Kernel Methods for Pattern Analysis
Assignment 3

Jilt SebastianCS13D020
Jom KuriakoseCS13S028
Abil N GeorgeCS13M002

April 26, 2014

Contents

Part I
Classification Tasks

1 Data-sets

1.0.1 Data-set 1: 2-dimensional input data:

(a) Linearly separable classes, (b) Non-linearly separable classes, (c) Overlapping class

1.0.2 Data-set 2 : Image data

Models

2 Experiment No.1: C-SVM with Linear Kernel

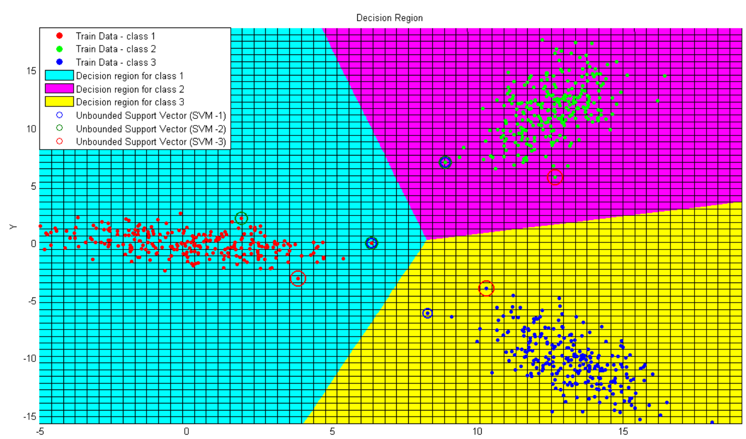
In this experiment,we were suppose to build C-SVM based classifier for the various types of data-sets given. One-against-Rest approach is used for all multiclass classification tasks.

2.1 Linearly Separable classes(Data-set 1(a))

Aim was to build classifier for the given data-set of Linearly Separable classes. Data is bi-variate and the classification task was a three class problem. A C-SVM Model is built for the given data-set using libsvm tool and examples in test data-set are classified using the trained model. One C-SVM model is built for each class and O-V-R classification is performed.

2.1.1 Decision Region Plots

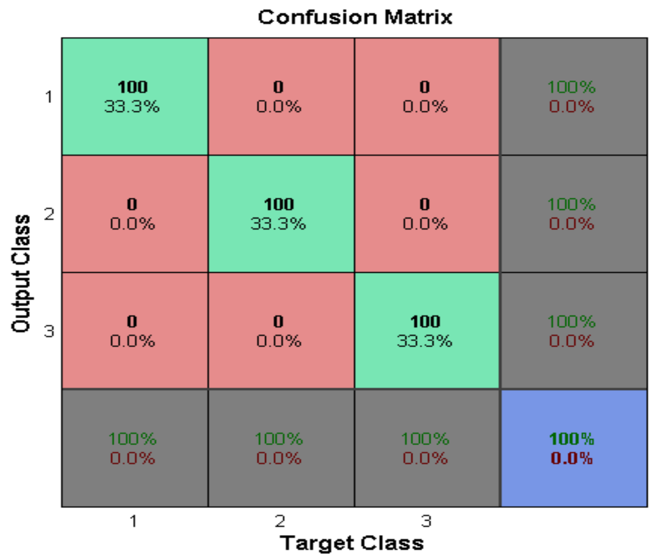
Below shown are decision region plots for the C-SVM Classifier with Linear Kernel.



(a) C-SVM Classifier with Linear Kernel

Figure 1: Decision Region Plots for Linearly Separable classes

2.1.2 Confusion Matrix and Average Classification Accuracy



(a) Confusion Matrix

Figure 2: C-SVM Classifier with Linear Kernel

Classification Accuracy

1	1	1

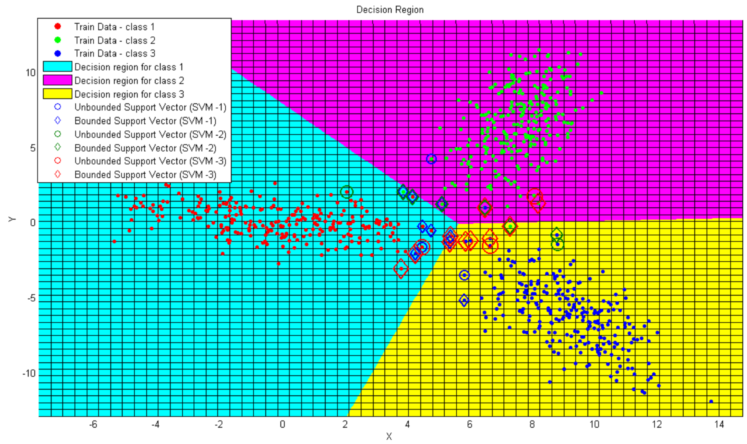
Average Accuracy=1

2.2 Overlapping classes(Data-set 1(c))

Aim was to build C-SVM Classifier with Linear Kernel for the given data-set of Overlapping Classes. Data is bi-variate and the classification task was a three class problem. Model is built for each class to classify a given element to any of the three classes and to find the Classification Accuracy, Confusion Matrix and Decision Region Plot.

2.2.1 Decision Region Plots

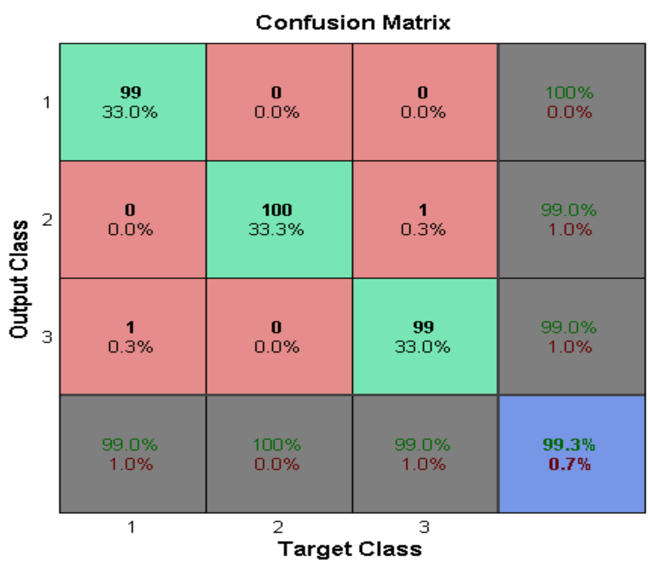
Below shown is the decision region plot for the C-SVM Classifier using Linear Kernel



(a) Decision Region Plots of C-SVM Classifier

Figure 3: Performance plots for Overlapping classes

2.2.2 Confusion Matrix and Average Classification Accuracy



(a) Confusion Matrix

Figure 4: C-SVM Classifier with Linear Kernel

Classification Accuracy

1	1	0.99

Average Accuracy=0.993

2.3 Observations

- For linear separable data, value of C is 0.5 and number of support vectors is 3.
- Same example acts as support vectors for multiple classes when using OVR approach of classification.
- For overlapping classes, value of C is 0.5 is and number of support vectors is 32.

3 Experiment No.2: C-SVM with Polynomial Kernel

3.1 Linearly Separable classes(Data-set 1(a))

Aim was to build SVM classifier for the given data-set of Linearly Separable classes using Polynomial Kernel. Data is bi-variate and the classification task was a three class problem. A C-SVM Model is built for the given data-set using libsvm tool and examples in test data-set are classified using the trained model. One C-SVM model is built for each class and O-V-R classification is performed

3.1.1 Kernel Gram Matrix

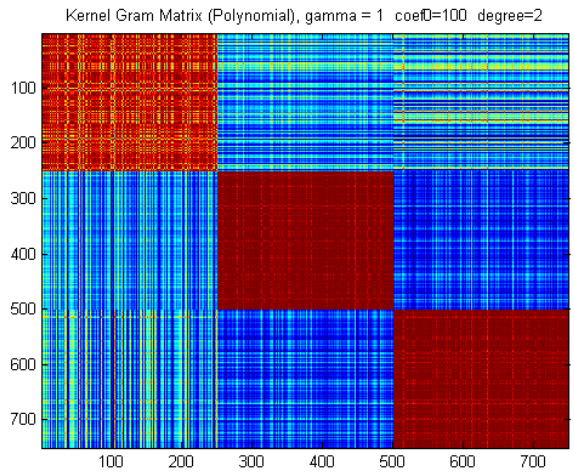
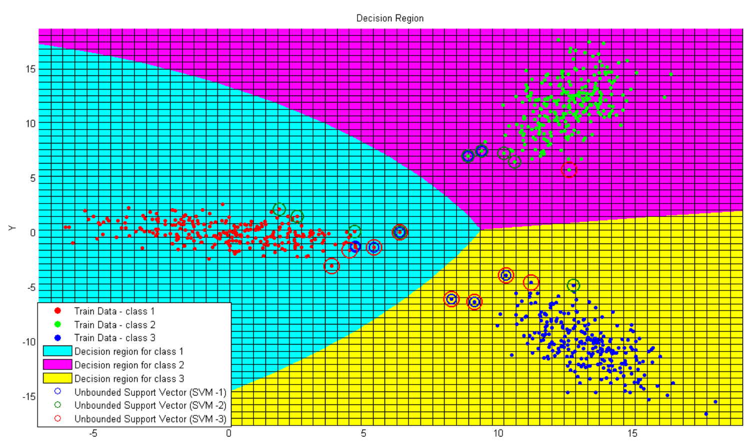


Figure 5: Kernel Gram Matrix

3.1.2 Decision Region Plots

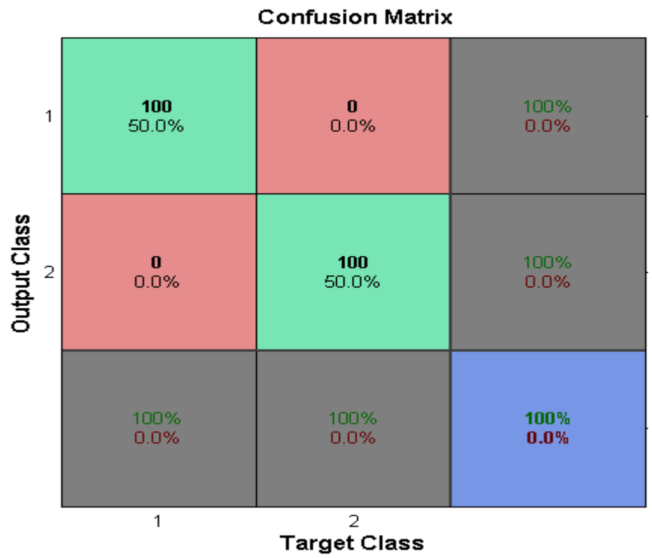
Below is the decision region plot for the Linearly Separable classes.



