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 GMM Bayes classifier for different cluster sizes and observe the results.

Data types provided –

- Non linearly separable data
  - Interlocking
  - Spiral
  - Ring
- Overlapping data
- Real world speech data
- Real world Image data

- All data except last contains 2/3 classes and has only two attributes.
- Last data has 3 classes. Each image has been divided into 36 segments and each of these segments has 23 attributes.
- Classifier to implement –
  - Bayes Classifier using GMM(Gaussian Mixture Modal)
    - Uses K-means Initialization
    - Initial points chosen randomly

Learning Objective:

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

- **Dataset I:** 2-dimensional artificial data of 2 classes that are **nonlinearly separable**

INTERLOCKING CLASSES –

1) Bayes Classifier using Gaussian Mixture Model –

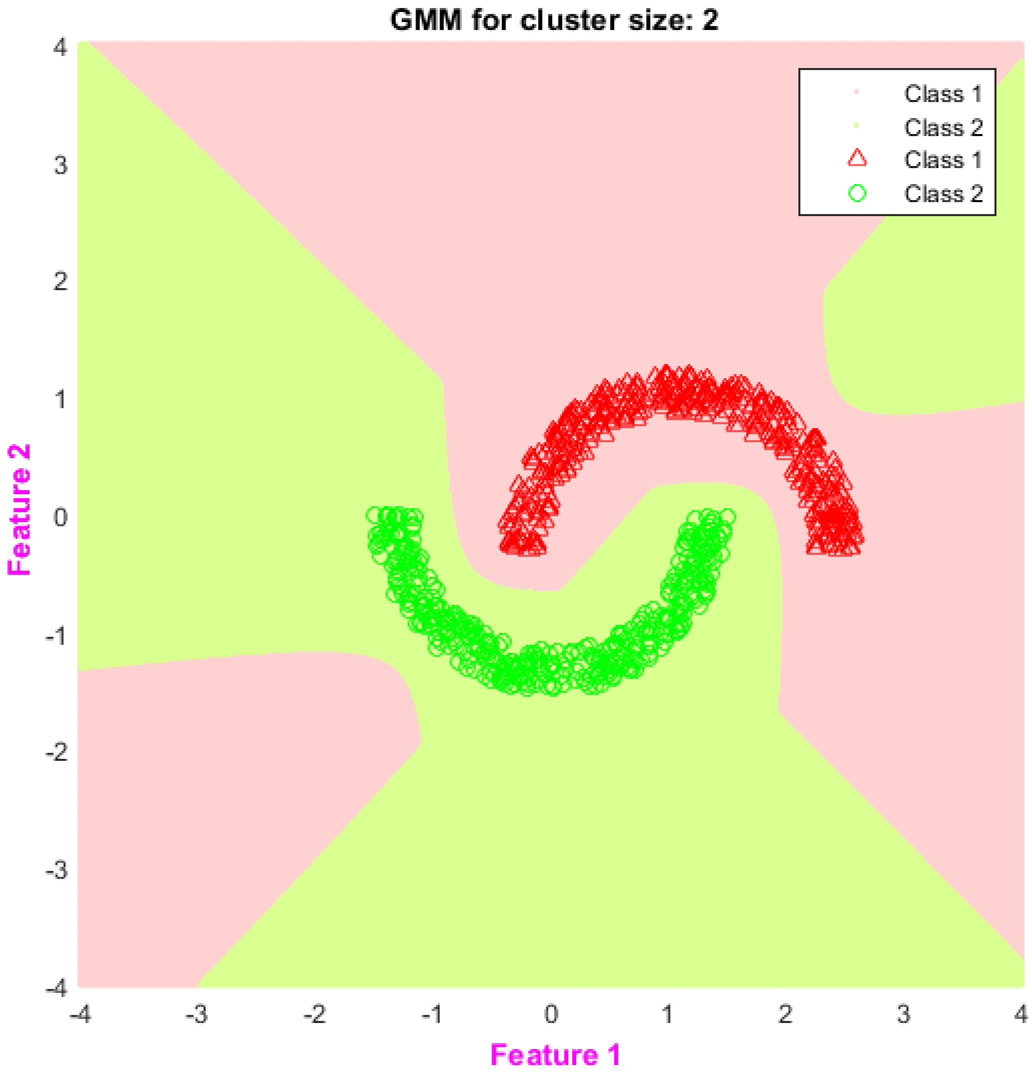

Fig -1a *Decision region plot for all the **interlocking** classes together with the training data superposed for cluster size **2***

Confusion Matrix based on performance for test data-

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

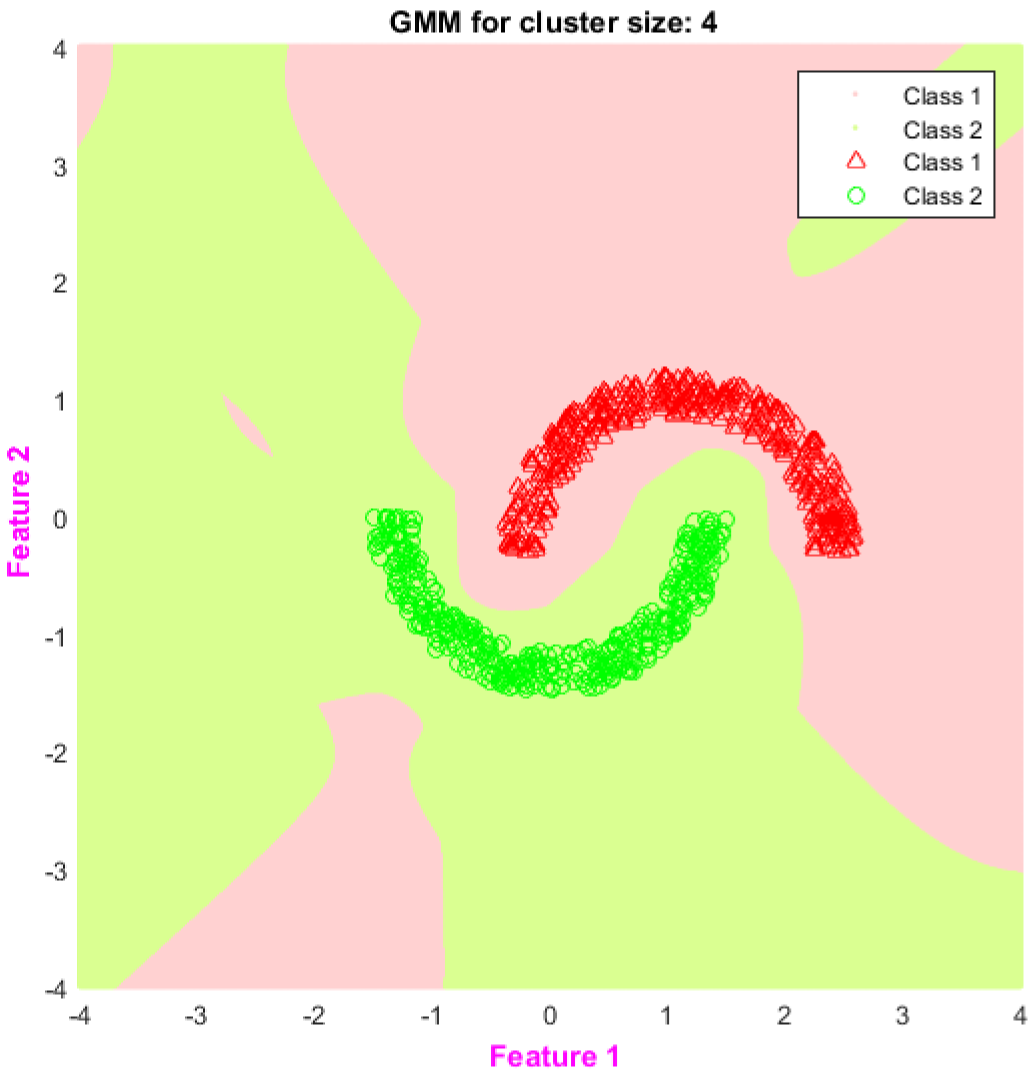

Fig -1b *Decision region plot for all the* ***interlocking*** *classes together with the training data superposed for cluster size* ***4***

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

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

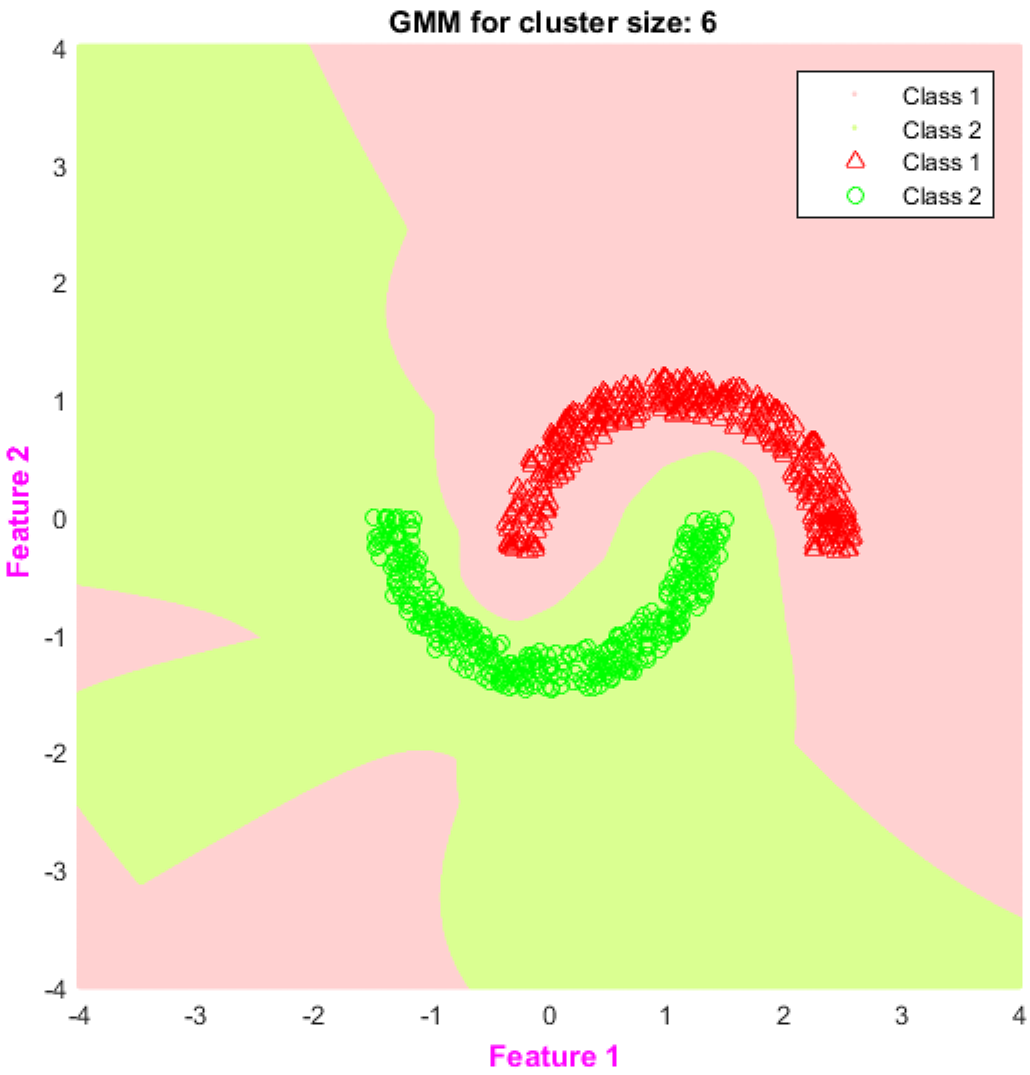

Fig -1c *Decision region plot for all the **interlocking** classes together with the training data superposed for cluster size **6***

