Machine Learning (CS60050)

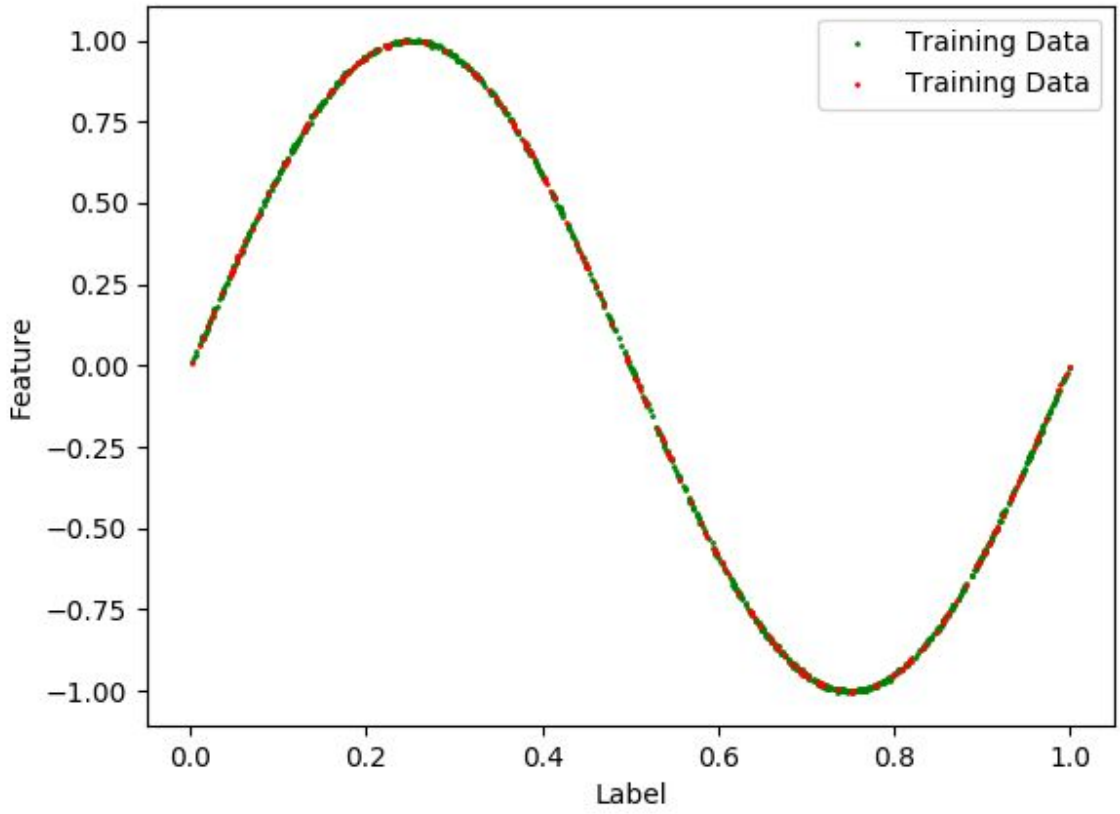
Assignment 1: Linear Regression

Name : Sanket Meshram

Roll NO.: 17CS30030

Part 1

Feature VS. Lable Plot *:*



Parameters for degree from 1 to 9:

Degree = 1 : [0.916085945138095, -1.8551469696207938]
Degree = 2 : [0.9737442801835667, -2.200960660808183, 0.3405230762705207]
Degree = 3 : [-0.0767107622444632, 10.506474965090034, -31.22341989014763, 20.91009547657465]
Degree = 4 : [0.0830115193715623, 7.1789094988642725, -15.648234832635653, -3.949813465235262, 12.653836818948276]
Degree = 5 : [0.19214933532351658, 5.409107429635385, -10.246347927458828, -4.97039620428527, 2.5245300301296223, 7.54786321066668]

Degree = 6 : [0.07188411988445503, 7.232111300967344, -15.800752702779072, -2.2153445033059707, 7.143565739411272, 6.203138432158996, -2.3073032846351706]
Degree = 7 : [0.034094258090568, 7.646632933024831, -16.094316704277727, -3.7246187987922417, 6.588480485119026, 8.273741039046474, 3.2233205566807532, -5.752931875305976]
Degree = 8 : [0.036915992631697155, 7.446733468235627, -14.807237270859819, -4.961867156594991, 4.648066523434921, 7.862705552940834, 5.78870314623487, 0.48898897853438483, -6.414490117455645]
Degree = 9 : [0.056302115740820786, 7.054577620218062, -13.32569986899939, -5.573456046397797, 2.9354429400924467, 6.661375897484483, 6.242060454084398, 3.2162208799395717, -1.1709198384755746, -6.09583832577531]

