# PATTERN RECOGNITION AND MACHINE LEARNING

**CS5691 (JUL-NOV 2018)**


## Indian Institute of Technology, Madras

**M. Tech (2018-19) Computer Science and Engineering**

**SUBMITTED TO**

**Prof. C. Chandra Sekhar (CSE Dept, IITM)**

**TEAM 8**

**Neha Sah (CS18M033)**

**Avani Shukla (CS18M052)**

**Aarti Tapdiya (CS18M055)**

## Dataset 2: 2-dimensional artificial data

## (a) Linearly separable data set for static pattern classification

### Classifier 1 : Multi-class logistic regression based classifier using polynomial basis functions

| M | Learning rate (η) | Training Accuracy (%) | Validation Accuracy (%) |
|---|---|---|---|
| 1 | 0.0001 | 100 | 100 |
| 1 | 0.00001 | 95.3333 | 96.2222 |
| 2 | 0.00001 | 98.9333 | 98.4444 |

*Table 1.1 classification accuracies for different values of hyperparameter*

- Best Configuration model is when M=1 and η = 0.0001
- Test accuracy : 100%

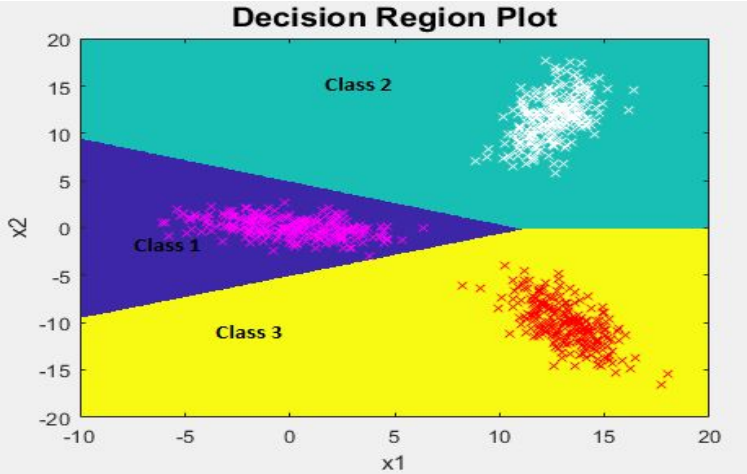


*Figure 1.1 plot for best performing model of training data N=750 for M=1 and η = 0.0001*

| | Class 1 | Class 2 | Class 3 |
|---|---|---|---|
| **Class 1** | 250 | 0 | 0 |
| **Class 2** | 0 | 250 | 0 |
| **Class 3** | 0 | 0 | 250 |

*Table 1.2 Confusion matrix for training data*

| | Class 1 | Class 2 | Class 3 |
|---|---|---|---|
| **Class 1** | 100 | 0 | 0 |
| **Class 2** | 0 | 100 | 0 |
| **Class 3** | 0 | 0 | 100 |

*Tabel 1.3 Confusion matrix for test data*

### Classifier 2 : Multi-class logistic regression based classifier using Gaussian basis functions

| k | σ | Training Accuracy (%) | Validation Accuracy (%) |
|---|---|---|---|
| 3 | 5 | 100 | 100 |
| 4 | 5 | 100 | 100 |
| 3 | 9 | 78.5333 | 78.8889 |
| 5 | 6 | 100 | 100 |

*Table 1.4 classification accuracies for different values of hyperparameter*

- Best Configuration model is when k=3 and σ=5
- Test accuracy : 100%

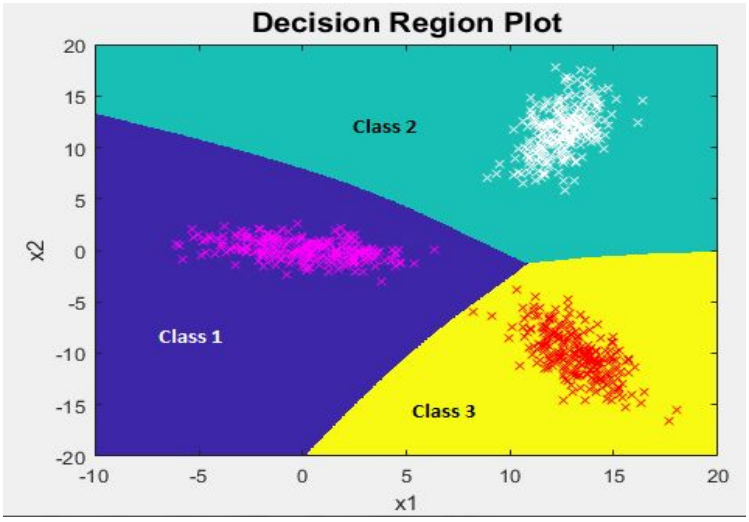


*Figure 1.2 plot for best performing model of training data N=750 for k=3 and σ=5*

| | Class 1 | Class 2 | Class 3 |
|---|---|---|---|
| **Class 1** | 250 | 0 | 0 |
| **Class 2** | 0 | 250 | 0 |
| **Class 3** | 0 | 0 | 250 |