(a) C-SVM with Polynomial Kernel

Figure 6: Decision Region Plots for Linearly Separable classes

3.1.3 Confusion Matrix and Average Classification Accuracy



(a) Confusion Matrix

Figure 7: C-SVM Classifier using Polynomial Kernel

Classification Accuracy

1	1	0.99

Average Accuracy=0.993

3.2 Non-linearly Separable classes(Data-set 1(b))

Aim was to build C-SVM Classifier with Polynomial Kernel for the given data-set of Non-linearly Separable Classes. Data is bi-variate and the classification task was a two class problem. Model is built for each class to classify a given element to any of the three classes and to find the Classification Accuracy, Confusion Matrix and Decision Region Plot.

3.2.1 Kernel Gram Matrix

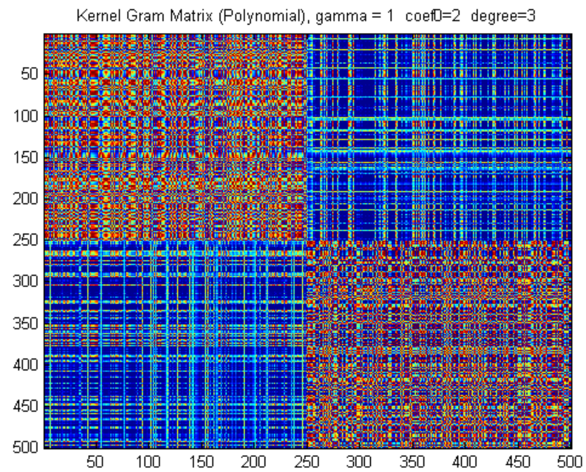
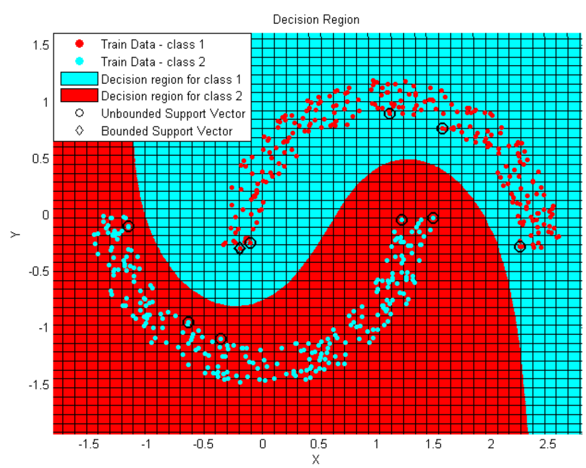


Figure 8: Kernel Gram Matrix

3.2.2 Decision Region Plots

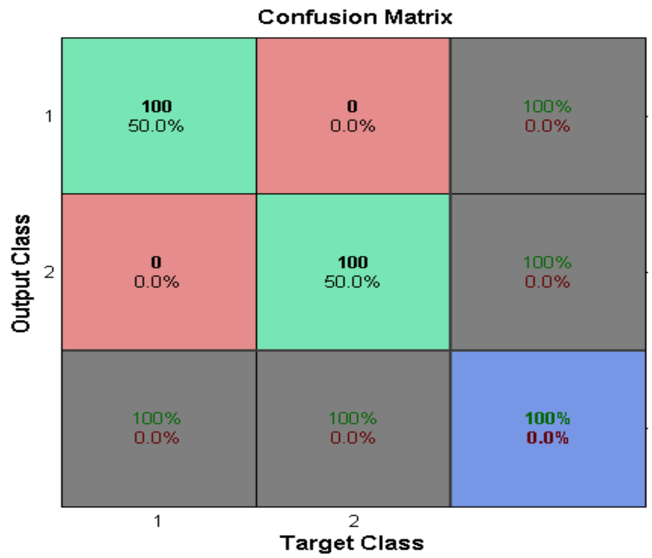
Below is the decision region plot for the Non-linearly Separable classes.



(a) C-SVM with Polynomial Kernel

Figure 9: Decision Region Plots for Non-linearly Separable classes

3.2.3 Confusion Matrix and Average Classification Accuracy



(a) Confusion Matrix

Figure 10: C-SVM Classifier using Polynomial Kernel for Non linearly Separable classes

Classification Accuracy

1	1	1

Average Accuracy=1

3.3 Overlapping classes(Data-set 1(c))

Aim was to build C-SVM Classifier with Polynomial Kernel for the given data-set of Non-linearly Separable Classes. Data is bi-variate and the classification task was a three class problem. Model is built for each class to classify a given element to any of the three classes and to find the Classification Accuracy, Confusion Matrix and Decision Region Plot.

3.3.1 Kernel Gram Matrix

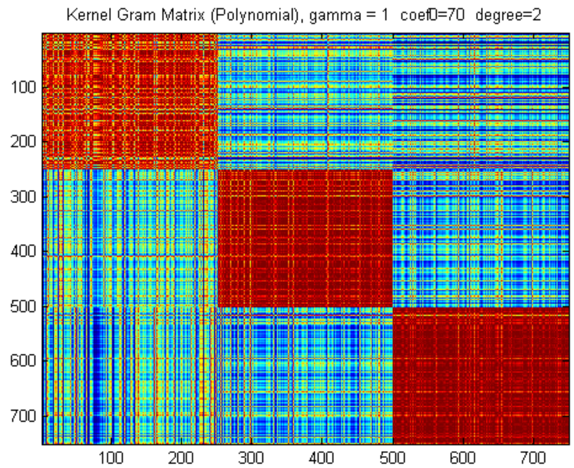
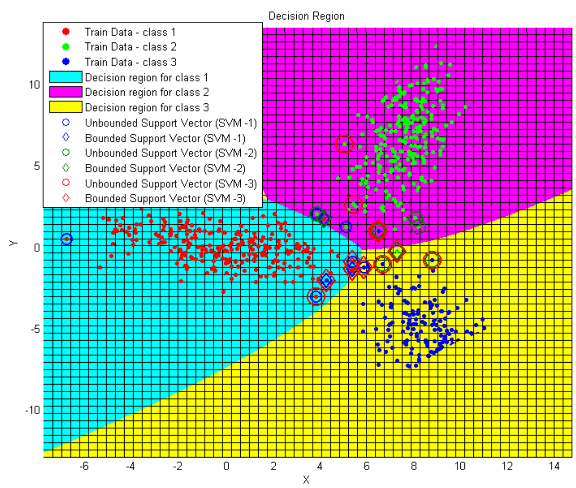


Figure 11: Kernel Gram Matrix

3.3.2 Decision Region Plots

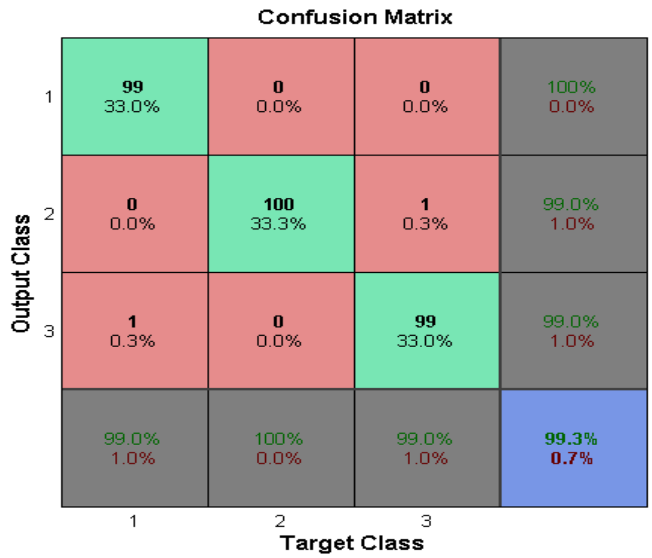
Below is the decision region plot for the Overlapping classes.



(a) C-SVM with Polynomial Kernel

Figure 12: Decision Region Plots for Overlapping classes

3.3.3 Confusion Matrix and Average Classification Accuracy



(a) Confusion Matrix

Figure 13: C-SVM Classifier using Polynomial Kernel for Overlapping classes

Classification Accuracy

1	1	0.99

Average Accuracy=0.993

3.4 Observations

- For linear separable data using Polynomial Kernel(a=100 b=1 ,degree=2), value of C is 30 and number of support vectors is 6.
- For Nonlinearly separable data,(a=2 b=1 ,degree=3), value of C is 0.5 and number of support vectors is 10.
- For overlapping data,(a=70 b=1 ,degree=2), value of C is 30 and number of support vectors is 9.

4 Experiment No.2: C-SVM with Gaussian Kernel

4.1 Linearly Separable classes(Data-set 1(a))

Aim was to build C-SVM classifier for the given data-set of Linearly Separable classes using Gaussian Kernel. Data is bi-variate and the classification task was a three class problem. The SVM Model is built for the given data-set and examples in test data-set are classified using the trained model. One SVM model is built for each class and O-V-R classification is performed

4.1.1 Kernel Gram Matrix

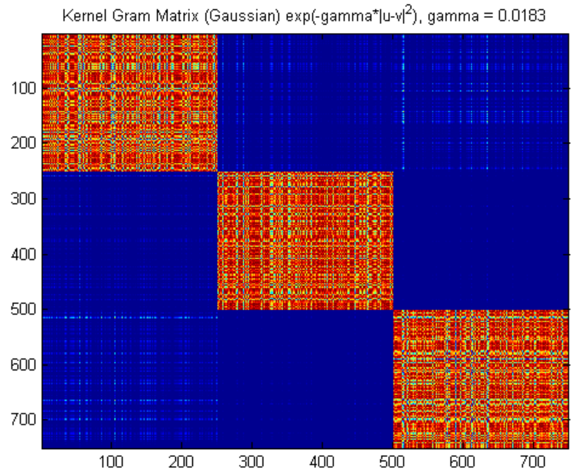
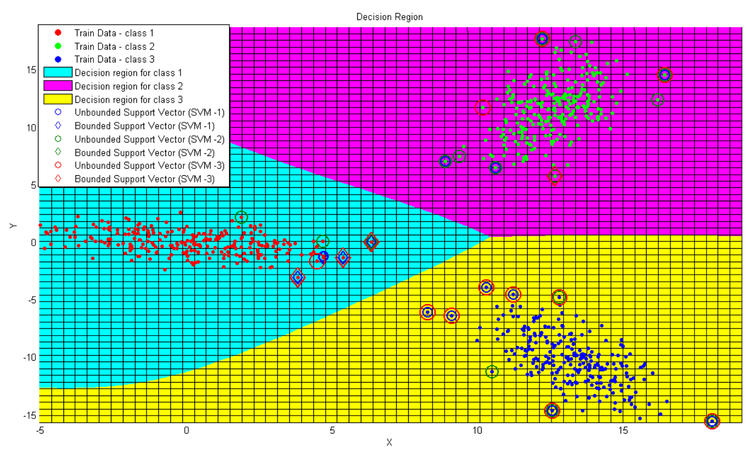