Test Error:

Degree 1 :

squared error on test data : 0.09553046711906686

Degree 2 :

squared error on test data : 0.09579851086810237

Degree 3 :

squared error on test data : 0.0032488537705092917

Degree 4 :

squared error on test data : 0.004675332769394 05

Degree 5 :

squared error on test data : 0.0088601840120000 28

Degree 6 :

squared error on test data : 0.0045907791323408 06

Degree 7 :

squared error on test data : 0.0023335021565723587

Degree 8 :

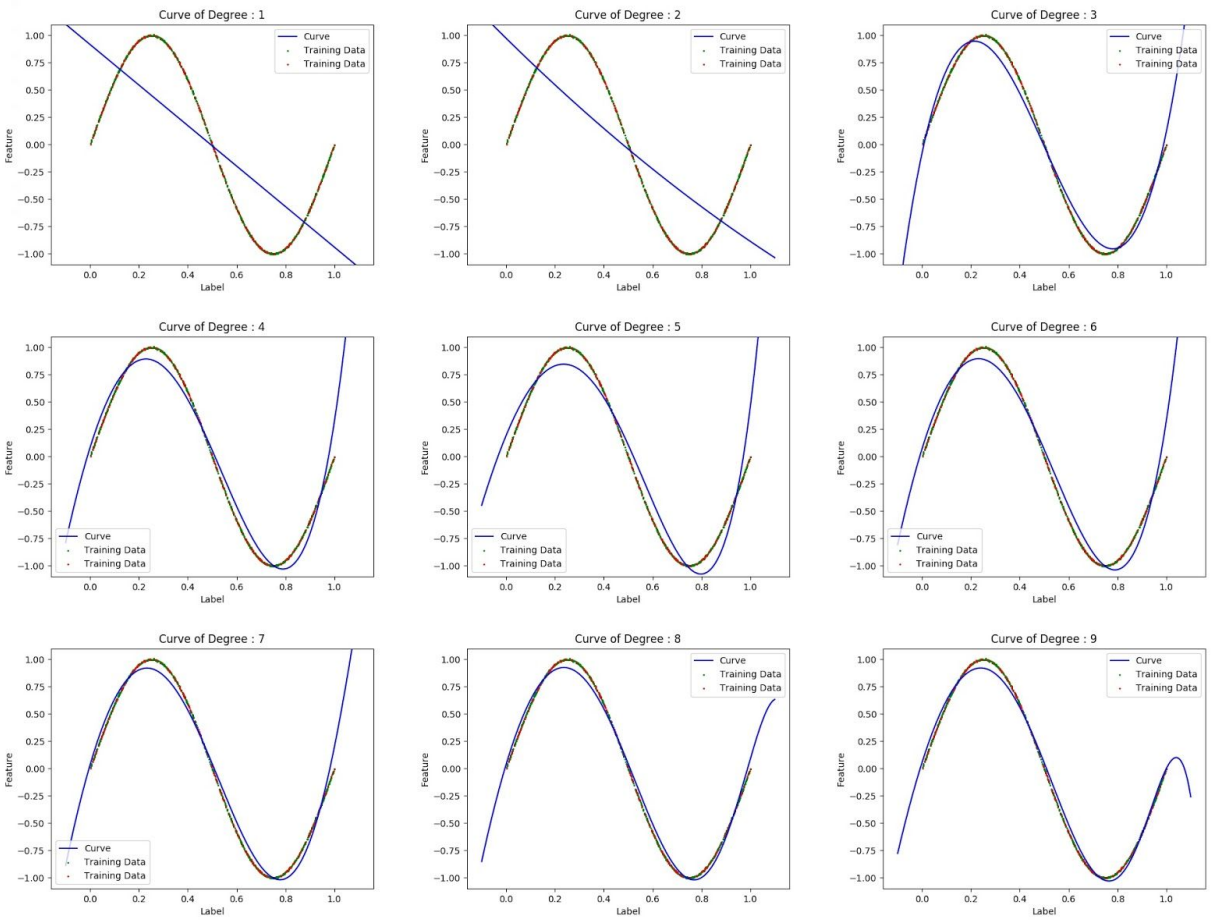
squared error on test data : 0.0014136056818152085

