

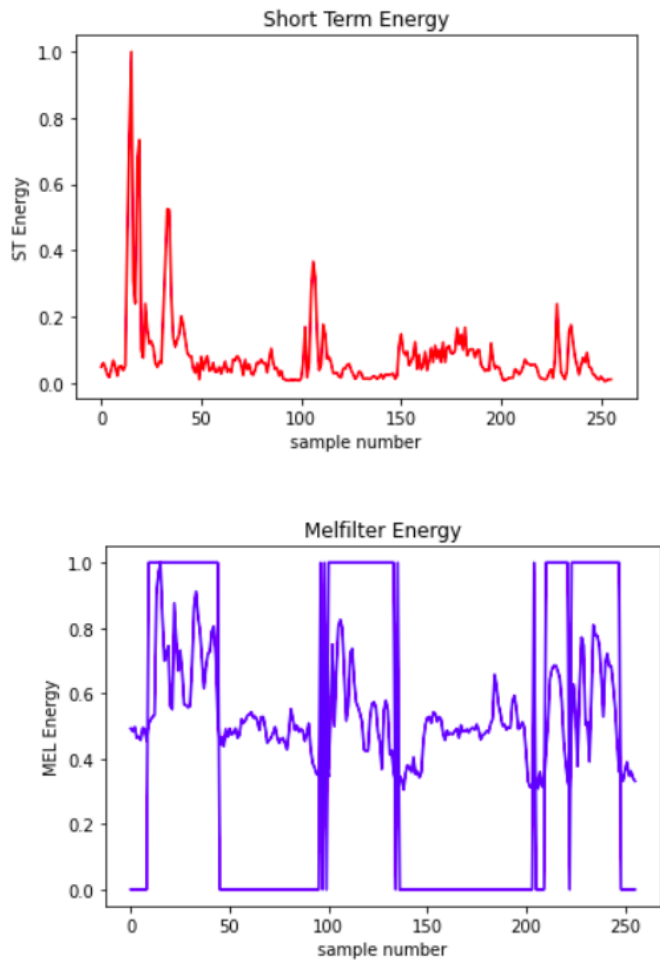
CS669 Pattern Recognition

Assignment 1

Rustam Narayan, B20128

Google Colab Notebook : [link](#)

#1 Outputs and Inferences for Question 1:



The following statistics were observed for the dataset:

Statistics for ST Energy Model:

