PATTERN RECOGNITION ASSIGNMENT- 4
REPORT

15h December 2018

PRATYUSH GAURAV
B16026

Faculty
Dr. A.D Dileep
Indian Institute of Technology, Mandi

Table of Contents

1. Introduction

Principal Component Analysis:

Principal Component Analysis (PCA) is a dimension-reduction tool that can be used to reduce a large set of variables to a small set that still contains most of the information in the large set. PCA transforms a number of correlated variables into smaller number of uncorrelated variables called principal components.

PCA only takes those directions in which the information content is maximum. While doing so it does not take into account the separability of data. It might so happen that, the direction in which separability of 2 classes is maximum is not the direction of maximum variance of projected data.

Data Set :

3 class scene image dataset whose 32-dimensional BoVW representation is taken which was generated in Assignment-2.

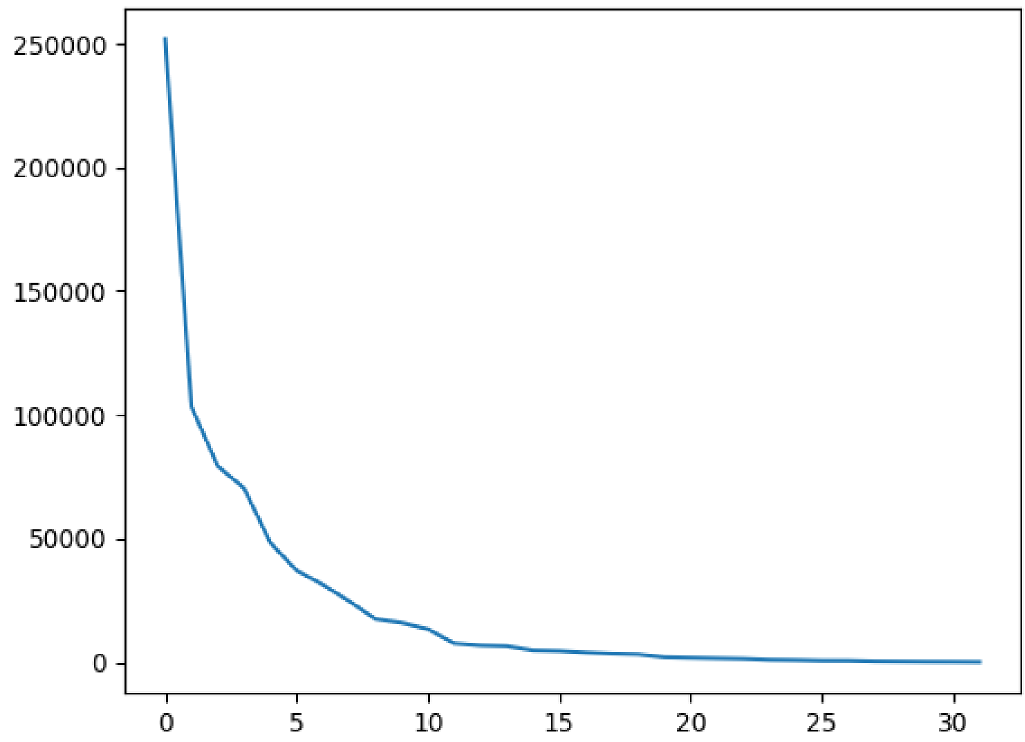
Experimental Observations

2. Principal Component Analysis

Train and test samples are classified into 3 classes. Each class contains samples of particular images. Class 1 contains images of candy store, class 2 contains images of forest broadleaf and class 3 contains images of football stadium. 32 dimesional BoVW representation of these images are taken as inputs.

2.1 Results

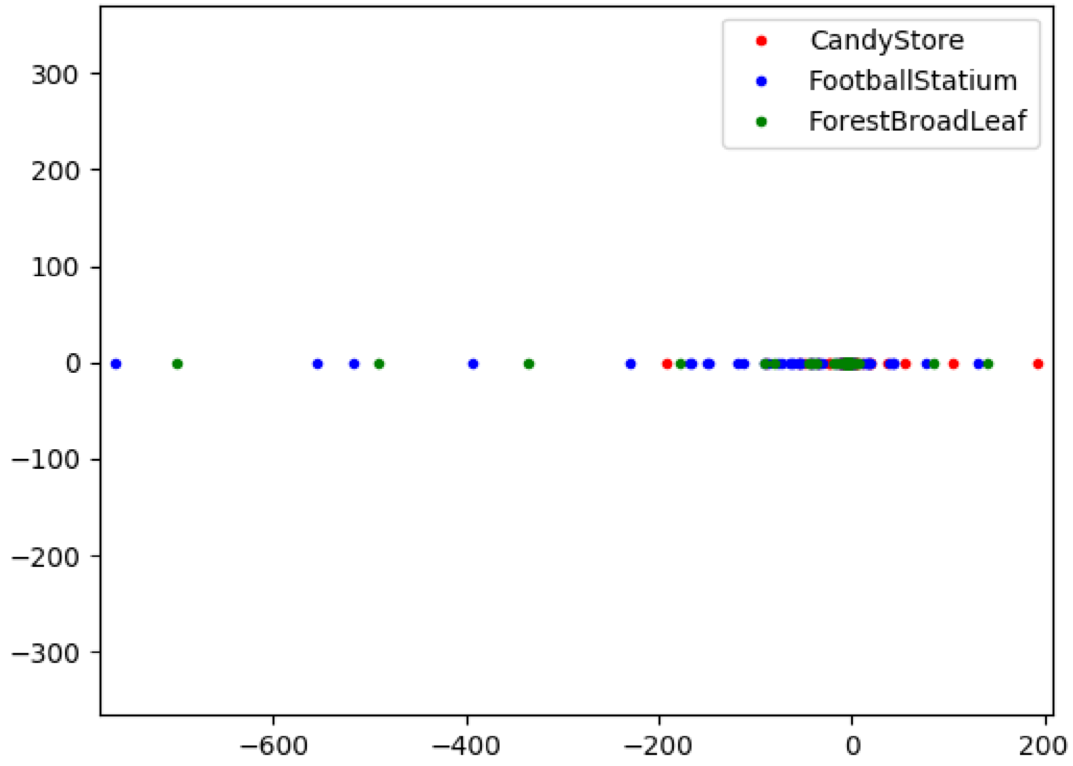
Plot of Eigen Values (Sorted) :-



Results of GMM for different values of l and K.

Class1:- Candy Store
Class2: Football Stadium
Class3: Forest Broadleaf

2.1.1 l = 1



Plot for L=1(Direction of Max Variance)

2.1.1.1 K=1

Number of Data Points = 50
Number of Parameters to be estimated = 2

Classification Accuracy (%)	38
Precision for Class 1 Precision for Class 2 Precision for Class 3	0.362 0.818 0.166
Mean Precision	0.449
Recall for Class 1 Recall for Class 2 Recall for Class 3	0.92 0.18 0.04
Mean Recall	0.38
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	0.519 0.295 0.064
Mean F-measure	0.293

Confusion Matrix :

$$C = \begin{matrix} 46 & 1 & 3 \\ 34 & 9 & 7 \\ 47 & 1 & 2 \end{matrix}$$

2.1.1.2 K=2

Number of Data Points = 50
Number of Parameters to be estimated = 6

Classification Accuracy (%)	43.3
Precision for Class 1 Precision for Class 2 Precision for Class 3	0.457 0.431 0.422
Mean Precision	0.437
Recall for Class 1 Recall for Class 2 Recall for Class 3	0.32 0.38 0.60
Mean Recall	0.433
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	0.376 0.404 0.495
Mean F-measure	0.424

Confusion Matrix :

$$C = \begin{matrix} 16 & 13 & 21 \\ 11 & 19 & 20 \\ 8 & 12 & 30 \end{matrix}$$

2.1.1.3 K=4

Number of Data Points = 50
Number of Parameters to be estimated = 12

Classification Accuracy (%)	44
Precision for Class 1 Precision for Class 2 Precision for Class 3	0.413 0.478 0.431
Mean Precision	0.440
Recall for Class 1 Recall for Class 2 Recall for Class 3	0.38 0.44 0.50
Mean Recall	0.44
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	0.395 0.458 0.463

Mean F-measure	0.480

Confusion Matrix :

$$\mathbf{C} = \begin{matrix} 19 & 14 & 17 \\ 12 & 22 & 16 \\ 15 & 10 & 25 \end{matrix}$$

