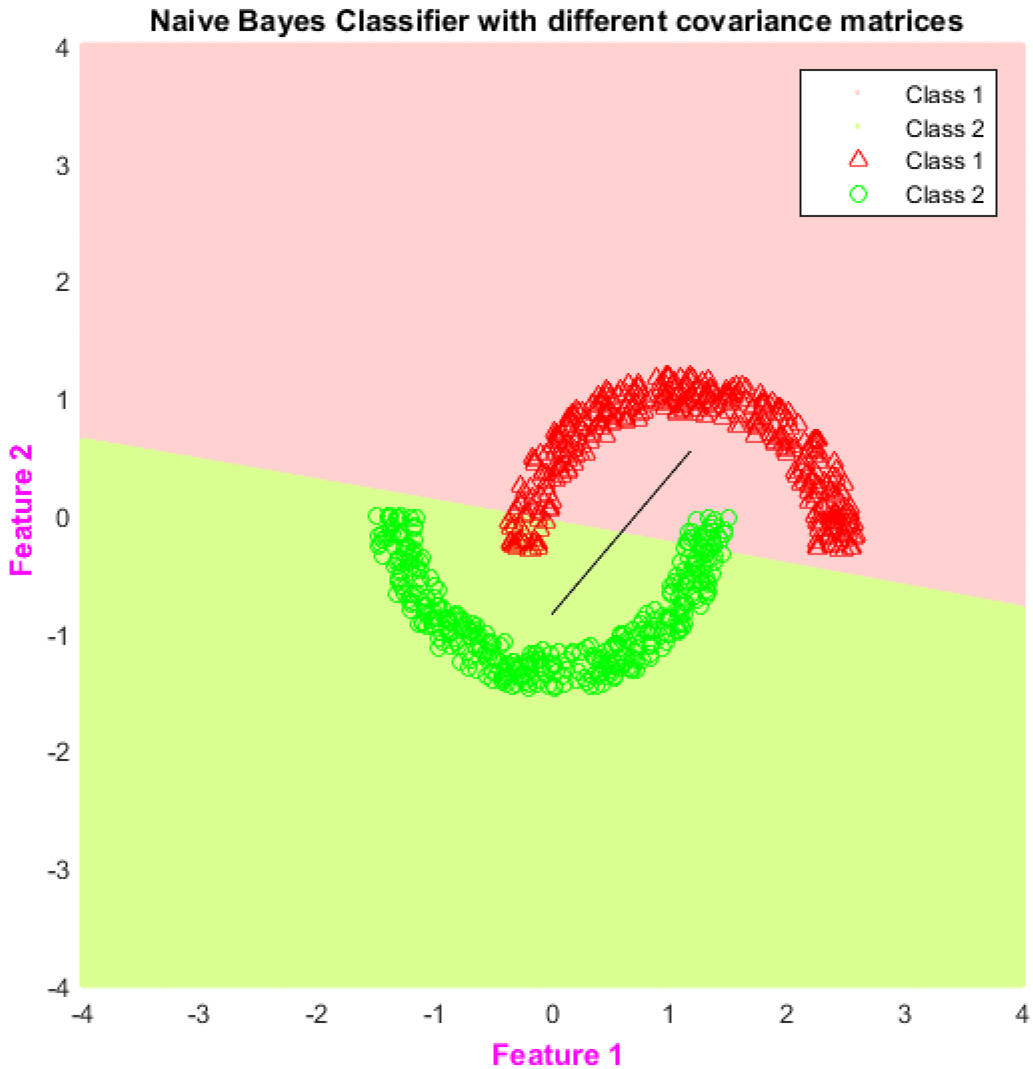
- *Confusion Matrix based on performance for test data-*

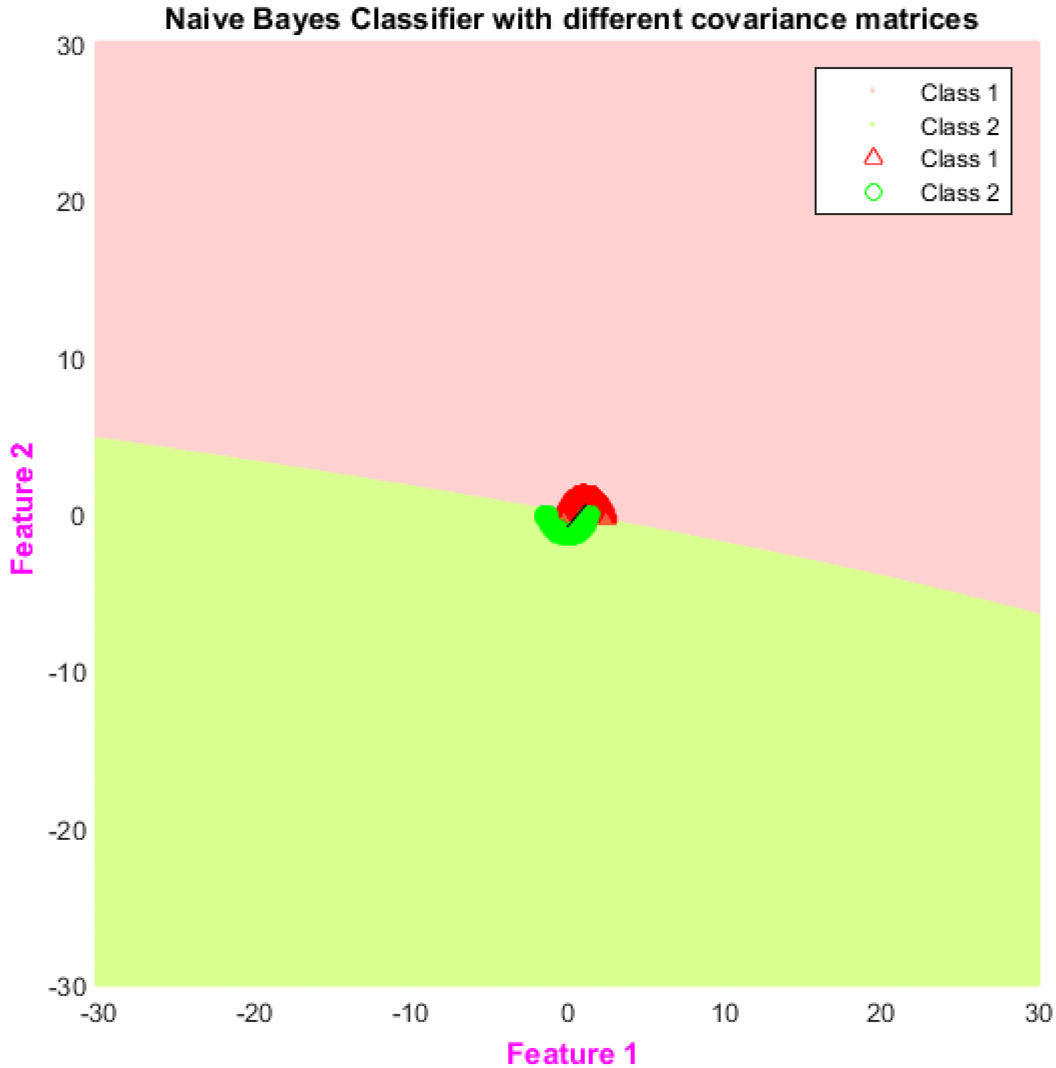
Predicted Class ⇨	CLASS 1	CLASS 2
Actual Class ⇩		
Class 1	118	7
Class 2	3	122

- *Classification accuracy on test data –*
  Overall Accuracy – 96.0
  Classifier Accuracy for class 1 – 94.4
  Classifier Accuracy for class 2 – 97.6

b) Covariance matrix for each class is different

- *Decision region plot for all the classes together with the training data superposed -*




Naive Bayes Classifier with different covariance matrices
Class 1
Class 2
Class 1
Class 2
Feature 2
Feature 1
30
20
10
0
-10
-20
-30
-30
-20
-10
0
10
20
30

- *Confusion Matrix based on performance for test data-*

Predicted Class ⇨	CLASS 1	CLASS 2
Actual Class ⇩		
Class 1	118	7
Class 2	3	122

- *Classification accuracy on test data –*
  Overall Accuracy – 96.0
  Classifier Accuracy for class 1 – 94.4
  Classifier Accuracy for class 2 – 97.6

♦ Observations –

1) Bayes Classifier –

a) Same covariance matrices (by taking average)

The decision boundaries are all of **linear** nature. They pass through the **midpoint** of line joining the means of two classes. The boundaries are **not necessarily** perpendicular to this line.

b) Same covariance matrices (from training data of all classes combined)

The decision boundaries are all of **linear** nature. They pass through the **midpoint** of line joining the means of two classes. The boundaries are **not necessarily** perpendicular to this line.

c) Covariance matrix for each class is different

The decision boundaries are all of **Quadratic** nature. For this dataset they turn out to be either **Parabola**. They do **not necessarily** pass through the **midpoint** of line joining the means of two classes.

2) Naïve Bayes Classifier –

a) Same covariance matrices (σ^2 I)

The decision boundaries are all of **linear** nature. They pass through the **midpoint** of line joining the means of two classes. The boundaries are perpendicular to this line.

b) Same covariance matrices and is C

The decision boundaries are all of **linear** nature. They pass through the **midpoint** of line joining the means of two classes. The boundaries are **not necessarily** perpendicular to this line.

c) Covariance matrix for each class is different

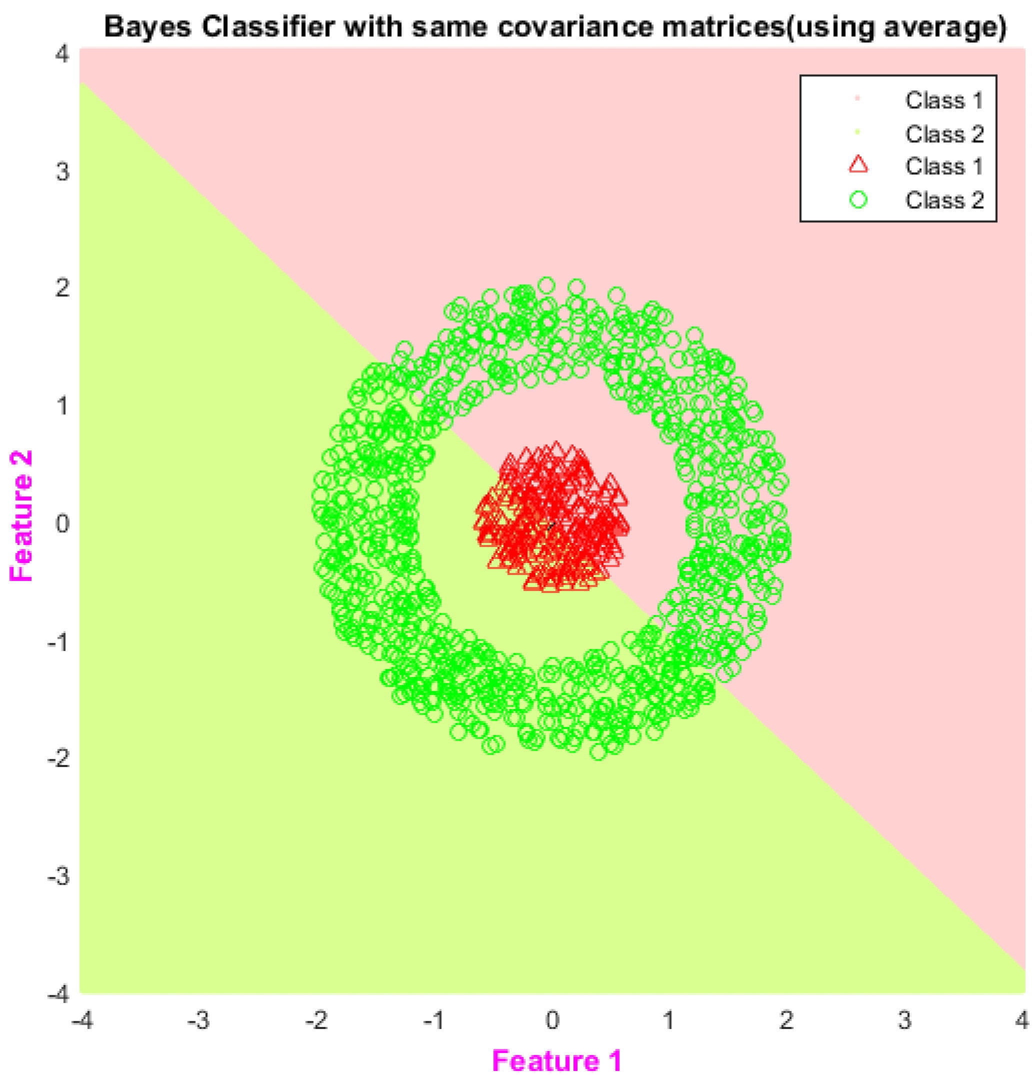
The decision boundaries are all of **Quadratic** nature. For this dataset they turn out to be either **Parabola**. They do **not necessarily** pass through the **midpoint** of line joining the means of two classes.

A RING WITH A CENTRAL MASS –

1) Bayes Classifier –

a) Same covariance matrices (by taking average)

- *Decision region plot for all the classes together with the training data superposed -*



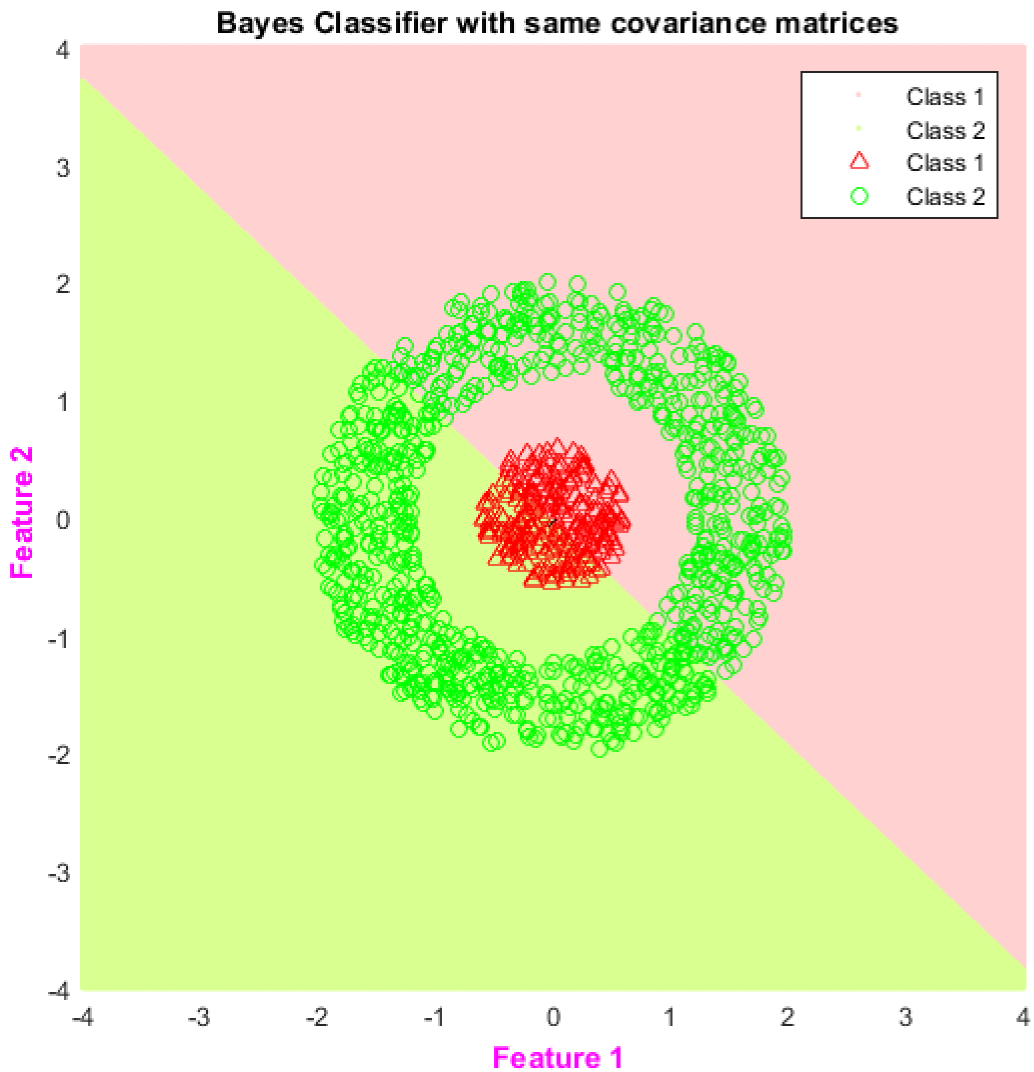
- *Confusion Matrix based on performance for test data-*

Predicted Class ⇨	CLASS 1	CLASS 2
Actual Class ⇩		
Class 1	46	29
Class 2	158	142

- *Classification accuracy on test data –*
  Overall Accuracy – 50.1333
  Classifier Accuracy for class 1 – 61.3333
  Classifier Accuracy for class 2 – 47.3333

b) Same covariance matrices (from training data of all classes combined)

- *<u>Decision region plot for all the classes together with the training data superposed -</u>*



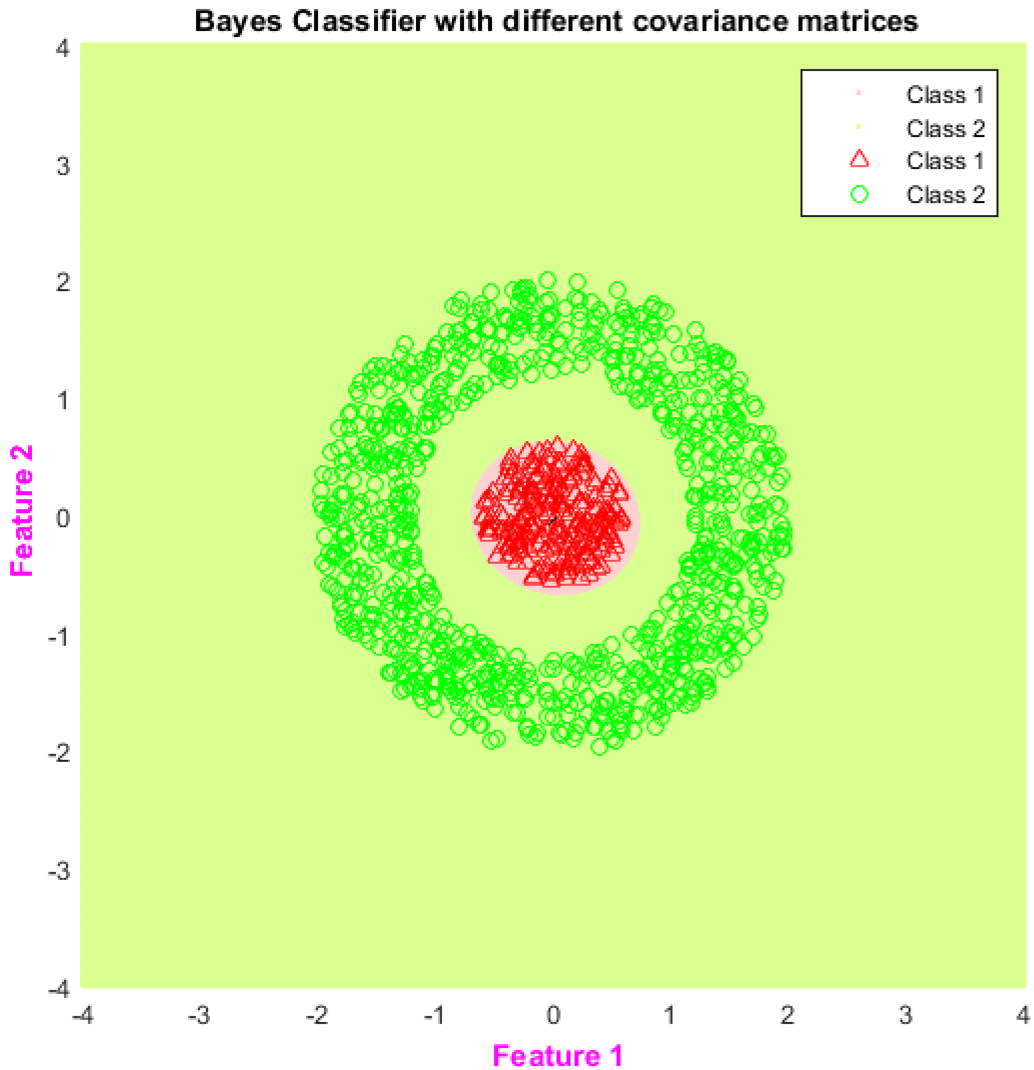
- *<u>Confusion Matrix based on performance for test data-</u>*

Predicted Class ⇨	CLASS 1	CLASS 2
Actual Class ⇩		
Class 1	46	29
Class 2	158	142

- *<u>Classification accuracy on test data –</u>*
  Overall Accuracy – 50.1333
  Classifier Accuracy for class 1 – 61.3333
  Classifier Accuracy for class 2 – 47.3333

b) Covariance matrix for each class is different

- *Decision region plot for all the classes together with the training data superposed -*



- *Confusion Matrix based on performance for test data-*

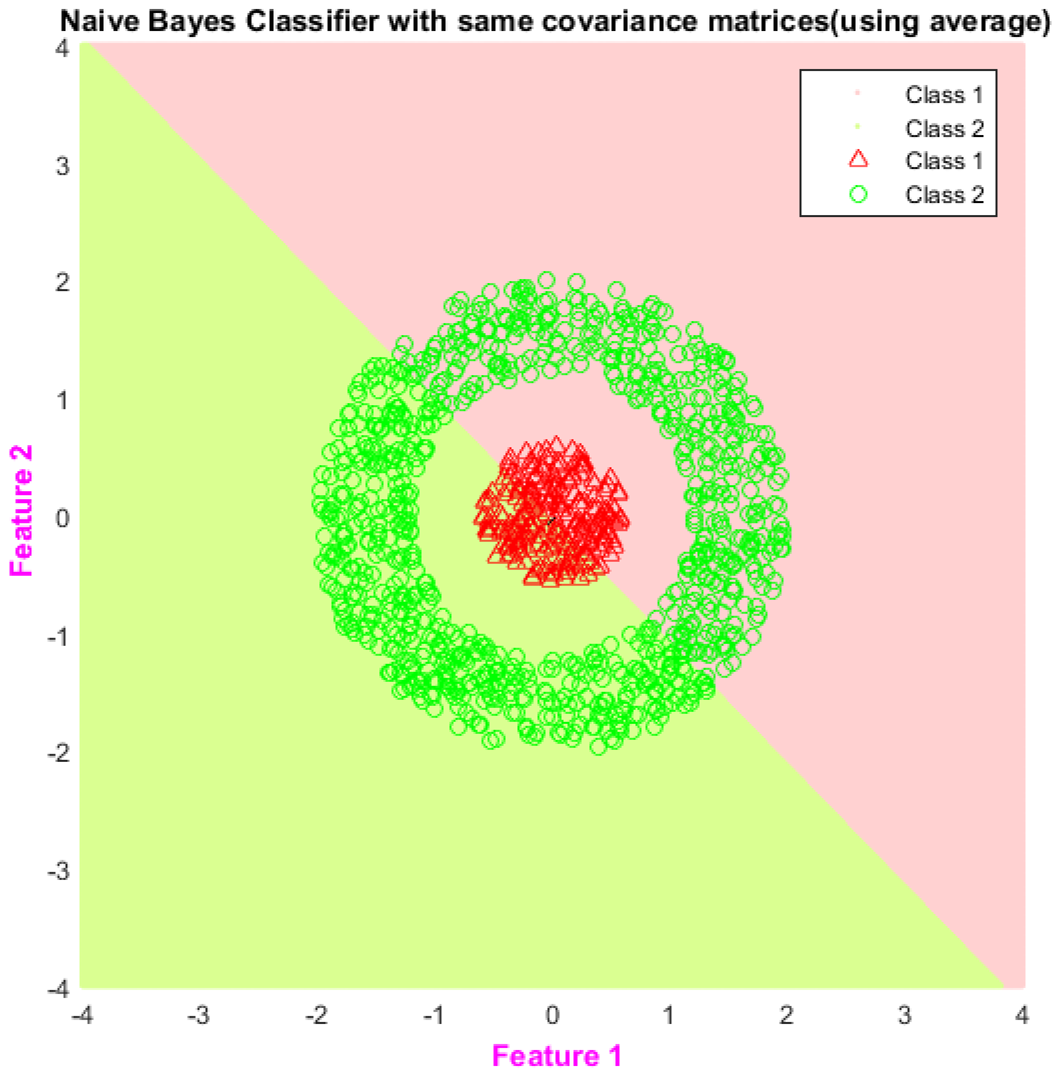
Predicted Class ⇨	CLASS 1	CLASS 2
Actual Class ⇩		
Class 1	75	0
Class 2	0	300

- *Classification accuracy on test data –*
  Overall Accuracy – 100.0
  Classifier Accuracy for class 1 – 100.0
  Classifier Accuracy for class 2 – 100.0

2) Naive Bayes Classifier –

a) Same covariance matrices (σ^2 I)

- *Decision region plot for all the classes together with the training data superposed -*



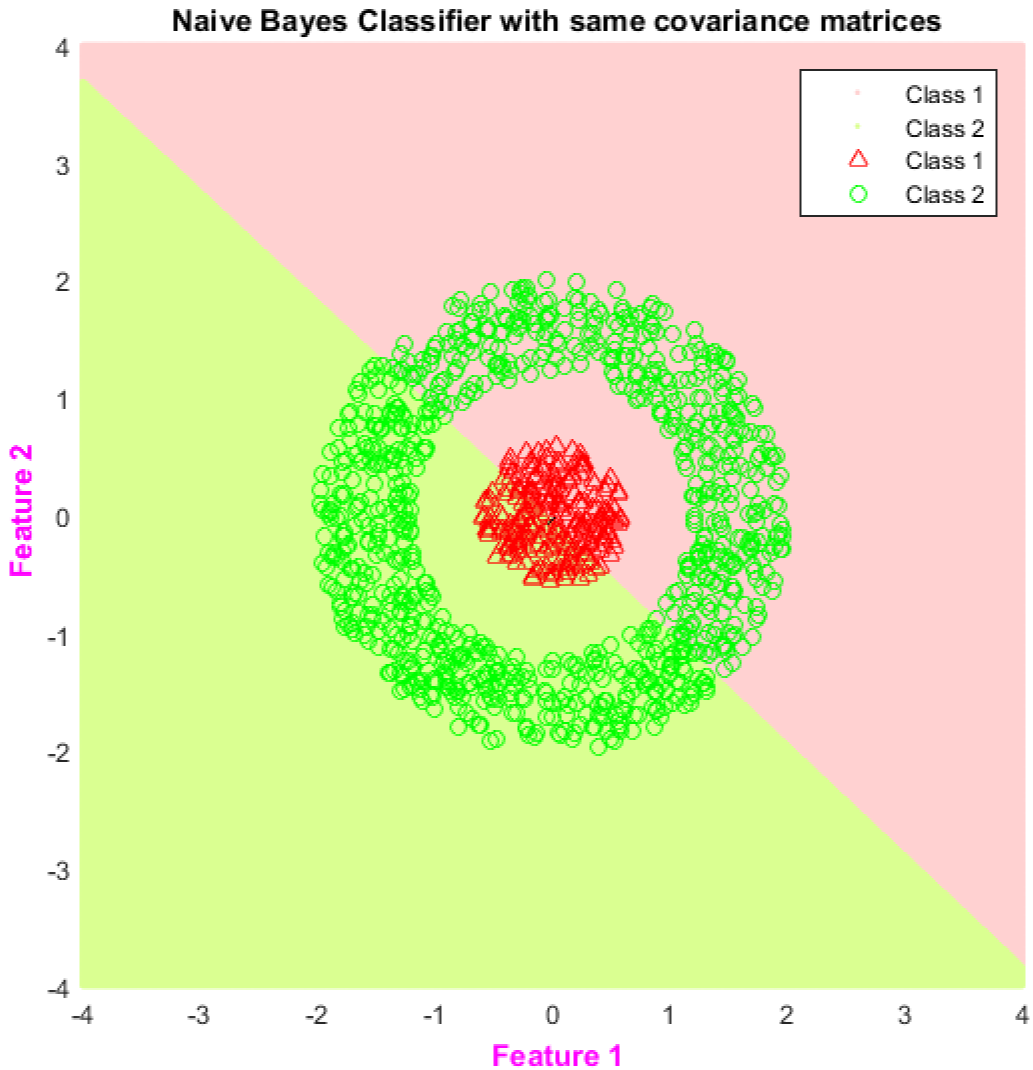
- *Confusion Matrix based on performance for test data-*

Predicted Class ⇨	CLASS 1	CLASS 2
Actual Class ⇩		
Class 1	47	28
Class 2	155	145

- *Classification accuracy on test data –*
  Overall Accuracy – 51.2
  Classifier Accuracy for class 1 – 62.6667
  Classifier Accuracy for class 2 – 48.3333

b) Same covariance matrices and is C

- *Decision region plot for all the classes together with the training data superposed -*



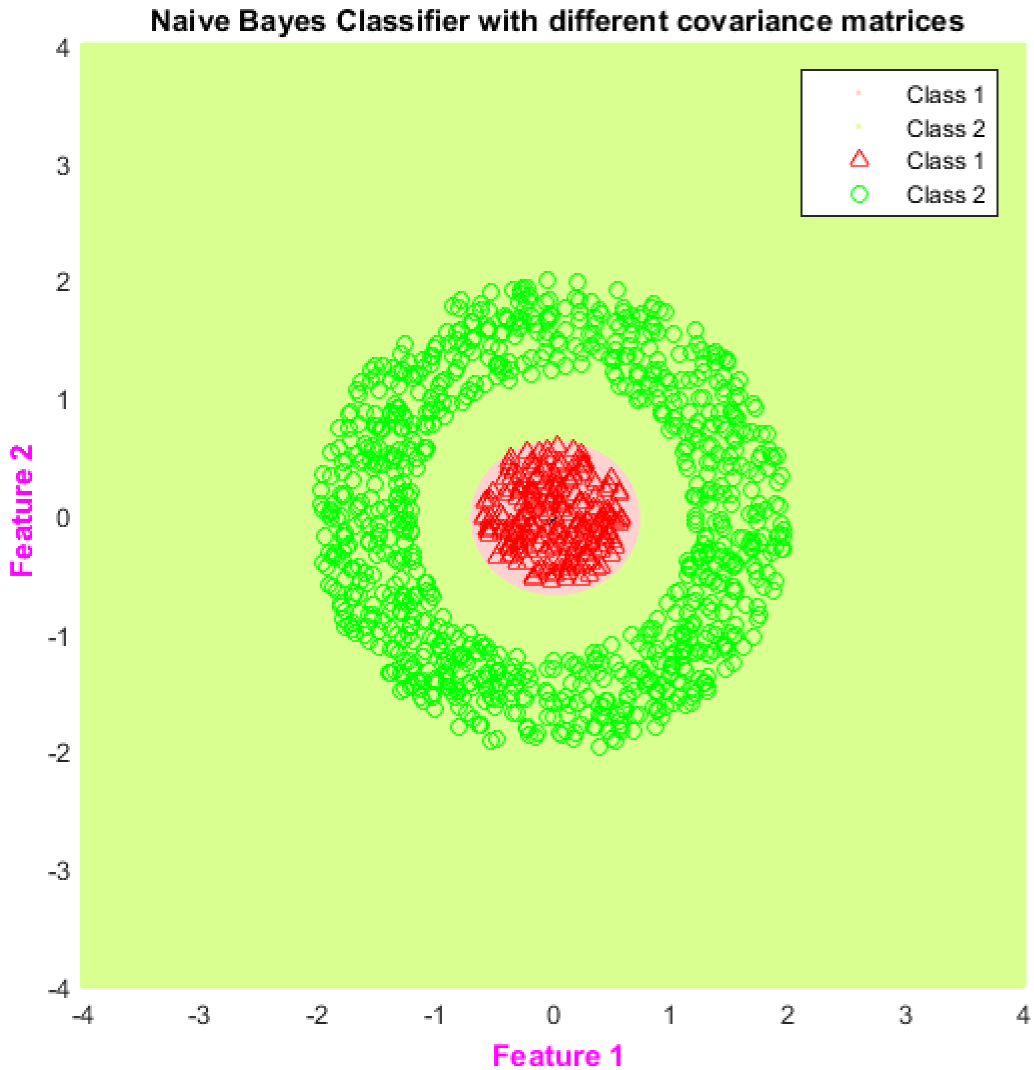
- *Confusion Matrix based on performance for test data-*

Predicted Class ⇨	CLASS 1	CLASS 2
Actual Class ⇩		
Class 1	46	29
Class 2	158	142

- *Classification accuracy on test data –*
  Overall Accuracy – 50.1333
  Classifier Accuracy for class 1 – 61.3333
  Classifier Accuracy for class 2 – 47.3333

c) Covariance matrix for each class is different

- *Decision region plot for all the classes together with the training data superposed -*



- *Confusion Matrix based on performance for test data-*

Predicted Class ⇨	CLASS 1	CLASS 2
Actual Class ⇩		
Class 1	75	0
Class 2	0	300

- *Classification accuracy on test data –*
  Overall Accuracy – 100.0
  Classifier Accuracy for class 1 – 100.0
  Classifier Accuracy for class 2 – 100.0

♦ Observations –

1) **Bayes Classifier –**
   a) **Same covariance matrices (by taking average)**
   The decision boundaries are all of **linear** nature. They pass through the **midpoint** of line joining the means of two classes. The boundaries are **not necessarily** perpendicular to this line.

   b) **Same covariance matrices (from training data of all classes combined)**
   The decision boundaries are all of **linear** nature. They pass through the **midpoint** of line joining the means of two classes. The boundaries are **not necessarily** perpendicular to this line.

   c) **Covariance matrix for each class is different**
   The decision boundaries are all of **Quadratic** nature. For this dataset they turn out to be **Circle**.