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

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

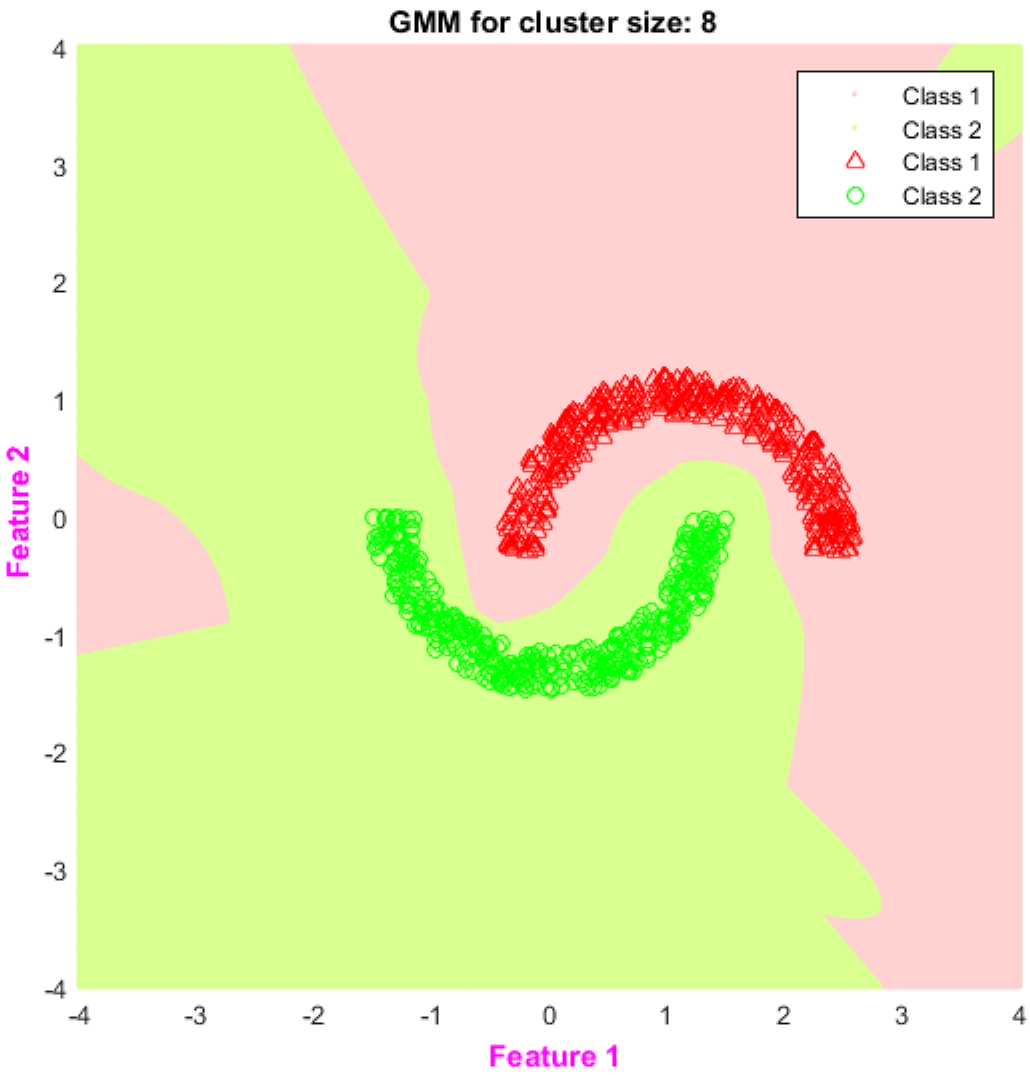

Fig -1d *Decision region plot for all the* ***interlocking*** *classes together with the training data superposed for cluster size* ***8***

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

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

Observations –

- Bayes classifier using Gaussian Mixture model was built and run for cluster sizes 2 to 10 of which, few plots of decision boundaries are shown above.
- The decision boundaries for all cluster sizes are not of any specific nature of linear or quadratic but superposition of several Gaussian curves.
- In case of Bayes and Naïve Bayes they were linear in case of same covariance matrices and quadratic in case of different covariance matrices.
- As far as the performance is concerned the maximum accuracy achieved for uni-modal case was 96%.
- Whereas in case of Bayes classifier using Gaussian mixture model here the maximum accuracy is 100% for all the cluster sizes.
- Accuracy is 100% even for cluster size 2 and remains same for all other cluster sizes.
- The possible reason for accuracy not changing on increasing cluster size is that, the data is **not overlapping** so even if the cluster size is increased, the overall likelihood of a point in a given class obtained by summation of product of pi(k)'s and gaussian distribution value with respect to parameters of that cluster.

A RING WITH A CENTRAL MASS –

1) Bayes Classifier using Gaussian Mixture Model –

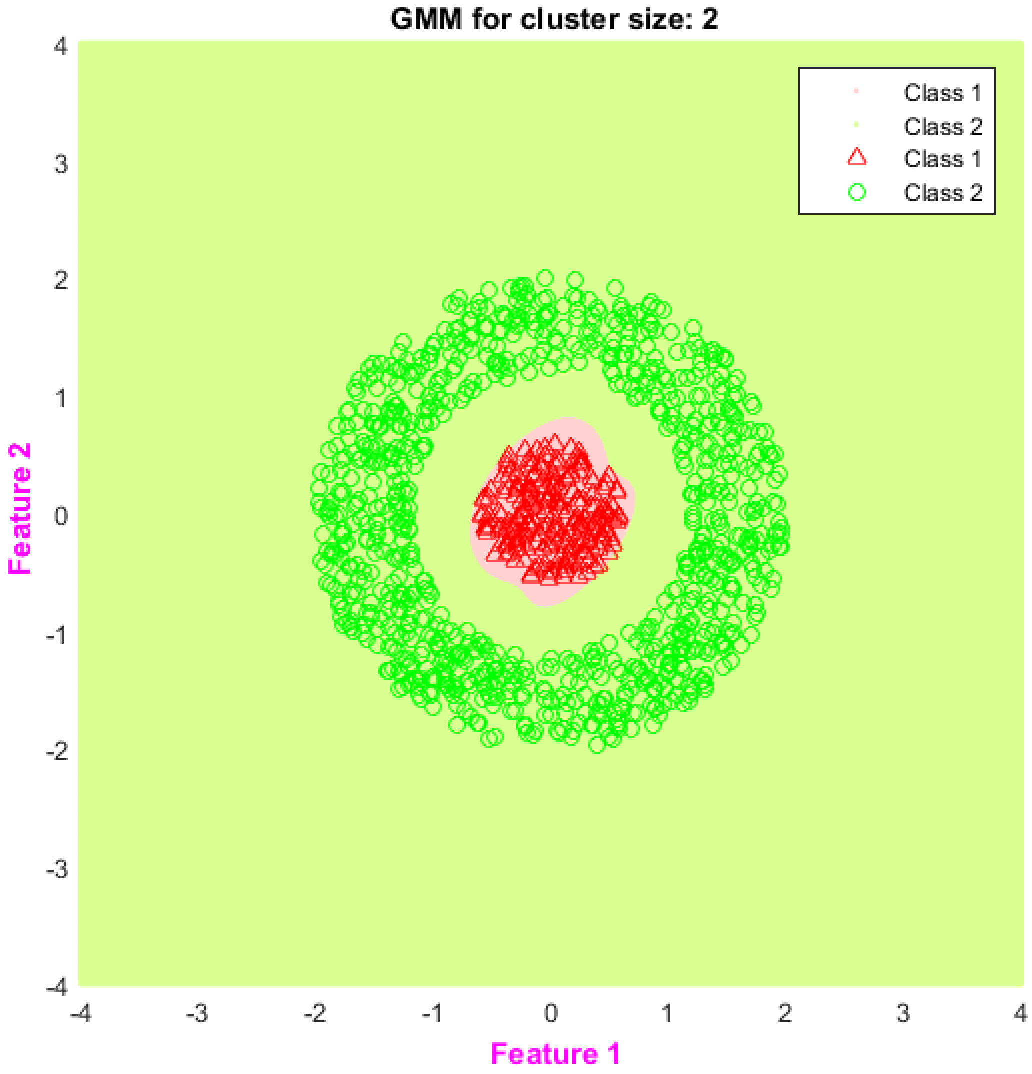

Fig -2a *Decision region plot for all the **ring** classes together with the training data superposed for cluster size **2***

- *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

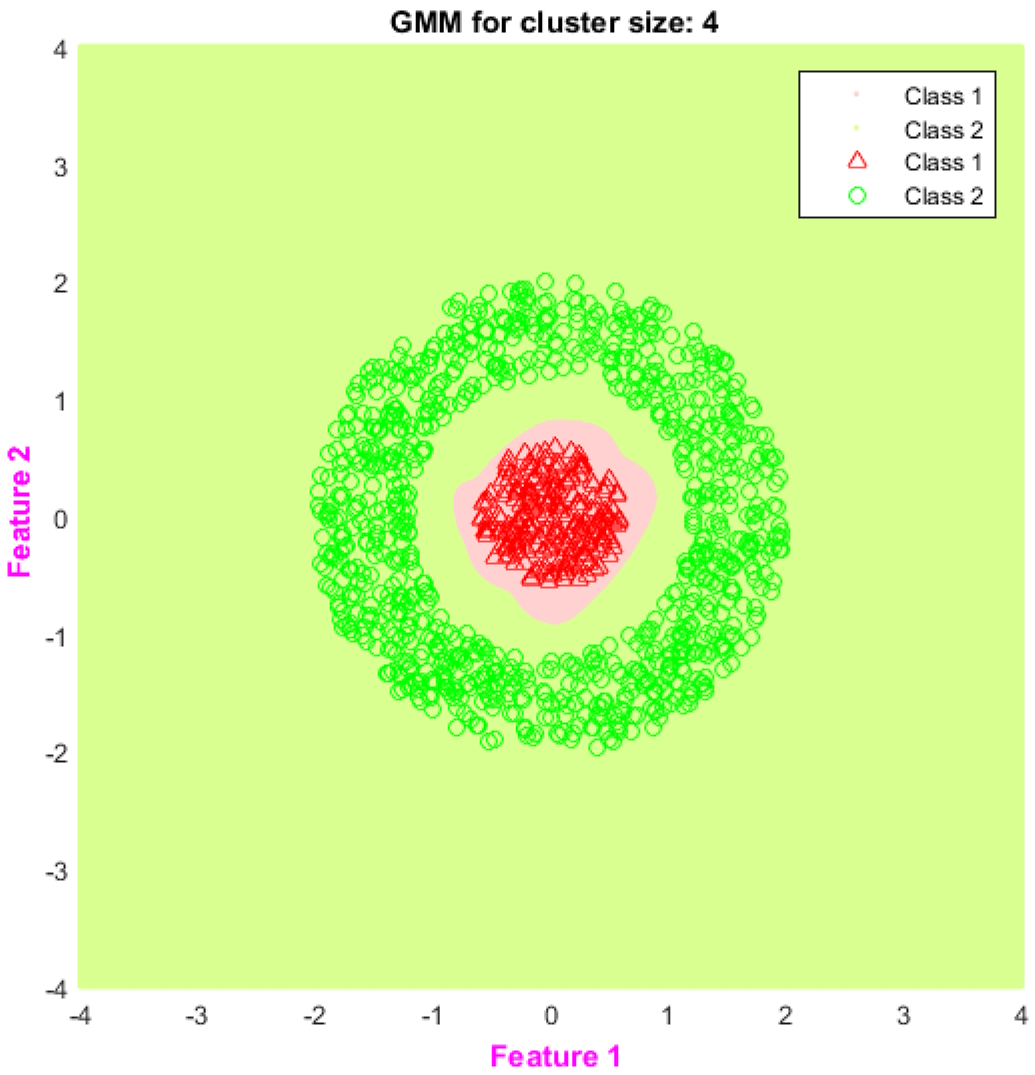

Fig -2b *Decision region plot for all the **ring** classes together with the training data superposed for cluster size **4***