2.1.1.4 K=8

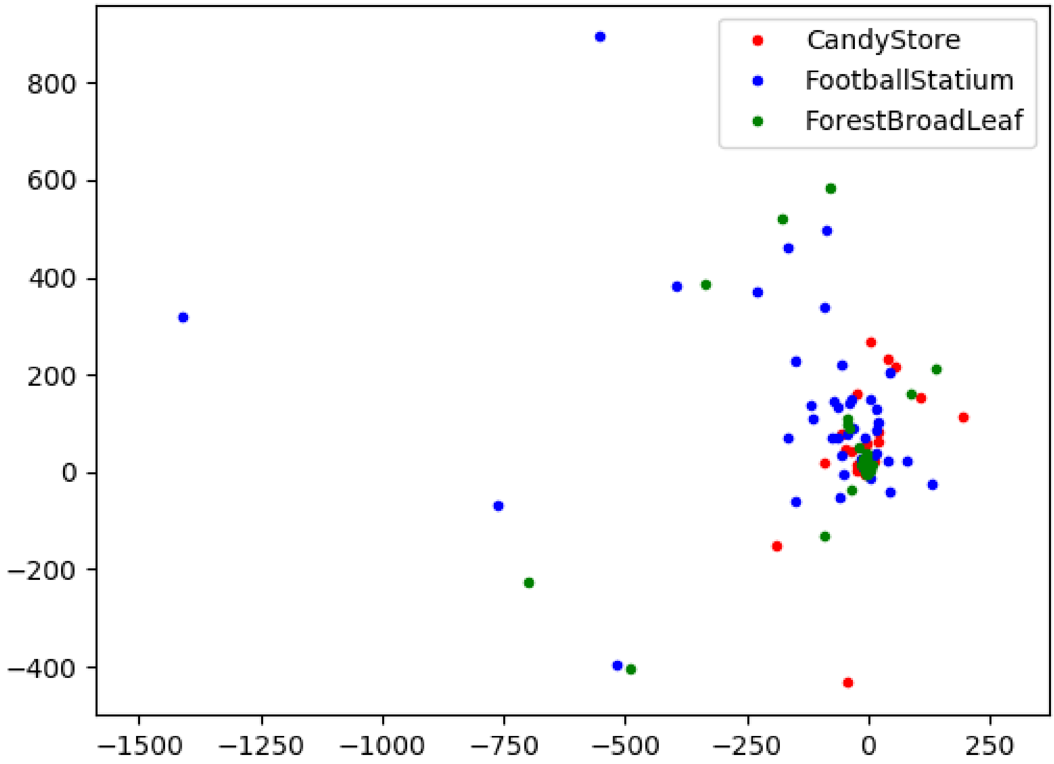
Number of Data Points = 50
Number of Parameters to be estimated = 24

Classification Accuracy (%)	48.667
Precision for Class 1 Precision for Class 2 Precision for Class 3	0.484 0.574 0.428
Mean Precision	0.495
Recall for Class 1 Recall for Class 2 Recall for Class 3	0.32 0.54 0.60
Mean Recall	0.4866
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	0.385 0.556 0.50
Mean F-measure	0.481

Confusion Matrix :

$$\mathbf{C} = \begin{matrix} 16 & 8 & 26 \\ 9 & 27 & 14 \\ 8 & 12 & 30 \end{matrix}$$

2.1.2 l = 2



Plot for L=2

2.1.2.1 K=1

Number of Data Points = 50
Number of Parameters to be estimated = 5

Classification Accuracy (%)	40.66
Precision for Class 1 Precision for Class 2 Precision for Class 3	0.394 0.608 0.153
Mean Precision	0.385
Recall for Class 1 Recall for Class 2 Recall for Class 3	0.90 0.28 0.04
Mean Recall	0.406
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	0.548 0.383 0.063
Mean F-measure	0.330

Confusion Matrix :

$$C = \begin{matrix} 45 & 3 & 2 \\ 27 & 14 & 7 \\ 42 & 6 & 2 \end{matrix}$$

2.1.2.2 K=2

Number of Data Points = 50
Number of Parameters to be estimated = 12

Classification Accuracy (%)	43.3
Precision for Class 1 Precision for Class 2 Precision for Class 3	0.398 0.588 0.0
Mean Precision	0.328
Recall for Class 1 Recall for Class 2 Recall for Class 3	0.90 0.40 0.0
Mean Recall	0.433
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	0.55 0.47 0.0
Mean F-measure	0.34

Confusion Matrix :

$$C = \begin{matrix} 45 & 4 & 1 \\ 28 & 20 & 2 \\ 40 & 10 & 0 \end{matrix}$$

2.1.2.3 K=4

Number of Data Points = 50
Number of Parameters to be estimated = 24

Classification Accuracy (%)	49.33
Precision for Class 1 Precision for Class 2 Precision for Class 3	0.478 0.478 0.517
Mean Precision	0.491
Recall for Class 1 Recall for Class 2 Recall for Class 3	0.44 0.44 0.6
Mean Recall	0.493
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	0.458 0.458 0.555

Mean F-measure	0.49

Confusion Matrix :

$$C = \begin{matrix} 22 & 15 & 13 \\ 13 & 22 & 15 \\ 11 & 9 & 30 \end{matrix}$$

2.1.2.4 K=8

Number of Data Points = 50
Number of Parameters to be estimated = 48

Classification Accuracy (%)	53.33
Precision for Class 1 Precision for Class 2 Precision for Class 3	0.50 0.545 0.56
Mean Precision	0.535
Recall for Class 1 Recall for Class 2 Recall for Class 3	0.56 0.48 0.56
Mean Recall	0.533
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	0.528 0.510 0.56
Mean F-measure	0.523

Confusion Matrix :

$$C = \begin{matrix} 28 & 10 & 12 \\ 16 & 24 & 10 \\ 12 & 10 & 28 \end{matrix}$$

2.1.3 l = 3

2.1.3.1 K=1
Number of Data Points = 50
Number of Parameters to be estimated = 9

Classification Accuracy (%)	41.6666666667
Precision for Class 1 Precision for Class 2 Precision for Class 3	0.383 0.652 0.20
Mean Precision	0.413
Recall for Class 1 Recall for Class 2 Recall for Class 3	0.86 0.30 0.06
Mean Recall	0.4066

F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	0.53 0.41 0.09
Mean F-measure	0.344

Confusion Matrix :

$$C = \begin{matrix} 43 & 3 & 4 \\ 27 & 15 & 8 \\ 42 & 5 & 3 \end{matrix}$$

2.1.3.2 K=2

Number of Data Points = 50
Number of Parameters to be estimated = 20

Classification Accuracy (%)	47.33
Precision for Class 1 Precision for Class 2 Precision for Class 3	0.50 0.515 0.443
Mean Precision	0.486
Recall for Class 1 Recall for Class 2 Recall for Class 3	0.38 0.34 0.7
Mean Recall	0.4733
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	0.431 0.409 0.542
Mean F-measure	0.46

Confusion Matrix :

$$C = \begin{matrix} 16 & 13 & 21 \\ 11 & 19 & 20 \\ 8 & 12 & 30 \end{matrix}$$

2.1.3.3 K=4

Number of Data Points = 50
Number of Parameters to be estimated = 40

Classification Accuracy (%)	41.33
Precision for Class 1 Precision for Class 2 Precision for Class 3	0.50 0.441 0.382
Mean Precision	0.441
Recall for Class 1 Recall for Class 2	0.22 0.30

Recall for Class 3	0.72
Mean Recall	0.413
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	0.305 0.357 0.50
Mean F-measure	0.396

Confusion Matrix :

$$\mathbf{C} = \begin{matrix} 11 & 8 & 31 \\ 8 & 5 & 27 \\ 3 & 11 & 36 \end{matrix}$$

2.1.3.4 K=8

Number of Data Points = 50
Number of Parameters to be estimated (Diagonal Cov Matrix)= 56

Classification Accuracy (%)	42
Precision for Class 1 Precision for Class 2 Precision for Class 3	0.464 0.714 0.382
Mean Precision	0.519
Recall for Class 1 Recall for Class 2 Recall for Class 3	0.12 0.2 0.94
Mean Recall	0.42
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	0.19 0.312 0.5433
Mean F-measure	0.348

Confusion Matrix :

$$\mathbf{C} = \begin{matrix} 6 & 2 & 42 \\ 6 & 10 & 34 \\ 1 & 2 & 47 \end{matrix}$$

2.1.4 l = 4

2.1.4.1 K=1

Number of Data Points = 50
Number of Parameters to be estimated = 14

Classification Accuracy (%)	40
Precision for Class 1 Precision for Class 2 Precision for Class 3	0.3909 0.5517 0.0909

Mean Precision	0.344
Recall for Class 1 Recall for Class 2 Recall for Class 3	0.86 0.32 0.02
Mean Recall	0.4
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	0.537 0.405 0.032
Mean F-measure	0.384

Confusion Matrix :

$$C = \begin{matrix} 43 & 5 & 2 \\ 26 & 16 & 8 \\ 41 & 8 & 1 \end{matrix}$$

2.1.4.2 K=2

Number of Data Points = 50
Number of Parameters to be estimated = 30

Classification Accuracy (%)	40.667
Precision for Class 1 Precision for Class 2 Precision for Class 3	0.4074 0.421 0.40
Mean Precision	0.409
Recall for Class 1 Recall for Class 2 Recall for Class 3	0.22 0.32 0.68
Mean Recall	0.40667
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	0.285 0.363 0.503
Mean F-measure	0.384

Confusion Matrix :

$$C = \begin{matrix} 11 & 14 & 25 \\ 8 & 16 & 26 \\ 8 & 8 & 34 \end{matrix}$$

2.1.4.3 K=4

Number of Data Points = 50
Number of Parameters to be estimated(Diagonal Cov Matrix) = 36

Classification Accuracy (%)	47.33

Precision for Class 1 Precision for Class 2 Precision for Class 3	0.456 0.489 0.472
Mean Precision	0.473
Recall for Class 1 Recall for Class 2 Recall for Class 3	0.42 0.48 0.52
Mean Recall	0.4733
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	0.4375 0.4848 0.495
Mean F-measure	0.472

Confusion Matrix :

$$\mathbf{C} = \begin{matrix} 21 & 14 & 15 \\ 12 & 24 & 12 \\ 13 & 11 & 26 \end{matrix}$$