2) **Naïve Bayes Classifier –**
   a) **Same covariance matrices (σ^2 I)**
   The decision boundaries are all of **linear** nature. They pass through the **midpoint** of line joining the means of two classes. The boundaries are perpendicular to this line.

   b) **Same covariance matrices and is C**
   The decision boundaries are all of **linear** nature. They pass through the **midpoint** of line joining the means of two classes. The boundaries are **not necessarily** perpendicular to this line.

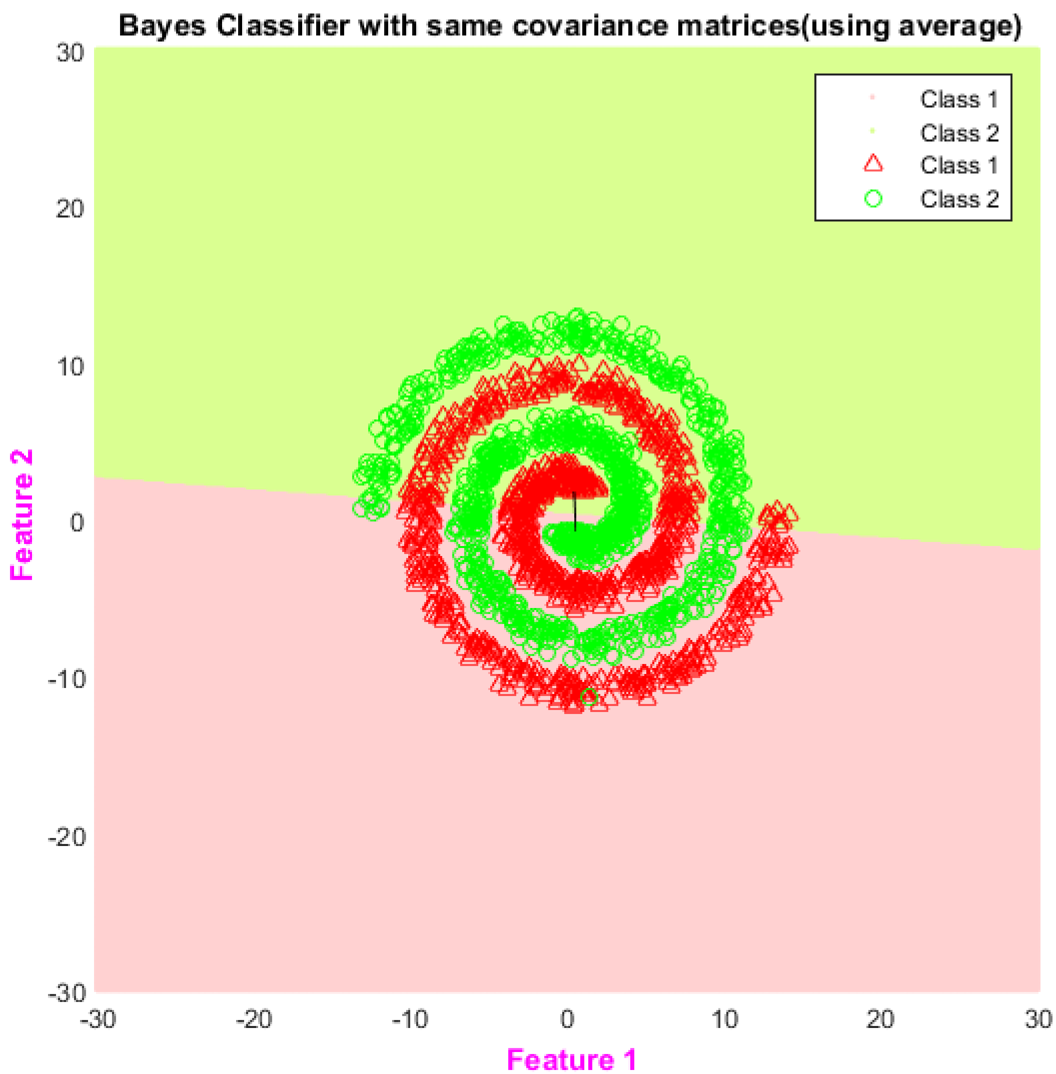
   c) **Covariance matrix for each class is different**
   The decision boundaries are all of **Quadratic** nature. For this dataset they turn out to be **Circle**. The classification results are more or less similar to that of Bayes.

SPIRAL DATASET –

1) Bayes Classifier –

a) Same covariance matrices (by taking average)

- *Decision region plot for all the classes together with the training data superposed -*



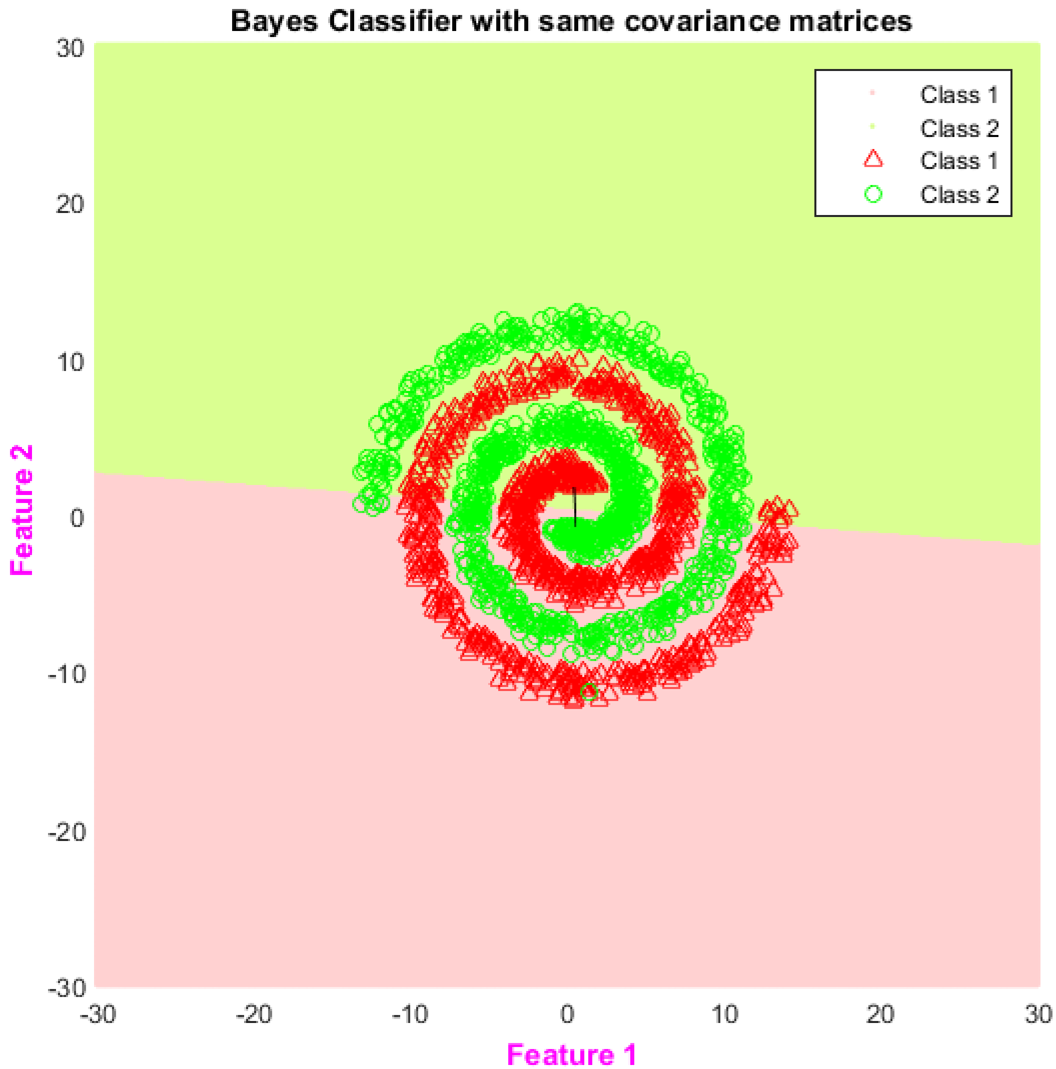
- *Confusion Matrix based on performance for test data-*

Predicted Class ⇨	CLASS 1	CLASS 2
Actual Class ⇩		
Class 1	175	151
Class 2	151	175

- *Classification accuracy on test data –*
  Overall Accuracy – 53.6810
  Classifier Accuracy for class 1 - 53.6810
  Classifier Accuracy for class 2 - 53.6810

b) Same covariance matrices (from training data of all classes combined)

- *Decision region plot for all the classes together with the training data superposed -*



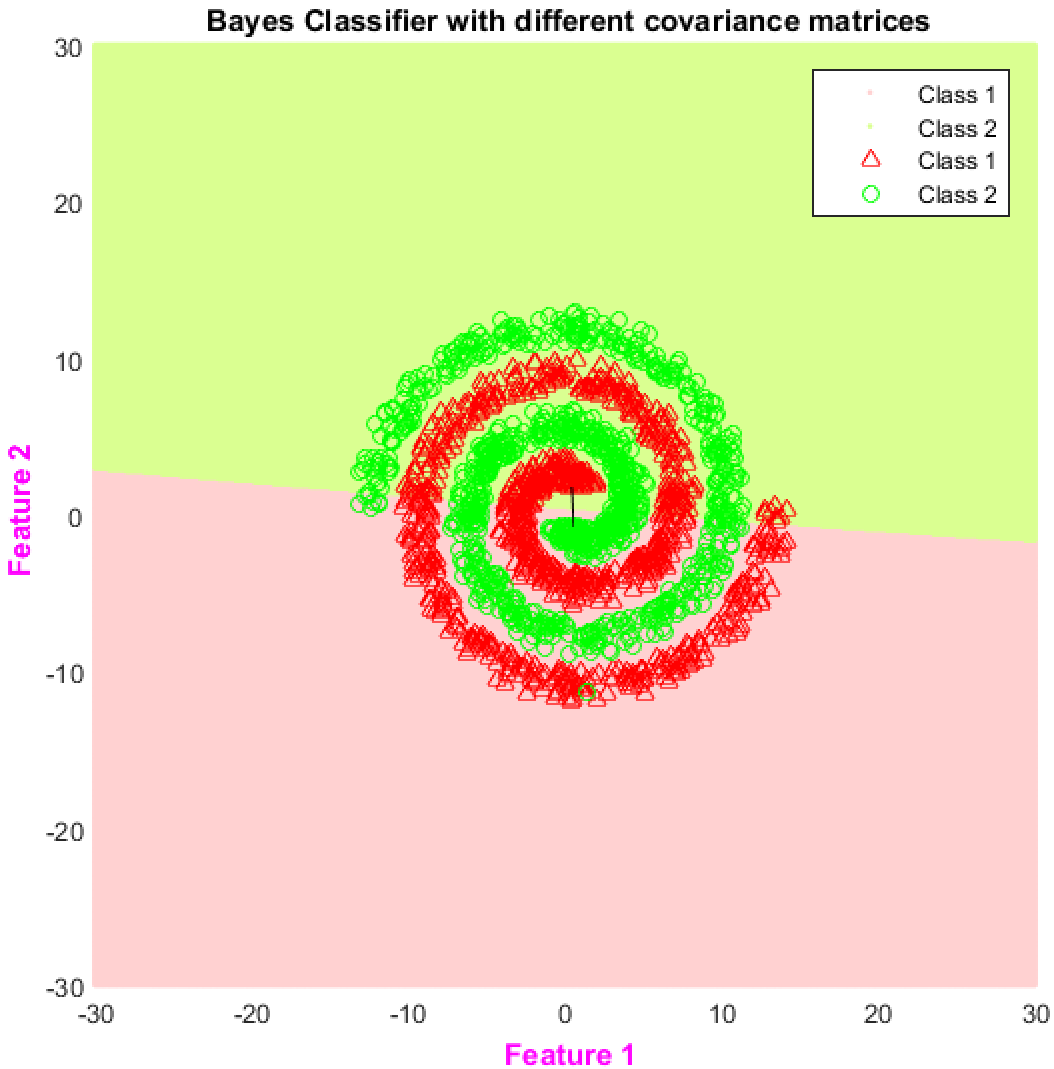
- *Confusion Matrix based on performance for test data-*

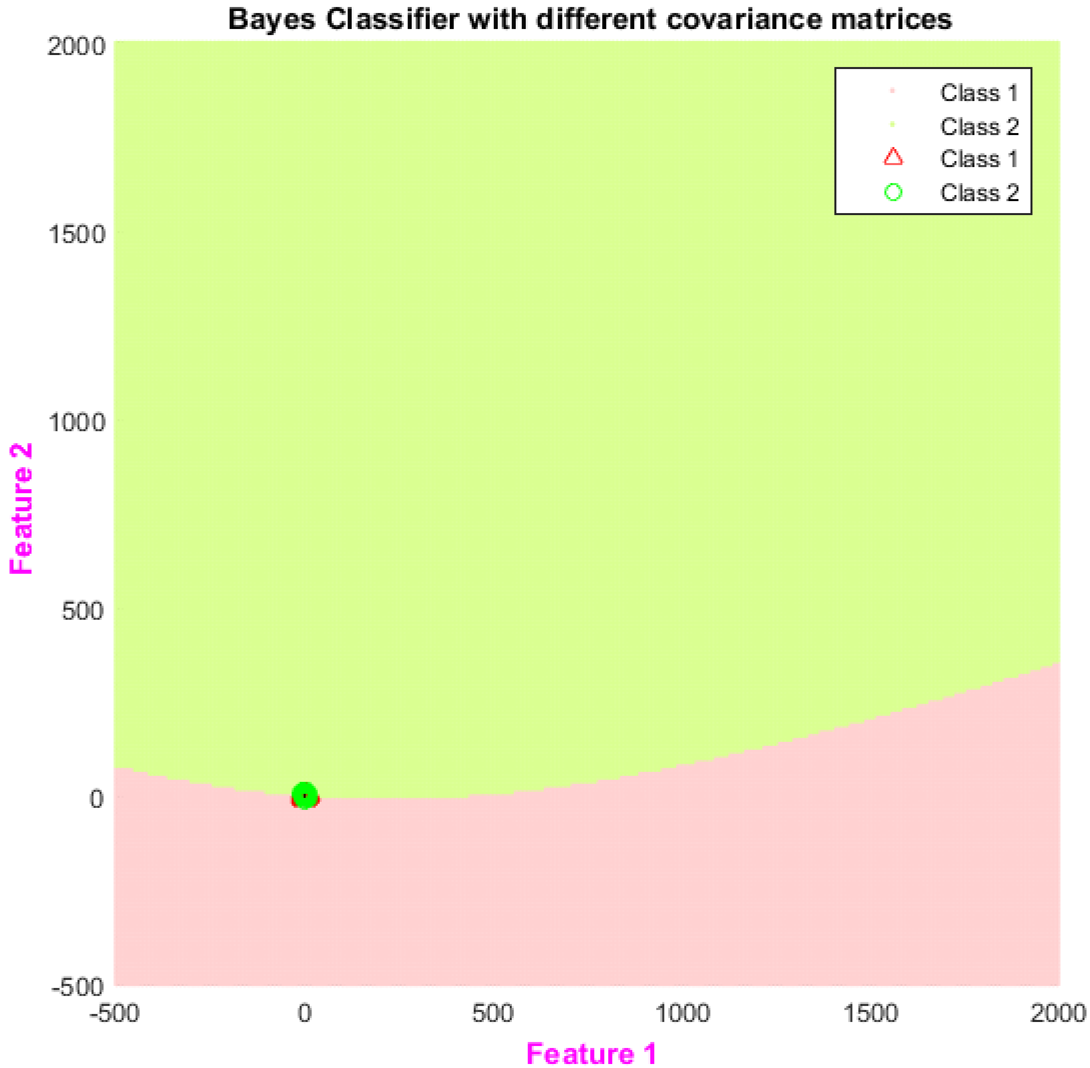
Predicted Class ⇨	CLASS 1	CLASS 2
Actual Class ⇩		
Class 1	175	151
Class 2	151	175

- *Classification accuracy on test data –*
  Overall Accuracy – 53.6810
  Classifier Accuracy for class 1 – 53.6810
  Classifier Accuracy for class 2 – 53.6810

c) Covariance matrix for each class is different

- *Decision region plot for all the classes together with the training data superposed -*




Bayes Classifier with different covariance matrices
Class 1
Class 2
Class 1
Class 2
Feature 1
Feature 2
-500
0
500
1000
1500
2000

- *Confusion Matrix based on performance for test data-*

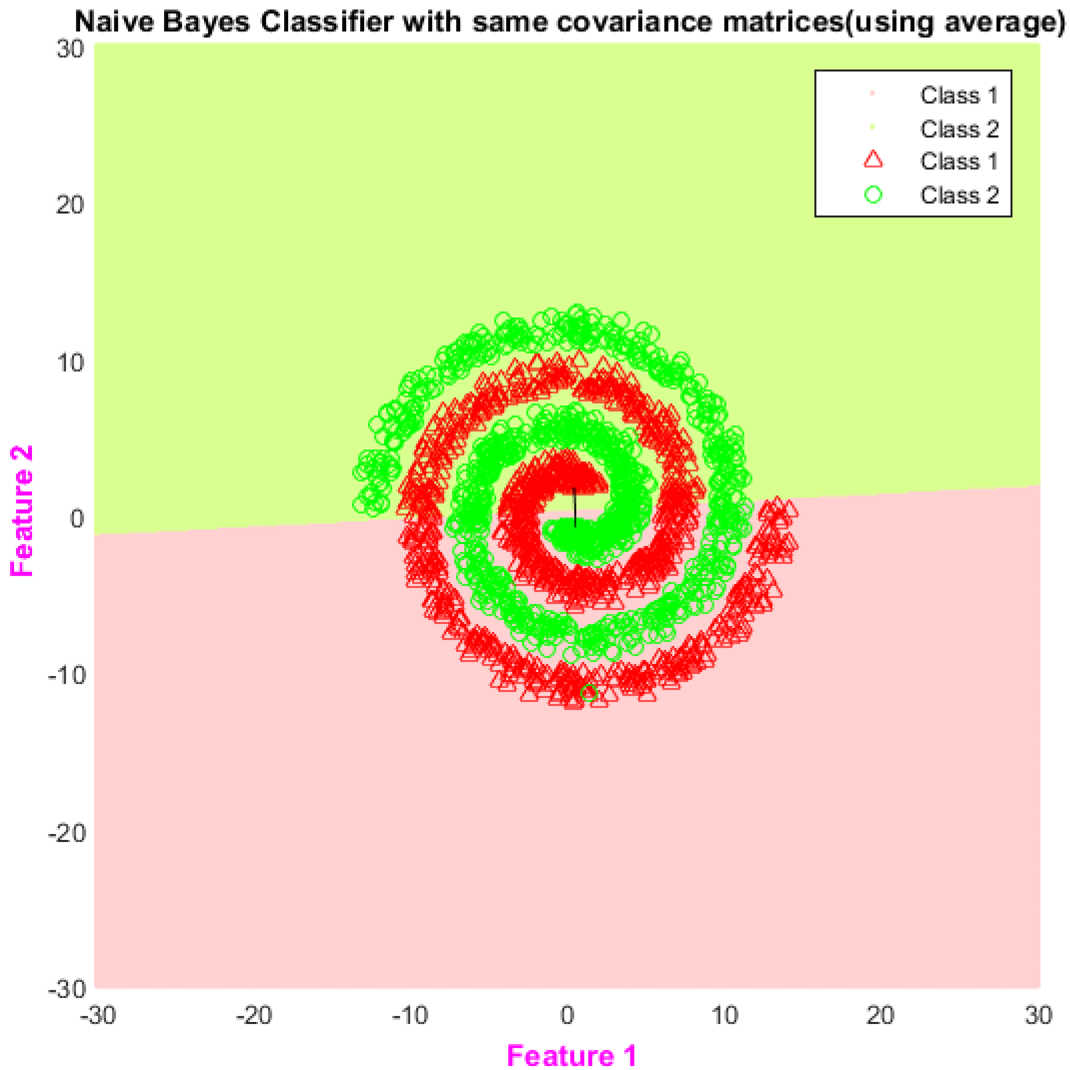
Predicted Class ⇨	CLASS 1	CLASS 2
Actual Class ⇩		
Class 1	175	151
Class 2	151	175

- *Classification accuracy on test data –*
  Overall Accuracy – 53.6810
  Classifier Accuracy for class 1 – 53.6810
  Classifier Accuracy for class 2 – 53.6810

2) Naive Bayes Classifier –

a) Same covariance matrices (σ^2 I)

- *Decision region plot for all the classes together with the training data superposed -*



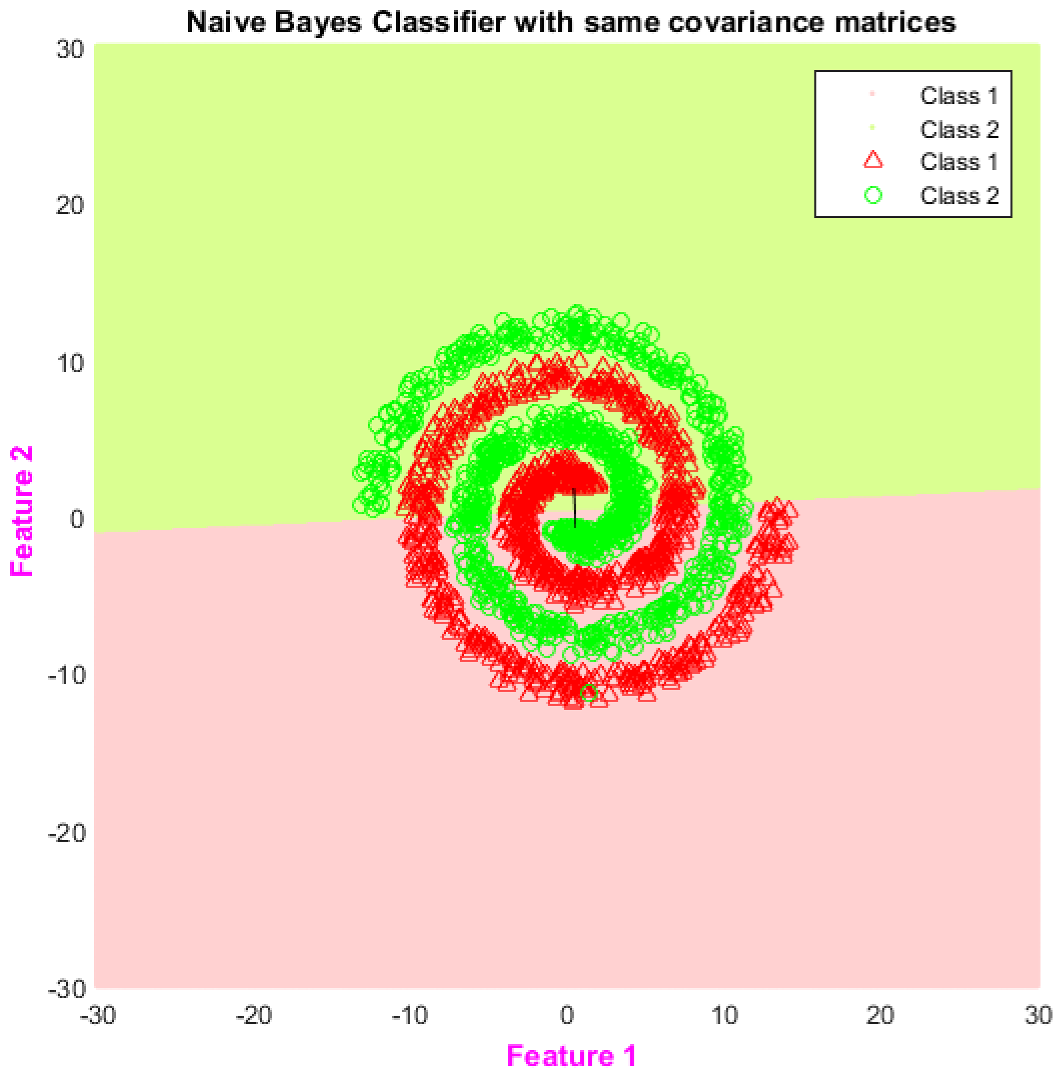
- *<u>Confusion Matrix based on performance for test data-</u>*

Predicted Class ⇨	CLASS 1	CLASS 2
Actual Class ⇩		
Class 1	176	150
Class 2	149	177

- *<u>Classification accuracy on test data –</u>*
  Overall Accuracy – 54.1411
  Classifier Accuracy for class 1 – 53.9877
  Classifier Accuracy for class 2 – 54.2945

b) Same covariance matrices and is C

- *<u>Decision region plot for all the classes together with the training data superposed -</u>*



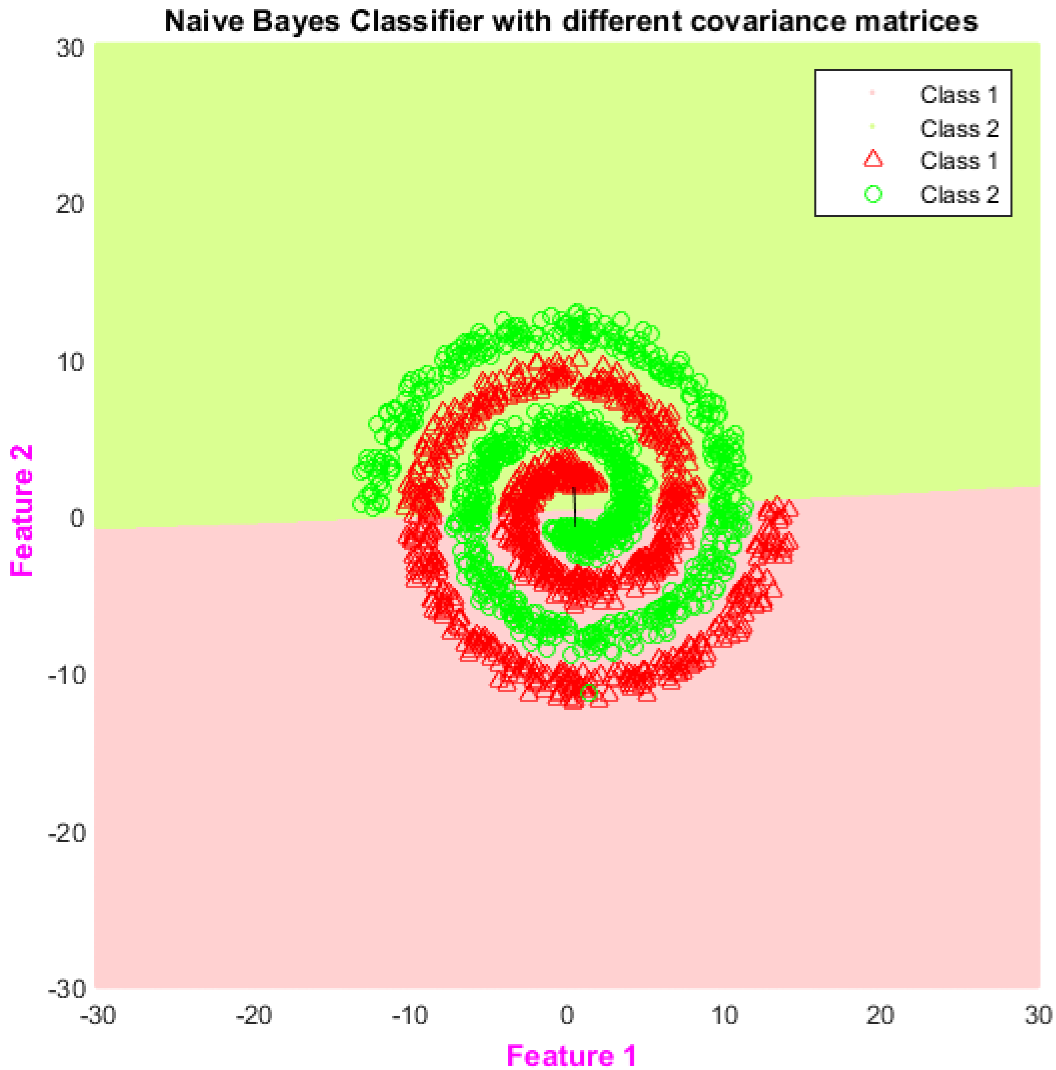
- *Confusion Matrix based on performance for test data-*

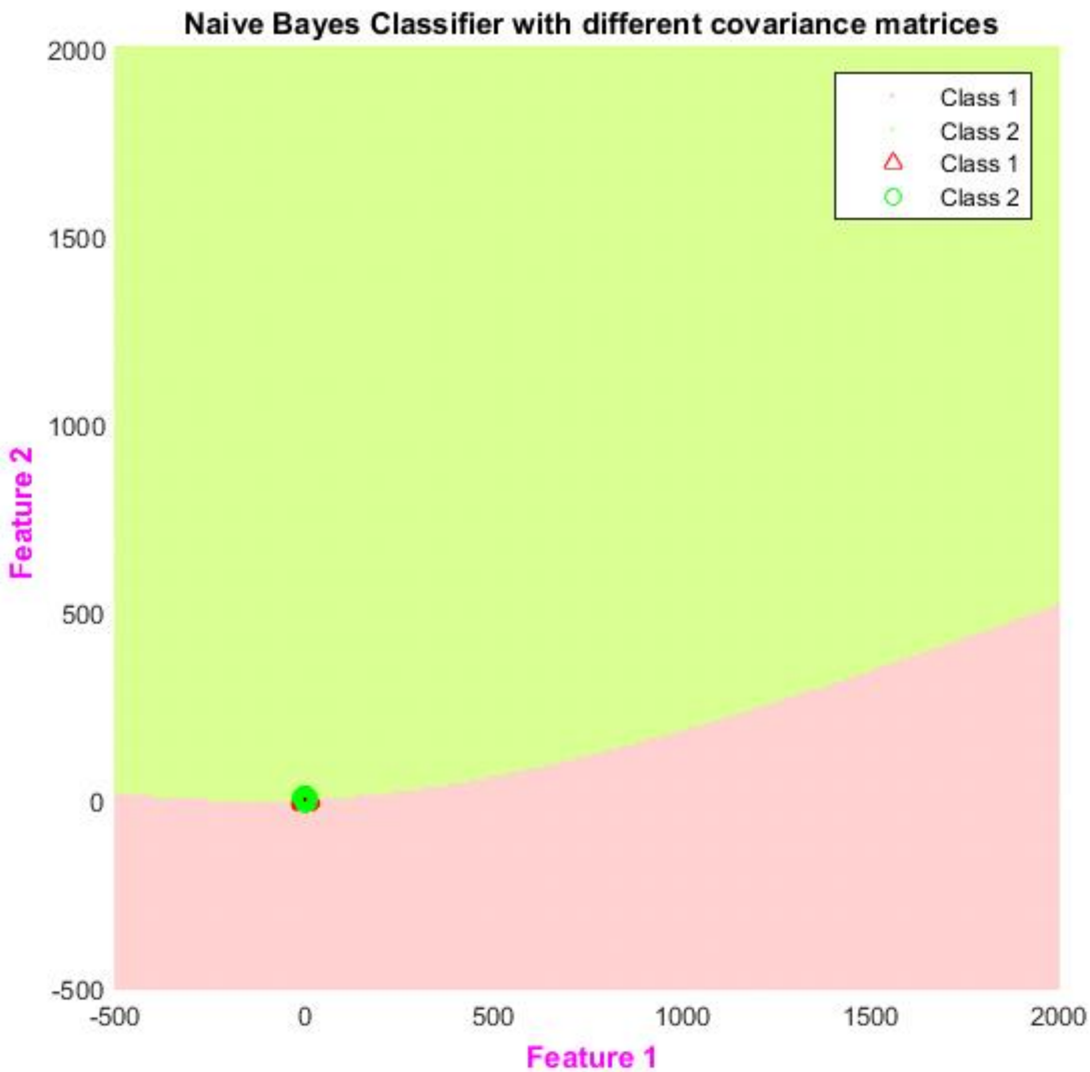
Predicted Class ⇨	CLASS 1	CLASS 2
Actual Class ⇩		
Class 1	177	149
Class 2	149	177

- *Classification accuracy on test data –*
  Overall Accuracy – 54.2945
  Classifier Accuracy for class 1 – 54.2945
  Classifier Accuracy for class 2 – 54.2945

c) Covariance matrix for each class is different

- *Decision region plot for all the classes together with the training data superposed -*




Naive Bayes Classifier with different covariance matrices
Class 1
Class 2
Class 1
Class 2
Feature 1
Feature 2
2000
1500
1000
500
0
-500
-500
0
500
1000
1500
2000

- *Confusion Matrix based on performance for test data-*

Predicted Class ⇨	CLASS 1	CLASS 2
Actual Class ⇩		
Class 1	177	149
Class 2	149	177

- *Classification accuracy on test data –*
  Overall Accuracy – 54.2945
  Classifier Accuracy for class 1 – 54.2945
  Classifier Accuracy for class 2 – 54.2945

♦ Observations –

1) Bayes Classifier –

a) Same covariance matrices (by taking average)
The decision boundaries are all of **linear** nature. They pass through the **midpoint** of line joining the means of two classes. The boundaries are **not necessarily** perpendicular to this line.

b) Same covariance matrices (from training data of all classes combined)
The decision boundaries are all of **linear** nature. They pass through the **midpoint** of line joining the means of two classes. The boundaries are **not necessarily** perpendicular to this line.

c) Covariance matrix for each class is different
The decision boundaries are all of **Quadratic** nature. For this dataset they turn out to be **Parabola**. They do **not necessarily** pass through the **midpoint** of line joining the means of two classes.

