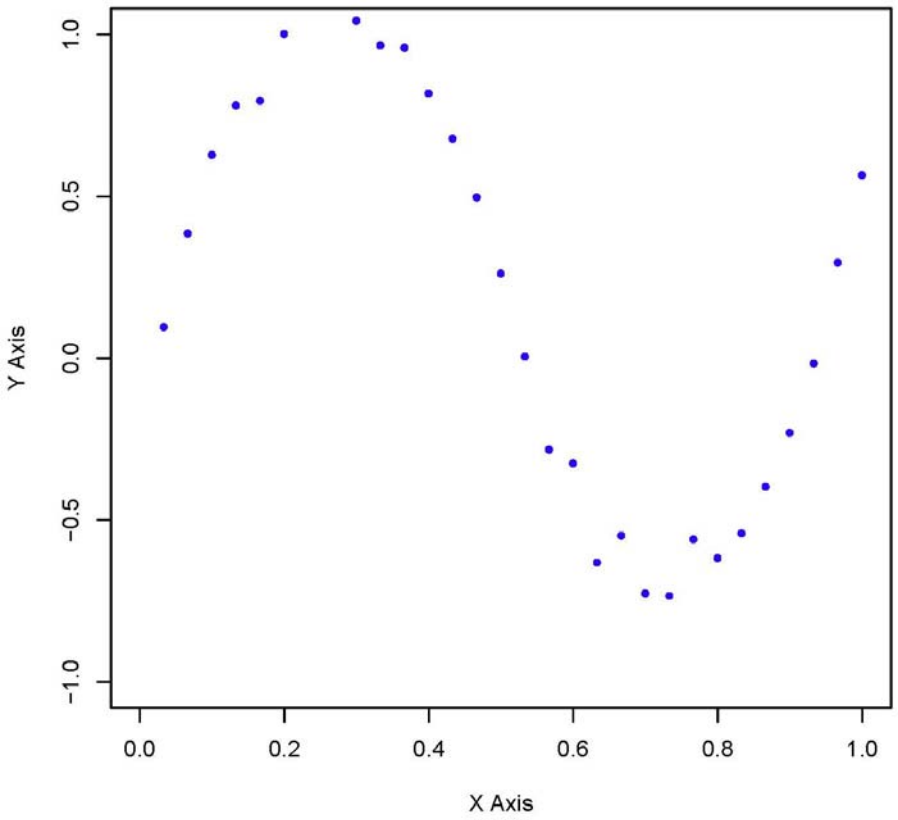
Machine Learning -- Assignment 1

Chinmaya R Naguri – cn262114

5.