Figure 14: Kernel Gram Matrix

4.1.2 Decision Region Plots

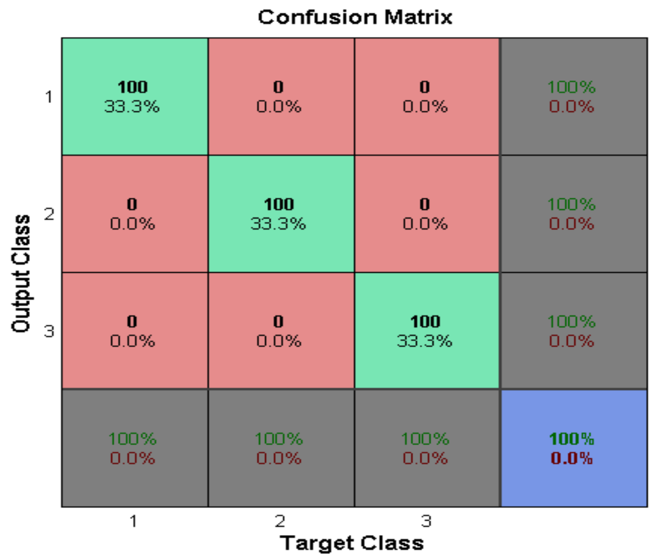
Below is the decision region plot for the Linearly Separable classes.



(a) C-SVM with Gaussian Kernel

Figure 15: Decision Region Plots for Linearly Separable classes

4.1.3 Confusion Matrix and Average Classification Accuracy



(a) Confusion Matrix

Figure 16: C-SVM Classifier using Gaussian Kernel for Linearly Separable classes

Classification Accuracy

1	1	1

Average Accuracy=1

4.2 Non-linearly Separable classes(Data-set 1(b))

Aim was to build C-SVM Classifier with Gaussian Kernel for the given data-set of Non-linearly Separable Classes. Data is bi-variate and the classification task was a two class problem. Model is built for each class to classify a given element to any of the three classes and to find the Classification Accuracy, Confusion Matrix and Decision Region Plot.

4.2.1 Kernel Gram Matrix

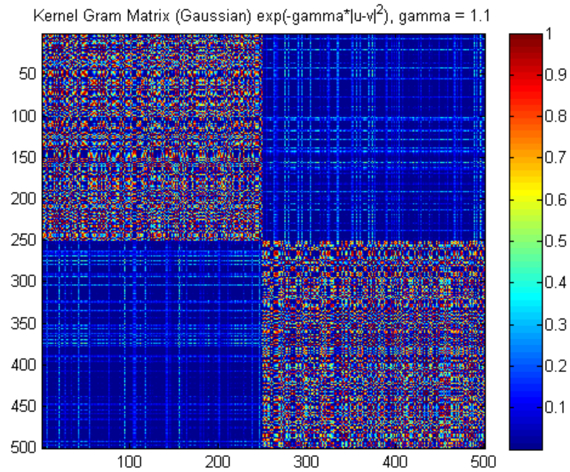
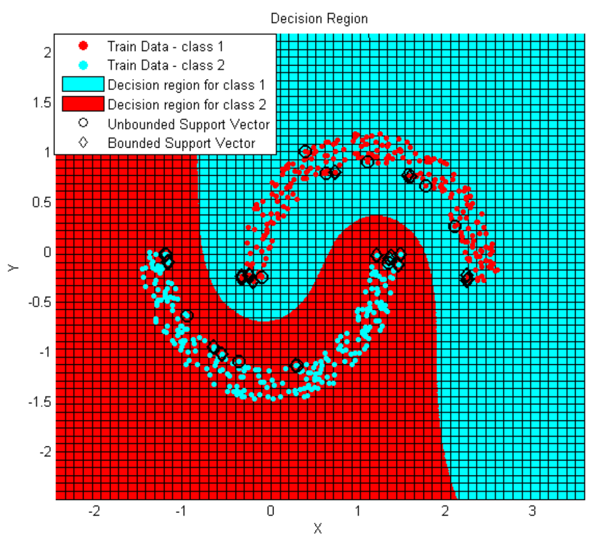


Figure 17: Kernel Gram Matrix

4.2.2 Decision Region Plots

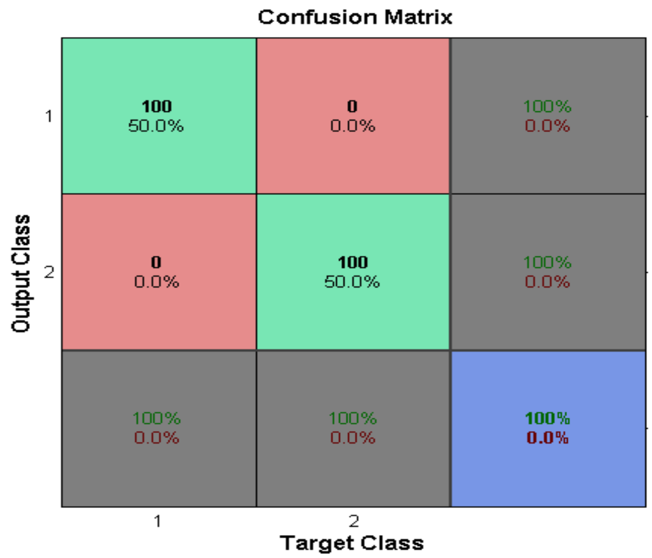
Below is the decision region plot for the Non-linearly Separable classes.



(a) C-SVM with Gaussian Kernel

Figure 18: Decision Region Plots for Non-linearly Separable classes

4.2.3 Confusion Matrix and Average Classification Accuracy



(a) Confusion Matrix

Figure 19: C-SVM Classifier using Gaussian Kernel for Non-linearly Separable classes

Classification Accuracy

1	1	1

Average Accuracy=1

4.3 Overlapping classes(Data-set 1(c))

Aim was to build C-SVM Classifier with Gaussian Kernel for the given data-set of Overlapping Classes. Data is bi-variate and the classification task was a three class problem. Model is built for each class to classify a given element to any of the three classes and to find the Classification Accuracy, Confusion Matrix and Decision Region Plot.

4.3.1 Kernel Gram Matrix

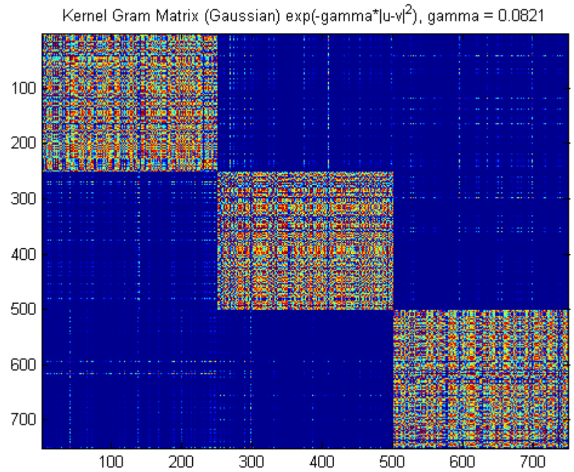
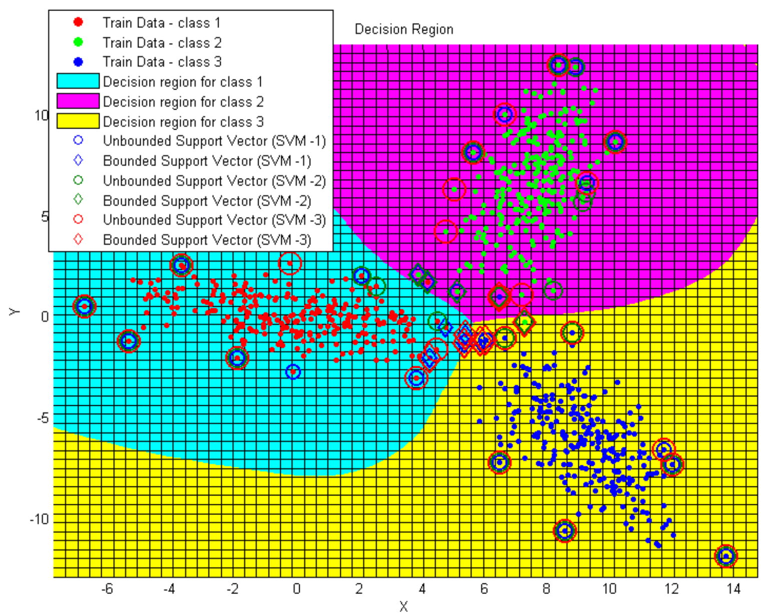


Figure 20: Kernel Gram Matrix

4.3.2 Decision Region Plots

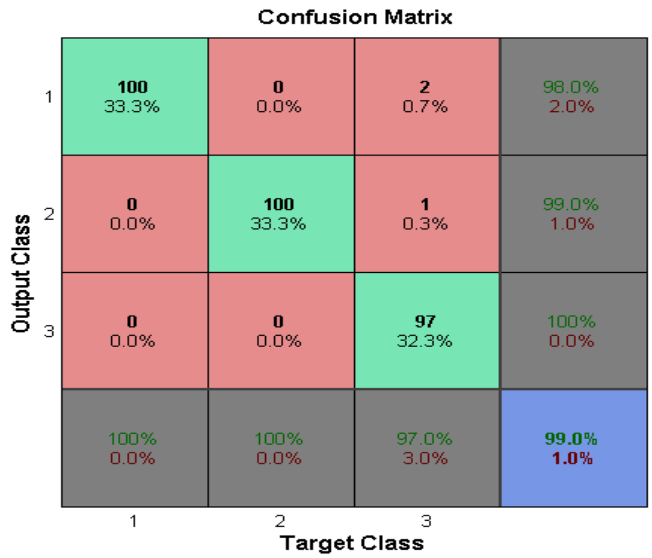
Below is the decision region plot for the Overlapping classes.