2) Naïve Bayes Classifier –

a) Same covariance matrices (σ^2 I)
The decision boundaries are all of **linear** nature. They pass through the **midpoint** of line joining the means of two classes. The boundaries are perpendicular to this line.

b) Same covariance matrices and is C
The decision boundaries are all of **linear** nature. They pass through the **midpoint** of line joining the means of two classes. The boundaries are **not necessarily** perpendicular to this line.

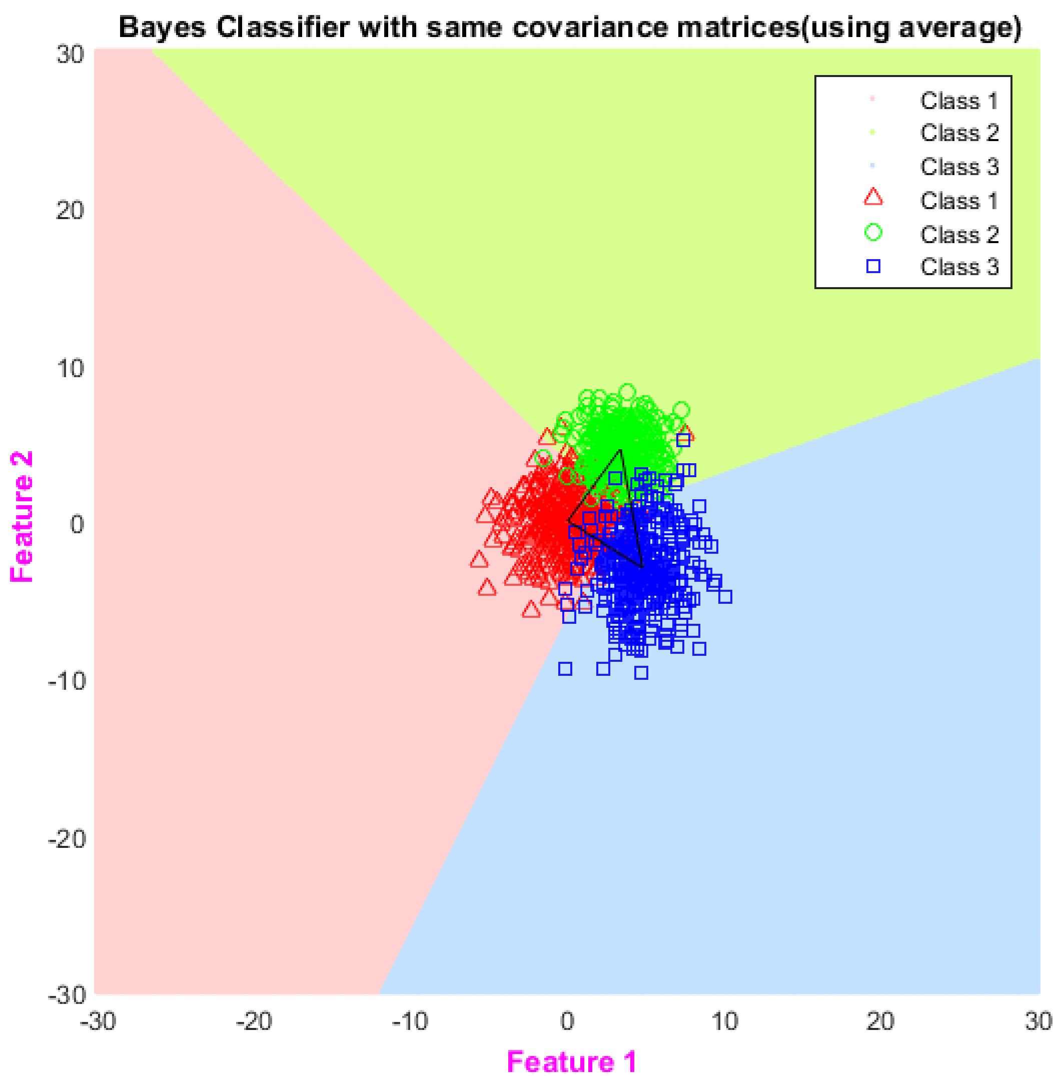
c) Covariance matrix for each class is different
The decision boundaries are all of **Quadratic** nature. For this dataset they turn out to be **Parabola**. They do **not necessarily** pass through the **midpoint** of line joining the means of two classes. The classification results are more or less similar to that of Bayes.

♦ **Dataset I-(c)**: 2-dimensional artificial data of 3 or 4 classes that are **overlapping**

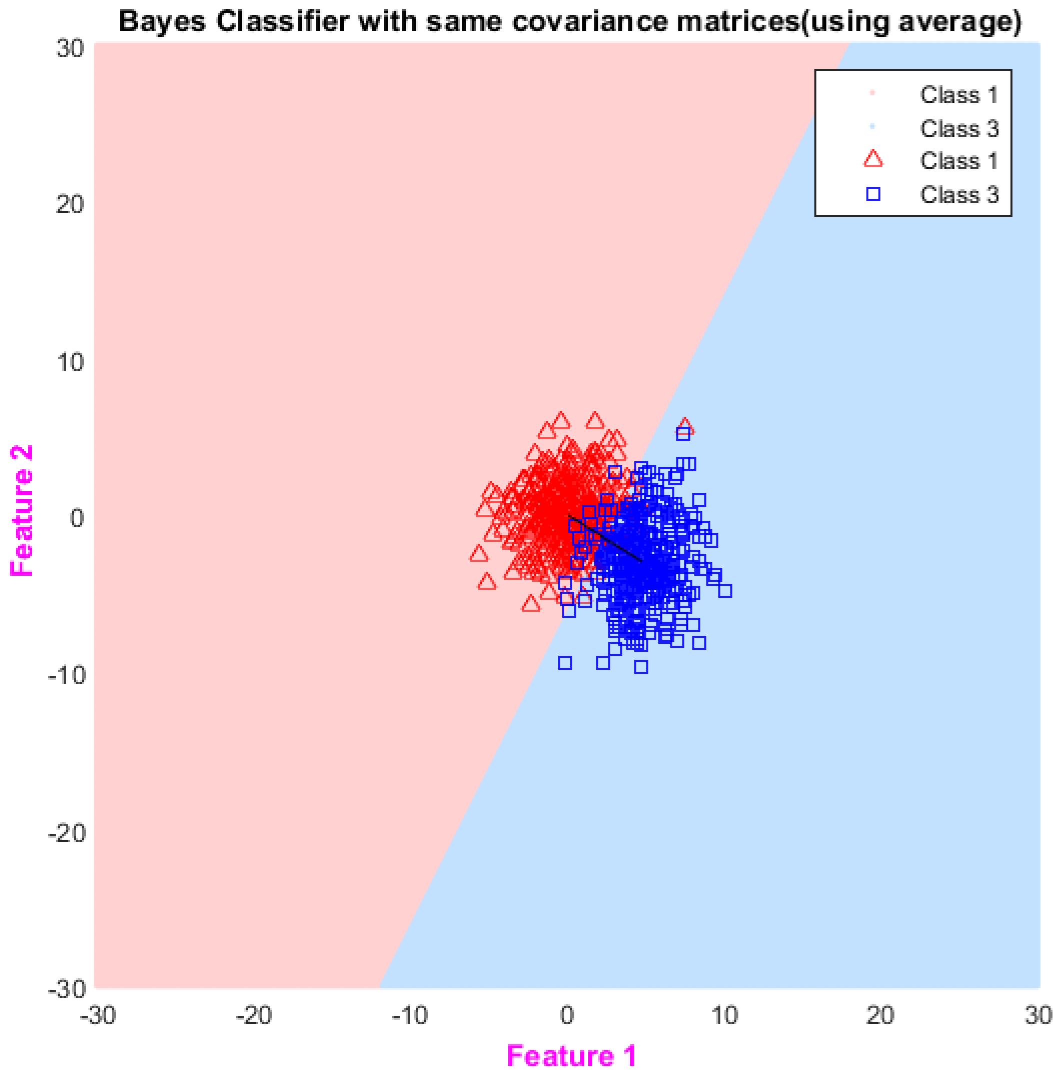
1) Bayes Classifier –

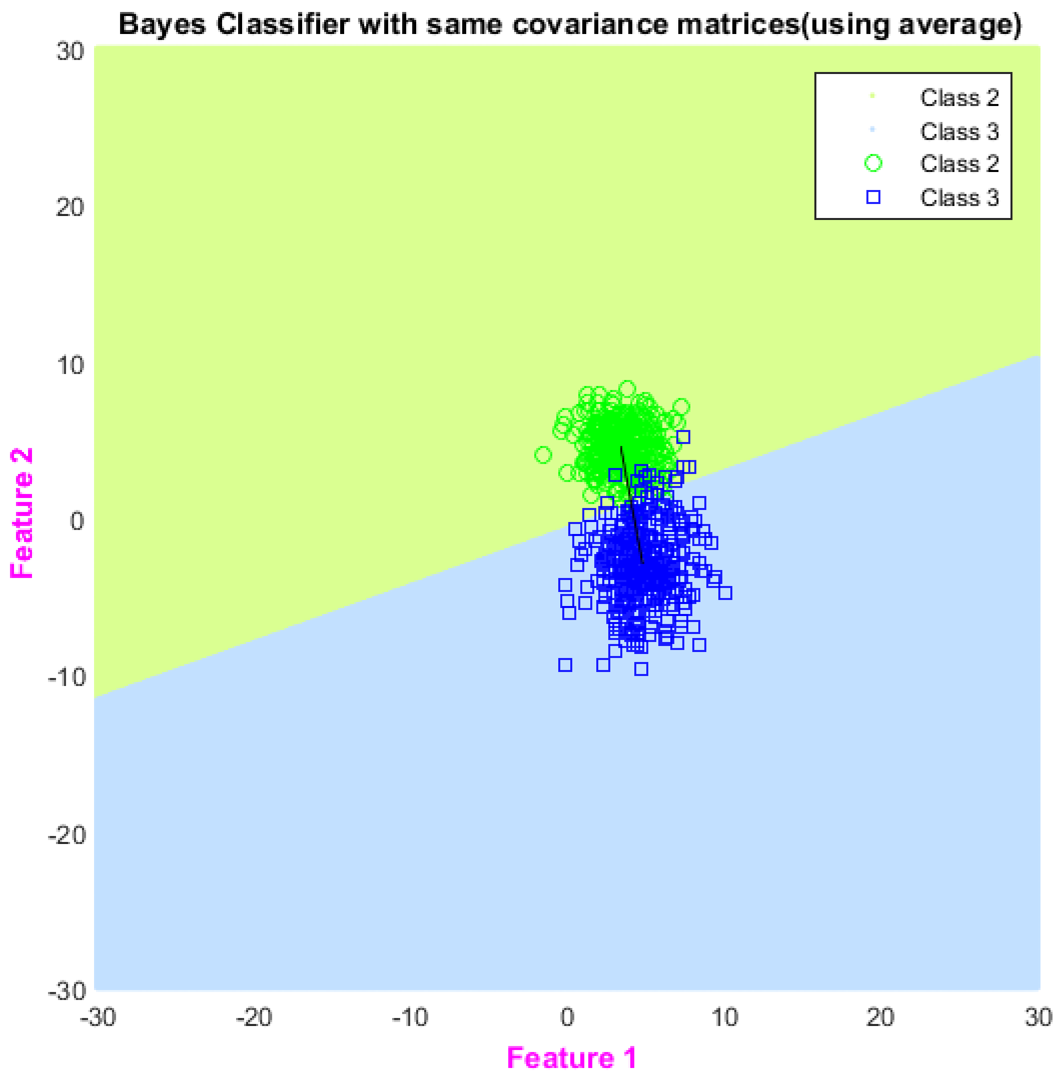
a) Same covariance matrices (by taking average)

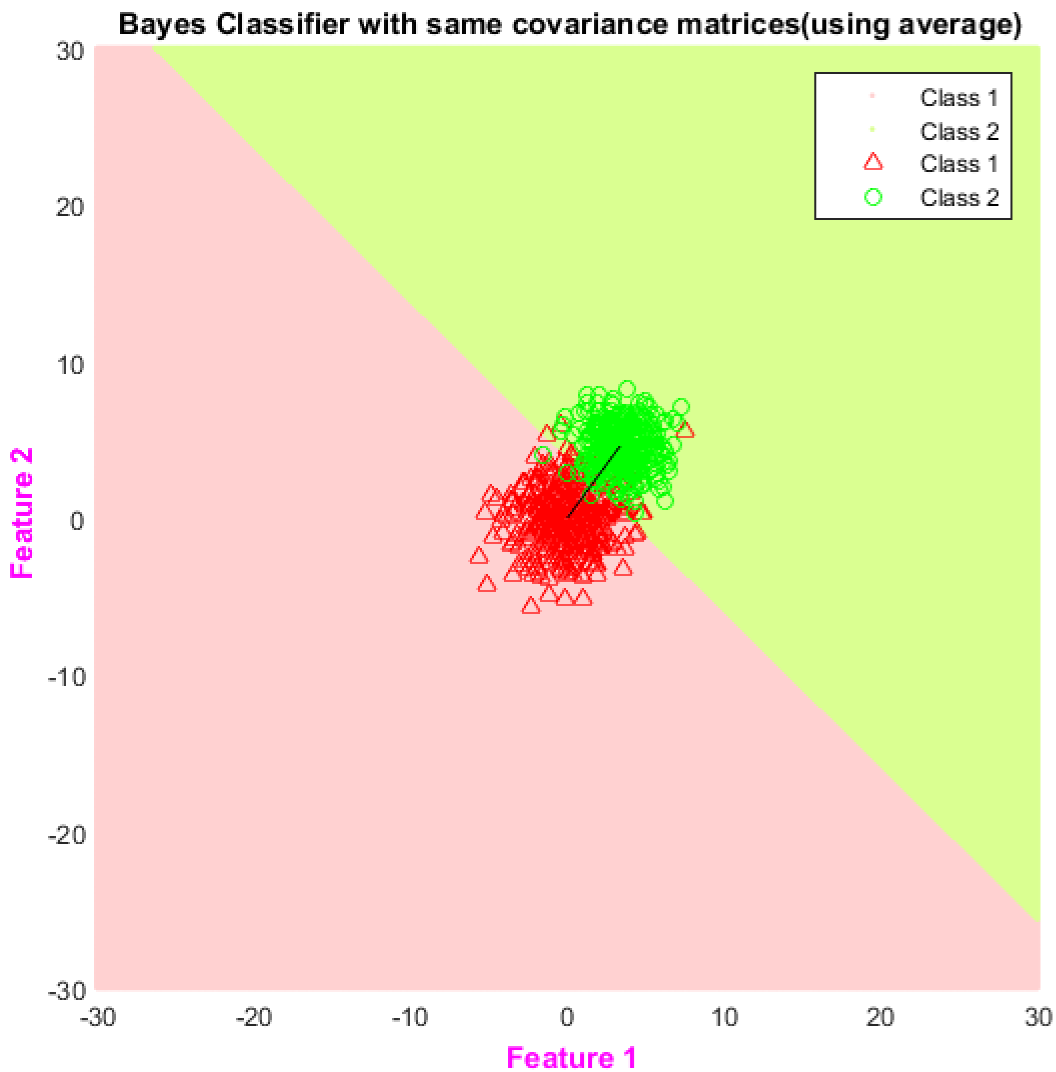
- *Decision region plot for all the classes together with the training data superposed -*



- *Decision region plot for every pair of classes -*




Bayes Classifier with same covariance matrices(using average)
Class 2
Class 3
Class 2
Class 3
30
20
10
0
-10
-20
-30
-30
-20
-10
0
10
20
30
Feature 1
Feature 2



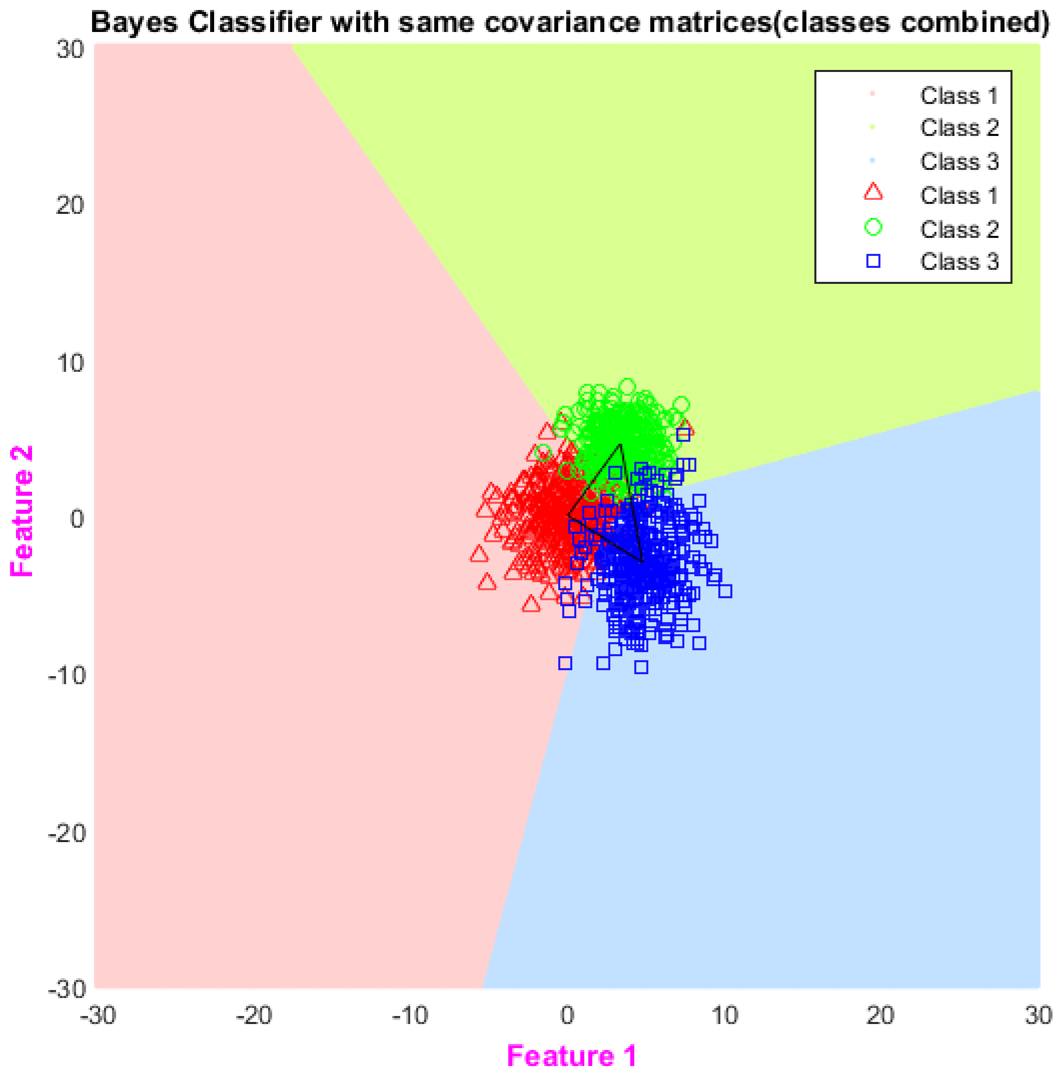
- *Confusion Matrix based on performance for test data-*

Predicted Class ⇨	CLASS 1	CLASS 2	CLASS 3
Actual Class ⇩			
Class 1	109	13	3
Class 2	3	121	1
Class 3	8	5	112

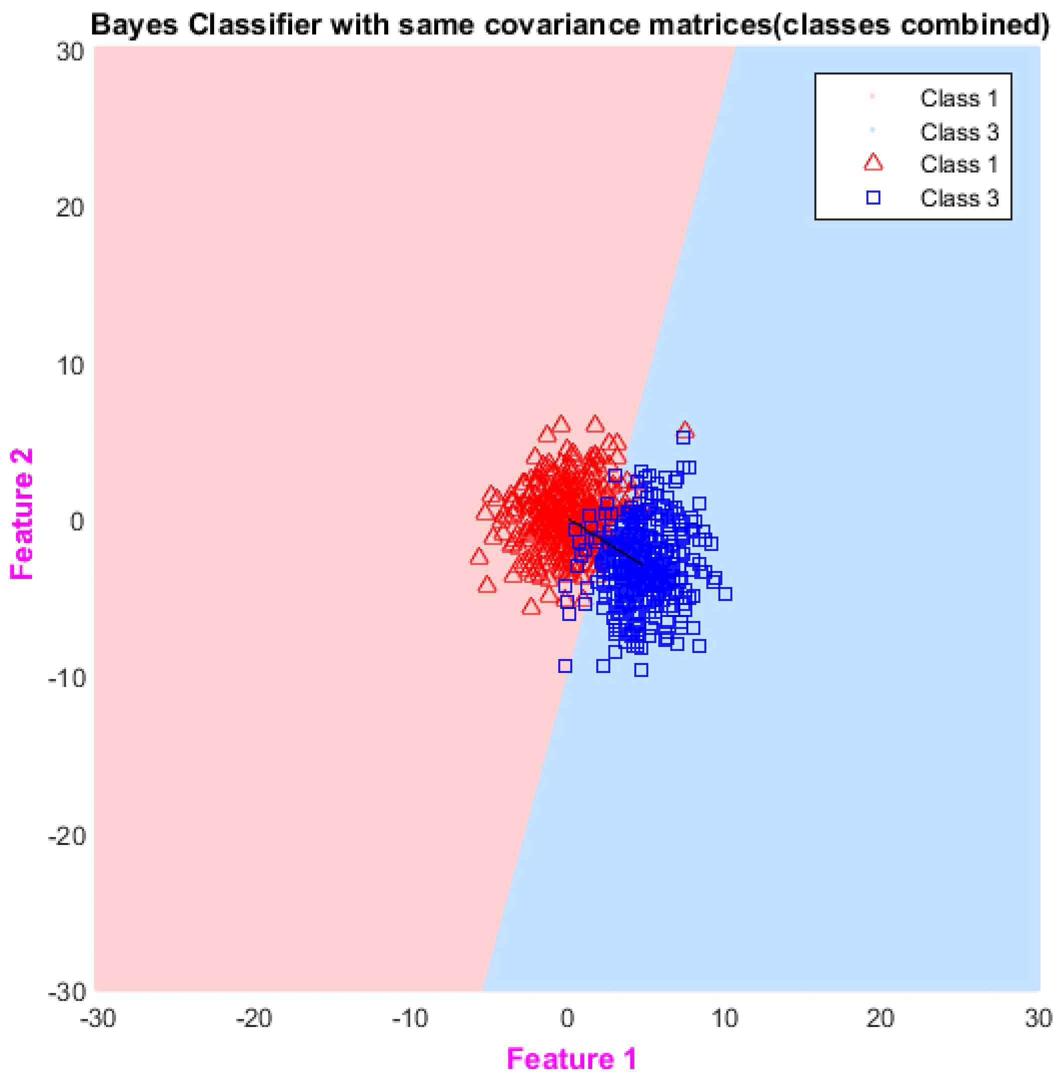
- *Classification accuracy on test data –*
  Overall Accuracy – 91.2
  Classifier Accuracy for class 1 – 87.2
  Classifier Accuracy for class 2 – 96.8
  Classifier Accuracy for class 3 – 89.6

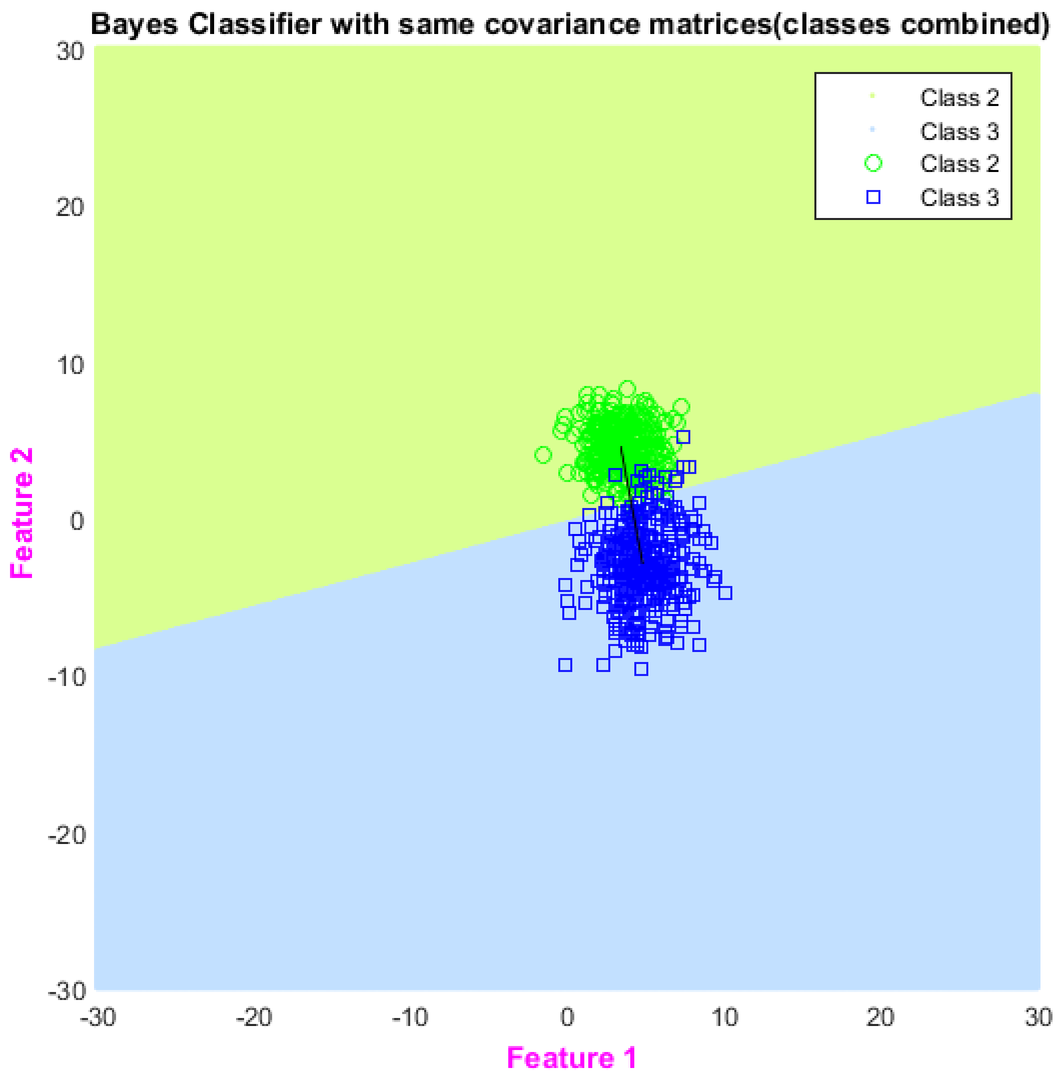
b) Same covariance matrices (from training data of all classes combined)

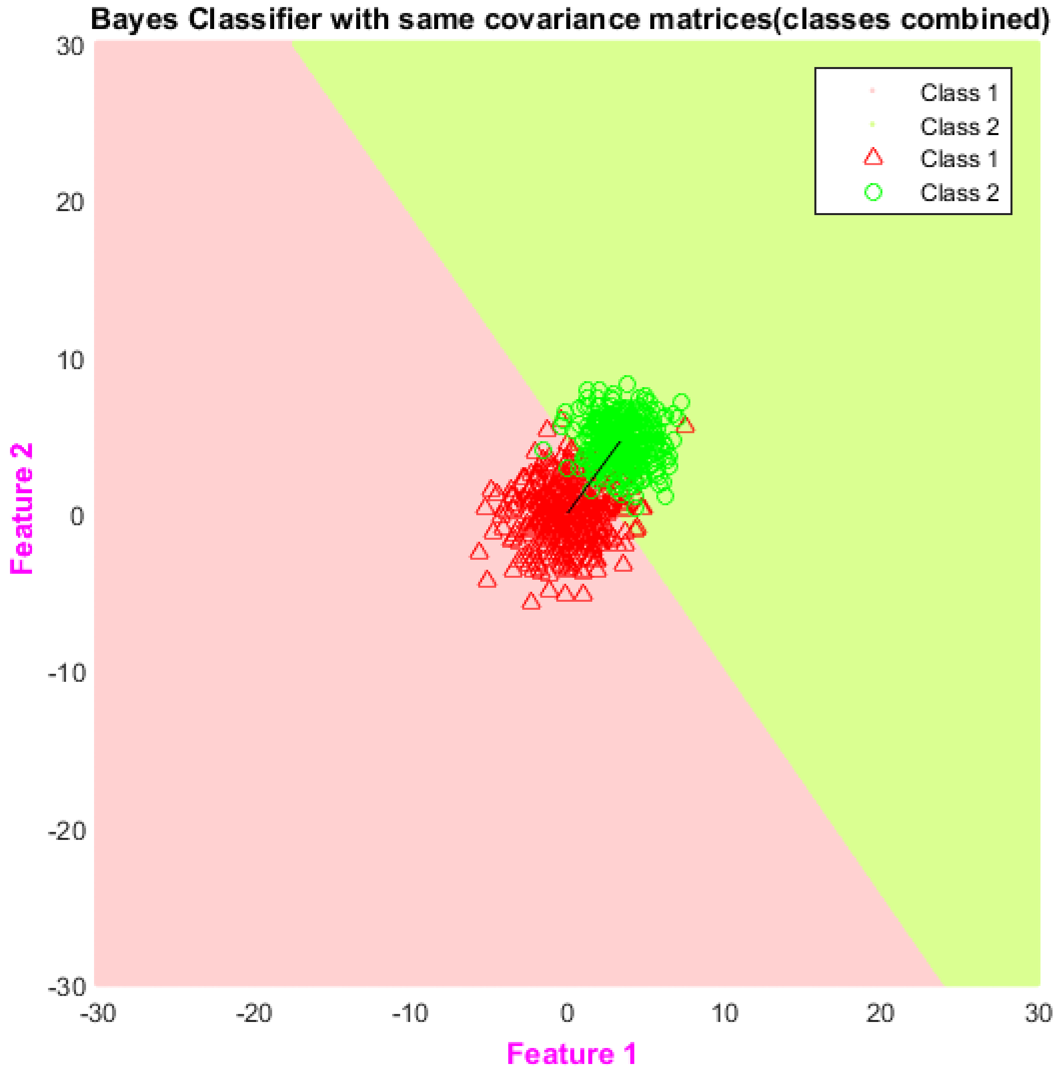
- *Decision region plot for all the classes together with the training data superposed -*



