For this question what I did was implementin the EM algorithm using gaussian mixtures. For this purpose, I had to initialize some parameters which I had to use them during the training process. First I am going to explain the way I have written my code, then I am going to show the result based on tables and plots.

My algorithm contains four functions which I am going to explain them one by one. First function is the initialization part which is named "**__init__**" . In this function I initilize mean, covariance matrix and the prioir probabilty or alpha.

1. Initilizing mean :
   For initializing mean I try to devide my training data in to equal parts (for example if we have 3 Gaussian mixtures for that class, I am going to devide my training data in to 3 parts). Then I find the mean of that subset and it becoms my mean for that Guassian Distribution used in the code. I do the same thing for other classes and Gaussian distributions.
2. Initilizing Covariance matrix :
   For this one I try to find a covariance based on the calculated mean for each subset of training data.
3. Initializing prior probability or alpha :
   For this one I try to have equally probable Gaussians. So for example if I use 2 Gaussians to train my data for a class, I have alphas equal to 0.5.

For initializing one thing should be defind in my algorithm and it is which kind of covariance we are going to use. Another important thing is how I stop learning in this algorithm. I stop learning by using a threshold. If the log likelihood chanes with respect to the previous one is less than 0.0001 I stop learning in that point. So with this initialization we can have different number of iterations and they are not fixed.

Same spherical covariance matrix :

Below in Table 1, I have written my choice for same spherical covariance matrix case.

	Class 1	Class 2	Class 3
Number of Gaussians	5	4	3

Table 1 Number of gaussian

First calss Covariances :

$$\Sigma_{blue1} = \begin{pmatrix} 0.01008928 & 0 \\ 0 & 0.01006928 \end{pmatrix} \quad , \quad \Sigma_{blue\ 2} = \begin{pmatrix} 0.001886886 & 0 \\ 0 & 0.001886886 \end{pmatrix}$$

$$\Sigma_{blue\ 3} = \begin{pmatrix} 0.00109292 & 0 \\ 0 & 0.00109292 \end{pmatrix} \quad , \quad \Sigma_{blue\ 4} = \begin{pmatrix} 0.00266107 & 0 \\ 0 & 0.00266107 \end{pmatrix}$$

$$\Sigma_{blue\ 4} = \begin{pmatrix} 0.00094142 & 0 \\ 0 & 0.00094142 \end{pmatrix}$$

Second Class Covariance :

Sepehr Bakhshi

$$\Sigma_{red1} = \begin{pmatrix} 0.00161563 & 0 \\ 0 & 0.00161563 \end{pmatrix} \quad , \quad \Sigma_{red\,2} = \begin{pmatrix} 0.0014733 & 0 \\ 0 & 0.0014733 \end{pmatrix}$$

$$\Sigma_{red\,3} = \begin{pmatrix} 0.00251004 & 0 \\ 0 & 0.00251004 \end{pmatrix} \quad , \quad \Sigma_{red\,4} = \begin{pmatrix} 0.00161461 & 0 \\ 0 & 0.00161461 \end{pmatrix}$$

Third Class Covaraince :

$$\Sigma_{green1} = \begin{pmatrix} 0.00183658 & 0 \\ 0 & 0.00183658 \end{pmatrix} \quad , \quad \Sigma_{green\,2} = \begin{pmatrix} 0.00253923 & 0 \\ 0 & 0.00253923 \end{pmatrix}$$

$$\Sigma_{green\,3} = \begin{pmatrix} 0.00172039 & 0 \\ 0 & 0.00172039 \end{pmatrix}$$

Below I have written the mean I have found for the Gaussians. The first element in each row of matrix corresponds to mean of a feature vector in that gaussian.