2.1.4.4 K=8

Number of Data Points = 50
Number of Parameters to be estimated(Diagonal Cov Matrix) = 72

Classification Accuracy (%)	42
Precision for Class 1 Precision for Class 2 Precision for Class 3	0.365 0.0 0.875
Mean Precision	0.4133
Recall for Class 1 Recall for Class 2 Recall for Class 3	0.98 0.0 0.28
Mean Recall	0.42
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	0.532 0.0 0.424
Mean F-measure	0.318

Confusion Matrix :

$$\mathbf{C} = \begin{matrix} 49 & 0 & 1 \\ 49 & 0 & 1 \\ 36 & 0 & 14 \end{matrix}$$

2.1.5 l = 10

2.1.5.1 K=1

Number of Data Points = 50

Number of Parameters to be estimated = 55

Classification Accuracy (%)	52.66
Precision for Class 1 Precision for Class 2 Precision for Class 3	0.50 0.596 0.484
Mean Precision	0.526
Recall for Class 1 Recall for Class 2 Recall for Class 3	0.34 0.62 0.62
Mean Recall	0.5266
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	0.404 0.607 0.543
Mean F-measure	0.518

Confusion Matrix :

$$C = \begin{matrix} 17 & 9 & 24 \\ 10 & 31 & 9 \\ 7 & 12 & 31 \end{matrix}$$

2.1.5.2 K=2

Number of Data Points = 50
Number of Parameters to be estimated(Diagonal Cov Matrix) = 42

Classification Accuracy (%)	61.33
Precision for Class 1 Precision for Class 2 Precision for Class 3	0.526 0.666 0.769
Mean Precision	0.654
Recall for Class 1 Recall for Class 2 Recall for Class 3	0.80 0.64 0.40
Mean Recall	0.6133
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	0.634 0.653 0.526
Mean F-measure	0.604

Confusion Matrix :

$$C = \begin{matrix} 40 & 7 & 3 \\ 15 & 32 & 3 \\ 21 & 9 & 20 \end{matrix}$$

2.1.5.3 K=4

Number of Data Points = 50
Number of Parameters to be estimated(Diagonal Cov Matrix) = 84

Classification Accuracy (%)	53.333
Precision for Class 1 Precision for Class 2 Precision for Class 3	0.667 0.472 0.60
Mean Precision	0.579
Recall for Class 1 Recall for Class 2 Recall for Class 3	0.32 0.86 0.42
Mean Recall	0.533
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	0.432 0.6099 0.494
Mean F-measure	0.52

Confusion Matrix :

$$\mathbf{C} = \begin{matrix} 16 & 25 & 9 \\ 2 & 43 & 5 \\ 6 & 23 & 21 \end{matrix}$$

2.1.5.4 K=8

Number of Data Points = 50
Number of Parameters to be estimated(Diagonal Cov Matrix) = 150+

Not enough data Points to estimate 150+ parameters.

2.2 Overall Observations:

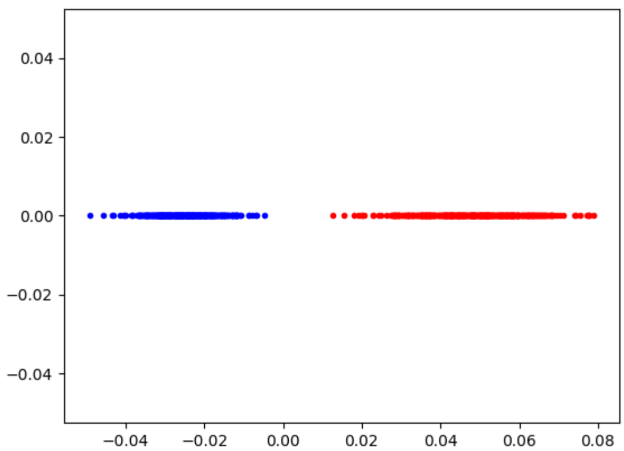
- Lesser is the value of l, more is the loss of information. If we take l equal to one, then the accuracy is quite less as only one projection direction is considered out of d possible directions in a d dimensional space.
- Transformation is defined in such a way that the first principal component has the largest possible variance and the plot for L=1 is shown.
- The increase in accuracy between successive values of l is directly proportional to the Eigen value of the new direction taken. Since the magnitude of eigenvalues keeps on decreasing so the percentage increase in the accuracy keeps on decreasing with increasing l.
- After a certain l, the value of Eigen values become very less or negligible compared to initial Eigen values then on increasing l no increase in accuracy is obtained.
- For larger values of l, curse of dimensionality becomes prominent.
- For a certain l the accuracy is maximum and on deviating form that value of l the accuracy will decrease.

3.Fisher linear discriminant analysis

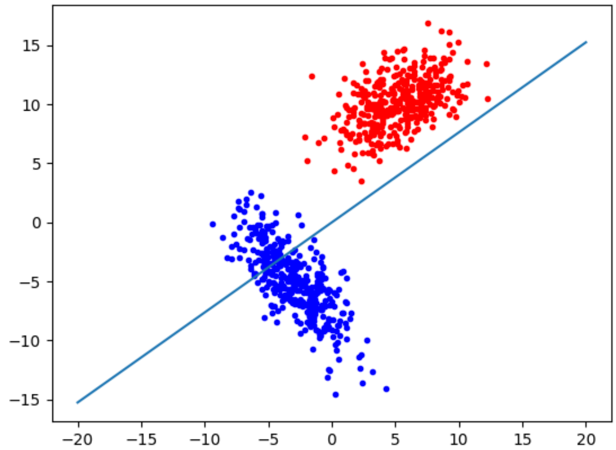
3.1 Dataset 1(a):-

2-dimensional artificial data Linearly separable

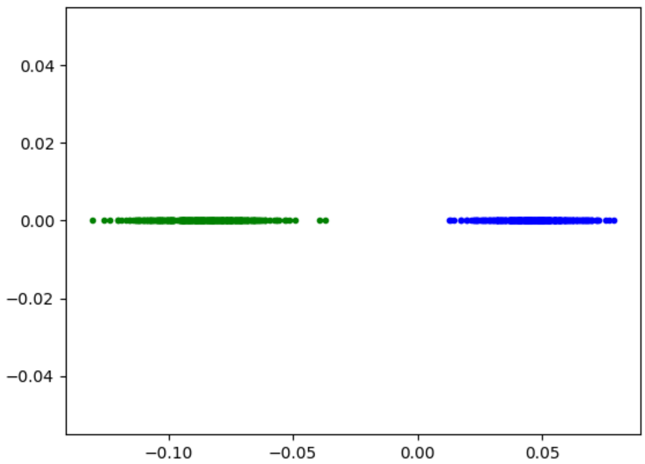
Train and test data is classified into 3 classes. One vs One approach was used to classify test data points.



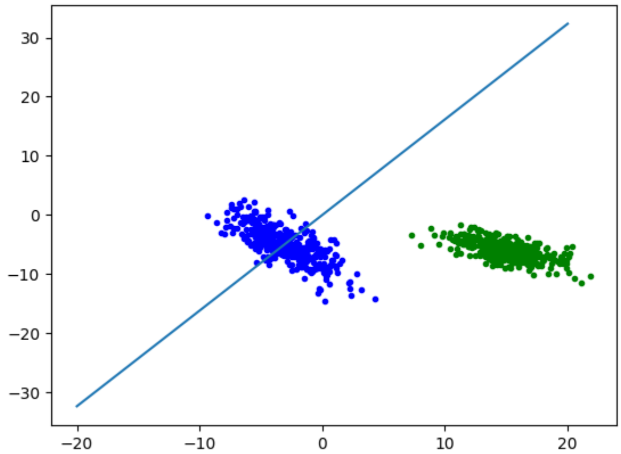
Projected Data for Class1 and Class2



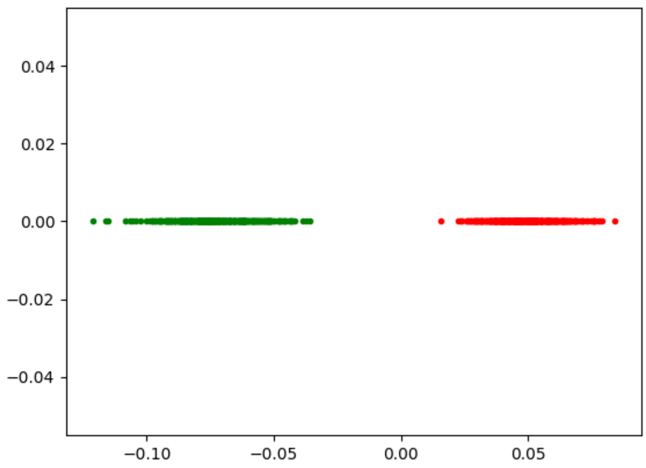
Vector for Class1 and Class2



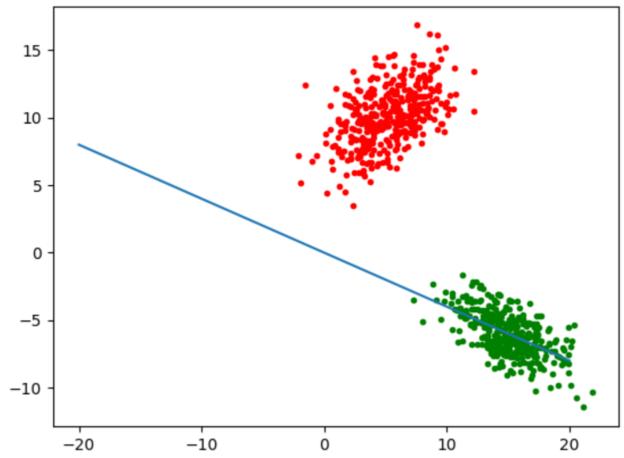
Projected Data for Class2 and Class3



Vector for Class2 and Class3



Projected Data for Class1 and Class3



Vector for Class1 and Class3

3.1.1 Unimodal Gaussian(K=1):-

Classification Accuracy (%)	100%
Precision for Class 1 Precision for Class 2 Precision for Class 3	1.0 1.0 1.0
Mean Precision	1.0
Recall for Class 1 Recall for Class 2 Recall for Class 3	1.0 1.0 1.0
Mean Recall	1.0
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	1.0 1.0 1.0
Mean F-measure	1.0