Degree 9 :

squared error on test data : 0.0012161062834012406

Part 2

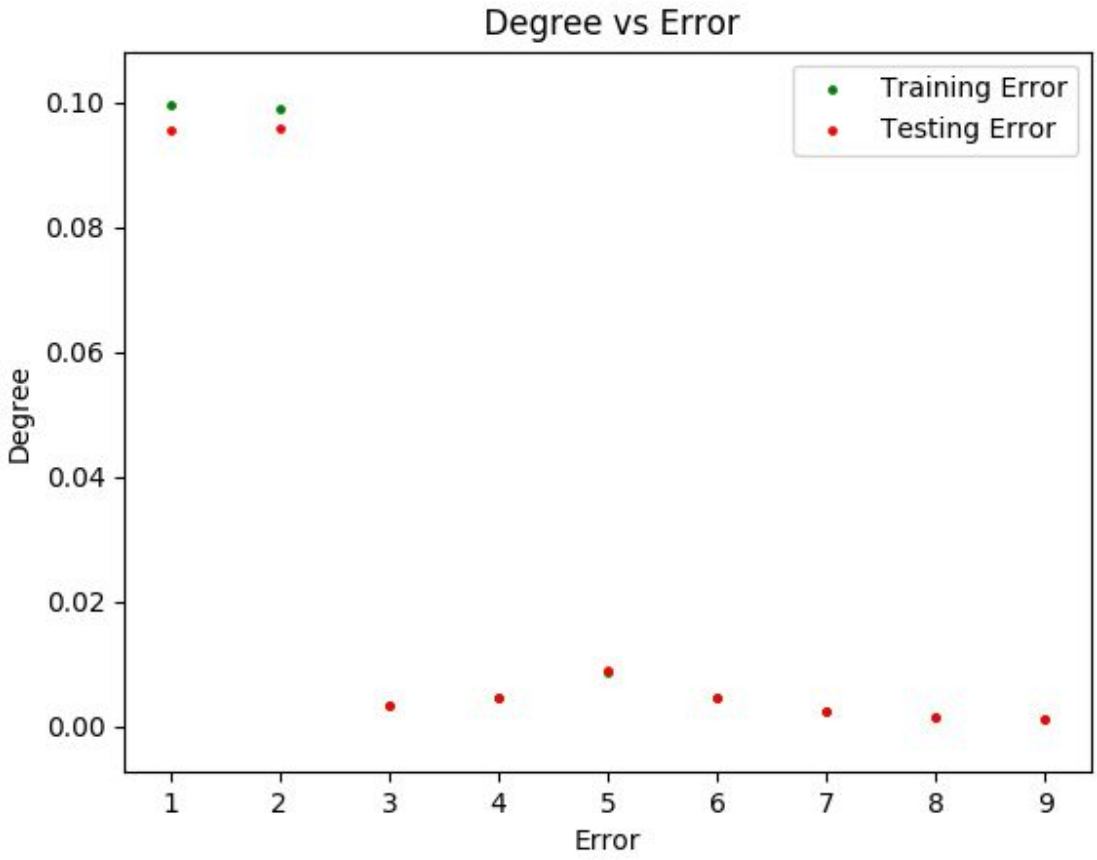
Plots for Degree from 1 to 9 :



Squared error on both train and test data for Degree from 1 to 9 :

Degree 1 :
squared error on train data : 0.099680541924774 88
squared error on test data : 0.09553046711906686
Degree 2 :
squared error on train data : 0.09914021585733448
squared error on test data : 0.09579851086810237
Degree 3 :
squared error on train data : 0.00323924405013444 87
squared error on test data : 0.00324885377050929 17
Degree 4 :
squared error on train data : 0.00461775461929982 9
squared error on test data : 0.00467533276939405
Degree 5 :
squared error on train data : 0.00865273768685701 6
squared error on test data : 0.00886018401200002 8
Degree 6 :
squared error on train data : 0.00454417418698067 6
squared error on test data : 0.00459077913234080 6
Degree 7 :
squared error on train data : 0.00233754525188824 84
squared error on test data : 0.00233350215657235 87
Degree 8 :
squared error on train data : 0.00143208746113048 41
squared error on test data : 0.00141360568181520 85
Degree 9 :
squared error on train data : 0.00122281874054321 17
squared error on test data : 0.00121610628340124 06

PLOT for Squared error VS Degree :



N=9 is Most suitable for this data.
Because it has least Squared Error on Test set as well as Train set.
And From all plots also we can see that it's (N=9) Best fitted plot.