- *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

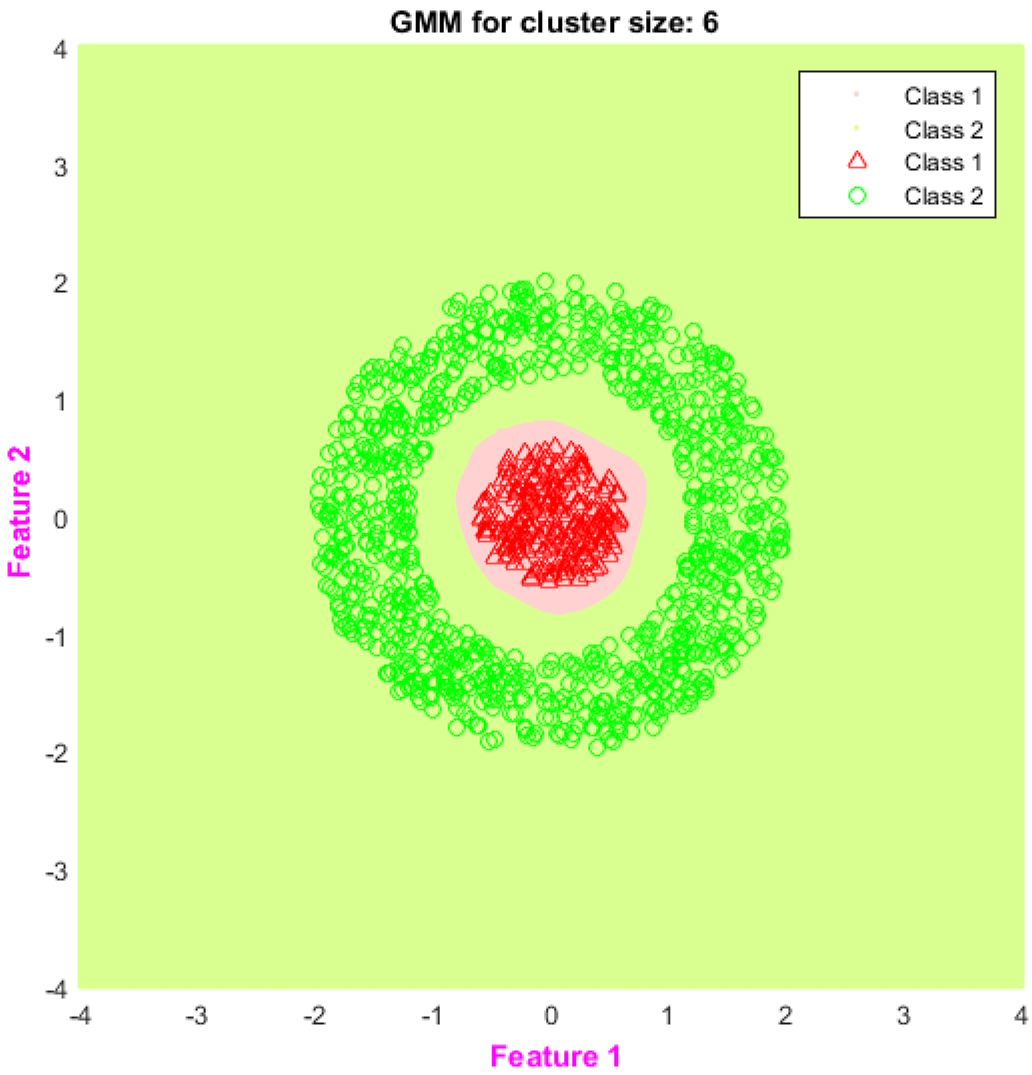

Fig -2c *<u>Decision region plot for all the **ring** classes together with the training data superposed for cluster size **6**</u>*

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

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

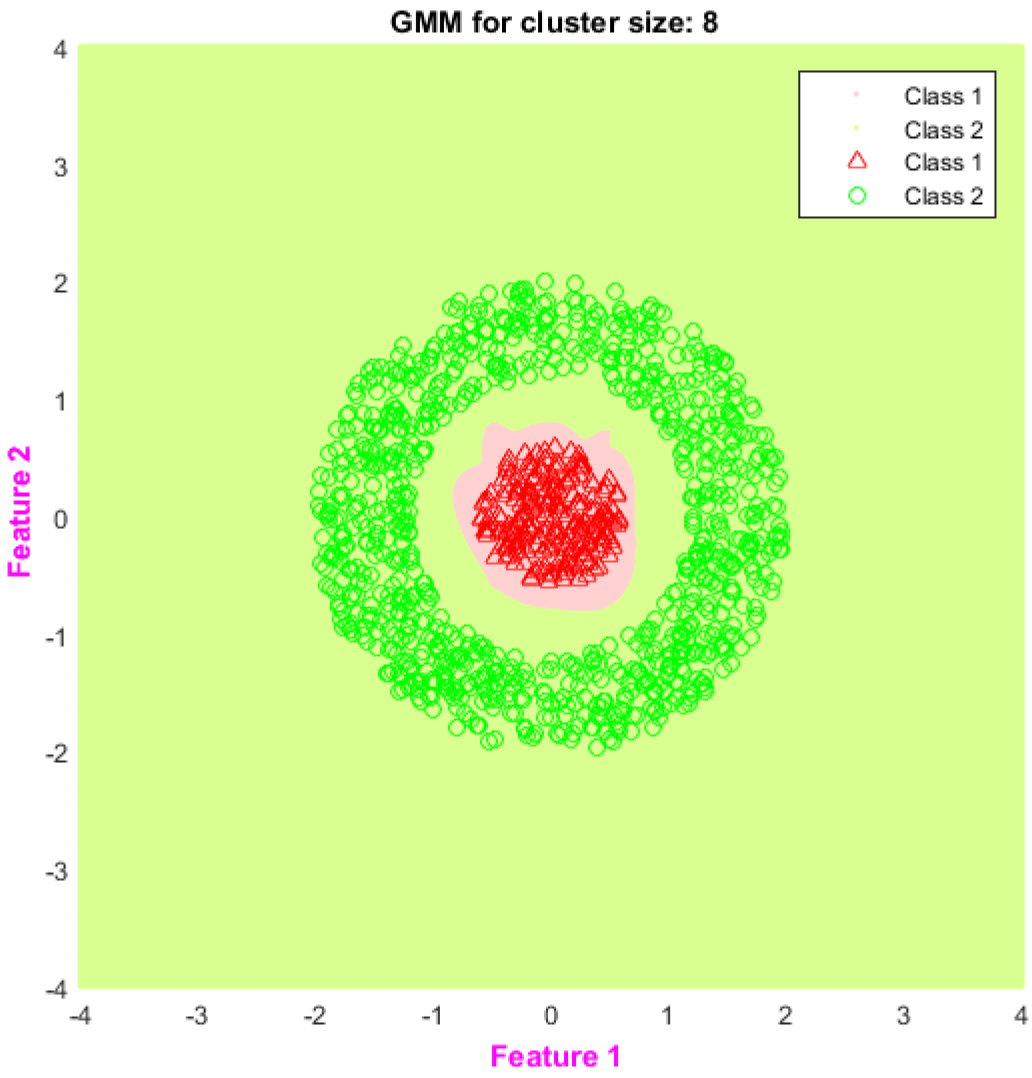

Fig -2d *Decision region plot for all the* ***ring*** *classes together with the training data superposed for cluster size* ***8***

- *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 –

- Bayes classifier using Gaussian Mixture model was built and run for cluster sizes 2 to 10 of which, few plots of decision boundaries are shown above.
- The decision boundaries for all cluster sizes are not of any specific nature of linear or quadratic but superposition of several Gaussian curves.
- In case of Bayes and Naïve Bayes decision boundaries were linear in case of same covariance matrices and circles in case of different covariance matrices.
- As far as the performance is concerned the maximum accuracy achieved for uni-modal case was 100%.
- Whereas in case of Bayes classifier using Gaussian mixture model, here the maximum accuracy is 100% for all the cluster sizes.
- Accuracy is 100% even for cluster size 2 and remains same for all other cluster sizes.
- The possible reason for accuracy not changing on increasing cluster size is that, the data is **not overlapping** so even if the cluster size is increased, the overall likelihood of a point in a given class obtained by summation of product of pi(k)’s and gaussian distribution value with respect to parameters of that cluster.