First calss Mean :

$\hat{\mu}_{blue1} = (0.76713018 \quad 0.41302846)$, $\hat{\mu}_{blue2} = (0.33184061 \quad 0.83725206)$,

$\hat{\mu}_{blue3} = (0.27272166 \quad 0.68529588)$, $\hat{\mu}_{blue4} = (0.47705503 \quad 0.53565068)$,

$\hat{\mu}_{blue5} = (0.31829926 \quad 0.60421954)$

Second calss Mean :

$\hat{\mu}_{red1} = (0.40536321 \quad 0.85207882)$, $\hat{\mu}_{red2} = (0.47427513 \quad 0.47207211)$,

$\hat{\mu}_{red3} = (0.32193684 \quad 0.69636005)$, $\hat{\mu}_{red4} = (0.32242729 \quad 0.51805658)$

Third calss Mean :

$\hat{\mu}_{green1} = (0.86062781 \quad 0.20225708)$, $\hat{\mu}_{green2} = (0.6061375 \quad 0.42874682)$,

$\hat{\mu}_{green3} = (0.75290484 \quad 0.34415748)$

First Class Alpha(Prior Probability):

$\alpha_{blue1} = 0.4100$, $\alpha_{blue2} = 0.1437$ $\alpha_{blue3} = 0.1738$, $\alpha_{blue4} = 0.1764$,
$\alpha_{blue5} = 0.0959$

Second Class Alpha(Prior Probability):

$\alpha_{red1} = 0.2761$, $\alpha_{red2} = 0.2254$, $\alpha_{red3} = 0.2138$, $\alpha_{red4} = 0.2846$

Third Class Alpha(Prior Probability):

$$\alpha_{green1} = 0.3181\ ,\ \alpha_{green2} = 0.3307,\ \alpha_{green3} = 0.3510$$

In the Table 2 I have written tha last Log Likelihood and the number of iterations for this experiment on the EM Algorithm with spherical covariance matrix.

	Class 1 (blue)	Class 2 (red)	Class 3 (green)
Last Loglikelihood	748.9370	1123.6210	1210.7078
Number of Iterations	103	92	121

Table 2

Below in Table 3 I have wirtten the confusion matrix which shows the number of elements belonging to each class after the prediction for **<u>Training data</u>**.

		Predicted		
		Class1 (Blue)	Class2 (Red)	Class3 (Green)
Actual	Class1 (Blue)	320	78	102
	Class2 (Red)	64	424	12
	Class3 (Green)	13	24	463

Table 3 Confusion Matrix for Training data(Spherical covariance Matrix)

Below in Table 4 I have written the accuracy for each Guassian mixture in **<u>Training data</u>**.

	First Class Guassian mixture	Second Class Guassian mixture	Third Class Guassian mixture
Accuracy	0.640	0.848	0.926

Table 4 Accuracy Table for Training data (Spherical Covariance Matrix)

The overall accuracy in Training Data (Spherical Covariance Matrix) : 0.804

Below in Table 5 I have wirtten the confusion matrix which shows the number of elements belonging to each class after the prediction for **Test data**.

		Predicted		
		Class1 (Blue)	Class2 (Red)	Class3 (Green)
Actual	Class1 (Blue)	298	95	107
	Class2 (Red)	94	396	10
	Class3 (Green)	28	23	449

Table 5 Confusion Matrix for Test data(Spherical covariance Matrix)

Below in Table 6 I have written the accuracy for each Guassian mixture in **Test data**.

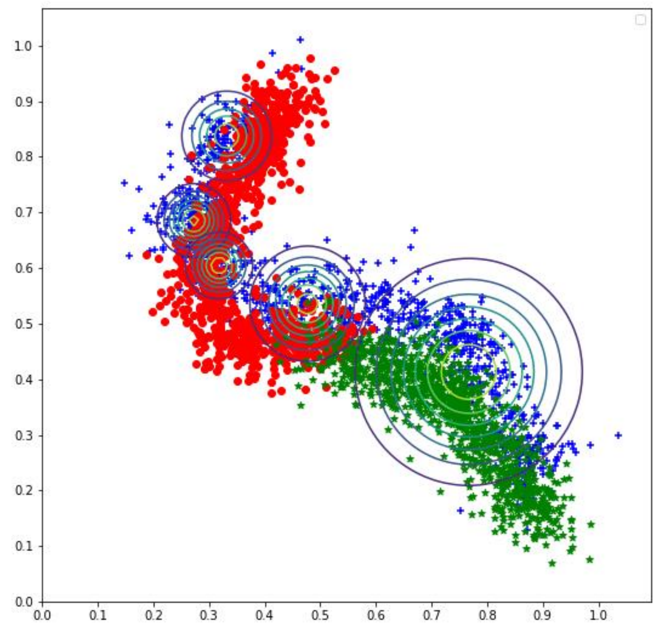
	First Class Guassian mixture	Second Class Guassian mixture	Third Class Guassian mixture
Accuracy	0.596	0.792	0.898