*Table 1.5 Confusion matrix for training data*

| | Class 1 | Class 2 | Class 3 |
|---|---|---|---|
| **Class 1** | 100 | 0 | 0 |
| **Class 2** | 0 | 100 | 0 |
| **Class 3** | 0 | 0 | 100 |

*Table 1.6 Confusion matrix for test data*

**Classifier 3 : Perceptron**

Hyperparameter : Number of epochs E

| E | Training Accuracy (%) | Validation Accuracy (%) |
|---|---|---|
| 1 | 68.93 | 69.55 |
| 3 | 91.20 | 93.11 |
| 4 | 95.60 | 97.55 |
| 5 | 100 | 100 |

*Table 1.7 classification accuracies for different values of hyperparameter*

- Best configuration of the model when E=5
- Testing Accuracy 99.66%

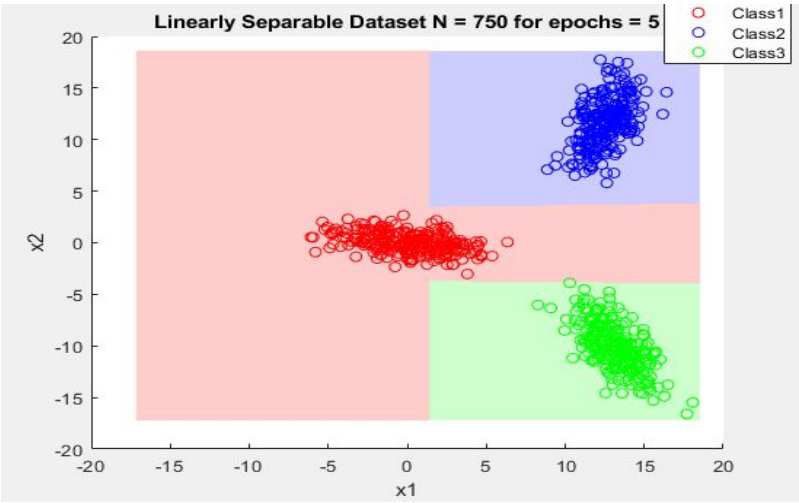


*Figure 1.3 plot for best performing model of training data N=750 for epochs = 5*

| | Class 1 | Class 2 | Class 3 |
|---|---|---|---|
| **Class 1** | 250 | 0 | 0 |
| **Class 2** | 0 | 250 | 0 |
| **Class 3** | 0 | 0 | 250 |

*Table 1.8 Confusion matrix for training data*

| | Class 1 | Class 2 | Class 3 |
|---|---|---|---|
| **Class 1** | 100 | 0 | 0 |
| **Class 2** | 0 | 100 | 0 |
| **Class 3** | 1 | 0 | 99 |

*Table 1.9 Confusion matrix for test data*

**Classifier 4 : Multilayer feedforward neural network (MLFFNN) with 2 hidden layers**

Hyperparameter : Number of epochs E, number of nodes in hidden layer 1 and 2, J1 and J2 respectively

| E | J1 | J2 | Training Accuracy (%) | Validation Accuracy (%) |
|---|---|---|---|---|
| 3000 | 2 | 3 | 66.40 | 66.44 |
| 9000 | 2 | 3 | 100 | 99.77 |
| 2000 | 3 | 4 | 90.26 | 90 |
| 3000 | 3 | 4 | 99.86 | 99.77 |
| 2000 | 4 | 4 | 99.86 | 99.77 |

*Table 1.10 classification accuracies for different values of hyperparameter*

- Best configuration of the model when E=9000, J1=2, J2=3,
- Testing Accuracy 99.66%

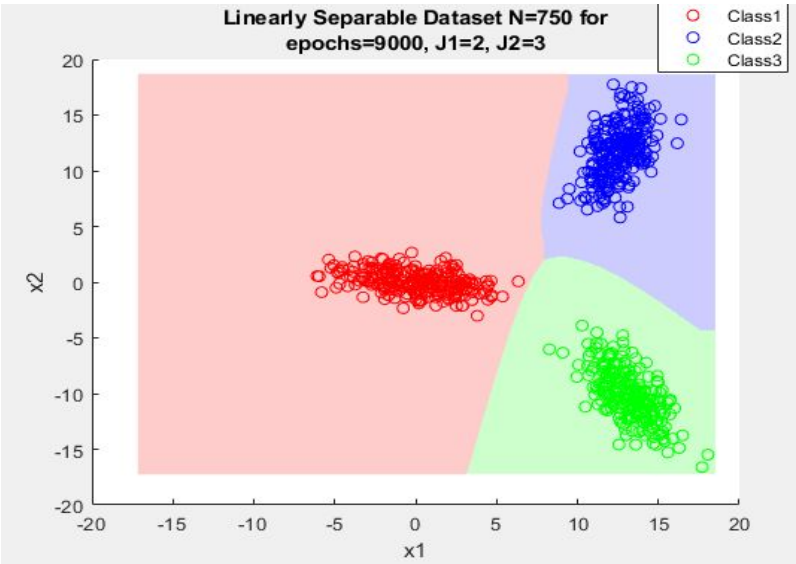


*Figure 1.4 plot for best performing model of training data N=750 for E=9000, J1=2 and J2=3*

| | Class 1 | Class 2 | Class 3 |
|---|---|---|---|
| **Class 1** | 250 | 0 | 0 |
| **Class 2** | 0 | 250 | 0 |
| **Class 3** | 0 | 0 | 250 |