Part 3

Degree 1 and 9 have maximum and minimum training error Respectively.
We will apply Regulerisation on Degree 1 and 9.

Error for Lasso Regression :

lambda = 0.25

squared error on train data when deg = 1 is : 0.09969039597658721
squared error on test data when deg = 1 is : 0.09547685812533711
squared error on train data when deg = 9 is : 0.008339583907523 58
squared error on test data when deg = 9 is : 0.008674067531898472

lambda = 0.5

squared error on train data when deg = 1 is : 0.09968771024528741
squared error on test data when deg = 1 is : 0.09548584359464295
squared error on train data when deg = 9 is : 0.008328697086353483
squared error on test data when deg = 9 is : 0.008670620369371 01

lambda = 0.75

squared error on train data when deg = 1 is : 0.09968536369618507
squared error on test data when deg = 1 is : 0.0954953911756938
squared error on train data when deg = 9 is : 0.008320493681063953
squared error on test data when deg = 9 is : 0.008669776731771826

lambda = 1

squared error on train data when deg = 1 is : 0.09968349788979701
squared error on test data when deg = 1 is : 0.09550480050902876
squared error on train data when deg = 9 is : 0.008314873979916327
squared error on test data when deg = 9 is : 0.008671431348204252

Error for ridge Regression

lambda = 0.25

squared error on train data when deg = 1 is : 0.099702940753057 04
squared error on test data when deg = 1 is : 0.095447402501404 71
squared error on train data when deg = 9 is : 0.023422233739543 79
squared error on test data when deg = 9 is : 0.024675549454413 18

lambda = 0.5

squared error on train data when deg = 1 is : 0.099714695453038 99
squared error on test data when deg = 1 is : 0.095429088834074 03
squared error on train data when deg = 9 is : 0.027833516424162 98
squared error on test data when deg = 9 is : 0.029390701313720 194

