Assignment 4

Archit Kumar,Avinav Sanyal,Sakhile Naga Koti Reddy

Monday 20$^{\text{th}}$ November, 2017

Contents

1 Theory

1.1 Principle Component Analysis

PCA is mathematically defined as an orthogonal linear transformation that transforms the data to a new coordinate system such that the greatest variance by some projection of the data comes to lie on the first coordinate (called the first principal component), the second greatest variance on the second coordinate, and so on.

1.1.1 Algorithm

- Mean center the data (optional)
- Compute the covariavce of the dimensions
- Find eigenvectors of covariance matrix
- Sort eigenvectors in desending order of eigenvalues
- Project onto eigenvectors in order

1.2 Fischer Discriminant Analysis

In Fischer Discriminant Analysis we find the direction in which separation in maximised. The maximum discriminant vector is the eigenvector of maximum eigen-value of Matrix D
s.t. $D = S_w^-1 * S_B$,
where $S_w and S_B$ are within class Scatter Matrix and between Class Scatter Matrix

1.3 Support Vector Analysis

Support vector machine are supervised learning models. We calculate support vectors for the data, the data maybe linearly separable or not. For linear sep data we can use linear SVM. For non-linear data, we have to transform it to higher dimensions using different kernels. We are using kernels such as quadratic, cubic, Radial-Basis Function.p In this assignment we used sklearn-SVM.

1.4 Perceptron

Perceptron is a supervised binary classifier. It is a linear classifier i.e. it can classify non-linear data. In this we find an equation of a line which separates the given data. At each step error of points lying on the wrong side of line is taken into account to change the equation of line.

2 Observations

2.1 Principal Component Analysis

2.1.1 Bag of Visual Words

- dim = 8
- clusters = 1

$$\begin{bmatrix} 9 & 13 & 28 \\ 10 & 8 & 32 \\ 13 & 3 & 34 \end{bmatrix}$$

Accuracy: 34%

Table 1: Results

	Precision	Recall	F-measure
Class 1	0.18	0.281	0.219
Class 2	0.16	0.333	0.216
Class 3	0.68	0.361	0.472
Average	0.34	0.325	0.302

- dim = 8
- cluster =2

$$\begin{bmatrix} 20 & 18 & 12 \\ 22 & 8 & 20 \\ 19 & 5 & 26 \end{bmatrix}$$

Accuracy : 36%

Table 2: Results

	Precision	Recall	F-measure
Class 1	0.4	0.327	0.360
Class 2	0.16	0.258	0.197
Class 3	0.52	0.448	0.481
Average	0.360	0.344	0.346

- dim = 8
- clusters = 4

$$\begin{bmatrix} 28 & 10 & 12 \\ 26 & 13 & 11 \\ 34 & 1 & 15 \end{bmatrix}$$

Accuracy:37.33%

Table 3: Results

	Precision	Recall	F-measure
Class 1	0.56	0.318	0.405
Class 2	0.26	0.541	0.351
Class 3	0.3	0.394	0.341
Average	0.373	0.418	0.366

- dim = 8
- clusters = 8

$$\begin{bmatrix} 13 & 4 & 33 \\ 15 & 3 & 32 \\ 5 & 2 & 43 \end{bmatrix}$$

Accuracy:39.33%

Table 4: Results

	Precision	Recall	F-measure
Class 1	0.26	0.393	0.313
Class 2	0.06	0.333	0.102
Class 3	0.86	0.398	0.544
Average	0.393	0.375	0.319

- dim = 4
- clusters = 1

$$\begin{bmatrix} 26 & 12 & 12 \\ 33 & 6 & 11 \\ 25 & 14 & 11 \end{bmatrix}$$

Accuracy:28.66%

Table 5: Results

	Precision	Recall	F-measure
Class 1	0.520	0.310	0.388
Class 2	0.120	0.188	0.146
Class 3	0.220	0.324	0.262
Average	0.287	0.274	0.265

- dim = 4
- clusters = 2

$$\begin{bmatrix} 28 & 15 & 7 \\ 22 & 7 & 21 \\ 15 & 15 & 20 \end{bmatrix}$$

Accuracy:36.66%

Table 6: Results

	Precision	Recall	F-measure
Class 1	0.56	0.430	0.486
Class 2	0.14	0.189	0.161
Class 3	0.4	0.416	0.408
Average	0.366	0.345	0.352

- dim = 4
- clusters = 4

$$\begin{bmatrix} 12 & 4 & 34 \\ 10 & 4 & 36 \\ 4 & 4 & 42 \end{bmatrix}$$

Accuracy:38.66%

Table 7: Results

	Precision	Recall	F-measure
Class 1	0.240	0.462	0.316
Class 2	0.080	0.333	0.129
Class 3	0.840	0.375	0.519
Average	0.387	0.390	0.321

- dim = 4
- clusters = 8

$$\begin{bmatrix} 21 & 11 & 18 \\ 13 & 5 & 32 \\ 13 & 3 & 34 \end{bmatrix}$$

Accuracy : 40%

Table 8: Results

	Precision	Recall	F-measure
Class 1	0.420	0.447	0.433
Class 2	0.100	0.263	0.145
Class 3	0.680	0.405	0.507
Average	0.400	0.372	0.362

- dim = 2
- clusters = 1

$$\begin{bmatrix} 28 & 9 & 13 \\ 27 & 12 & 11 \\ 32 & 7 & 11 \end{bmatrix}$$

Accuracy:34%

Table 9: Results

	Precision	Recall	F-measure
Class 1	0.560	0.322	0.409
Class 2	0.240	0.429	0.308
Class 3	0.220	0.314	0.259
Average	0.340	0.355	0.325

- dim = 2
- clusters = 2

$$\begin{bmatrix} 16 & 12 & 22 \\ 10 & 21 & 19 \\ 15 & 17 & 18 \end{bmatrix}$$

Accuracy: 36.66%

Table 10: Results

	Precision	Recall	F-measure
Class 1	0.320	0.390	0.352
Class 2	0.420	0.420	0.420
Class 3	0.360	0.305	0.330
Average	0.367	0.372	0.367

- dim = 2
- clusters = 4

$$\begin{bmatrix} 9 & 3 & 38 \\ 6 & 7 & 37 \\ 6 & 1 & 43 \end{bmatrix}$$

Accuracy:39.33%

Table 11: Results

	Precision	Recall	F-measure
Class 1	0.180	0.429	0.254
Class 2	0.140	0.636	0.230
Class 3	0.860	0.364	0.512
Average	0.393	0.476	0.332

- dim = 2
- clusters = 8

$$\begin{bmatrix} 17 & 15 & 18 \\ 10 & 27 & 13 \\ 15 & 18 & 17 \end{bmatrix}$$

Accuracy : 40.66%

Table 12: Results

	Precision	Recall	F-measure
Class 1	0.340	0.405	0.370
Class 2	0.540	0.450	0.491
Class 3	0.340	0.354	0.347
Average	0.407	0.403	0.402

- dim = 1
- clusters = 1

$$\begin{bmatrix} 4 & 13 & 33 \\ 7 & 0 & 43 \\ 4 & 11 & 35 \end{bmatrix}$$

Accuracy:31.33

Table 13: Results

	Precision	Recall	F-measure
Class 1	0.08	0.267	0.123
Class 2	0	0	not def 0.356
Class 3	0.70	0.315	0.434
Average	0.26	0.194	not def

