**CS 557 STATISTICAL PATTERN RECOGNITION AND LEARNING**

**FALL 2014**

**ASSIGNMENT 7**

**DUE: Tuesday December 2, 2014.**

**PROBLEM**

1. There are three datasets trainData, validData, testData. The format of all three files is the same. The first columns is X and the second column is the target. It is univariate data.

2. Fit the following regression models to the **training data only**:

a. linear

b. quadratic

c. cubic

d. fourth degree

e. fifth degree

f. sixth degree

3. Plot the original training data along with predictions on the training data for all the regression models. You can use separate graphs or the same graph.

4. Find the error rate of your predictions. You can use one error rate of your choice (optional) and the other error rate should be: sum(|prediction-target|)/totalExamples

5. Use the validation data to evaluate all models (i.e. find the error rate of your predictions) and select the best one.

6. Use the test data to evaluate all models and note your observations. Do you get the best accuracy on the test set for the model you selected using the validation set?

Note: You can use matlab’s helper functions like load, sum, plot etc. but **NOT** the curve fitting functions provided by it. You have to write all the regression routines yourself.

**TO SUBMIT**

1. Make a folder with your roll number as folder name. Put Matlab’s source code in it and place it in the ‘submit assign1’ folder on xeon
2. Hard copy of a report which is **not more than two pages** long that shows the plots and mean square error values for the three datasets. Present your results clearly using tables.

**IMPORTANT**

There is a strict policy against plagiarism and cheating. The penalty can be an F grade.

**TO SUBMIT**

1. Make a folder with your roll number as folder name. Put Matlab’s/Python’s source code in it and place it on xeon. PLEASE DO NOT EMAIL
2. **Hard** **copy** of a report which is **not more than two pages** long that describes all the results of your experiments AND YOUR CONCLUSION and COMMENTS ON THE RESULTS.