Confusion Matrix :

$$C = \begin{matrix} 125 & 0 & 0 \\ 0 & 125 & 0 \\ 0 & 0 & 125 \end{matrix}$$

3.1.2 Multimodal Gaussian(K=2):-

Classification Accuracy (%)	100%
Precision for Class 1	1.0

Precision for Class 2 Precision for Class 3	1.0 1.0
Mean Precision	1.0
Recall for Class 1 Recall for Class 2 Recall for Class 3	1.0 1.0 1.0
Mean Recall	1.0
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	1.0 1.0 1.0
Mean F-measure	1.0

Confusion Matrix :

$$C = \begin{matrix} 125 & 0 & 0 \\ 0 & 125 & 0 \\ 0 & 0 & 125 \end{matrix}$$

3.1.3 Multimodal Gaussian(K=4):-

Classification Accuracy (%)	100%
Precision for Class 1 Precision for Class 2 Precision for Class 3	1.0 1.0 1.0
Mean Precision	1.0
Recall for Class 1 Recall for Class 2 Recall for Class 3	1.0 1.0 1.0
Mean Recall	1.0
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	1.0 1.0 1.0
Mean F-measure	1.0

Confusion Matrix :

$$C = \begin{matrix} 125 & 0 & 0 \\ 0 & 125 & 0 \\ 0 & 0 & 125 \end{matrix}$$

3.1.4 Multimodal Gaussian(K=8):-

Classification Accuracy (%)	100%
Precision for Class 1 Precision for Class 2 Precision for Class 3	1.0 1.0 1.0

Mean Precision	1.0
Recall for Class 1 Recall for Class 2 Recall for Class 3	1.0 1.0 1.0
Mean Recall	1.0
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	1.0 1.0 1.0
Mean F-measure	1.0

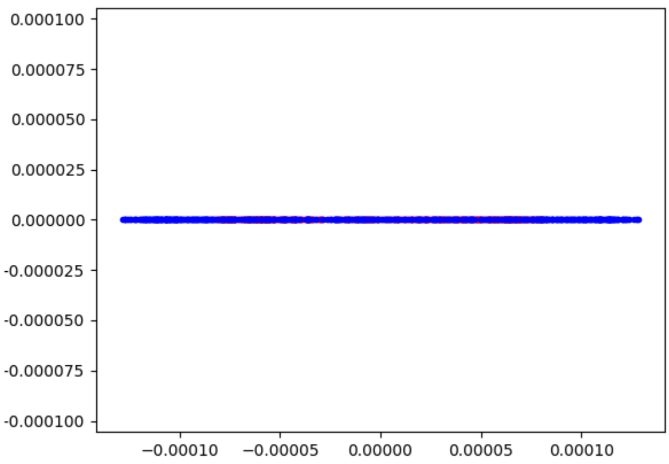
Confusion Matrix :

$$C = \begin{matrix} 125 & 0 & 0 \\ 0 & 125 & 0 \\ 0 & 0 & 125 \end{matrix}$$

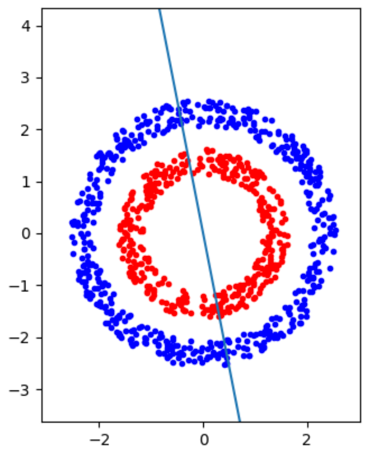
3.2 Dataset 1(b):-

2-dimensional artificial data Nonlinearly separable

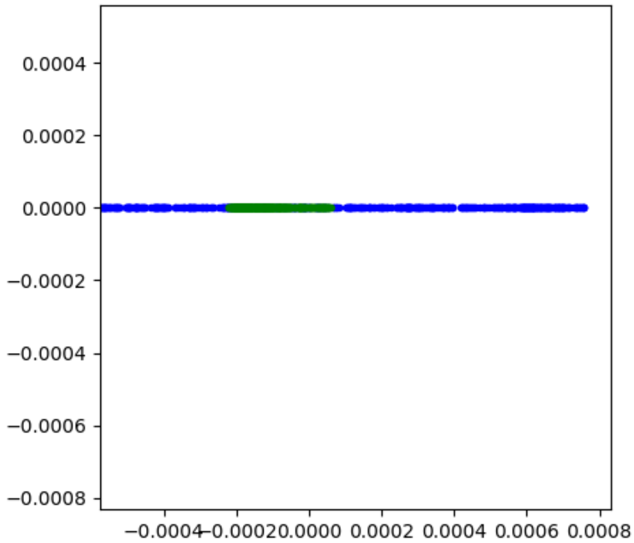
Train and test data is classified into 3 classes. One vs One approach was used to classify test data points.



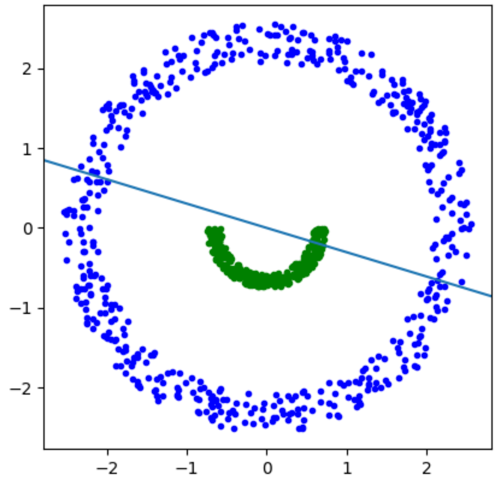
Projected data for Class1 and Class2



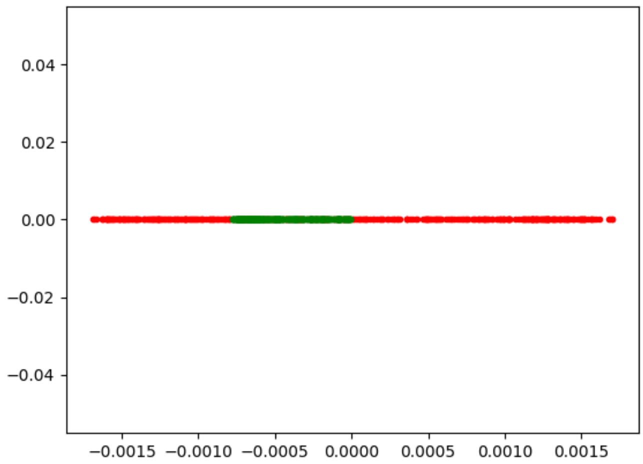
Vector for Class1 and Class2



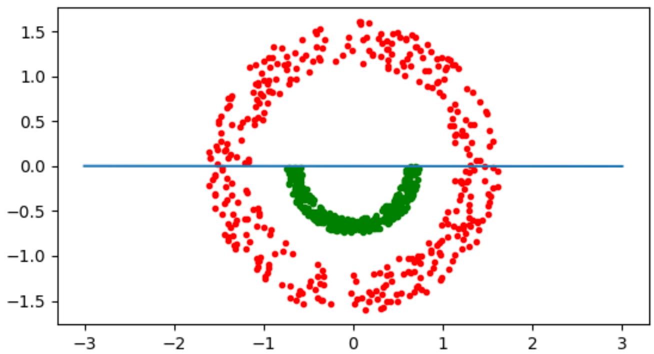
Projected Data for Class2 and Class3



Vector for Class2 and Class3



Projected Data Class1 and Class3



Vector for Class1 and Class3

3.2.1 Unimodal Gaussian(K=1):-

Classification Accuracy (%)	65.88%
Precision for Class 1 Precision for Class 2 Precision for Class 3	0.4712 0.628 0.830
Mean Precision	0.6439
Recall for Class 1 Recall for Class 2 Recall for Class 3	0.328 0.7485 0.864
Mean Recall	0.646
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	0.386 0.684 0.847
Mean F-measure	0.639

Confusion Matrix :

$$C = \begin{matrix} 41 & 65 & 19 \\ 41 & 131 & 3 \\ 5 & 12 & 108 \end{matrix}$$

3.2.2 Multimodal Gaussian(K=2):-

Classification Accuracy (%)	67.7%
Precision for Class 1 Precision for Class 2 Precision for Class 3	0.464 0.7 0.83
Mean Precision	0.667
Recall for Class 1 Recall for Class 2 Recall for Class 3	0.472 0.6 0.992
Mean Recall	0.688
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	0.468 0.646 0.908
Mean F-measure	0.674