*Table 1.11 Confusion matrix for training data*

| | Class 1 | Class 2 | Class 3 |
|---|---|---|---|
| **Class 1** | 99 | 0 | 1 |
| **Class 2** | 0 | 100 | 0 |
| **Class 3** | 1 | 0 | 99 |

*Table 1.12 Confusion matrix for test data*

**C-SVM:**

**Objective Function:**

$$min \ \frac{1}{2} \| w \|^2 + C \sum_i \xi_i$$

**Subject to**

$$y_i (w x_i + b) \geq 1 - \xi_i \ , \ \forall x_i$$

$$\xi_i \geq 0$$

**Kernel Function:**

**Linear:** u'*v

**Polynomial:** (gamma*u'*v + coef0)^degree

**Radial basis function(Gaussian):** exp(-gamma*|u-v|^2)

-d degree : set degree in kernel function

-g gamma : set gamma in kernel function

-r coef0 : set coef0 in kernel function

-c cost : set the parameter C of C-SVC

**Classifier 5: Linear kernel based C-SVM**

Hyperparameter : Cost C

| C | Training Accuracy (%) | Validation Accuracy (%) |
|---|---|---|
| 0.01 | 100 | 100 |
| 1 | 100 | 100 |

*Table 1.13 classification accuracies for different values of hyperparameter*

- Best configuration of the model when C=1
- Testing Accuracy 100%

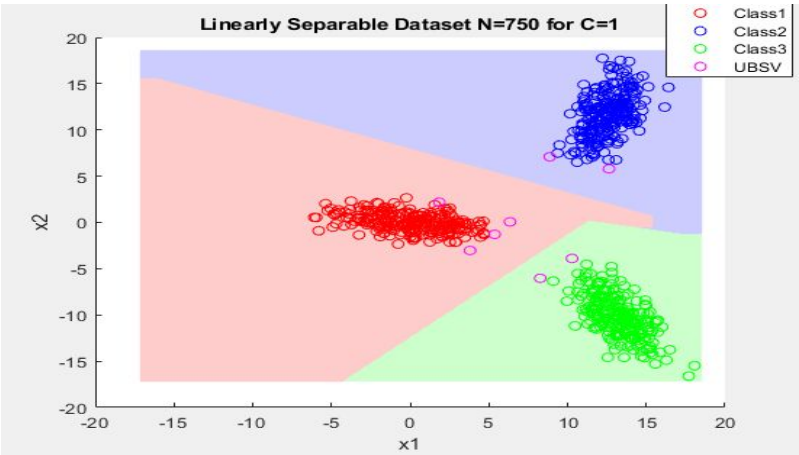


*Figure 1.5 plot for best performing model of training data N=750 for C=1*

| | Class 1 | Class 2 | Class 3 |
|---|---|---|---|
| **Class 1** | 250 | 0 | 0 |
| **Class 2** | 0 | 250 | 0 |
| **Class 3** | 0 | 0 | 250 |

*Table 1.14 Confusion matrix for training data*

| | Class 1 | Class 2 | Class 3 |
|---|---|---|---|
| **Class 1** | 100 | 0 | 0 |
| **Class 2** | 0 | 100 | 0 |
| **Class 3** | 0 | 0 | 100 |

*Table 1.15 Confusion matrix for test data*

**Classifier 6 : Nonlinear kernel based C-SVM using polynomial kernel**

Hyperparameter : Cost C, degree d, gamma g

| C | d | g | Training Accuracy (%) | Validation Accuracy (%) |
|---|---|---|---|---|
| 0.01 | 2 | 0.01 | 99.86 | 100 |
| 1 | 2 | 0.01 | 100 | 100 |
| 10 | 2 | 0.01 | 100 | 100 |
| 1 | 3 | 0.01 | 100 | 100 |

