
Pattern Recognition

CS669

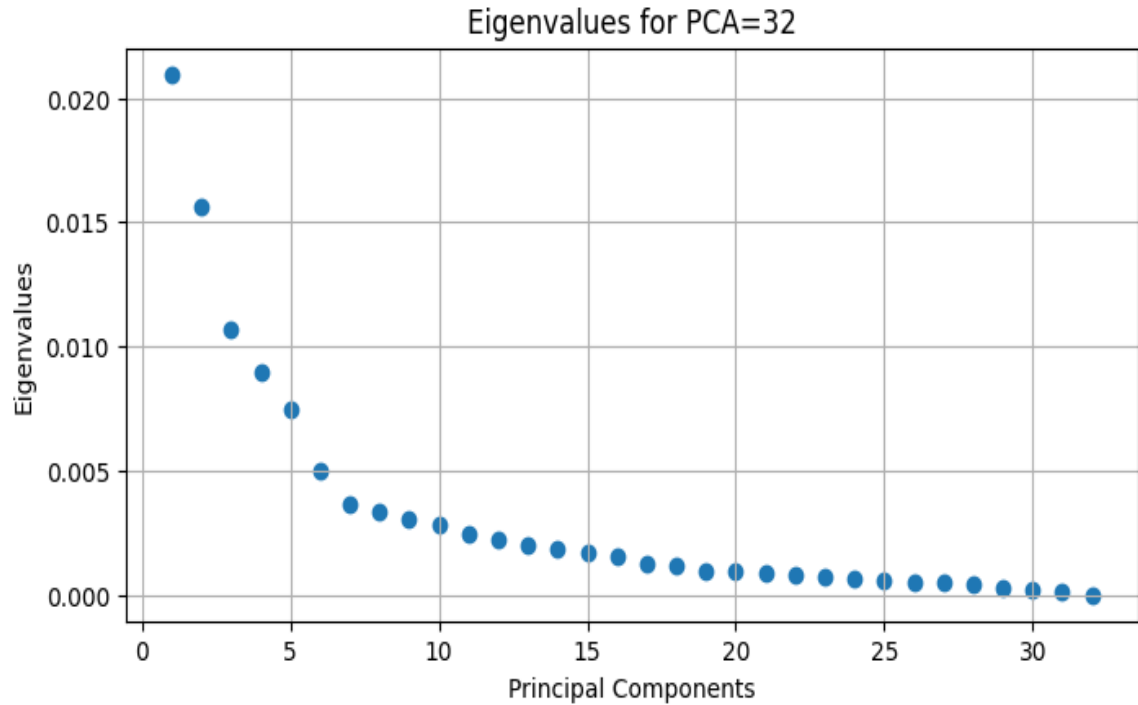
ASSIGNMENT 5



Group Number 18

Abhishek Singh Rawat	T23191
Arpit Dua	T23192
Rohit Kumar Roy	S22048

1. Build Bayes classifier using Gaussian mixture model (GMM) with 1, 2, 4 and 8 mixtures on the reduced dimensional representations of Dataset-2 obtained using PCA. • Perform the experiments on different values of l (including $l=1$), the reduced dimensions in PCA.

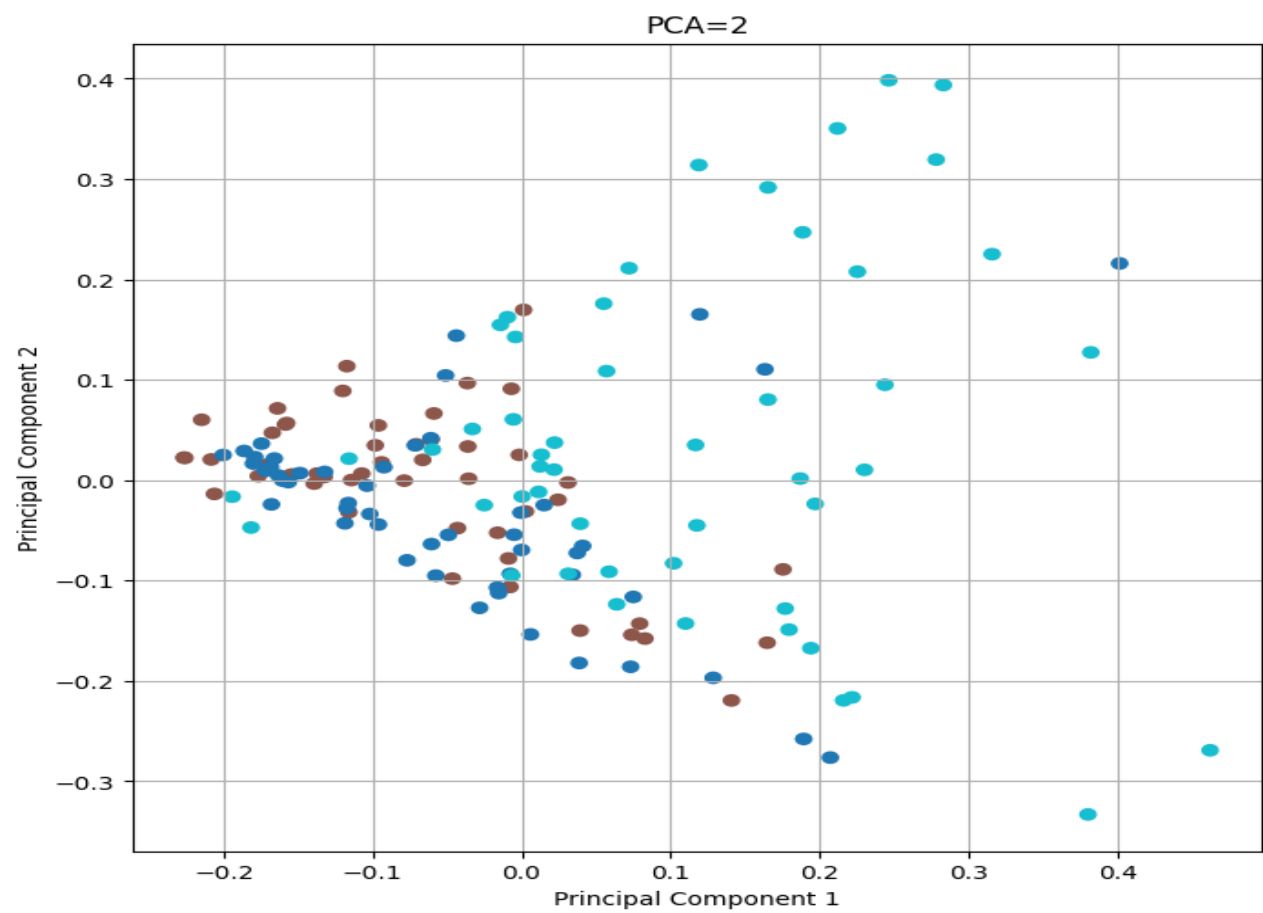


Plot of eigenvalues in ascending order during PCA

Inference from the graph:

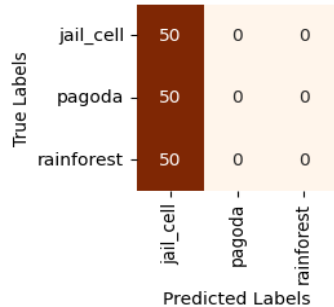
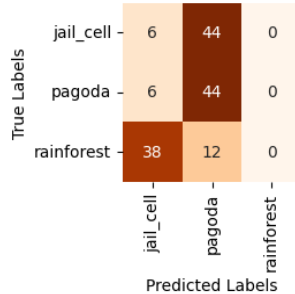
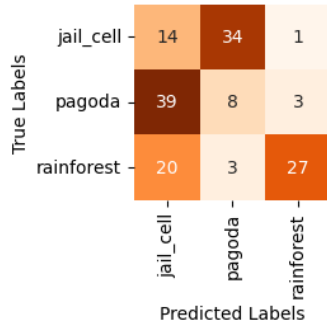
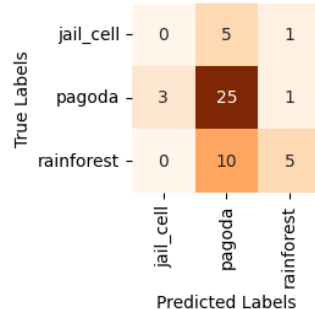
- 20 seems to be the ideal number of principal components for us to represent the 32 dimensional original data to a lower dimension while retaining as much variance as possible.
- Since we need to consider distinct values of the lower dimensions/principal components hence the graph does not have a continuous line, it has a discrete nature.

Plot of 2-dimensional reduced dimensional representations using PCA.



- Results and Observations for the different number of principal components:

A) PCA components = 32

GMM = 1	GMM = 2																																
<p>Accuracy = 0.33 Precision = 0.77 Recall = 0.33 F1 Score = 0.16</p> <p>Confusion Matrix for PCA=32, GMM Mixtures=1</p>  <table><tr><th>True Labels \ Predicted Labels</th><th>jail_cell</th><th>pagoda</th><th>rainforest</th></tr><tr><th>jail_cell</th><td>50</td><td>0</td><td>0</td></tr><tr><th>pagoda</th><td>50</td><td>0</td><td>0</td></tr><tr><th>rainforest</th><td>50</td><td>0</td><td>0</td></tr></table>	True Labels \ Predicted Labels	jail_cell	pagoda	rainforest	jail_cell	50	0	0	pagoda	50	0	0	rainforest	50	0	0	<p>Accuracy = 0.33 Precision = 0.52 Recall = 0.33 F1 Score = 0.23</p> <p>Confusion Matrix for PCA=32, GMM Mixtures=2</p>  <table><tr><th>True Labels \ Predicted Labels</th><th>jail_cell</th><th>pagoda</th><th>rainforest</th></tr><tr><th>jail_cell</th><td>6</td><td>44</td><td>0</td></tr><tr><th>pagoda</th><td>6</td><td>44</td><td>0</td></tr><tr><th>rainforest</th><td>38</td><td>12</td><td>0</td></tr></table>	True Labels \ Predicted Labels	jail_cell	pagoda	rainforest	jail_cell	6	44	0	pagoda	6	44	0	rainforest	38	12	0
True Labels \ Predicted Labels	jail_cell	pagoda	rainforest																														
jail_cell	50	0	0																														
pagoda	50	0	0																														
rainforest	50	0	0																														
True Labels \ Predicted Labels	jail_cell	pagoda	rainforest																														
jail_cell	6	44	0																														
pagoda	6	44	0																														
rainforest	38	12	0																														
GMM = 4	GMM = 8																																
<p>Accuracy = 0.32 Precision = 0.41 Recall = 0.32 F1 Score = 0.35</p> <p>Confusion Matrix for PCA=32, GMM Mixtures=4</p>  <table><tr><th>True Labels \ Predicted Labels</th><th>jail_cell</th><th>pagoda</th><th>rainforest</th></tr><tr><th>jail_cell</th><td>14</td><td>34</td><td>1</td></tr><tr><th>pagoda</th><td>39</td><td>8</td><td>3</td></tr><tr><th>rainforest</th><td>20</td><td>3</td><td>27</td></tr></table>	True Labels \ Predicted Labels	jail_cell	pagoda	rainforest	jail_cell	14	34	1	pagoda	39	8	3	rainforest	20	3	27	<p>Accuracy = 0.20 Precision = 0.44 Recall = 0.20 F1 Score = 0.57</p> <p>Confusion Matrix for PCA=32, GMM Mixtures=8</p>  <table><tr><th>True Labels \ Predicted Labels</th><th>jail_cell</th><th>pagoda</th><th>rainforest</th></tr><tr><th>jail_cell</th><td>0</td><td>5</td><td>1</td></tr><tr><th>pagoda</th><td>3</td><td>25</td><td>1</td></tr><tr><th>rainforest</th><td>0</td><td>10</td><td>5</td></tr></table>	True Labels \ Predicted Labels	jail_cell	pagoda	rainforest	jail_cell	0	5	1	pagoda	3	25	1	rainforest	0	10	5
True Labels \ Predicted Labels	jail_cell	pagoda	rainforest																														
jail_cell	14	34	1																														
pagoda	39	8	3																														
rainforest	20	3	27																														
True Labels \ Predicted Labels	jail_cell	pagoda	rainforest																														
jail_cell	0	5	1																														
pagoda	3	25	1																														
rainforest	0	10	5																														

B) PCA components = 19

GMM = 1	GMM = 2																																												
<p>Accuracy = 0.33 Precision = 0.77 Recall = 0.33 F1 Score = 0.16</p> <p>Confusion Matrix for PCA=19, GMM Mixtures=1</p> <table><tr><td rowspan="4">True Labels</td><td>jail_cell</td><td>50</td><td>0</td><td>0</td></tr><tr><td>pagoda</td><td>50</td><td>0</td><td>0</td></tr><tr><td>rainforest</td><td>50</td><td>0</td><td>0</td></tr><tr><td></td><td>jail_cell</td><td>pagoda</td><td>rainforest</td></tr><tr><td></td><td></td><td colspan="3">Predicted Labels</td></tr></table>	True Labels	jail_cell	50	0	0	pagoda	50	0	0	rainforest	50	0	0		jail_cell	pagoda	rainforest			Predicted Labels			<p>Accuracy = 0.34 Precision = 0.51 Recall = 0.34 F1 Score = 0.23</p> <p>Confusion Matrix for PCA=19, GMM Mixtures=2</p> <table><tr><td rowspan="4">True Labels</td><td>jail_cell</td><td>5</td><td>45</td><td>0</td></tr><tr><td>pagoda</td><td>4</td><td>46</td><td>0</td></tr><tr><td>rainforest</td><td>39</td><td>11</td><td>0</td></tr><tr><td></td><td>jail_cell</td><td>pagoda</td><td>rainforest</td></tr><tr><td></td><td></td><td colspan="3">Predicted Labels</td></tr></table>	True Labels	jail_cell	5	45	0	pagoda	4	46	0	rainforest	39	11	0		jail_cell	pagoda	rainforest			Predicted Labels		
True Labels		jail_cell	50	0	0																																								
		pagoda	50	0	0																																								
		rainforest	50	0	0																																								
		jail_cell	pagoda	rainforest																																									
		Predicted Labels																																											
True Labels	jail_cell	5	45	0																																									
	pagoda	4	46	0																																									
	rainforest	39	11	0																																									
		jail_cell	pagoda	rainforest																																									
		Predicted Labels																																											
GMM = 4	GMM = 8																																												
<p>Accuracy = 0.66 Precision = 0.68 Recall = 0.66 F1 Score = 0.66</p> <p>Confusion Matrix for PCA=19, GMM Mixtures=4</p> <table><tr><td rowspan="4">True Labels</td><td>jail_cell</td><td>39</td><td>4</td><td>5</td></tr><tr><td>pagoda</td><td>15</td><td>28</td><td>7</td></tr><tr><td>rainforest</td><td>10</td><td>7</td><td>32</td></tr><tr><td></td><td>jail_cell</td><td>pagoda</td><td>rainforest</td></tr><tr><td></td><td></td><td colspan="3">Predicted Labels</td></tr></table>	True Labels	jail_cell	39	4	5	pagoda	15	28	7	rainforest	10	7	32		jail_cell	pagoda	rainforest			Predicted Labels			<p>Accuracy = 0.3 Precision = 0.48 Recall = 0.3 F1 Score = 0.53</p> <p>Confusion Matrix for PCA=19, GMM Mixtures=8</p> <table><tr><td rowspan="4">True Labels</td><td>jail_cell</td><td>0</td><td>43</td><td>2</td></tr><tr><td>pagoda</td><td>0</td><td>45</td><td>0</td></tr><tr><td>rainforest</td><td>0</td><td>12</td><td>0</td></tr><tr><td></td><td>jail_cell</td><td>pagoda</td><td>rainforest</td></tr><tr><td></td><td></td><td colspan="3">Predicted Labels</td></tr></table>	True Labels	jail_cell	0	43	2	pagoda	0	45	0	rainforest	0	12	0		jail_cell	pagoda	rainforest			Predicted Labels		
True Labels		jail_cell	39	4	5																																								
		pagoda	15	28	7																																								
		rainforest	10	7	32																																								
		jail_cell	pagoda	rainforest																																									
		Predicted Labels																																											
True Labels	jail_cell	0	43	2																																									
	pagoda	0	45	0																																									
	rainforest	0	12	0																																									
		jail_cell	pagoda	rainforest																																									
		Predicted Labels																																											