SPIRAL DATASET –

1) Bayes Classifier using Gaussian Mixture Model –

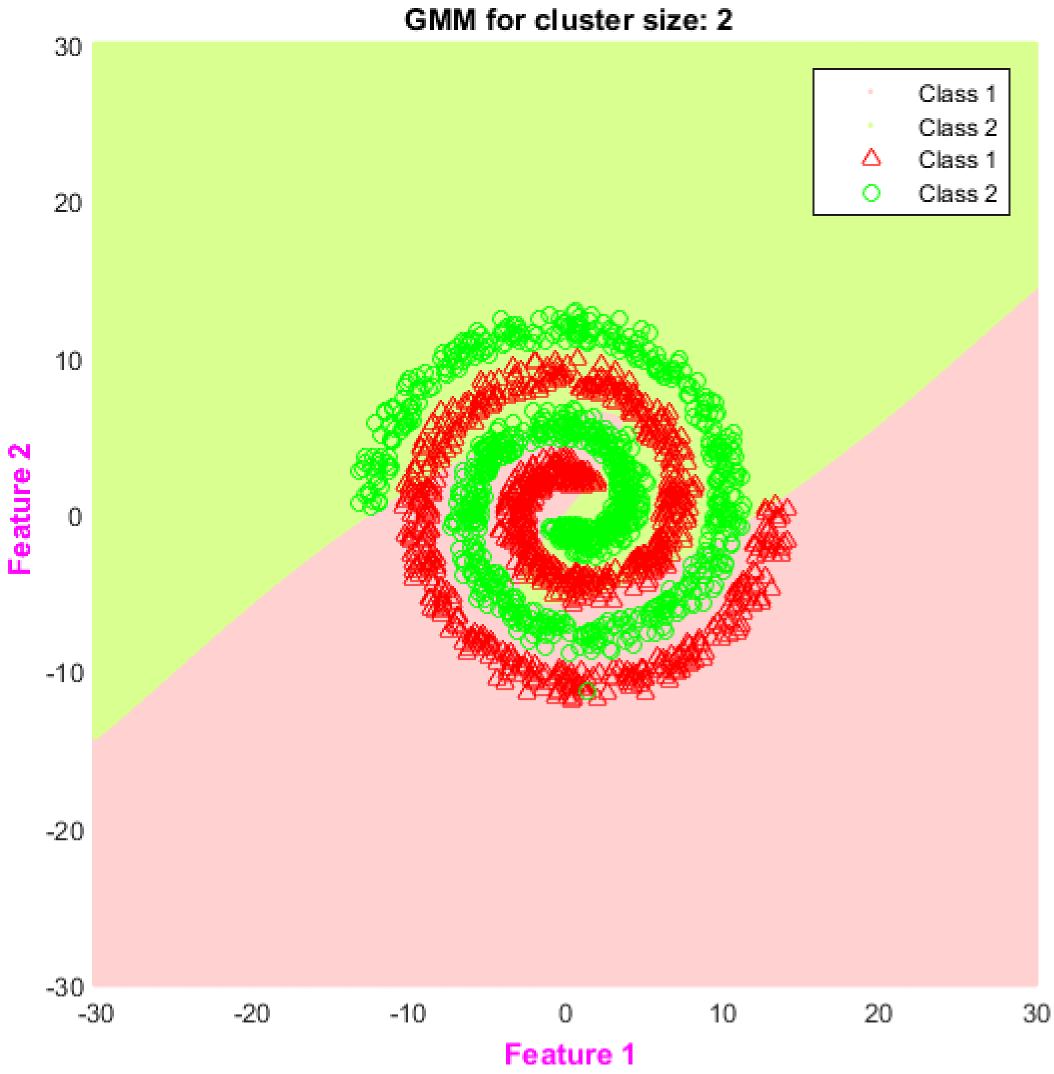

Fig -3a *Decision region plot for all the* ***spiral*** *classes together with the training data superposed for cluster size* ***2***

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

Predicted Class ⇨	CLASS 1	CLASS 2
Actual Class ⇩		
Class 1	200	126
Class 2	126	200

- *Classification accuracy on test data –*
  Overall Accuracy – 61.3497
  Classifier Accuracy for class 1 - 61.3497
  Classifier Accuracy for class 2 - 61.3497

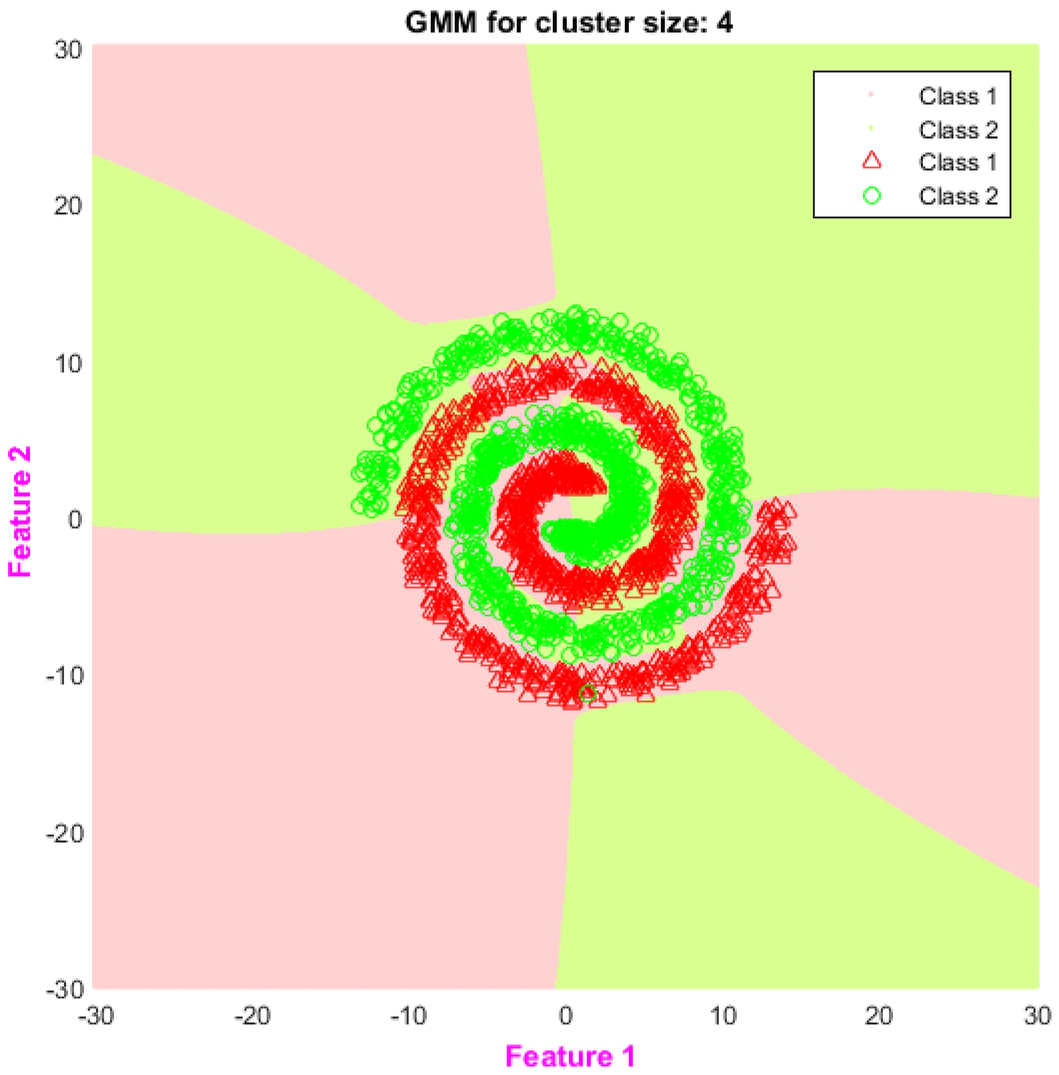

Fig -3b *Decision region plot for all the **spiral** classes together with the training data superposed for cluster size **4***

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

Predicted Class ⇨	CLASS 1	CLASS 2
Actual Class ⇩		
Class 1	220	106
Class 2	101	225

- *Classification accuracy on test data –*
  Overall Accuracy – 68.2515
  Classifier Accuracy for class 1 – 67.4847
  Classifier Accuracy for class 2 – 69.0184

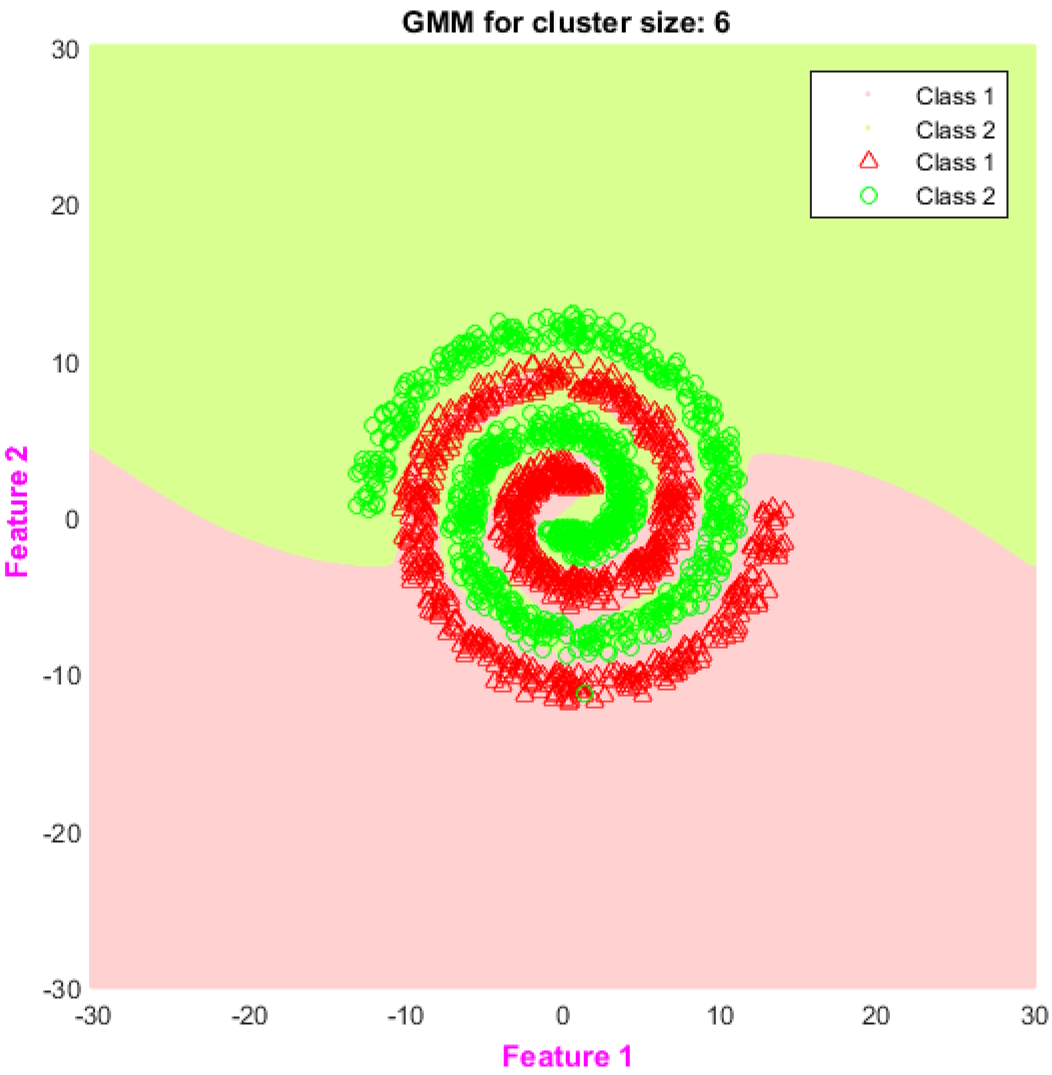

Fig -3c *Decision region plot for all the **spiral** classes together with the training data superposed for cluster size **6***

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

Predicted Class ⇨	CLASS 1	CLASS 2
Actual Class ⇩		
Class 1	296	30
Class 2	28	298

- *Classification accuracy on test data –*
  Overall Accuracy – 91.1043
  Classifier Accuracy for class 1 – 90.7975
  Classifier Accuracy for class 2 – 91.4110