Table 6 Accuracy Table Fo Test data

The overall accuracy in Test Data (Spherical Covariance Matrix) : 0.762

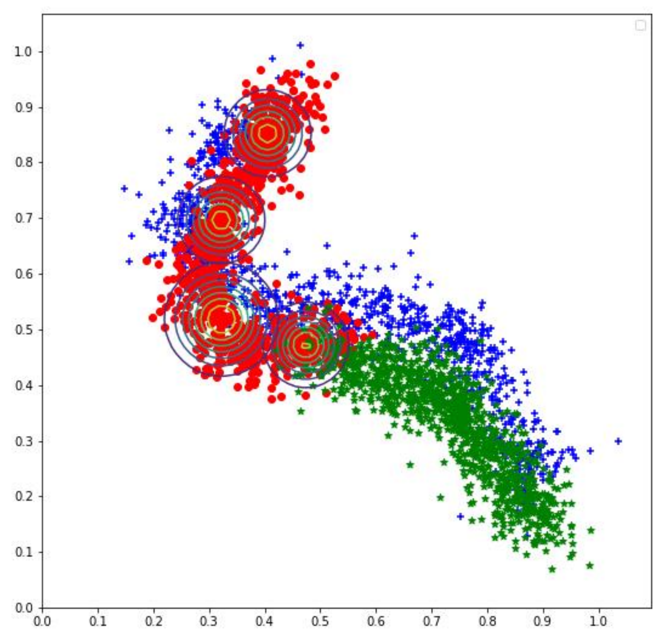
Below in Figure 1 , Figure 2, Figure 3 we see the result of plotting these Guassian mixtures with Spherical covariance on our data. Each figure is showing the result for one class.

Sepehr Bakhshi



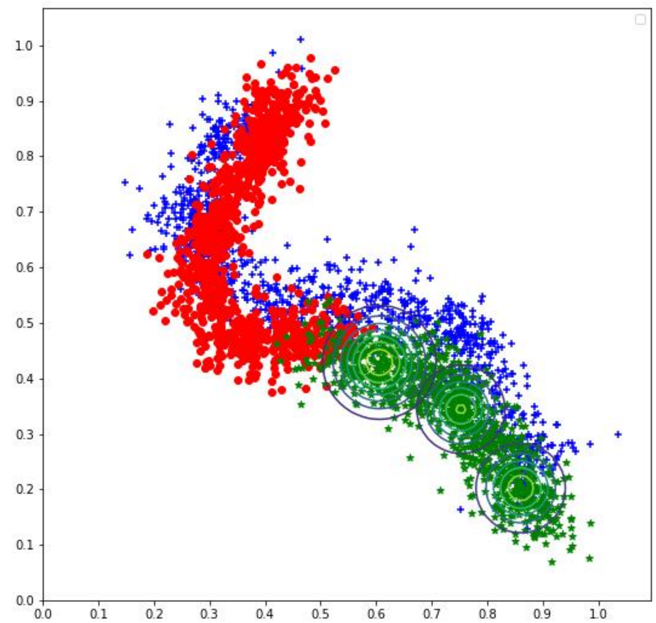
First class with spherical covariance matrix

Figure 1



Second class with spherical covariance matrix

Figure 2



Third class with spherical covariance matrix

Figure 3

Differenct diagonal covariance matrix :

Below in Table 7, I have written my choice for diagonal covariance matrix case.

