**January-May 2015 Semester**

**CS6011: Kernel Methods for Pattern Analysis**

**Programming Assignment I**

**Date: January 20, 2015**

**Deadline for submission of report (PDF file) on the moodle**: **4PM, Thursday, February 05, 2015**

**Note**: **Each team must use the datasets identified for that team**

**Function approximation**

**Datasets**

**Dataset 1:**  1-dimensional (Univariate) input data

**Dataset 2:** 2-dimensional (Bivariate) input data

**Models:**

1. Polynomial curve fitting for Dataset 1
2. Linear model for regression using Gaussian basis functions for Dataset 2. Use the centers of clusters as the centers of Gaussian basis functions. Clusters may be formed using the K-means clustering method on the training data.

**Presentation of Results for each Dataset:**

1. Dataset 1: Plot of the approximated functions obtained using training datasets of different sizes, for different model complexities and for different values of regularization parameter. (Similar to Figures 1.4, 1.6 and 1.7 of Bishop’s book).
2. Dataset 2: Plot of target output and model output for training data, for different model complexities and for different values of regularization parameter.
3. Datasets 1and 2: Plot of the root mean squared error (RMSE) on training data, validation data and test data, for different model complexities and for different values of regularization parameter. (Similar to Figures 1.5 and 1.8 of Bishop’s book).
4. Datasets 1 and 2: Scatter plot with target output on *x*-axis and model output on *y*-axis, for training data, validation data and test data, for different model complexities and for different values of regularization parameter.

**Report by a team should include the plots and observations about the results of studies**