- *Decision region plot for every pair of classes -*




Bayes Classifier with same covariance matrices(classes combined)
Class 2
Class 3
Class 2
Class 3
Feature 2
Feature 1
30
20
10
0
-10
-20
-30
-30
-20
-10
0
10
20
30



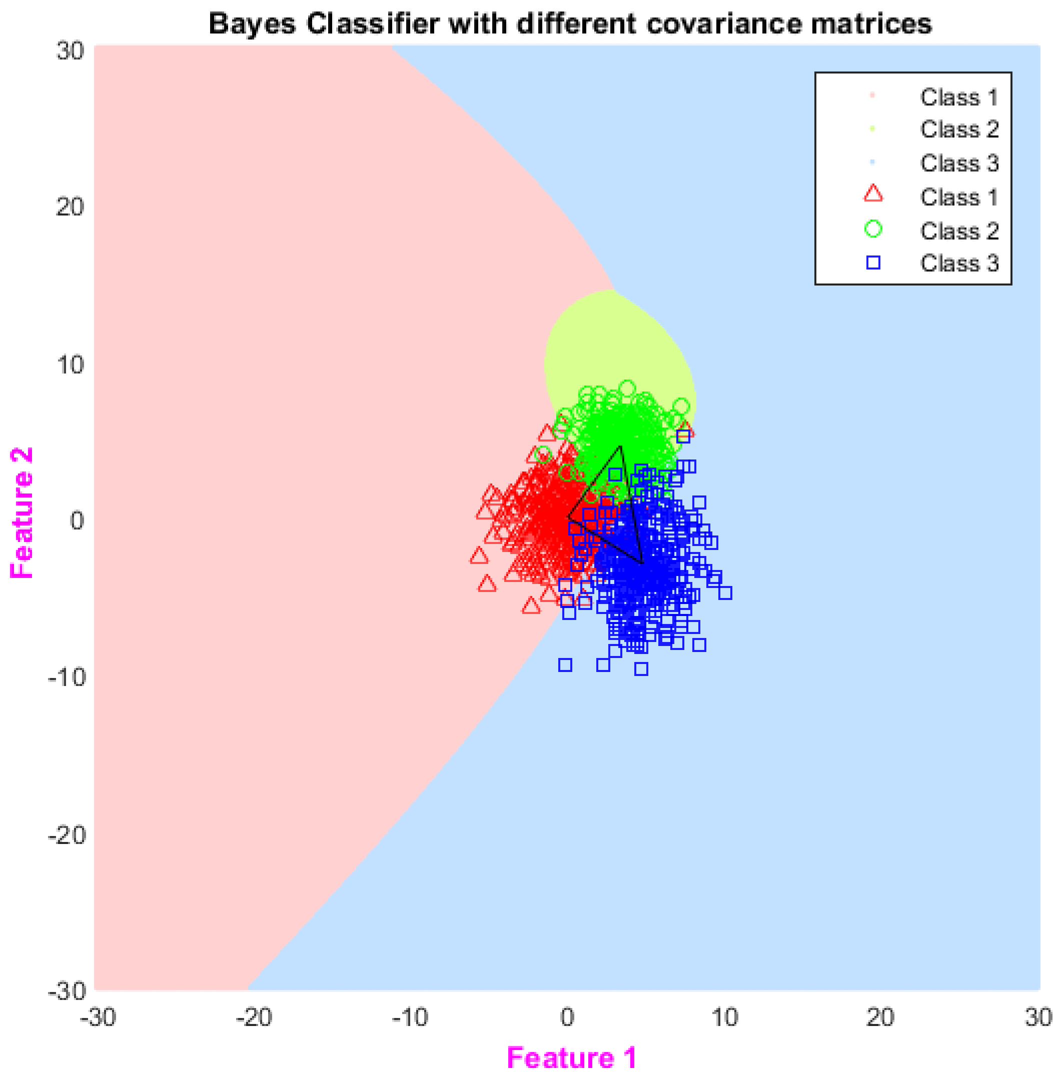
- *<u>Confusion Matrix based on performance for test data-</u>*

Predicted Class ⇨	CLASS 1	CLASS 2	CLASS 3
Actual Class ⇩			
Class 1	106	14	5
Class 2	3	121	1
Class 3	6	7	112

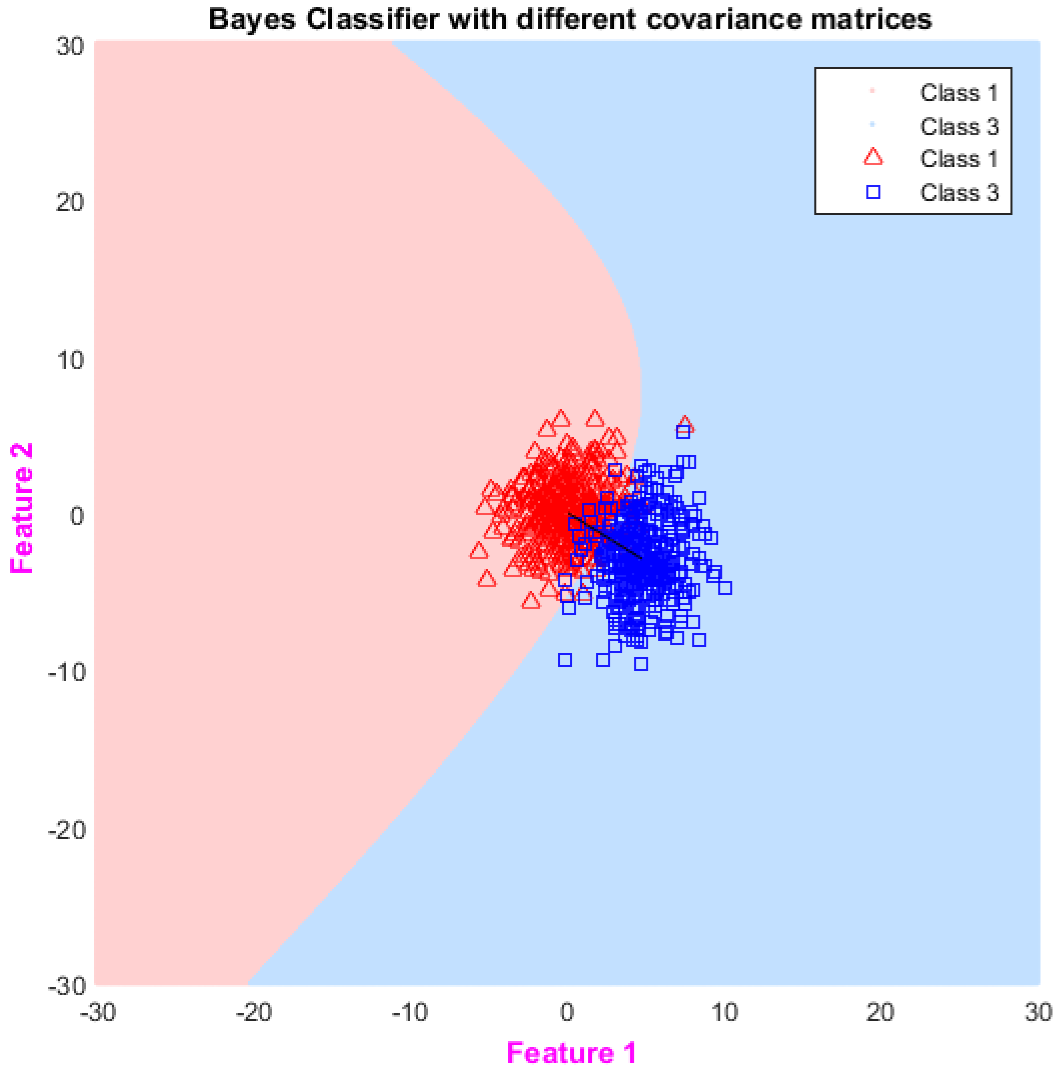
- *<u>Classification accuracy on test data –</u>*
  Overall Accuracy – 90.4
  Classifier Accuracy for class 1 – 84.8
  Classifier Accuracy for class 2 – 96.8
  Classifier Accuracy for class 3 – 89.6

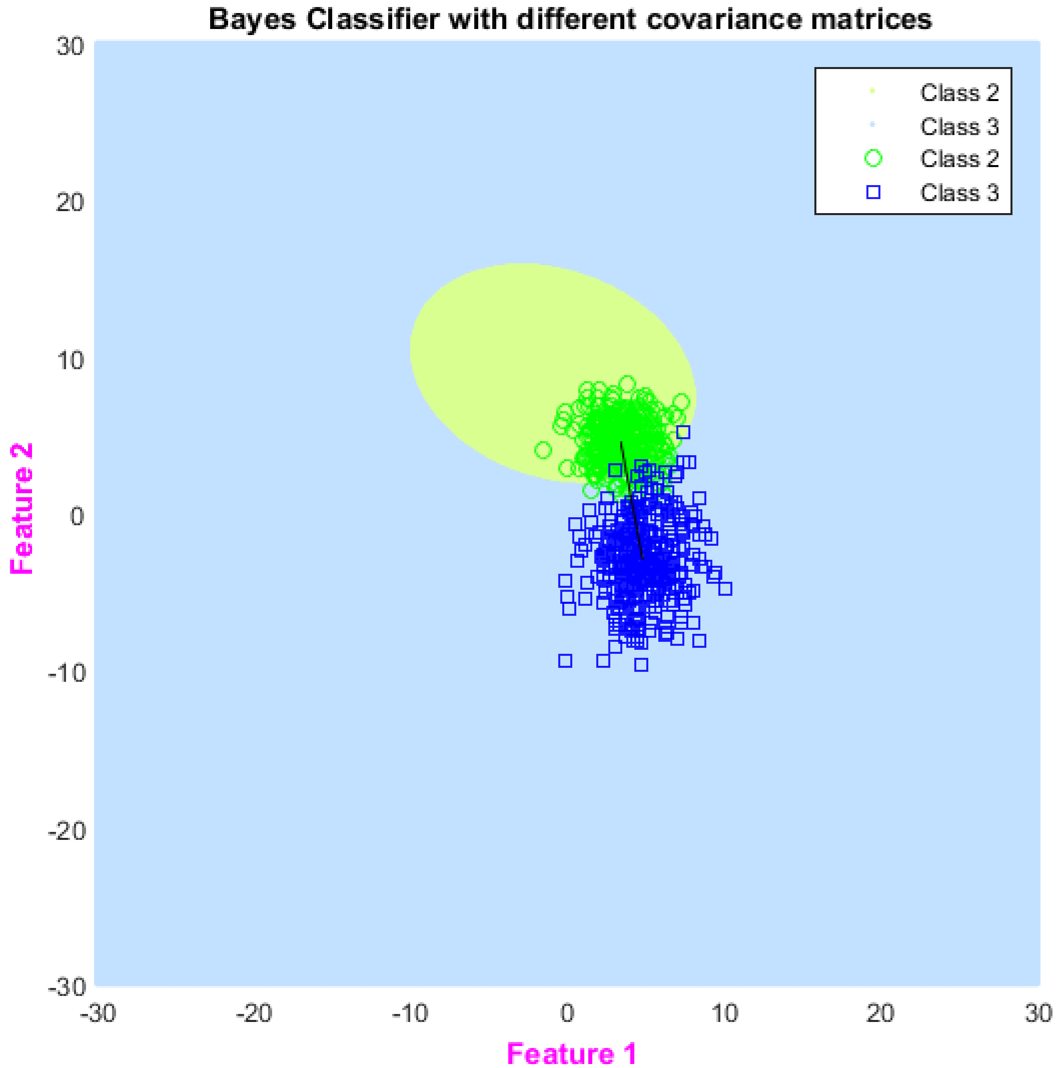
c) Covariance matrix for each class is different

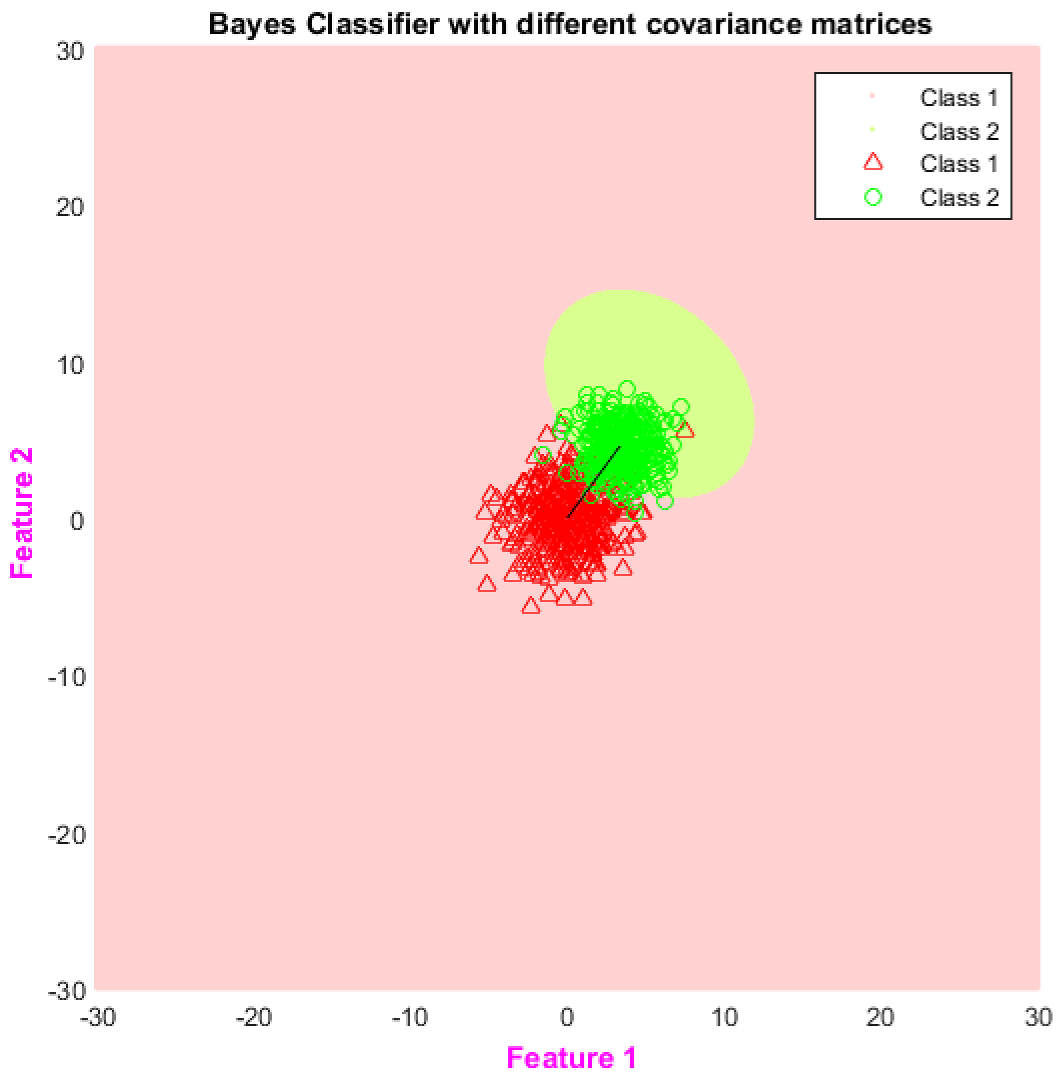
- *<u>Decision region plot for all the classes together with the training data superposed -</u>*



- *Decision region plot for every pair of classes -*


Bayes Classifier with different covariance matrices
Class 1
Class 3
Class 1
Class 3
Feature 1
Feature 2
30
20
10
0
-10
-20
-30
-30
-20
-10
0
10
20
30


Bayes Classifier with different covariance matrices
Class 2
Class 3
Class 2
Class 3
Feature 2
Feature 1
30
20
10
0
-10
-20
-30
-30
-20
-10
0
10
20
30



- *Confusion Matrix based on performance for test data-*

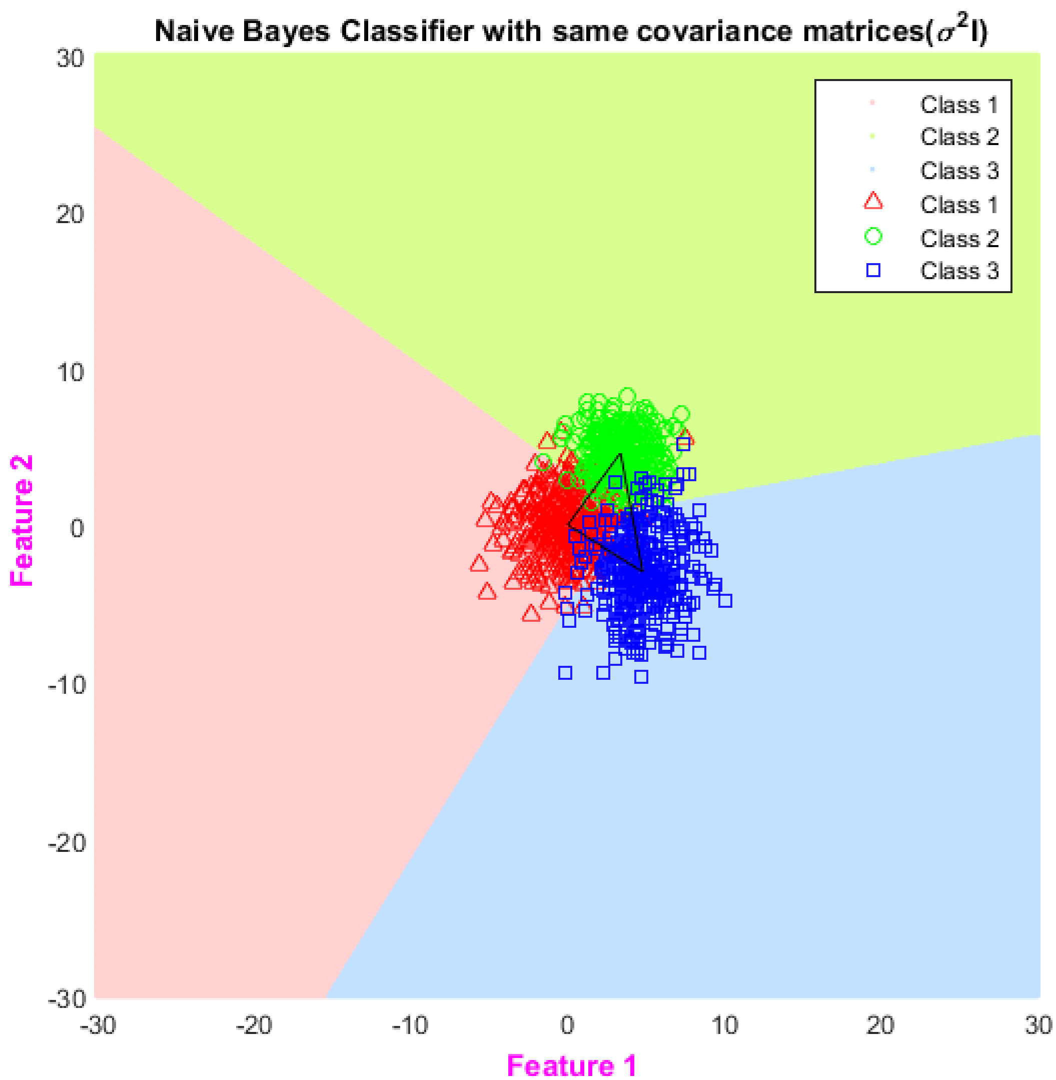
Predicted Class ⇨	CLASS 1	CLASS 2	CLASS 3
Actual Class ⇩			
Class 1	112	10	3
Class 2	3	121	1
Class 3	8	2	115

- *Classification accuracy on test data –*
  Overall Accuracy – 92.8
  Classifier Accuracy for class 1 – 89.6
  Classifier Accuracy for class 2 – 96.8
  Classifier Accuracy for class 3 – 92.0

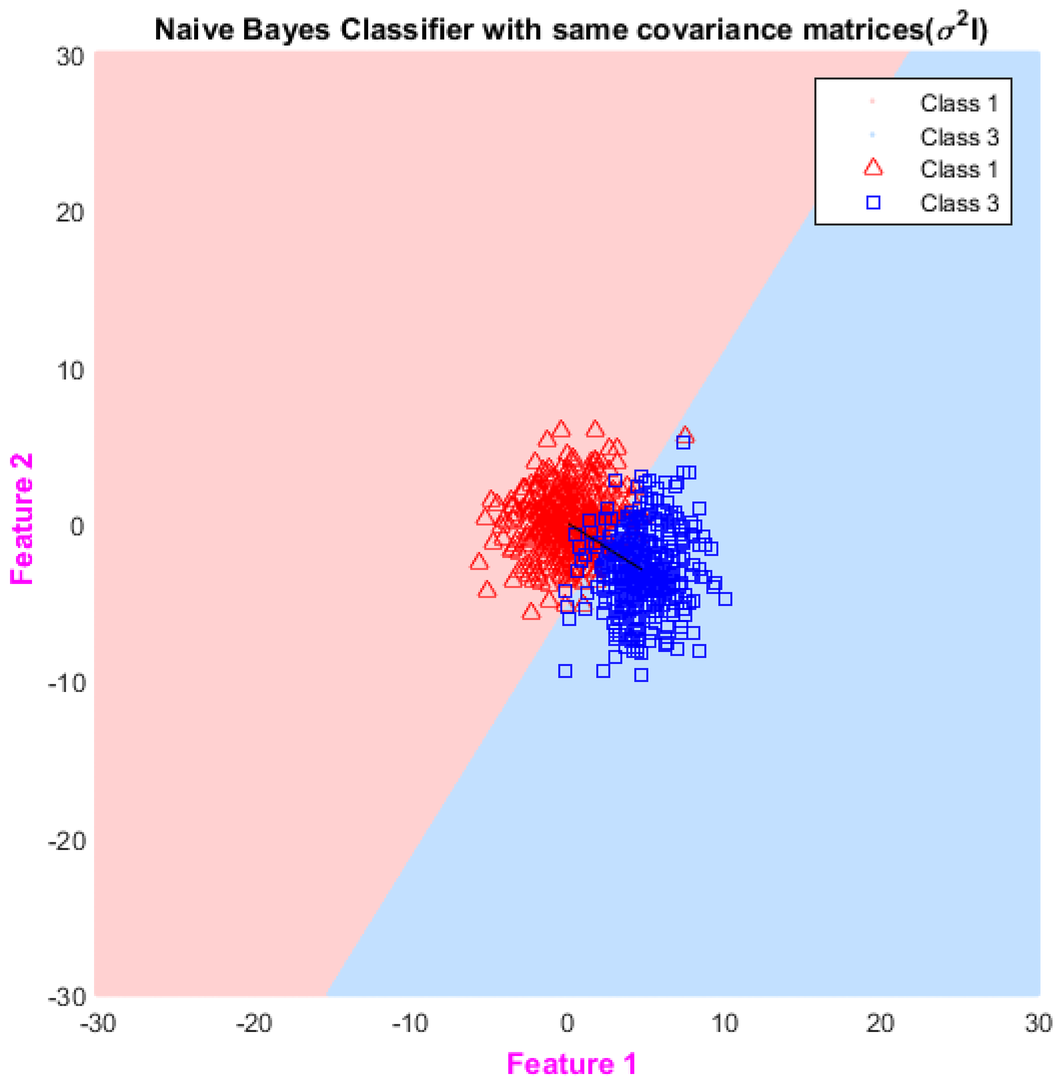
2) Naive Bayes Classifier –

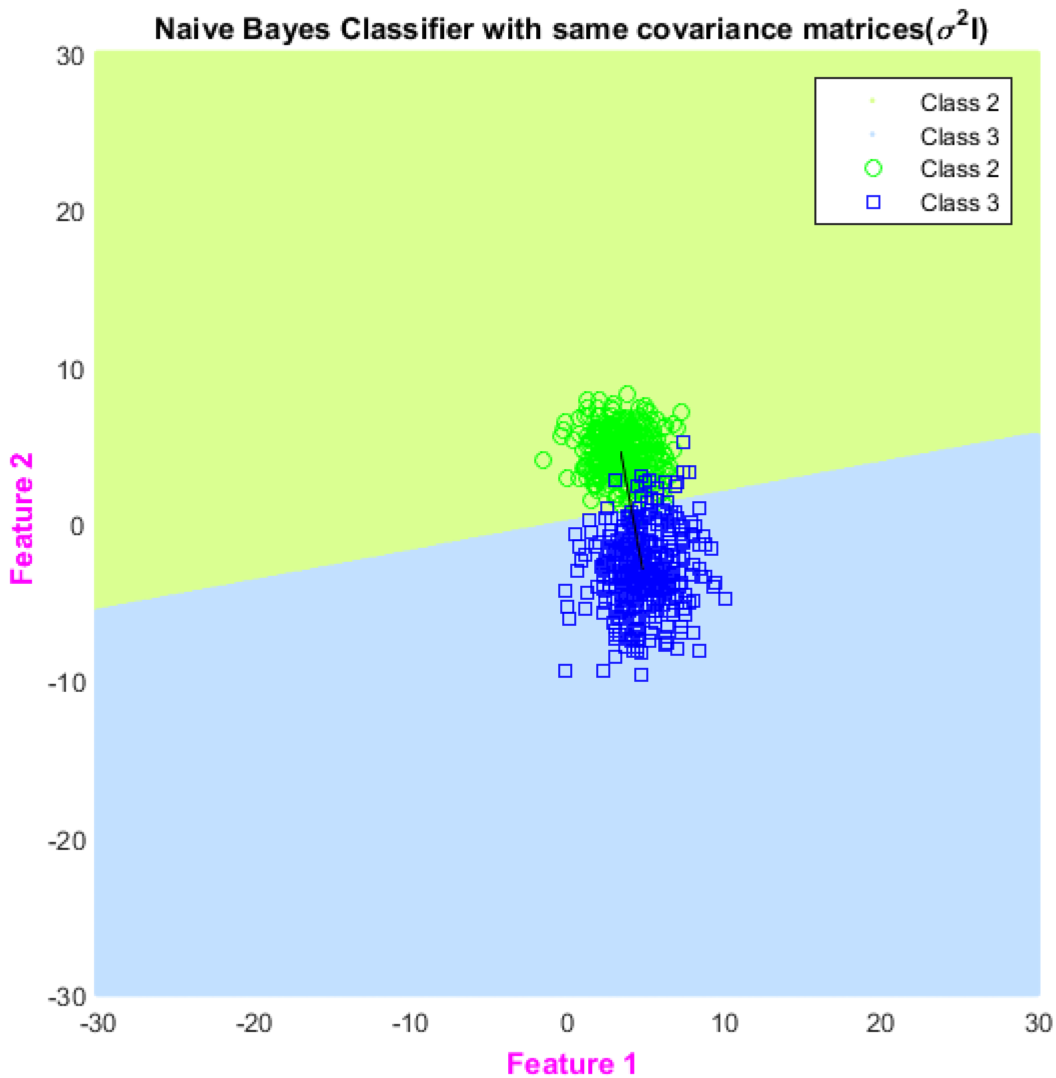
a) Same covariance matrices (σ^2 I)

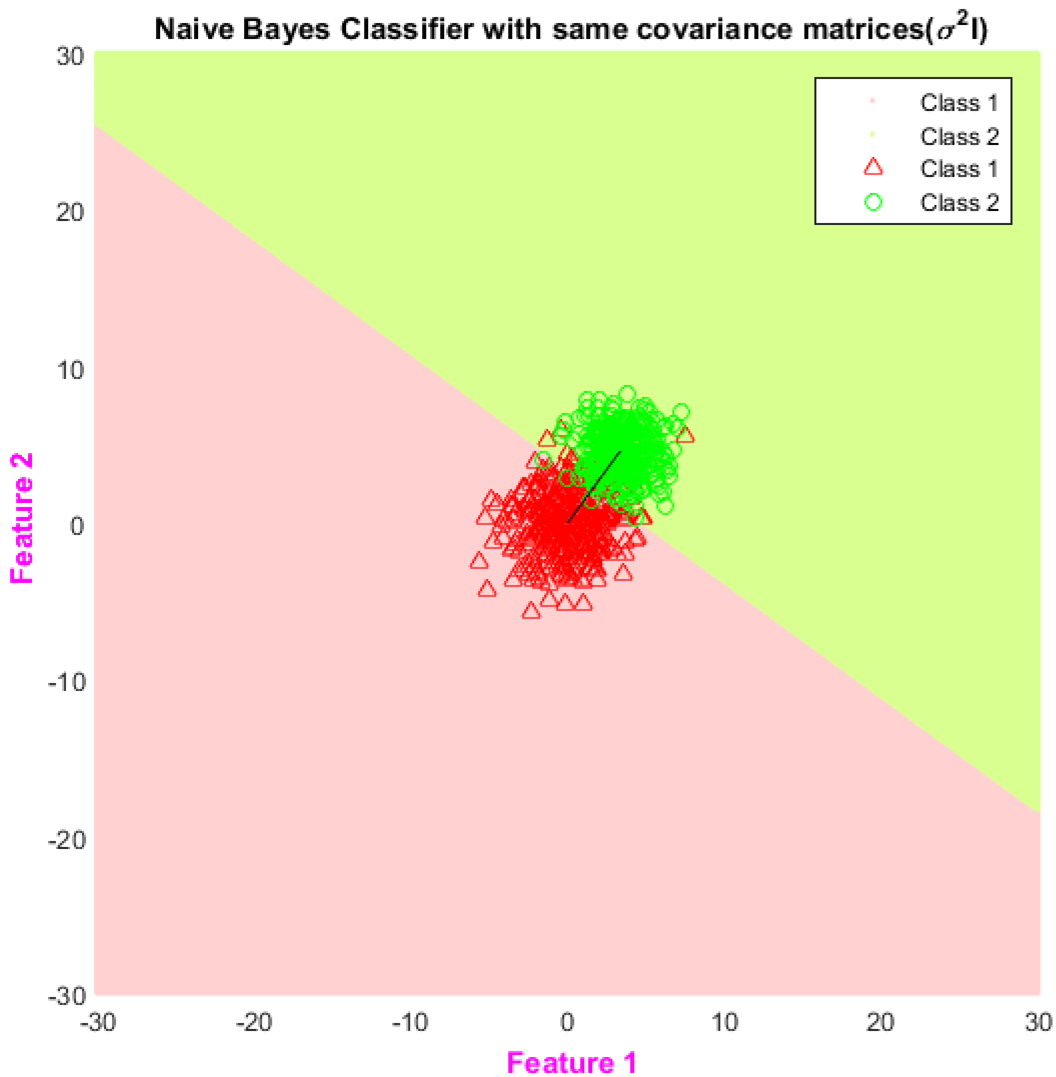
- *Decision region plot for all the classes together with the training data superposed -*