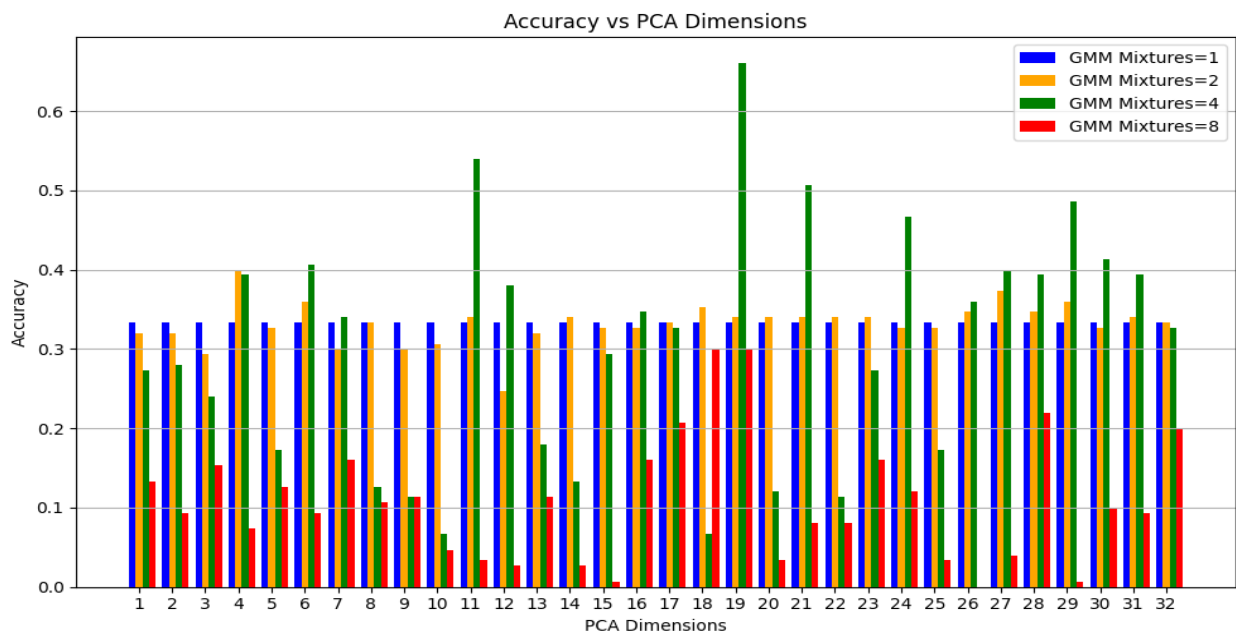
C) PCA component = 10

GMM = 1	GMM = 2																																														
<p>Accuracy = 0.32 Precision = 0.49 Recall = 0.32 F1 Score = 0.20</p> <p>Confusion Matrix for PCA=10, GMM Mixtures=1</p> <table><tr><td rowspan="3">True Labels</td><td>jail_cell</td><td>50</td><td>0</td><td>0</td></tr><tr><td>pagoda</td><td>50</td><td>0</td><td>0</td></tr><tr><td>rainforest</td><td>50</td><td>0</td><td>0</td></tr><tr><td></td><td></td><td>jail_cell</td><td>pagoda</td><td>rainforest</td></tr><tr><td></td><td></td><td colspan="3">Predicted Labels</td></tr></table>	True Labels	jail_cell	50	0	0	pagoda	50	0	0	rainforest	50	0	0			jail_cell	pagoda	rainforest			Predicted Labels			<p>Accuracy = 0.30 Precision = 0.51 Recall = 0.30 F1 Score = 0.22</p> <p>Confusion Matrix for PCA=10, GMM Mixtures=2</p> <table><tr><td rowspan="3">True Labels</td><td>jail_cell</td><td>39</td><td>11</td><td>0</td></tr><tr><td>pagoda</td><td>43</td><td>7</td><td>0</td></tr><tr><td>rainforest</td><td>8</td><td>42</td><td>0</td></tr><tr><td></td><td></td><td>jail_cell</td><td>pagoda</td><td>rainforest</td></tr><tr><td></td><td></td><td colspan="3">Predicted Labels</td></tr></table>	True Labels	jail_cell	39	11	0	pagoda	43	7	0	rainforest	8	42	0			jail_cell	pagoda	rainforest			Predicted Labels		
True Labels		jail_cell	50	0	0																																										
		pagoda	50	0	0																																										
	rainforest	50	0	0																																											
		jail_cell	pagoda	rainforest																																											
		Predicted Labels																																													
True Labels	jail_cell	39	11	0																																											
	pagoda	43	7	0																																											
	rainforest	8	42	0																																											
		jail_cell	pagoda	rainforest																																											
		Predicted Labels																																													
GMM = 4	GMM = 8																																														
<p>Accuracy = 0.06 Precision = 0.065 Recall = 0.066 F1 Score = 0.063</p> <p>Confusion Matrix for PCA=10, GMM Mixtures=4</p> <table><tr><td rowspan="3">True Labels</td><td>jail_cell</td><td>2</td><td>10</td><td>36</td></tr><tr><td>pagoda</td><td>11</td><td>2</td><td>37</td></tr><tr><td>rainforest</td><td>16</td><td>26</td><td>6</td></tr><tr><td></td><td></td><td>jail_cell</td><td>pagoda</td><td>rainforest</td></tr><tr><td></td><td></td><td colspan="3">Predicted Labels</td></tr></table>	True Labels	jail_cell	2	10	36	pagoda	11	2	37	rainforest	16	26	6			jail_cell	pagoda	rainforest			Predicted Labels			<p>Accuracy = 0.046 Precision = 0.39 Recall = 0.0466 F1 Score = 0.38</p> <p>Confusion Matrix for PCA=10, GMM Mixtures=8</p> <table><tr><td rowspan="3">True Labels</td><td>jail_cell</td><td>0</td><td>0</td><td>13</td></tr><tr><td>pagoda</td><td>0</td><td>0</td><td>19</td></tr><tr><td>rainforest</td><td>0</td><td>15</td><td>7</td></tr><tr><td></td><td></td><td>jail_cell</td><td>pagoda</td><td>rainforest</td></tr><tr><td></td><td></td><td colspan="3">Predicted Labels</td></tr></table>	True Labels	jail_cell	0	0	13	pagoda	0	0	19	rainforest	0	15	7			jail_cell	pagoda	rainforest			Predicted Labels		
True Labels		jail_cell	2	10	36																																										
		pagoda	11	2	37																																										
	rainforest	16	26	6																																											
		jail_cell	pagoda	rainforest																																											
		Predicted Labels																																													
True Labels	jail_cell	0	0	13																																											
	pagoda	0	0	19																																											
	rainforest	0	15	7																																											
		jail_cell	pagoda	rainforest																																											
		Predicted Labels																																													

D) PCA component = 1

GMM = 1	GMM = 2																																												
<p>Accuracy = 0.33 Precision = 0.77 Recall = 0.33 F1 Score = 0.16</p> <p>Confusion Matrix for PCA=1, GMM Mixtures=1</p> <table><tr><td rowspan="4">True Labels</td><td>jail_cell</td><td>50</td><td>0</td><td>0</td></tr><tr><td>pagoda</td><td>50</td><td>0</td><td>0</td></tr><tr><td>rainforest</td><td>50</td><td>0</td><td>0</td></tr><tr><td></td><td>jail_cell</td><td>pagoda</td><td>rainforest</td></tr><tr><td></td><td></td><td colspan="3">Predicted Labels</td></tr></table>	True Labels	jail_cell	50	0	0	pagoda	50	0	0	rainforest	50	0	0		jail_cell	pagoda	rainforest			Predicted Labels			<p>Accuracy = 0.32 Precision = 0.49 Recall = 0.32 F1 Score = 0.20</p> <p>Confusion Matrix for PCA=1, GMM Mixtures=2</p> <table><tr><td rowspan="4">True Labels</td><td>jail_cell</td><td>44</td><td>6</td><td>0</td></tr><tr><td>pagoda</td><td>46</td><td>4</td><td>0</td></tr><tr><td>rainforest</td><td>29</td><td>21</td><td>0</td></tr><tr><td></td><td>jail_cell</td><td>pagoda</td><td>rainforest</td></tr><tr><td></td><td></td><td colspan="3">Predicted Labels</td></tr></table>	True Labels	jail_cell	44	6	0	pagoda	46	4	0	rainforest	29	21	0		jail_cell	pagoda	rainforest			Predicted Labels		
True Labels		jail_cell	50	0	0																																								
		pagoda	50	0	0																																								
		rainforest	50	0	0																																								
		jail_cell	pagoda	rainforest																																									
		Predicted Labels																																											
True Labels	jail_cell	44	6	0																																									
	pagoda	46	4	0																																									
	rainforest	29	21	0																																									
		jail_cell	pagoda	rainforest																																									
		Predicted Labels																																											
GMM = 4	GMM = 8																																												
<p>Accuracy = 0.27 Precision = 0.25 Recall = 0.27 F1 Score = 0.24</p> <p>Confusion Matrix for PCA=1, GMM Mixtures=4</p> <table><tr><td rowspan="4">True Labels</td><td>jail_cell</td><td>5</td><td>20</td><td>25</td></tr><tr><td>pagoda</td><td>2</td><td>29</td><td>18</td></tr><tr><td>rainforest</td><td>13</td><td>26</td><td>7</td></tr><tr><td></td><td>jail_cell</td><td>pagoda</td><td>rainforest</td></tr><tr><td></td><td></td><td colspan="3">Predicted Labels</td></tr></table>	True Labels	jail_cell	5	20	25	pagoda	2	29	18	rainforest	13	26	7		jail_cell	pagoda	rainforest			Predicted Labels			<p>Accuracy = 0.13 Precision = 0.54 Recall = 0.13 F1 Score = 0.19</p> <p>Confusion Matrix for PCA=1, GMM Mixtures=8</p> <table><tr><td rowspan="4">True Labels</td><td>jail_cell</td><td>6</td><td>9</td><td>0</td></tr><tr><td>pagoda</td><td>8</td><td>9</td><td>0</td></tr><tr><td>rainforest</td><td>9</td><td>7</td><td>5</td></tr><tr><td></td><td>jail_cell</td><td>pagoda</td><td>rainforest</td></tr><tr><td></td><td></td><td colspan="3">Predicted Labels</td></tr></table>	True Labels	jail_cell	6	9	0	pagoda	8	9	0	rainforest	9	7	5		jail_cell	pagoda	rainforest			Predicted Labels		
True Labels		jail_cell	5	20	25																																								
		pagoda	2	29	18																																								
		rainforest	13	26	7																																								
		jail_cell	pagoda	rainforest																																									
		Predicted Labels																																											
True Labels	jail_cell	6	9	0																																									
	pagoda	8	9	0																																									
	rainforest	9	7	5																																									
		jail_cell	pagoda	rainforest																																									
		Predicted Labels																																											

E) PCA components vs. Accuracy plots:



Maximum Accuracy	0.66
------------------	------

GMM Mixtures	4
--------------	---

PCA components	19
----------------	----

2. Bayes classifier using the density estimated from K-nearest neighbor (KNN) method for Dataset-1 (a)

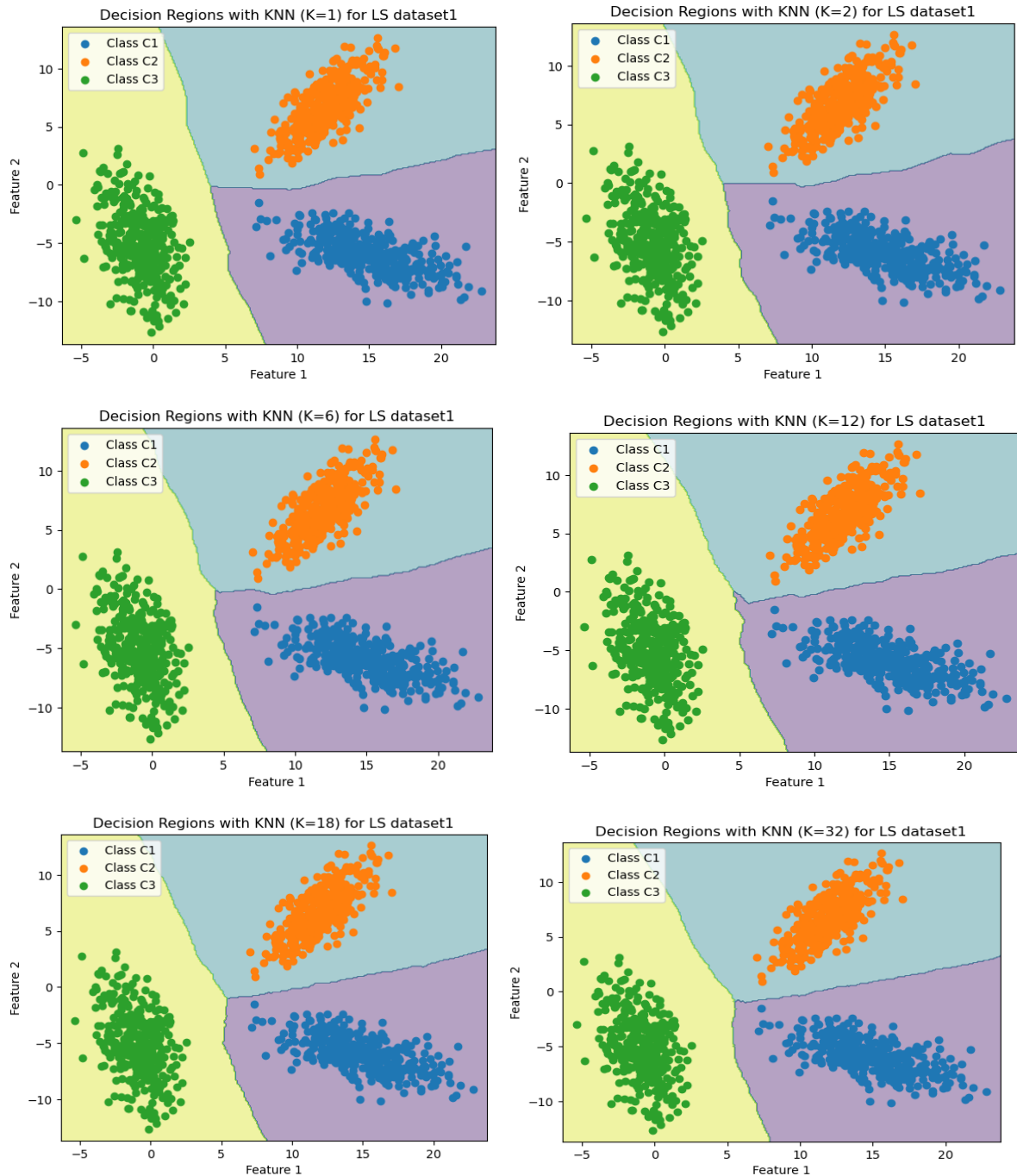
(a) LS data

Dataset description:

- Linearly separable dataset
- Total sample size: 1500
- Total classes: 3
- Samples in each class: 500
- Dimensionality: each point is 2D

Note: This data description will be valid and hence not mentioned again whenever LS dataset term is used in the report.

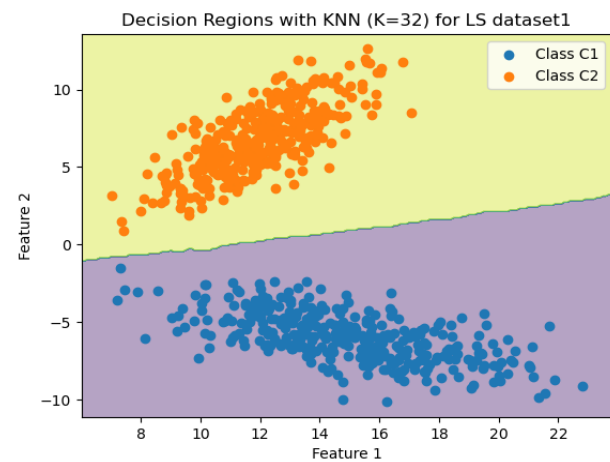
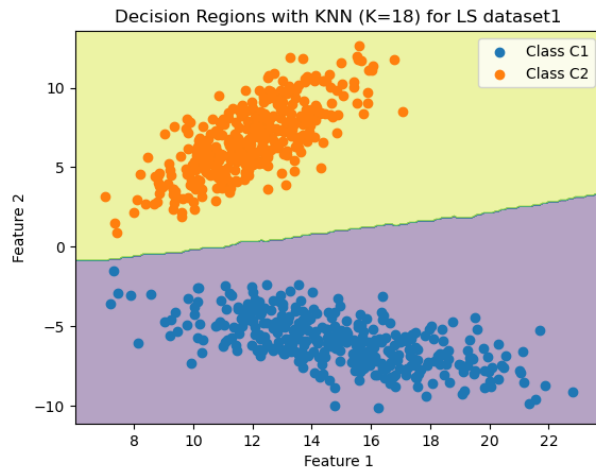
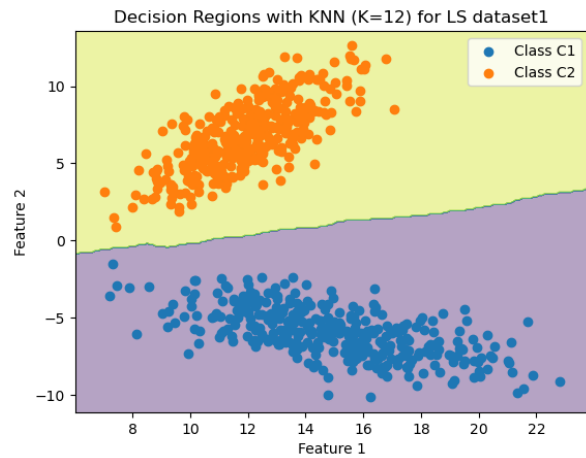
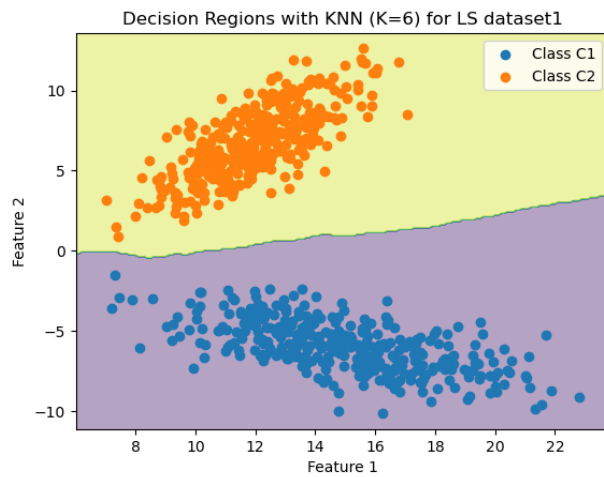
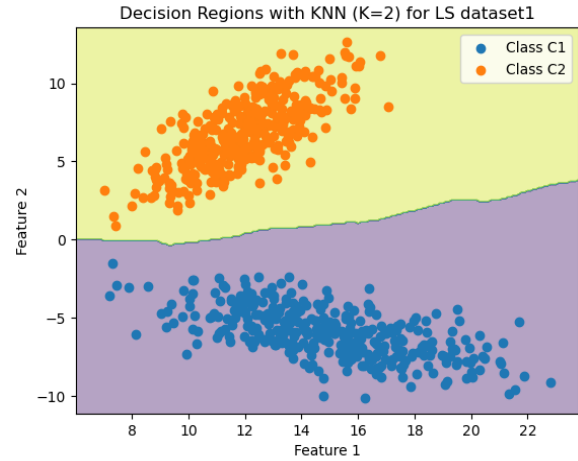
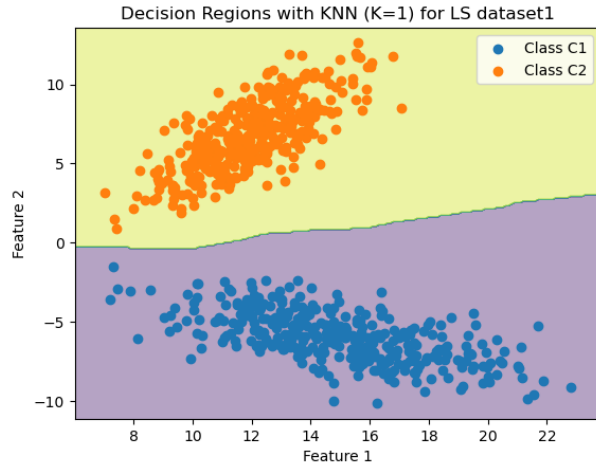
- **Bayes classifier using the density estimated from K-nearest neighbor (KNN) method for Dataset-1 (a)/ linearly separable dataset for class 1 Vs class 2 Vs class 3 as the values of k varies (1, 2, 6, 12, 18, 32).**