a. Plot for the complete Dataset

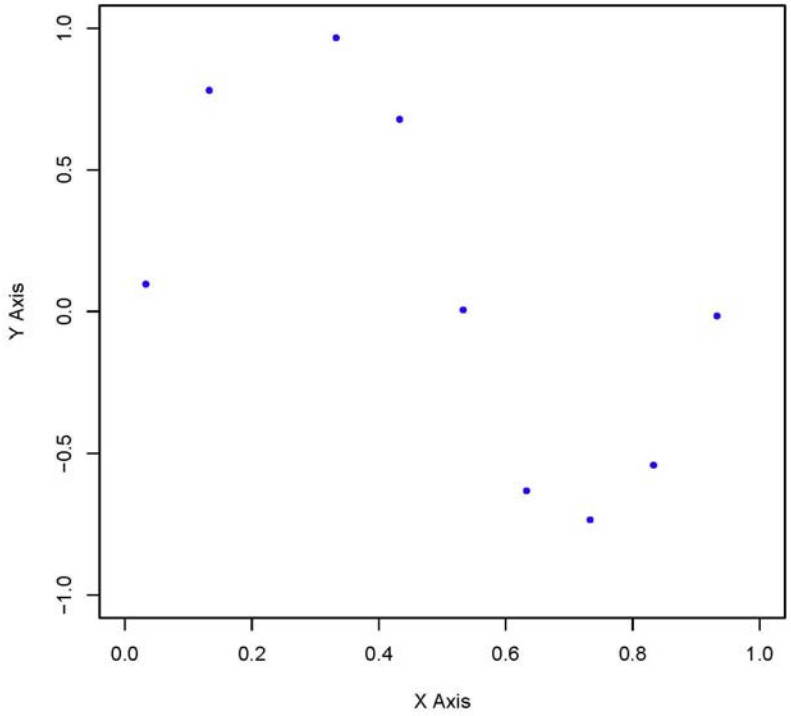
Rplots_Dataset.pdf



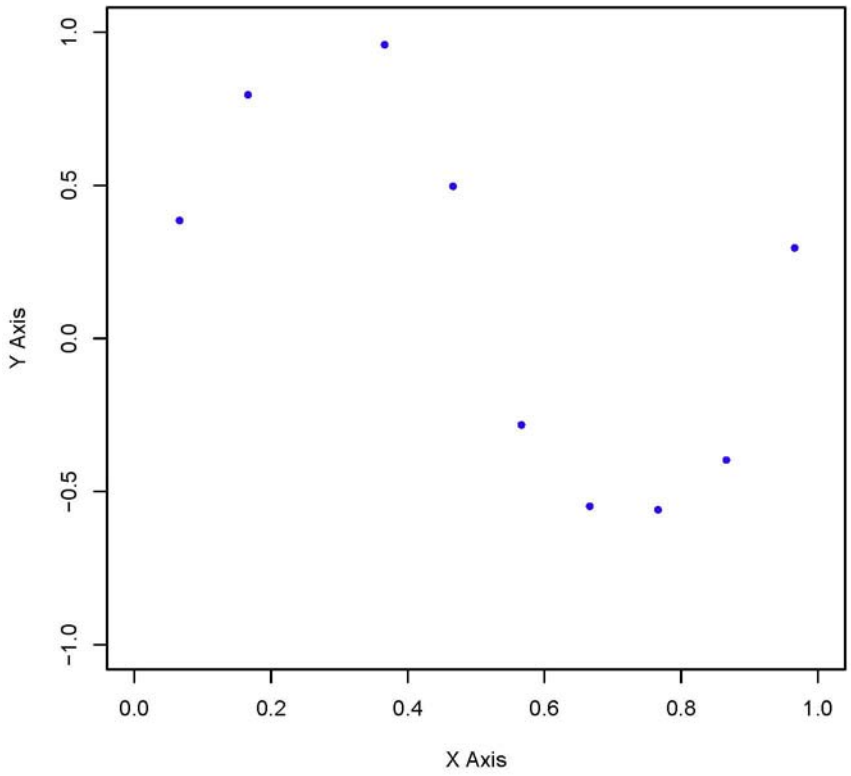
b. Plot for Train, Test and Validation data