- *Decision region plot for every pair of classes -*




Naive Bayes Classifier with same covariance matrices(σ^2I)
Class 2
Class 3
Class 2
Class 3
Feature 2
Feature 1
30
20
10
0
-10
-20
-30
-30
-20
-10
0
10
20
30



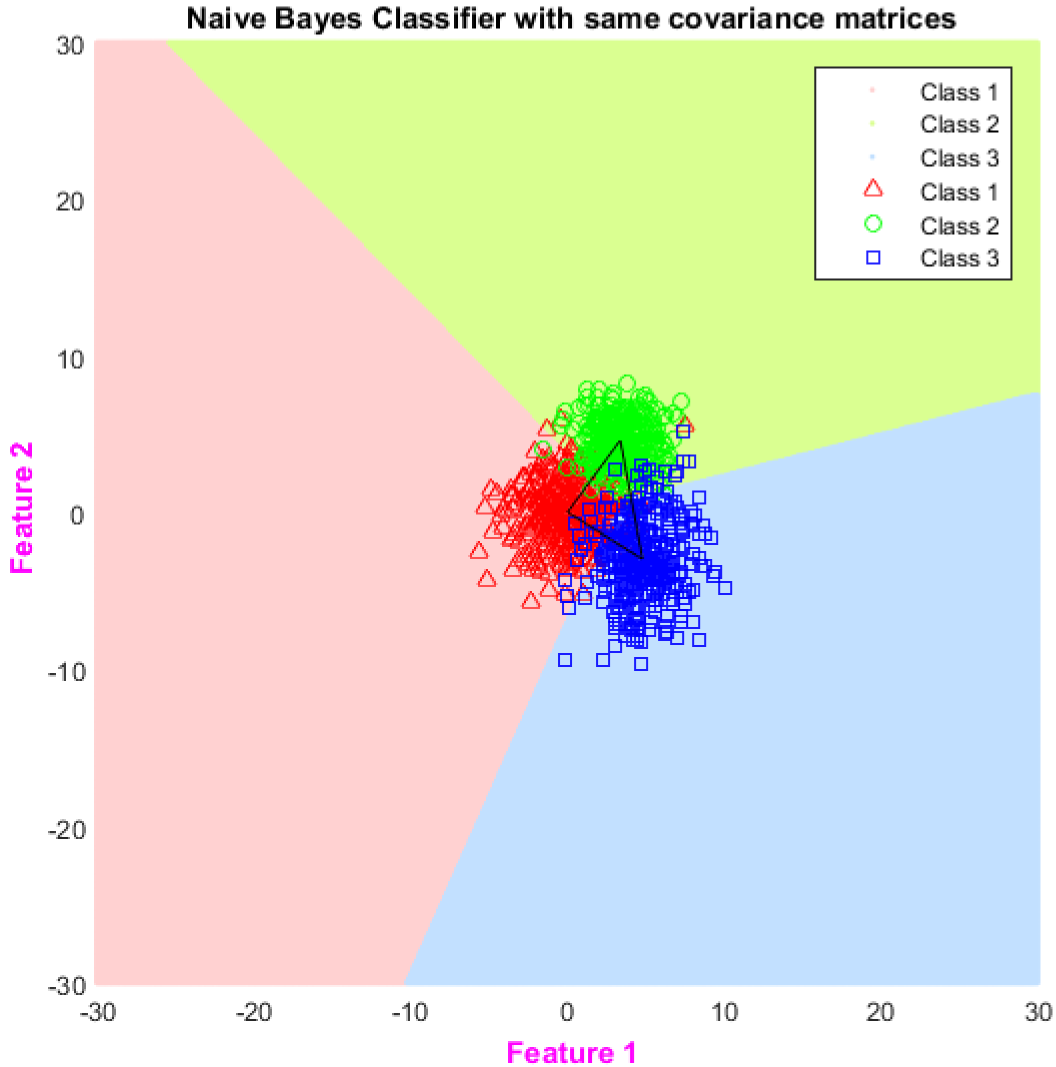
- *Confusion Matrix based on performance for test data-*

Predicted Class ⇨	CLASS 1	CLASS 2	CLASS 3
Actual Class ⇩			
Class 1	108	14	3
Class 2	3	121	1
Class 3	9	6	110

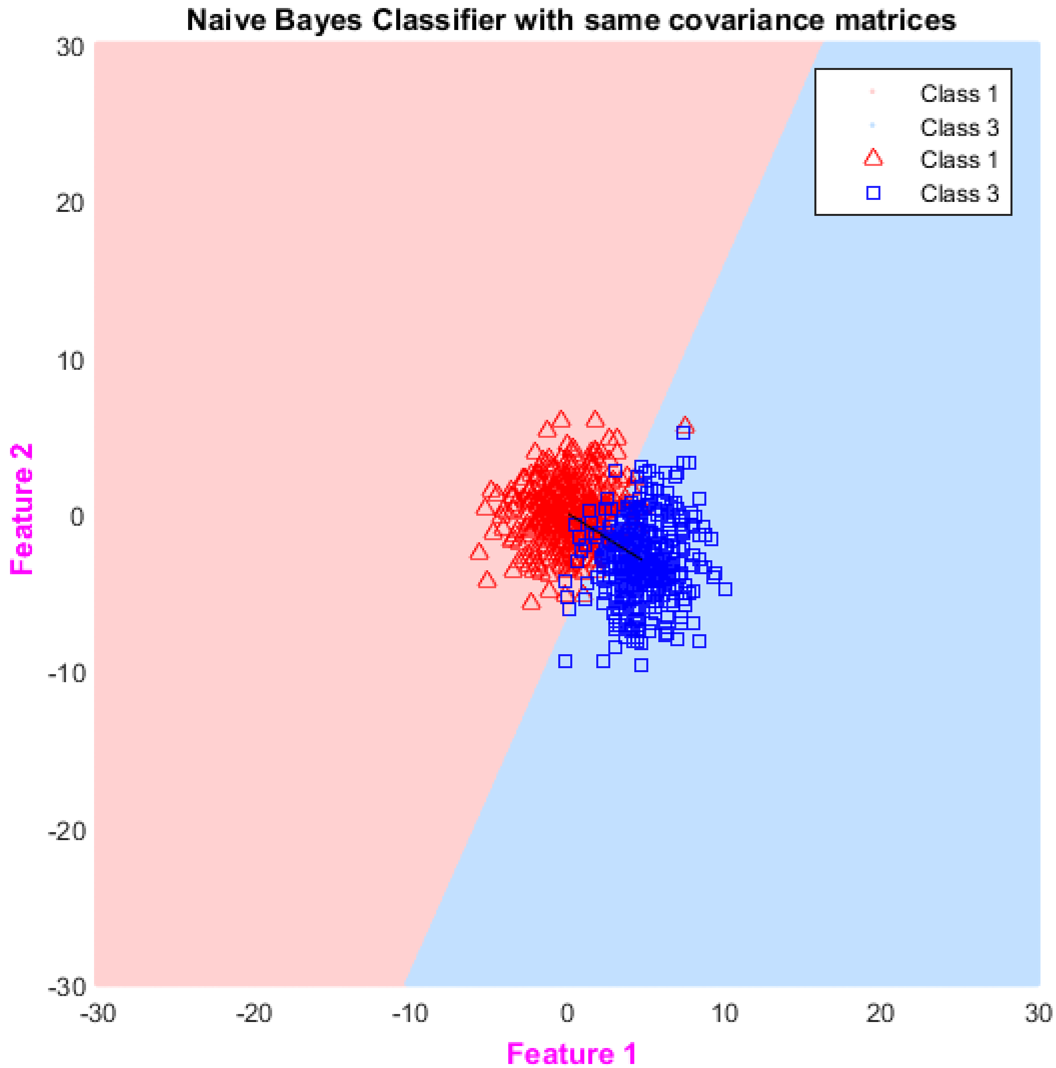
- *Classification accuracy on test data –*
  Overall Accuracy – 90.4
  Classifier Accuracy for class 1 – 86.4
  Classifier Accuracy for class 2 – 96.8
  Classifier Accuracy for class 3 – 88.0

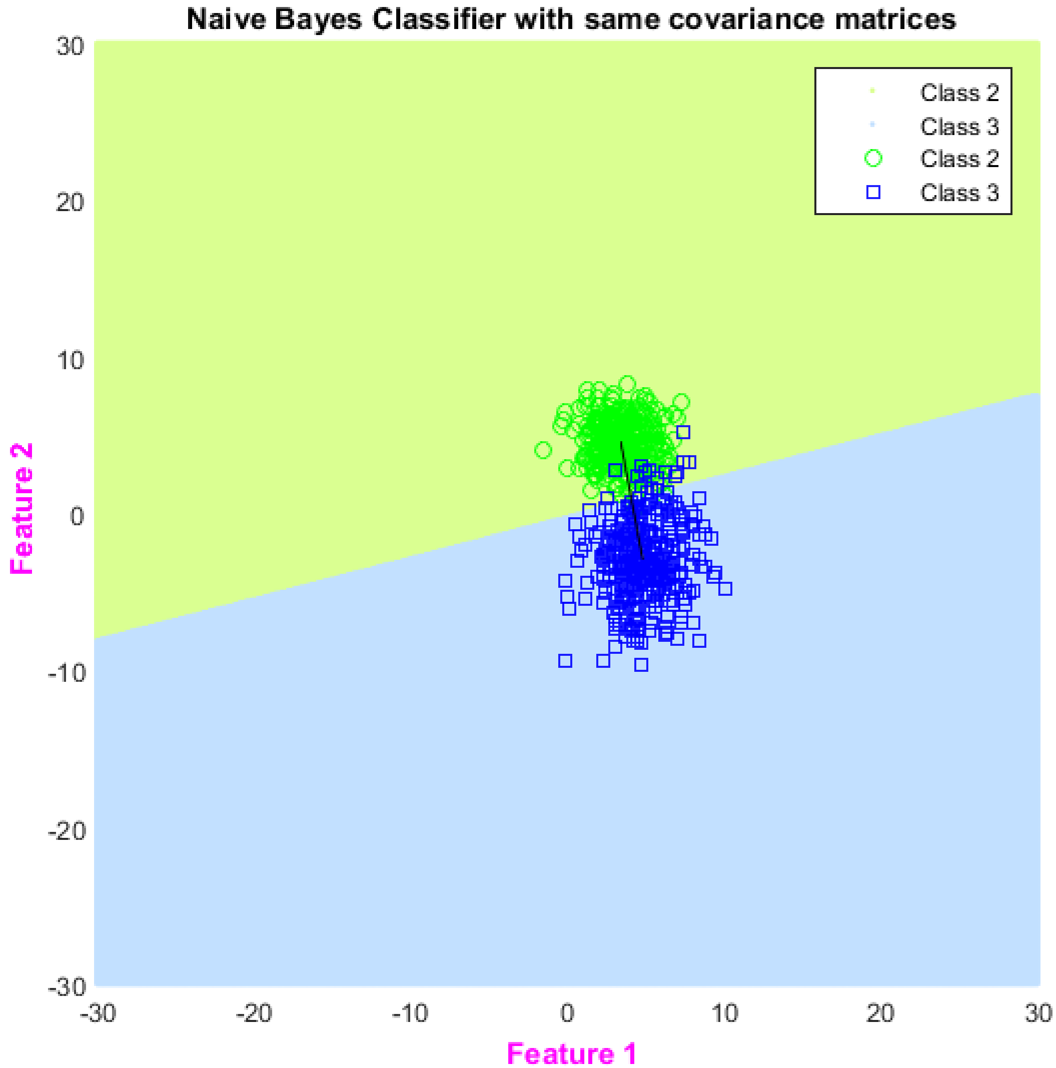
b) Same covariance matrices and is C

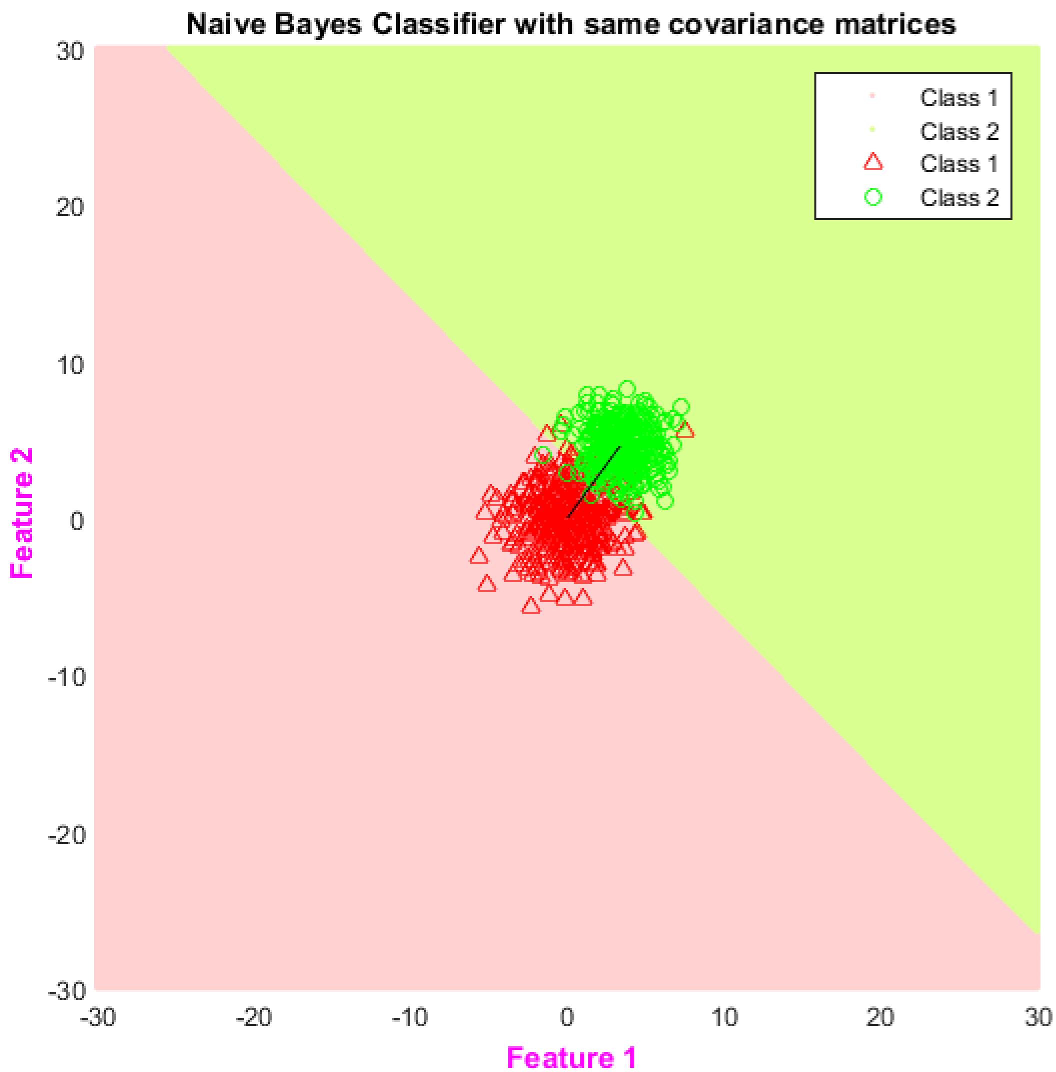
- *Decision region plot for all the classes together with the training data superposed -*



- *Decision region plot for every pair of classes -*




Naive Bayes Classifier with same covariance matrices
Class 2
Class 3
Class 2
Class 3
Feature 2
Feature 1
30
20
10
0
-10
-20
-30
-30
-20
-10
0
10
20
30



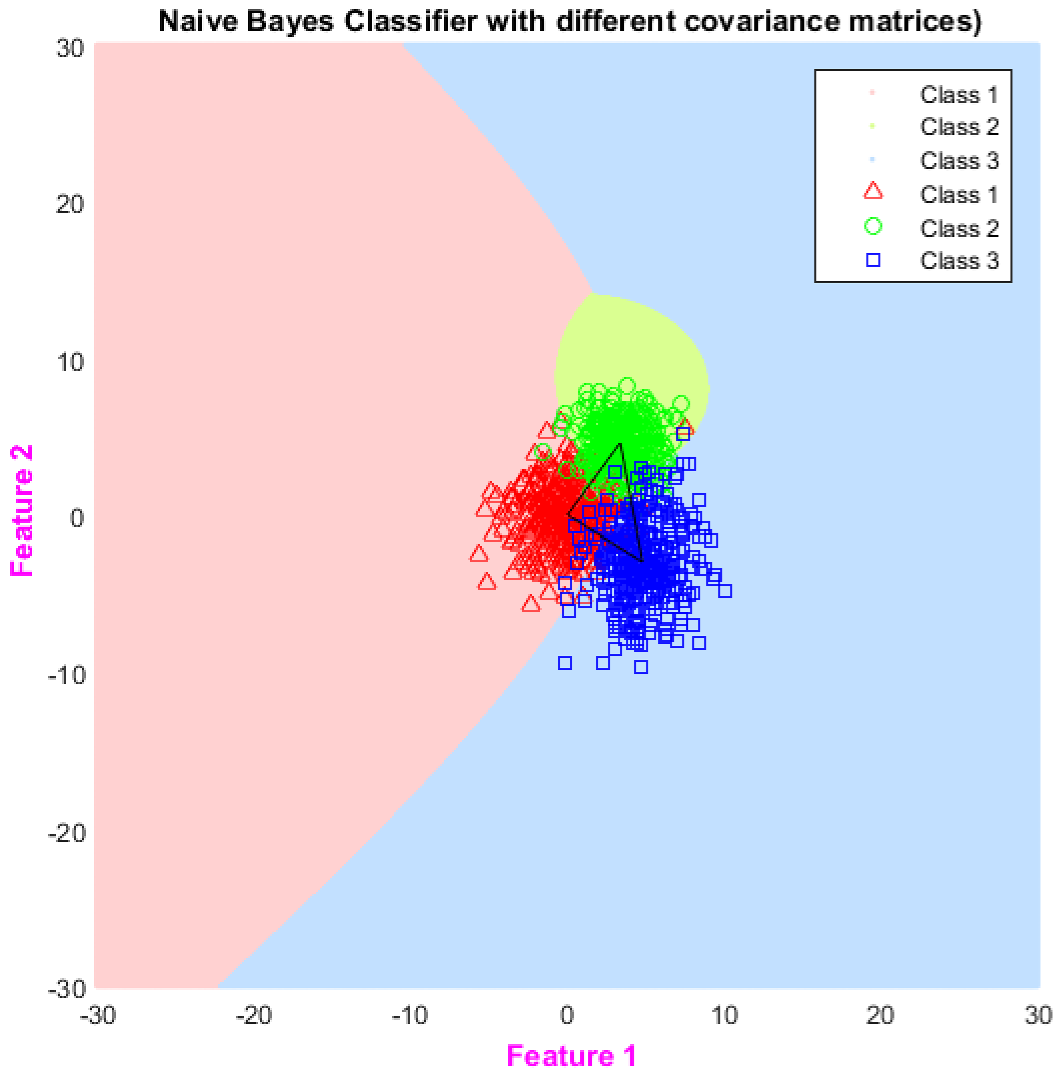
- *<u>Confusion Matrix based on performance for test data-</u>*

Predicted Class ⇨	CLASS 1	CLASS 2	CLASS 3
Actual Class ⇩			
Class 1	109	13	3
Class 2	3	121	1
Class 3	7	6	112

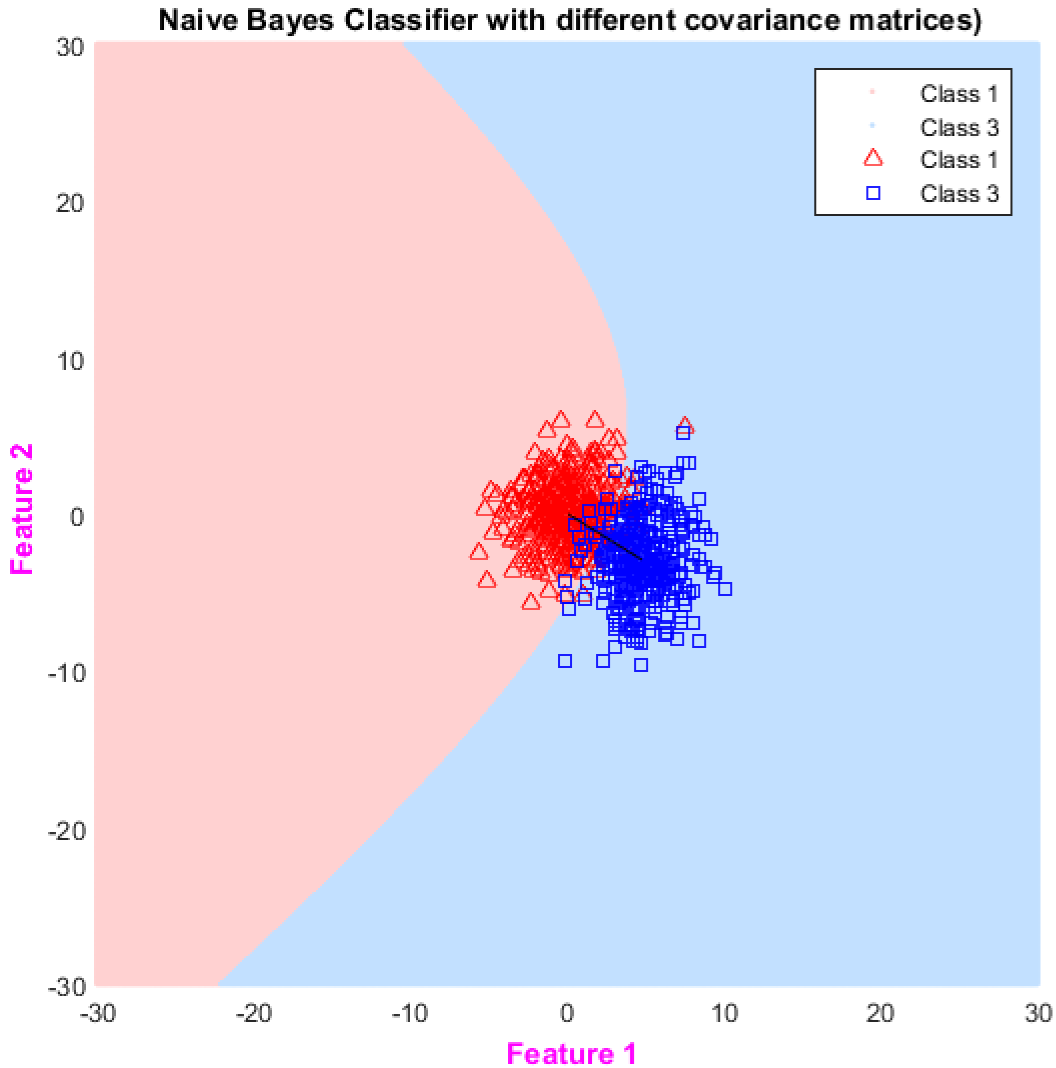
- *<u>Classification accuracy on test data –</u>*
  Overall Accuracy – 91.2
  Classifier Accuracy for class 1 – 87.2
  Classifier Accuracy for class 2 – 96.8
  Classifier Accuracy for class 3 – 89.6

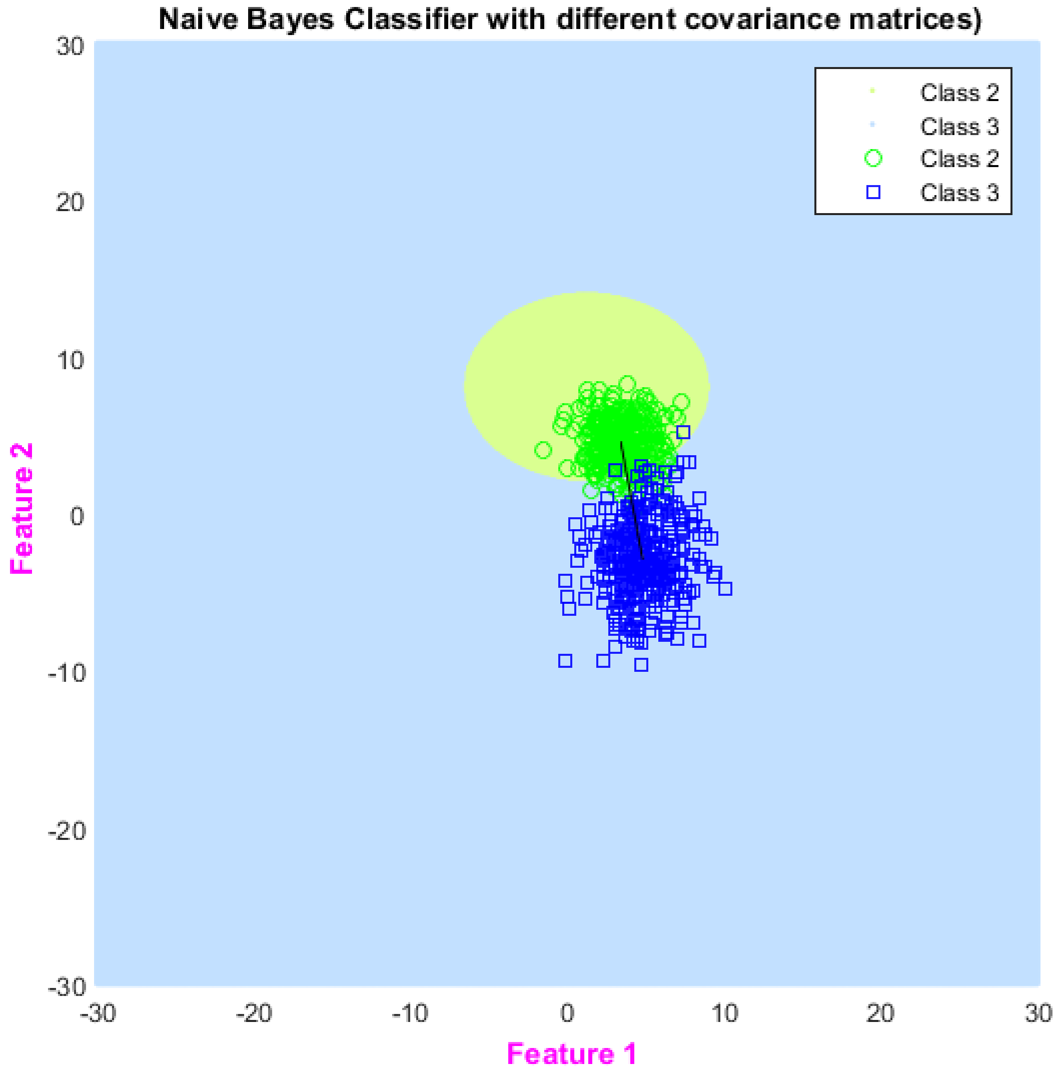
c) Covariance matrix for each class is different

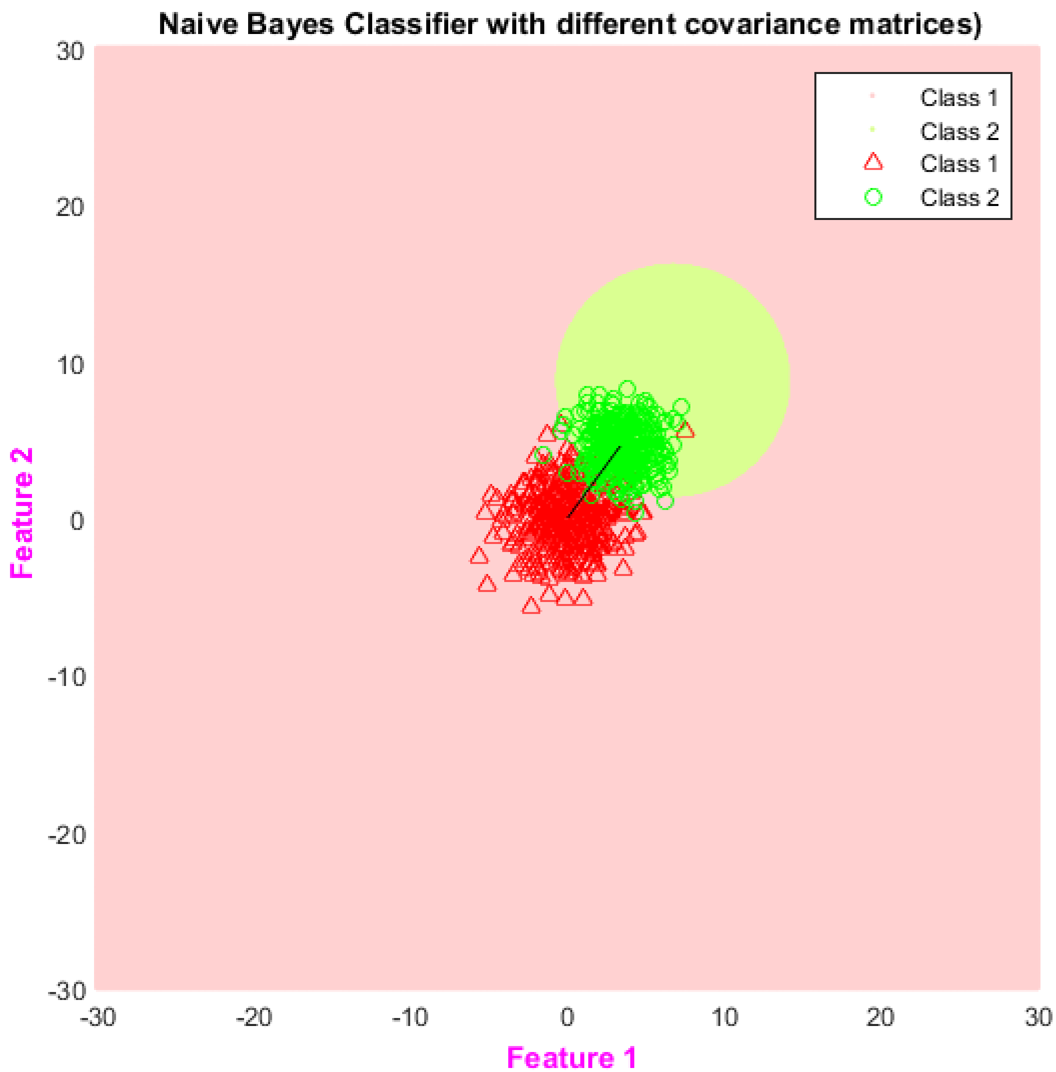
- *Decision region plot for all the classes together with the training data superposed -*



- *Decision region plot for every pair of classes -*




Naive Bayes Classifier with different covariance matrices)
Class 2
Class 3
Class 2
Class 3
Feature 2
Feature 1
30
20
10
0
-10
-20
-30
-30
-20
-10
0
10
20
30



- *<u>Confusion Matrix based on performance for test data-</u>*

Predicted Class ⇨	CLASS 1	CLASS 2	CLASS 3
Actual Class ⇩			
Class 1	110	11	4
Class 2	3	121	1
Class 3	8	2	115

- *<u>Classification accuracy on test data –</u>*
  Overall Accuracy – 92.2667
  Classifier Accuracy for class 1 – 88.0
  Classifier Accuracy for class 2 – 96.8
  Classifier Accuracy for class 3 – 92.0

- <u>**Observations –**</u>

1) **Bayes Classifier –**

   a) **Same covariance matrices (by taking average)**
   The decision boundaries are all of **linear** nature. They pass through the **midpoint** of line joining the means of two classes. The boundaries are **not necessarily** perpendicular to this line.

   b) **Same covariance matrices (from training data of all classes combined)**
   The decision boundaries are all of **linear** nature. They pass through the **midpoint** of line joining the means of two classes. The boundaries are **not necessarily** perpendicular to this line.

   c) **Covariance matrix for each class is different**
   The decision boundaries are all of **Quadratic** nature. For this dataset they turn out to be either **Parabola** or **Ellipse**. They do **not necessarily** pass through the **midpoint** of line joining the means of two classes.

2) **Naïve Bayes Classifier –**

   a) **Same covariance matrices (σ^2 I)**
   The decision boundaries are all of **linear** nature. They pass through the **midpoint** of line joining the means of two classes. The boundaries are perpendicular to this line.

   b) **Same covariance matrices and is C**
   The decision boundaries are all of **linear** nature. They pass through the **midpoint** of line joining the means of two classes. The boundaries are **not necessarily** perpendicular to this line.