	Class 1	Class 2	Class 3
Number of Gaussians	5	2	3

Table 7 Number of Guassians

First calss Covariances :

$$\Sigma_{blue1} = \begin{pmatrix} 0.00245285 & 0 \\ 0 & 0.00593946 \end{pmatrix} \quad , \quad \Sigma_{blue\ 2} = \begin{pmatrix} 0.01630953 & 0 \\ 0 & 0.00127038 \end{pmatrix}$$

$$\Sigma_{blue\ 3} = \begin{pmatrix} 0.00154051 & 0 \\ 0 & 0.00195092 \end{pmatrix} \quad , \quad \Sigma_{blue\ 4} = \begin{pmatrix} 0.00080201 & 0 \\ 0 & 0.00199278 \end{pmatrix}$$

$$\Sigma_{blue\ 4} = \begin{pmatrix} 0.00080279 & 0 \\ 0 & 0.00096472 \end{pmatrix}$$

Second Class Covariance :

$$\Sigma_{red1} = \begin{pmatrix} 0.00332072 & 0 \\ 0 & 0.01657487 \end{pmatrix} \quad , \quad \Sigma_{red\ 2} = \begin{pmatrix} 0.00590115 & 0 \\ 0 & 0.00075463 \end{pmatrix}$$

Third Class Covaraince :

$$\Sigma_{green1} = \begin{pmatrix} 0.00130356 & 0 \\ 0 & 0.00217451 \end{pmatrix} \quad , \quad \Sigma_{green\ 2} = \begin{pmatrix} 0.00504687 & 0 \\ 0 & 0.00107789 \end{pmatrix}$$

$$\Sigma_{green\ 3} = \begin{pmatrix} 0.00087436 & 0 \\ 0 & 0.00169517 \end{pmatrix}$$

Below I have written the mean I have found for the Gaussians in Diagonal mode. The first element in each row of matrix corresponds to mean of a feature vector in that gaussian.

First calss Mean :

$\hat{\mu}_{blue1} = (0.8451913 \quad 0.31696693)$, $\hat{\mu}_{blue2} = (0.59802521 \quad 0.51512528)$,

$\hat{\mu}_{blue3} = (0.33511543 \quad 0.84066231)$, $\hat{\mu}_{blue4} = (0.27452504 \quad 0.68285776)$,

$\hat{\mu}_{blue5} = (0.32705805 \quad 0.60070352)$

Second calss Mean :

$\hat{\mu}_{red1} = (0.3537664 \quad 0.72846768)$, $\hat{\mu}_{red2} = (0.42600135 \quad 0.47414327)$,

Third calss Mean :

$\hat{\mu}_{green1} = (0.86315467 \quad 0.20055065)$, $\hat{\mu}_{green2} = (0.62409755 \quad 0.42345349)$,

$\hat{\mu}_{green3} = (0.76398794 \quad 0.3284692)$

First Class Alpha(Prior Probability):

α_{blue1}= 0.1931 , α_{blue2}= 0.3939 α_{blue3}= 0.1356 , α_{blue4}= 0.2020, α_{blue5}= 0.0752

Second Class Alpha(Prior Probability):

α_{red1}= 0.6444 , α_{red2}= 0.3555,

Third Class Alpha(Prior Probability):

α_{green1}= 0.3086 , α_{green2}= 0.4000, α_{green3}= 0.2913

In the Table 8 I have written tha last Log Likelihood and the number of iterations for this experiment on the EM Algorithm with Diagonal covariance matrix.

	Class 1 (blue)	Class 2 (red)	Class 3 (green)
Last Loglikelihood	880.0421	988.2451	1269.8494
Number of Iterations	111	36	62

Table 8

Table 9

Below in Table 9 I have wirtten the confusion matrix which shows the number of elements belonging to each class after the prediction for **Training data**.