Rplots_Train.pdf Rplots_Test.pdf Rplots_Valid.pdf

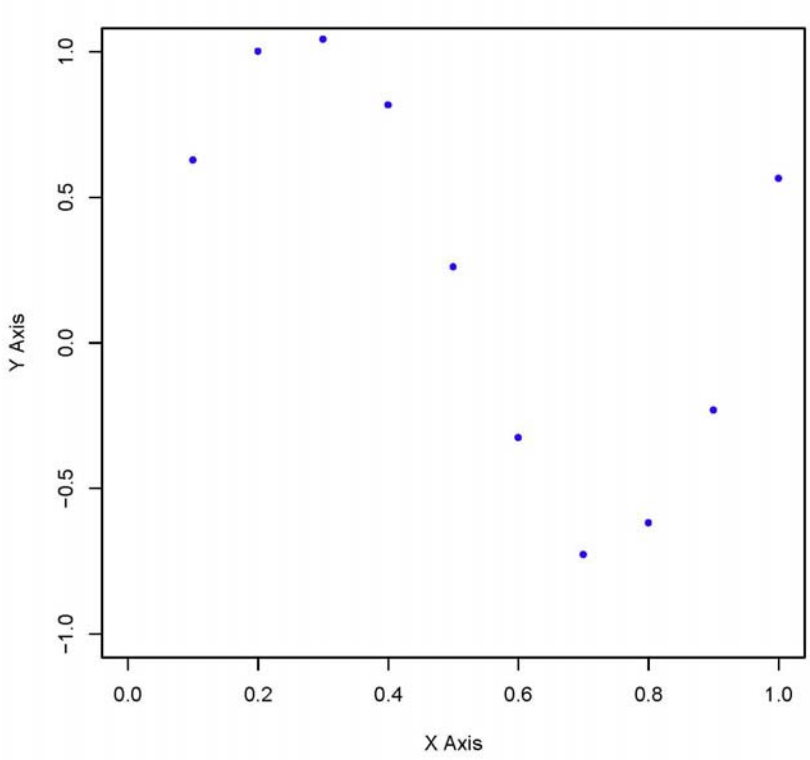
Training Data



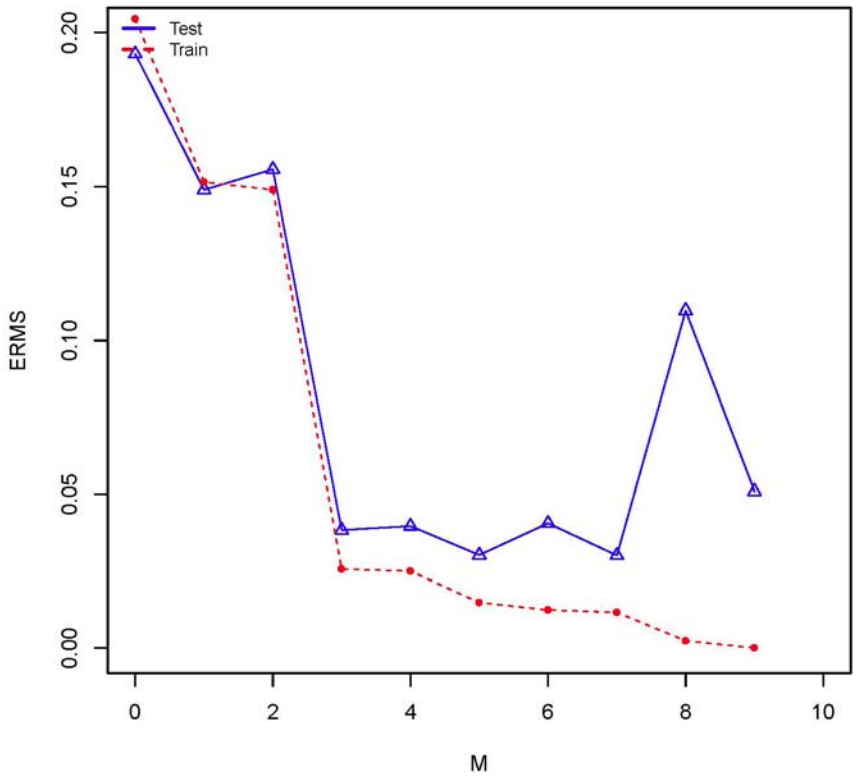
Test Data



Validation Data

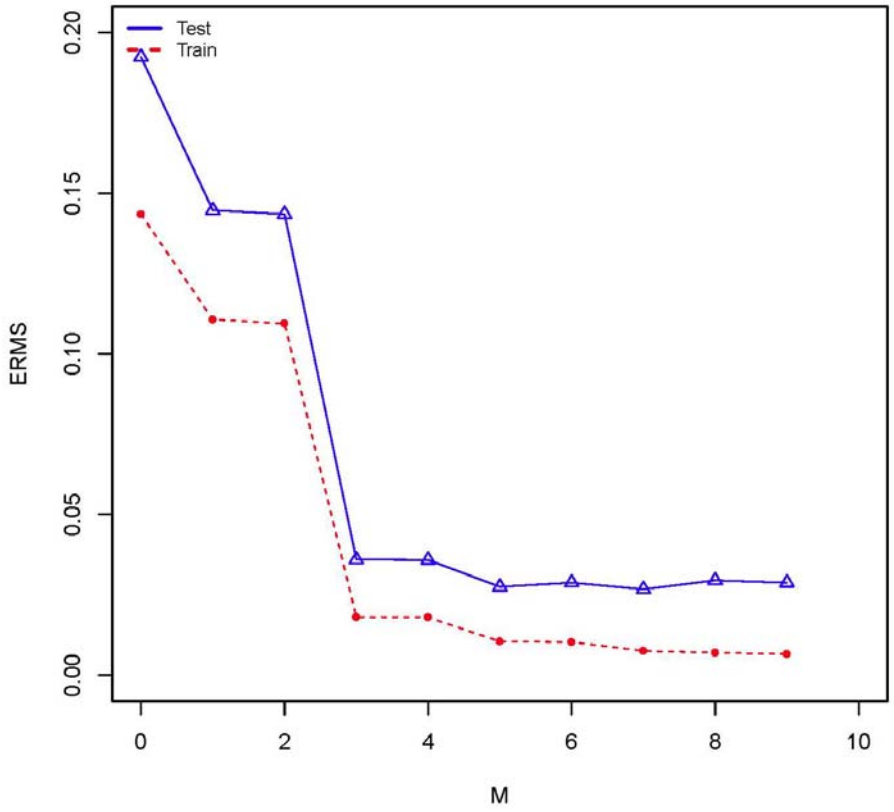


c. 1. **Without Regularization – M vs ERMS values for Training and Test Data**



Rplots_MvsERMS_
WithoutReg_Trainin

c. 2. **Without Regularization – M vs ERMS values for Training + Validation Data and Test Data**