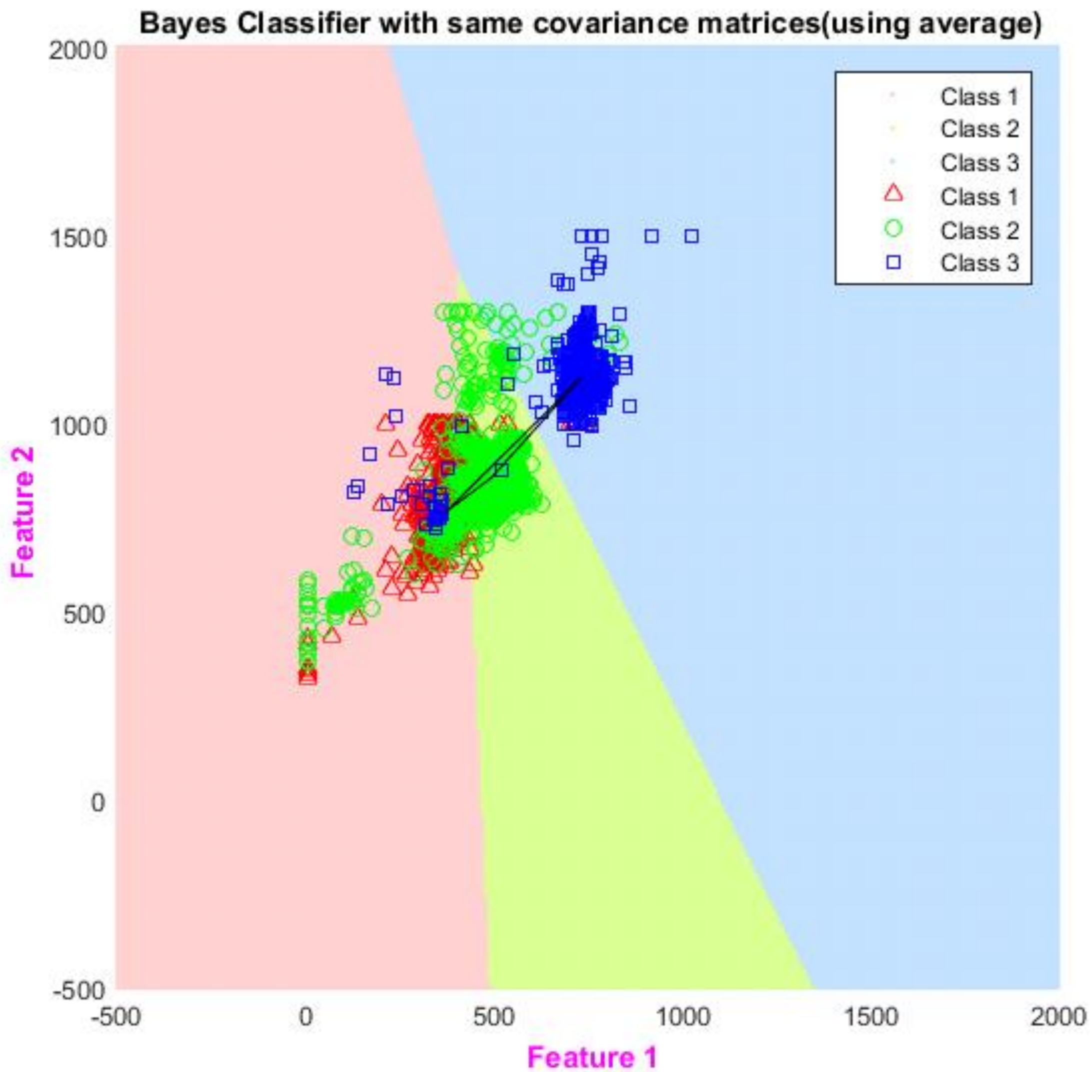
   c) **Covariance matrix for each class is different**
   The decision boundaries are all of **Quadratic** nature. For this dataset they turn out to be either **Parabola** or **Ellipse**. They do **not necessarily** pass through the **midpoint** of line joining the means of two classes. The classification results are more or less similar to that of Bayes.

♦ **Dataset II** : Real world data of 3 classes
(The real world data sets correspond to the formant frequencies F1 and F2 for vowel utterances.)

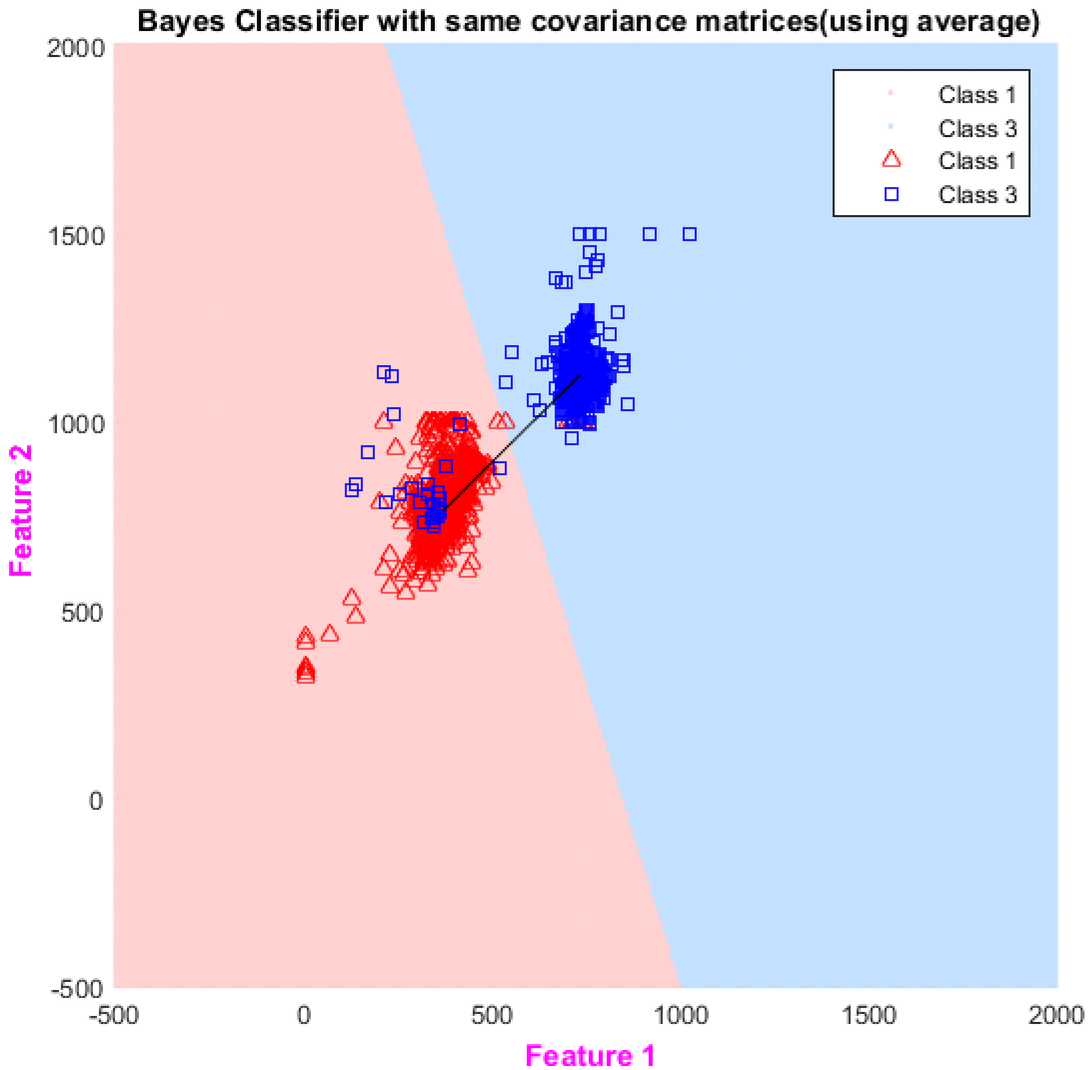
1) Bayes Classifier –

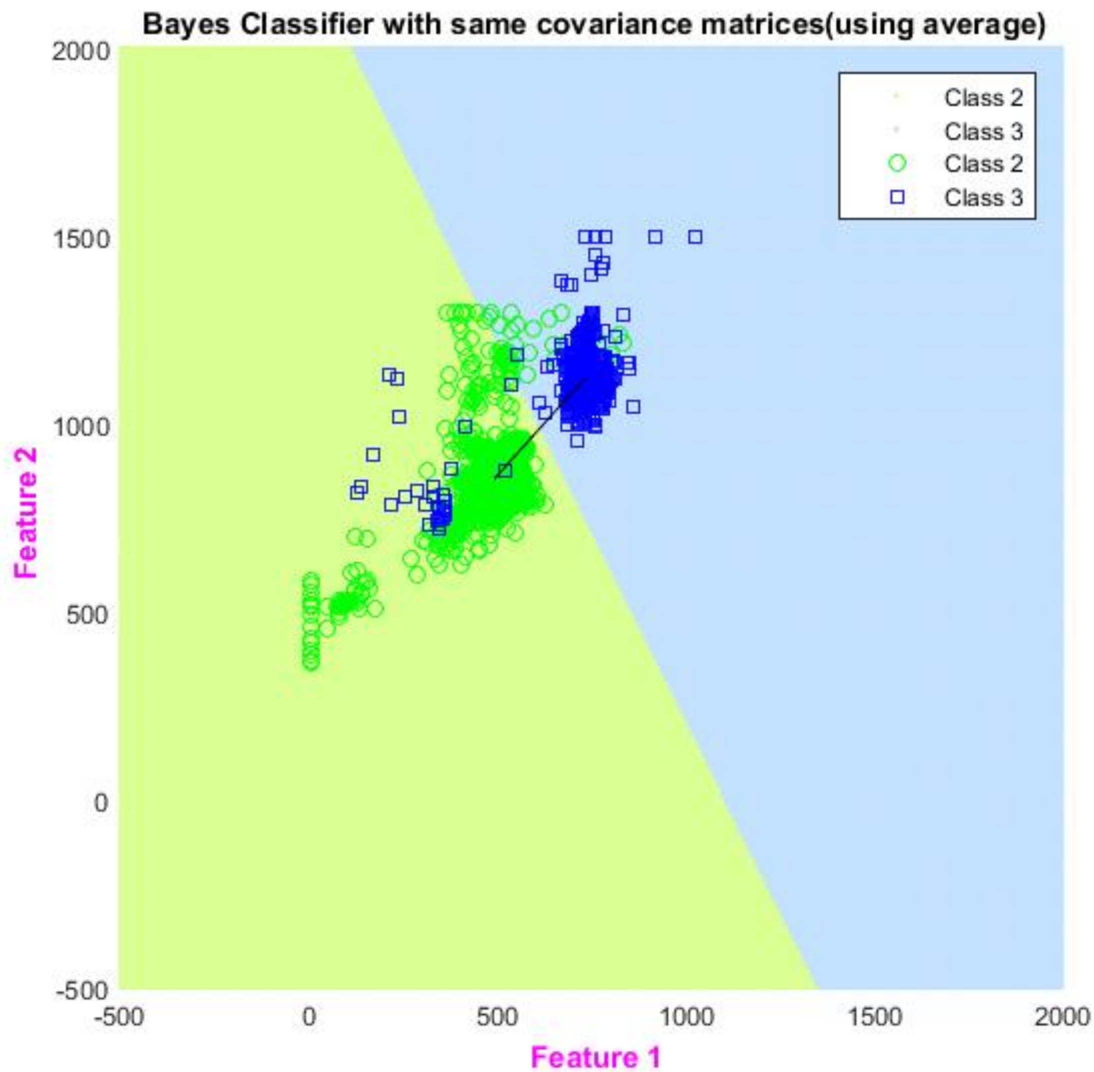
a) Same covariance matrices (by taking average)

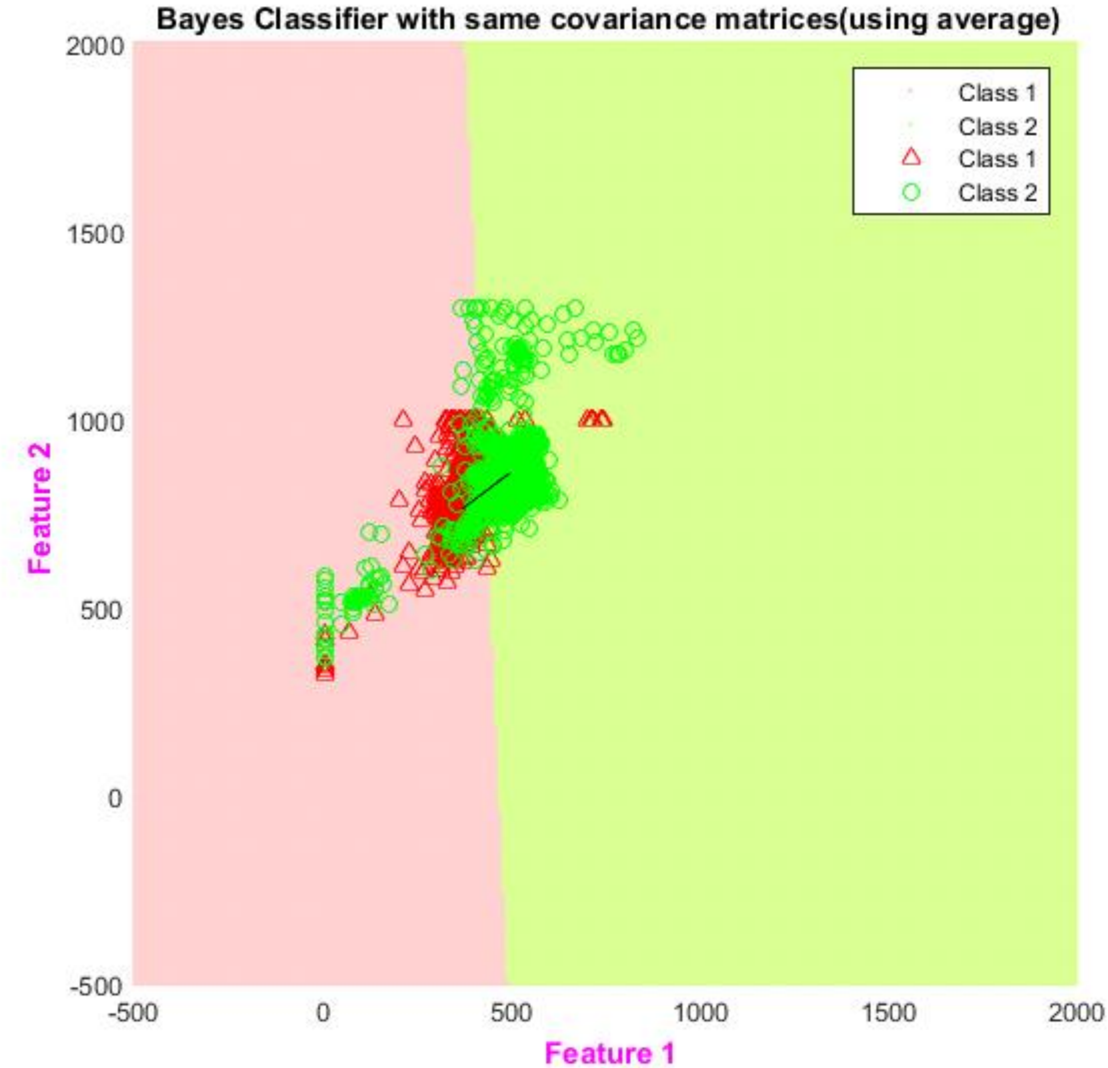
- *Decision region plot for all the classes together with the training data superposed -*



- *Decision region plot for every pair of classes -*




Bayes Classifier with same covariance matrices(using average)
Class 2
Class 3
Class 2
Class 3
2000
1500
1000
500
0
-500
-500
0
500
1000
1500
2000
Feature 2
Feature 1



- *Confusion Matrix based on performance for test data-*

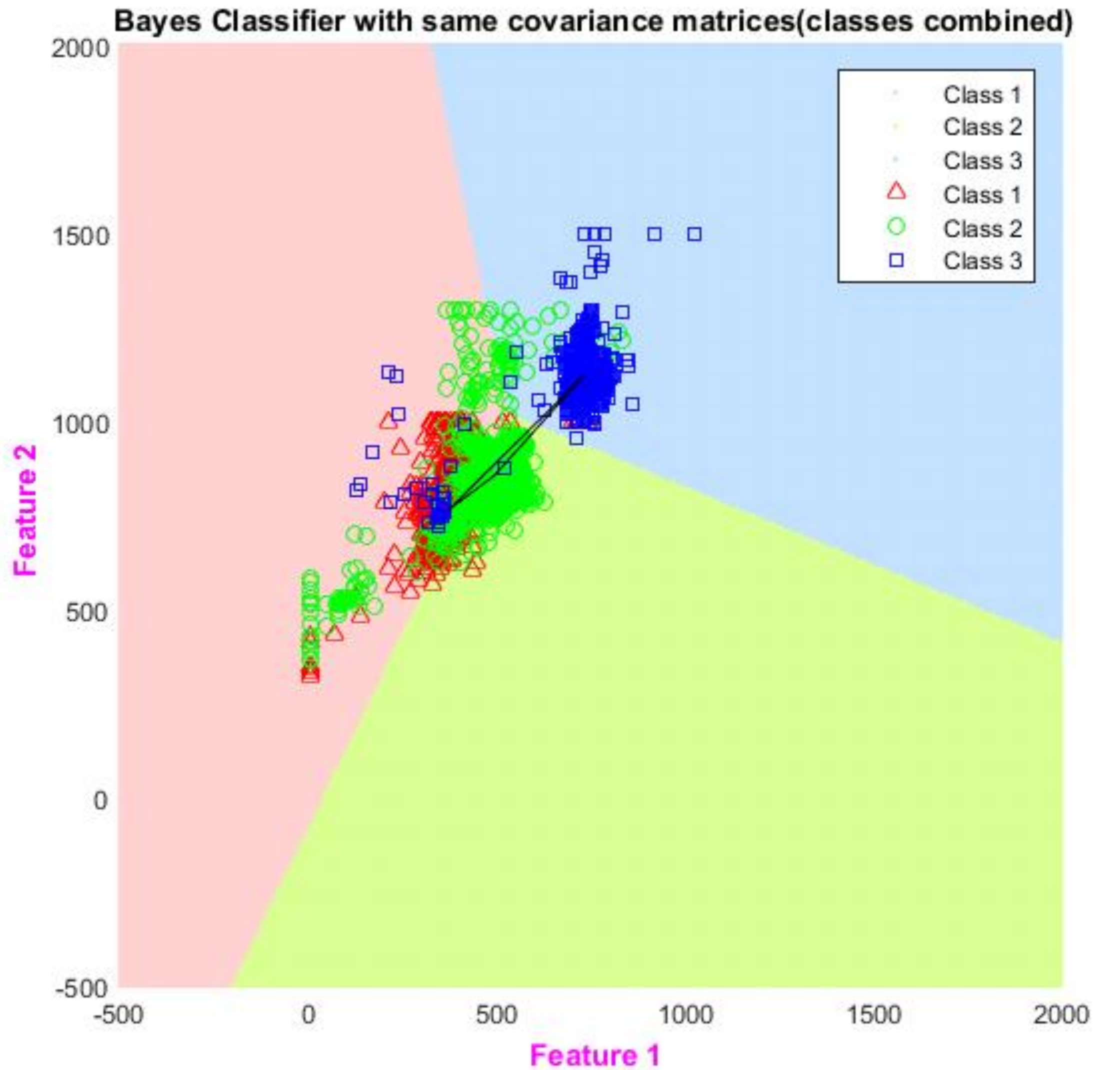
Predicted Class ⇨	CLASS 1	CLASS 2	CLASS 3
Actual Class ⇩			
Class 1	602	13	7
Class 2	197	397	20
Class 3	26	5	510

- *Classification accuracy on test data –*

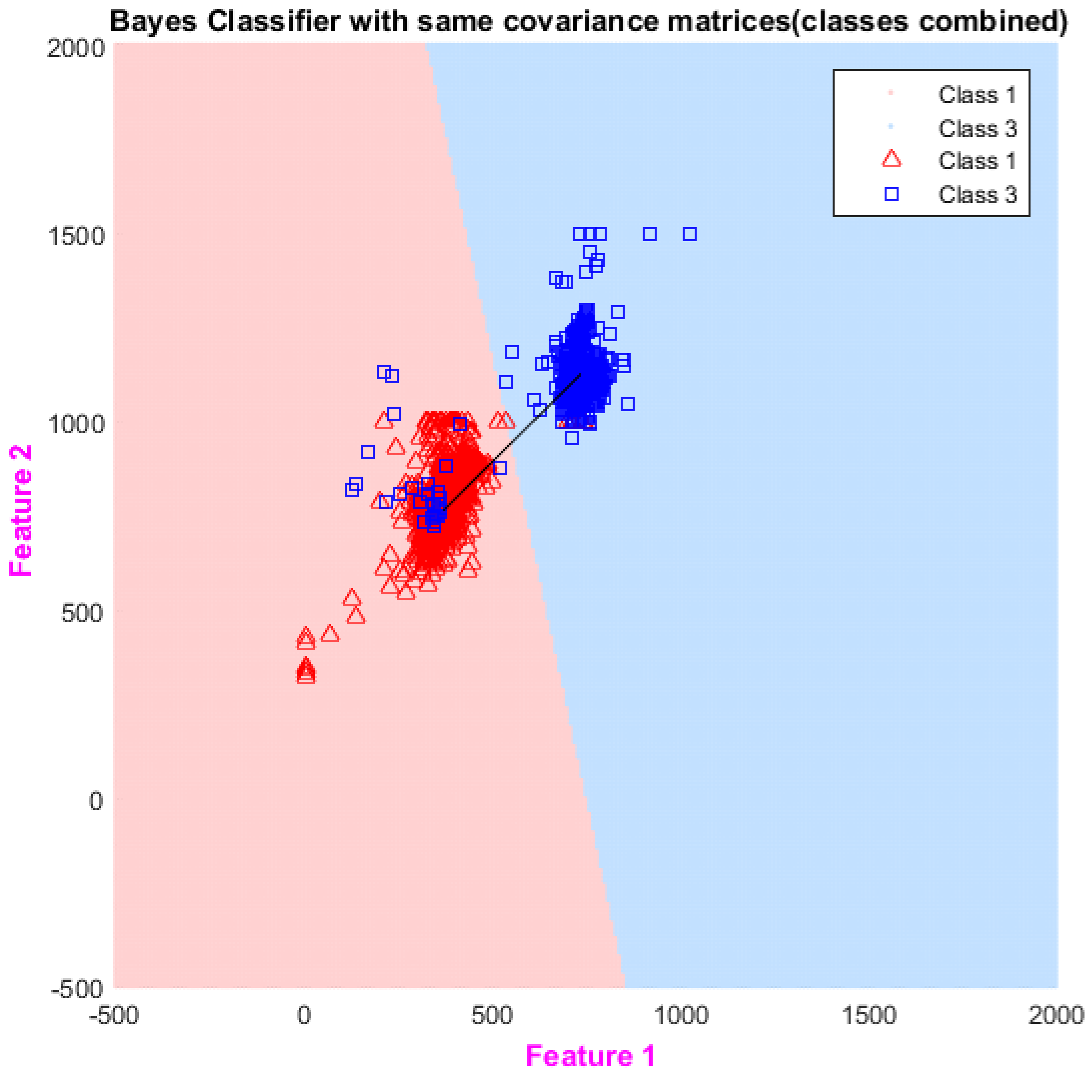
Overall Accuracy – 84.9184
Classifier Accuracy for class 1 – 96.7846
Classifier Accuracy for class 2 – 64.6580
Classifier Accuracy for class 3 – 94.2699

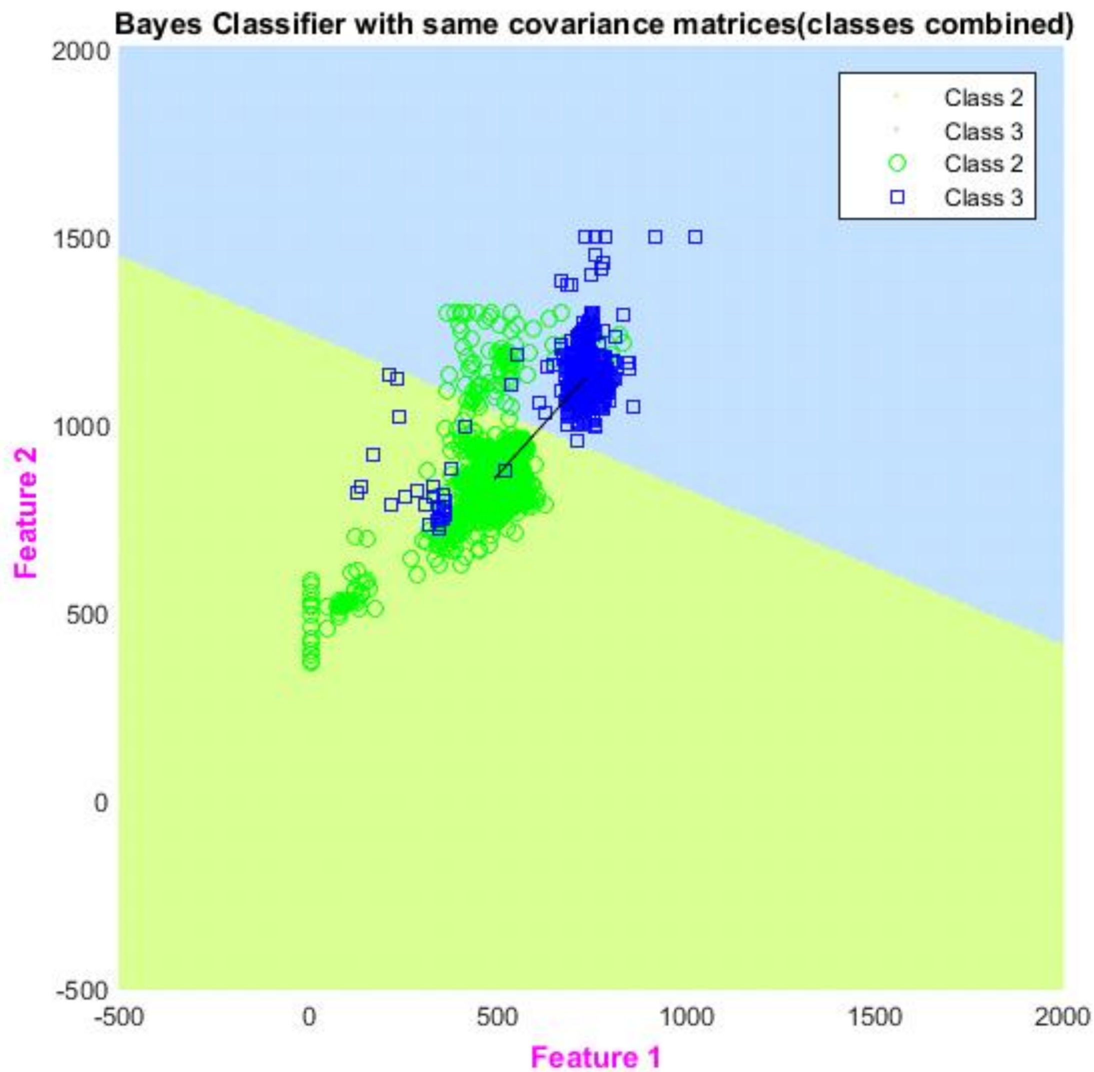
b) Same covariance matrices (from training data of all classes combined)

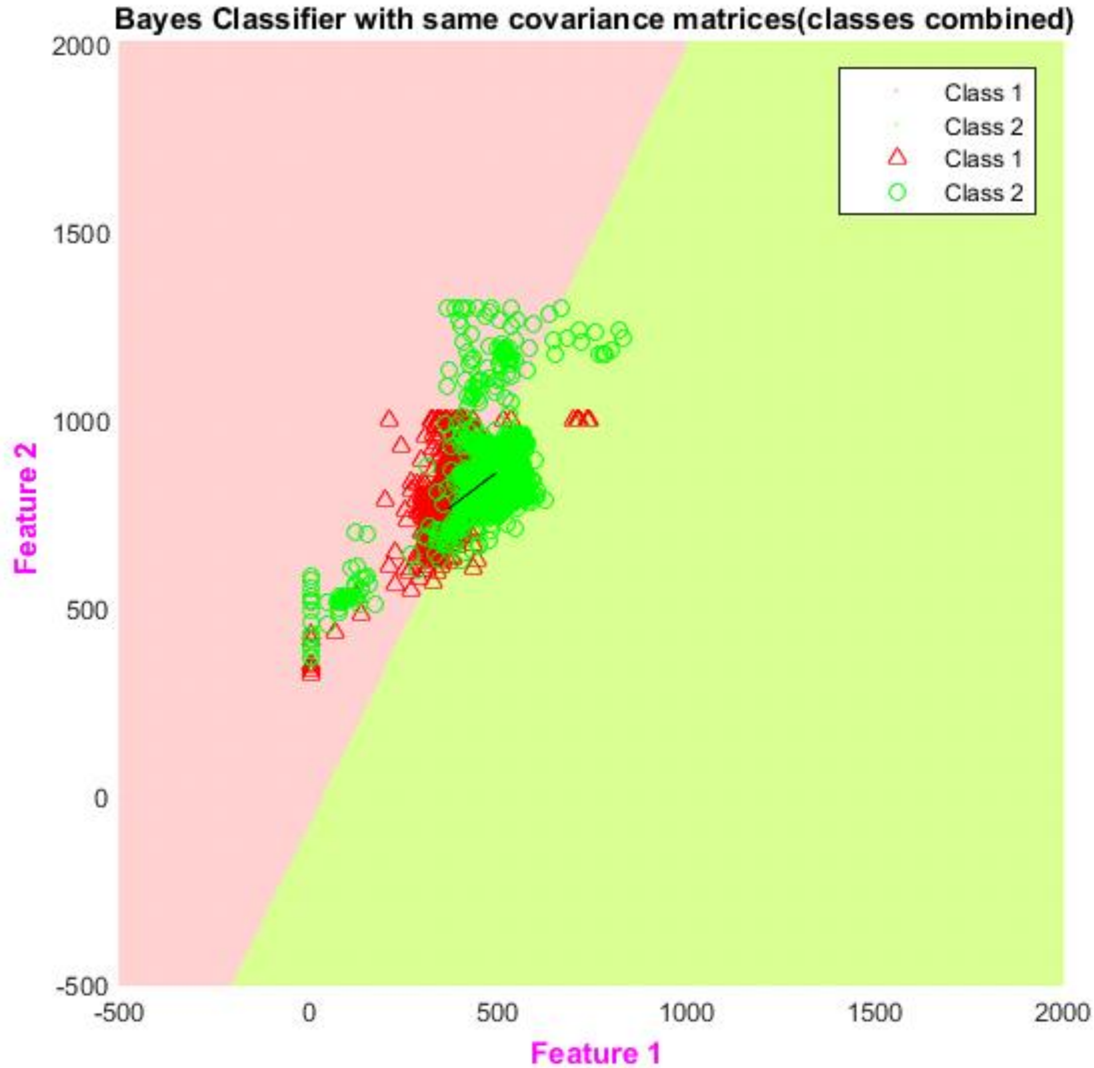
- *<u>Decision region plot for all the classes together with the training data superposed -</u>*



- *Decision region plot for every pair of classes -*




Bayes Classifier with same covariance matrices(classes combined)
Class 2
Class 3
Class 2
Class 3
Feature 1
Feature 2
2000
1500
1000
500
0
-500
-500
0
500
1000
1500
2000