Confusion Matrix :

$$C = \begin{matrix} 59 & 45 & 21 \\ 67 & 105 & 3 \\ 1 & 0 & 124 \end{matrix}$$

3.2.3 Multimodal Gaussian(K=4):-

Classification Accuracy (%)	67.9%
Precision for Class 1 Precision for Class 2 Precision for Class 3	0.47 0.71 0.83
Mean Precision	0.67
Recall for Class 1 Recall for Class 2 Recall for Class 3	0.512 0.571 0.992
Mean Recall	0.691
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	0.49 0.633 0.908
Mean F-measure	0.677

Confusion Matrix :

$$C = \begin{matrix} 64 & 40 & 21 \\ 72 & 100 & 3 \\ 0 & 1 & 124 \end{matrix}$$

3.2.4 Multimodal Gaussian(K=8):-

Classification Accuracy (%)	67.8%
Precision for Class 1 Precision for Class 2 Precision for Class 3	0.456 0.715 0.842
Mean Precision	0.671
Recall for Class 1 Recall for Class 2 Recall for Class 3	0.544 0.531 0.984
Mean Recall	0.686
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	0.496 0.609 0.907
Mean F-measure	0.671

Confusion Matrix :

$$C = \begin{matrix} 68 & 37 & 20 \\ 79 & 93 & 3 \\ 2 & 0 & 123 \end{matrix}$$

3.3 Dataset 2:-

3 class scene image dataset
Class1:-CandyStore
Class2:-Football Stadium
Class3:-Forest BroadLeaf

One vs One approach was used to classify test data points.

3.3.1 Unimodal Gaussian(K=1):-

Classification Accuracy (%)	49.33%
Precision for Class 1 Precision for Class 2 Precision for Class 3	0.435 0.633 0.482
Mean Precision	0.517
Recall for Class 1 Recall for Class 2 Recall for Class 3	0.54 0.38 0.56
Mean Recall	0.493
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	0.482 0.475 0.518
Mean F-measure	0.4918

Confusion Matrix :

$$C = \begin{matrix} 27 & 6 & 17 \\ 18 & 19 & 13 \\ 17 & 5 & 28 \end{matrix}$$

3.3.2 Multimodal Gaussian(K=2):-

Classification Accuracy (%)	46.6%
Precision for Class 1 Precision for Class 2 Precision for Class 3	0.447 0.571 0.436
Mean Precision	0.485
Recall for Class 1 Recall for Class 2	0.6 0.32

Recall for Class 3	0.48
Mean Recall	0.466
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	0.512 0.410 0.457
Mean F-measure	0.460

Confusion Matrix :

$$C = \begin{matrix} 30 & 6 & 14 \\ 17 & 16 & 17 \\ 20 & 6 & 24 \end{matrix}$$

3.3.3 Multimodal Gaussian(K=4):-

Classification Accuracy (%)	45.33%
Precision for Class 1 Precision for Class 2 Precision for Class 3	0.45 0.545 0.426
Mean Precision	0.473
Recall for Class 1 Recall for Class 2 Recall for Class 3	0.54 0.28 0.58
Mean Recall	0.453
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	0.4909 0.333 0.491
Mean F-measure	0.438

Confusion Matrix :

$$C = \begin{matrix} 27 & 5 & 18 \\ 17 & 12 & 21 \\ 16 & 5 & 29 \end{matrix}$$

3.3.4 Multimodal Gaussian(K=8):-

Classification Accuracy (%)	44%
Precision for Class 1 Precision for Class 2 Precision for Class 3	0.419 0.653 0.370
Mean Precision	0.48
Recall for Class 1 Recall for Class 2 Recall for Class 3	0.52 0.34 0.46

Mean Recall	0.44
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	0.464 0.447 0.410
Mean F-measure	0.440

Confusion Matrix :

$$C = \begin{matrix} 26 & 5 & 19 \\ 13 & 17 & 20 \\ 23 & 4 & 23 \end{matrix}$$

3.4 Overall Observations:-

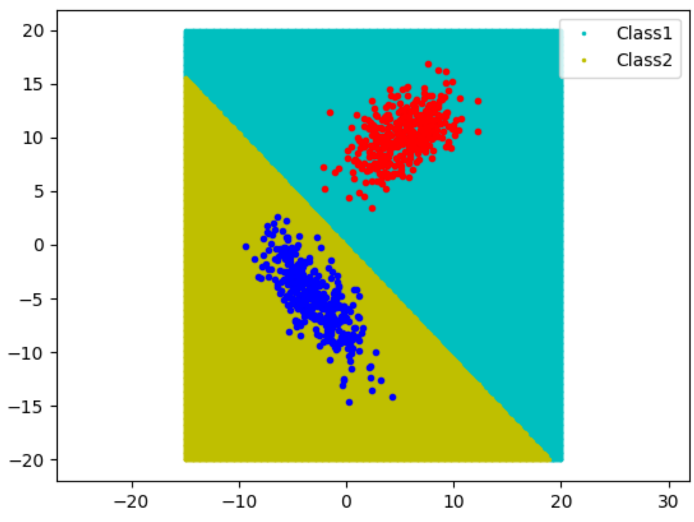
- For Linearly separable data, LDA gives direction in which seperation between 2 classes is maximum.
- For Non-Linearly separable data, there is no such direction hence accuracy is very low
- In Non-linear separable classes (1 and 2) and classes (2 and 3),projecting the data on the given line overlaps the 2 classes causing a lot of confusion between them when bayes classifier is used to classify them.
- For Non-Linearly separable data, as number of clusters increase, Accuracy increases.

4. Perceptron Learning

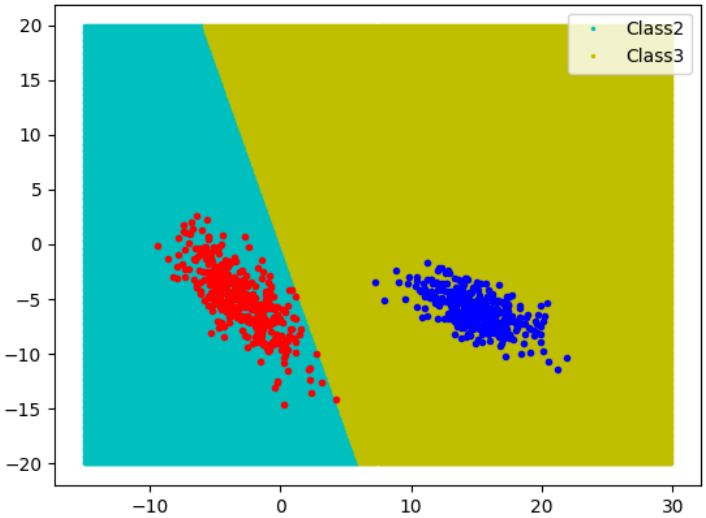
4.1 Dataset 1(a):-

2-dimensional artificial data Linearly separable

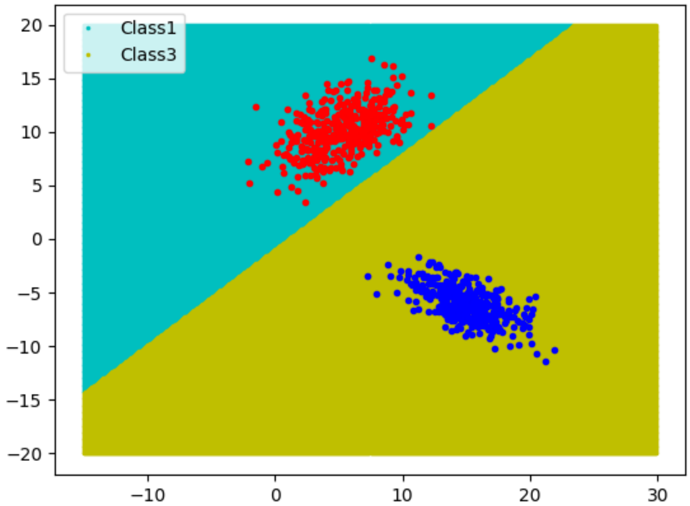
Train and test data is classified into 3 classes. One vs One approach was used to classify test data points.