		Predicted		
		Class1 (Blue)	Class2 (Red)	Class3 (Green)
Actual	Class1 (Blue)	328	93	79
	Class2 (Red)	101	391	8
	Class3 (Green)	15	15	470

Table 10 Confusion Matrix for Training data(Diagonal covariance Matrix)

Below in Table 11 I have written the accuracy for each Guassian mixture in **Training data**.

	First Class Guassian mixture	Second Class Guassian mixture	Third Class Guassian mixture
Accuracy	0.656	0.782	0.94

Table 11 Accuracy Table for Training data (Diagonal Covariance Matrix)

The overall accuracy in Training Data (Diagonal Covariance Matrix) : 0.792

Below in Table 12 I have wirtten the confusion matrix which shows the number of elements belonging to each class after the prediction for **Test data**.

		Predicted		
Actual		Class1 (Blue)	Class2 (Red)	Class3 (Green)
	Class1 (Blue)	323	103	74
	Class2 (Red)	110	383	7
	Class3 (Green)	34	17	449

Table 12 Confusion Matrix for Test data(Diagonal covariance Matrix)

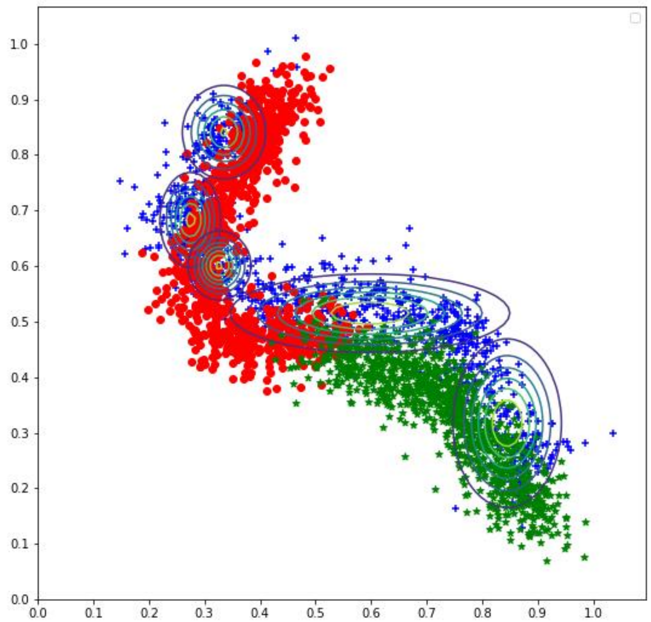
elow in Table 6 I have written the accuracy for each Guassian mixture in **<u>Test data</u>**.

	First Class Guassian mixture	Second Class Guassian mixture	Third Class Guassian mixture
Accuracy	0.646	0.766	0.898

Table 13 Accuracy Table Fo Test data

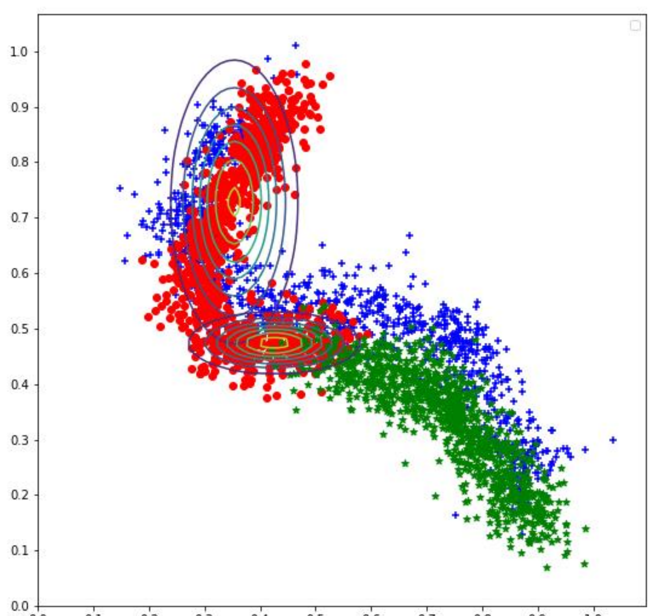
<u>The overall accuracy in Test Data (Diagonal Covariance Matrix) : 0.770</u>

Below in Figure 4 , Figure 5, Figure 6 we see the result of plotting these Guassian mixtures on our data. Each figure is showing the result for one class.