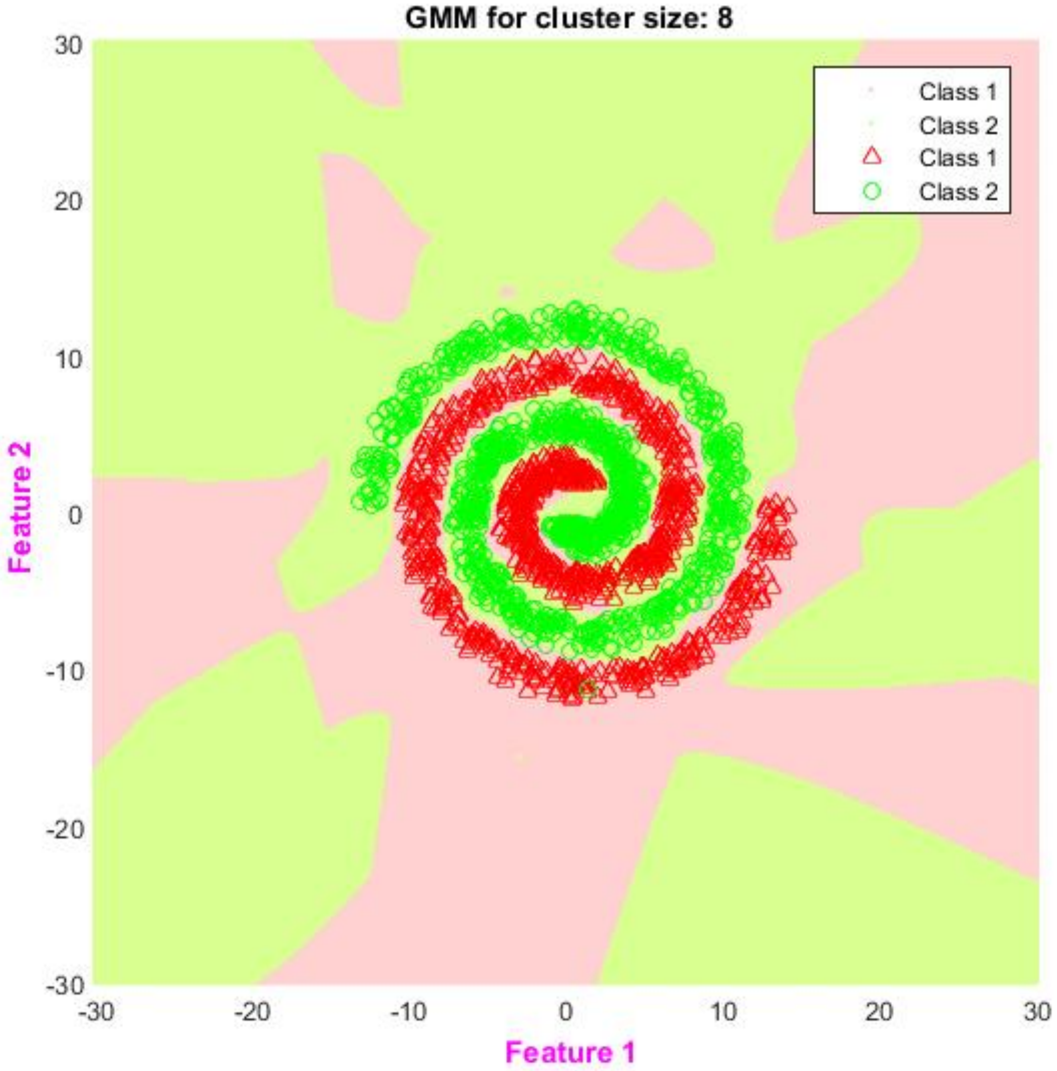

Fig -3d *Decision region plot for all the **spiral** classes together with the training data superposed for cluster size **8***

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

Predicted Class ⇨	CLASS 1	CLASS 2
Actual Class ⇩		
Class 1	300	17
Class 2	7	319

- *Classification accuracy on test data –*
  Overall Accuracy – 96.3190
  Classifier Accuracy for class 1 – 94.7853
  Classifier Accuracy for class 2 – 97.8528

♦ Observations –

- Bayes classifier using Gaussian Mixture model was built and run for cluster sizes 2 to 10 of which, few plots of decision boundaries are shown above.
- The decision boundaries for all cluster sizes are not of any specific nature of linear or quadratic but superposition of several Gaussian curves.
- In case of Bayes and Naïve Bayes decision boundaries were linear in case of same covariance matrices and parabola in case of different covariance matrices.
- As far as the performance is concerned the maximum accuracy achieved for uni-modal case was 54.29%.
- Whereas in case of Bayes classifier using Gaussian mixture model, here the maximum accuracy is 96.3% for all the cluster sizes.
- Accuracy is ~61% e for cluster size 2 and increases as the number of cluster increases. That is accuracy for all other cluster sizes greater than 2 is more than accuracy for cluster size 2.
- The reason for accuracy increasing on increasing the cluster size is that, the classes are intermingled in a spiral which can never be covered with a single Gaussian curve and even small cluster sizes are not sufficient to cover the whole data so as number of cluster increases the data is covered more effectively resulting in increase of accuracy for more cluster sizes.

ARTIFICIAL OVERLAPPING DATA OF 3 CLASSES -

1) Bayes Classifier using Gaussian Mixture Model –

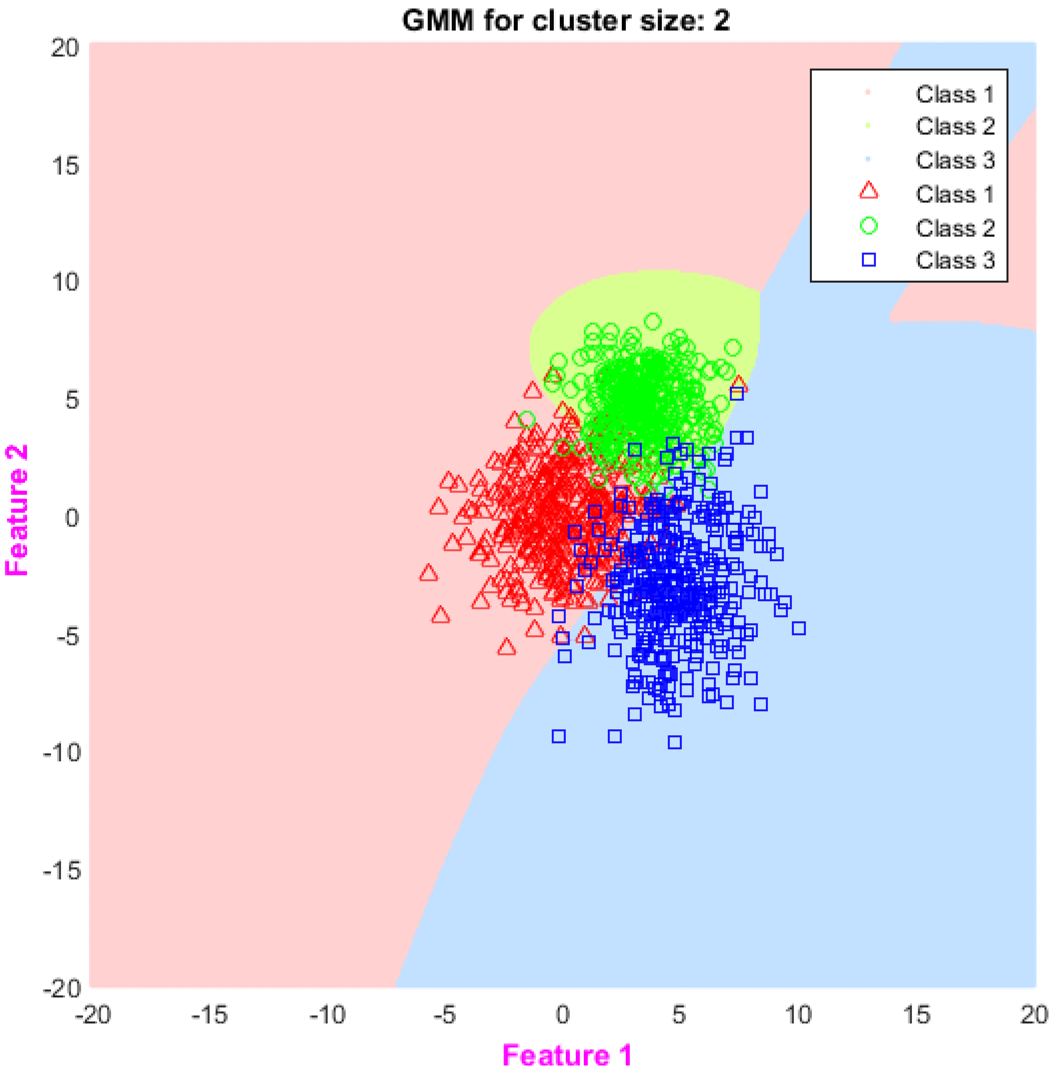

Fig -4a *<u>Decision region plot for all the **overlapping data** of 3 classes together with the training data superposed for cluster size **2**</u>*

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

Predicted Class ⇨	CLASS 1	CLASS 2	CLASS 3
Actual Class ⇩			
Class 1	112	10	3
Class 2	3	120	2
Class 3	8	2	115

- *Classification accuracy on test data –*
  Overall Accuracy – 92.5333
  Classifier Accuracy for class 1 – 89.6000
  Classifier Accuracy for class 2 – 96.0000
  Classifier Accuracy for class 3 – 92.0000