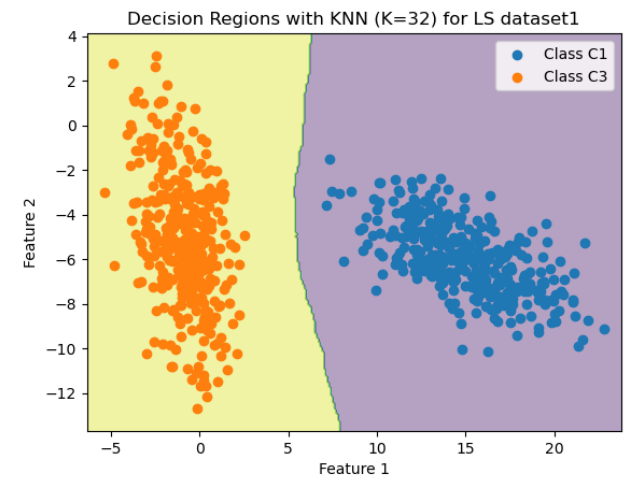
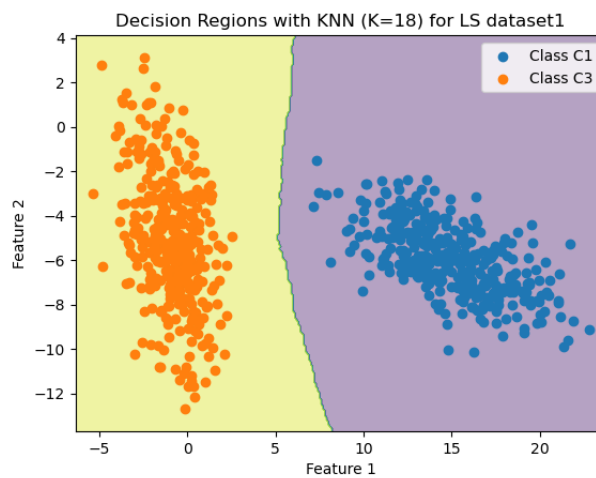
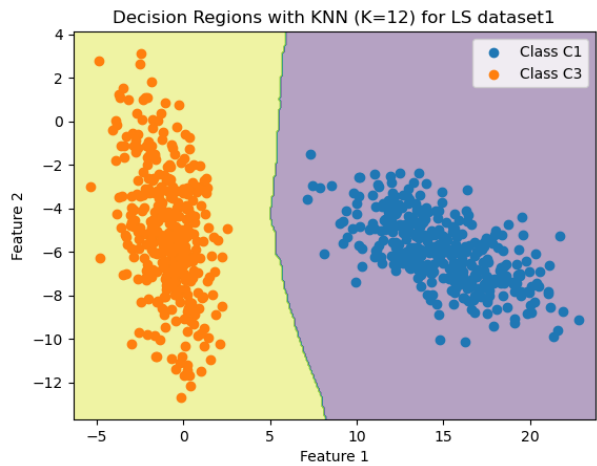
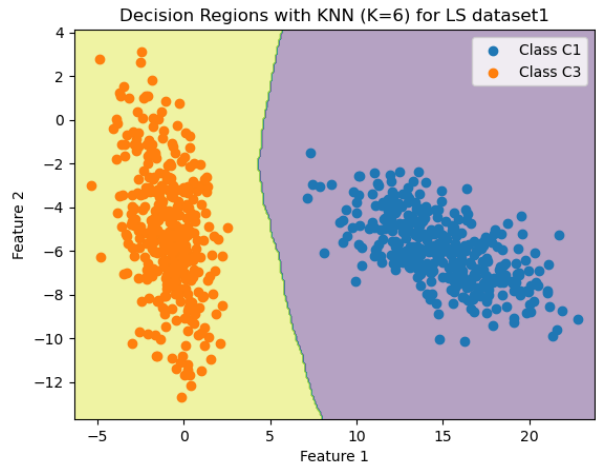
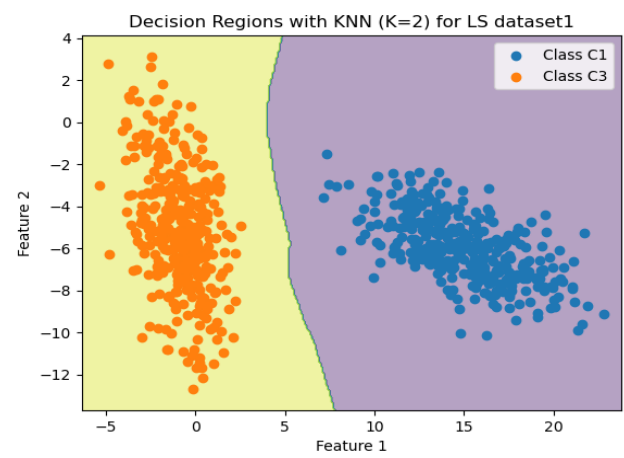
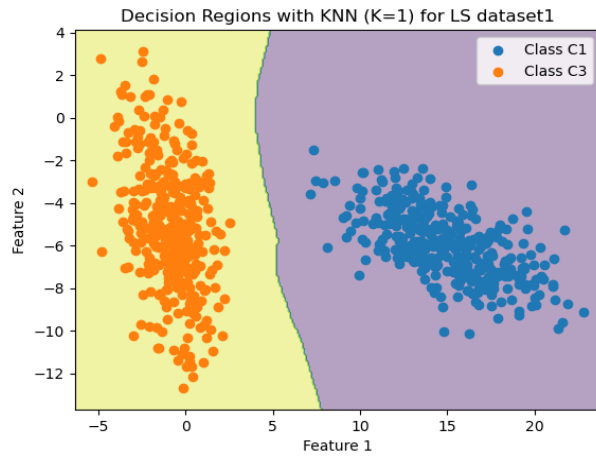
Inferences:

- At $k=1$ we have the decision boundary as the perpendicular bisector of the nearest points. Overall the shape is non linear but it is piecewise linear.
- As the k value increases the decision boundary between the classes becomes more smooth.

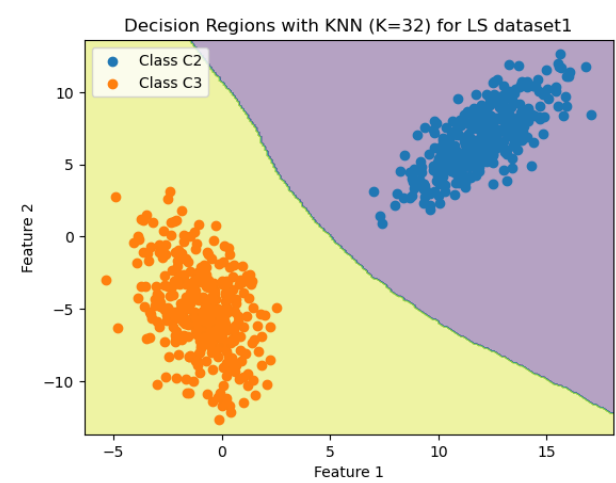
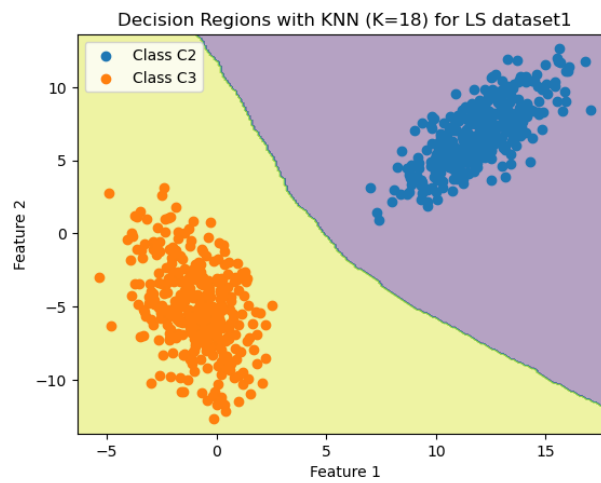
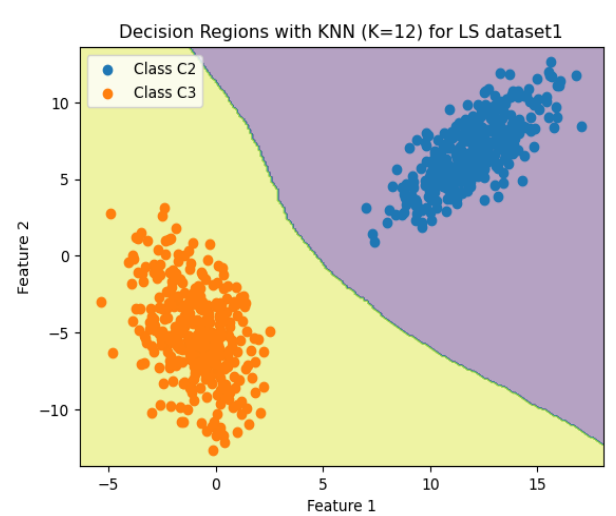
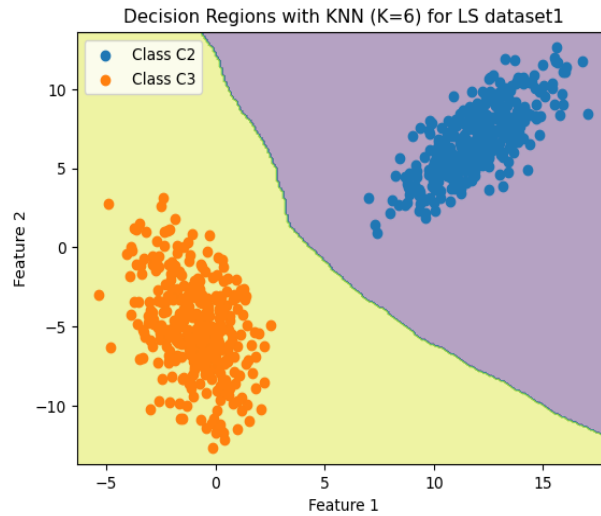
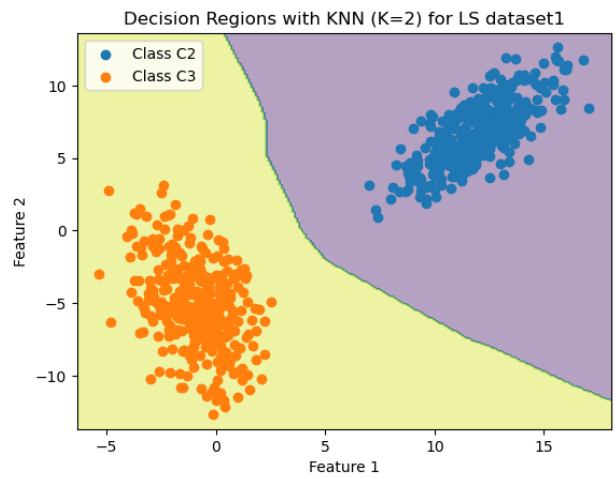
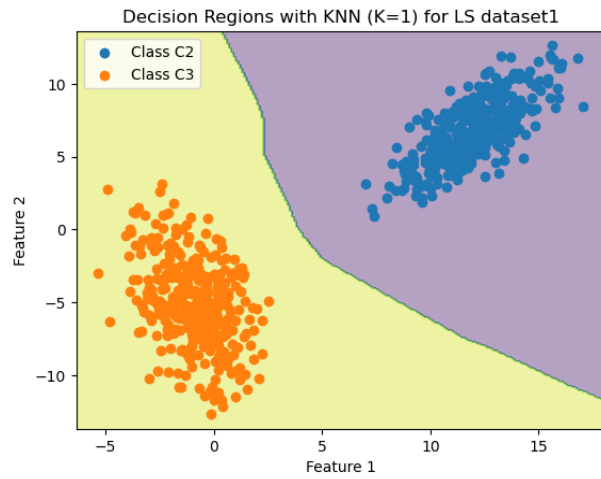
- Bayes classifier using the density estimated from K-nearest neighbor (KNN) method for Dataset-1 (a)/ linearly separable dataset for class 1 Vs class 2 as the values of k varies (1, 2, 6, 12, 18, 32).



- Bayes classifier using the density estimated from K-nearest neighbor (KNN) method for Dataset-1 (a)/ linearly separable dataset for class 1 Vs class 3 as the values of k varies (1, 2, 6, 12, 18, 32).



- Bayes classifier using the density estimated from K-nearest neighbor (KNN) method for Dataset-1 (a)/ linearly separable dataset for class 2 Vs class 3 as the values of k varies (1, 2, 6, 12, 18, 32).



b) Bayes classifier using the density estimated from K-nearest neighbor (KNN) method for NLS Dataset-1 (b).

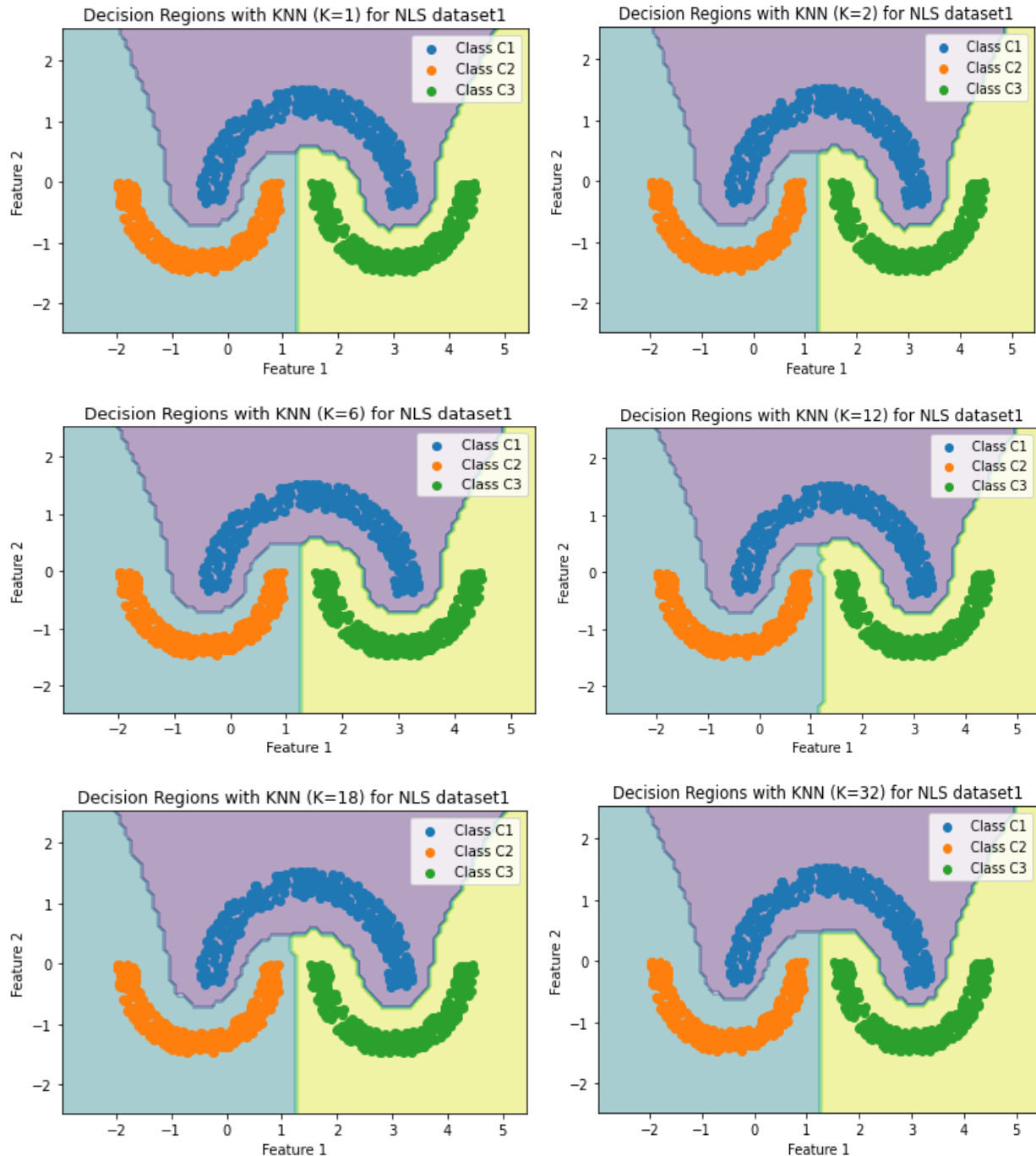
(b) NLS data

Dataset description:

- Non Linearly separable dataset
- Total sample size: 1500
- Total classes: 3
- Samples in each class: 500
- Dimensionality: each point is 2D

Note: This data description will be valid and hence not mentioned again whenever NLS dataset term is used in the report.