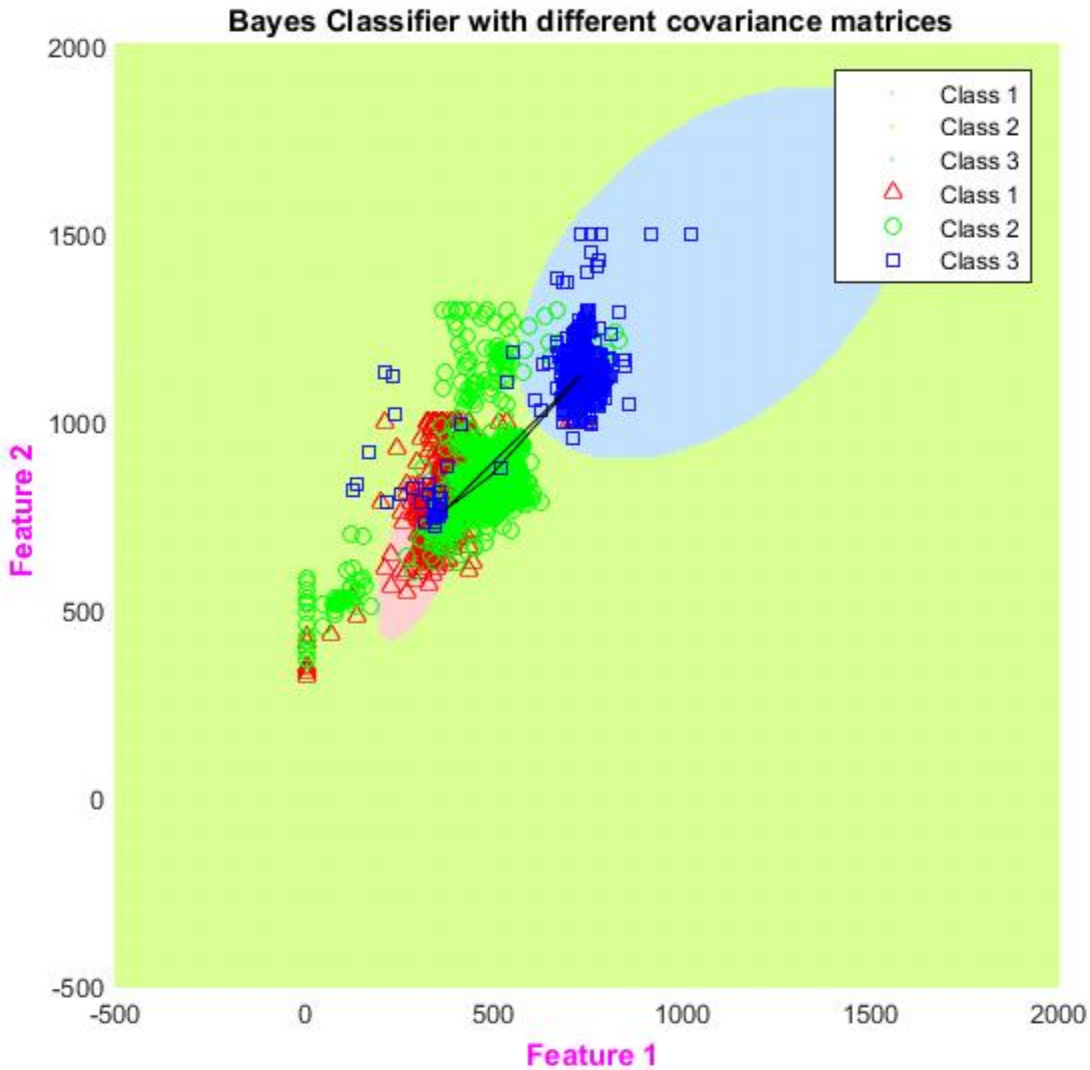
- *<u>Confusion Matrix based on performance for test data-</u>*

Predicted Class ⇨	CLASS 1	CLASS 2	CLASS 3
Actual Class ⇩			
Class 1	544	71	7
Class 2	261	332	21
Class 3	30	1	510

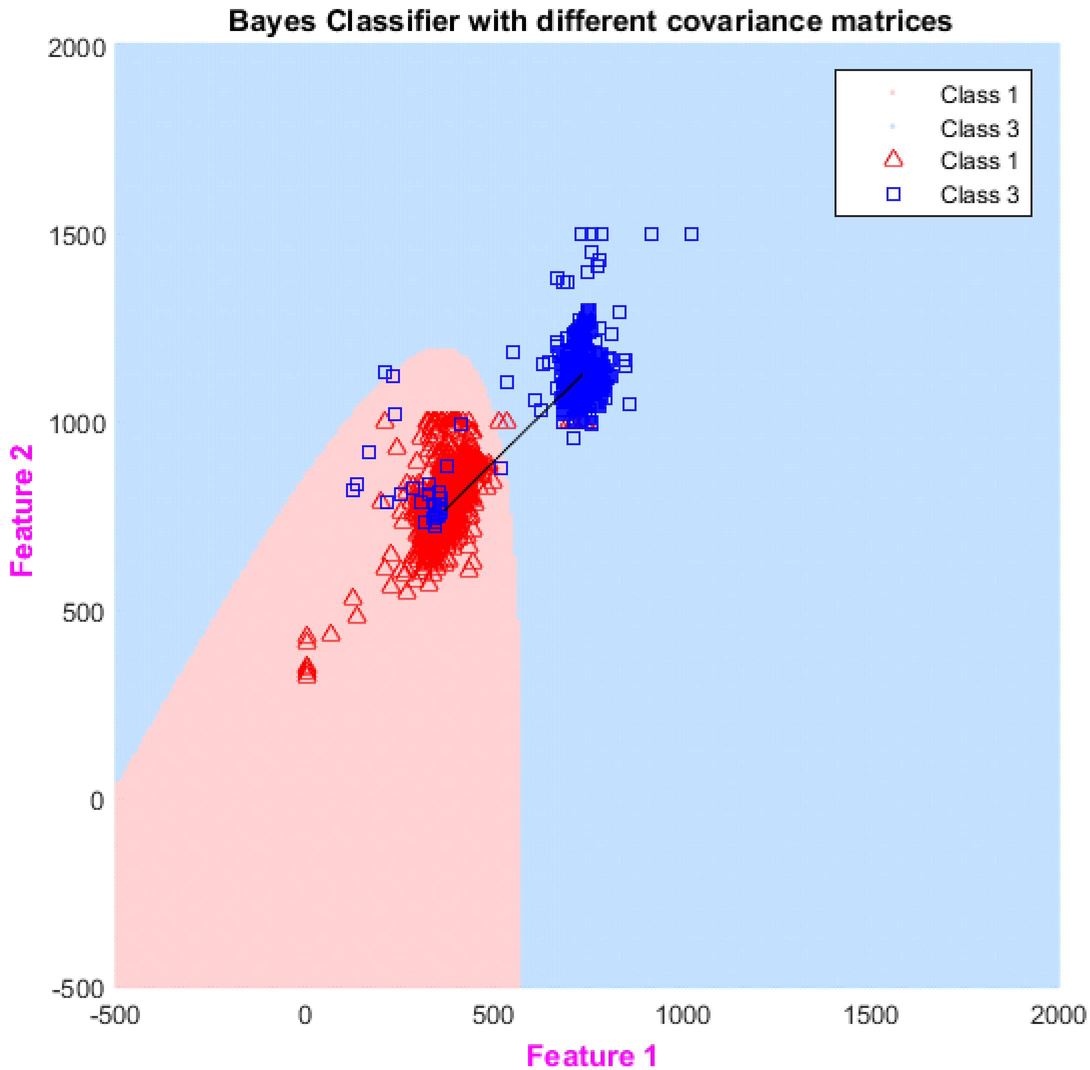
- *<u>Classification accuracy on test data –</u>*
  Overall Accuracy – 77.9966
  Classifier Accuracy for class 1 – 87.4598
  Classifier Accuracy for class 2 – 54.0717
  Classifier Accuracy for class 3 – 94.2699

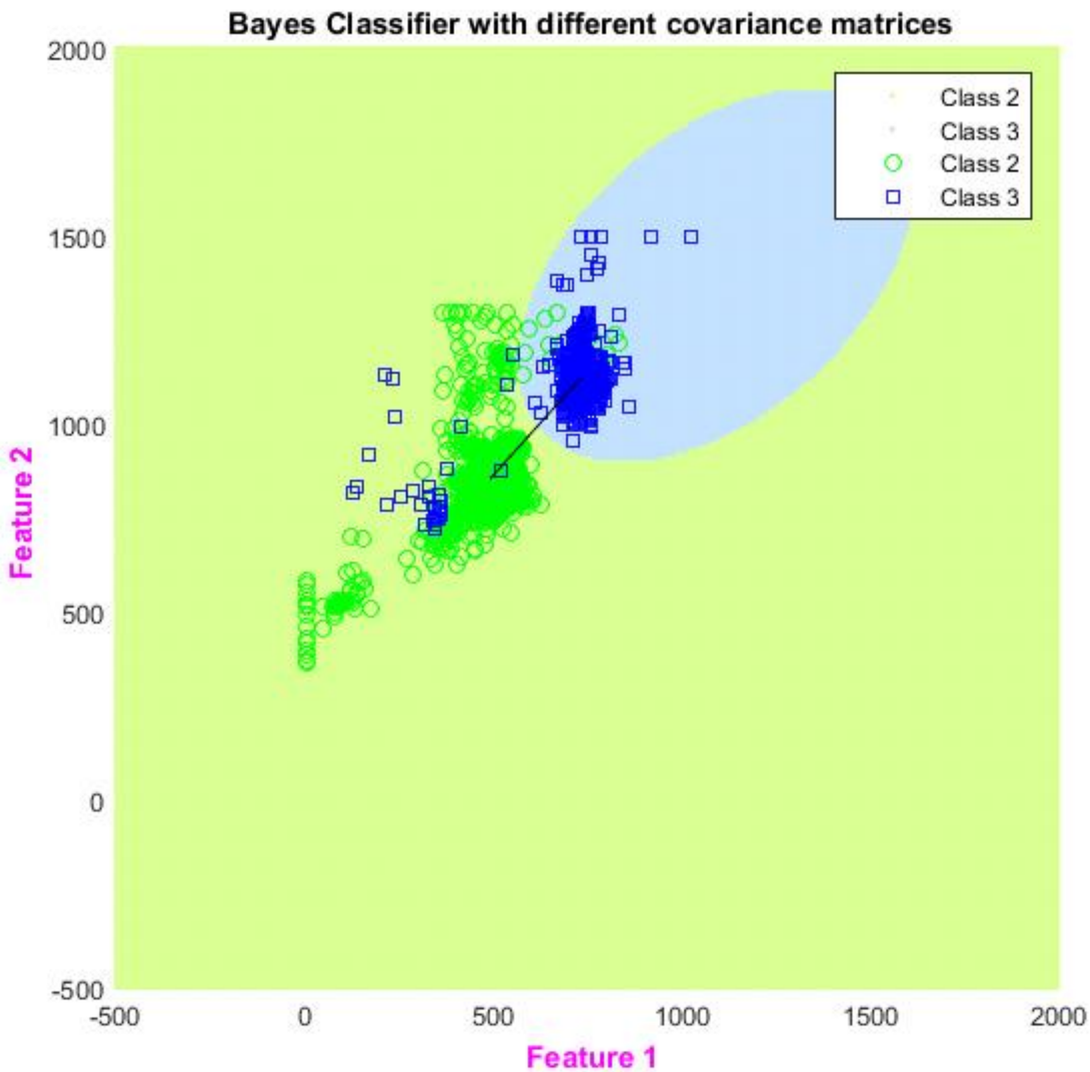
c) Covariance matrix for each class is different

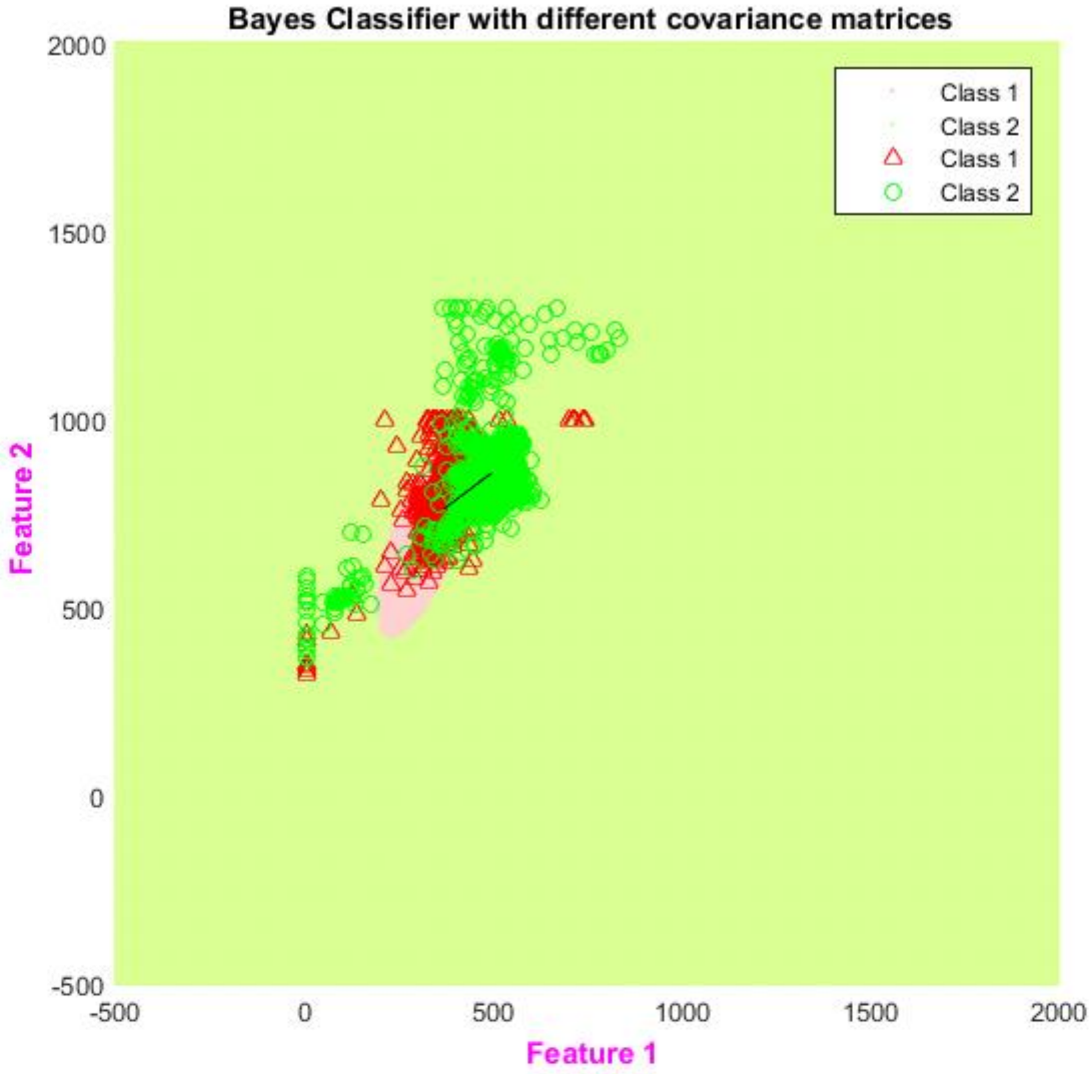
- *Decision region plot for all the classes together with the training data superposed -*



- *Decision region plot for every pair of classes -*




Bayes Classifier with different covariance matrices
Class 2
Class 3
Class 2
Class 3
Feature 2
Feature 1
2000
1500
1000
500
0
-500
-500
0
500
1000
1500
2000



- *<u>Confusion Matrix based on performance for test data-</u>*

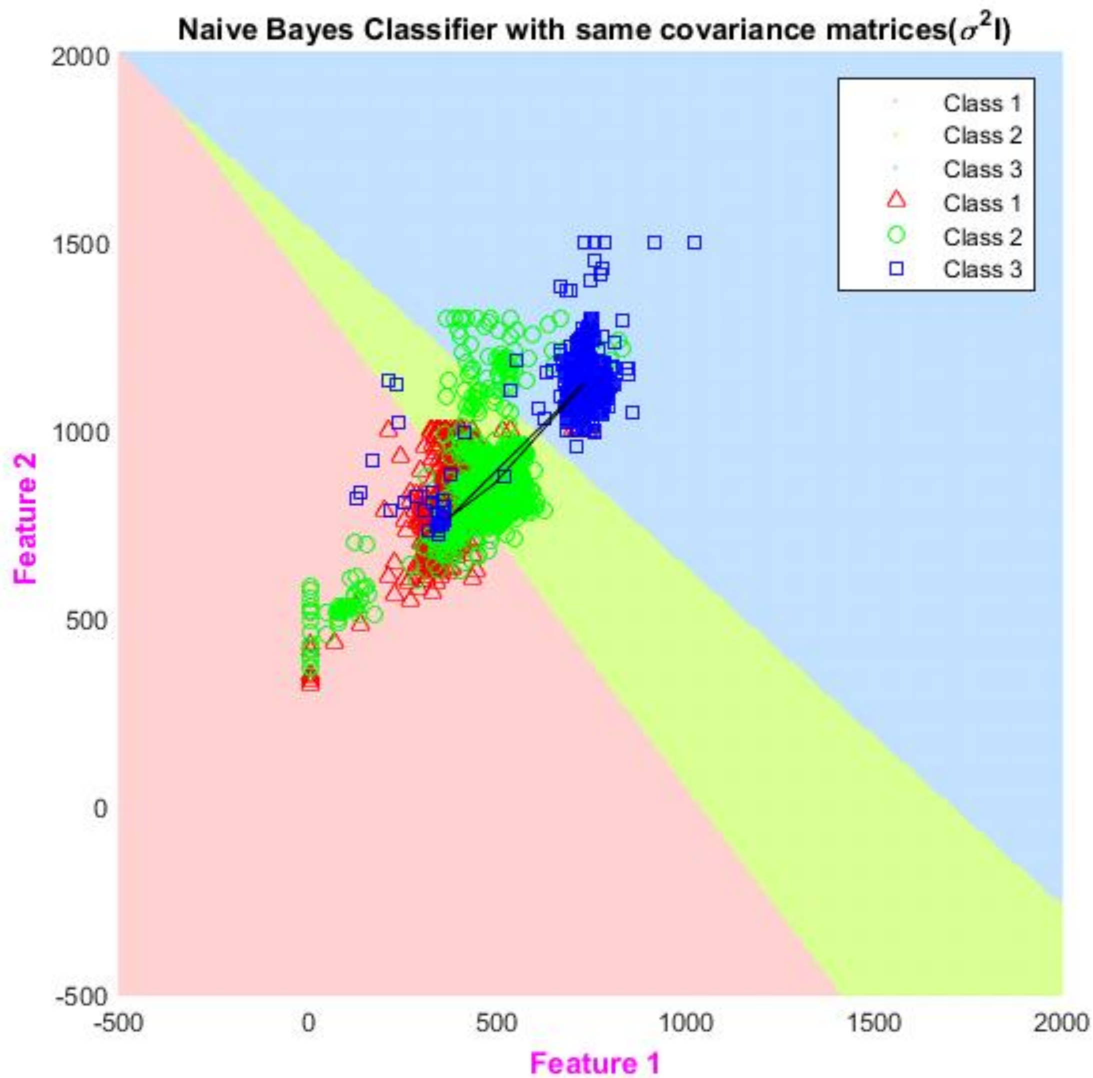
Predicted Class ⇨	CLASS 1	CLASS 2	CLASS 3
Actual Class ⇩			
Class 1	531	84	7
Class 2	203	406	5
Class 3	13	19	509

- *<u>Classification accuracy on test data –</u>*
  Overall Accuracy – 81.3731
  Classifier Accuracy for class 1 – 85.3698
  Classifier Accuracy for class 2 – 66.1238
  Classifier Accuracy for class 3 – 94.0850

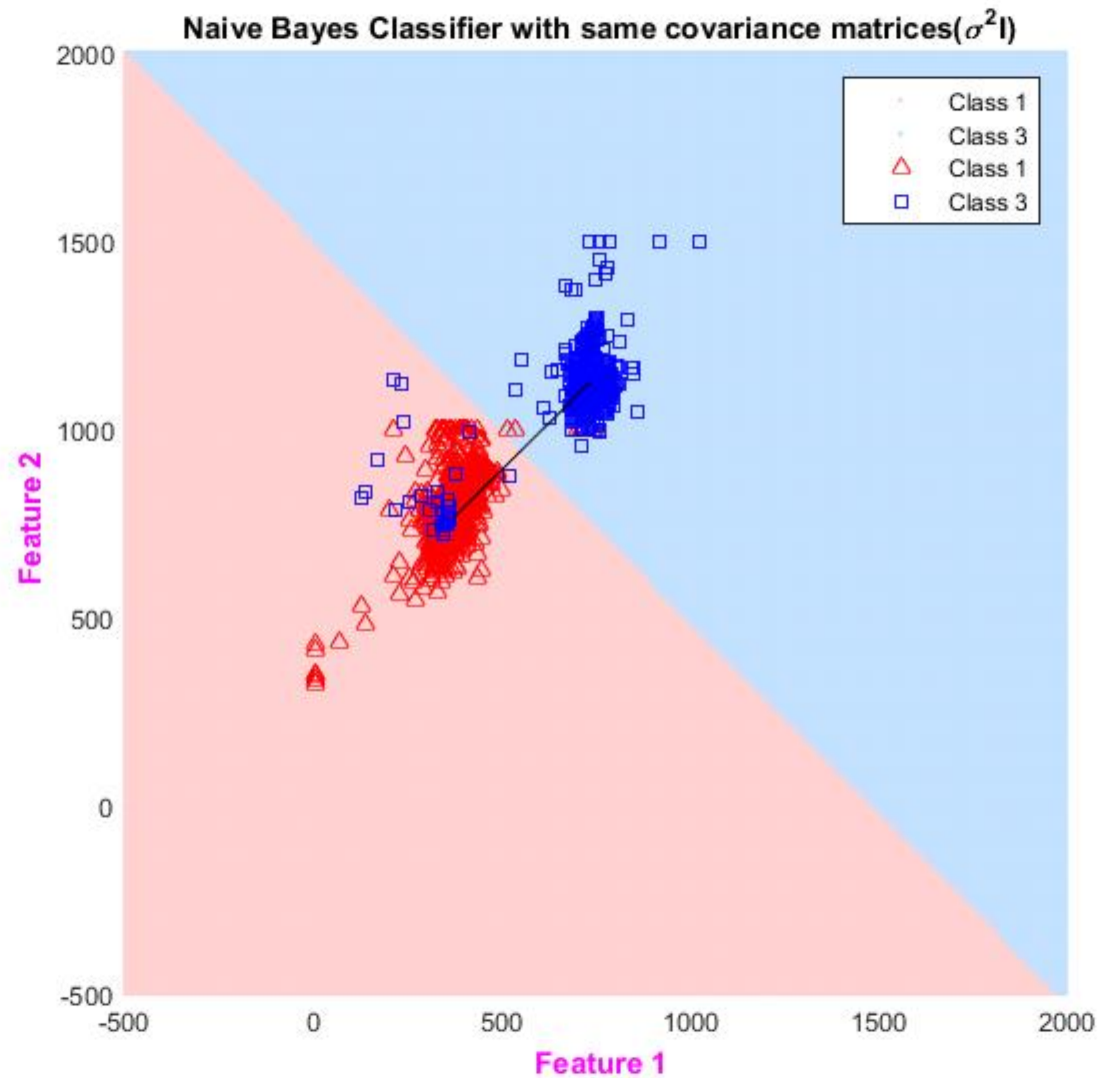
2) Naive Bayes Classifier –

a) Same covariance matrices (σ^2 I)

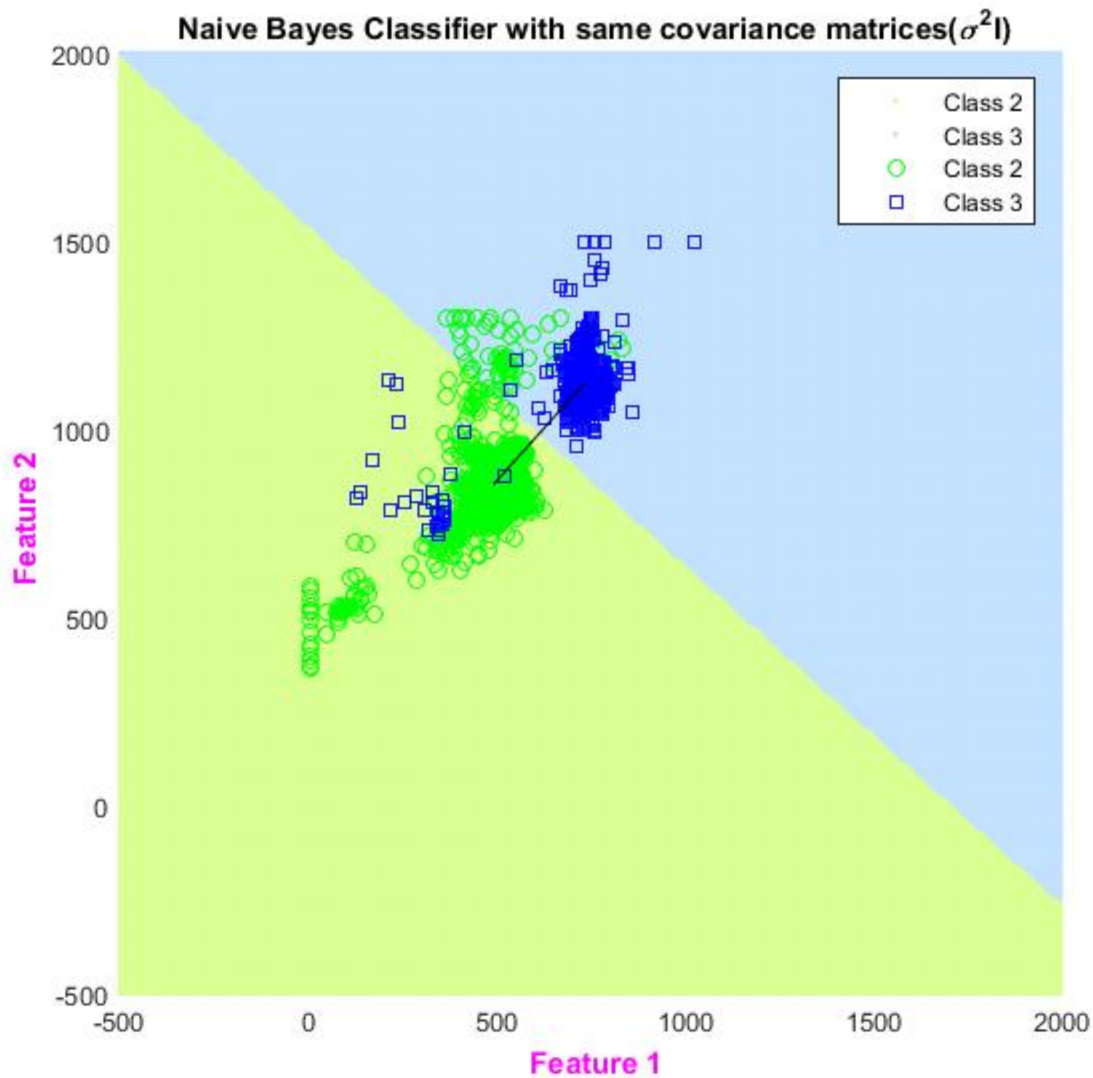
- *Decision region plot for all the classes together with the training data superposed -*

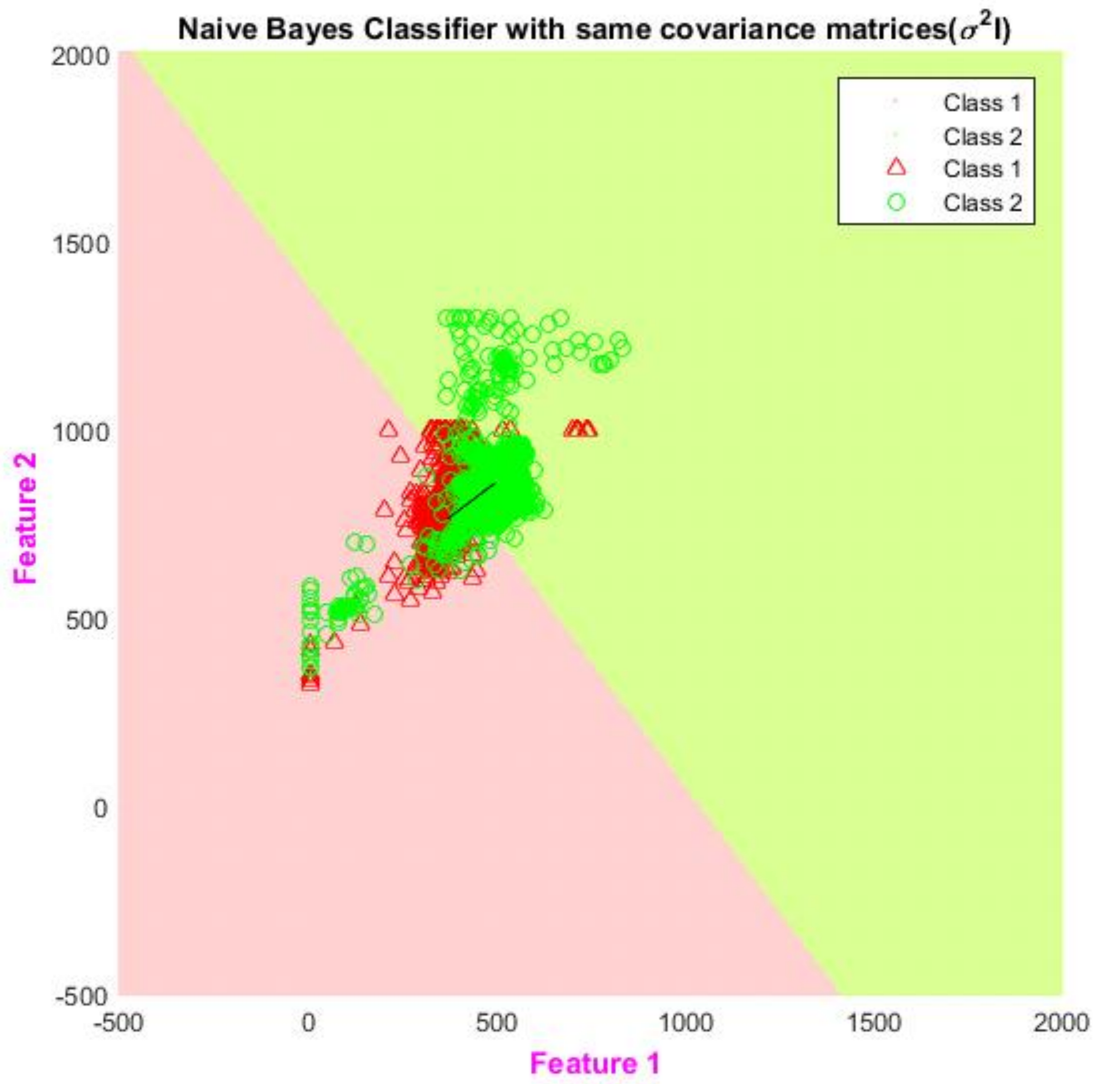


- *Decision region plot for every pair of classes -*



u


Naive Bayes Classifier with same covariance matrices(σ^2I)
Class 2
Class 3
Class 2
Class 3
2000
1500
1000
500
0
-500
-500
0
500
1000
1500
2000
Feature 2
Feature 1