- dim = 1
- clusters = 2

$$\begin{bmatrix} 17 & 22 & 11 \\ 7 & 22 & 21 \\ 13 & 23 & 14 \end{bmatrix}$$

Accuracy:35.33%

Table 14: Results

	Precision	Recall	F-measure
Class 1	0.340	0.459	0.391
Class 2	0.440	0.328	0.376
Class 3	0.280	0.304	0.292
Average	0.353	0.364	0.353

- dim = 1
- clusters = 4

$$\begin{bmatrix} 36 & 3 & 11 \\ 20 & 4 & 26 \\ 31 & 1 & 18 \end{bmatrix}$$

Accuracy: 38.66%

Table 15: Results

	Precision	Recall	F-measure
Class 1	0.720	0.414	0.526
Class 2	0.080	0.500	0.138
Class 3	0.360	0.327	0.343
Average	0.387	0.414	0.335

- dim = 1
- clusters = 8

$$\begin{bmatrix} 21 & 17 & 12 \\ 5 & 26 & 19 \\ 10 & 23 & 17 \end{bmatrix}$$

Accuracy: 42.66%

Table 16: Results

	Precision	Recall	F-measure
Class 1	0.420	0.583	0.488
Class 2	0.520	0.394	0.448
Class 3	0.340	0.354	0.347
Average	0.427	0.444	0.428

2.1.2 Observations

Here we can see the clasification accuracy didnot follow any order with the dimension as we know pca destroys clasification accuracy. As we see from the data that if we increase the number of clusters the accuracy is generally increasing as we can now better parameterise the data.

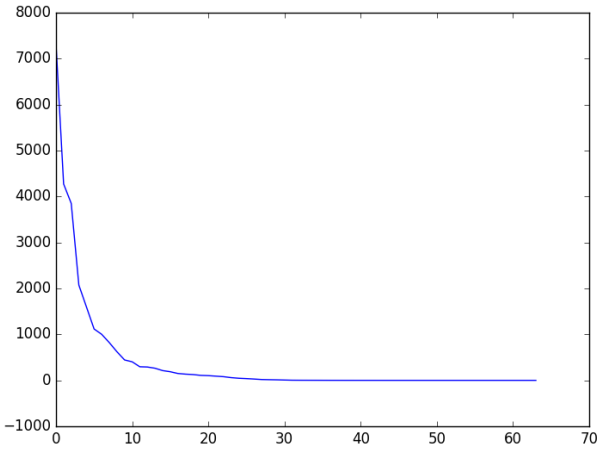


Figure 1: Eigen Values for Arch Dataset

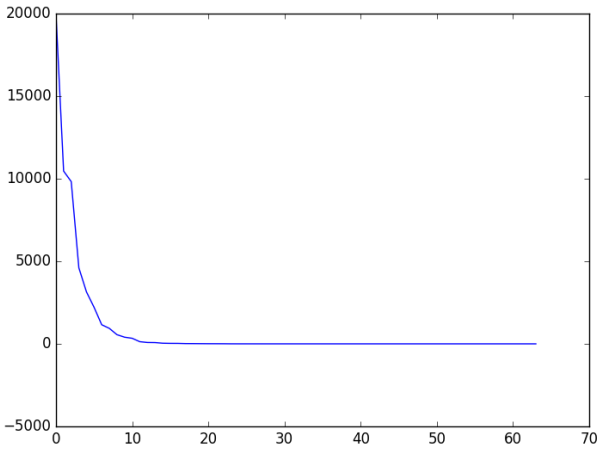


Figure 2: Eigen Values for Forest Path Dataset

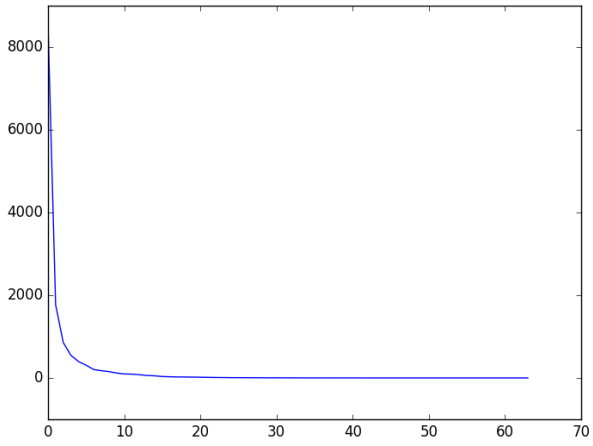
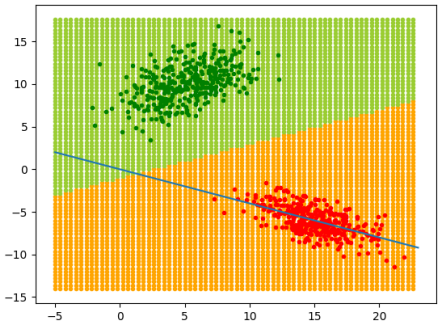


Figure 3: Eigen Values for Highway Dataset

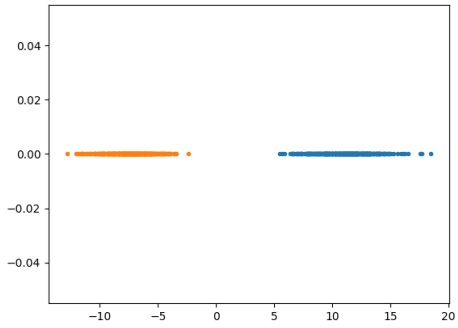
2.2 Fischer Discriminant Analysis

2.2.1 For linearly separable data

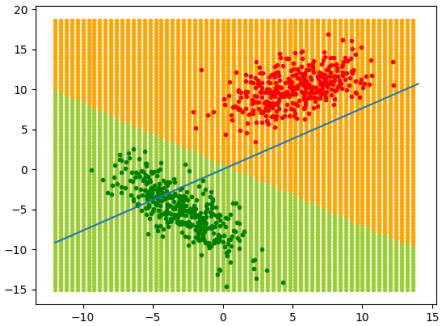
Pair-wise FDA Gaussian



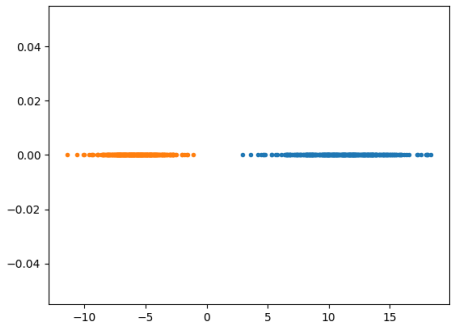
FDA on Class 1 and 2



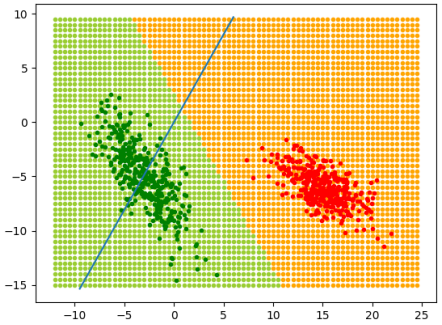
Projection of Class 1 and 2



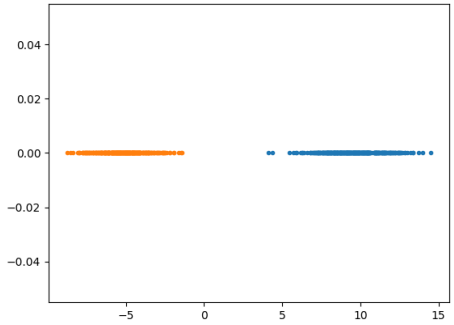
FDA on Class 2 and 3



Projection of Class 2 and 3



FDA on Class 1 and 3



Projection of Class 1 and 3

Using FDA Gaussian classifier

$$\begin{bmatrix} 125 & 0 & 0 \\ 0 & 125 & 0 \\ 0 & 0 & 125 \end{bmatrix}$$

Accuracy: 100%

Table 17: Results

	Precision	Recall	F-measure
Class 1	1	1	1
Class 2	1	1	1
Class 3	1	1	1
Average	1	1	1