Rplots_MvsERMS_
WithoutReg_Train&

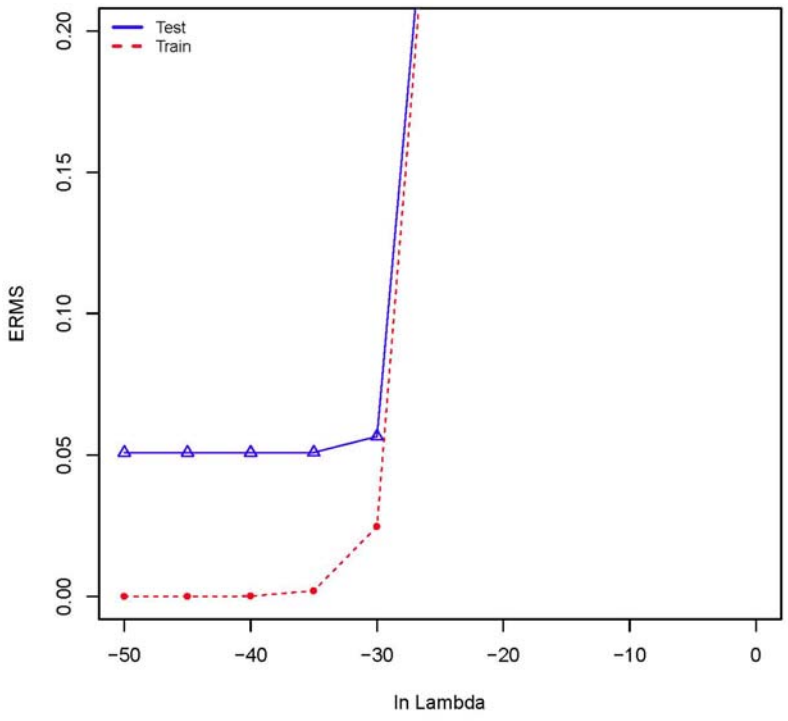
In Without Regularization, with just use *Train.dat* (i.e. training data). We can see that M-ERMS Graph for Test data is consistent in M = [3 to 7] with some distortion. But when the training data is increased to Train + Validation data these distortions are gone & is smooth.