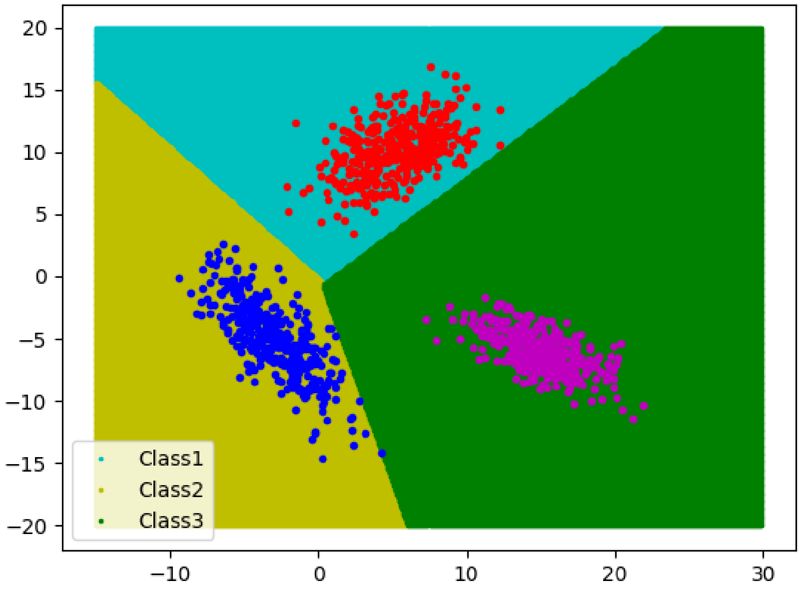
Decision Boundary between class 1 and class 2



Decision Boundary between class 2 and class 3



Decision Boundary between class 1 and class 3



Plot for all 3 classes

Classification Accuracy (%)	99.7%
Precision for Class 1 Precision for Class 2 Precision for Class 3	1.0 1.0 0.992
Mean Precision	0.997
Recall for Class 1 Recall for Class 2 Recall for Class 3	1.0 0.992 1.0
Mean Recall	0.997
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	1.0 0.996 0.995
Mean F-measure	0.997

Confusion Matrix :

$$C = \begin{matrix} 125 & 0 & 0 \\ 0 & 124 & 1 \\ 0 & 0 & 125 \end{matrix}$$

4.2 Observations:-

- The decision boundary found by perceptron between classes 2 and 3 is very close to class 2 and hence some test data is misclassified.
- Similarly for decision boundary between classes 1 and 3

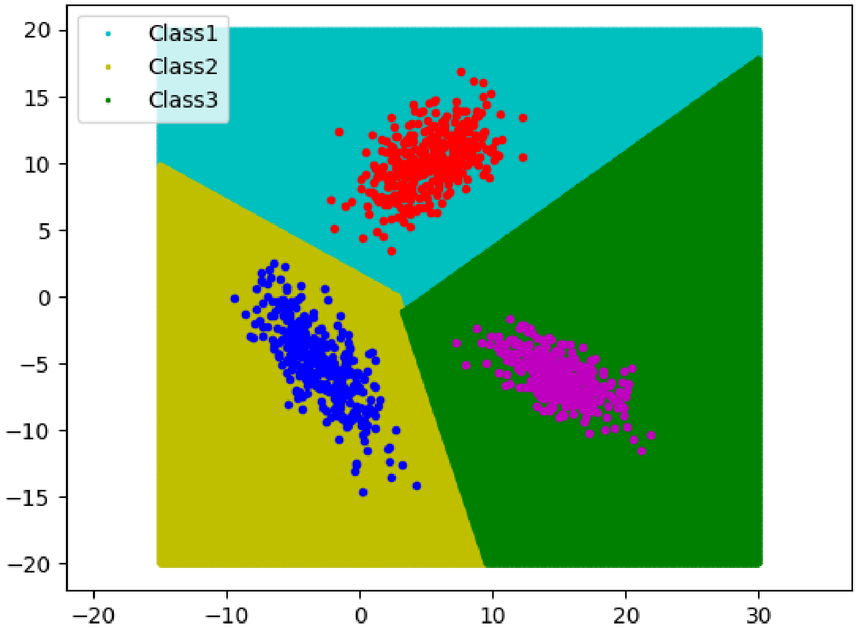
5.Support Vector Machine

5.1 Dataset 1(a):-

2-dimensional artificial data Linearly separable

Train and test data is classified into 3 classes. One vs One approach was used to classify test data points.

5.1.1 Linear Kernel:-



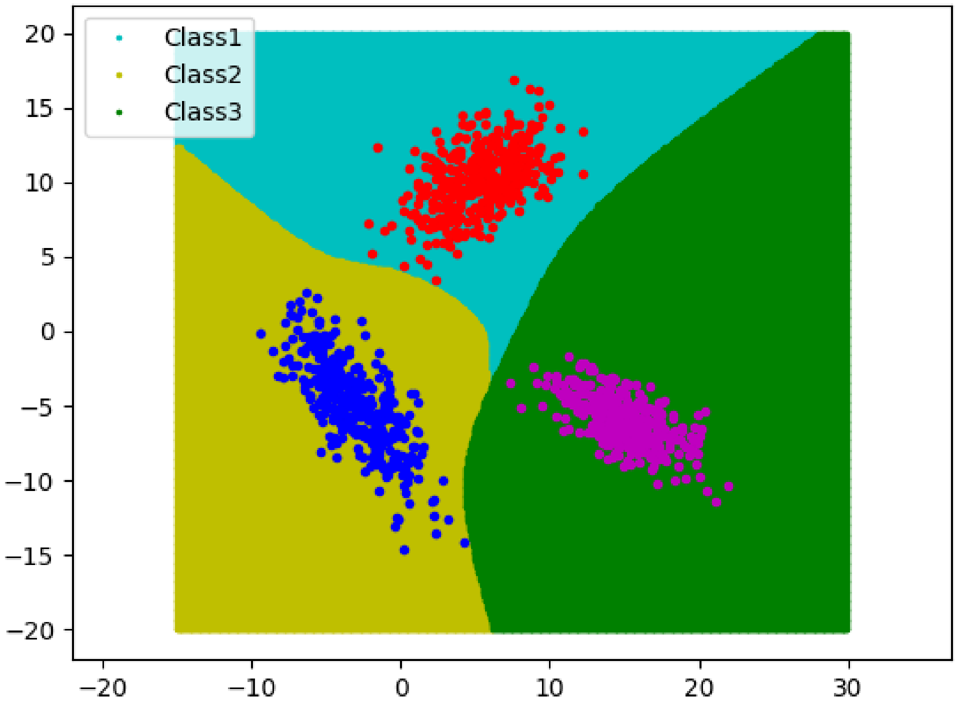
Classification Accuracy (%)	100%
Precision for Class 1 Precision for Class 2 Precision for Class 3	1.0 1.0 1.0
Mean Precision	1.0
Recall for Class 1 Recall for Class 2 Recall for Class 3	1.0 1.0 1.0
Mean Recall	1.0

F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	1.0 1.0 1.0
Mean F-measure	1.0

Confusion Matrix :

$$C = \begin{matrix} 125 & 0 & 0 \\ 0 & 125 & 0 \\ 0 & 0 & 125 \end{matrix}$$

5.1.2 Polynomial(degree=5) Kernel:-



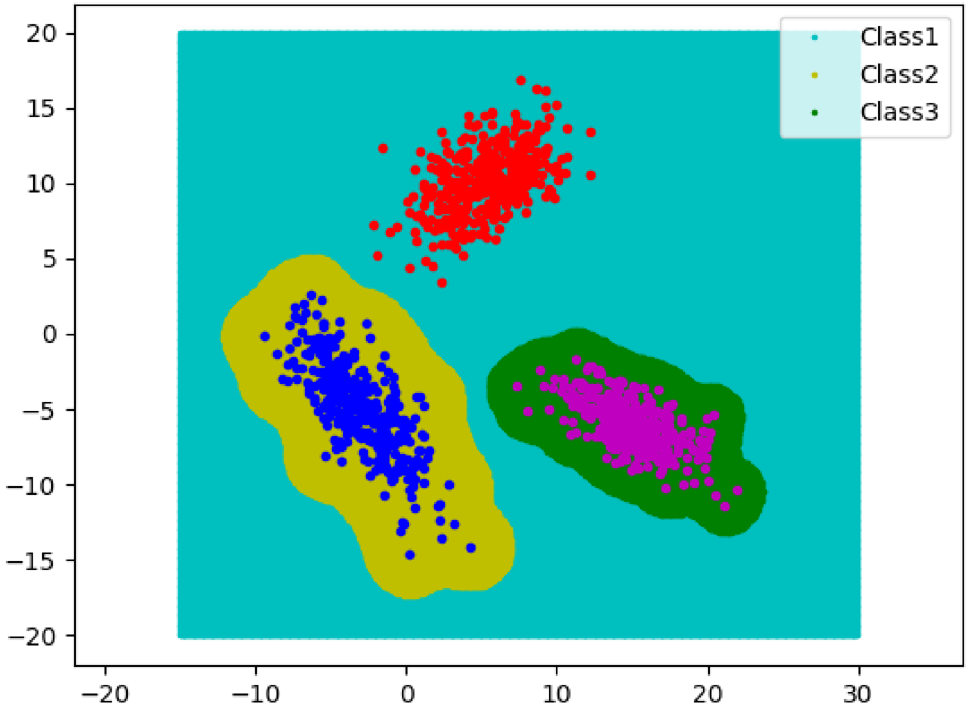
Classification Accuracy (%)	100%
Precision for Class 1 Precision for Class 2 Precision for Class 3	1.0 1.0 1.0
Mean Precision	1.0
Recall for Class 1 Recall for Class 2 Recall for Class 3	1.0 1.0 1.0

Mean Recall	1.0
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	1.0 1.0 1.0
Mean F-measure	1.0

Confusion Matrix :

$$C = \begin{matrix} 125 & 0 & 0 \\ 0 & 125 & 0 \\ 0 & 0 & 125 \end{matrix}$$

5.1.3 Gaussian Kernel:-



Classification Accuracy (%)	100%
Precision for Class 1 Precision for Class 2 Precision for Class 3	1.0 1.0 1.0
Mean Precision	1.0