Using FDA GMM classifier

$$\begin{bmatrix} 125 & 0 & 0 \\ 0 & 125 & 0 \\ 0 & 0 & 125 \end{bmatrix}$$

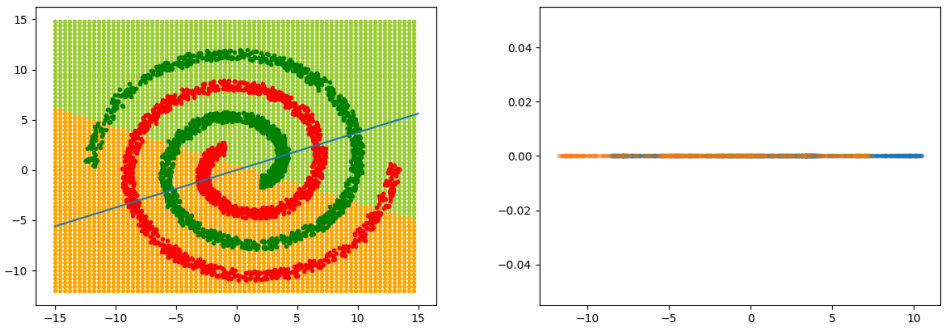
Accuracy: 100%

Table 18: Results

	Precision	Recall	F-measure
Class 1	1	1	1
Class 2	1	1	1
Class 3	1	1	1
Average	1	1	1

2.2.2 For Non Linearly Separable data

Pair-wise FDA Gaussian



Projection of Class 1 and 2

Using Gaussian FDA

$$\begin{bmatrix} 377 & 234 \\ 231 & 380 \end{bmatrix}$$

Accuracy: 61.94%

Table 19: Results

	Precision	Recall	F-measure
Class 1	0.617	0.620	0.619
Class 2	0.622	0.619	0.620
Average	0.413	0.413	0.413

Using FDA GMM Classifier

$$\begin{bmatrix} 408 & 203 \\ 148 & 463 \end{bmatrix}$$

Accuracy: 71.27%

Table 20: Results

	Precision	Recall	F-measure
Class 1	0.668	0.734	0.699
Class 2	0.758	0.695	0.725
Average	0.475	0.476	0.475

2.2.3 For Bovw

Projections Pair-wise

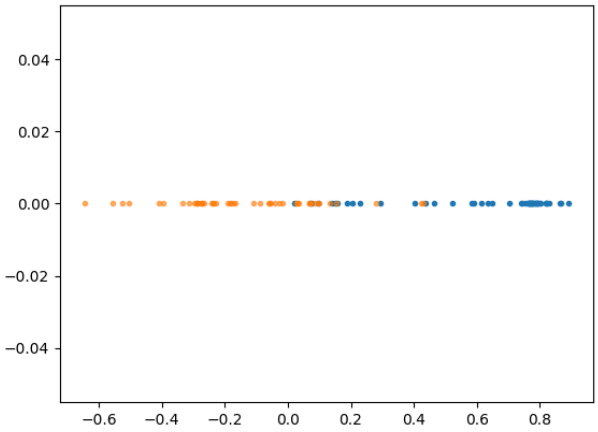


Figure 4: Projection of Class 1 and 2

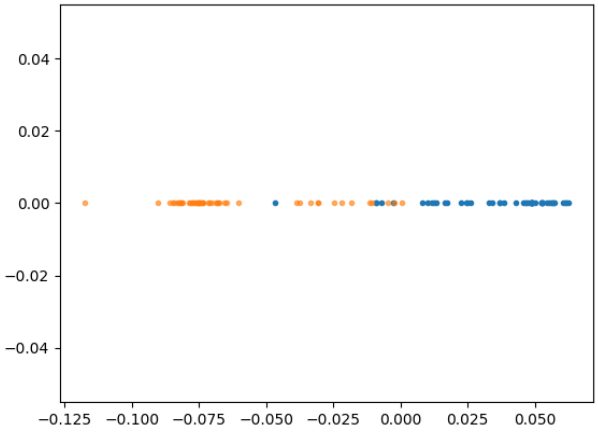


Figure 5: Projection of Class 2 and 3

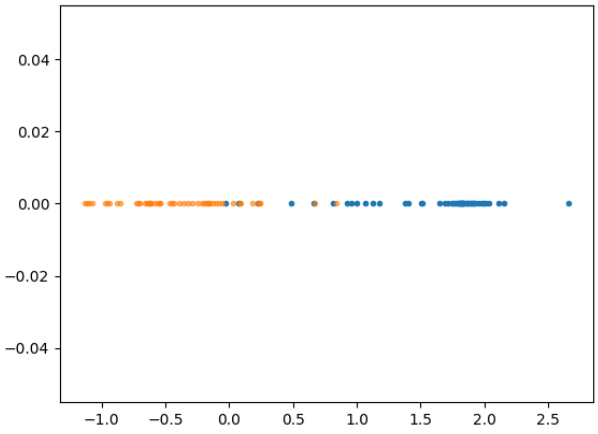


Figure 6: Projection of Class 1 and 3

Using FDA Gaussian Classifier

$$\begin{bmatrix} 19 & 9 & 22 \\ 16 & 28 & 6 \\ 17 & 1 & 32 \end{bmatrix}$$

Accuracy: 52.66%

Table 21: Results

	Precision	Recall	F-measure
Class 1	0.380	0.365	0.373
Class 2	0.560	0.737	0.636
Class 3	0.640	0.533	0.582
Average	0.527	0.545	0.530

Using GMM classifier using clusters = 2

$$\begin{bmatrix} 25 & 6 & 19 \\ 10 & 17 & 23 \\ 14 & 5 & 31 \end{bmatrix}$$

Accuracy: 48.66%

Table 22: Results

	Precision	Recall	F-measure
Class 1	0.500	0.510	0.505
Class 2	0.340	0.607	0.436
Class 3	0.620	0.425	0.504
Average	0.487	0.514	0.482

2.3 Support Vector Machine

2.3.1 Linearly Separable Data

Using Linear Kernel

$$\begin{bmatrix} 125 & 0 & 0 \\ 0 & 125 & 0 \\ 0 & 0 & 125 \end{bmatrix}$$

Accuracy: 100%

Table 23: Results

	Precision	Recall	F-measure
Class 1	1	1	1
Class 2	1	1	1
Class 3	1	1	1
Average	1	1	1