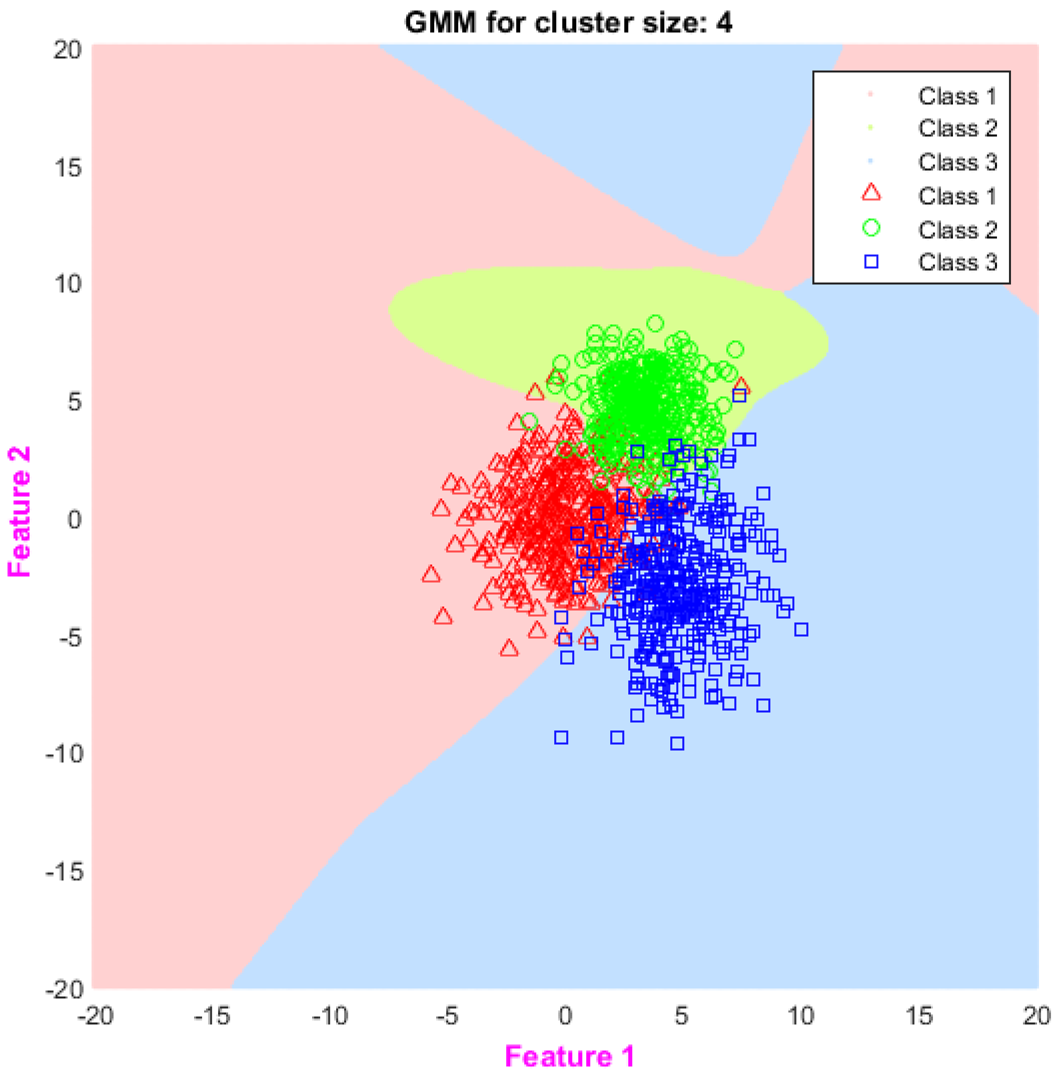

Fig -4b *Decision region plot for all the **overlapping data** of 3 classes together with the training data superposed for cluster size **4***

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

Predicted Class ⇨	CLASS 1	CLASS 2	CLASS 3
Actual Class ⇩			
Class 1	111	11	3
Class 2	4	120	1
Class 3	9	2	114

- *Classification accuracy on test data –*
  Overall Accuracy – 92.0000
  Classifier Accuracy for class 1 – 88.8000
  Classifier Accuracy for class 2 – 96.0000
  Classifier Accuracy for class 3 – 91.2000

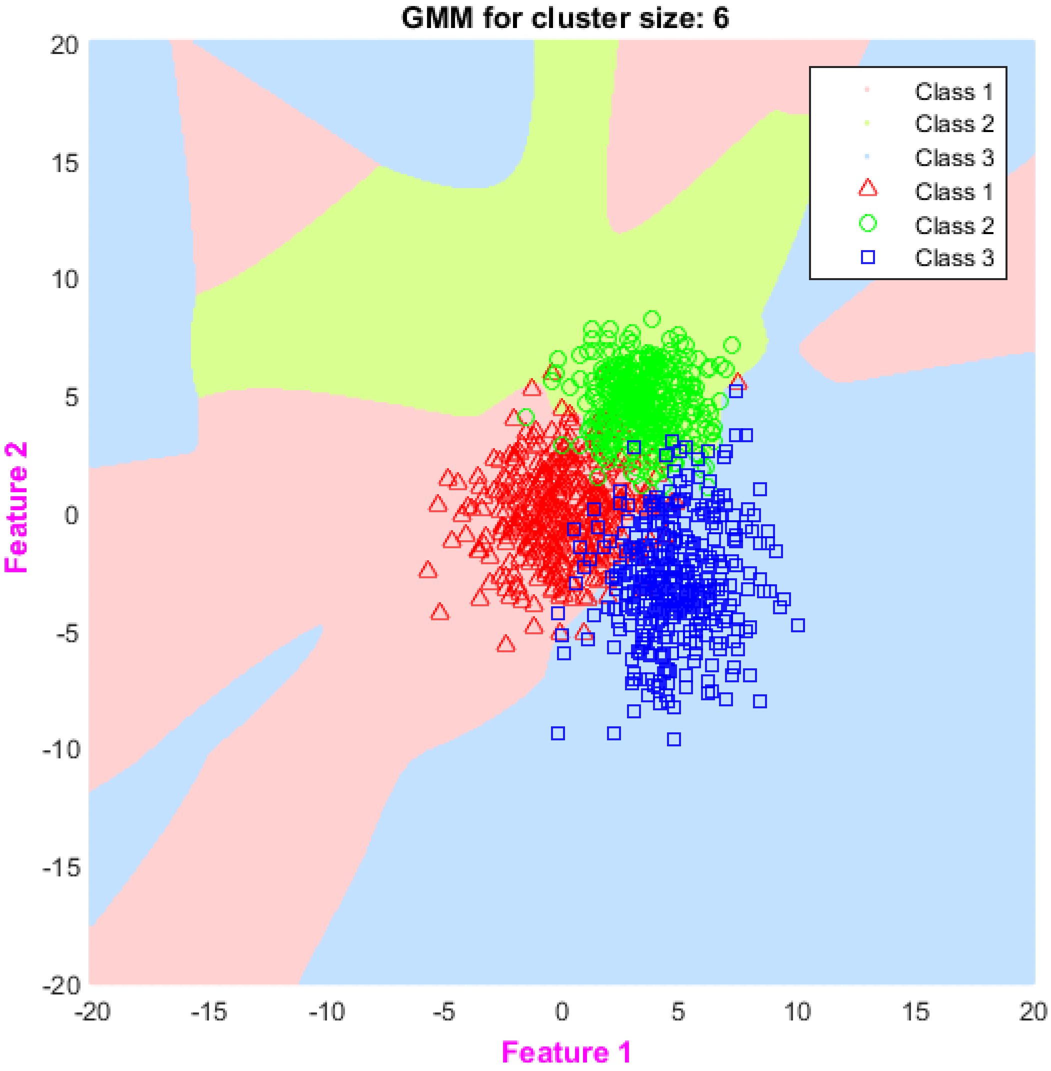

Fig -4c *<u>Decision region plot for all the **overlapping data** of 3 classes together with the training data superposed for cluster size **6**</u>*

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

Predicted Class ⇨	CLASS 1	CLASS 2	CLASS 3
Actual Class ⇩			
Class 1	112	10	3
Class 2	5	119	1
Class 3	10	2	113

- *Classification accuracy on test data –*
  Overall Accuracy – 91.7333
  Classifier Accuracy for class 1 – 89.6000
  Classifier Accuracy for class 2 – 95.2000
  Classifier Accuracy for class 3 – 90.4000

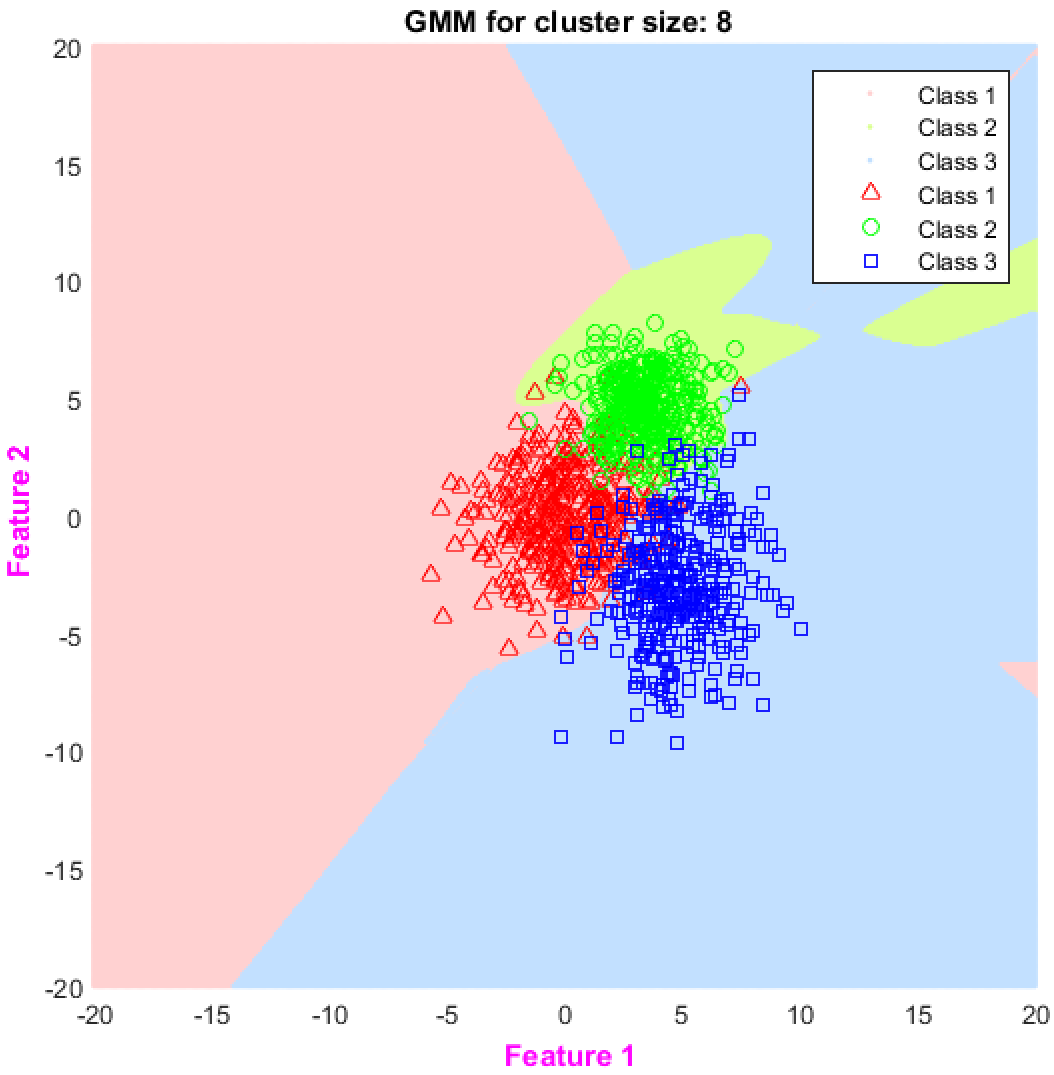

Fig -4d *Decision region plot for all the **overlapping data** of 3 classes together with the training data superposed for cluster size **8***

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

Predicted Class ⇨	CLASS 1	CLASS 2	CLASS 3
Actual Class ⇩			
Class 1	110	12	3
Class 2	5	118	2
Class 3	8	2	115

- *Classification accuracy on test data –*
  Overall Accuracy – 91.4667
  Classifier Accuracy for class 1 – 88.0000
  Classifier Accuracy for class 2 – 94.4000
  Classifier Accuracy for class 3 – 92.0000

- **Observations –**

- Bayes classifier using Gaussian Mixture model was built and run for cluster sizes 2 to 10 of which, few plots of decision boundaries are shown above.
- The decision boundaries for all cluster sizes are not of any specific nature of linear or quadratic but superposition of several Gaussian curves.
- In case of Bayes and Naïve Bayes decision boundaries were linear in case of same covariance matrices and circles in case of different covariance matrices.
- As far as the performance is concerned the maximum accuracy achieved for uni-modal case was 92.8%.
- Whereas in case of Bayes classifier using Gaussian mixture model, here the maximum accuracy is 92.55% for all the cluster sizes.
- Accuracy is 92.55% for cluster size 2 and very slight decrease is observed as the number of clusters is increased from 2 to 8.
- The possible reason for decrease in accuracy on increasing number of clusters is that, the data is **overlapping** so if the cluster size is increased, the overall likelihood of a point in a given class obtained by summation of product of pi(k)'s and Gaussian distribution value with respect to parameters of that cluster which can result in summation of small probability values if the point lies far away from mean of all the clusters of a that class .So the point may be classified to other class.