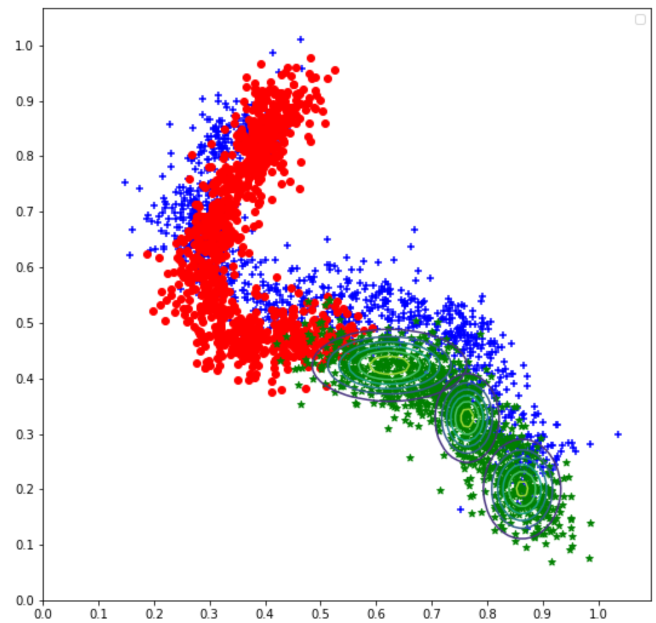
First class with Diagonal covariance matrix

Figure 4



Second class with Diagonal covariance matrix

Figure 5



Third class with Diagonal covariance matrix

Figure 6

Differenct arbitrary covariance matrix :

Below in Table 7, I have written my choice for diagonal covariance matrix case.

	Class 1	Class 2	Class 3
Number of Gaussians	4	2	2

Table 14 Number of Guassians

First calss Covariances :

$$\Sigma_{blue1} = \begin{pmatrix} 0.00378509 & -0.00500064 \\ -0.00500064 & 0.00947633 \end{pmatrix} \quad , \quad \Sigma_{blue\,2} = \begin{pmatrix} 0.00252569 & 0.0027831 \\ 0.0027831 & 0.00481327 \end{pmatrix}$$

$$\Sigma_{blue\,3} = \begin{pmatrix} 0.00095582 & -0.0007843 \\ -0.0007843 & 0.00189877 \end{pmatrix} \quad , \quad \Sigma_{blue\,4} = \begin{pmatrix} 0.01681603 & -0.0025977 \\ -0.0025977 & 0.00124885 \end{pmatrix}$$

Second Class Covariance :

$$\Sigma_{red1} = \begin{pmatrix} 0.00337512 & 0.00579541 \\ 0.00579541 & 0.0128586 \end{pmatrix} \quad , \quad \Sigma_{red\,2} = \begin{pmatrix} 0.0063525 & -0.00070607 \\ -0.00070607 & 0.00104543 \end{pmatrix}$$

Third Class Covaraince :

$$\Sigma_{green1} = \begin{pmatrix} 0.00423596 & -0.004765 \\ -0.004765 & 0.00703049 \end{pmatrix} \quad , \quad \Sigma_{green\,2} = \begin{pmatrix} 0.00581329 & -0.001371 \\ -0.001371 & 0.00124422 \end{pmatrix}$$

Below I have written the mean I have found for the Gaussians. The first element in each row of matrix corresponds to mean of a feature vector in that gaussian.