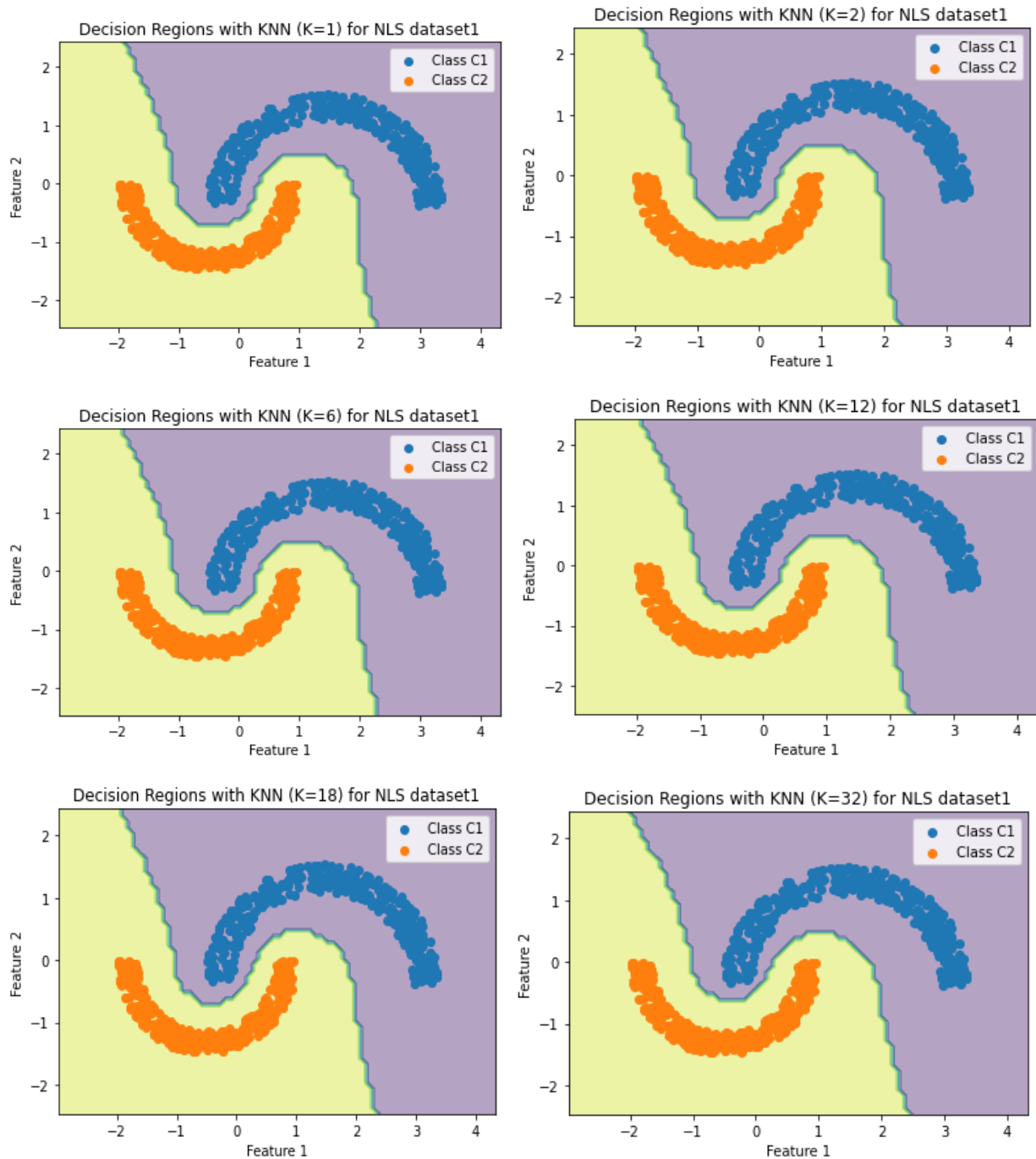
- Bayes classifier using the density estimated from K-nearest neighbor (KNN) method for Dataset-1 (b)/ non linearly separable dataset for class 1 Vs class 2 Vs class 3 as the values of k varies (1, 2, 6, 12, 18, 32).



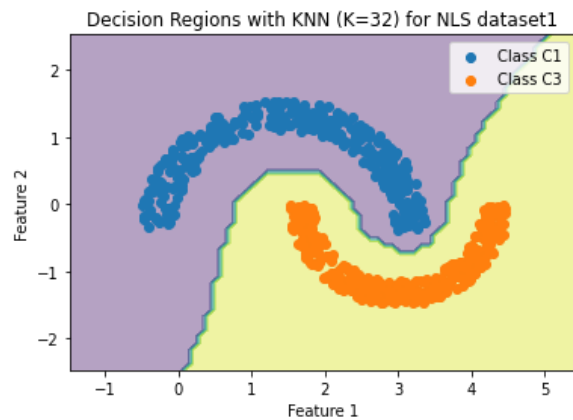
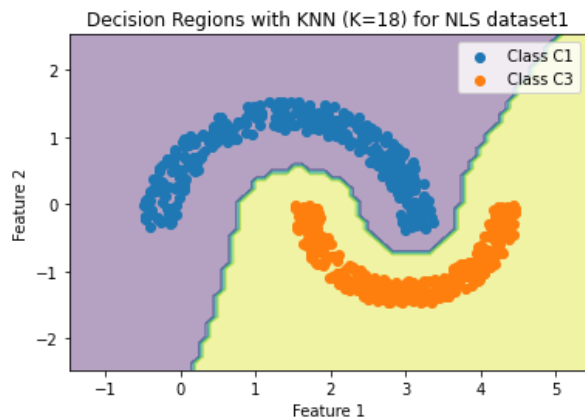
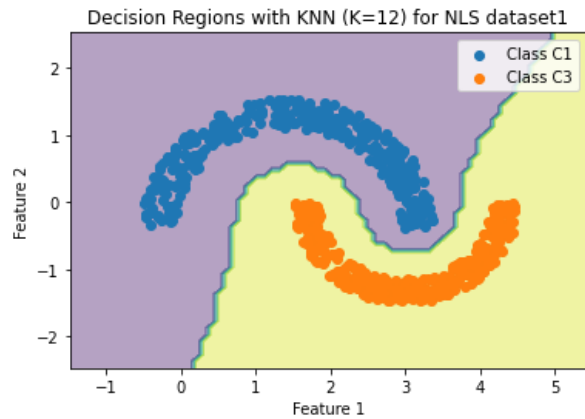
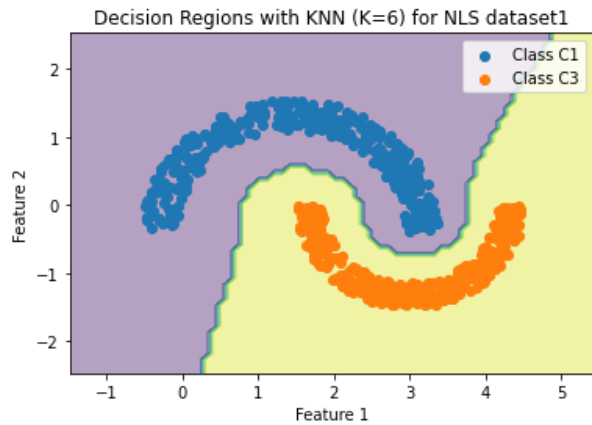
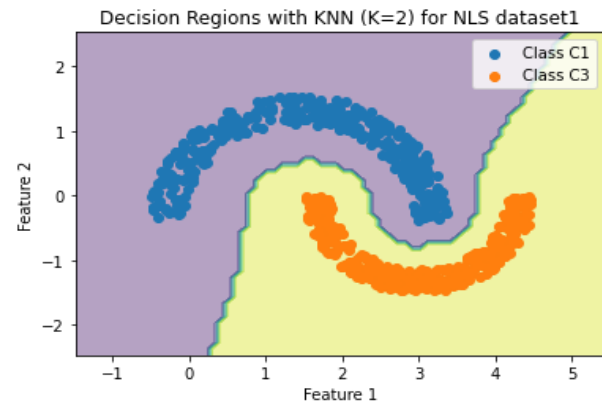
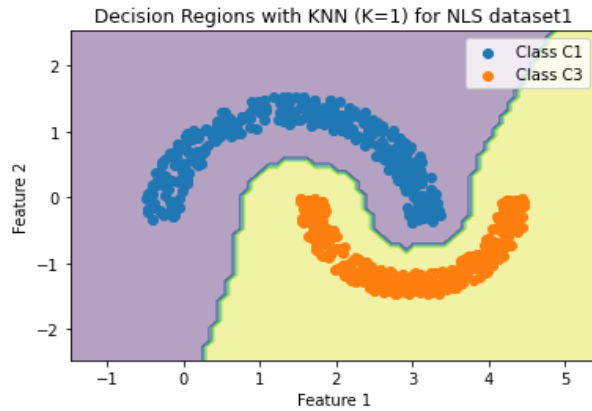
Inferences:

- At $k=1$ we have the decision boundary as the perpendicular bisector of the nearest points. Overall the shape is non linear but it is piecewise linear.
- As the k value increases the decision boundary between classes becomes more smooth.

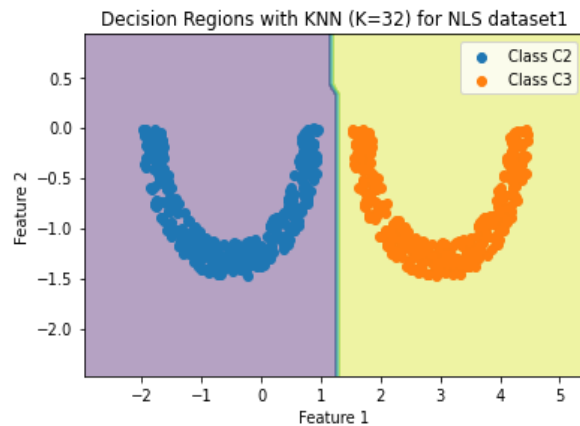
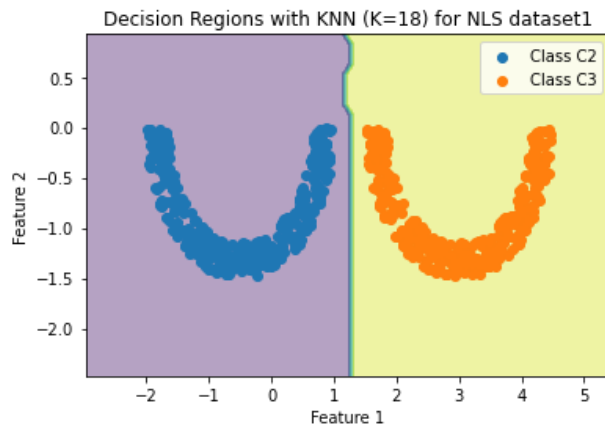
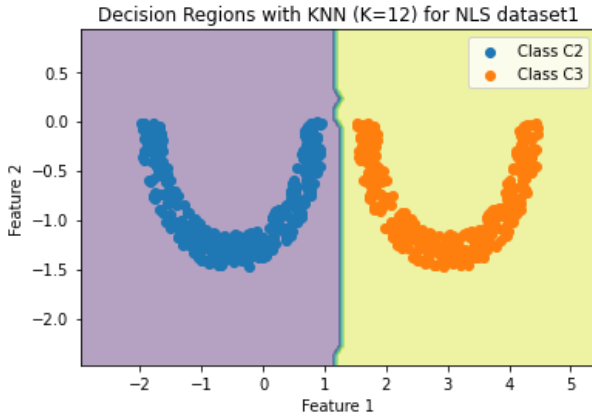
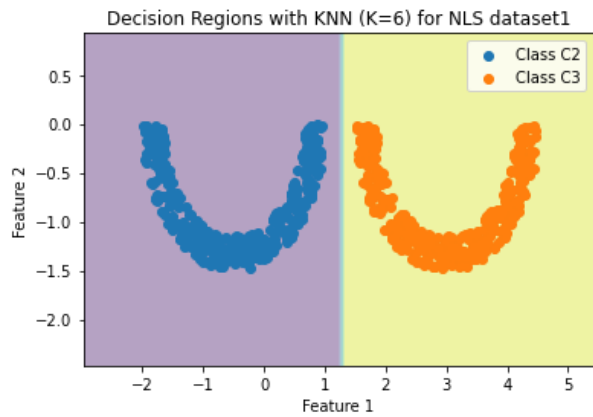
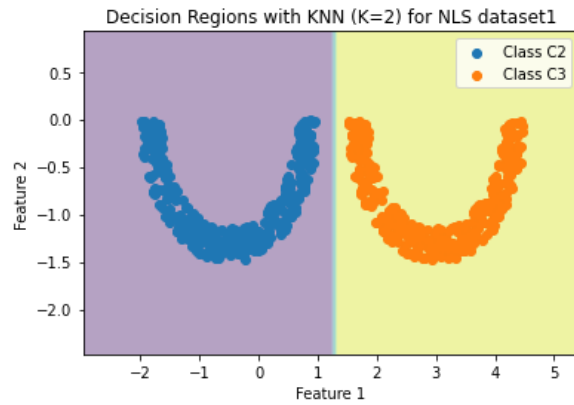
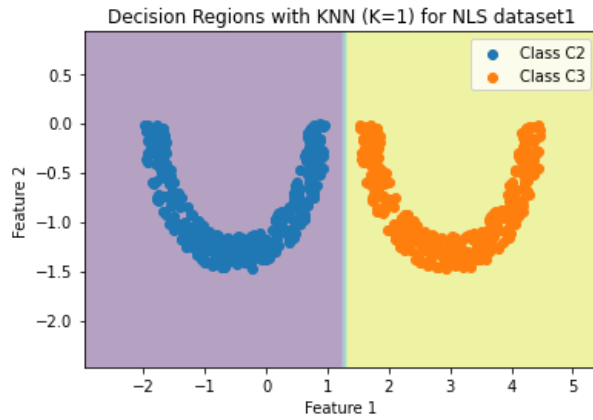
- Bayes classifier using the density estimated from K-nearest neighbor (KNN) method for Dataset-1 (b)/non linearly separable dataset for class 1 Vs class 2 as the values of k varies (1, 2, 6, 12, 18, 32)



- Bayes classifier using the density estimated from K-nearest neighbor (KNN) method for Dataset-1 (b)/non linearly separable dataset for class 1 Vs class 3 as the values of k varies (1, 2, 6, 12, 18, 32)



- Bayes classifier using the density estimated from K-nearest neighbor (KNN) method for Dataset-1 (b)/non linearly separable dataset for class 2 Vs class 3 as the values of k varies (1, 2, 6, 12, 18, 32)



(c) KNN for 32D BoVW data:

For k-nearest neighbors = 1 the specifics are:

Classification Accuracy: 0.72

Classification Report:

Class 1:

Precision: 0.7955

Recall: 0.7000

F1-Score: 0.7447

Support: 50.0

Class 2:

Precision: 0.6667

Recall: 0.6800

F1-Score: 0.6733

Support: 50.0

Class 3:

Precision: 0.7091

Recall: 0.7800

F1-Score: 0.7429

Support: 50.0

Mean Precision: 0.7237373737373737

Mean Recall: 0.7200000000000001

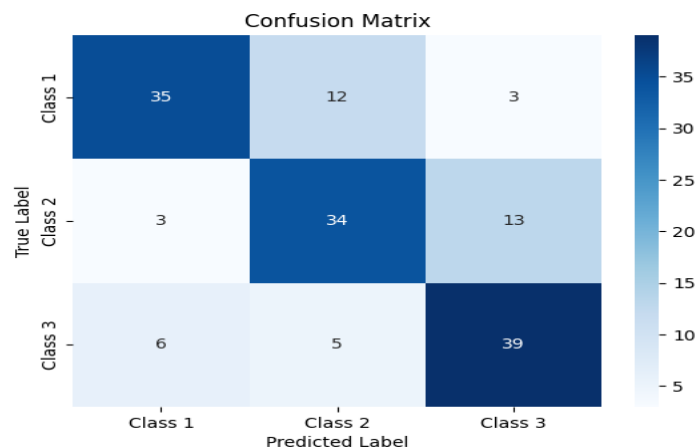
Mean F1-Score: 0.720268440217882

Confusion Matrix:

[[35 12 3]

[3 34 13]

[6 5 39]]



For k-nearest neighbors = 3 the specifics are:

Classification Accuracy: 0.7066666666666667

Classification Report:

Class 1:
Precision: 0.6939
Recall: 0.6800
F1-Score: 0.6869
Support: 50.0

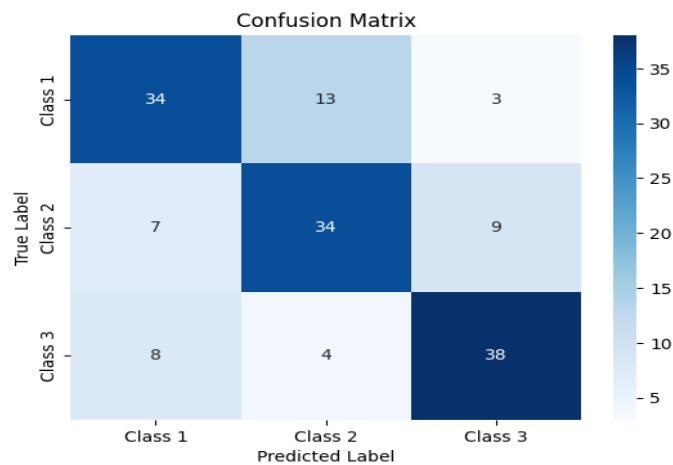
Class 2:
Precision: 0.6667
Recall: 0.6800
F1-Score: 0.6733
Support: 50.0