(a) C-SVM with Gaussian Kernel

Figure 21: Decision Region Plots for Overlapping classes

4.3.3 Confusion Matrix and Average Classification Accuracy



(a) Confusion Matrix

Figure 22: C-SVM Classifier using Gaussian Kernel for Overlapping classes

Classification Accuracy

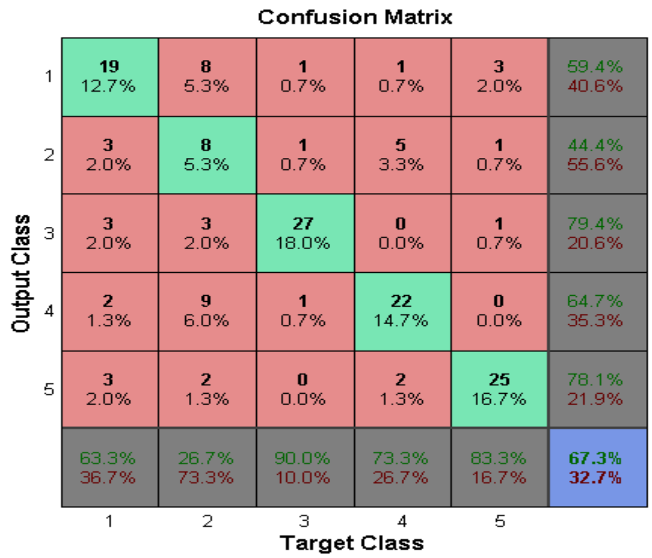
1	1	0.97

Average Accuracy=0.99

4.4 Image Data(Data-set 2)

Aim was to build classifier for the given data-set of Classes. Data is multivariate with 48 dimensions and the classification task was a five class problem. Image data-set is splitted in ratio 70:10:20 for Train:Validation:Test . C-SVM Model using Gaussian Kernel is built to classify a given element to any of the five classes and to find the Classification Accuracy, Confusion Matrix.

4.4.1 Confusion Matrix and Average Classification Accuracy



(a) Confusion Matrix

Figure 23: C-SVM Classifier using Gaussian Kernel for Image Dataset

Classification Accuracy

0.633	0.267	0.90	0.733	0.833

Average Accuracy=0.673

4.5 Observations

- For linear separable data using Gaussian Kernel(gamma=0.0821), value of C is 0.5 and number of support vectors is 25.
- For Nonlinearly separable data,(gamma=1.1), value of C is 0.5 and number of support vectors is 31.
- For overlapping data,(gamma=0.0183), value of C is 8 and number of support vectors is 38.
- For image data,(gamma=exp(-22)), value of C is 50 and number of support vectors is 135.

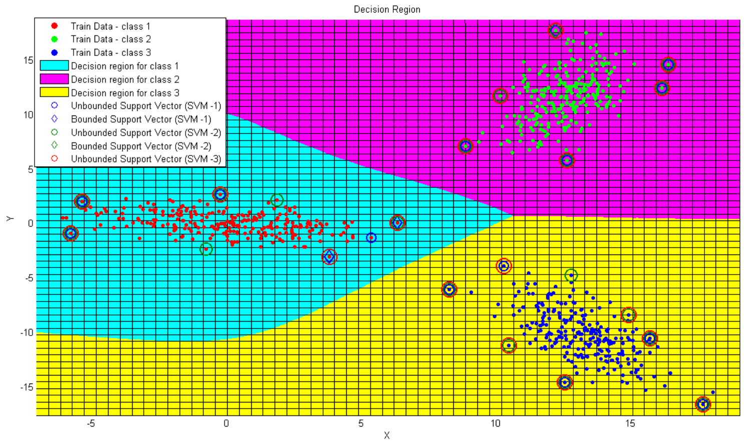
5 Experiment No.3: ν-SVM with Gaussian Kernel

5.1 Linearly Separable classes(Data-set 1(a))

Aim was to build ν-SVM classifier for the given data-set of Linearly Separable classes using Gaussian Kernel. Data is bi-variate and the classification task was a three class problem. The SVM Model is built for the given data-set and examples in test data-set are classified using the trained model. One SVM model is built for each class and O-V-R classification is performed.

5.1.1 Decision Region Plots

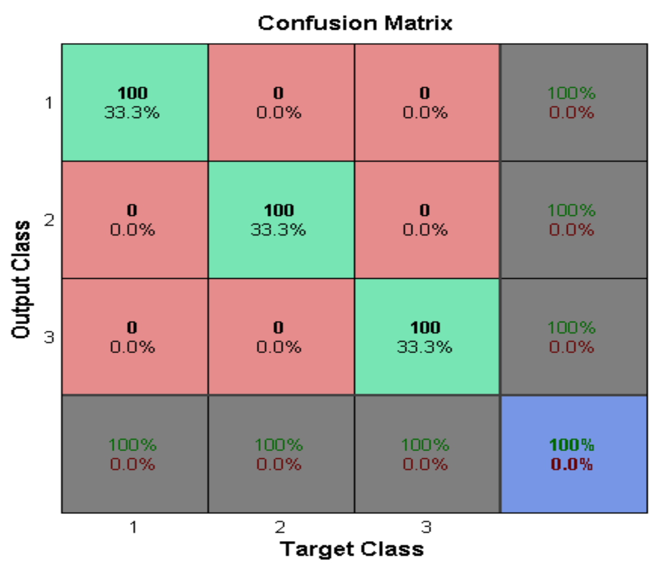
Below is the decision region plot for the Linearly Separable classes.



(a) ν-SVM with Gaussian Kernel

Figure 24: Decision Region Plots for Linearly Separable classes

5.1.2 Confusion Matrix and Average Classification Accuracy



(a) Confusion Matrix

Figure 25: ν-SVM Classifier using Gaussian Kernel for Linearly Separable classes

Classification Accuracy

1	1	1

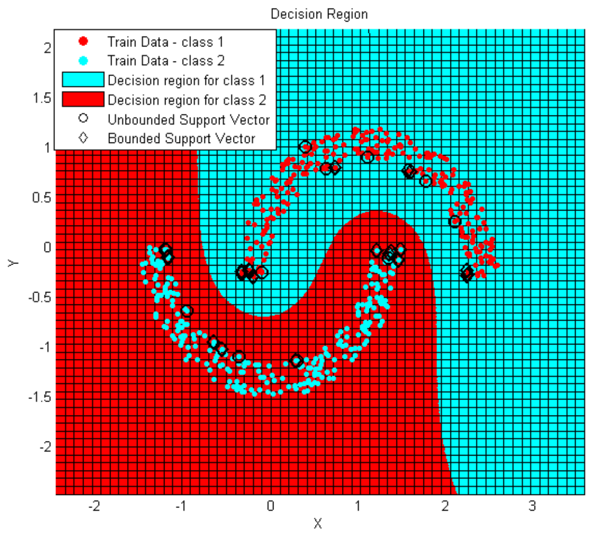
Average Accuracy=1

5.2 Non-linearly Separable classes(Data-set 1(b))

Aim was to build ν-SVM Classifier with Gaussian Kernel for the given data-set of Non-linearly Separable Classes. Data is bi-variate and the classification task was a two class problem. Model is built for each class to classify a given element to any of the three classes and to find the Classification Accuracy, Confusion Matrix and Decision Region Plot.

5.2.1 Decision Region Plots

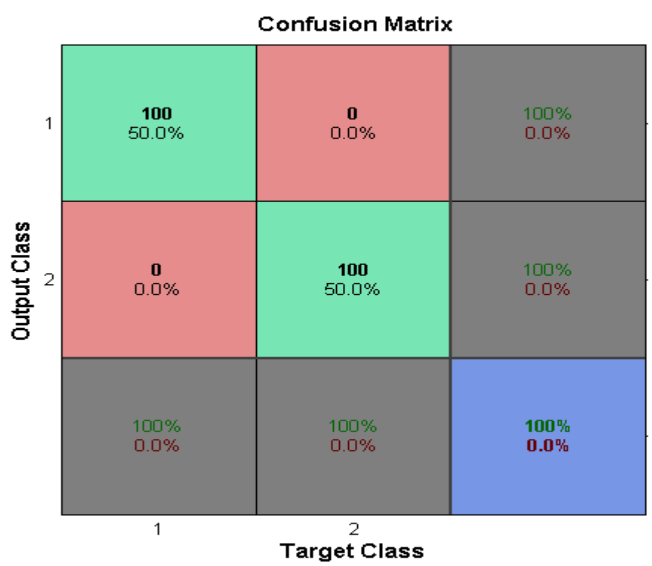
Below is the decision region plot for the Non-linearly Separable classes.



(a) ν-SVM with Gaussian Kernel

Figure 26: Decision Region Plots for Non-linearly Separable classes

5.2.2 Confusion Matrix and Average Classification Accuracy



(a) Confusion Matrix

Figure 27: ν-SVM Classifier using Gaussian Kernel for Non-linearly Separable classes

Classification Accuracy

1	1	1

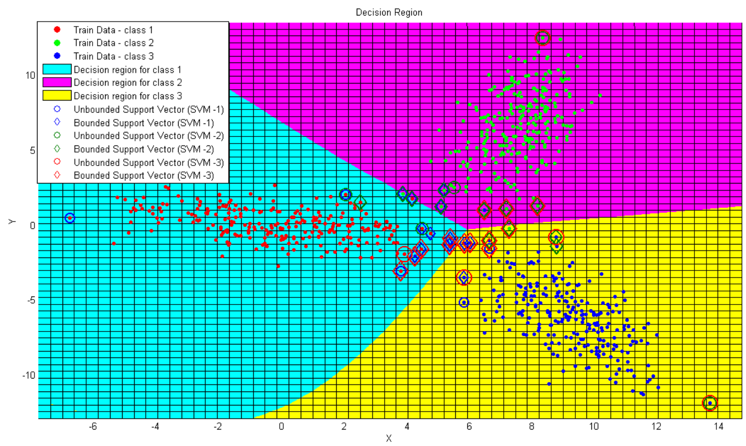
Average Accuracy=1

5.3 Overlapping classes(Data-set 1(c))

Aim was to build ν-SVM Classifier with Gaussian Kernel for the given data-set of Overlapping Classes. Data is bi-variate and the classification task was a three class problem. Model is built for each class to classify a given element to any of the three classes and to find the Classification Accuracy, Confusion Matrix and Decision Region Plot.