- ♦ **Dataset II (b):** 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 using Gaussian Mixture Model –

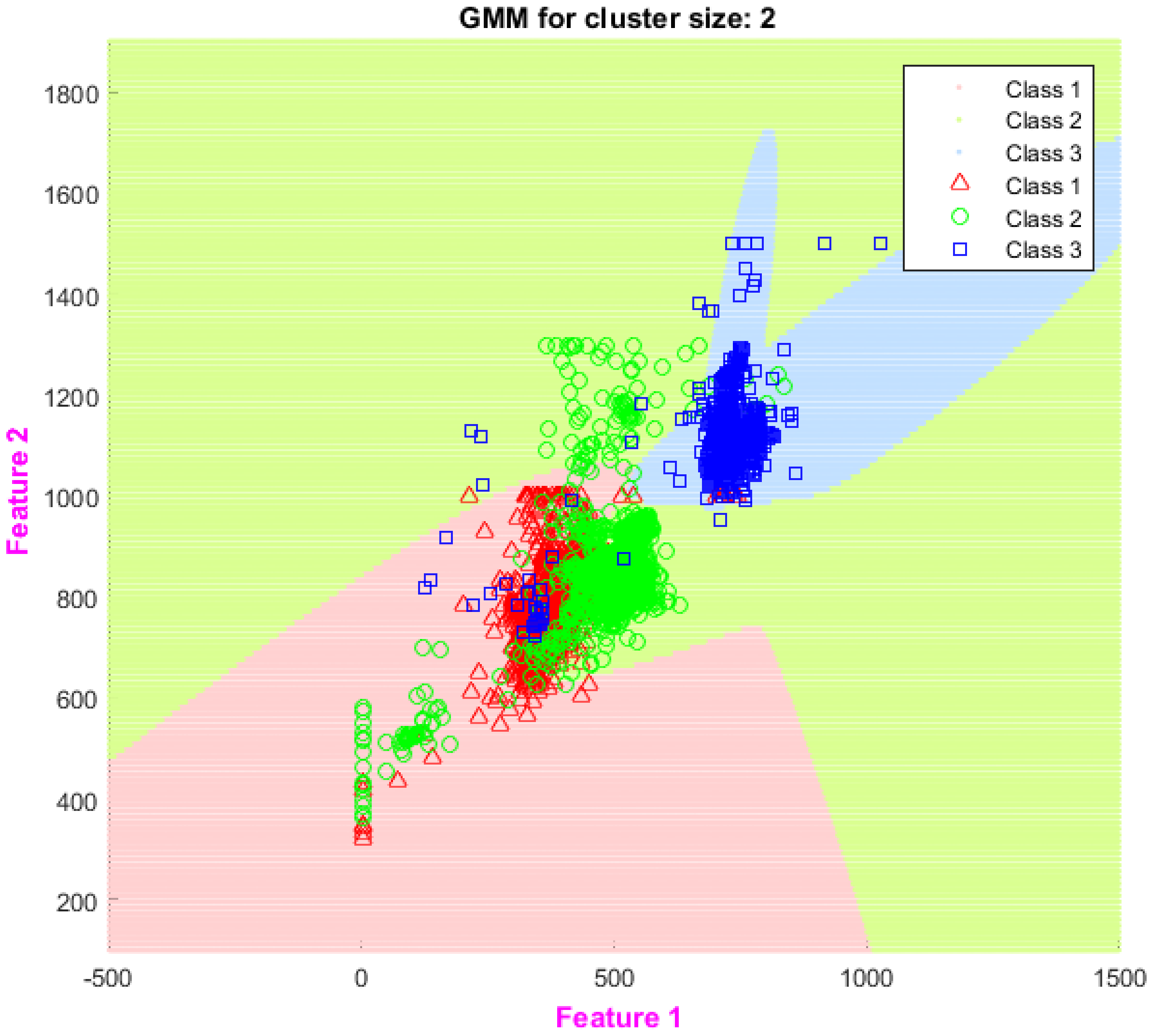

Fig -5a *Decision region plot for all the **real world data** of 3 classes together with the training data superposed for cluster size **2***

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

Predicted Class ⇨	CLASS 1	CLASS 2	CLASS 3
Actual Class ⇩			
Class 1	583	32	7
Class 2	135	477	2
Class 3	24	16	501

- *Classification accuracy on test data –*
  Overall Accuracy – 87.8447
  Classifier Accuracy for class 1 – 93.7299
  Classifier Accuracy for class 2 – 77.6873
  Classifier Accuracy for class 3 – 92.6063

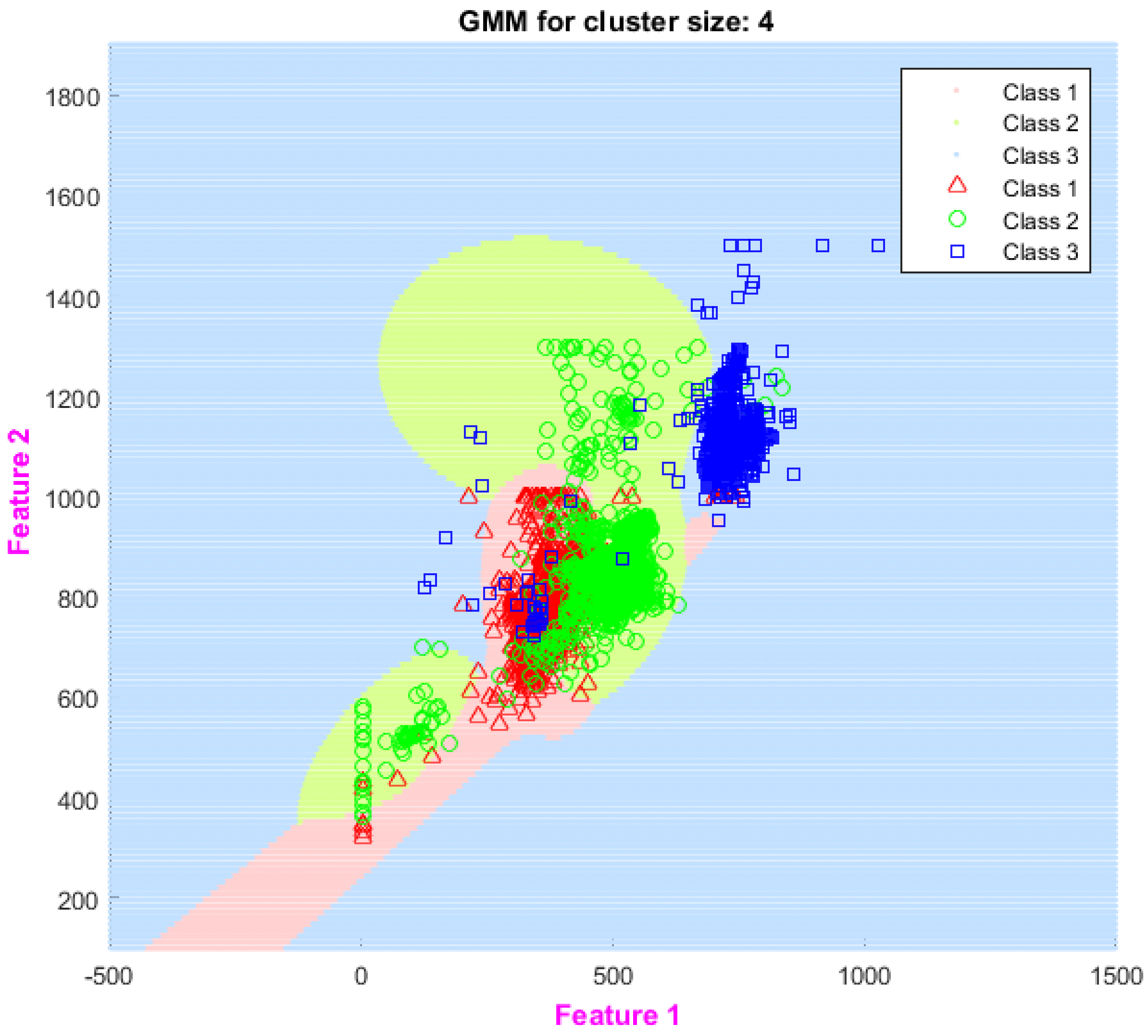

Fig -5b *Decision region plot for all the **real world data** of 3 classes together with the training data superposed for cluster size **4***

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

Predicted Class ⇨	CLASS 1	CLASS 2	CLASS 3
Actual Class ⇩			
Class 1	572	38	12
Class 2	236	375	3
Class 3	16	13	512

- *Classification accuracy on test data –*
  Overall Accuracy – 82.1047
  Classifier Accuracy for class 1 – 91.9614
  Classifier Accuracy for class 2 – 61.0749
  Classifier Accuracy for class 3 – 94.6396