*Table 1.16 classification accuracies for different values of hyperparameter*

- Best configuration of the model when C=1, d=2, g=0.01
- Testing Accuracy 100%

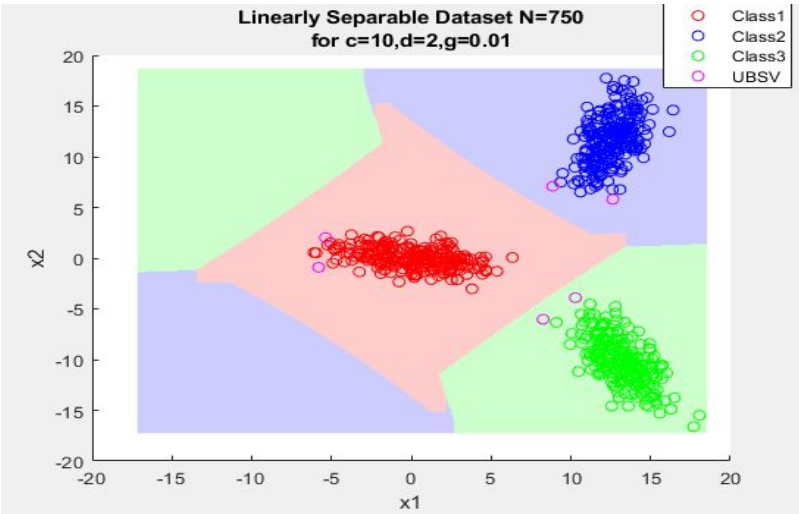


*Figure 1.6 plot for best performing model of training data N=750 for C=1, d=2 and g=0.01*

| | Class 1 | Class 2 | Class 3 |
|---|---|---|---|
| **Class 1** | 250 | 0 | 0 |
| **Class 2** | 0 | 250 | 0 |
| **Class 3** | 0 | 0 | 250 |

*Table 1.17 Confusion matrix for training data*

| | Class 1 | Class 2 | Class 3 |
|---|---|---|---|
| **Class 1** | 100 | 0 | 0 |
| **Class 2** | 0 | 100 | 0 |
| **Class 3** | 0 | 0 | 100 |

*Table 1.18 Confusion matrix for test data*

**Classifier 7 : Nonlinear kernel based C-SVM using Gaussian kernel.**

Hyperparameter : Cost C, gamma g

| C | g | Training Accuracy (%) | Validation Accuracy (%) |
|---|---|---|---|
| 0.01 | 0.01 | 100 | 100 |
| 1 | 0.01 | 100 | 100 |
| 10 | 0.01 | 100 | 100 |

*Table 1.19 classification accuracies for different values of hyperparameter*

- Best configuration of the model when C=10, g=0.01
- Testing Accuracy 100%

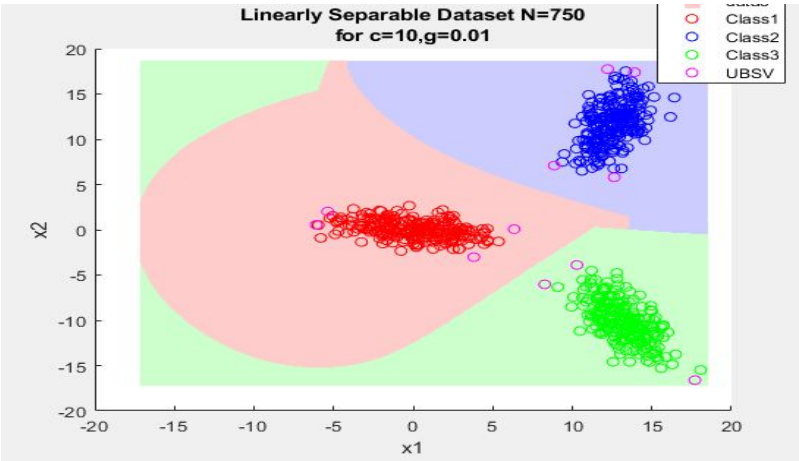


*Figure 1.7 plot for best performing model of training data N=750 for C=10 and g=0.01*

| | Class 1 | Class 2 | Class 3 |
|---|---|---|---|
| **Class 1** | 250 | 0 | 0 |
| **Class 2** | 0 | 250 | 0 |
| **Class 3** | 0 | 0 | 250 |

*Table 1.20 Confusion matrix for training data*

| | Class 1 | Class 2 | Class 3 |
|---|---|---|---|
| **Class 1** | 100 | 0 | 0 |
| **Class 2** | 0 | 100 | 0 |
| **Class 3** | 0 | 0 | 100 |

*Table 1.21 Confusion matrix for test data*

## (b) NonLinearly separable data set for static pattern classification

### Classifier 1 : Multi-class logistic regression based classifier using polynomial basis functions

| M | η | Training Accuracy (%) | Validation Accuracy (%) |
|---|---|---|---|
| 2 | 0.00001 | 92.60 | 95 |
| 1 | 0.00001 | 91.40 | 93 |
| 2 | 0.0001 | 94.20 | 96.67 |
| 1 | 0.0001 | 91.40 | 94 |
| 2 | 0.000001 | 84.20 | 88 |
| 1 | 0.000001 | 86 | 88.33 |