5.3.1 Decision Region Plots

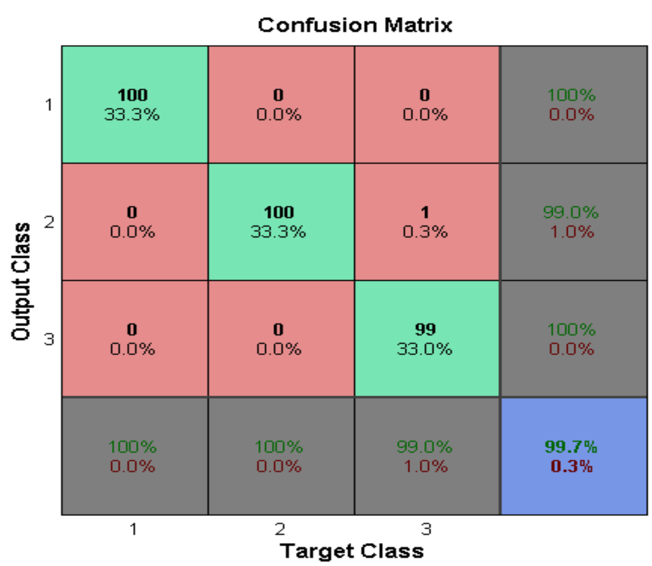
Below is the decision region plot for the Overlapping classes.



(a) ν-SVM with Gaussian Kernel

Figure 28: Decision Region Plots for Overlapping classes

5.3.2 Confusion Matrix and Average Classification Accuracy



(a) Confusion Matrix

Figure 29: ν-SVM Classifier using Gaussian Kernel for Overlapping classes

Classification Accuracy

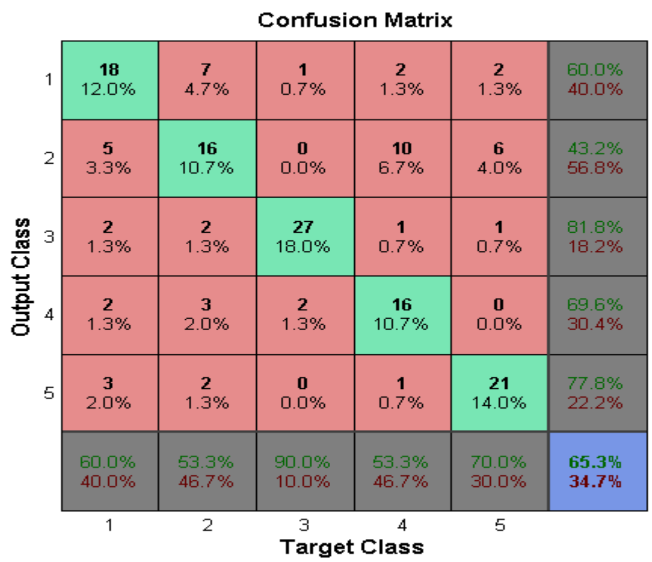
1	1	0.99

Average Accuracy=0.997

5.4 Image Data(Data-set 2)

Aim was to build classifier for the given data-set of Classes. Data is multivariate with 48 dimensions and the classification task was a five class problem. Image data-set is splitted in ratio 70:10:20 for Train:Validation:Test . ν-SVM Model using Gaussian Kernel is built to classify a given element to any of the five classes and to find the Classification Accuracy, Confusion Matrix.

5.4.1 Confusion Matrix and Average Classification Accuracy



(a) Confusion Matrix

Figure 30: ν-SVM Classifier using Gaussian Kernel for Image Dataset

Classification Accuracy

0.60	0.533	0.90	0.533	0.70

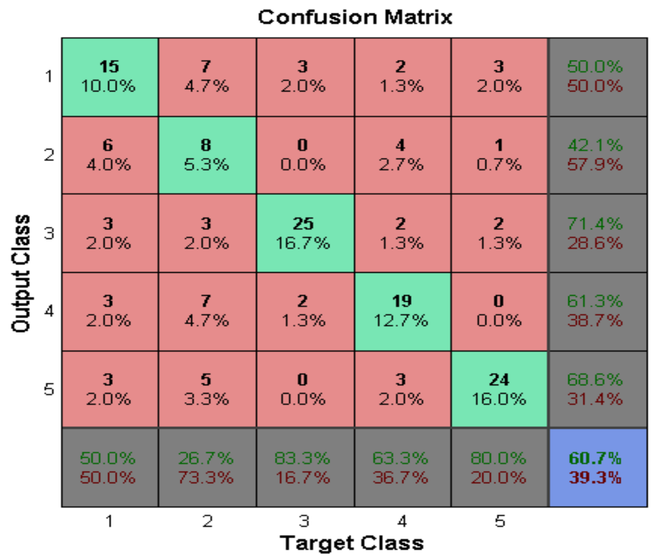
Average Accuracy=0.653

5.5 C-SVM with Histogram Intersection Kernel (Image Data)

Here, histogram intersection is used as the kernel for Support Vector Machines. Color histogram intersection $Kint$ measures the similarity between two color histograms. If A and B are the histograms of two images A_img and B_img, and both contains m number of bins, then the histogram intersection is given by,

$K_{int}(A,B) = \sum min(a_i, b_i)$, where a_i and b_i are i'th bins.

5.5.1 Confusion Matrix and Average Classification Accuracy



(a) Confusion Matrix

Figure 31: C-SVM with Histogram Intersection Kernel

Classification Accuracy

0.50	0.267	0.833	0.633	0.80

Average Accuracy=0.607

5.6 Observations

For ν-SVM using Gaussian Kernel:

- For linear separable data using Gaussian Kernel(gamma=exp(-3.5)), value of $\nu$is 0.05 and number of support vectors is 87.
- For Nonlinearly separable data,(gamma=1.1), value ofν is 0.1 and number of support vectors is 56.
- For overlapping data,(gamma=,exp(-2.7)), value of ν is 0.2 and number of support vectors is 91
- For image data,(gamma=,exp(-3.5)), value of ν is 0.2 and number of support vectors is 138.

For C-SVM using Histogram Kernel:

- The value of C is 16 and Number of Support Vectors is 185.

Part II
Regression Tasks

6 Data-sets

6.0.1 Data-set 1: 1-dimensional (Uni-variate) input data

6.0.2 Data-set 2: 2-dimensional (Bi-variate) input data

6.0.3 Data-set 3: Multivariate input data

Models

7 Experiment No.1: ϵ-SVR using Gaussian Kernel

In this Experiment, the aim was to build regression model based on Support Vector Machines.ϵ-SVM with Gaussian Kernel is used here to build the regression model for different tpes of dataset.

7.1 Uni-variate Data (Data-set 1)

Aim was to build ϵ-SVM regression model for the given data-set. Data is uni-variate and the regression task was to approximate the underlying function for the test data-set using the trained model

7.1.1 Plots of ϵ-tube and output and approximated function for ϵ-SVR

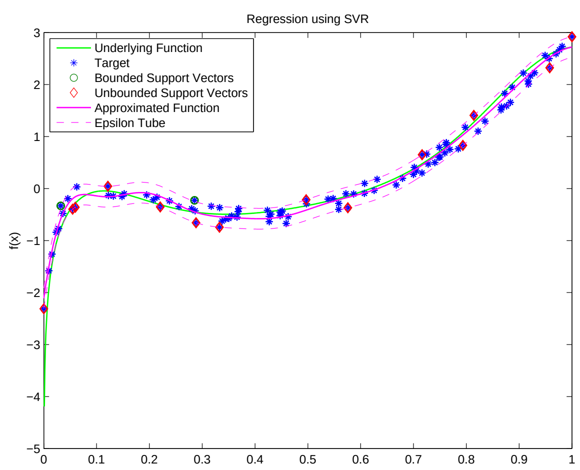
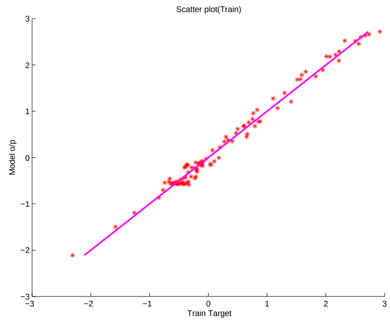
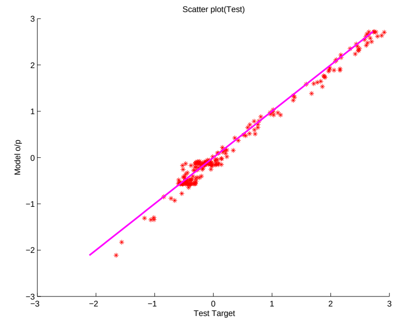


Figure 32: Univariate Data-epsilon tube and underlying function

7.1.2 Scatter Plot of Target versus Model Output



(b) Target vs Model o/p for train data



(d) Target vs Model o/p for test data

Figure 33: Scatter Plot of Target vs Model output

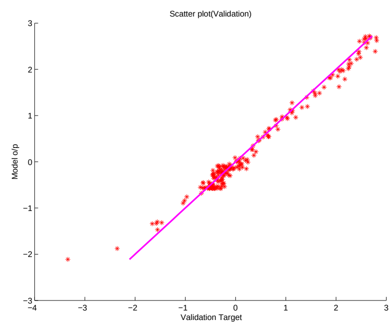
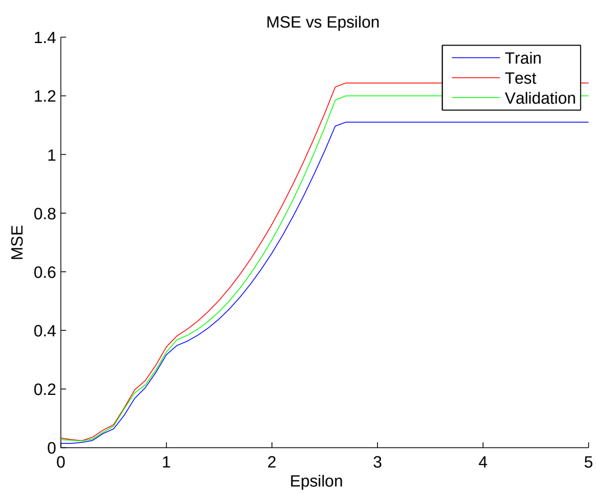


Figure 34: Scatter Plot of Target vs Model o/p for validation data

7.1.3 Plots of MSE Error vs ϵ on Training data, Validation data and Test data.

Following are the plots of Mean Squared Error for different values of ϵ.



