Homework Assignment # 1 due on 13 Sep 2017 (Wednesday)

1. Write a source code (in python) for investigating the over-fitting and model selection problems using the function sin ($2\pi x$) in the range [0,1].

a) First generate the training data points (with N = 10) using the above function by adding some level of random noise having Gaussian distribution with $\mu=0$, $\sigma^2=1$. (Training Data $(x_N, t_N,)$).

b) Fit the model using the M[th] polynomial $y = \sum_{i=0}^{M} w_i x^i$ where M = 0, 1, 2, and 8. Plot the results obtained by the fitting the model and compute the RMS error in each case. Generate the 50 test data points and compute the RMS error and plot the results of both training error and testing error and discuss it with the values of M=0,1,2,8.

c) Repeat (a) and (b) with N = 20 and 100

Discuss the model selection and overfitting problem in the context with number of data points and order of polynomial (M=0,1,2,8) for each N.