*Table 2.1 classification accuracies for different values of hyperparameter*

*Best model has test accuracy: 92.5%*
*with M = 2 and η=0.0001*

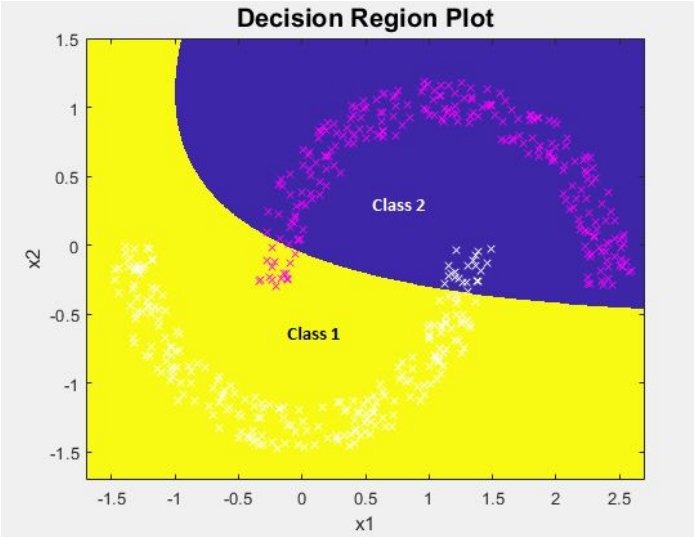


*Figure 2.1 plot for best performing model of training data N=750 for M = 2 and η=0.0001*

| | Class 1 | Class 2 |
|---|---|---|
| **Class 1** | 236 | 14 |
| **Class 2** | 15 | 235 |

*Table 2.2 Confusion matrix for training data*

| | Class 1 | Class 2 |
|---|---|---|
| **Class 1** | 90 | 10 |
| **Class 2** | 5 | 95 |

*Table 2.3 Confusion matrix for test data*

### Classifier 2 : Multi-class logistic regression based classifier using polynomial basis functions Gaussian basis functions

| k | σ | η | Training Accuracy (%) | Validation Accuracy (%) |
|---|---|---|---|---|
| 4 | 5 | 0.00001 | 93.2000 | 95 |
| 4 | 6 | 0.0001 | 91.6000 | 93 |
| 2 | 6 | 0.0001 | 82.6000 | 80.3333 |
| 5 | 6 | 0.0001 | 83.8000 | 89.6667 |

*Table 2.4 classification accuracies for different values of hyperparameter*

- Best model has test accuracy : 92.5000
  When k=4 , σ= 5 and η = 0.00001

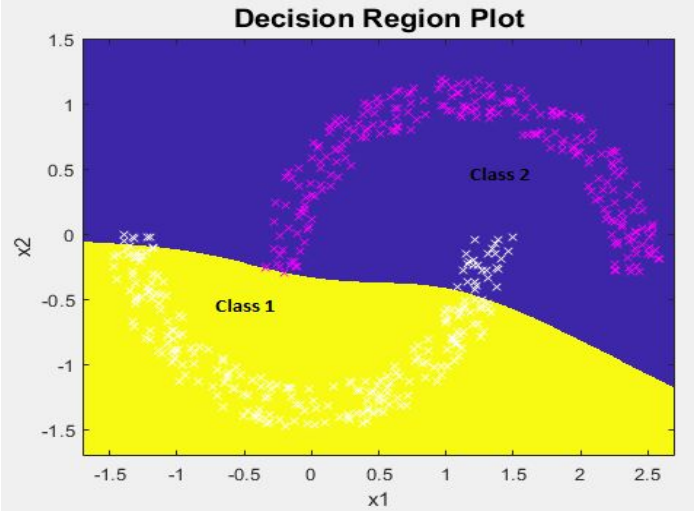


*Figure 2.2 plot for best performing model of training*

*data N=750 for k=4 , σ= 5 and η = 0.00001*

| | Class 1 | Class 2 |
|---|---|---|
| **Class 1** | 250 | 0 |
| **Class 2** | 34 | 216 |

*Table 2.5 Confusion matrix for training data*

| | Class 1 | Class 2 |
|---|---|---|
| **Class 1** | 96 | 4 |
| **Class 2** | 11 | 89 |

*Table 2.6 Confusion matrix for test data*

**Classifier 3 : Multilayer feedforward neural network (MLFFNN) with 2 hidden layers**

Hyperparameter : Number of epochs E, number of nodes in hidden layer 1 and 2, J1 and J2 respectively

| E | J1 | J2 | Training Accuracy (%) | Validation Accuracy (%) |
|---|---|---|---|---|
| 1000 | 3 | 4 | 83 | 86 |
| 1000 | 5 | 5 | 93.80 | 95.33 |
| 2000 | 3 | 4 | 90 | 92 |
| 5000 | 7 | 6 | 97.80 | 99.33 |
| 10000 | 5 | 5 | 99.20 | 99.66 |
| 10000 | 8 | 7 | 100 | 100 |

*Table 2.7 classification accuracies for different values of hyperparameter*

- Best configuration of the model when E=10000, J1=8, J2=7
- Testing Accuracy 100%