First calss Mean :

$\hat{\mu}_{blue1}$ = (0.82406541 0.35215986) , $\hat{\mu}_{blue2}$ = (0.31376872 0.80565892) ,

$\hat{\mu}_{blue3}$ = (0.29171623 0.65058792) , $\hat{\mu}_{blue4}$ = (0.56364283 0.52302667) ,

Second calss Mean :

$\hat{\mu}_{red1}$ = (0.35528836 0.75035681) , $\hat{\mu}_{red2}$ = (0.4135624 0.47946999) ,

Third calss Mean :

$\hat{\mu}_{green1}$ = (0.80992846 0.26825899) , $\hat{\mu}_{green2}$ = (0.62572793 0.41999152) ,

First Class Alpha(Prior Probability):

α_{blue1}= 0.2453 , α_{blue2}= 0.1923 α_{blue3}= 0.2025 , α_{blue4}= 0.3596

Second Class Alpha(Prior Probability):

α_{red1}= 0.5853 , α_{red2}= 0.4146,

Third Class Alpha(Prior Probability):

α_{green1}= 0.6129 , α_{green2}= 0.3870

In the Table 15 I have written tha last Log Likelihood and the number of iterations for this experiment on the EM Algorithm with Arbitrary covariance matrix.

	Class 1 (blue)	Class 2 (red)	Class 3 (green)
Last Loglikelihood	987.6565	1197.2992	1327.8457
Number of Iterations	74	52	52

Table 15

Below in Table 16 I have wirtten the confusion matrix which shows the number of elements belonging to each class after the prediction for **Training data**.

		Predicted		
		Class1 (Blue)	Class2 (Red)	Class3 (Green)
Actual	Class1 (Blue)	397	59	44
	Class2 (Red)	72	417	11
	Class3 (Green)	22	14	464

Table 16 Confusion Matrix for Training data(Arbitrary covariance Matrix)

Below in Table 11 I have written the accuracy for each Guassian mixture in **Training data**.

	First Class Guassian mixture	Second Class Guassian mixture	Third Class Guassian mixture
Accuracy	0.794	0.834	0.928

Table 17 Accuracy Table for Training data (Arbitrary Covariance Matrix)

The overall accuracy in Training Data (Arbitrary Covariance Matrix) : 0.852

Below in Table 18 I have wirtten the confusion matrix which shows the number of elements belonging to each class after the prediction for **Test data**.

	Predicted			
Actual		Class1 (Blue)	Class2 (Red)	Class3 (Green)
	Class1 (Blue)	372	70	58
	Class2 (Red)	57	386	18
	Class3 (Green)	57	18	425

Table 18 Confusion Matrix for Test data(Arbitrary covariance Matrix)

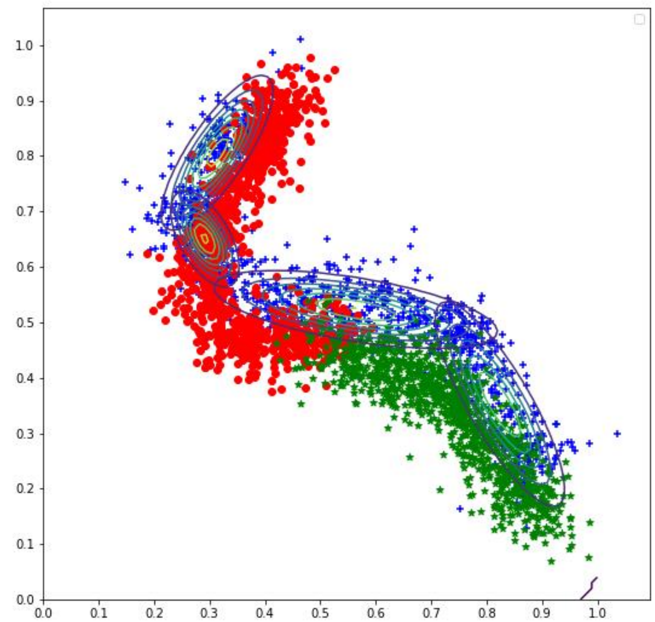
Below in Table 6 I have written the accuracy for each Guassian mixture in **Test data**.