Accuracy : 56.72%

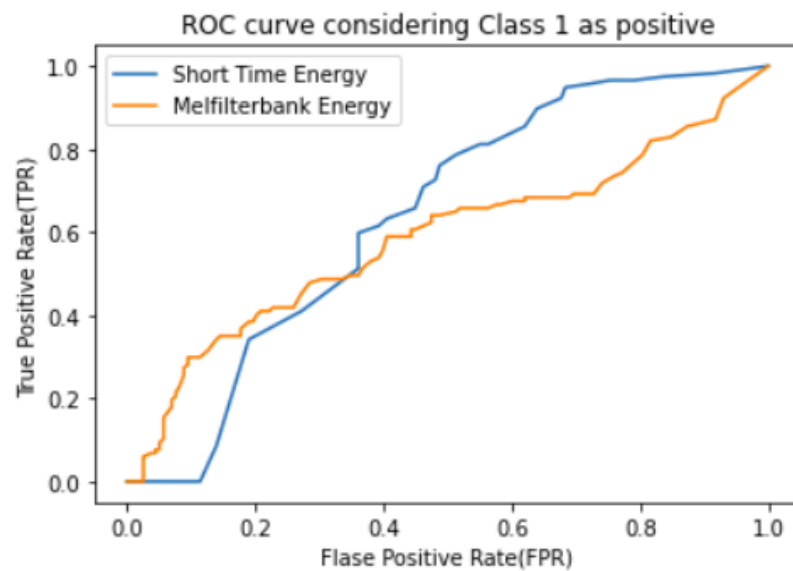
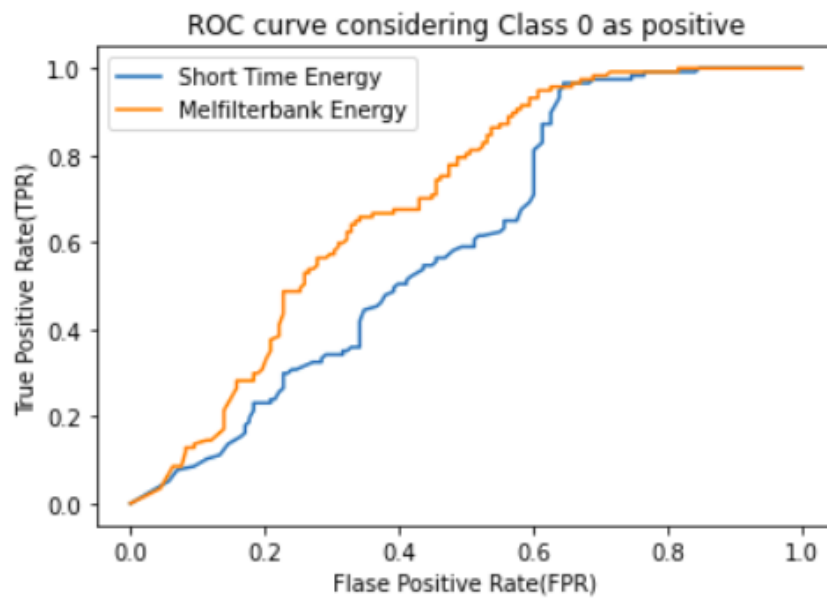
	Mean	variance
class 0	0.053	0.001
class 1	0.12	0.027

Statistics for Mel Energy Model:

Accuracy : 62.54

	Mean	variance
class 0	0.458	0.005
class 1	0.610	0.023

ROC Curves and Inferences:



Inferences:

1. Melfilterbank Energy is a better feature for detecting non speech data (classlabel = 0) as area under the MEL ROC curve is more. (fig: ROC for class 0 as positive).
2. Similarly, ST Energy is a better feature for detecting speech data.
3. The accuracies are less in both the models because the training and test data do not come from a normal distribution.

#2 Outputs and Inferences for Question 2:

Linearly Separable Data:

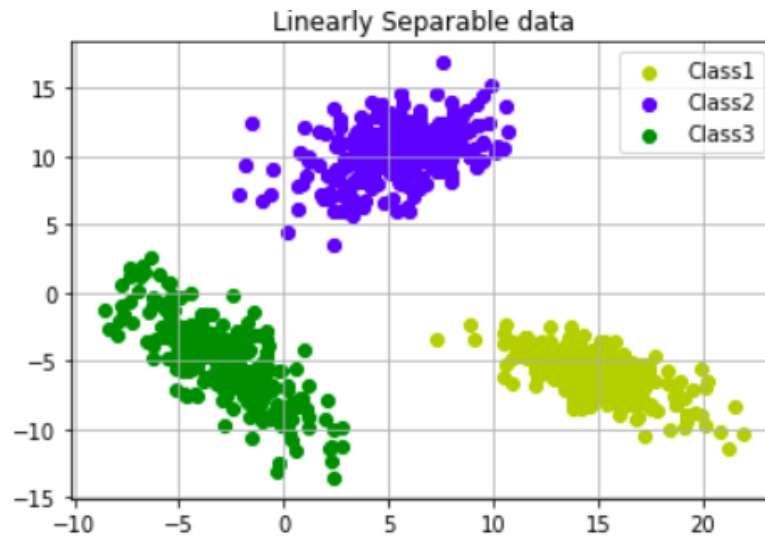


Fig: Data plot for linearly separable data

The following statistics have been observed for the different classifiers.