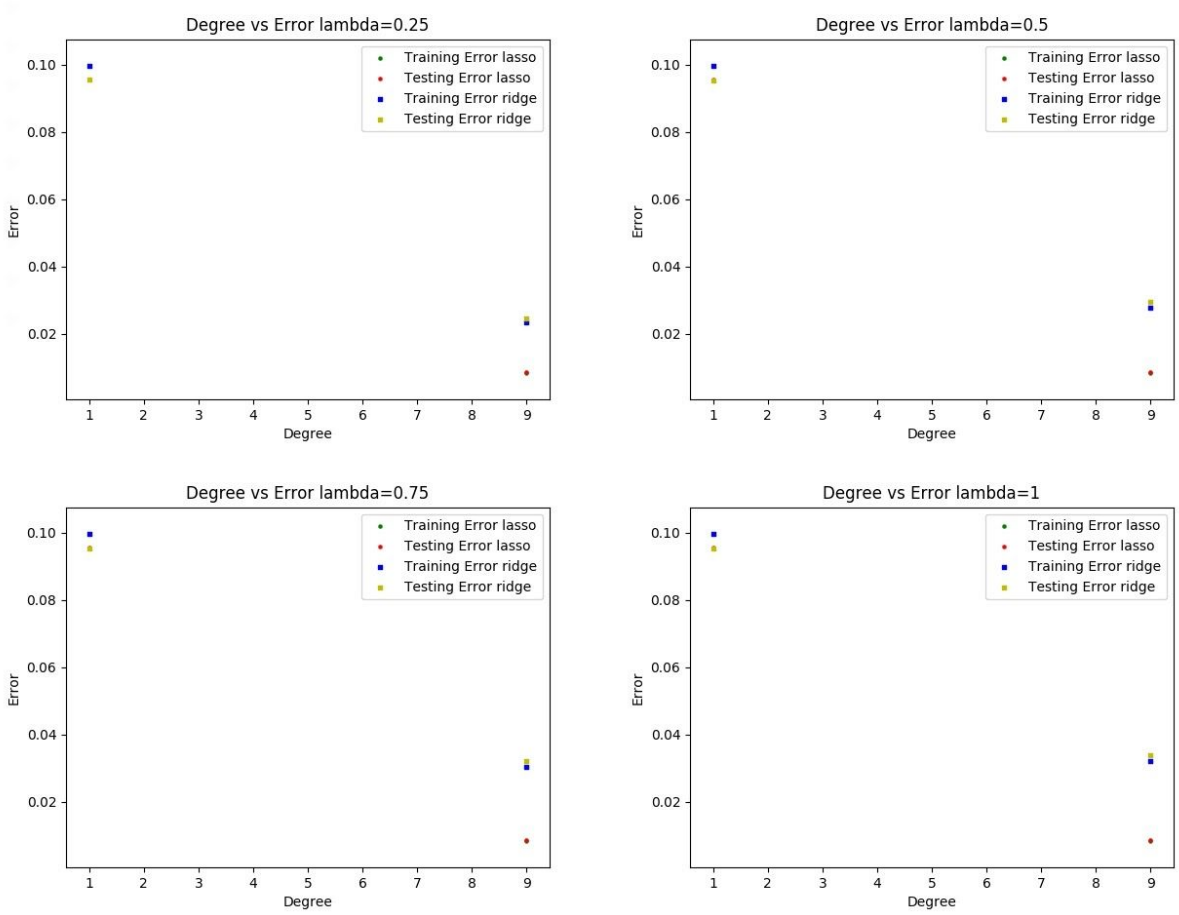
lambda = 0.75

squared error on train data when deg = 1 is : 0.099728551713804 36
squared error on test data when deg = 1 is : 0.095413324689783 43
squared error on train data when deg = 9 is : 0.030247561068265 43
squared error on test data when deg = 9 is : 0.031967315840444 585

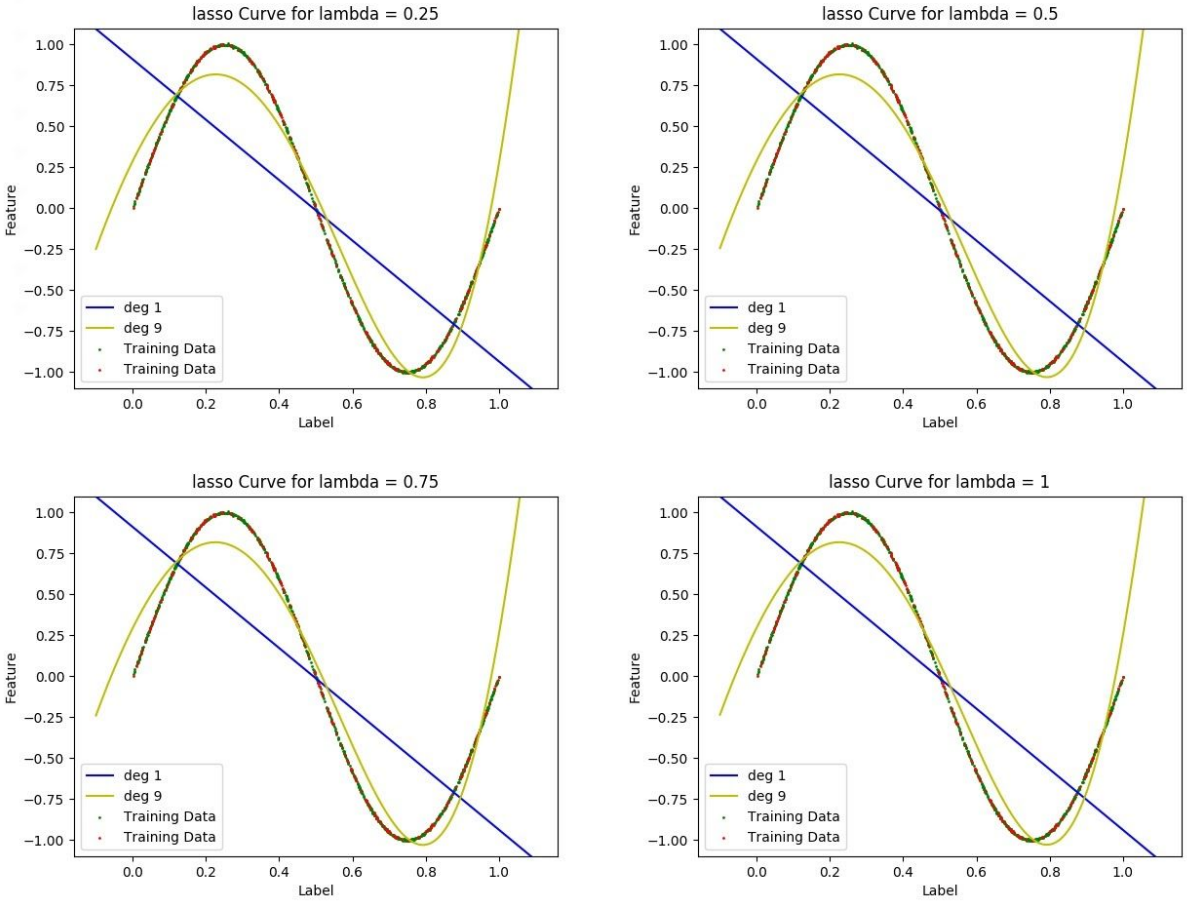
lambda = 1

squared error on train data when deg = 1 is : 0.099744812461988 3
squared error on test data when deg = 1 is : 0.095399647830831 32
squared error on train data when deg = 9 is : 0.031996045171430 836
squared error on test data when deg = 9 is : 0.033819723428435 29