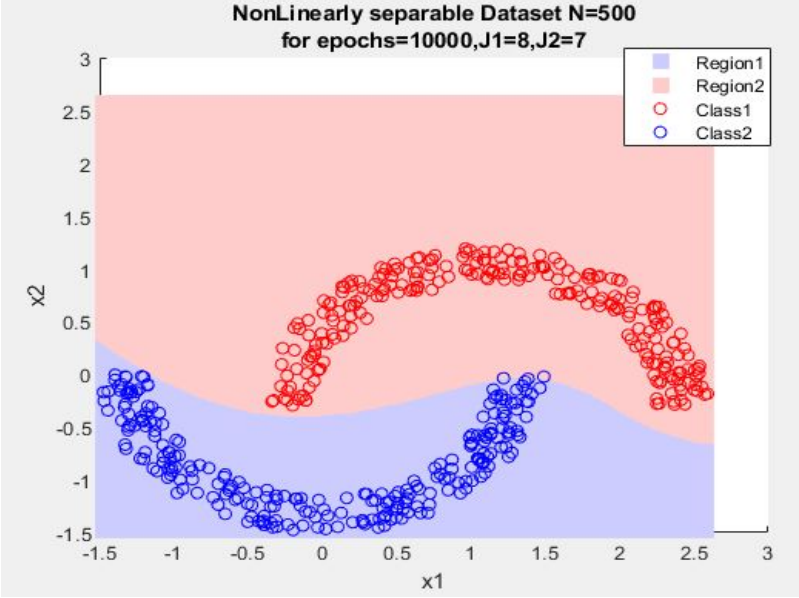


*Figure 2.3 plot for best performing model of training data N=750 for E=10000, J1=8 and J2=7*

| | Class 1 | Class 2 |
|---|---|---|
| **Class 1** | 250 | 0 |
| **Class 2** | 0 | 250 |

*Table 2.8 Confusion matrix for training data*

| | Class 1 | Class 2 |
|---|---|---|
| **Class 1** | 100 | 0 |
| **Class 2** | 0 | 100 |

*Table 2.9 Confusion matrix for test data*

**Classifier 4 : Nonlinear kernel based C-SVM using polynomial kernel**

Hyperparameter : Cost C, degree d, gamma g

| C | d | g | Training Accuracy (%) | Validation Accuracy (%) |
|---|---|---|---|---|
| 0.01 | 3 | 1 | 96.6 | 97.33 |
| 1 | 3 | 0.2 | 90.20 | 91.66 |
| 1 | 3 | 2 | 97.6 | 98 |
| 10 | 3 | 2 | 97.6 | 98 |
| 100 | 4 | 0.5 | 77.8 | 78 |

*Table 2.10 classification accuracies for different values of hyperparameter*

- Best configuration of the model when C=10, d=3, g=2
- Testing Accuracy 98.50%

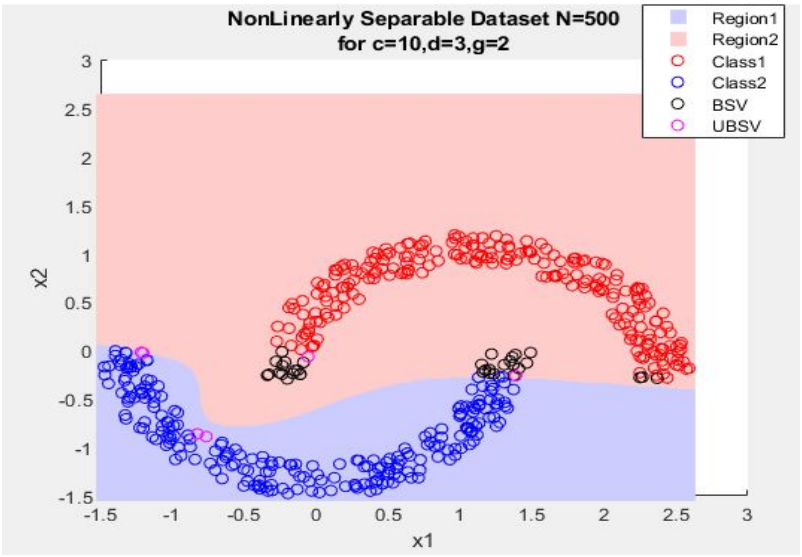


*Figure 2.4 plot for best performing model of training data N=750 for* C=10, d=3 and g=2

| | Class 1 | Class 2 |
|---|---|---|
| **Class 1** | 250 | 0 |
| **Class 2** | 12 | 238 |

*Table 2.11 Confusion matrix for training data*

| | Class 1 | Class 2 |
|---|---|---|
| **Class 1** | 100 | 0 |
| **Class 2** | 3 | 97 |

*Table 2.12 Confusion matrix for test data*

**Classifier 5 : Nonlinear kernel based C-SVM using Gaussian kernel.**

Hyperparameter : Cost C, gamma g

| C | g | Training Accuracy (%) | Validation Accuracy (%) |
|---|---|---|---|
| 1 | 0.5 | 100 | 100 |
| 10 | 0.5 | 100 | 100 |
| 10 | 0.01 | 94 | 95.66 |

