CS6011: Kernel Methods for Pattern Analysis
Programming Assignment III

Date: March 24, 2017   **Deadline for submission of report**: Monday, May 08, 2017

Task1 Regression :
    Dataset 1: 1-dimensional input data of curve fitting
    Dataset 2: 2-dimensional (Bivariate) input data
Model : ν-SVR using Gaussian kernel
Presentation of Results :
1. Plots of underlying function, ε-tube, target output and approximated function obtained for Dataset 1. Mark the unbounded and bounded support vectors.
2. Plots of model output and target output for training data for Dataset 2
3. Scatter plots for training data, validation data and test data, for Datasets 1 and 2
4. Comparison of performance of SVR with that of Linear Model for Regression, RBF and MLFFNN for Datasets 1 and 2 including the surfaces realized.

Task2 Novelty detection :
    Dataset 3: 2-dimensional input data of overlapping classes
    Dataset 4: Multivariate input data
Model : ν-SVDD using Gaussian kernel
Presentation of Results :
1. Mark the bounded and unbounded support vectors for Dataset 3 and plot decision regions.
2. Percentage of true positives and false alarms for test data in Datasets 3 and 4.

Task3 Clustering :
    Dataset 5 : 2-dimensional data of nonlinearly separable classes
Methods:
1. K-means clustering on Dataset 5 and Dataset 3
2. Kernel K-means clustering using Gaussian kernel on Dataset 5 and Dataset 3
3. Kernel based soft clustering using Gaussian kernel on Dataset 5 and Dataset 3
Presentation of Results :
1. Decision region plots for each of the methods and for each of the datasets after initialization, 2nd iteration , 5th iteration and after convergence.

Task4 Dimension Reduction:
    Dataset 6: Image data set
Methods:
a. Principal Component Analysis (PCA)
b. Autoassociative neural network (AANN)
c. Kernel PCA
Presentation of Results :
1. Evaluate the reduced dimension representations obtained using the above methods on Dataset 6 by building the ν-SVM classifiers and comparing the classification accuracies.

Task5- Classification and Clustering using kernels for structured data
    Dataset 7: Graph representation of data
Kernel to use : Random Walk Graph Kernel
Methods using a kernel on graphs:
a. ν-SVM classifier
b. Kernel K-means clustering (with number of clusters same as the number of classes)
Presentation of Results :
1. Confusion matrix for classification task
2. Cluster purity for clustering task