Model Details for Classifier 1

	Mean Vector	
class 1	14.869	-6.013
class 2	5.269	10.218
class 3	-2.796	-5.222

Covariance matrix for class 1:

5.63	0
0	5.63

Covariance matrix for class 2:

5.63	0
0	5.63

Covariance matrix for class 3:

5.63	0
0	5.63

Model Details for Classifier 2

	Mean Vector	
class 1	14.869	-6.013
class 2	5.269	10.218
class 3	-2.796	-5.222

Covariance matrix for class 1:

5.858	-2.012
-2.012	5.401

Covariance matrix for class 2:

5.858	-2.012
-2.012	5.401

Covariance matrix for class 3:

5.858	-2.012
-2.012	5.401

Model Details for Classifier 3

	Mean Vector	
class 1	14.869	-6.013
class 2	5.269	10.218
class 3	-2.796	-5.222

Covariance matrix for class 1:

5.779	0
0	2.801

Covariance matrix for class 2:

6.118	0
0	4.223

Covariance matrix for class 3:

5.677	0
0	9.179

Model Details for Classifier 4

	Mean Vector	
class 1	14.869	-6.013
class 2	5.269	10.218
class 3	-2.796	-5.222

Covariance matrix for class 1:

5.779	-2.605
-2.605	2.801

Covariance matrix for class 2:

6.118	2.1
2.1	4.223

Covariance matrix for class 3:

5.677	-5.532
-5.532	9.179

The performance of the different classifiers:

	Accuracy	Precision	Recall	F-score
C1	1	1	1	1
C2	1	1	1	1
C3	0.999	0.998	0.998	0.998
C4	1	1	1	1

The plot of contours and data is given below.