*Table 2.13 classification accuracies for different values of hyperparameter*

- Best configuration of the model when C=10, g=0.5
- Testing Accuracy 100%

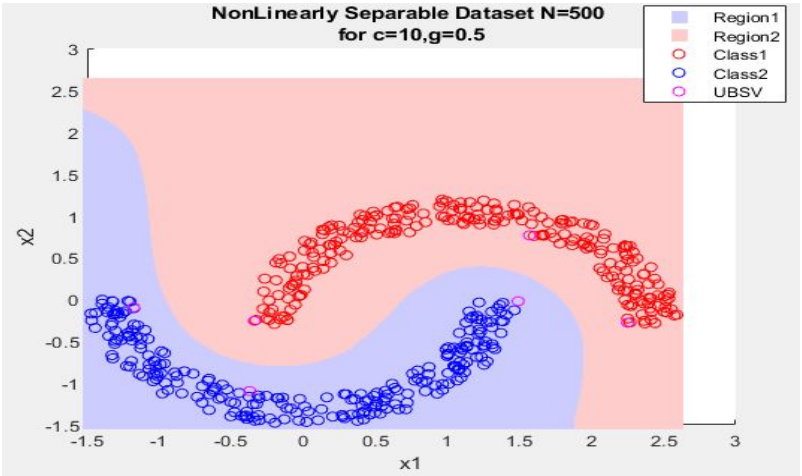


*Figure 2.5 plot for best performing model of training data N=750 for C=10, g=0.5*

| | Class 1 | Class 2 |
|---|---|---|
| **Class 1** | 250 | 0 |
| **Class 2** | 0 | 250 |

*Table 2.14 Confusion matrix for training data*

| | Class 1 | Class 2 |
|---|---|---|
| **Class 1** | 100 | 0 |
| **Class 2** | 0 | 100 |

*Table 2.15 Confusion matrix for test data*

## Dataset 3: Image data set

### Classifier 1 : Multi-class logistic regression based classifier using polynomial basis functions

| M | σ | η | Training Accuracy (%) | Validation Accuracy (%) |
|---|---|---|---|---|
| 2 | 9 | 0.00001 | 43.7795 | 38.89 |
| 6 | 9 | 0.00001 | 60.7874 | 42.43 |

*Table 3.1 classification accuracies for different values of hyperparameter*

- Best Model has Test accuracy : 34.44%
  With M= 6 , σ = 9 and η = 0.00001

| | Class 1 | Class 2 | Class 3 |
|---|---|---|---|
| **Class 1** | 182 | 0 | 0 |
| **Class 2** | 249 | 0 | 0 |
| **Class 3** | 0 | 0 | 204 |

*Table 3.2 Confusion matrix for training data*

| | Class 1 | Class 2 | Class 3 |
|---|---|---|---|
| **Class 1** | 12 | 13 | 1 |
| **Class 2** | 25 | 10 | 0 |
| **Class 3** | 13 | 7 | 9 |

*Table 3.3 Confusion matrix for test data*

### Classifier 2 : Multi-class logistic regression based classifier using gaussian basis functions

| k | σ | η | Training Accuracy (%) | Validation Accuracy (%) |
|---|---|---|---|---|
| 30 | 7 | 0.00001 | 62.04 | 60.43 |
| 50 | 15 | 0.00005 | 50.55 | 51.11 |

*Table 3.4 classification accuracies for different values of hyperparameter*

- Best Model has Test accuracy :38.79 %
  k = 30 , σ= 7 and η = 0.00001
- Model with k=50 , σ= 15 and η = 0.00005 has test accuracy : 51.91%

| | Class 1 | Class 2 | Class 3 |
|---|---|---|---|
| **Class 1** | 152 | 20 | 10 |
| **Class 2** | 85 | 126 | 38 |
| **Class 3** | 60 | 28 | 116 |

*Table 3.5 Confusion matrix for training data*

| | Class 1 | Class 2 | Class 3 |
|---|---|---|---|
| **Class 1** | 33 | 19 | 0 |
| **Class 2** | 34 | 38 | 0 |
| **Class 3** | 42 | 17 | 0 |

*Table 3.6 Confusion matrix for test data*

**Classifier 3 : Multilayer feedforward neural network (MLFFNN) with 2 hidden layers**

Hyperparameter : Number of epochs E, number of nodes in hidden layer 1 and 2, J1 and J2 respectively

| E | J1 | J2 | Training Accuracy (%) | Validation Accuracy (%) |
|---|---|---|---|---|
| 5000 | 20 | 15 | 80.31 | 92.22 |
| 10000 | 20 | 15 | 82.51 | 92.86 |
| 10000 | 25 | 23 | 84.72 | 93.33 |
| 10000 | 35 | 40 | 86.45 | 92.22 |
| 10000 | 60 | 70 | 85.35 | 93.33 |