Using Radial Basis Function Kernel

$$\begin{bmatrix} 125 & 0 & 0 \\ 0 & 125 & 0 \\ 0 & 0 & 125 \end{bmatrix}$$

Accuracy: 100%

Table 24: Results

	Precision	Recall	F-measure
Class 1	1	1	1
Class 2	1	1	1
Class 3	1	1	1
Average	1	1	1

Using Quadratic Kernel

$$\begin{bmatrix} 125 & 0 & 0 \\ 0 & 125 & 0 \\ 0 & 0 & 125 \end{bmatrix}$$

Accuracy: 100%

Table 25: Results

	Precision	Recall	F-measure
Class 1	1	1	1
Class 2	1	1	1
Class 3	1	1	1
Average	1	1	1

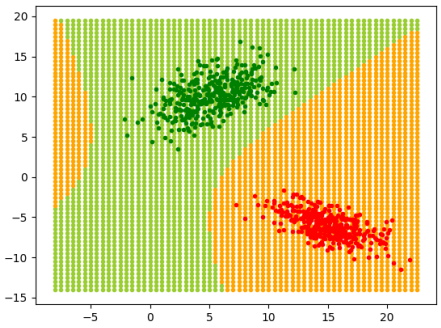


Figure 7: LS12_cubic

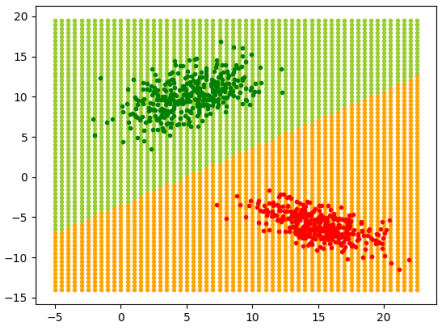


Figure 8: LS12_linear

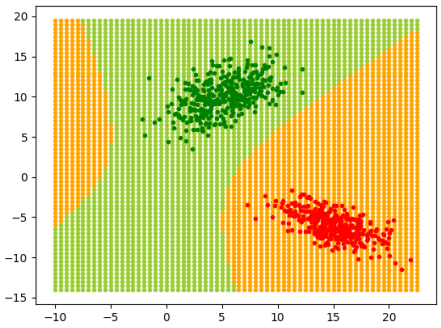


Figure 9: LS12_poly

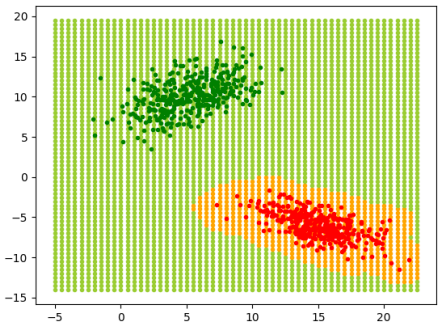


Figure 10: LS12_rbf

Figure 11: LS13_cubic

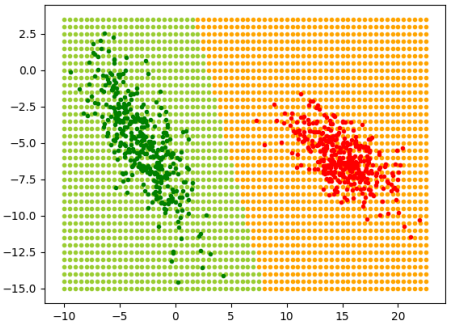


Figure 12: LS13_linear

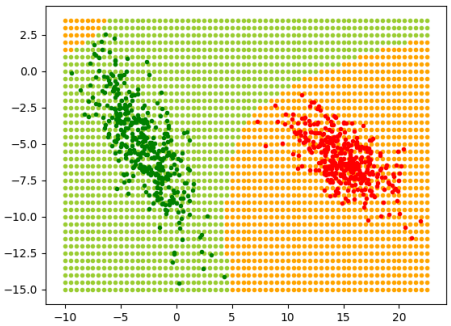


Figure 13: LS13_poly

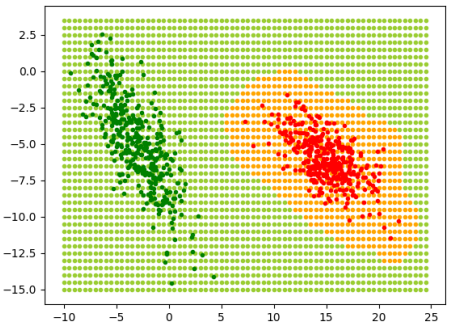


Figure 14: LS13_rbf

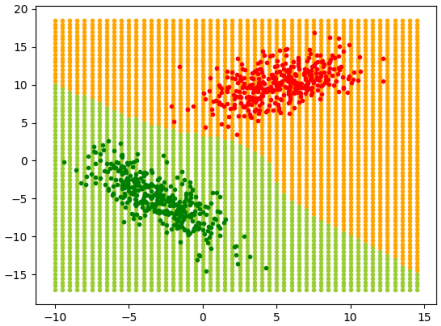


Figure 15: LS23_cubic

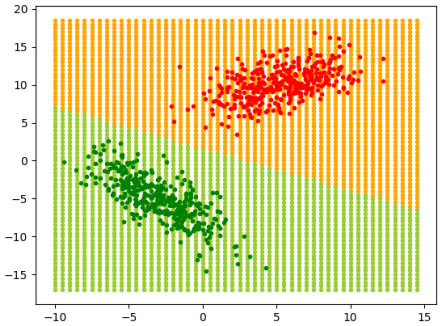


Figure 16: LS23_linear

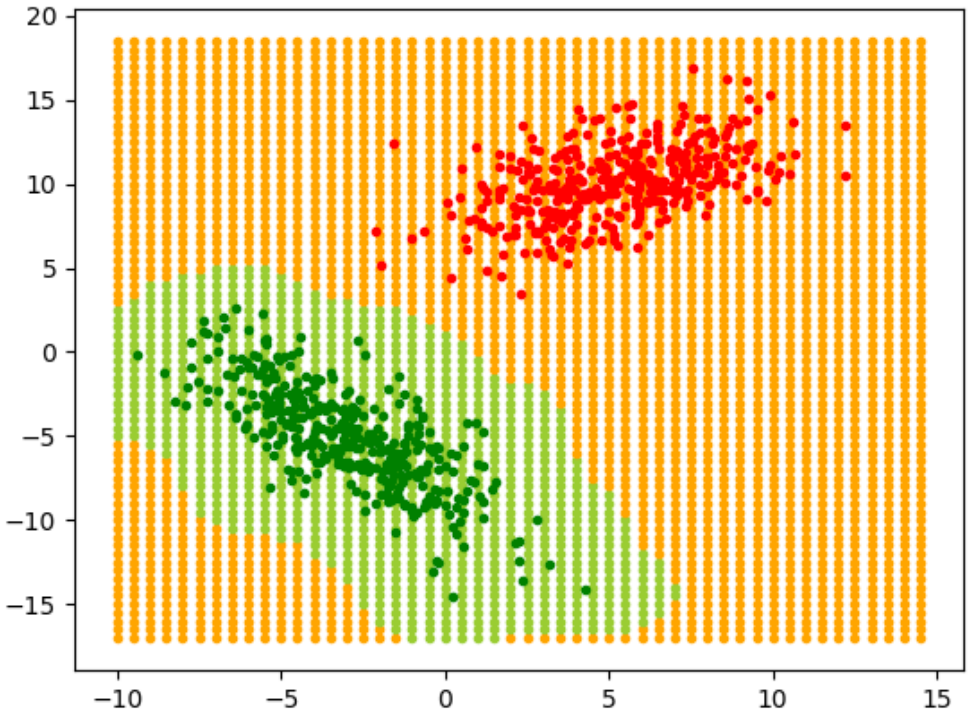


Figure 17: LS23_rbf

2.3.2 Non linearly separable Data

Using Linear Kernel

$$\begin{bmatrix} 372 & 239 \\ 228 & 383 \end{bmatrix}$$

Accuracy: 61.78%

Table 26: Results

	Precision	Recall	F-measure
Class 1	0.609	0.620	0.614
Class 2	0.627	0.616	0.621
Average	0.412	0.412	0.412