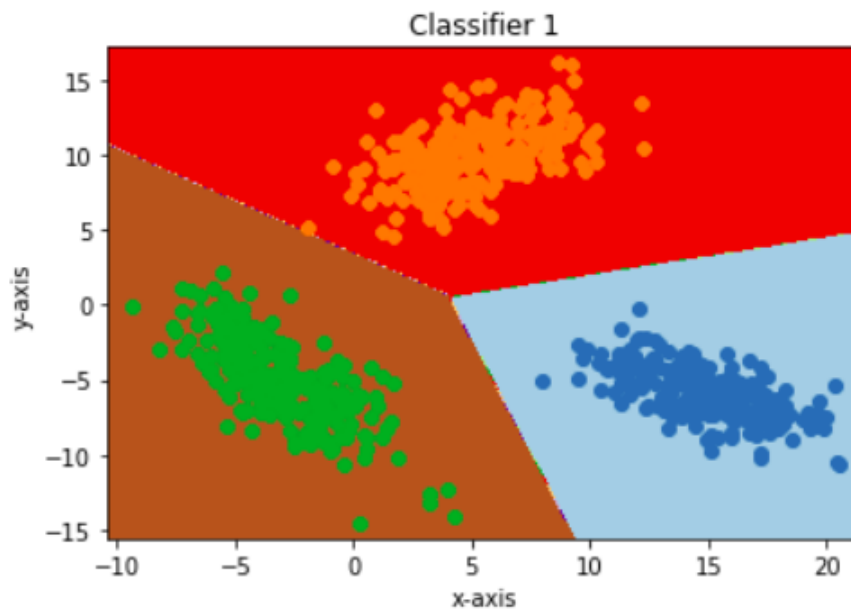


Fig: Decision boundaries with Classifier 1

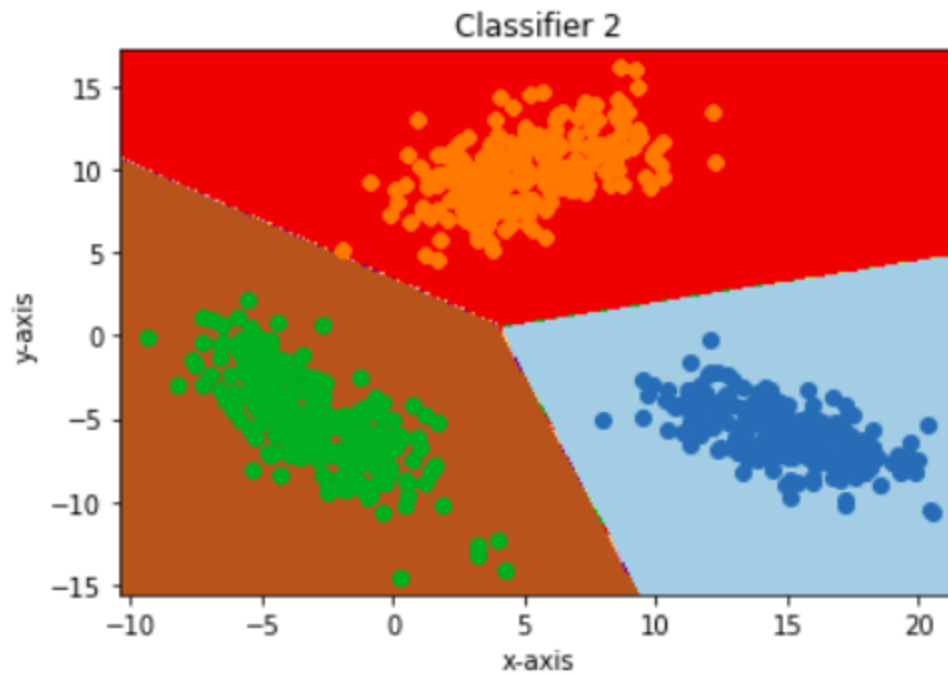


Fig: Decision boundaries with Classifier 2

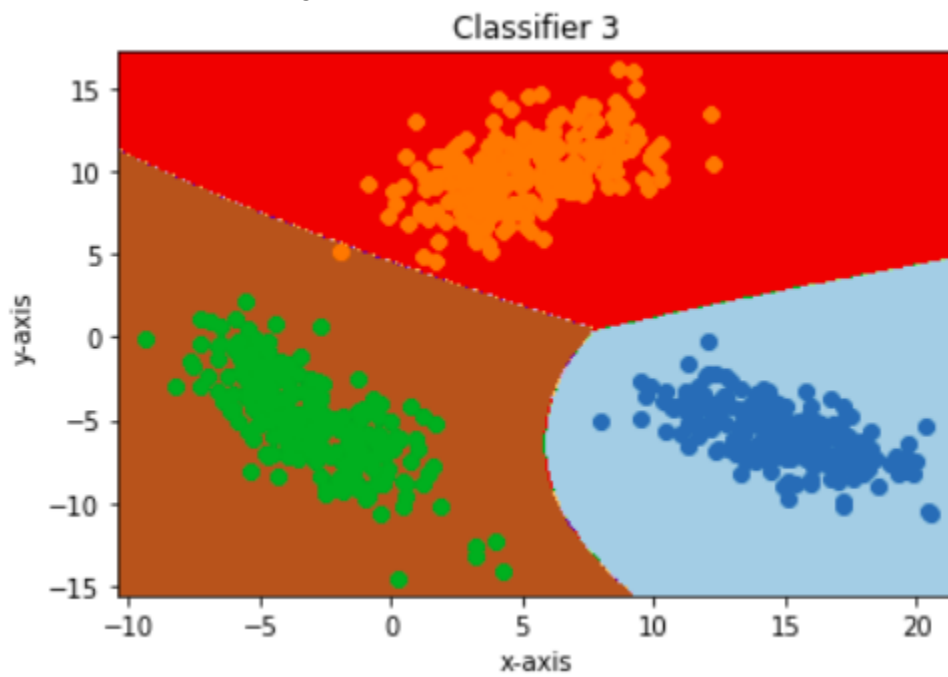


Fig: Decision boundaries with Classifier 3

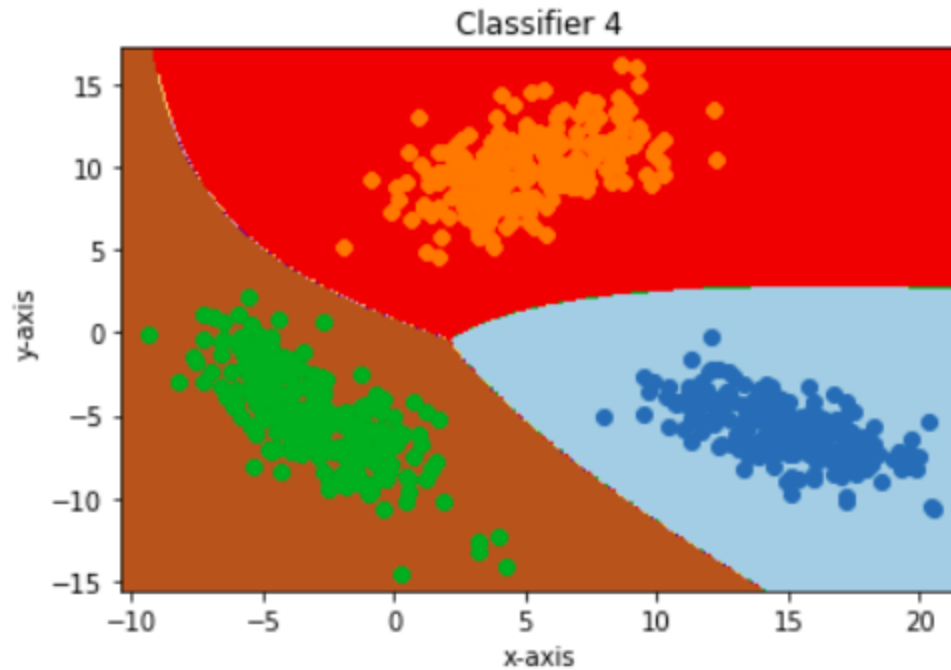


Fig: Decision boundaries with Classifier 4

Inferences:

1. The accuracy, precision and other performance parameters of these classifiers are quite high. Thus it can be deduced that the data sample collected, comes from some normal distribution.
2. The data contours are clearly visible. It supports the fact that the data given is linearly separable.
3. The decision boundaries are quite vivid.

Non-Linearly Separable Data:

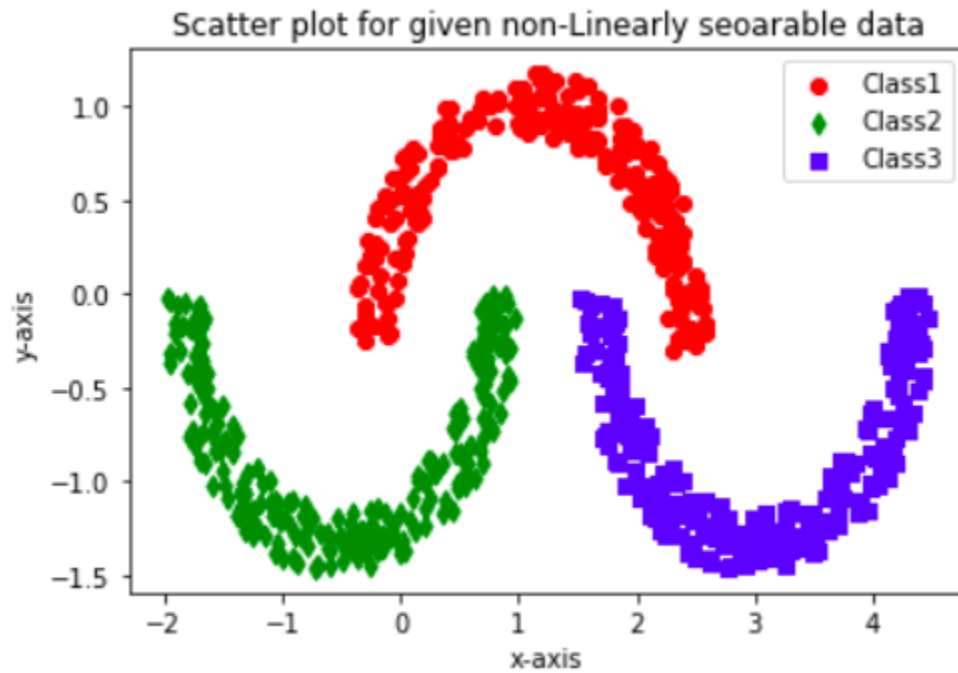


Fig: Data plot for linearly separable test data

The following statistics have been observed for the different classifiers:

Model details for Classifier 1

	Mean Vector	
class 1	1.223	0.543
class 2	-0.474	-0.833
class 3	3.025	-0.833

Covariance matrix for class 1:

0.539	0
0	0.539

Covariance matrix for class 2:

0.539	0
0	0.539

Covariance matrix for class 3:

0.539	0
0	0.539

Model details for Classifier 2

	Mean Vector	
class 1	1.223	0.543
class 2	-0.474	-0.833
class 3	3.025	-0.833

Covariance matrix for class 1:

0.901	0.015
0.015	0.177

Covariance matrix for class 2:

0.901	0.015
0.015	0.177

Covariance matrix for class 3:

0.901	0.015
0.015	0.177

Model details for Classifier 3

	Mean Vector	
class 1	1.223	0.543
class 2	-0.474	-0.833
class 3	3.025	-0.833

Covariance matrix for class 1:

0.908	0
0	0.153

Covariance matrix for class 2:

0.898	0
0	0.19

Covariance matrix for class 3:

0.898	0
0	0.19

Model details for Classifier 4

	Mean Vector	
class 1	1.223	0.543
class 2	-0.474	-0.833
class 3	3.025	-0.833

Covariance matrix for class 1:

0.908	-0.02
-0.02	0.153

Covariance matrix for class 2:

0.898	0.032
0.032	0.19

Covariance matrix for class 3:

0.898	0.032
0.032	0.19

The performance summary of the different classifiers:

	Accuracy	Precision	Recall	F-score
C1	0.88	0.822	0.82	0.821
C2	0.942	0.913	0.913	0.913
C3	0.944	0.915	0.916	0.915
C4	0.943	0.914	0.914	0.914

The plot of contours and data is given below.