*Table 3.7 classification accuracies for different values of hyperparameter*

- Best configuration of the model when E=10000, J1=25, J2=23
- Testing Accuracy 67.21%

| | Class 1 | Class 2 | Class 3 |
|---|---|---|---|
| **Class 1** | 229 | 2 | 18 |
| **Class 2** | 21 | 152 | 9 |
| **Class 3** | 31 | 5 | 168 |

*Table 3.8 Confusion matrix for training data*

| | Class 1 | Class 2 | Class 3 |
|---|---|---|---|
| **Class 1** | 62 | 1 | 9 |
| **Class 2** | 10 | 21 | 21 |
| **Class 3** | 18 | 1 | 40 |

*Table 3.9 Confusion matrix for test data*

**Classifier 4 : Nonlinear kernel based C-SVM using polynomial kernel**

Hyperparameter : Cost C, degree d, gamma g

| C | d | g | Training Accuracy (%) | Validation Accuracy (%) |
|---|---|---|---|---|
| 10 | 3 | 0.043 | 80 | 92.22 |
| 10 | 5 | 0.043 | 83.77 | 93.33 |

*Table 3.10 classification accuracies for different values of hyperparameter*

- Best configuration of the model when C=10, d=5, g=0.0435
- Testing Accuracy 62.30%

| | Class 1 | Class 2 | Class 3 |
|---|---|---|---|
| **Class 1** | 231 | 0 | 18 |
| **Class 2** | 30 | 137 | 15 |
| **Class 3** | 39 | 1 | 164 |

*Table 3.11 Confusion matrix for training data*

| | Class 1 | Class 2 | Class 3 |
|---|---|---|---|
| **Class 1** | 63 | 1 | 8 |
| **Class 2** | 18 | 11 | 23 |
| **Class 3** | 18 | 1 | 40 |

*Table 3.12 Confusion matrix for test data*

**Classifier 5 : Nonlinear kernel based C-SVM using Gaussian kernel.**

Hyperparameter : Cost C, gamma g

| C | g | Training Accuracy (%) | Validation Accuracy (%) |
|---|---|---|---|
| 1 | 0.0435 | 79.37 | 86.66 |
| 10 | 0.0435 | 81.41 | 93.33 |
| 100 | 0.0435 | 85.66 | 94.44 |

*Table 3.13 classification accuracies for different values of hyperparameter*

- Best configuration of the model when C=100, g=0.0435
- Testing Accuracy 61.74%

| | Class 1 | Class 2 | Class 3 |
|---|---|---|---|
| **Class 1** | 236 | 0 | 13 |
| **Class 2** | 29 | 141 | 12 |
| **Class 3** | 34 | 3 | 167 |

*Table 3.14 Confusion matrix for training data*

| | Class 1 | Class 2 | Class 3 |
|---|---|---|---|
| **Class 1** | 62 | 0 | 10 |
| **Class 2** | 20 | 11 | 21 |
| **Class 3** | 18 | 1 | 40 |

*Table 3.15 Confusion matrix for test data*

## Dataset 1 : Data for sequential pattern classification

## (a) On-line handwritten character data

**Classifier 1 : : Discrete HMM**

| k | No. of states | Test Accuracy (%) |
|---|---|---|
| 20 | 5 | 93.33 |
| 25 | 5 | 94.44 |
| 20 | 4 | 93.33 |

*Table 4.1 classification accuracies for different values of hyperparameter*

*Best Model has test Accuracy 94.44% with k=25 and no. of states = 5*

| | Class 1 | Class 2 | Class 3 |
|---|---|---|---|
| **Class 1** | 12 | 0 | 0 |
| **Class 2** | 1 | 11 | 0 |
| **Class 3** | 1 | 0 | 11 |

*Table 4.2 Confusion matrix for test data*

**Classifier 2: Linear kernel based C-SVM**

Hyperparameter : Cost C

| C | Training Accuracy (%) | Validation Accuracy (%) |
|---|---|---|
| 1 | 91.82 | 100 |
| 10 | 98.55 | 93.10 |
| 50 | 100 | 86.20 |

*Table 4.3 classification accuracies for different values of hyperparameter*

- Best configuration of the model when C=1
- Testing Accuracy 95.08%

| | Class 1 | Class 2 | Class 3 |
|---|---|---|---|
| **Class 1** | 65 | 1 | 4 |
| **Class 2** | 4 | 63 | 3 |
| **Class 3** | 5 | 0 | 63 |

*Table 4.4 Confusion matrix for training data*

| | Class 1 | Class 2 | Class 3 |
|---|---|---|---|
| **Class 1** | 19 | 0 | 1 |
| **Class 2** | 0 | 20 | 0 |
| **Class 3** | 2 | 0 | 19 |

*Table 4.5 Confusion matrix for test data*

**Observation:**

- On linear dataset all the classifiers give the best results .
- On non linear Multilayer feedforward neural network and C-SVM with Gaussian kernel give the best result.
- On image data set Multilayer feedforward neural network give the best result.