Class 3:
Precision: 0.7600
Recall: 0.7600
F1-Score: 0.7600
Support: 50.0

Mean Precision: 0.7068480725623583
Mean Recall: 0.7066666666666667
Mean F1-Score: 0.7067120045337868

Confusion Matrix:

```
[[34 13  3]
 [ 7 34  9]
 [ 8  4 38]]
```



For k-nearest neighbors = 5 the specifics are:

Classification Accuracy: 0.7266666666666667

Classification Report:

Class 1:
Precision: 0.7907
Recall: 0.6800
F1-Score: 0.7312
Support: 50.0

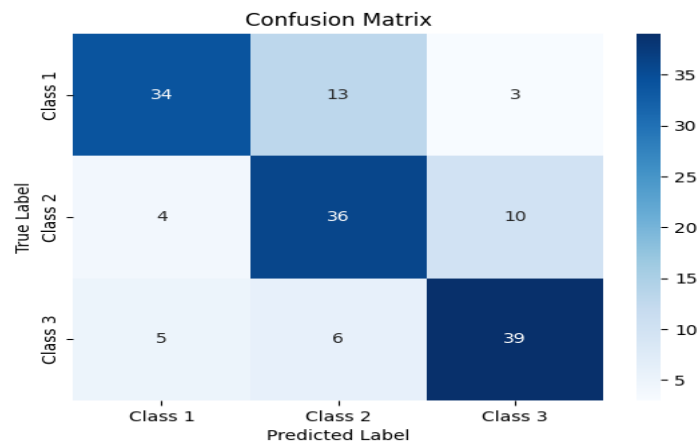
Class 2:
Precision: 0.6545
Recall: 0.7200
F1-Score: 0.6857
Support: 50.0

Class 3:
Precision: 0.7500
Recall: 0.7800
F1-Score: 0.7647
Support: 50.0

Mean Precision: 0.7317477096546864
Mean Recall: 0.7266666666666666
Mean F1-Score: 0.7272009879220506

Confusion Matrix:

```
[[34 13  3]
 [ 4 36 10]
 [ 5  6 39]]
```



For k-nearest neighbors = 7 the specifics are:

Classification Accuracy: 0.7466666666666667

Classification Report:

Class 1:
Precision: 0.8537
Recall: 0.7000
F1-Score: 0.7692
Support: 50.0

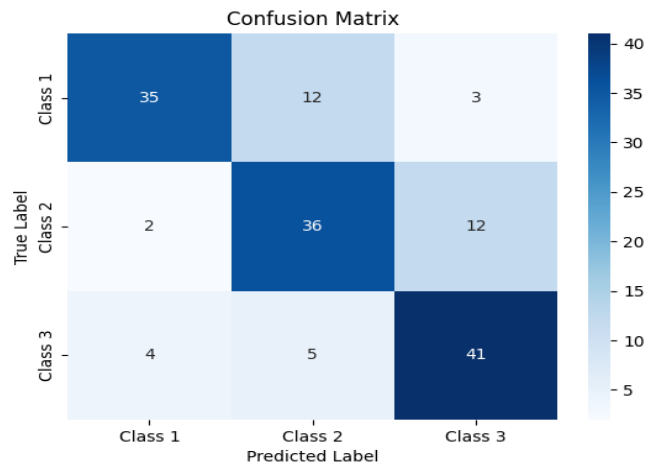
Class 2:
Precision: 0.6792
Recall: 0.7200
F1-Score: 0.6990
Support: 50.0

Class 3:
Precision: 0.7321
Recall: 0.8200
F1-Score: 0.7736
Support: 50.0

Mean Precision: 0.7550155589156969
Mean Recall: 0.7466666666666666
Mean F1-Score: 0.7472816003682462

Confusion Matrix:

```
[[35 12  3]
 [ 2 36 12]
 [ 4  5 41]]
```



For k-nearest neighbors = 9 the specifics are:

Classification Accuracy: 0.7266666666666667

Classification Report:

Class 1:

Precision: 0.8095

Recall: 0.6800

F1-Score: 0.7391

Support: 50.0

Class 2:

Precision: 0.6607

Recall: 0.7400

F1-Score: 0.6981

Support: 50.0

Class 3:

Precision: 0.7308

Recall: 0.7600

F1-Score: 0.7451

Support: 50.0

Mean Precision: 0.7336691086691087

Mean Recall: 0.7266666666666666

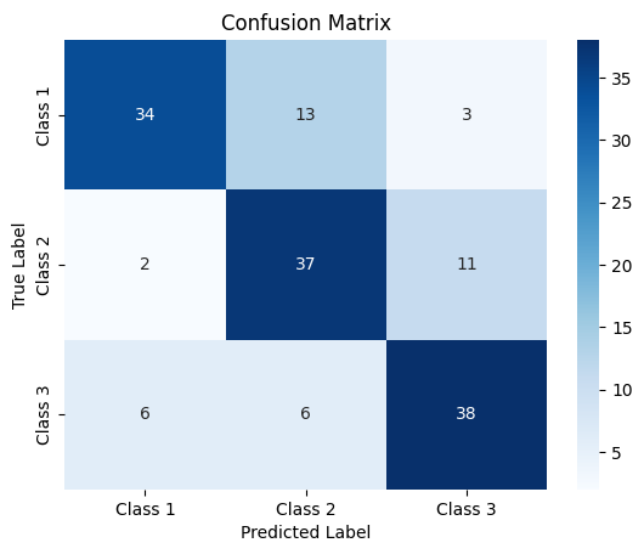
Mean F1-Score: 0.7274472271818216

Confusion Matrix:

[[34 13 3]

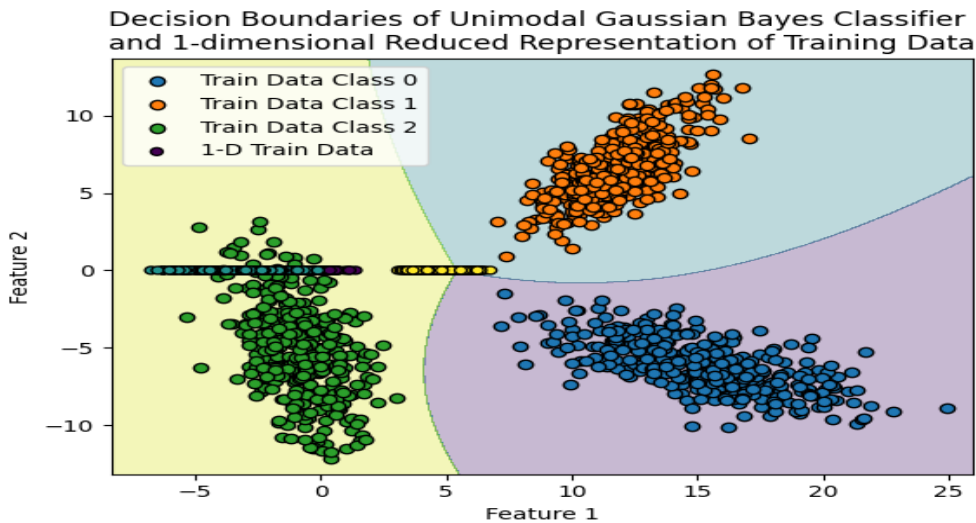
[2 37 11]

[6 6 38]]



3. Apply Fisher linear discriminant analysis (FDA) on Dataset-1 and Dataset-2. Use Bayes classifier using both unimodal Gaussian and GMM.

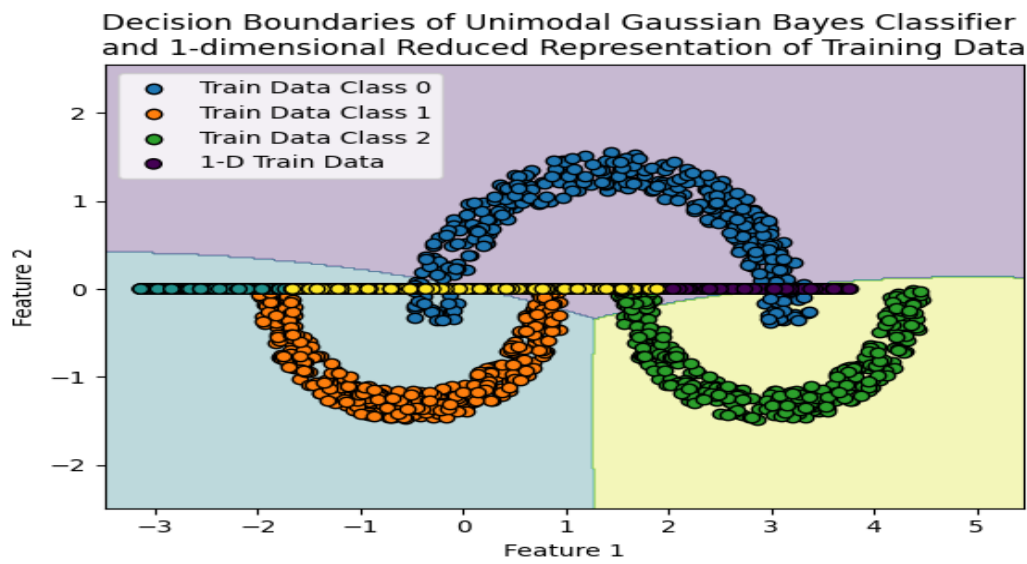
For LS dataset:



Metrics	Scores
Accuracy	0.9
Precision (Class-wise)	(0.84, 0.86, 1.00)
Mean Precision	0.9047
Recall (Class-wise)	(0.87, 0.84, 1.00)
Mean Recall	0.904
F-measure (Class-wise)	(0.85, 0.85, 1.00)
Mean F-measure	0.904

	Predicted output			
	Class 1	Class 2	Class 3	
	Class 1	131	19	0
Actual output	Class 2	24	126	0
	Class 3	0	0	150

For NLS Dataset:

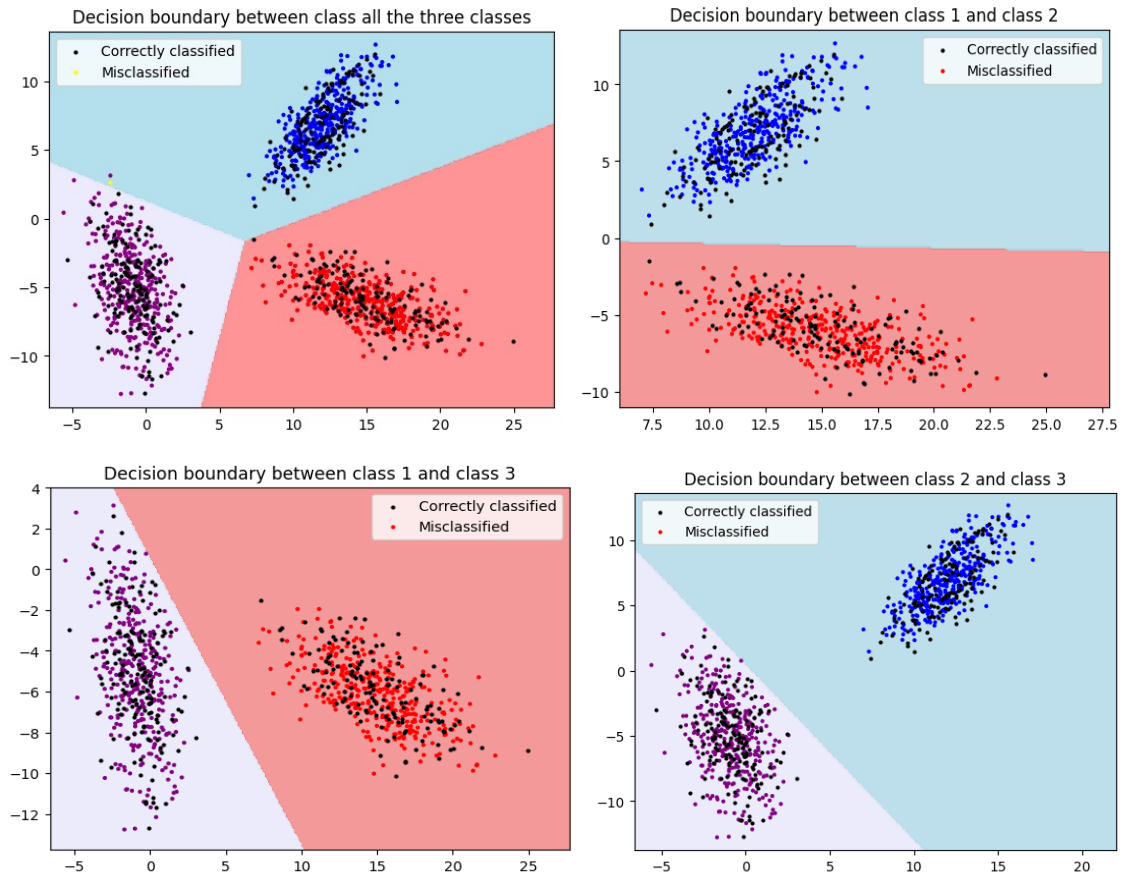


Metrics	Scores
Accuracy	0.72
Precision (Class-wise)	(0.93, 0.67, 0.58)
Mean Precision	0.73
Recall (Class-wise)	(0.80, 0.80, 0.56)
Mean Recall	0.72
F-measure (Class-wise)	(0.86, 0.73, 0.57)
Mean F-measure	0.904

	Predicted output			
		Class 1	Class 2	Class 3
Actual output	Class 1	121	0	29
	Class 2	0	120	30
	Class 3	8	58	84

4. Perceptron-based classifier:

4.a) LS Dataset:



Precision, Recall and F1-score on Test set for all LS classes independently:

Class Label	Precision	Recall	F1-score	Total instances
Class 1	1	0.99	1	148
Class 2	0.99	1	1	141
Class 3	1	1	1	161

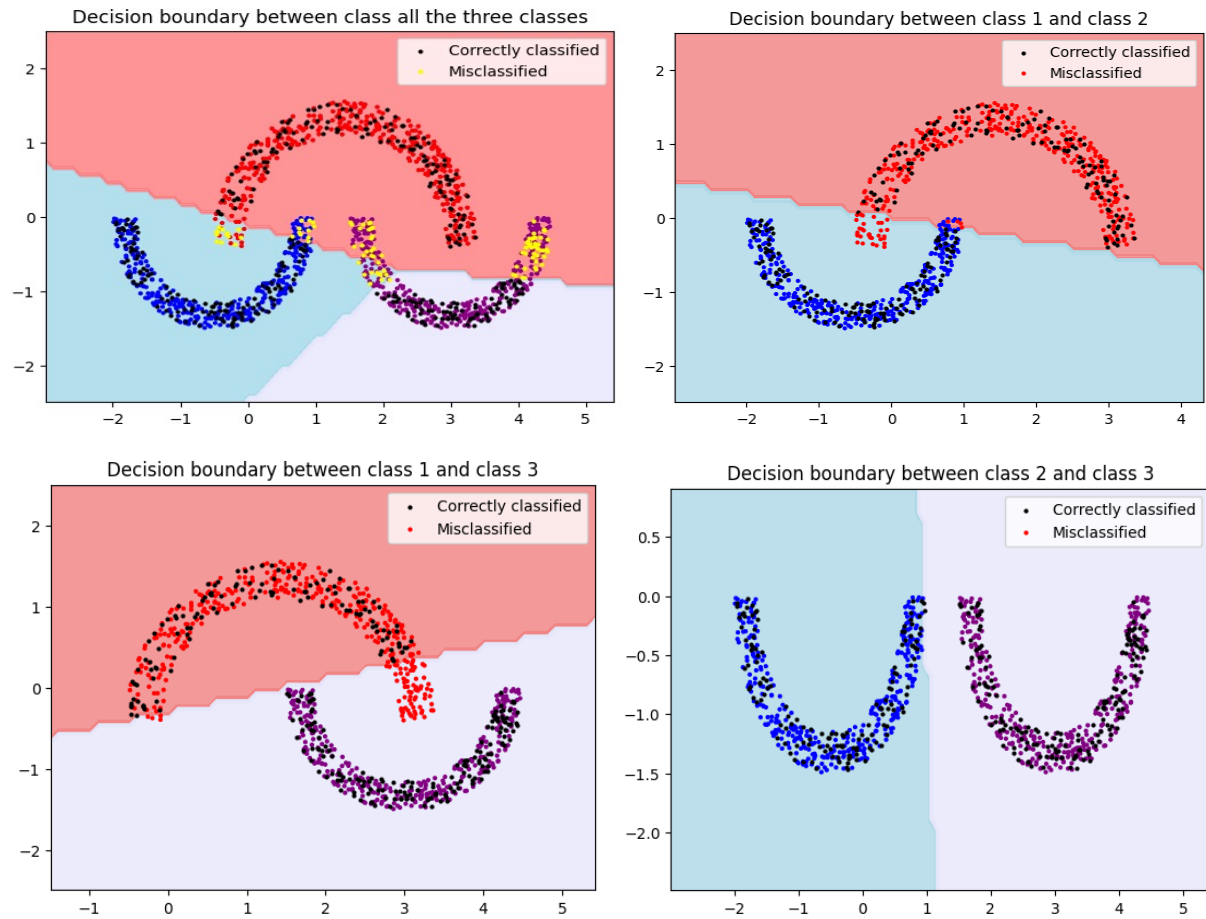
Confusion Matrix for 450 test data of LS dataset:

	Predicted output			
	Class 1	Class 2	Class 3	
	Class 1	147	1	0
Actual output	Class 2	0	141	0
	Class 3	0	0	161

Accuracy score for test data of all three classes of LS data:

Accuracy = 1 (100%).

b) NLS dataset:



Precision, Recall and F1-score on Test set for all NLS classes independently:

Class Label	Precision	Recall	F1-score	Total instances
Class 1	0.97	0.75	0.84	148
Class 2	0.91	0.98	0.94	141
Class 3	0.87	0.99	0.93	161