(a) Plot of MSE vs epsilon

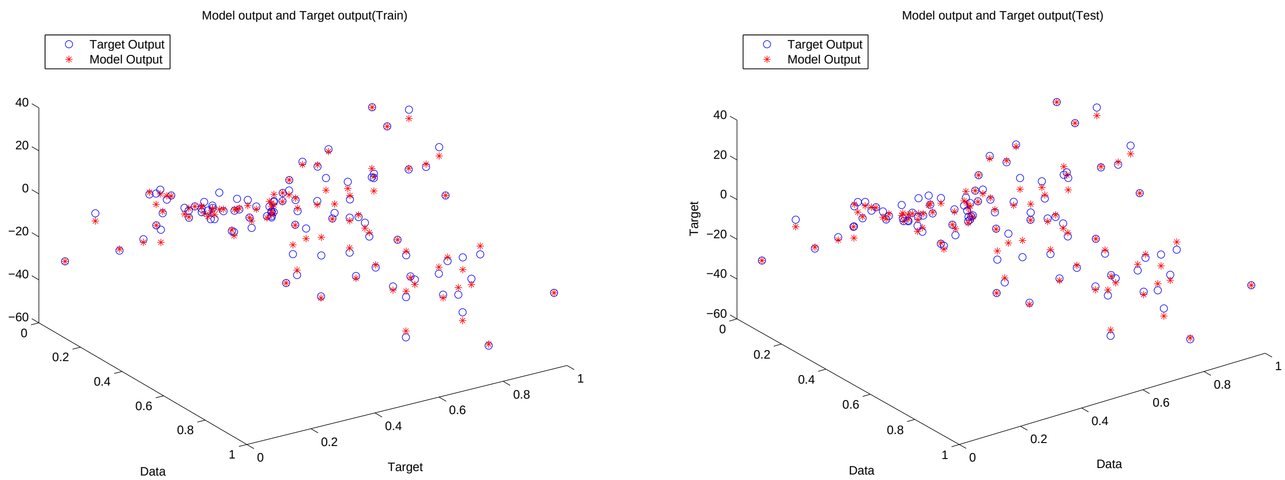
Figure 35: epsilon-SVR

7.2 Bi-variate Data (Data-set 2)

Aim was to build ϵ-SVM regression model for the given data-set. Data is bi-variate and the regression task was to approximate the underlying function for the test data-set using the trained model.

7.2.1 Plot of model output and target output

Below are the plots of model output and target output for training data, validation data and test data using ϵ-SVR Model.



(b) Train data

(d) Test data

Figure 36: Realization of model output and target output

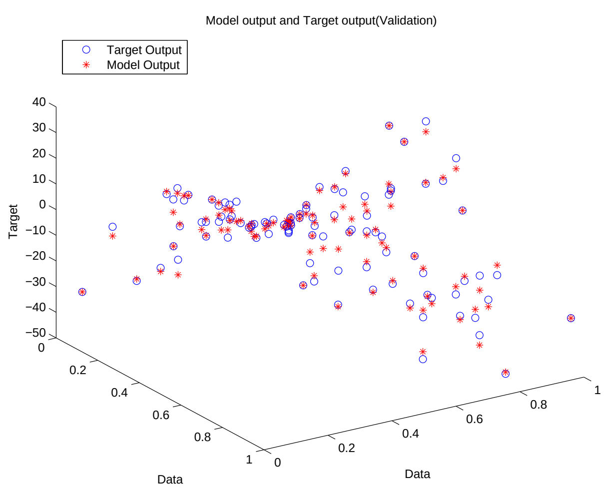
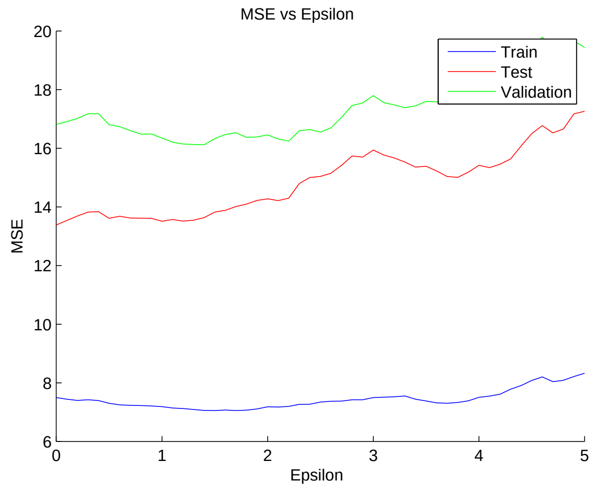


Figure 37: Validation data-Realization of model and target output

7.2.2 Plots of MSE Error vs ϵ on Training data, Validation data and Test data.

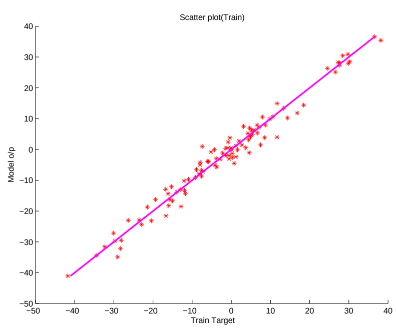
Following are the plots of Mean Squared Error for different values of ϵ.



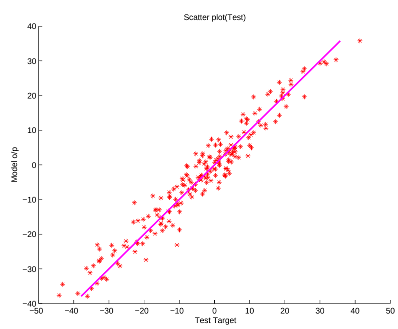
(a) MSE vs epsilon

Figure 38: epsilon-SVR for Bivariate Data

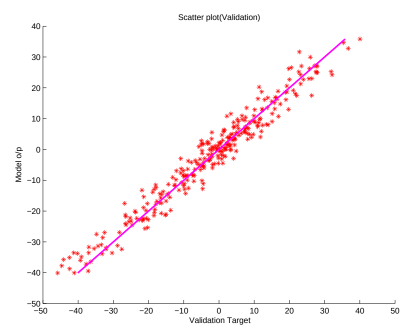
7.2.3 Scatter Plot of Target versus Model Output



(b) Target vs Model o/p for train data



(d) Target vs Model o/p for test data



(f) Target vs Model o/p for validation data

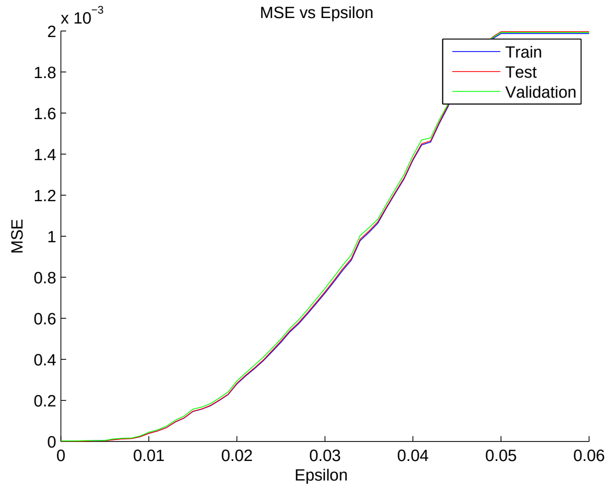
Figure 39: Target vs Model output

7.3 Multivariate Data (Data-set 3)

Aim was to build ϵ-SVM regression model for the given data-set. Data is multi-variate and the regression task was to approximate the underlying function for the test data-set using the trained model.

7.3.1 Plots of MSE Error vs ϵ on Training data, Validation data and Test data.

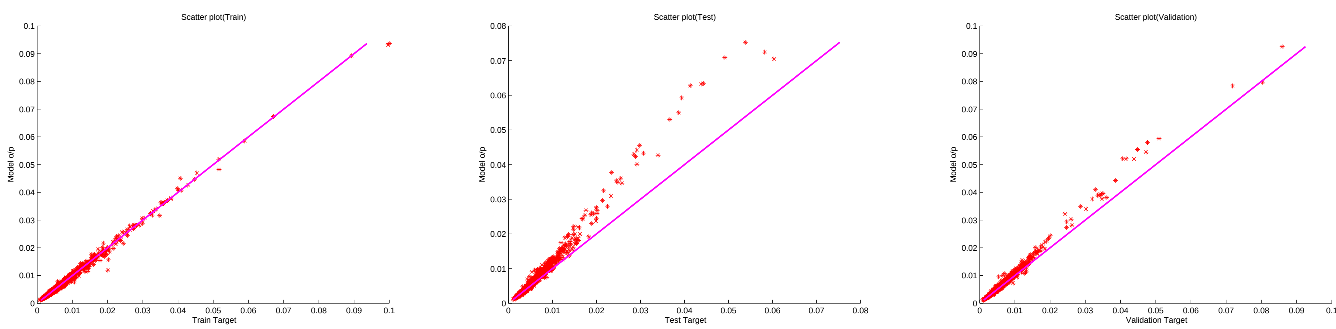
Following are the plots of Mean Squared Error for different values of ϵ.



(a) MSE vs epsilon

Figure 40: epsilon-SVR on Multivariate Data

7.3.2 Scatter Plot of Target versus Model Output



(b) Target vs Model o/p for train data (d) Target vs Model o/p for test data (f) Target vs Model o/p for validation data

Figure 41: Target vs Model output

8 Experiment No.2: ν-SVR using Gaussian Kernel

In this Experiment, the aim was to build regression model based on Support Vector Machines.ϵ-SVM with Gaussian Kernel is used here to build the regression model for different tpes of dataset.