Using Radial Basis Function Kernel
Default Gamma

$$\begin{bmatrix} 611 & 0 \\ 0 & 611 \end{bmatrix}$$

Accuracy: 100%

Table 27: Results

	Precision	Recall	F-measure
Class 1	1	1	1
Class 3	1	1	1
Average	1	1	1

Using Radial Basis Function Gamma = 0.1

$$\begin{bmatrix} 611 & 0 \\ 0 & 611 \end{bmatrix}$$

Accuracy: 100%

Table 28: Results

	Precision	Recall	F-measure
Class 1	1	1	1
Class 3	1	1	1
Average	1	1	1

Using RBF gamma = 1e-2

$$\begin{bmatrix} 417 & 194 \\ 164 & 447 \end{bmatrix}$$

Accuracy: 70.70%

Table 29: Results

	Precision	Recall	F-measure
Class 1	0.682	0.718	0.700
Class 2	0.732	0.697	0.714
Average 0.471	0.472	0.471	

Using RBF gamma = 1e-3

$$\begin{bmatrix} 300 & 311 \\ 179 & 432 \end{bmatrix}$$

Accuracy: 59.90%

Table 30: Results

	Precision	Recall	F-measure
Class 1	0.491	0.626	0.550
Class 2	0.707	0.581	0.638
Average	0.399	0.403	0.396

Using RBF kernel gamma = 1e-4

$$\begin{bmatrix} 372 & 239 \\ 226 & 385 \end{bmatrix}$$

Accuracy: 61.94%

Table 31: Results

	Precision	Recall	F-measure
Class 1	0.609	0.622	0.615
Class 2	0.630	0.617	0.623
Average	0.413	0.413	0.413

Using Quadratic Kernel

$$\begin{bmatrix} 525 & 86 \\ 502 & 109 \end{bmatrix}$$

Accuracy: 51.88%

Table 32: Results

	Precision	Recall	F-measure
Class 1	0.859	0.511	0.641
Class 2	0.178	0.559	0.270
Average	0.346	0.357	0.304

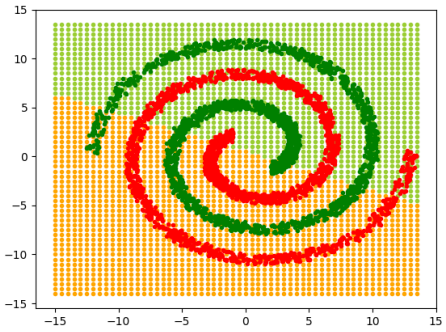


Figure 18: nls_linear

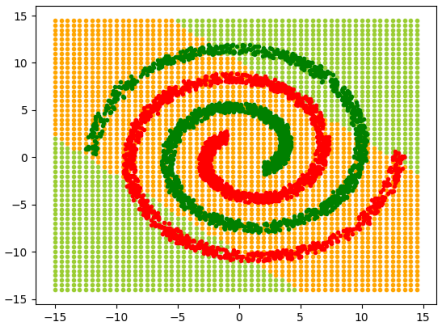


Figure 19: nls_quadratic

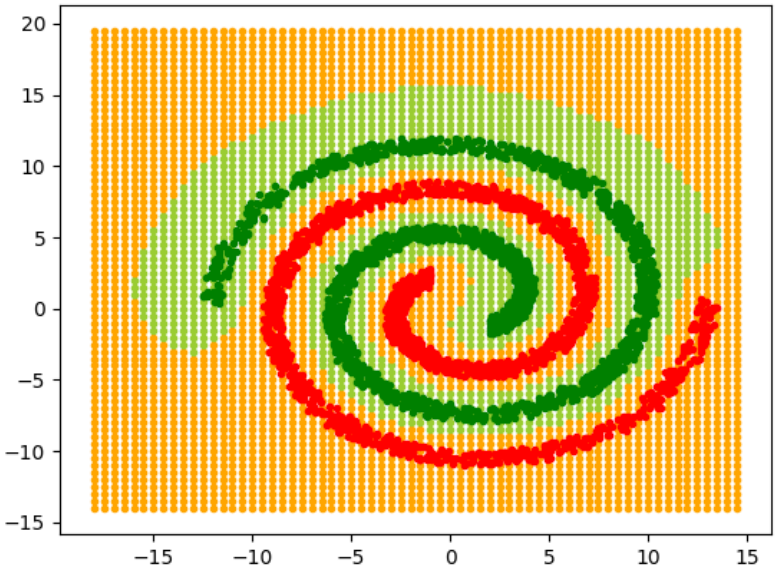


Figure 20: nls_rbf

Observations: As we can see from the Confusion Matrices, the best Kernel is Radial Basis Function.
Default value of Gamma is selected by Radial Basis func is (1 / number of features).

2.3.3 Bovw

Using linear kernel

$$\begin{bmatrix} 31 & 12 & 7 \\ 11 & 38 & 1 \\ 10 & 5 & 35 \end{bmatrix}$$

Accuracy: 69.3%

Table 33: Results

	Precision	Recall	F-measure
Class 1	0.620	0.596	0.608
Class 2	0.760	0.691	0.724
Class 3	0.700	0.814	0.753
Average	0.693	0.700	0.695

Using RBF kernel Using Default gamma

$$\begin{bmatrix} 46 & 2 & 2 \\ 32 & 18 & 0 \\ 27 & 2 & 21 \end{bmatrix}$$

Accuracy: 56.66%

Table 34: Results

	Precision	Recall	F-measure
Class 1	0.920	0.438	0.594
Class 2	0.360	0.818	0.500
Class 3	0.420	0.913	0.575
Average	0.567	0.723	0.556

Using gamma = 1e-4

$$\begin{bmatrix} 18 & 27 & 5 \\ 13 & 36 & 1 \\ 8 & 20 & 22 \end{bmatrix}$$

Accuracy: 56.66%

Table 35: Results

	Precision	Recall	F-measure
Class 1	0.360	0.462	0.404
Class 2	0.720	0.434	0.541
Class 3	0.440	0.786	0.564
Average	0.507	0.560	0.503

Using gamma=1e-3

$$\begin{bmatrix} 26 & 20 & 4 \\ 19 & 31 & 0 \\ 12 & 23 & 15 \end{bmatrix}$$

Accuracy: 48.0%

Table 36: Results

	Precision	Recall	F-measure
Class 1	0.520	0.456	0.486
Class 2	0.620	0.419	0.500
Class 3	0.300	0.789	0.435
Average	0.480	0.555	0.474