When M = 8/9 with just *Train.dat* data, ERMS value is high but after introducing validation data into training we can observe that it is smoothed out and is consistent with the rest of graph.

One more observation would be the ERMS values of the training data prediction is considerably reduced after introducing validation data into training data set.

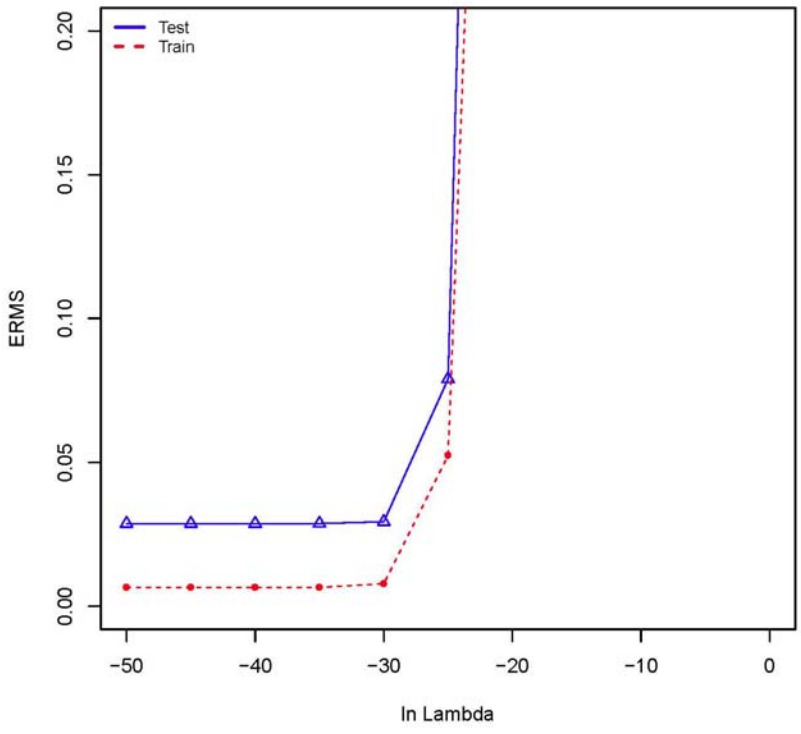
So, we can see that introducing more data reduced the ERMS values i.e. deviation of the predicted value from the original is getting reduced. It even stabilized the error rate as well from M = 3 to M = 9 for this data set.

c. 3. **With Regularization – ln λ vs ERMS value for Training Data & Test Data (M=9)**



Rplots_MvsERMS_
WithReg_TrainingDa

With Regularization – ln λ vs ERMS value for Training + Validation Data & Test Data (M=9)



Rplots_MvsERMS_
WithReg_Train&Vali

We introduce regularization if we need a good prediction even from few training examples. Here we can see for ln λ = -50 to ln λ = -30, ERMS values are consistent and then it shoots to really high values. After introducing the validation data into training data we observe that the ERMS value almost reduced by 50%. i.e. our purpose of introducing ln λ is served.

An interesting observation along with this is that the ERMS values of the training + validation data increased than that of the training data, this might be caused due to the extra data which might be introducing noise.

One more observation would be the ERMS value for ln λ = -25 is visible in the graph (i.e within range of 0.2) for the training + validation data.

If we have to choose an ln λ value then I would suggest to do it in between -50 to -30 as this range has the least & consistent ERMS values.