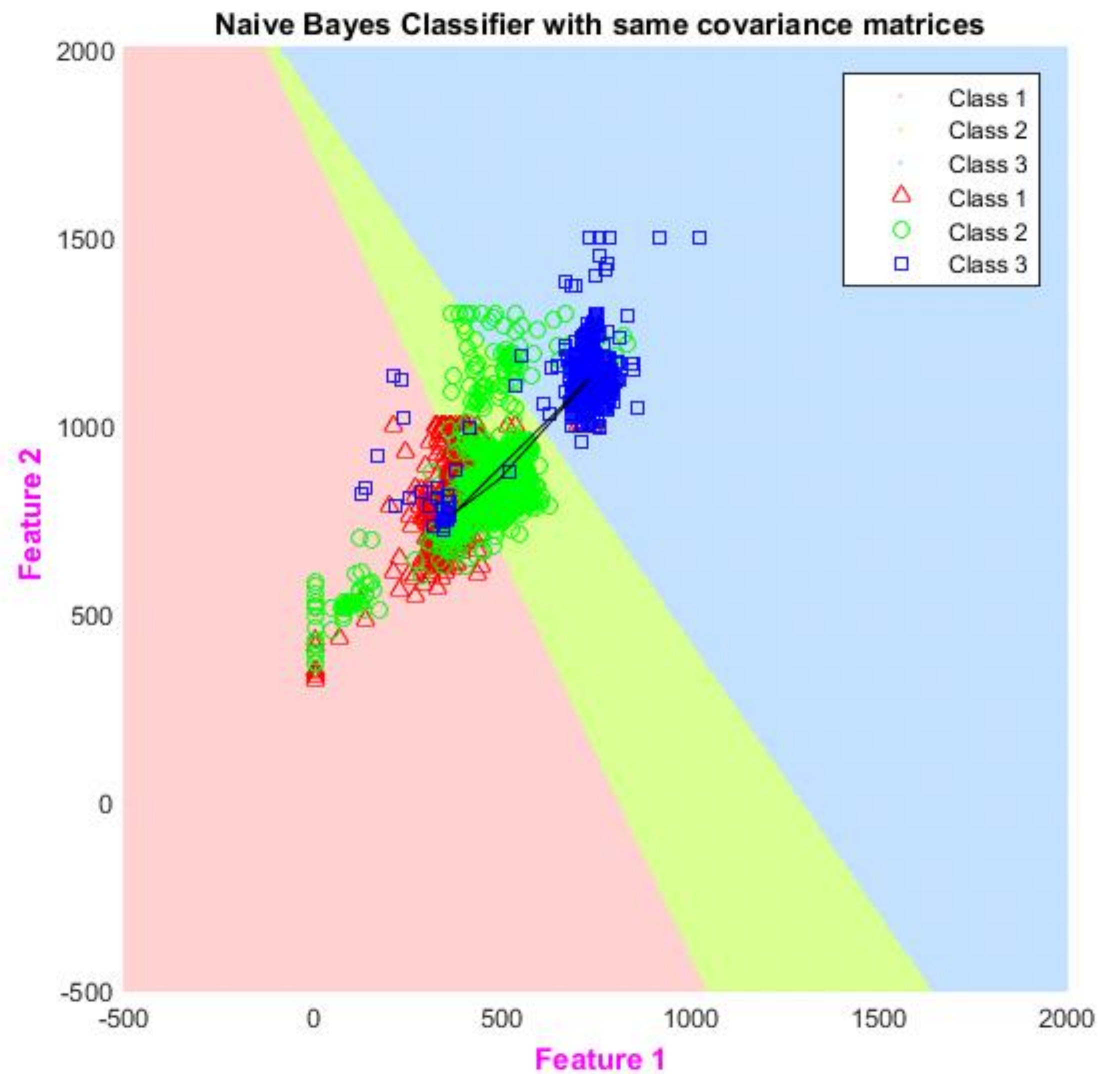
- *Confusion Matrix based on performance for test data-*

Predicted Class ⇨	CLASS 1	CLASS 2	CLASS 3
Actual Class ⇩			
Class 1	572	43	7
Class 2	172	398	44
Class 3	24	5	512

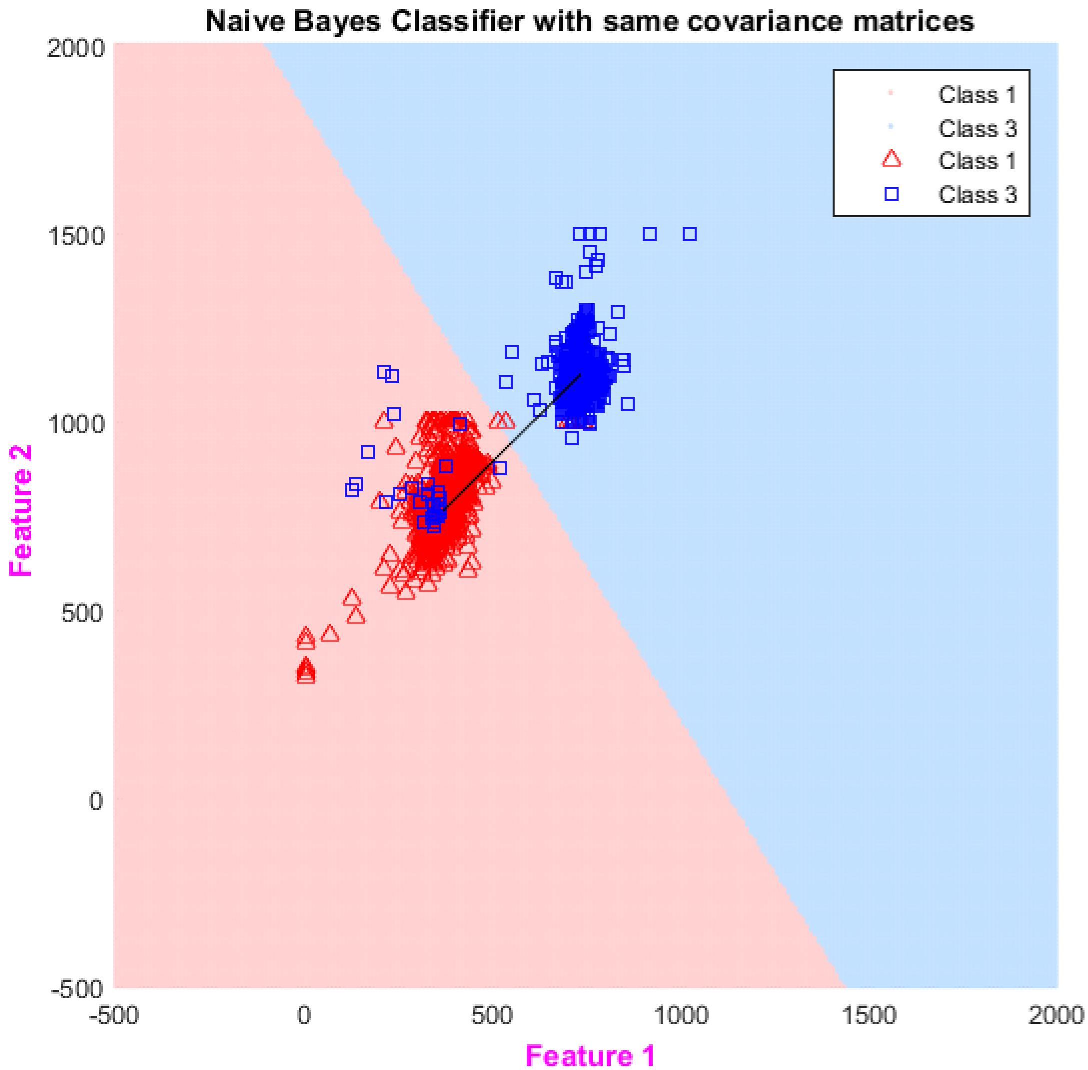
- *Classification accuracy on test data –*
  Overall Accuracy – 83.3990
  Classifier Accuracy for class 1 - -91.9614
  Classifier Accuracy for class 2 – 64.8208
  Classifier Accuracy for class 3 – 94.6396

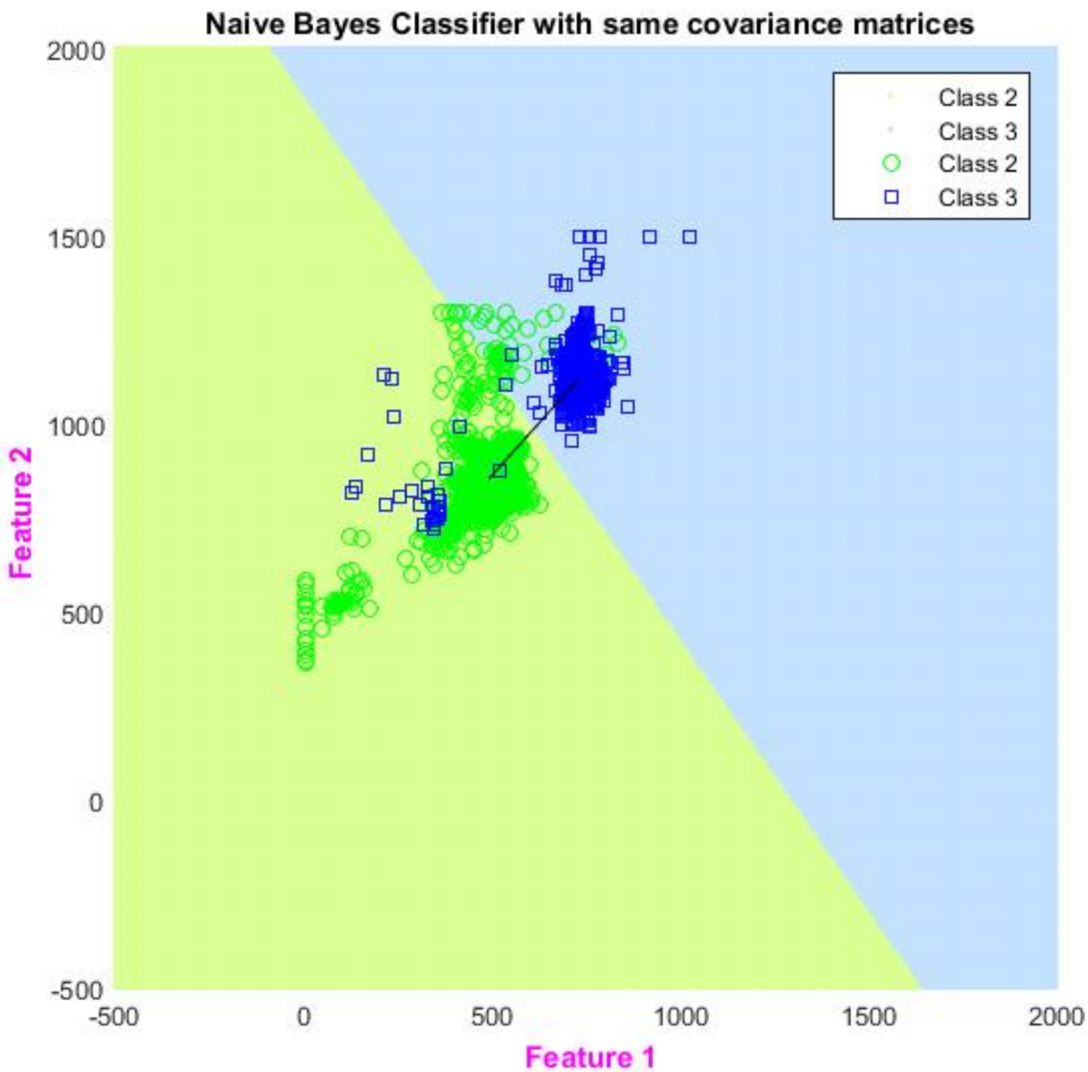
b) Same covariance matrices and is C

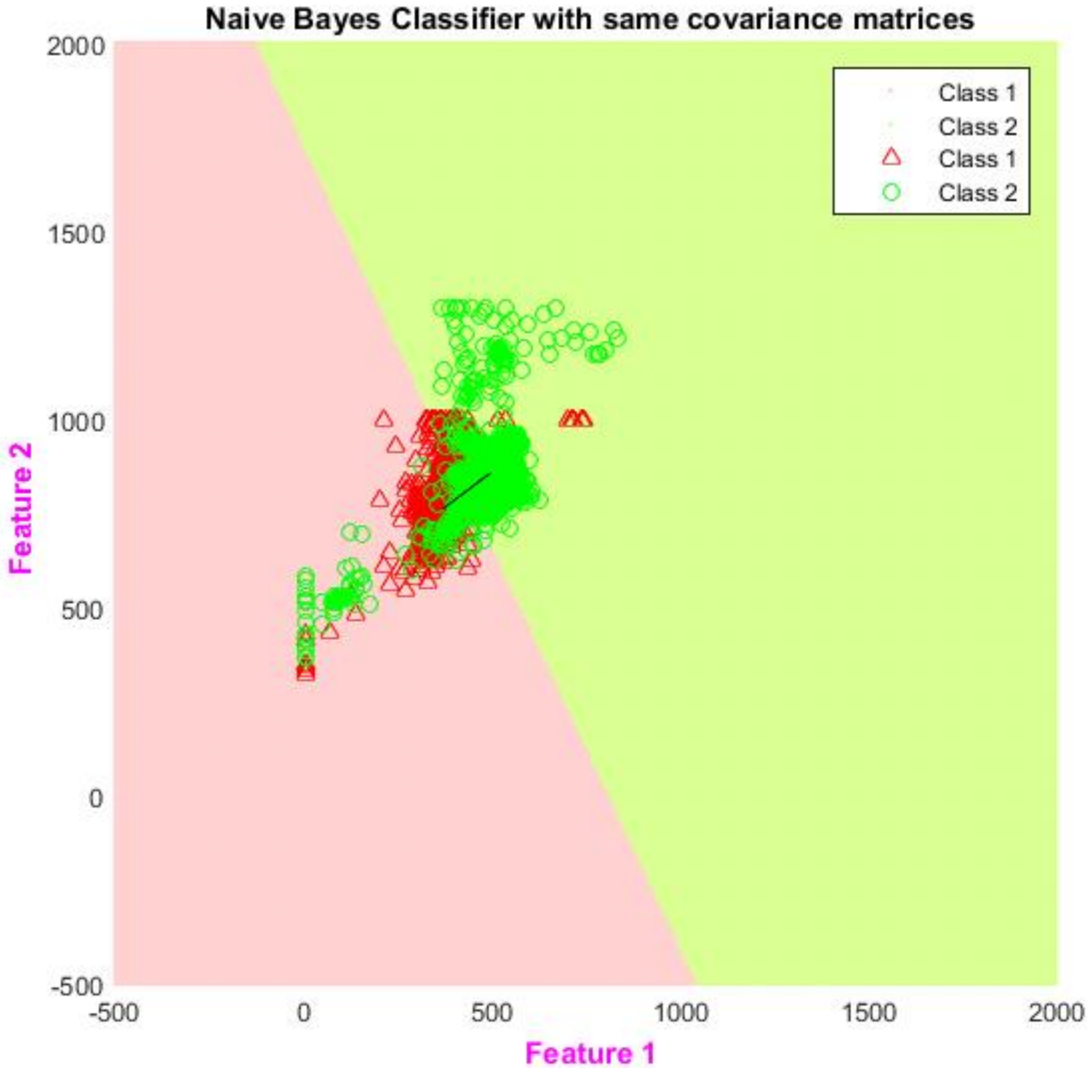
- *<u>Decision region plot for all the classes together with the training data superposed -</u>*



- *Decision region plot for every pair of classes -*




Naive Bayes Classifier with same covariance matrices
Class 2
Class 3
Class 2
Class 3
Feature 2
Feature 1
2000
1500
1000
500
0
-500
-500
0
500
1000
1500
2000



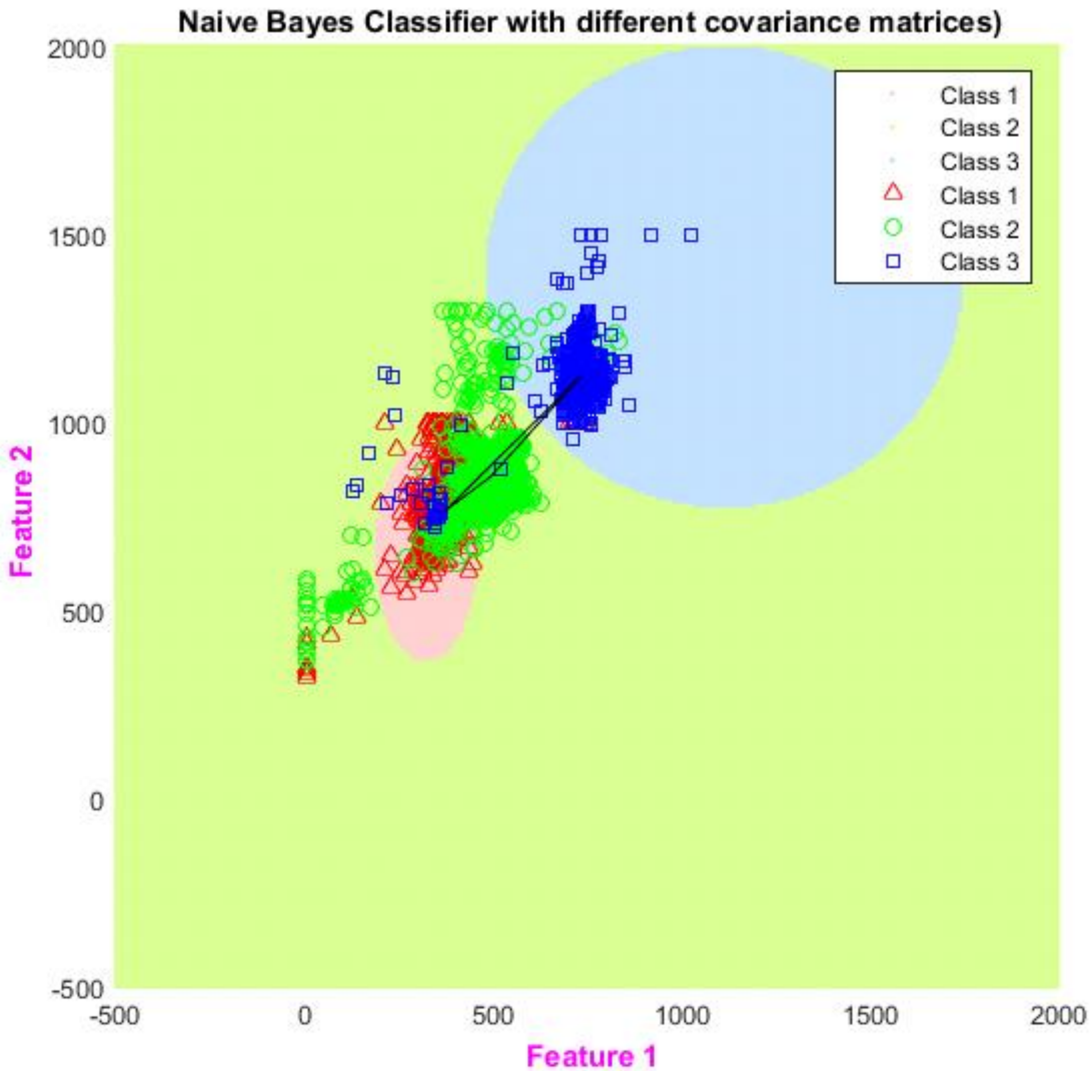
- *<u>Confusion Matrix based on performance for test data-</u>*

Predicted Class ⇨	CLASS 1	CLASS 2	CLASS 3
Actual Class ⇩			
Class 1	580	35	7
Class 2	181	402	31
Class 3	25	4	512

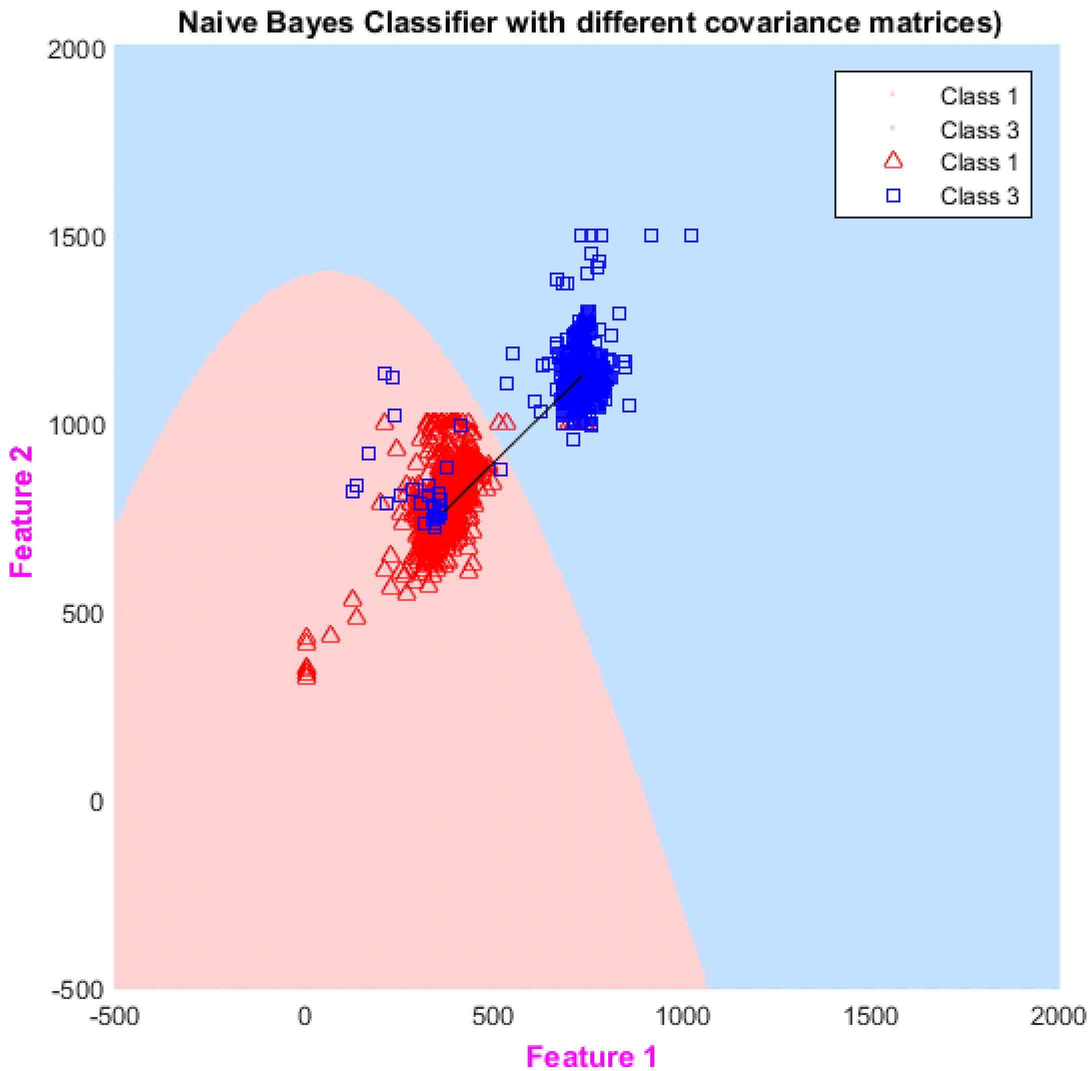
- *<u>Classification accuracy on test data –</u>*
  Overall Accuracy – 84.0743
  Classifier Accuracy for class 1 – 93.2476
  Classifier Accuracy for class 2 – 65.4723
  Classifier Accuracy for class 3 – 94.6396

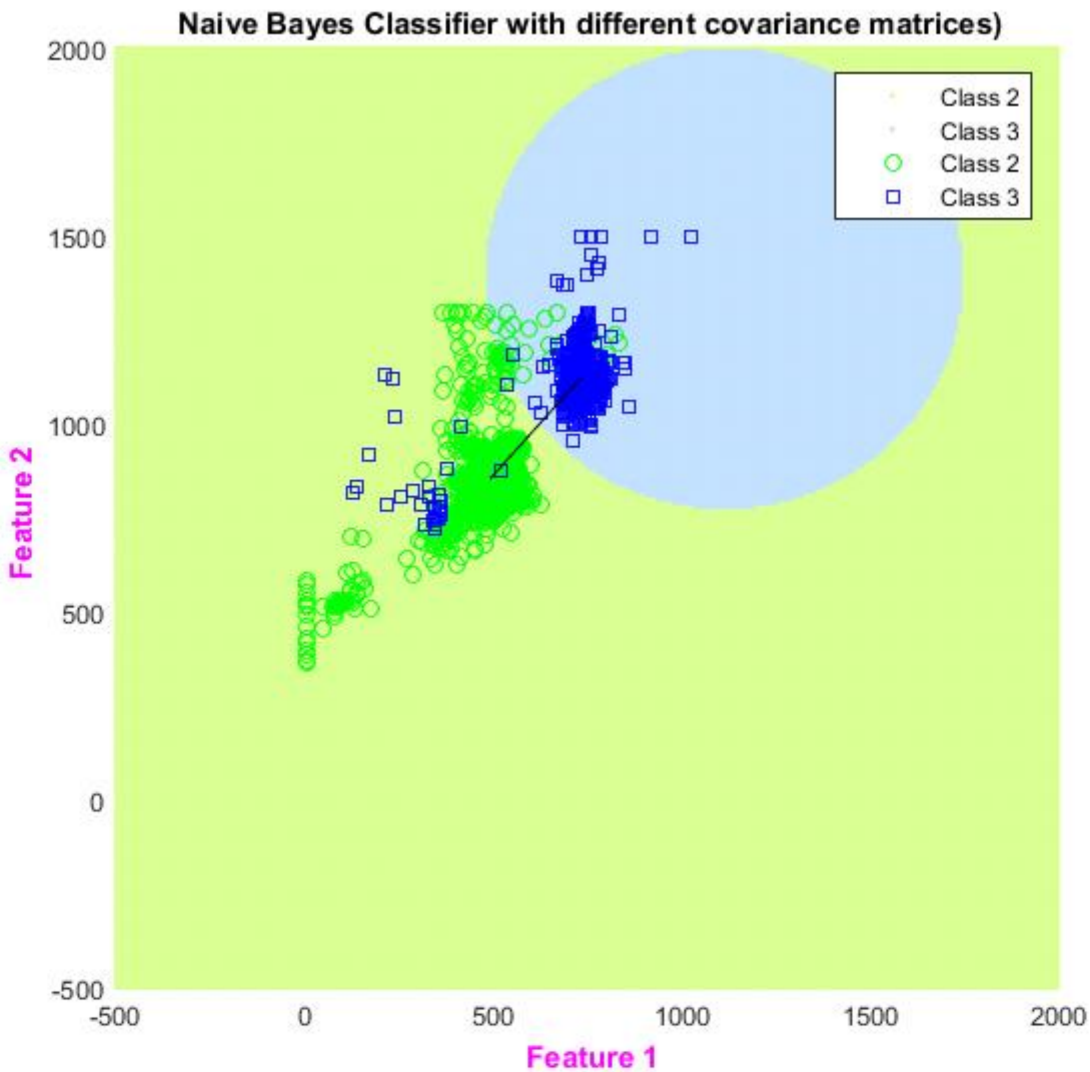
c) Covariance matrix for each class is different

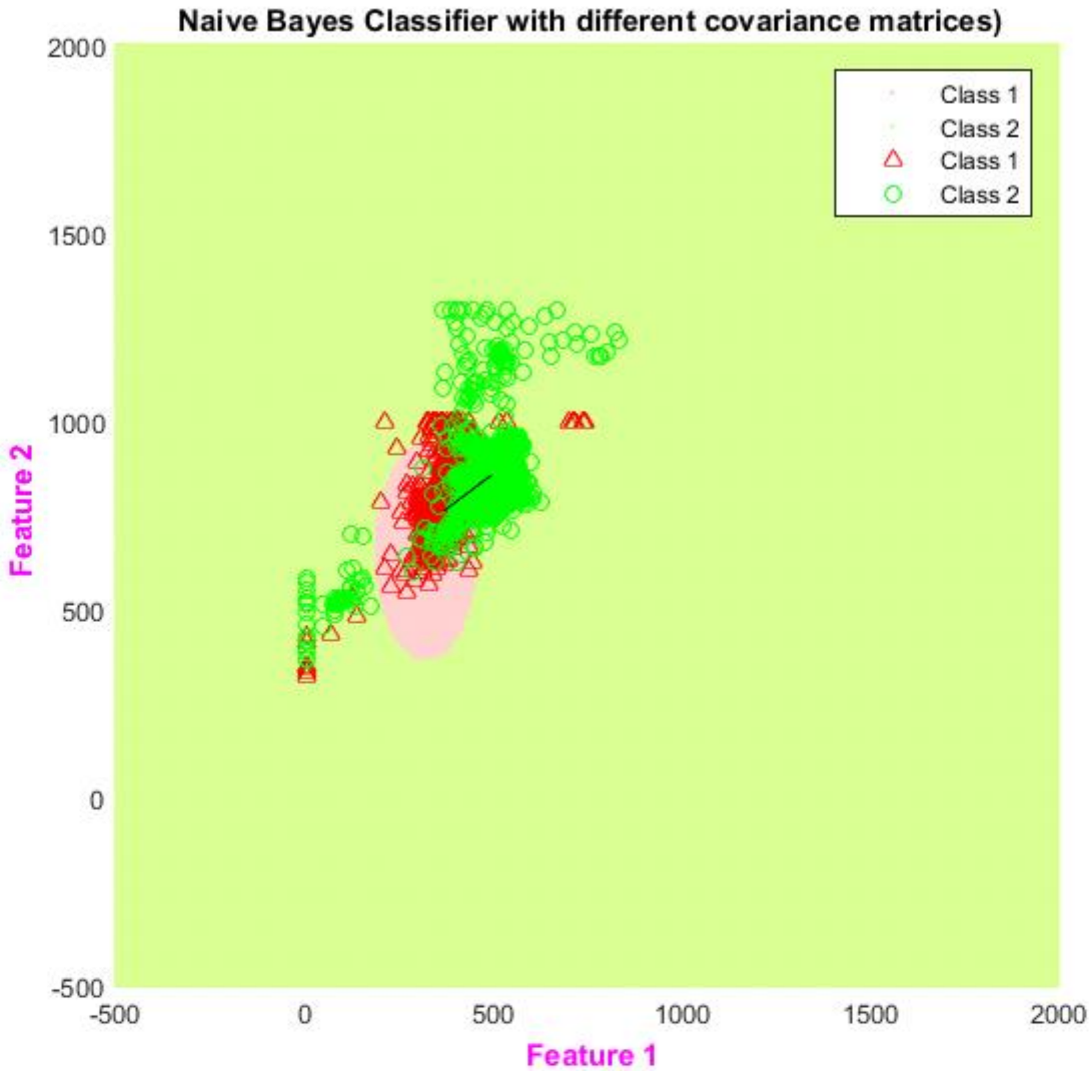
- *<u>Decision region plot for all the classes together with the training data superposed -</u>*



- *Decision region plot for every pair of classes -*




Naive Bayes Classifier with different covariance matrices)
Class 2
Class 3
Class 2
Class 3
Feature 1
Feature 2
2000
1500
1000
500
0
-500
-500
0
500
1000
1500
2000



- *<u>Confusion Matrix based on performance for test data-</u>*

Predicted Class ⇨	CLASS 1	CLASS 2	CLASS 3
Actual Class ⇩			
Class 1	546	69	7
Class 2	185	410	19
Class 3	15	16	510

- *<u>Classification accuracy on test data –</u>*
  Overall Accuracy – 82.4986
  Classifier Accuracy for class 1 – 87.7814
  Classifier Accuracy for class 2 – 66.7752
  Classifier Accuracy for class 3 – 94.2699

♦ Observations –

1) Bayes Classifier –

a) Same covariance matrices (by taking average)

The decision boundaries are all of **linear** nature. They pass through the **midpoint** of line joining the means of two classes. The boundaries are **not necessarily** perpendicular to this line.

b) Same covariance matrices (from training data of all classes combined)

The decision boundaries are all of **linear** nature. They pass through the **midpoint** of line joining the means of two classes. The boundaries are **not necessarily** perpendicular to this line.

c) Covariance matrix for each class is different

The decision boundaries are all of **Quadratic** nature. For this dataset they turn out to be either **Parabola** or **Ellipse**. They do **not necessarily** pass through the **midpoint** of line joining the means of two classes.

2) Naïve Bayes Classifier –

a) Same covariance matrices (σ^2 I)

The decision boundaries are all of **linear** nature. They pass through the **midpoint** of line joining the means of two classes. The boundaries are perpendicular to this line.

b) Same covariance matrices and is C

The decision boundaries are all of **linear** nature. They pass through the **midpoint** of line joining the means of two classes. The boundaries are **not necessarily** perpendicular to this line.

c) **Covariance matrix for each class is different**

The decision boundaries are all of **Quadratic** nature. For this dataset they turn out to be either **Parabola** or **Ellipse or Circle**. They do **not necessarily** pass through the **midpoint** of line joining the means of two classes. The classification results are more or less similar to that of Bayes.