8.1 Uni-variate Data (Data-set 1)

Aim was to build ν-SVM regression model for the given data-set. Data is uni-variate and the regression task was to approximate the underlying function for the test data-set using the trained model

8.1.1 Plot of ν-tube and Plot of target output and approximated function forν-SVR

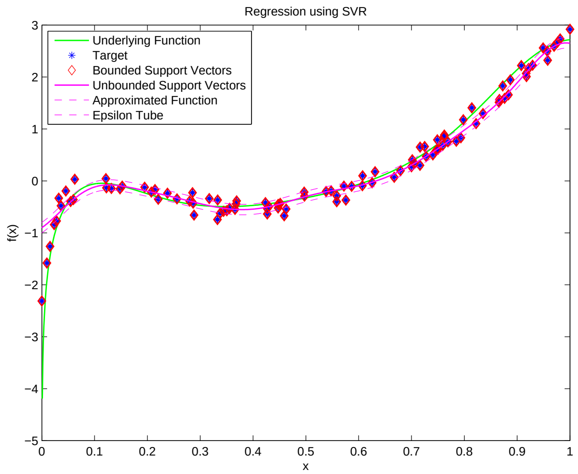
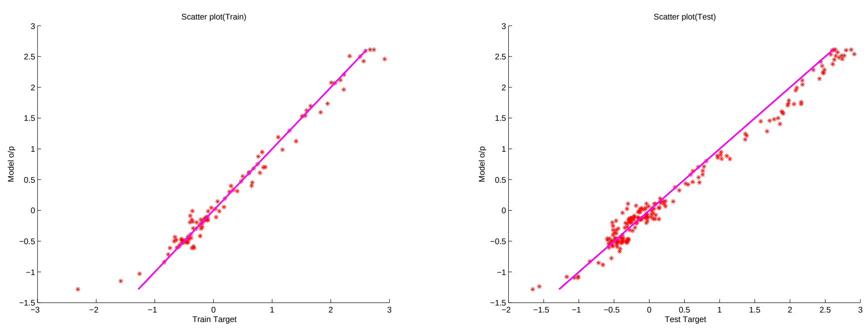


Figure 42: ν-SVR on univariate data

8.1.2 Scatter Plot of Target versus Model Output



(b) Target vs Model o/p for train data (d) Target vs Model o/p for test data

Figure 43: Scatter Plot of Target vs Model output

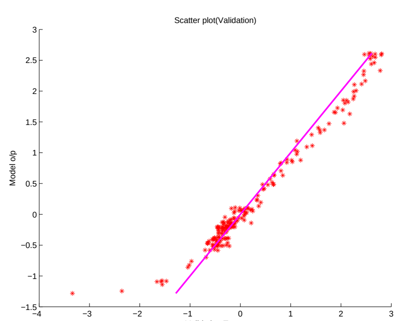
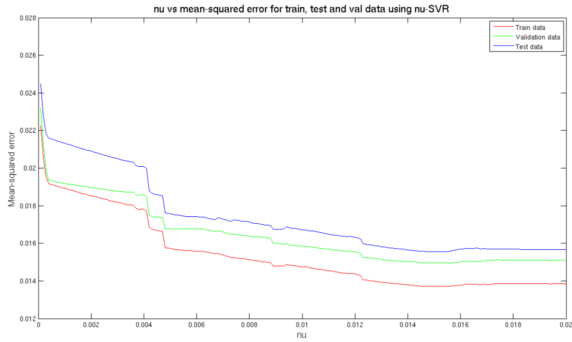


Figure 44: Scatter Plot of Target vs Model o/p for validation data

8.1.3 Plots of MSE Error vs ν- on Training data, Validation data and Test data.

Following are the plots of Mean Squared Error for different values of ν.



(a) Plot of MSE vs nu

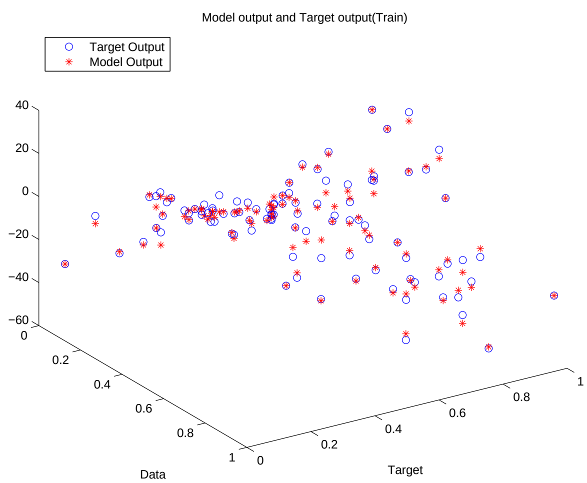
Figure 45: nu-SVR

8.2 Bi-variate Data (Data-set 2)

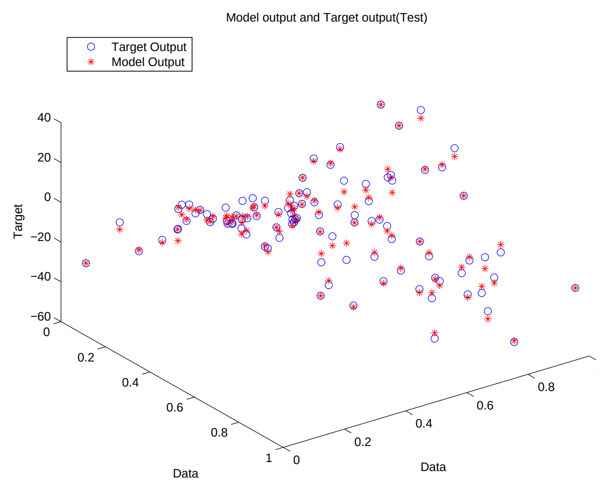
Aim was to build ν-SVM regression model for the given data-set. Data is bi-variate and the regression task was to approximate the underlying function for the test data-set using the trained model.

8.2.1 Plot of model output and target output

Below are the plots of model output and target output for training data, validation data and test data using ν-SVR Model.



(b) Train data



(d) Test data

Figure 46: Realization of model output and target output

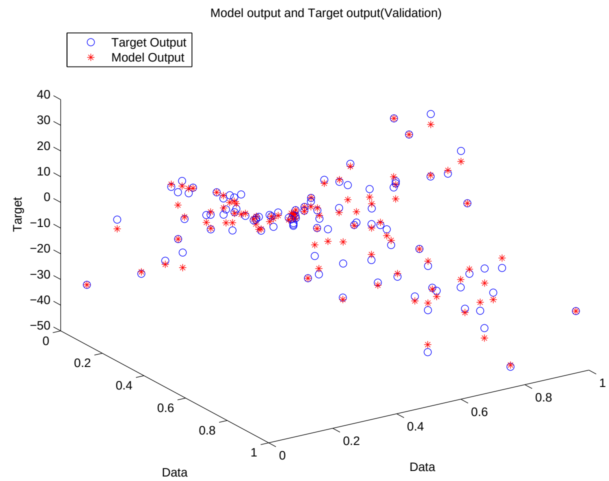
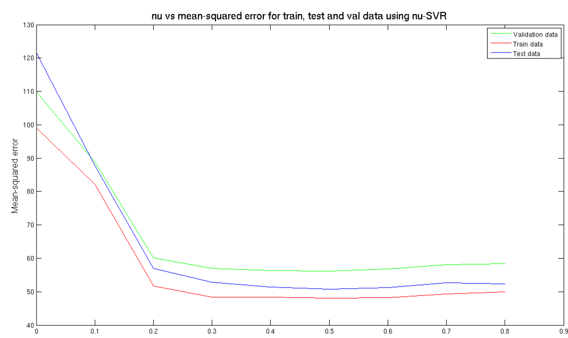


Figure 47: Validation data-Realization of model and target output

8.2.2 Plots of MSE Error vs ϵ on Training data, Validation data and Test data.

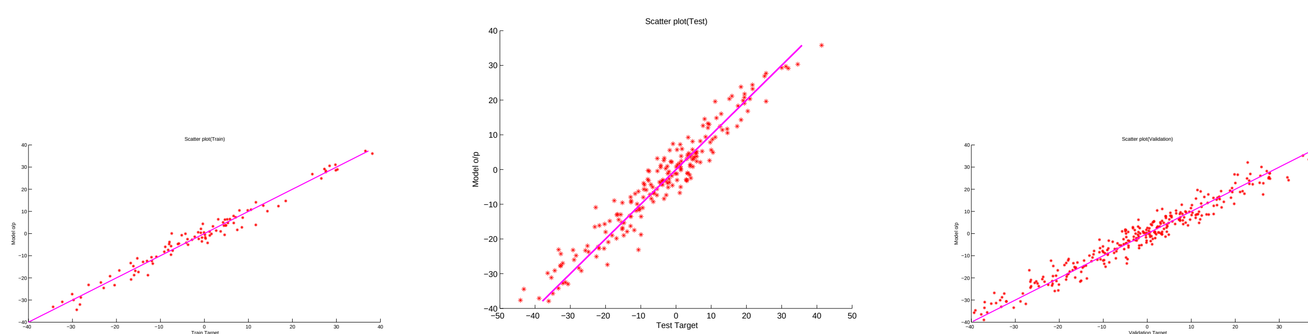
Following are the plots of Mean Squared Error for different values of ν.



(a) MSE vs nu

Figure 48: nu-SVR for Bi-variate Data

8.2.3 Scatter Plot of Target versus Model Output



(b) Target vs Model o/p for train data (d) Target vs Model o/p for test data (f) Target vs Model o/p for validation data

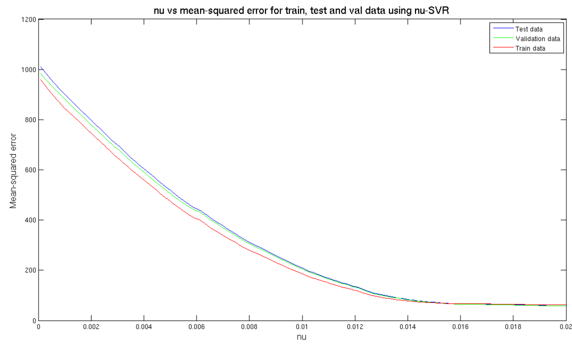
Figure 49: Target vs Model output

8.3 Multivariate Data (Data-set 3)

Aim was to build ν-SVR model for the given data-set. Data is multivariate and the regression task was to approximate the underlying function for the test data-set using the trained model.

8.3.1 Plots of MSE Error vs ν on Training data, Validation data and Test data.

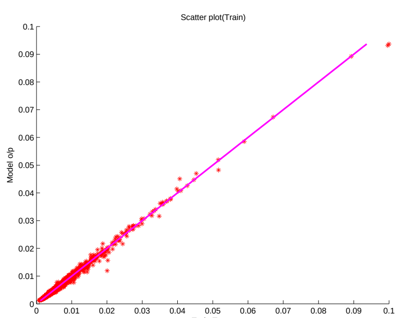
Following are the plots of Mean Squared Error for different values of ν.