Using gamma=1e-2

$$\begin{bmatrix} 18 & 3 & 29 \\ 4 & 18 & 28 \\ 2 & 3 & 45 \end{bmatrix}$$

Accuracy: 54.0%

Table 37: Results

	Precision	Recall	F-measure
Class 1	0.360	0.750	0.486
Class 2	0.360	0.750	0.486
Class 3	0.900	0.441	0.592
Average	0.540	0.647	0.522

Using Quadratic kernel

$$\begin{bmatrix} 25 & 13 & 12 \\ 10 & 39 & 1 \\ 5 & 4 & 41 \end{bmatrix}$$

Accuracy: 70.0%

Table 38: Results

	Precision	Recall	F-measure
Class 1	0.500	0.625	0.556
Class 2	0.780	0.696	0.736
Class 3	0.820	0.759	0.788
Average	0.700	0.694	0.693

Using polynomial kernel of degree 4

$$\begin{bmatrix} 24 & 10 & 16 \\ 9 & 34 & 7 \\ 5 & 3 & 42 \end{bmatrix}$$

Accuracy:66.66%

Table 39: Results

	Precision	Recall	F-measure
Class 1	0.480	0.632	0.545
Class 2	0.680	0.723	0.701
Class 3	0.840	0.646	0.730
Average	0.667	0.667	0.659

Higher degree polynomial kernel SVM for Dataset 2 is running much faster compared to Dataset 1.