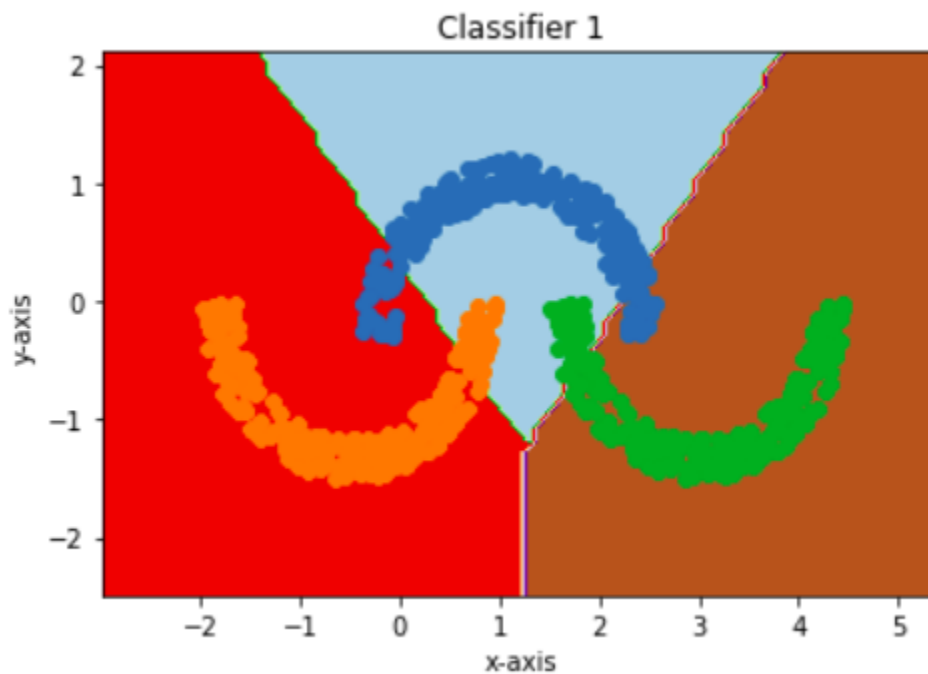


Fig: Decision boundaries with Classifier 1

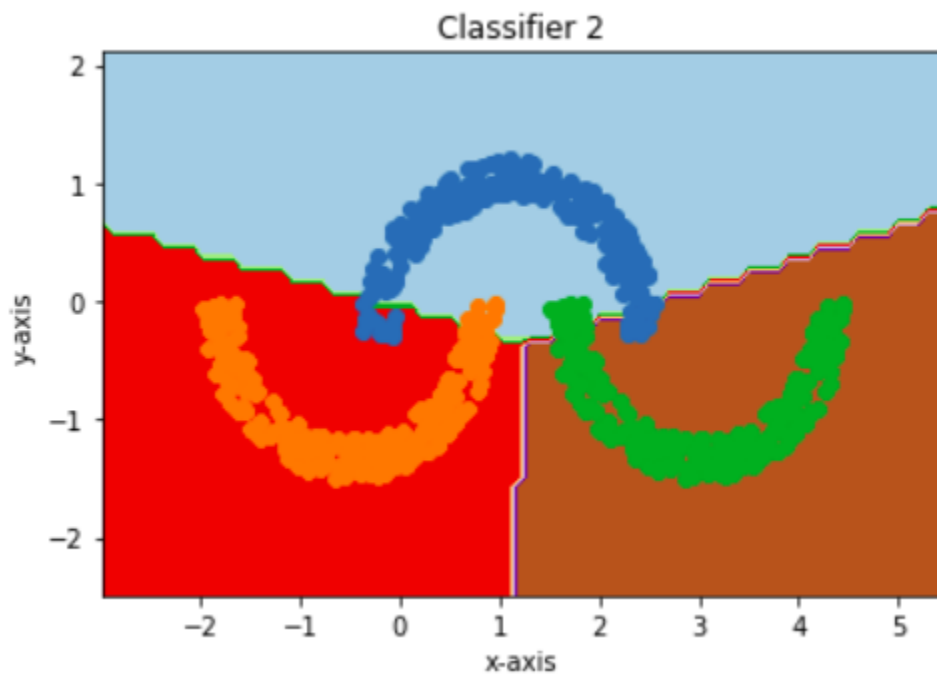


Fig: Decision boundaries with Classifier 2

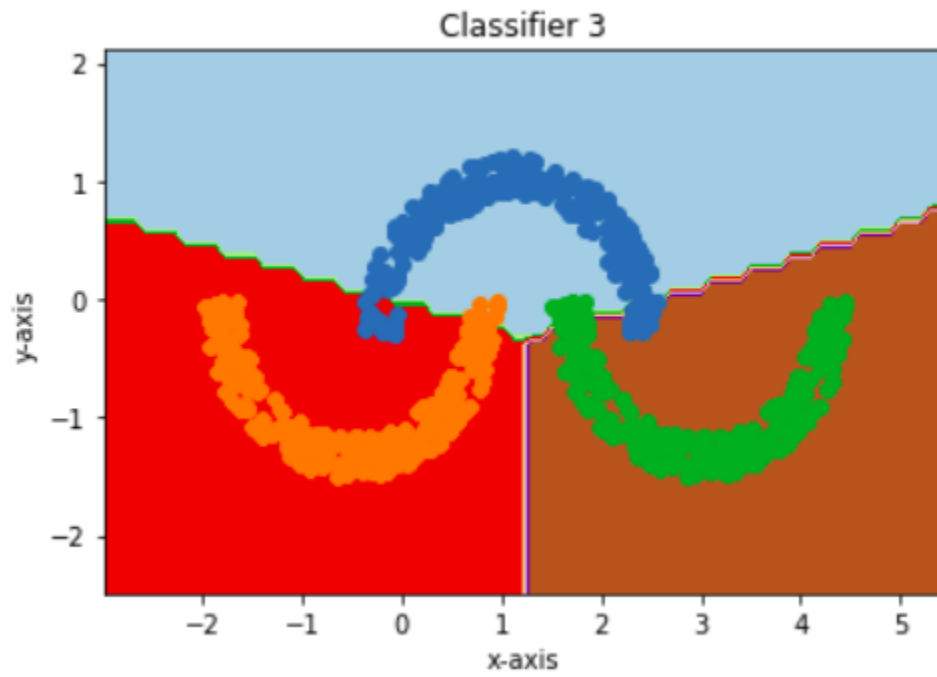


Fig: Decision boundaries with Classifier 3

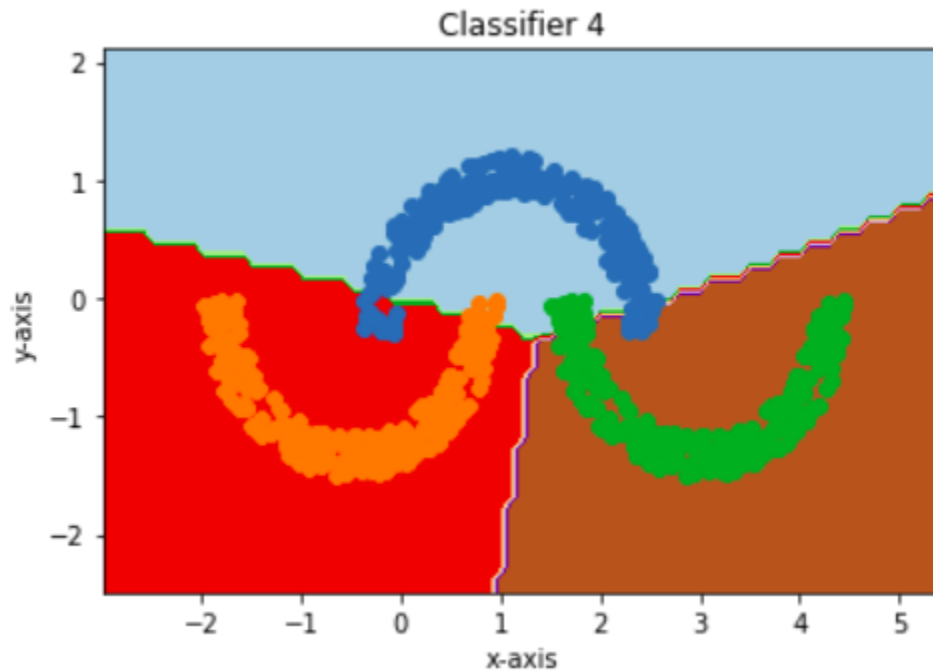


Fig: Decision boundaries with Classifier 4

Inferences:

1. Here, in this case also, the accuracy and precision of all the four classifiers are high (Clearly not as high as the previous data). This suggests that this non linearly separable data also belongs to some normal distribution. However, maximum likelihood estimation is not able to identify the parameters perfectly but gives an approximate of the distribution.
2. The data is non-linearly separable and thus we are unable to make a perfect (linear) decision boundary. Maybe a multimodal Gaussian Mixture Model can give a more accurate decision boundary.