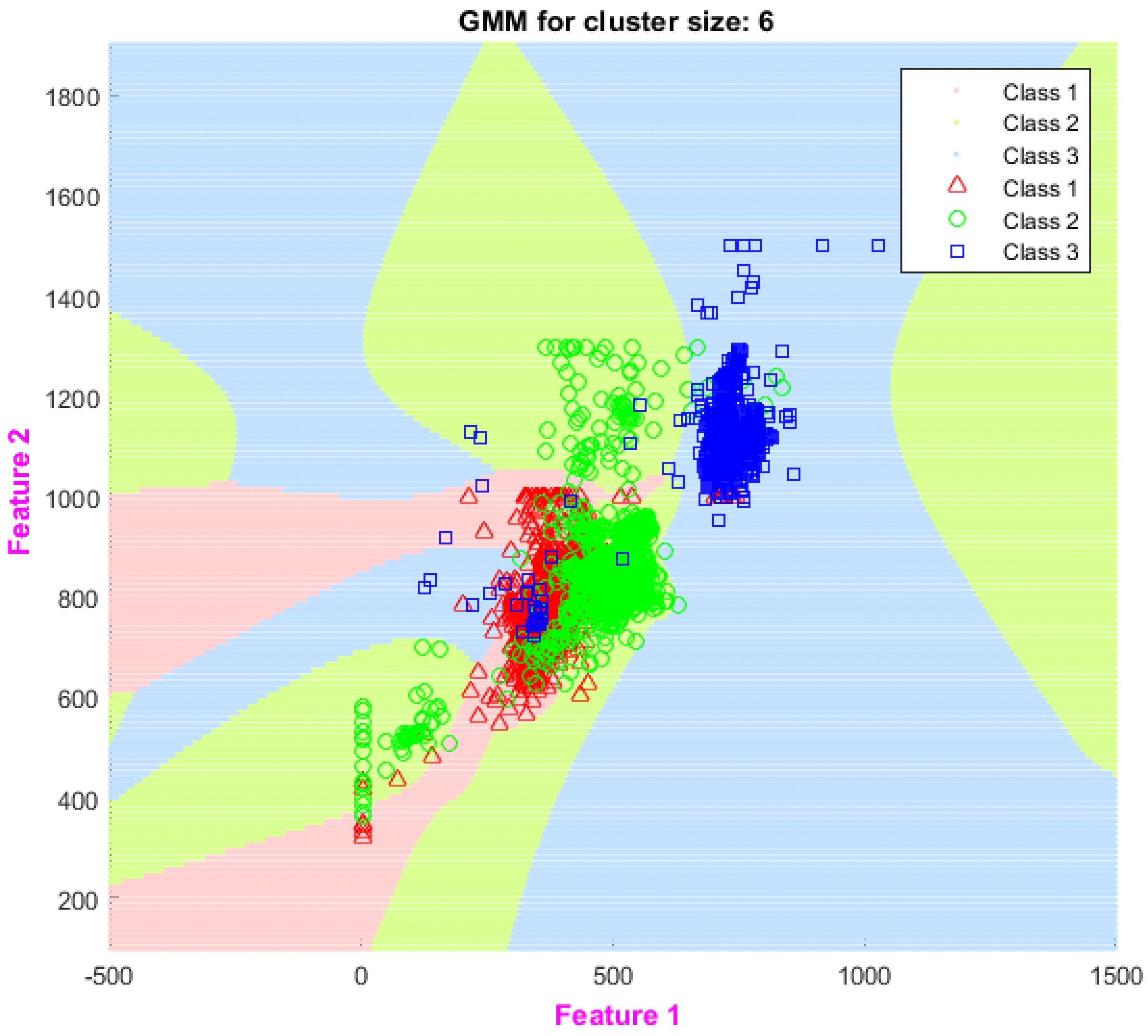

Fig -5c *Decision region plot for all the* ***real world data*** *of 3 classes together with the training data superposed for cluster size* ***6***

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

Predicted Class ⇨	CLASS 1	CLASS 2	CLASS 3
Actual Class ⇩			
Class 1	548	37	37
Class 2	209	402	3
Class 3	23	7	511

- *Classification accuracy on test data –*
  Overall Accuracy – 82.2172
  Classifier Accuracy for class 1 – 88.1029
  Classifier Accuracy for class 2 – 64.4723
  Classifier Accuracy for class 3 – 94.4547

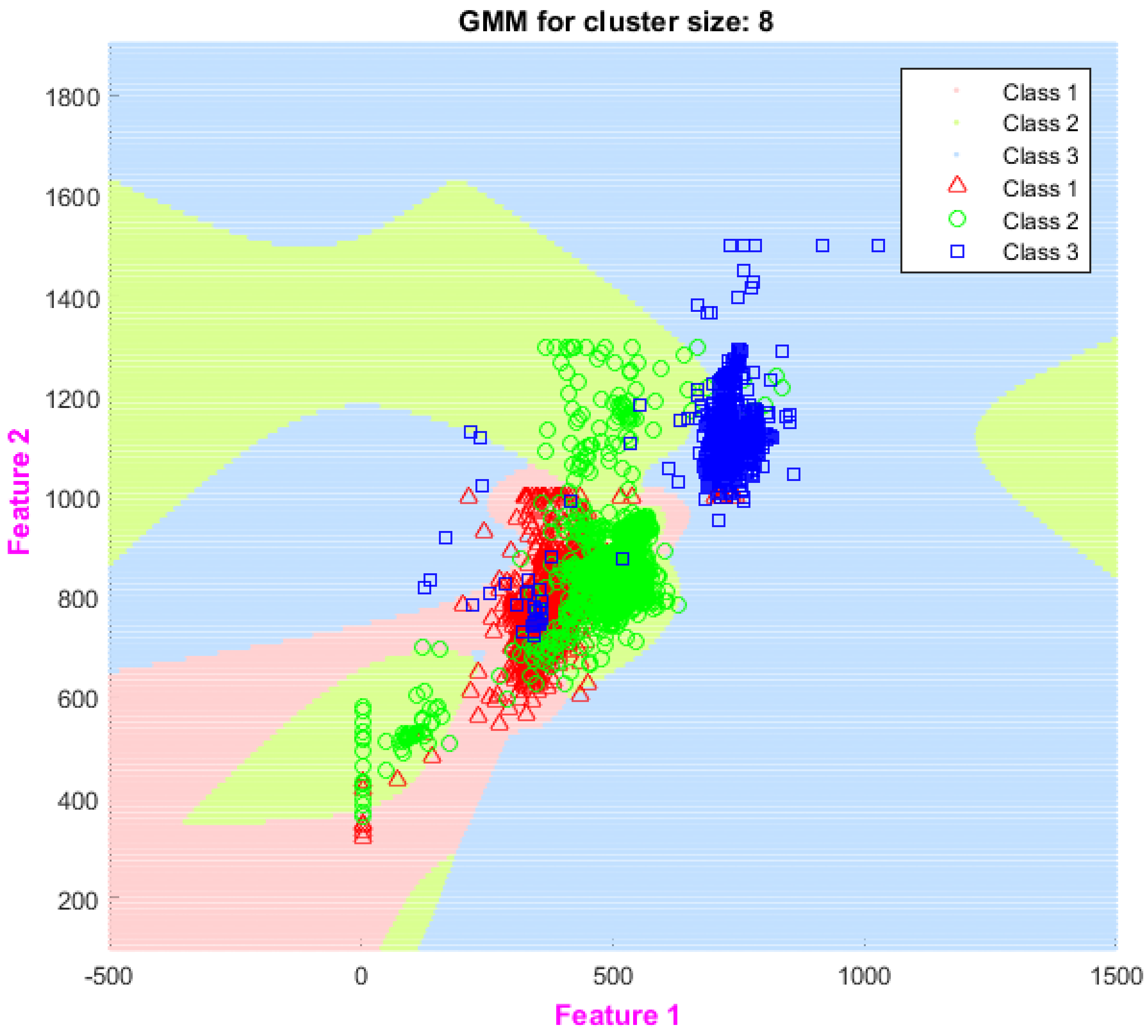

Fig -5d *Decision region plot for all the **real world data** of 3 classes together with the training data superposed for cluster size **8***

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

Predicted Class ⇨	CLASS 1	CLASS 2	CLASS 3
Actual Class ⇩			
Class 1	575	38	9
Class 2	195	418	1
Class 3	15	13	513

- *Classification accuracy on test data –*
  Overall Accuracy – 84.7496
  Classifier Accuracy for class 1 – 92.4437
  Classifier Accuracy for class 2 – 68.0782
  Classifier Accuracy for class 3 – 94.8244

- **Observations –**

- Bayes classifier using Gaussian Mixture model was built and run for cluster sizes 2 to 10 of which, few plots of decision boundaries are shown above.
- The decision boundaries for all cluster sizes are not of any specific nature of linear or quadratic but superposition of several Gaussian curves.
- In case of Bayes and Naïve Bayes decision boundaries were linear in case of same covariance matrices and circles in case of different covariance matrices.
- As far as the performance is concerned the maximum accuracy achieved for uni-modal case was 84.91%.
- Whereas in case of Bayes classifier using Gaussian mixture model, here the maximum accuracy is 87.8% for all the cluster sizes.
- Accuracy is 87.8% for cluster size 2 and no proper increase or decrease is observed upon increasing the number of clusters.
- The possible reason for no prominent increase/decrease in accuracy on increasing number of clusters is that, the data is **overlapping** so if the cluster size is increased, the overall likelihood of a point in a given class obtained by summation of product of pi(k)'s and Gaussian distribution value with respect to parameters of that cluster which can result in summation of small probability values if the point lies far away from mean of all the clusters of a that class .So the point may/may not be classified to other class.

- ♦ **Dataset II (c)** : Scene image data corresponding to 3 different classes
(A 23-dimensional feature vector is extracted from local blocks of an image for a particular scene. The 23-dimensional features include color histogram, edge directed histograms and entropy of wavelet coefficients. Each scene image is represented as a collection of 23-dimensional local feature vectors.)

FOR CLUSTER SIZE 2

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

Predicted Class ⇨	CLASS 1	CLASS 2	CLASS 3
Actual Class ⇩			
Class 1	40	19	6
Class 2	4	55	14
Class 3	27	16	46

- *Classification accuracy on test data –*
  Overall Accuracy – 62.1145
  Classifier Accuracy for class 1 – 61.5385
  Classifier Accuracy for class 2 – 75.3425
  Classifier Accuracy for class 3 – 51.6854

FOR CLUSTER SIZE 8

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

Predicted Class ⇨	CLASS 1	CLASS 2	CLASS 3
Actual Class ⇩			
Class 1	53	11	1
Class 2	5	58	10
Class 3	29	15	45

- *Classification accuracy on test data –*
  Overall Accuracy – 67.8414
  Classifier Accuracy for class 1 – 81.5385
  Classifier Accuracy for class 2 – 79.4521
  Classifier Accuracy for class 3 – 50.5618

FOR CLUSTER SIZE 16

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

Predicted Class ⇨	CLASS 1	CLASS 2	CLASS 3
Actual Class ⇩			
Class 1	47	15	3
Class 2	4	51	18
Class 3	10	08	71

- *Classification accuracy on test data –*
  Overall Accuracy – 74.4493
  Classifier Accuracy for class 1 – 72.3077
  Classifier Accuracy for class 2 – 69.8630
  Classifier Accuracy for class 3 – 79.7753

FOR CLUSTER SIZE 32

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

Predicted Class ⇨	CLASS 1	CLASS 2	CLASS 3
Actual Class ⇩			
Class 1	52	11	2
Class 2	7	50	16
Class 3	18	6	65

- *Classification accuracy on test data –*
  Overall Accuracy – 73.5683
  Classifier Accuracy for class 1 – 80.0000
  Classifier Accuracy for class 2 – 68.4932
  Classifier Accuracy for class 3 – 73.0337

FOR CLUSTER SIZE 64

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

Predicted Class ⇨	CLASS 1	CLASS 2	CLASS 3
Actual Class ⇩			
Class 1	56	9	0
Class 2	3	59	11
Class 3	20	9	60

- *Classification accuracy on test data –*
  Overall Accuracy – 77.0925
  Classifier Accuracy for class 1 – 86.1538
  Classifier Accuracy for class 2 – 80.8219
  Classifier Accuracy for class 3 – 67.4157