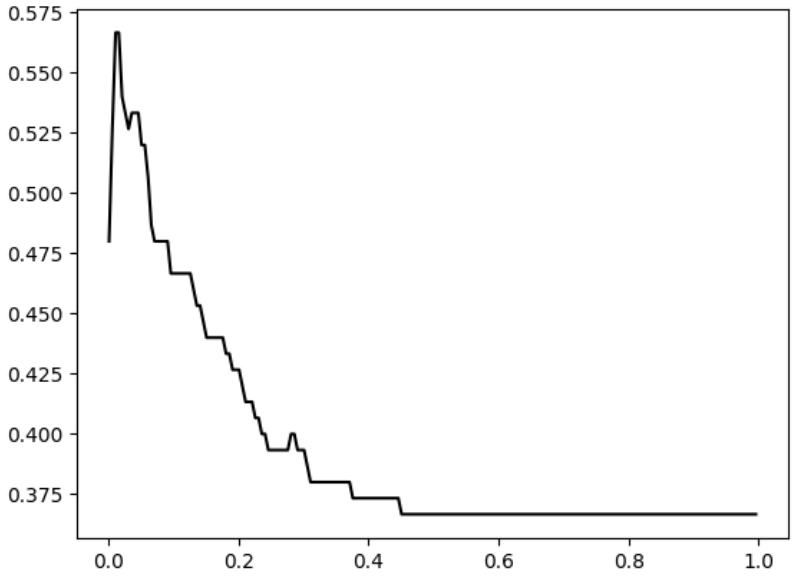


Figure 21: Varying Gamma parameter of SVM for BOVW dataset

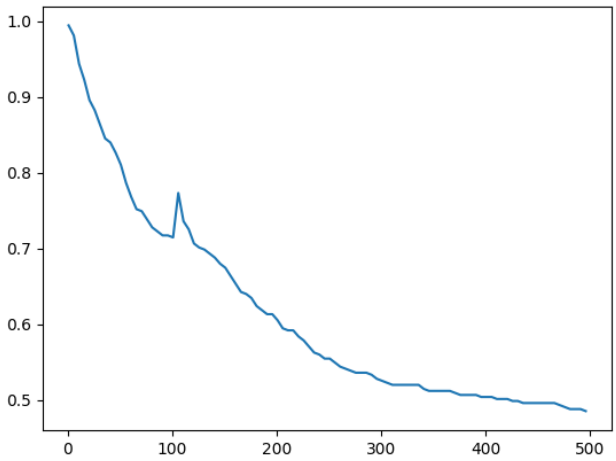


Figure 22: Varying Gamma parameter of SVM for LS dataset

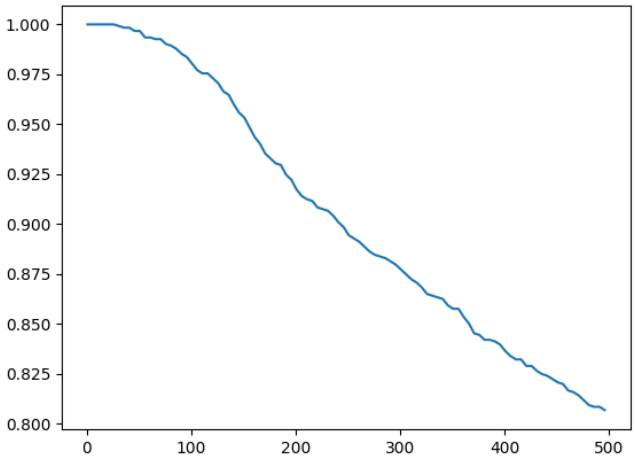


Figure 23: Varying Gamma parameter of SVM for NLS dataset

2.4 Peceptron

2.4.1 Linearly separable data

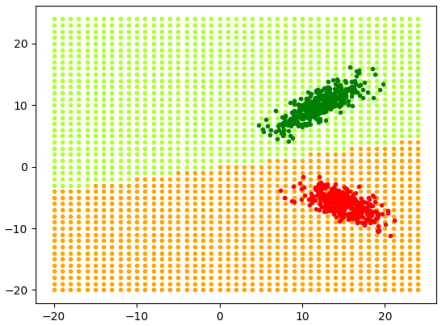


Figure 24: perc0_1_eta0.2

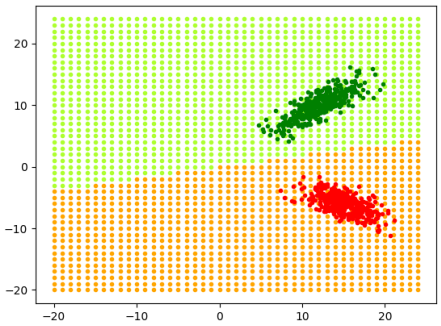


Figure 25: perc0_1_eta0.4

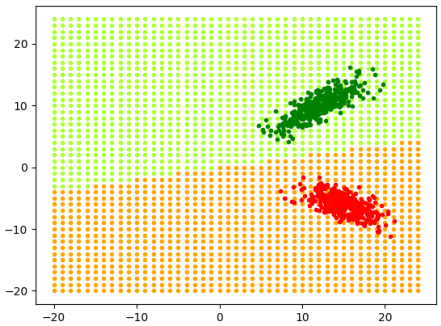


Figure 26: perc0_1_eta0.6

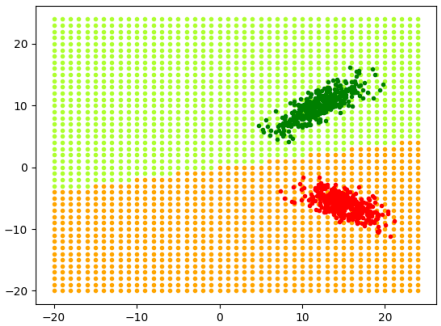


Figure 27: perc0_1_eta0.8

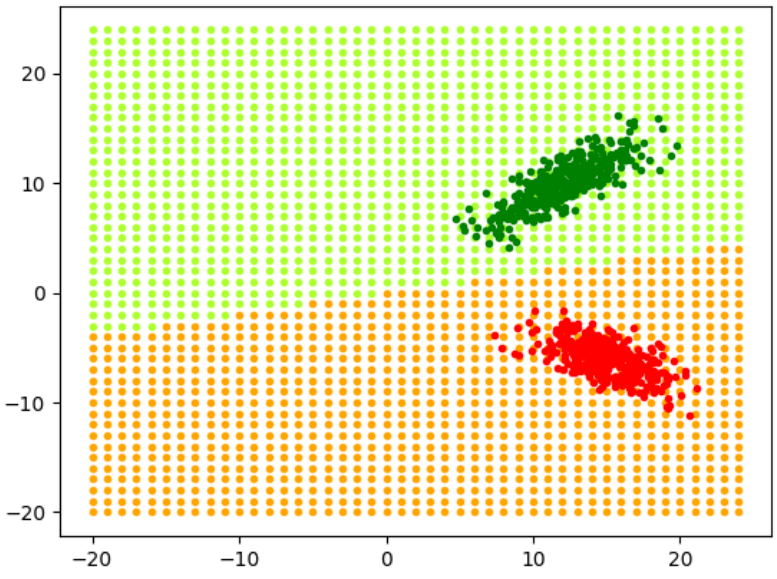


Figure 28: perc0_1_eta1.0

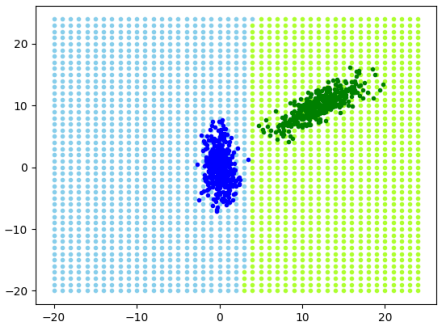


Figure 29: perc1_2_eta0.2

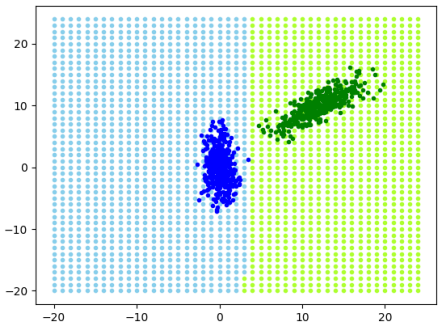


Figure 30: perc1_2_eta0.4

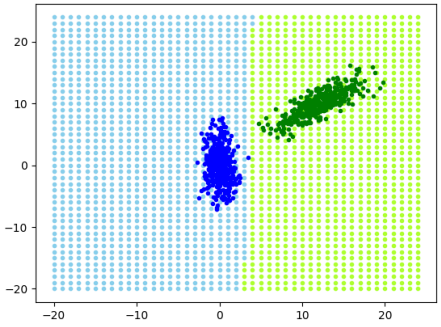


Figure 31: perc1_2_eta0.6

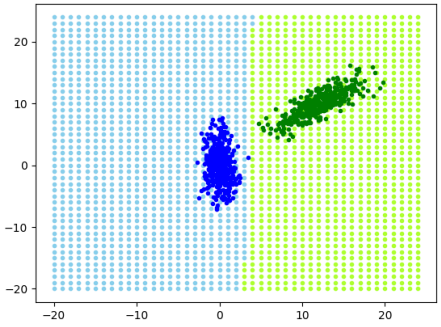


Figure 32: perc1_2_eta0.8

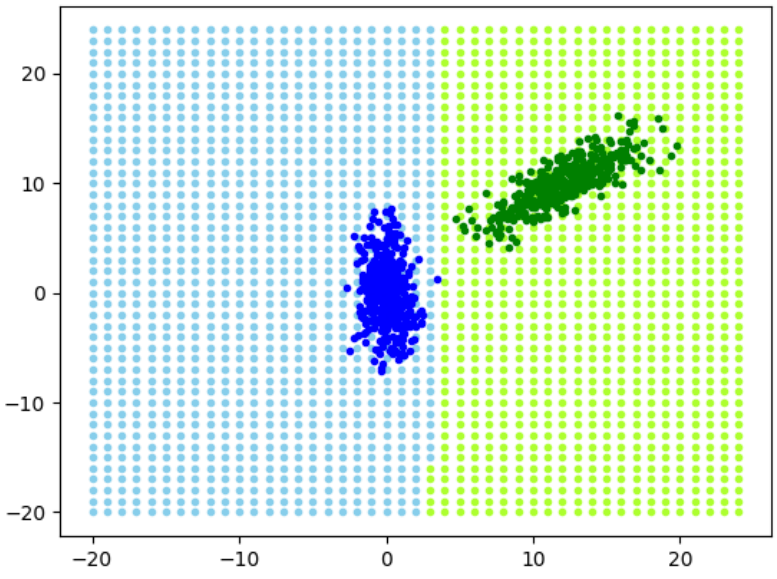


Figure 33: perc1_2_eta1.0

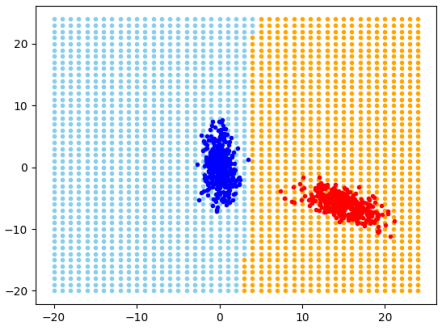


Figure 34: perc2_0_eta0.2

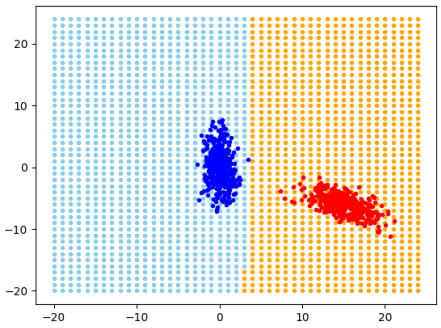


Figure 35: perc2_0_eta0.4

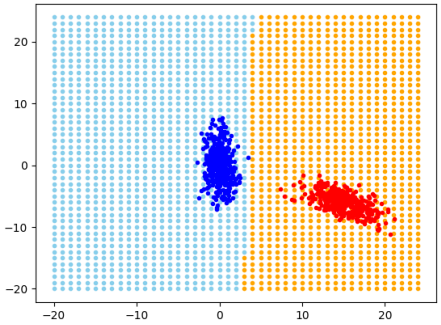


Figure 36: perc2_0_eta0.6

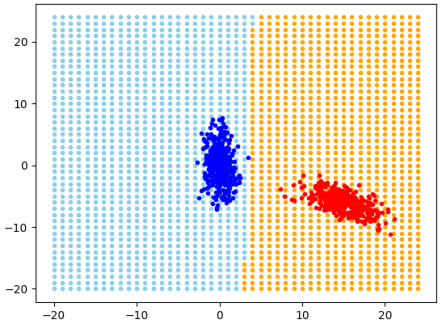


Figure 37: perc2_0_eta0.8

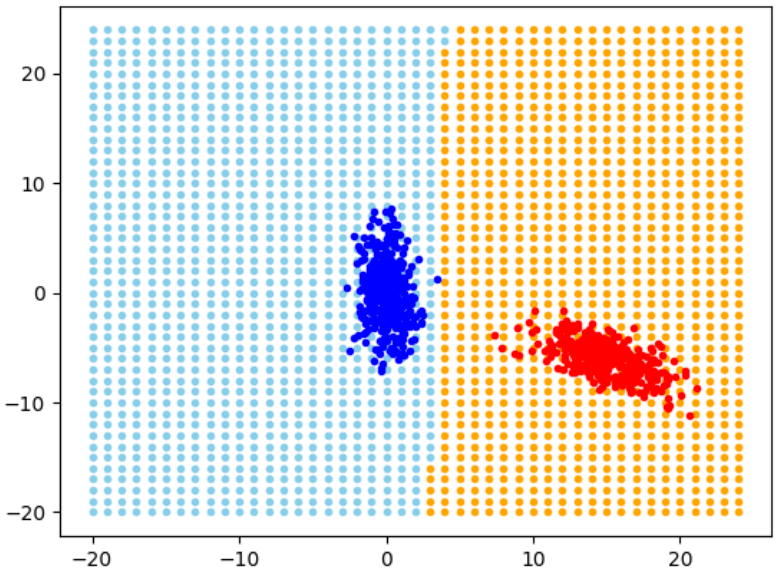


Figure 38: perc2_0_eta1.0

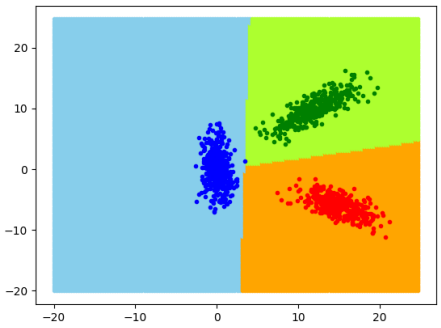


Figure 39: percAll_eta0.2

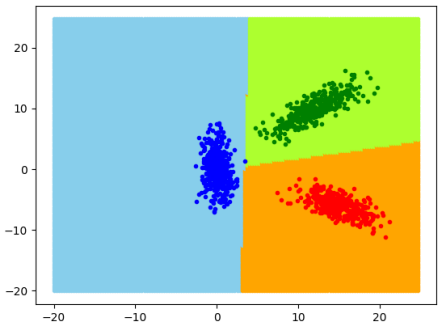


Figure 40: percAll_eta0.4

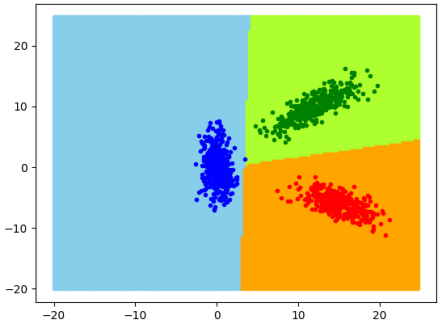


Figure 41: percAll_eta0.6

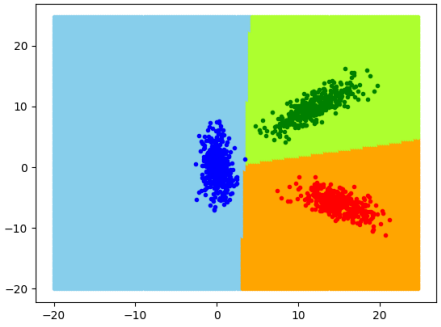


Figure 42: percAll_eta0.8