Confusion Matrix for 450 test data of NLS dataset:

	Predicted output			
	Class 1	Class 2	Class 3	
	Class 1	111	14	23
Actual output	Class 2	3	138	0
	Class 3	1	0	160

Accuracy score for test data of all three classes of NLS data:

Accuracy = 0.9

c) Perceptron based classifier for 32 Dimensional Bag OF Words representation:

Test Classification Report:

	precision	recall	f1-score	support
jail_cell	0.74	0.50	0.60	50
pagoda	0.47	0.90	0.62	50
rainforest	0.90	0.38	0.54	50
accuracy			0.59	150
macro avg	0.70	0.59	0.58	150
weighted avg	0.70	0.59	0.58	150

Test Confusion Matrix:

	Predicted output			
	Class 1	Class 2	Class 3	
	Class 1	25	23	2
Actual output	Class 2	5	45	0
	Class 3	4	27	19

Test Classification Report with PCA:

	precision	recall	f1-score	support
jail_cell	0.21	0.30	0.25	50
pagoda	0.32	0.50	0.39	50
rainforest	0.00	0.00	0.00	50
accuracy			0.27	150
macro avg	0.18	0.27	0.21	150
weighted avg	0.18	0.27	0.21	150

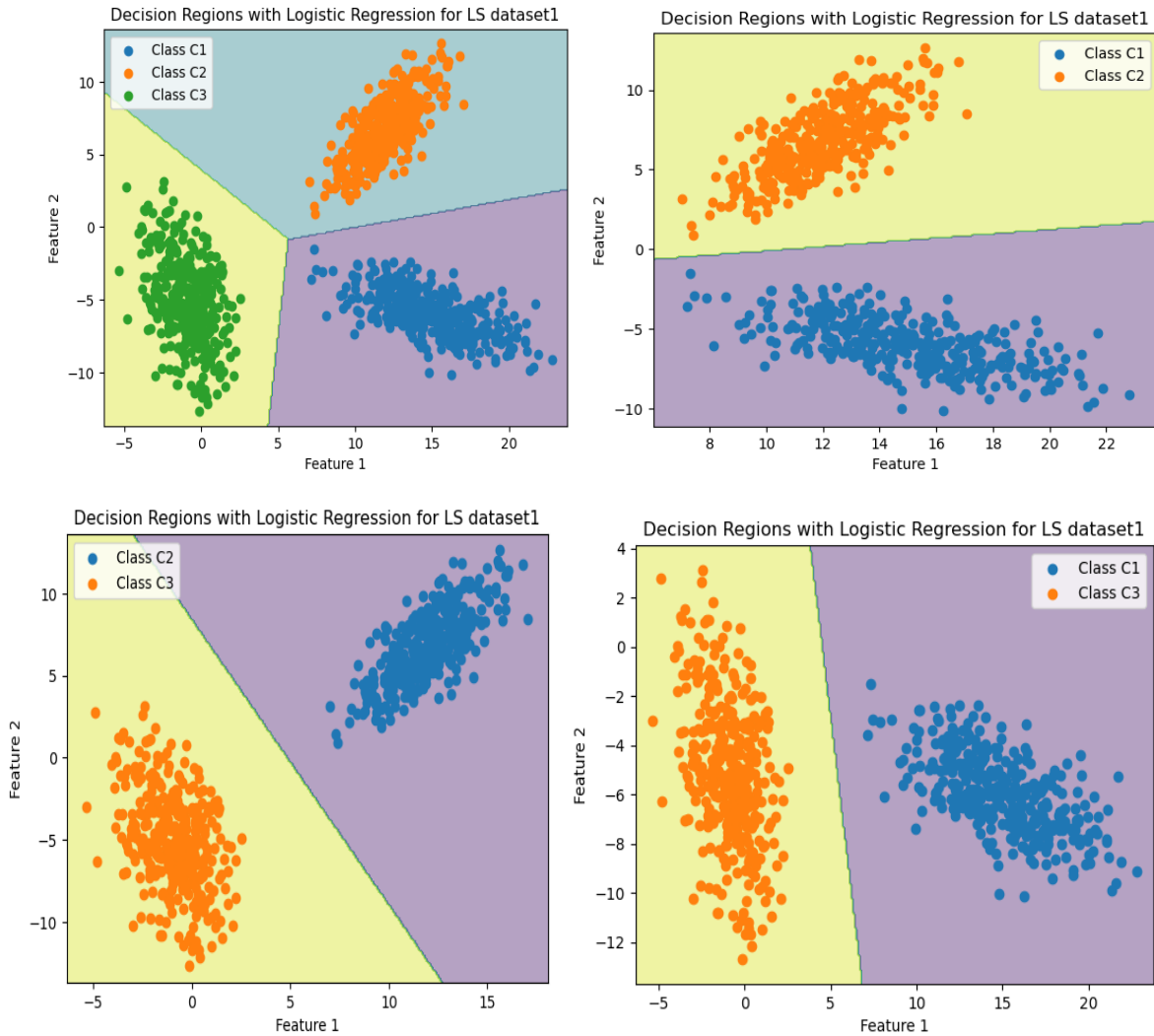
Test Confusion Matrix with PCA:

	Predicted output			
	Class 1	Class 2	Class 3	
	Class 1	15	34	1
Actual output	Class 2	25	25	0
	Class 3	31	19	0

5. Logistic regression classifier on LS Dataset-1 (a)

Dataset description:

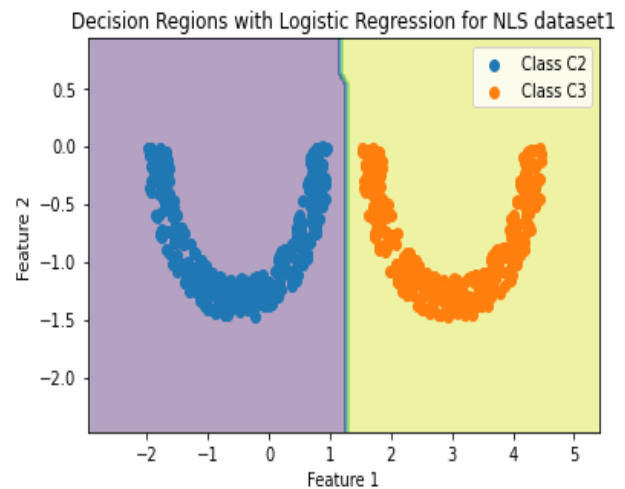
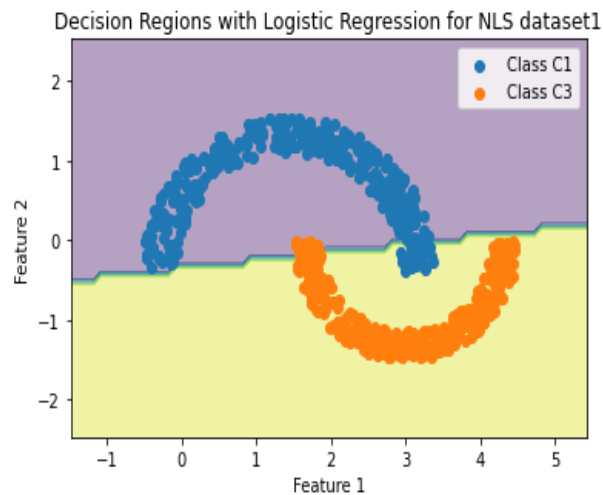
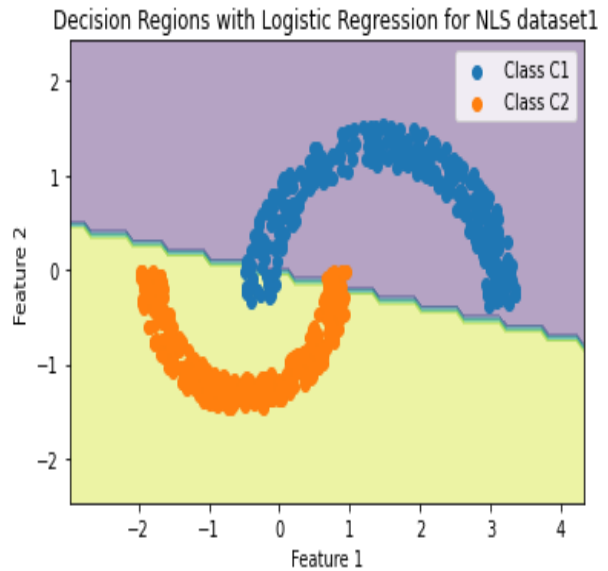
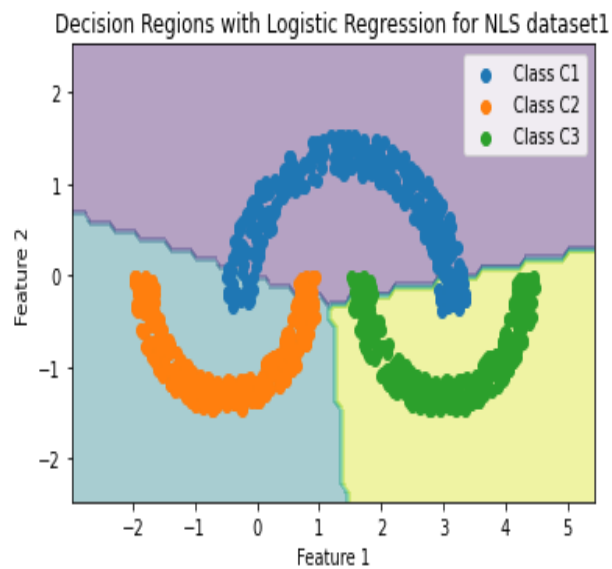
- Linearly separable dataset
- Total sample size: 1500
- Total classes: 3
- Samples in each class: 500, each point being 2D



Logistic regression classifier on NLS Dataset-1 (b)

Dataset description:

- Non Linearly separable dataset
- Total sample size: 1500
- Total classes: 3
- Samples in each class: 500
- Dimensionality: each point is 2D



Logistic regression for 32D BoVW representation.

LOGISTIC REGRESSION

Classification Accuracy: 0.7266666666666667

Classification Report:

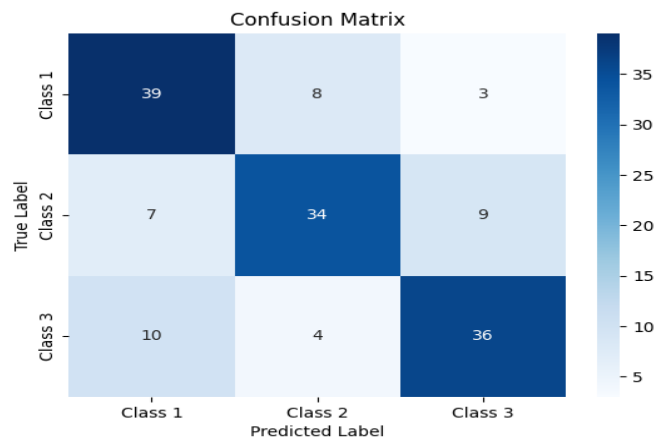
Class 1:
Precision: 0.6964
Recall: 0.7800
F1-Score: 0.7358
Support: 50.0

Class 2:
Precision: 0.7391
Recall: 0.6800
F1-Score: 0.7083
Support: 50.0

Class 3:
Precision: 0.7500
Recall: 0.7200
F1-Score: 0.7347
Support: 50.0

Mean Precision: 0.72851966873706
Mean Recall: 0.7266666666666666
Mean F1-Score: 0.7262920891627092

Confusion Matrix:
[[39 8 3]
[7 34 9]
[10 4 36]]



6. SVM-based classifier using (a) linear kernel, (b) polynomial kernel and (c) Gaussian/RBF kernel on Dataset-1 (a) Linearly Separable (LS).

Accuracy: 1.0

Precision per class: [1. 1. 1.]

Mean Precision: 1.0

Recall per class: [1. 1. 1.]

Mean Recall: 1.0

F1 Score per class: [1. 1. 1.]

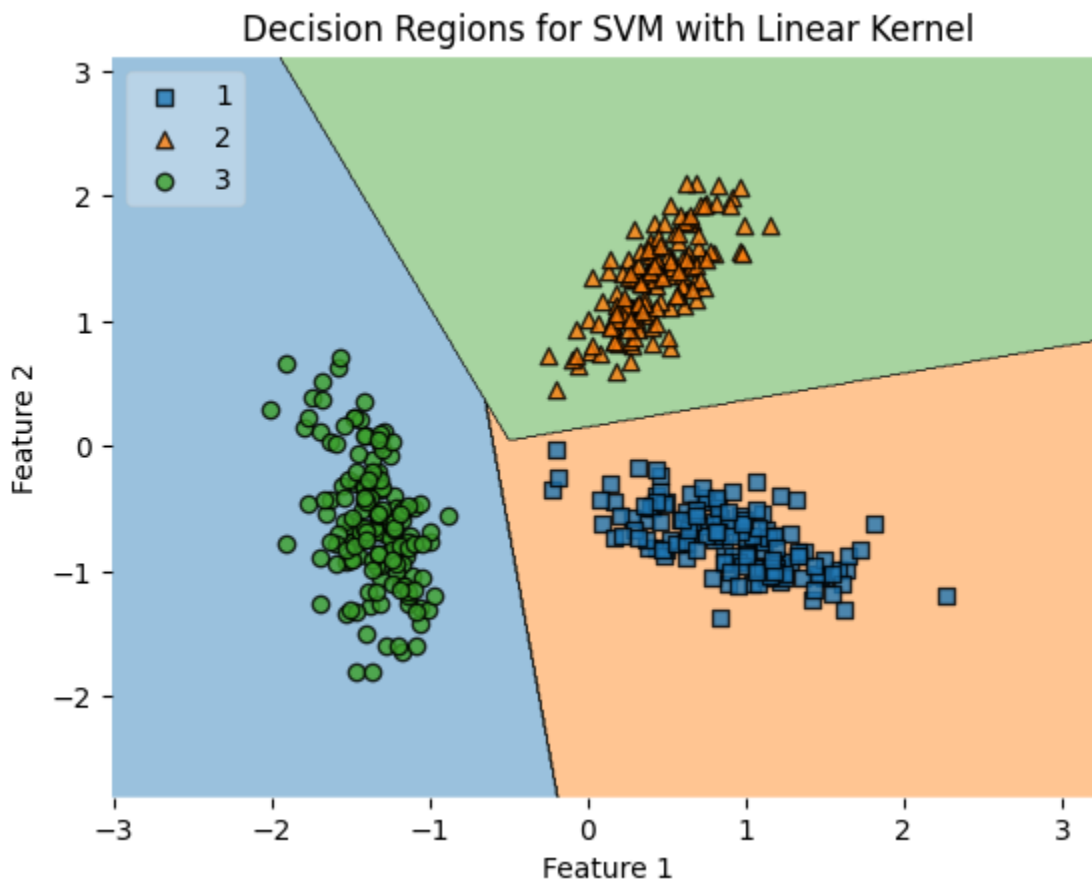
Mean F1 Score: 1.0

Confusion Matrix:

```
[[146  0  0]
```

```
 [ 0 147  0]
```

```
 [ 0  0 157]]
```



Accuracy: 0.9911111111111112

Precision per class: [0.97333333 1. 1.]

Mean Precision: 0.9911111111111112

Recall per class: [1. 0.97278912 1.]

Mean Recall: 0.9911111111111112

F1 Score per class: [0.98648649 0.9862069 1.]

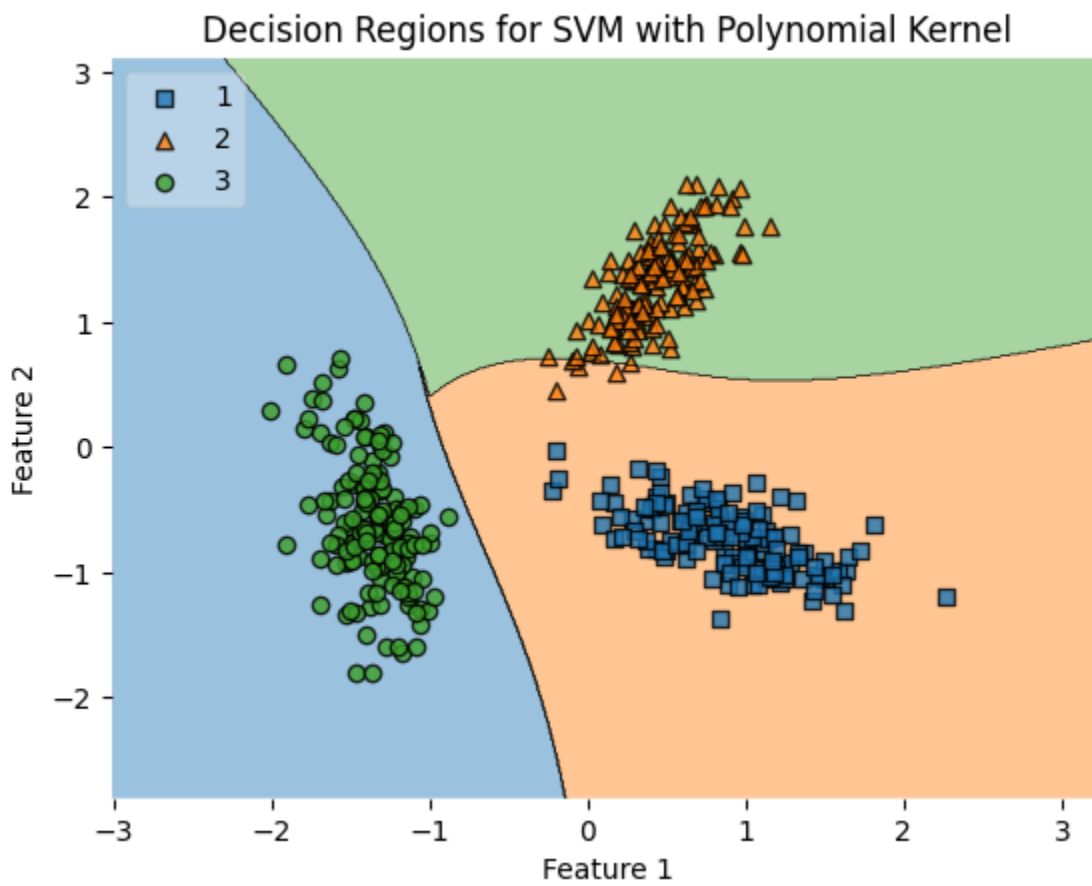
Mean F1 Score: 0.9911111111111112

Confusion Matrix:

```
[[146  0  0]
```

```
 [ 4 143  0]
```

```
 [ 0  0 157]]
```



Accuracy: 1.0

Precision per class: [1. 1. 1.]