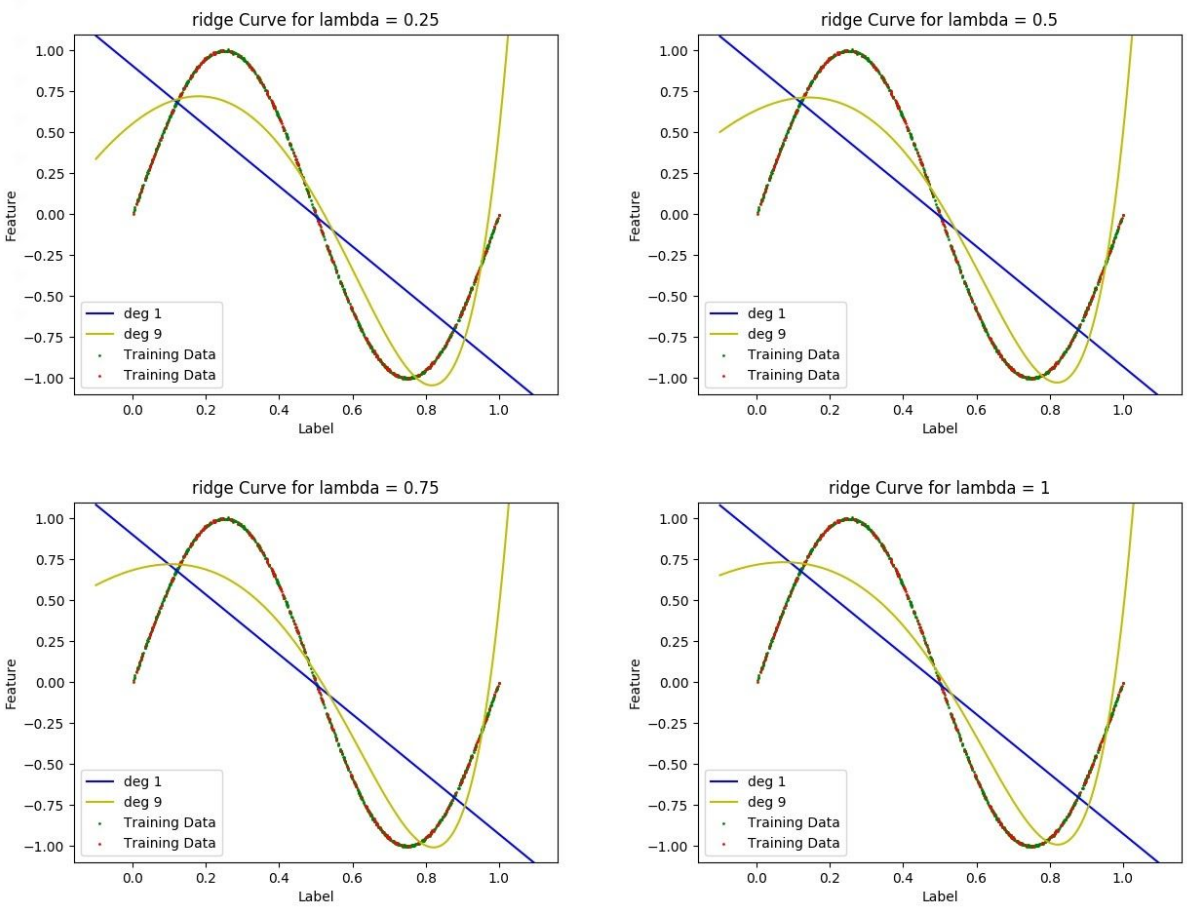
PLOT for Squared error VS Degree :



Plot for Lasso regression on degree 1 and 9 :



Plot for Ridge regression on degree 1 and 9 :



Lasso regression tries to reduce value of parameter i.e. w can be large -ve provided it’s local minima.

But Ridge regression tries to reduce magnitude of parameter.

I will prefer Lasso regression Because it has less Error than ridge and also from plots we can see Lasso regression Model fits data well.