Recall for Class 1 Recall for Class 2 Recall for Class 3	1.0 1.0 1.0
Mean Recall	1.0
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	1.0 1.0 1.0
Mean F-measure	1.0

Confusion Matrix :

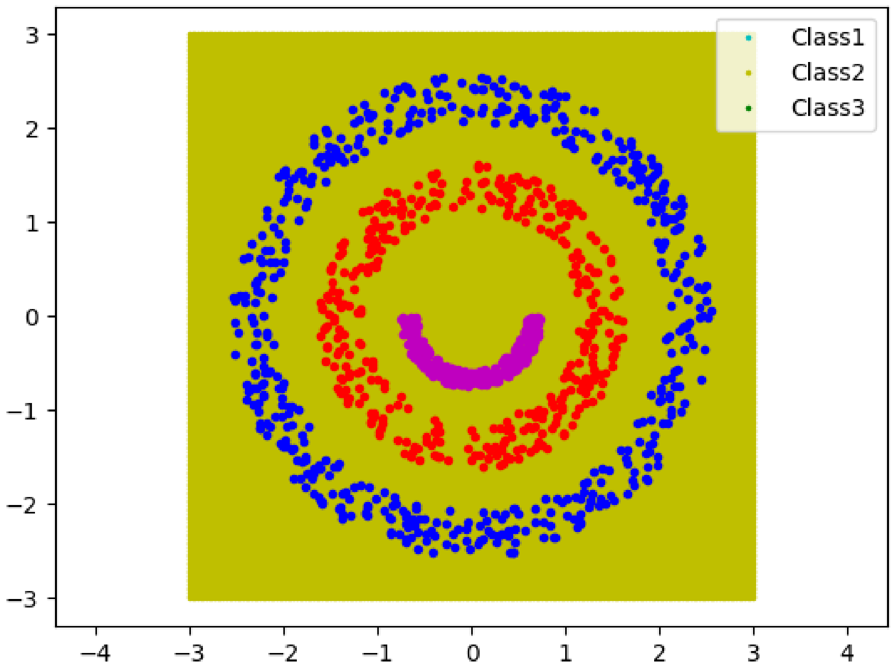
$$C = \begin{matrix} 125 & 0 & 0 \\ 0 & 125 & 0 \\ 0 & 0 & 125 \end{matrix}$$

5.2 Dataset 1(b):-

2-dimensional artificial data Non-Linearly separable

Train and test data is classified into 3 classes. One vs One approach was used to classify test data points.

5.2.1 Linear Kernel:-

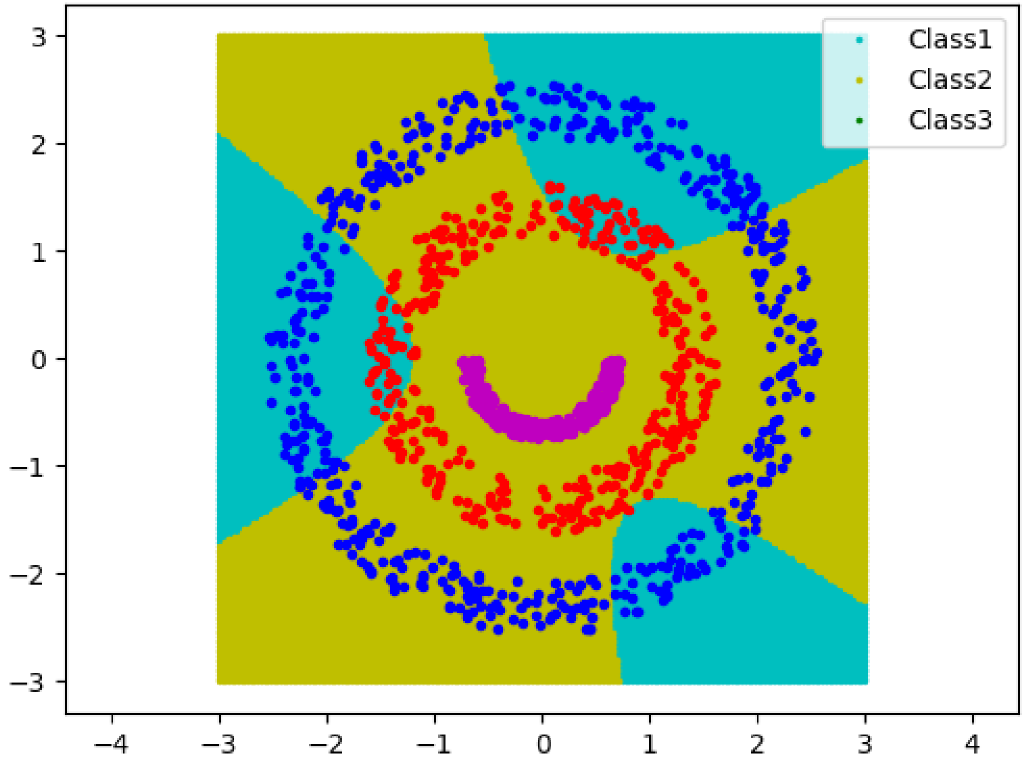


Classification Accuracy (%)	41.1%
Precision for Class 1 Precision for Class 2 Precision for Class 3	0.0 0.411 0.0
Mean Precision	0.13
Recall for Class 1 Recall for Class 2 Recall for Class 3	0.0 1.0 0.0
Mean Recall	0.333
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	0.00 0.583 0.0
Mean F-measure	0.194

Confusion Matrix :

$$C = \begin{matrix} 0 & 125 & 0 \\ 0 & 175 & 0 \\ 0 & 125 & 0 \end{matrix}$$

5.2.2 Polynomial(degree=3) Kernel:-



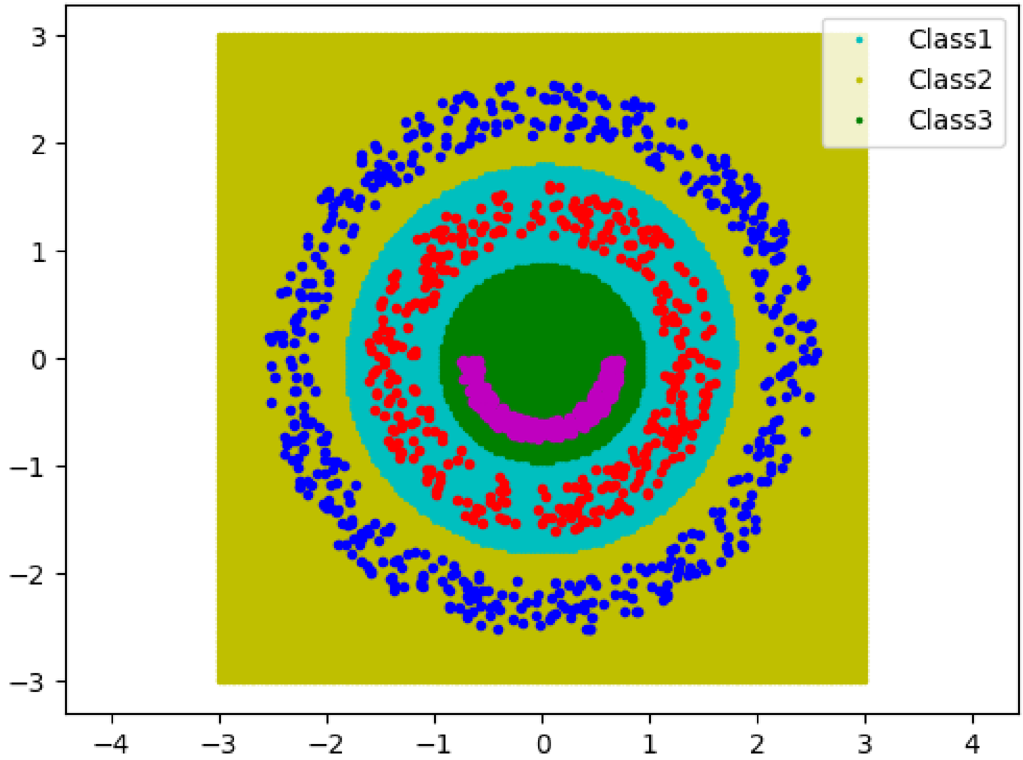
Classification Accuracy (%)	31%
Precision for Class 1 Precision for Class 2 Precision for Class 3	0.24 0.30 0
Mean Precision	0.18
Recall for Class 1 Recall for Class 2 Recall for Class 3	0.2 0.56 0
Mean Recall	0.255
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	0.22 0.396 0.0

Mean F-measure	0.206

Confusion Matrix :

$$C = \begin{matrix} 25 & 100 & 0 \\ 76 & 99 & 0 \\ 0 & 125 & 0 \end{matrix}$$

5.2.3 Gaussian Kernel:-



Classification Accuracy (%)	100%
Precision for Class 1 Precision for Class 2 Precision for Class 3	1.0 1.0 1.0
Mean Precision	1.0
Recall for Class 1 Recall for Class 2 Recall for Class 3	1.0 1.0 1.0
Mean Recall	1.0

F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	1.0 1.0 1.0
Mean F-measure	1.0

Confusion Matrix :

$$C = \begin{matrix} 125 & 0 & 0 \\ 0 & 125 & 0 \\ 0 & 0 & 125 \end{matrix}$$

5.3 Dataset 2:-

3 class scene image dataset

Class1:-CandyStore
Class2:-Football Stadium
Class3:-Forest BroadLeaf

One vs One approach was used to classify test data points.