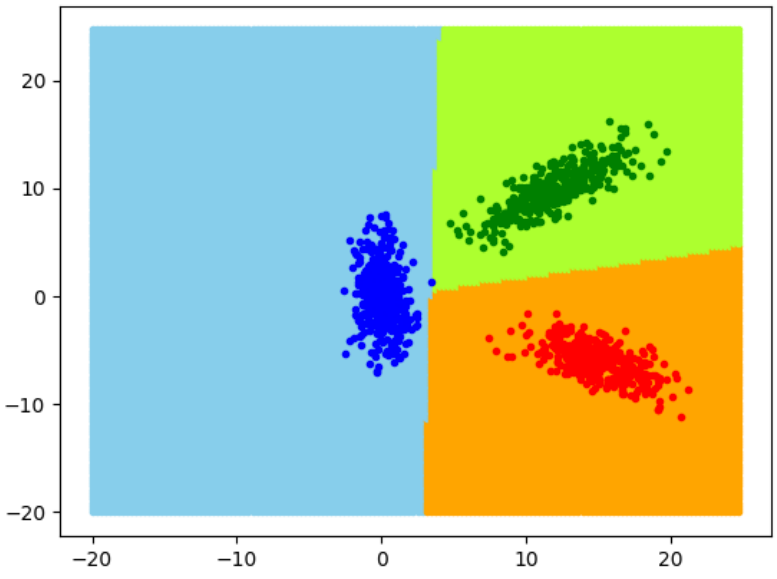


Figure 43: percAll_eta1.0

$$\begin{bmatrix} 125 & 0 & 0 \\ 0 & 125 & 0 \\ 0 & 0 & 125 \end{bmatrix}$$

Note : The accuracy is 100% in all the cases. The precision, Recall, F-measure is 1 for all the classes in all the classifiers built using Perceptron. The mean-precision, mean-recall and mean-F-measure is also 1 for all the classifiers built using Perceptron.

3 Conclusions

3.1 Linearly Separable

Always 100 % accuracy in all the assignments.

3.2 Non-linearly Separable Data

3.2.1 Assignement 1

Average accuracy using Naive Bayes : 63.09%

3.2.2 Assignment 2

For clusters >= 10 :
Accuracy : 100%

3.2.3 Assignment 4

- For FDA:
  Average Accuracy:66.63%
- For SVM:
  Linear Kernel : 61.78%
  RBF kernel : 100%

As we can see Bayes classifier is not a good way to classify non-linear data. GMM is the best way to classify. SVM with Radial Basis Function is also a good way to classify as it transforms to a better dimension.

3.3 Bag Of Visual Words

3.3.1 Assignment 2

For GMM:

- Clusters = [1,1,1]
  Accuracy : 59.4%
- Clusters = [2,2,2]
  Accuracy : 63.3%

3.3.2 Assignment 4

- SVM:
  Linear Kernel:
  Accuracy : 69.3%

  RBF kernel:
  Accuracy : 56.66%

  Quadratic Kernel:
  Accuracy : 70%

  Polynomial Kernel (degreee 4):
  Accuracy : 66.66%

- FDA:
  Gaussian :
  Accuracy : 52.66%

  GMM (cluster = 2):
  Accuracy : 48.66%

- PCA:
  Dim 1: 42.66%
  Dim 2: 40.66%
  Dim 4: 38.66%
  Dim 8: 39.33%

The classification accuracy of GM is maximum 63.3%. Linear and quadratic Kernels in SVM also catches up with Accuracy 69.3% and 70% respectively. FDA and PCA decreases the classification accuracy as some information to characterise data is lost.