	First Class Guassian mixture	Second Class Guassian mixture	Third Class Guassian mixture
Accuracy	0.744	0.772	0.85

Table 19 Accuracy Table Fo Test data

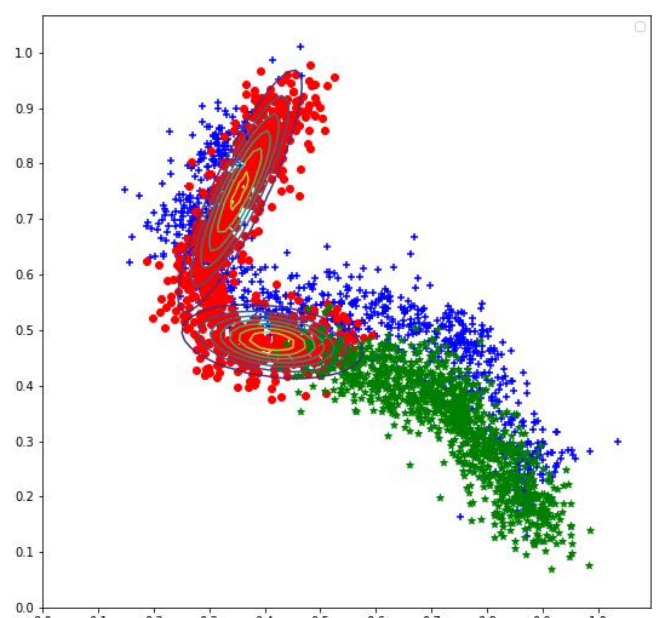
The overall accuracy in Test Data (Arbitrary Covariance Matrix) : 0.788

Below in Figure 7 , Figure 8, Figure 9 we see the result of plotting these Guassian mixtures on our data. Each figure is showing the result for one class.



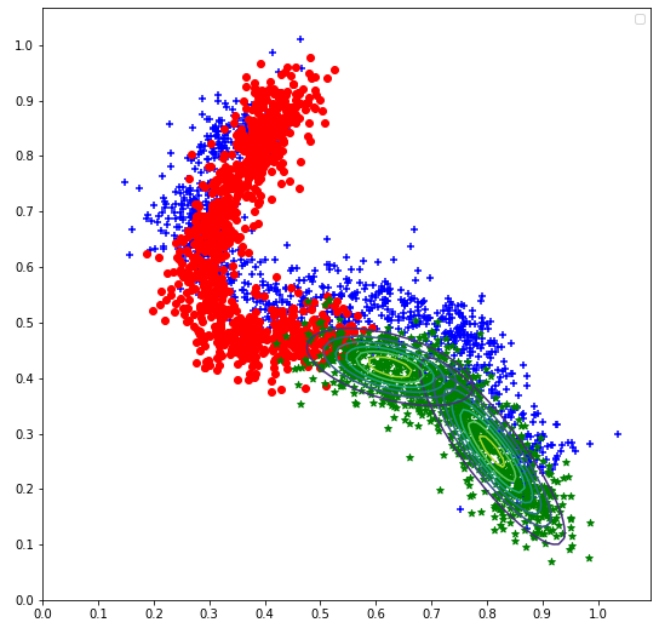
First class with Arbitrary covariance matrix

Figure 7



Second class with Arbitrary covariance matrix

Figure 8



Third class with Arbitrary covariance matrix

Figure 9

Conclusion :

After doing these experiments I have reached to a few conclusions which I am going to explain.

First is that with using Arbitrary covariance matrix I found better estimation and also I got better accuracy and higher log likelihood.

The classes which were not that much spread had higher log likelihood and also higher accuracy. Like the third class which is plotted in green.

Next, the classes which had some overlappings had lower accuracy because they also capture some of data in other class.

One other thing that captured my attention is that while using arbitrary covarince I used less numeber of guassians and despite using less number of guassians, I got better result. What I understand from this is that we can have better result using arbitrary covariance matrix.