Mean Precision: 1.0

Recall per class: [1. 1. 1.]

Mean Recall: 1.0

F1 Score per class: [1. 1. 1.]

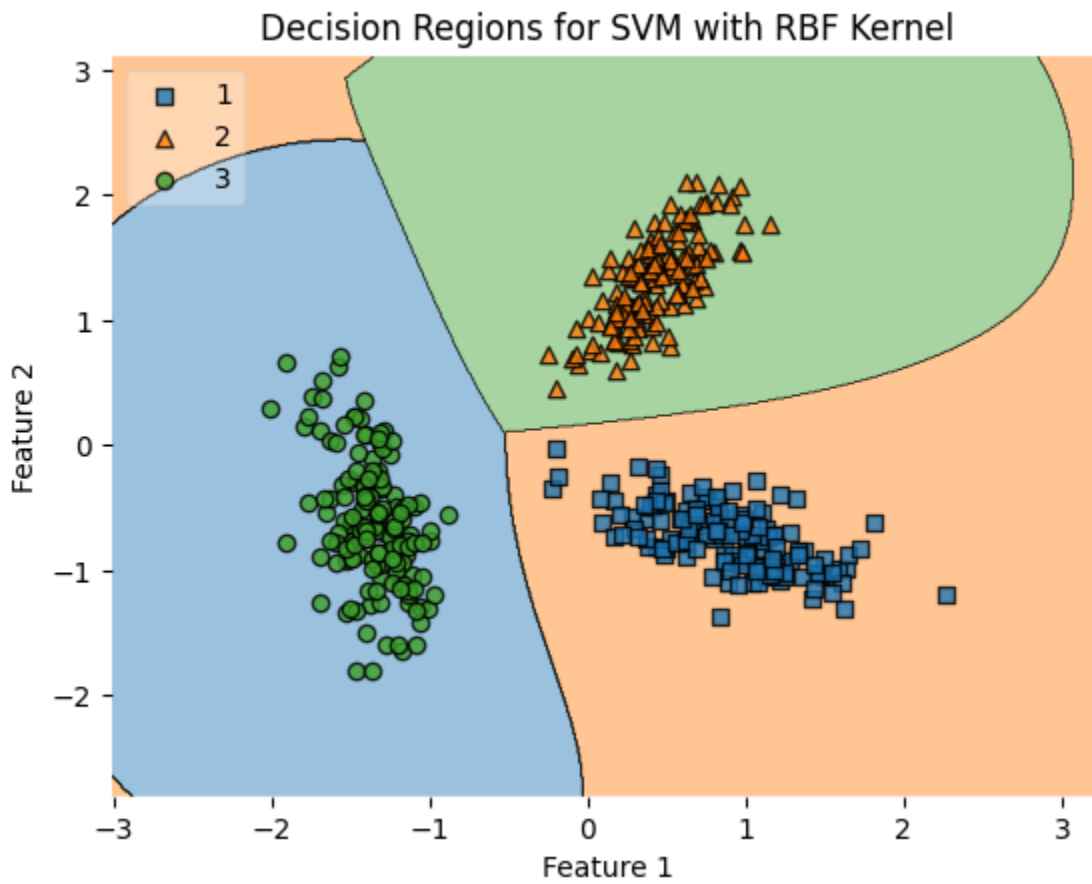
Mean F1 Score: 1.0

Confusion Matrix:

```
[[146  0  0]
```

```
 [ 0 147  0]
```

```
 [ 0  0 157]]
```



SVM-based classifier using (a) linear kernel, (b) polynomial kernel and (c) Gaussian/RBF kernel on Dataset-1 (b) Non Linearly Separable (NLS).

Accuracy: 0.9111111111111111

Precision per class: [0.88405797 0.92307692 0.92307692]

Mean Precision: 0.9111111111111111

Recall per class: [0.83561644 0.97959184 0.91719745]

Mean Recall: 0.9111111111111111

F1 Score per class: [0.85915493 0.95049505 0.9201278]

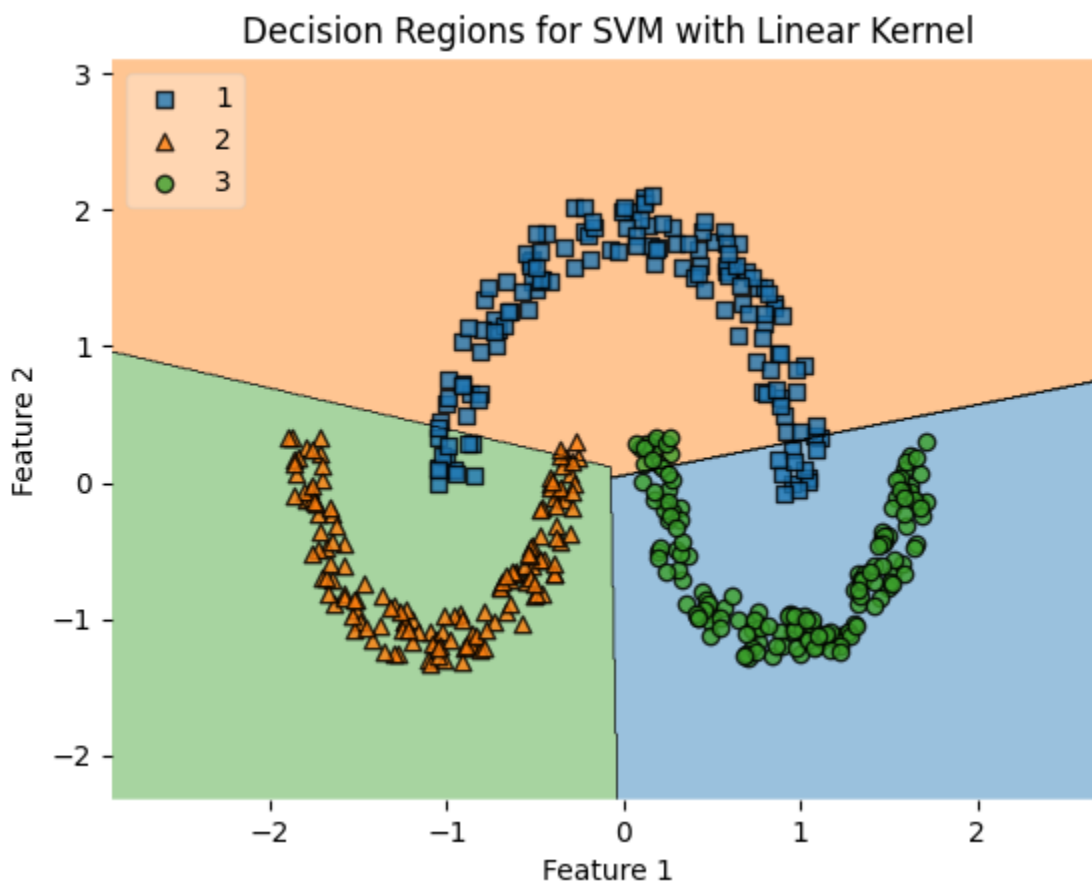
Mean F1 Score: 0.9111111111111111

Confusion Matrix:

[[122 12 12]

[3 144 0]

[13 0 144]]



Accuracy: 0.8711111111111111

Precision per class: [1. 0.77368421 0.89361702]

Mean Precision: 0.8711111111111111

Recall per class: [0.81506849 1. 0.80254777]

Mean Recall: 0.8711111111111111

F1 Score per class: [0.89811321 0.87240356 0.84563758]

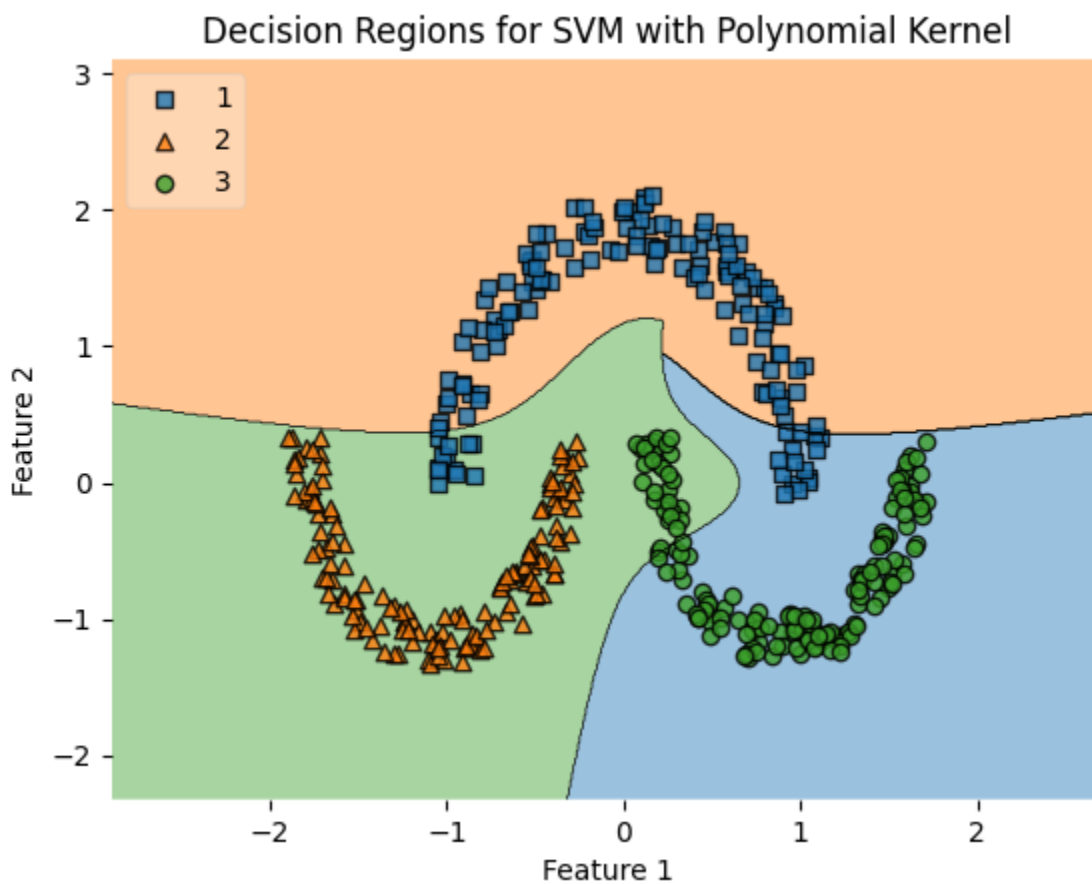
Mean F1 Score: 0.8711111111111111

Confusion Matrix:

[[119 12 15]

[0 147 0]

[0 31 126]]



Accuracy: 1.0

Precision per class: [1. 1. 1.]

Mean Precision: 1.0

Recall per class: [1. 1. 1.]

Mean Recall: 1.0

F1 Score per class: [1. 1. 1.]

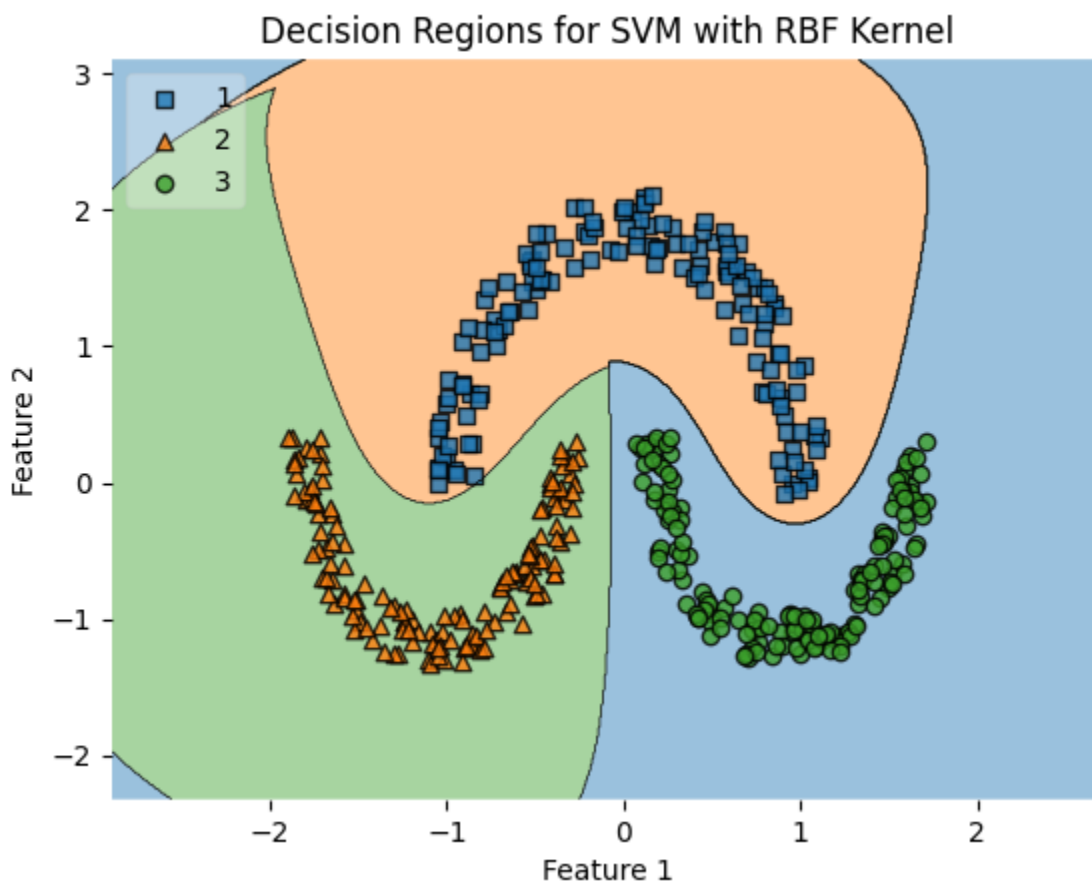
Mean F1 Score: 1.0

Confusion Matrix:

```
[[146  0  0]
```

```
 [  0 147  0]
```

```
 [  0  0 157]]
```



SVM-based classifier using (a) linear kernel, (b) polynomial kernel and (c) Gaussian/RBF kernel on Dataset-2 which has 32D BoVW representation.

SVM Linear Kernel

Classification Accuracy: 0.7

Classification Report:

Class 1:

Precision: 0.7111

Recall: 0.6400

F1-Score: 0.6737

Support: 50.0

Class 2:

Precision: 0.6316

Recall: 0.7200

F1-Score: 0.6729

Support: 50.0

Class 3:

Precision: 0.7708

Recall: 0.7400

F1-Score: 0.7551

Support: 50.0

Mean Precision: 0.7045077972709551

Mean Recall: 0.6999999999999998

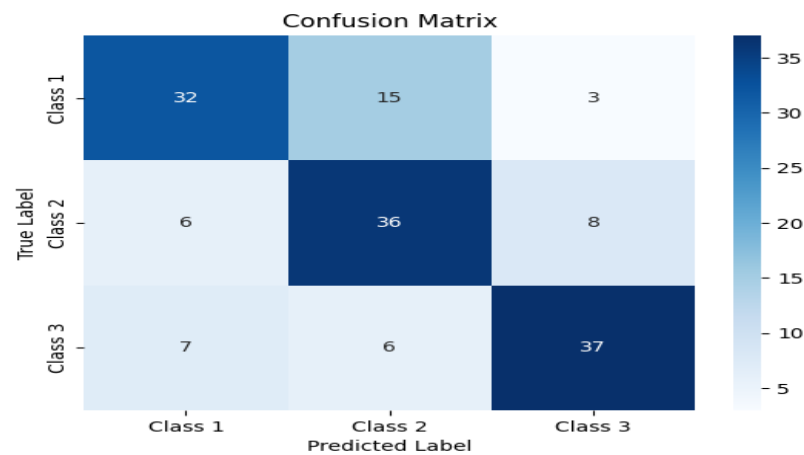
Mean F1-Score: 0.7005611492014415

Confusion Matrix:

[[32 15 3]

[6 36 8]

[7 6 37]]



SVM Polynomial Kernel

SVM Polynomial Kernel, degree = 1

Classification Accuracy: 0.7866666666666666

Classification Report:

Class 1:

Precision: 0.8163
Recall: 0.8000
F1-Score: 0.8081
Support: 50.0

Class 2:

Precision: 0.7708
Recall: 0.7400
F1-Score: 0.7551
Support: 50.0

Class 3:

Precision: 0.7736
Recall: 0.8200
F1-Score: 0.7961
Support: 50.0

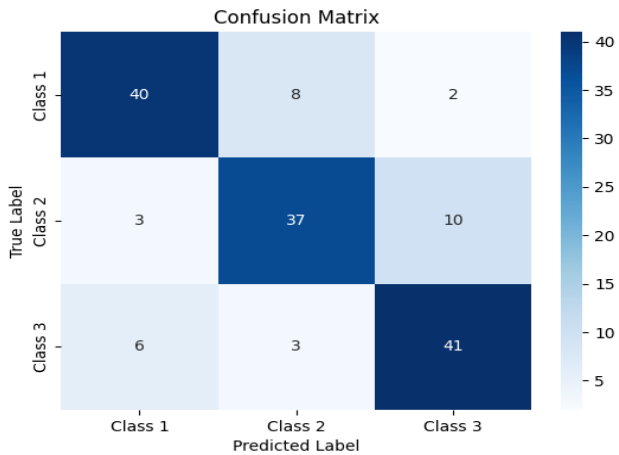
Mean Precision: 0.7869149232019853

Mean Recall: 0.7866666666666666

Mean F1-Score: 0.7864331179171679

Confusion Matrix:

```
[[40  8  2]
 [ 3 37 10]
 [ 6  3 41]]
```



SVM Polynomial Kernel, degree = 2

Classification Accuracy: 0.7666666666666667

Classification Report:

Class 1:

Precision: 0.7917

Recall: 0.7600

F1-Score: 0.7755

Support: 50.0

Class 2:

Precision: 0.7170

Recall: 0.7600

F1-Score: 0.7379

Support: 50.0

Class 3:

Precision: 0.7959

Recall: 0.7800

F1-Score: 0.7879

Support: 50.0

Mean Precision: 0.7681887220296924

Mean Recall: 0.7666666666666666

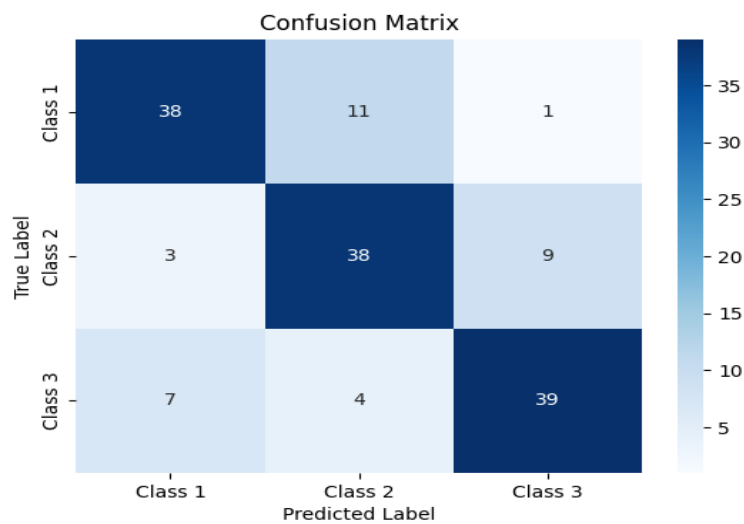
Mean F1-Score: 0.7670843565434411

Confusion Matrix:

[[38 11 1]

[3 38 9]

[7 4 39]]



SVM Polynomial Kernel, degree = 5

Classification Accuracy: 0.7

Classification Report:

Class 1:

Precision: 0.8108

Recall: 0.6000

F1-Score: 0.6897

Support: 50.0

Class 2:

Precision: 0.5915

Recall: 0.8400

F1-Score: 0.6942

Support: 50.0

Class 3:

Precision: 0.7857

Recall: 0.6600

F1-Score: 0.7174

Support: 50.0

Mean Precision: 0.7293581307665815

Mean Recall: 0.7000000000000001

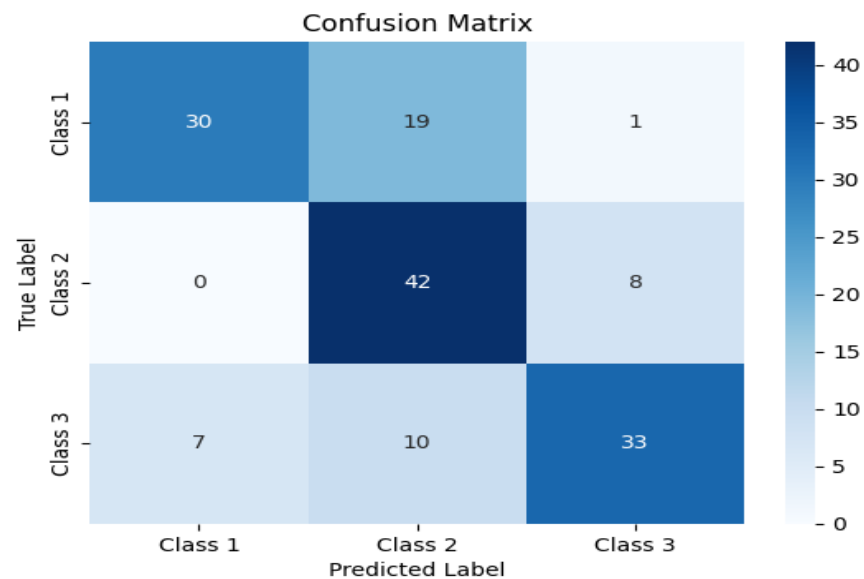
Mean F1-Score: 0.700420450931559

Confusion Matrix:

```
[[30 19  1]
```

```
[ 0 42  8]
```

```
[ 7 10 33]]
```



SVM RBF Kernel

SVM RBF Kernel: degree = 1, gamma = 0.001, width = 1000.0

Classification Accuracy: 0.64

Classification Report:

Class 1:

Precision: 0.8261

Recall: 0.3800

F1-Score: 0.5205

Support: 50.0

Class 2:

Precision: 0.5056

Recall: 0.9000

F1-Score: 0.6475

Support: 50.0

Class 3:

Precision: 0.8421

Recall: 0.6400

F1-Score: 0.7273

Support: 50.0

Mean Precision: 0.7246033990692412

Mean Recall: 0.64

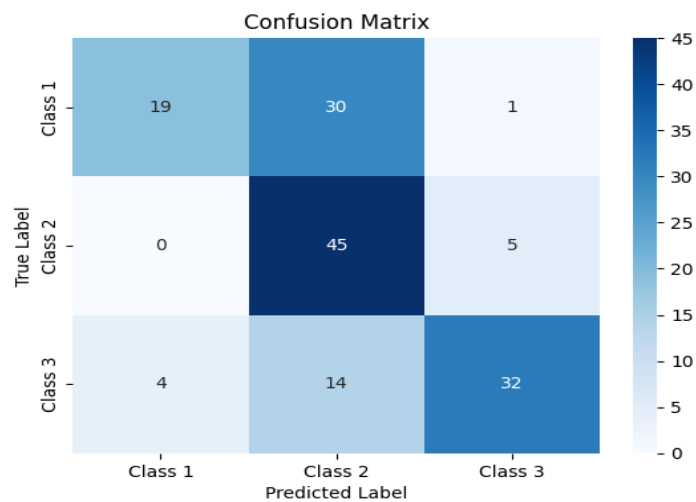
Mean F1-Score: 0.6317675622888985

Confusion Matrix:

```
[[19 30 1]
```

```
[ 0 45 5]
```

```
[ 4 14 32]]
```



SVM RBF Kernel: degree = 1, gamma = 0.01, width = 100.0

Classification Accuracy: 0.64

Classification Report:

Class 1:

Precision: 0.8261

Recall: 0.3800

F1-Score: 0.5205

Support: 50.0

Class 2:

Precision: 0.5056

Recall: 0.9000

F1-Score: 0.6475

Support: 50.0

Class 3:

Precision: 0.8421

Recall: 0.6400

F1-Score: 0.7273

Support: 50.0

Mean Precision: 0.7246033990692412

Mean Recall: 0.64

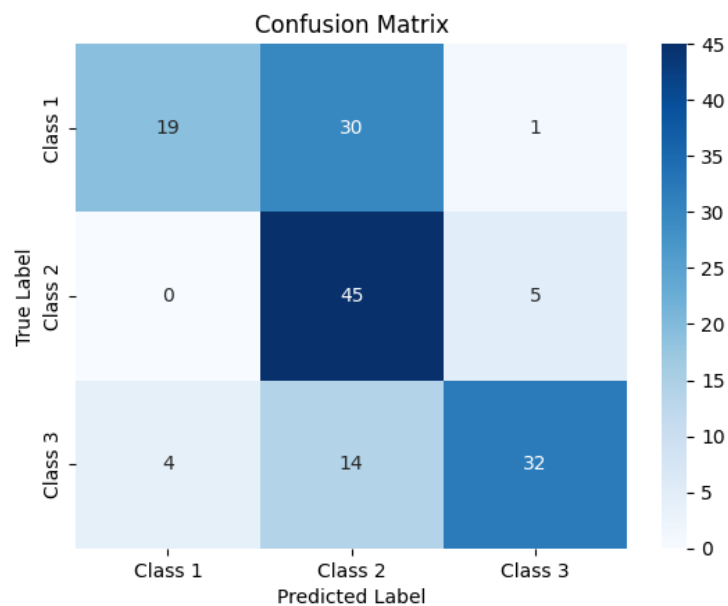
Mean F1-Score: 0.6317675622888985

Confusion Matrix:

[[19 30 1]

[0 45 5]

[4 14 32]]



SVM RBF Kernel: degree = 1, gamma = 0.5, width = 2.0

Classification Accuracy: 0.7133333333333334

Classification Report:

Class 1:

Precision: 0.8108

Recall: 0.6000

F1-Score: 0.6897

Support: 50.0

Class 2:

Precision: 0.6250

Recall: 0.8000

F1-Score: 0.7018

Support: 50.0

Class 3:

Precision: 0.7551

Recall: 0.7400

F1-Score: 0.7475

Support: 50.0

Mean Precision: 0.7303042838757126

Mean Recall: 0.7133333333333333

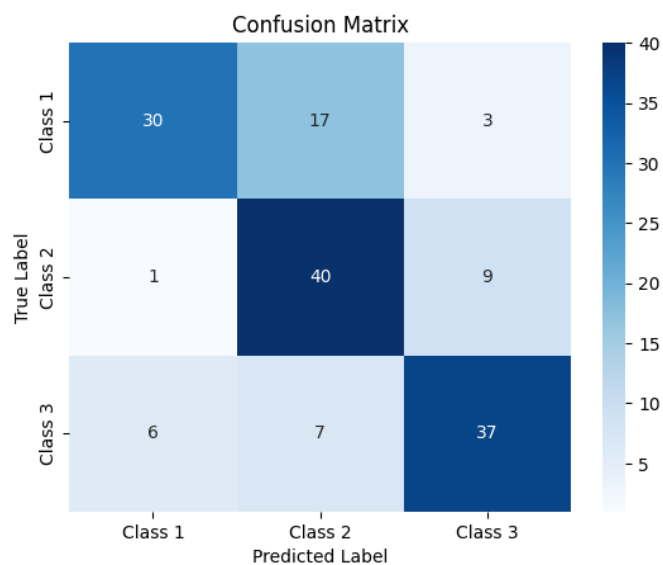
Mean F1-Score: 0.7129614352844843

Confusion Matrix:

[[30 17 3]

[1 40 9]

[6 7 37]]



SVM RBF Kernel: degree = 1, gamma = 0.1, width = 10.0

Classification Accuracy: 0.64

Classification Report:

Class 1:

Precision: 0.8261

Recall: 0.3800

F1-Score: 0.5205

Support: 50.0

Class 2:

Precision: 0.5056

Recall: 0.9000

F1-Score: 0.6475

Support: 50.0

Class 3:

Precision: 0.8421

Recall: 0.6400

F1-Score: 0.7273

Support: 50.0

Mean Precision: 0.7246033990692412

Mean Recall: 0.64

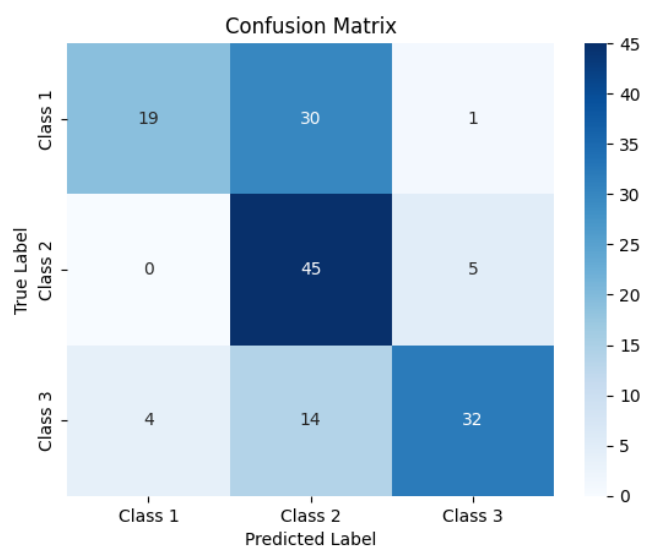
Mean F1-Score: 0.6317675622888985

Confusion Matrix:

```
[[19 30 1]
```

```
[ 0 45 5]
```

```
[ 4 14 32]]
```



SVM RBF Kernel: degree = 5, gamma = 0.01, width = 100.0

Classification Accuracy: 0.64

Classification Report:

Class 1:

Precision: 0.8261

Recall: 0.3800

F1-Score: 0.5205

Support: 50.0

Class 2:

Precision: 0.5056

Recall: 0.9000

F1-Score: 0.6475

Support: 50.0

Class 3:

Precision: 0.8421

Recall: 0.6400

F1-Score: 0.7273

Support: 50.0

Mean Precision: 0.7246033990692412

Mean Recall: 0.64

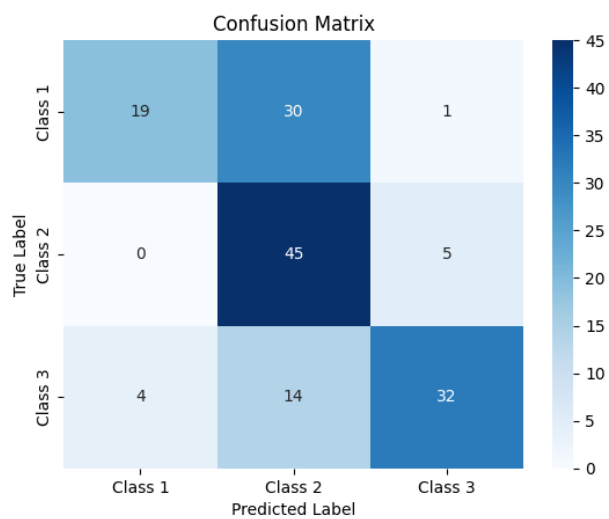
Mean F1-Score: 0.6317675622888985

Confusion Matrix:

```
[[19 30  1]
```

```
[ 0 45  5]
```

```
[ 4 14 32]]
```



SVM RBF Kernel: degree = 5, gamma = 0.5, width = 2.0

Classification Accuracy: 0.7133333333333334

Classification Report:

Class 1:

Precision: 0.8108
Recall: 0.6000
F1-Score: 0.6897
Support: 50.0

Class 2:

Precision: 0.6250
Recall: 0.8000
F1-Score: 0.7018
Support: 50.0

Class 3:

Precision: 0.7551
Recall: 0.7400
F1-Score: 0.7475
Support: 50.0

Mean Precision: 0.7303042838757126

Mean Recall: 0.7133333333333333

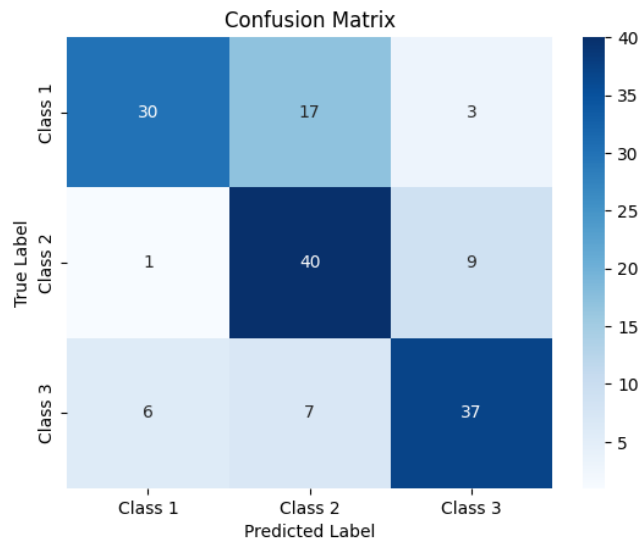
Mean F1-Score: 0.7129614352844843

Confusion Matrix:

[[30 17 3]

[1 40 9]

[6 7 37]]



SVM RBF Kernel: degree = 5, gamma = 0.1, width = 10.0

Classification Accuracy: 0.64

Classification Report:

Class 1:

Precision: 0.8261

Recall: 0.3800

F1-Score: 0.5205

Support: 50.0

Class 2:

Precision: 0.5056

Recall: 0.9000

F1-Score: 0.6475

Support: 50.0

Class 3:

Precision: 0.8421

Recall: 0.6400

F1-Score: 0.7273

Support: 50.0

Mean Precision: 0.7246033990692412

Mean Recall: 0.64

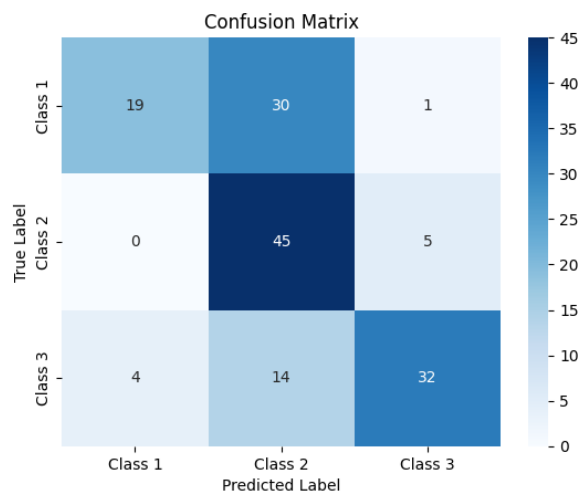
Mean F1-Score: 0.6317675622888985

Confusion Matrix:

```
[[19 30 1]
```

```
[ 0 45 5]
```

```
[ 4 14 32]]
```



The various values of parameters have been taken along with various degrees so essentially we have done experiments on various values of parameters.