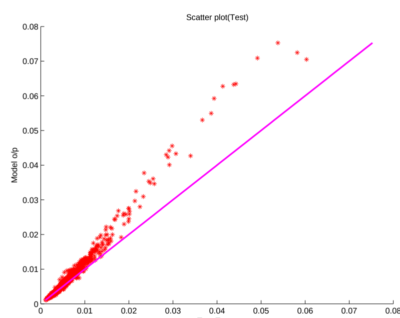
(a) MSE vs nu

Figure 50: nu-SVR on Multivariate Data

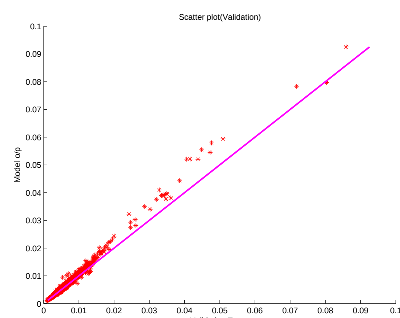
8.3.2 Scatter Plot of Target versus Model Output



(b) Target vs Model o/p for train data



(d) Target vs Model o/p for test data



(f) Target vs Model o/p for validation data

Figure 51: Target vs Model output

8.4 Observations

- For univariate data, value of epsilon is 0.2 and for nu SVR the value of nu is 0.1.
- Number of support vectors is 16 and 10 for epsilon and nu SVRs respectively.
- For bivariate data, value of epsilon is 0.001 and for nu SVR the value of nu is 1.
- Number of support vectors is 100 and 30 for epsilon and nu SVRs respectively.
- For multivariate data, value of epsilon is 0.2 and for nu SVM the value of nu is 0.01.
- Number of support vectors is 120 and 75 for epsilon and nu SVRs respectively.

Part III
Novelty detection tasks

9 Data-sets

9.0.1 Data-set 1: 2-dimensional input data of one of the classes in the overlapping data-set

9.0.2 Data-set 2: Multivariate input data

Models

10 Experiment No.1: C-SVDD using Gaussian Kernel

10.1 Data-set 1 (Overlapping Data)

Aim was to build C-SVDD model for one of the classes of overlapping data-set. This was built for Novelty detection of the selected class with respect to other classes. Data is bi-variate and the task was to detect the normal and abnormal classes from the test data.

10.1.1 Plot of Bounded and unbounded Support Vectors.

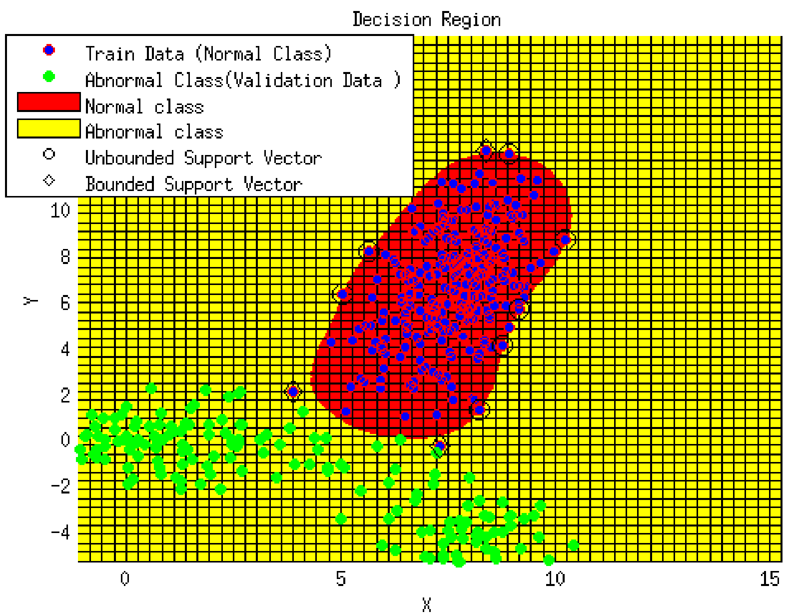


Figure 52: Overlapping data-Bounded and unbounded Support Vectors

10.1.2 Percentage of TPs.

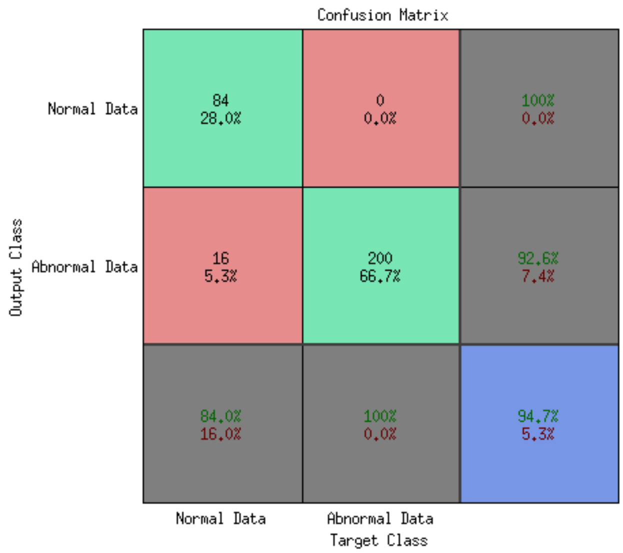


Figure 53: Overlapping data data-Percentage of TPs and FAs

10.2 Data-set 2 (Network Data)

Aim was to build C-SVDD model for given network data. This was built for Novelty detection of the selected class labeled 'normal' with respect to other classes. These were one of the features of the network data, which was actually a 42 dimensional data.Task was to detect the normal and abnormal classes from the test data.

10.2.1 Percentage of TPs and FAs.

Output/Target	Normal Data	Abnormal Data
Normal Data	98.6%	1.39J%
Abnormal Data	22.71%	77.29%

Table 1: Confusion Matrix for C-SVDD Novelty Detection

Accuracy = 81.4455%

11 Experiment No.2: ν-SVDD using Gaussian Kernel

Aim was to build ν-SVDD model for one of the classes of overlapping dataset. This was built for Novelty detection of the selected class with respect to other classes. Data is bi-variate and the task was to detect the normal and abnormal classes from the test data.

11.0.2 Plot of Bounded and unbounded Support Vectors.

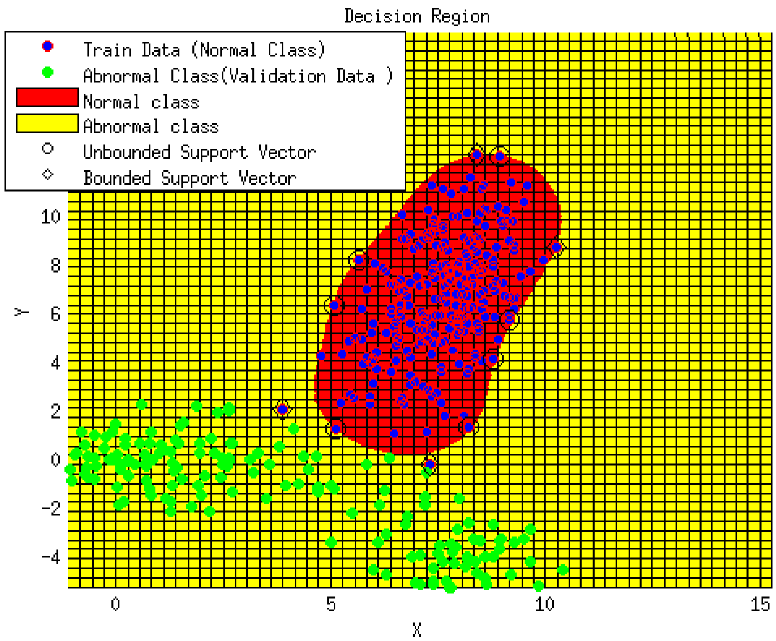


Figure 54: Overlapping data-Bounded and unbounded Support Vectors

11.0.3 Percentage of TPs and FAs.

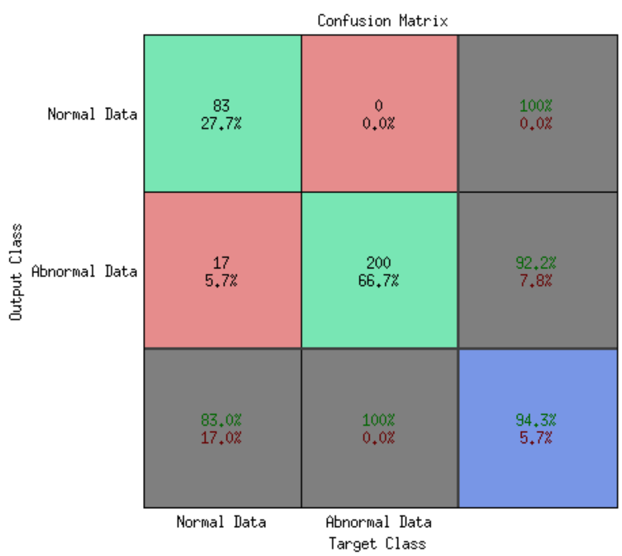


Figure 55: Overlapping data data-Percentage of TPs and FAs

11.1 Data-set 2 (Network Data)

Aim was to build ν-SVDD model for given network data. This was built for Novelty detection of the selected class labeled 'normal' with respect to other classes. These were one of the features of the network data, which was actually a 42 dimensional data.Task was to detect the normal and abnormal classes from the test data.

11.1.1 Percentage of TPs and FAs-Confusion Matrix Representation.

Output/Target	Normal Data	Abnormal Data
Normal Data	98.67%	1.32%
Abnormal Data	22.66%	77.33%

Table 2: Confusion Matrix for ν-SVDD Novelty Detection

Accuracy = 81.47%

11.2 Observations

- For Network data, value of ν for ν-SVDD is 0.0041.Value of gamma is 0.01 for the Gaussian Kernel and number of Support Vectors is 252 of which 35 are bounded.
- For overlapping data, value of ν for ν-SVDD is 0.03.Value of gamma is exp(-2.5) for the Gaussian Kernel and number of Support Vectors is 11 of which 4 are bounded.
- For Network data, value of C for C-SVDD is found to be 0.1.Value of gamma is 0.01 for the Gaussian Kernel and number of Support Vectors is 252 of which 32 are bounded SVs.
- For overlapping data, value of C for C-SVDD is found to be 0.15.Value of gamma is exp(-2.7) for the Gaussian Kernel and number of Support Vectors is 10 out of which 3 are bounded SVs.