5.3.1 Linear Kernel:-

Classification Accuracy (%)	67.3%
Precision for Class 1 Precision for Class 2 Precision for Class 3	0.589 0.833 0.694
Mean Precision	0.705
Recall for Class 1 Recall for Class 2 Recall for Class 3	0.92 0.6 0.5
Mean Recall	0.6733
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	0.718 0.697 0.581
Mean F-measure	0.665

Confusion Matrix :

$$C = \begin{matrix} 46 & 2 & 2 \\ 11 & 30 & 9 \\ 21 & 4 & 25 \end{matrix}$$

5.3.2 Polynomial(degree=2) Kernel:-

Classification Accuracy (%)	71.3%
Precision for Class 1	0.65

Precision for Class 2 Precision for Class 3	0.875 0.6875
Mean Precision	0.74
Recall for Class 1 Recall for Class 2 Recall for Class 3	0.92 0.56 0.66
Mean Recall	0.713
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	0.766 0.6829 0.610
Mean F-measure	0.676

Confusion Matrix :

$$C = \begin{matrix} 46 & 1 & 3 \\ 10 & 28 & 12 \\ 14 & 3 & 33 \end{matrix}$$

5.3.3 Gaussian Kernel:-

Classification Accuracy (%)	40%
Precision for Class 1 Precision for Class 2 Precision for Class 3	0.5 0.36 1.0
Mean Precision	0.62
Recall for Class 1 Recall for Class 2 Recall for Class 3	0.04 1.0 0.16
Mean Recall	0.4
F-measure for Class 1 F-measure for Class 2 F-measure for Class 3	0.07 0.54 0.275
Mean F-measure	0.293

Confusion Matrix :

$$C = \begin{matrix} 2 & 48 & 0 \\ 0 & 50 & 0 \\ 2 & 40 & 8 \end{matrix}$$

5.4 Observations:-

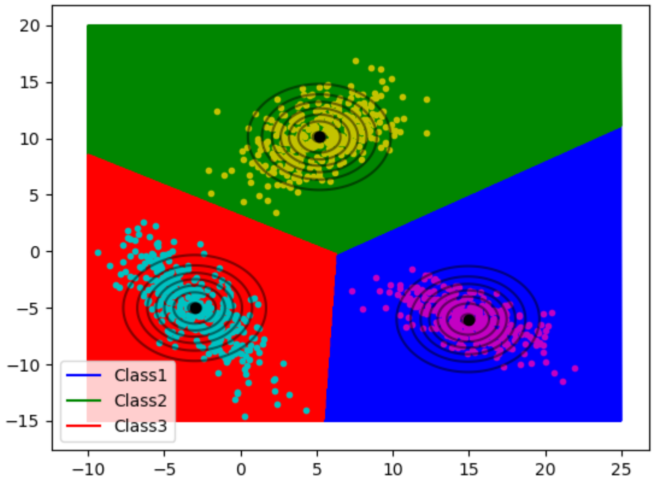
- For Linearly seperable dataset, 100% accuracy is observed irrespective of type of kernel

- For Non Linearly seperable dataset, for linear and polynomial kernel, accuracy is very poor, this is so because the error term C is very small and the kernel is not ideal for dataset.
- Linear decision boundary is obtained incase of linear kernel and non linear incase of gaussian and polynomial kernel.
- For nonlinearly separability, in gaussian kernel the decision boundary seems to perfectly fit the data.
- In Linearly separable dataset, the decision boundary incase of linear kernel is linear and maximises the margin.

6. Comparison with all Assignments.

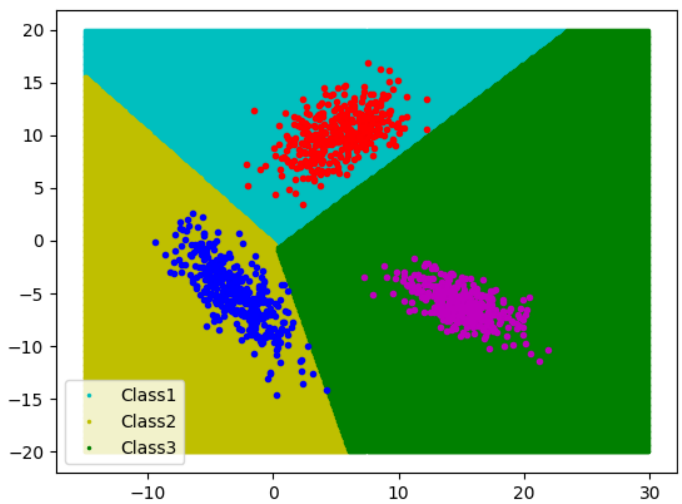
Dataset 1(a):-

Assignment1 Plots:-

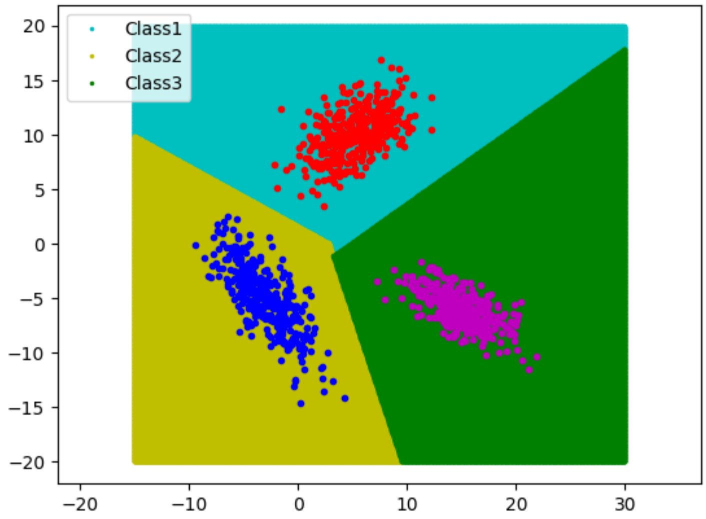


Assignment4 Best accuracy Plot:-

a) Perceptron:-



b) SVM:-



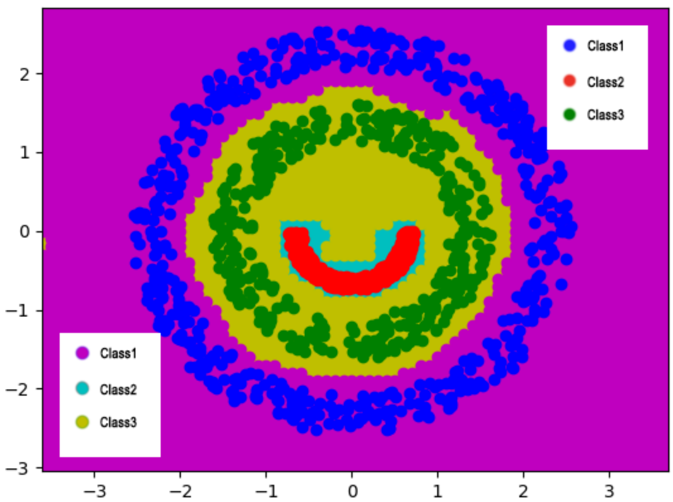
Observations:-

- For SVM, the linear decision boundary is in direction such that margin is maximized.
- Incase of Perceptron learning, The decision boundary is very close to class 2 and hence we do not observe 100% accuracy for it.
- For Bayes classifier in which class conditional density is determined by Unimodal Gaussian, the decision boundary is better than that found by perceptron but not as good as that found by SVM because it is not in direction of max margin between support vectors, it also has the disadvantage of curse of dimentionality problem.

Dataset 1(b):-

Assignment2 best Plot:-

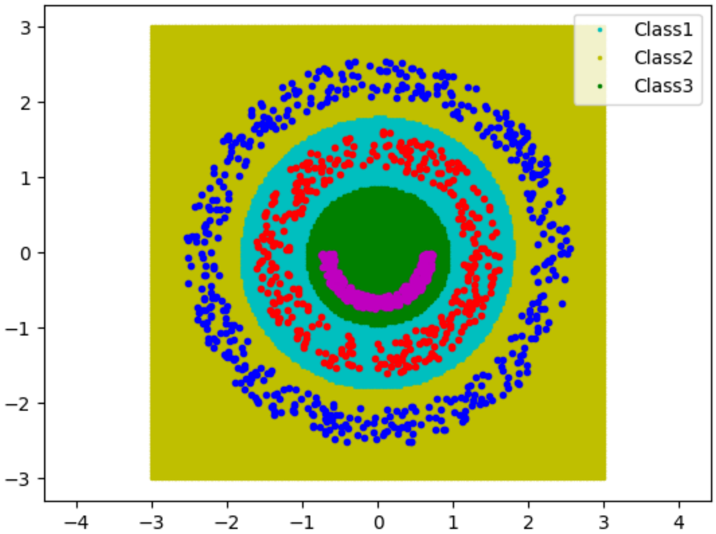
Multimodal Gaussian with K=8 clusters:-



Accuracy:-100%

Assignment 4 best Plot:-

SVM(RBF kernel):-



Accuracy:- 100%

6.2 Observations:-

- The Decision boundary found in 2nd assignment is not sharp and does not have a fixed shape. Although it does give 100% accuracies but it over fits the data, and the decision boundary is not sharp, on the other hand SVM with gaussian kernel gives a circular decision boundary which fits the data exactly and gives 100% accuracies.