Name: Sai Anish Garapati
UIN: 650208577

Assignment 2:

1) Question Answering

Q1) Overfitting: The case in which a learnt model fits very well on the training data but fails to generalize on the unseen new examples.
Underfitting: The case in which a learnt model fits poorly both on the training data as well as the test data.

Q2) Two strategies to help avoid overfitting:

- Regularization: Adding penalty term to the error function to minimize the values of learned weights.
- Cross-validation: Split the training data into groups of k and perform the training with one group as testset and the remaining as the training set on all the groups.

Q3) This statement is incorrect, as we can take linear combinations of a set of nonlinear functions(basis functions) of the input variables which can produce nonlinear models with respect to the input variables.

Q4) There are 4 learnable parameters in the model, which are the bias weight, weight corresponding to input feature ‘size’, weight corresponding to feature ‘built year’ and weight corresponding to ‘score of school district’.

Q5) Validation is the method in which we split the training data into multiple groups or folds and we train on all but one fold which is the validation fold which is used for testing. Cross-validation is the method in which we go through all the possible choices of validation folds and average the results. Cross-validation is used to prevent overfitting and is also used to tune hyperparameters of the model.

2) Programming

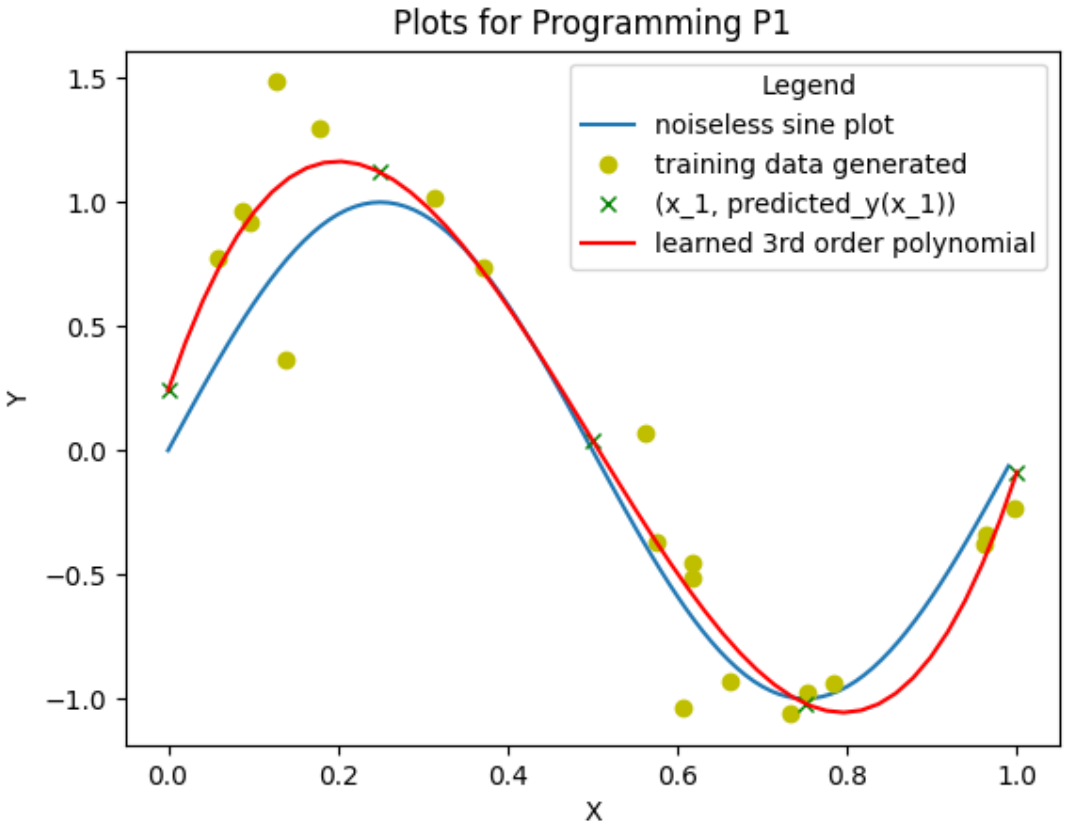
P1)
2. The learned parameters obtained from linear regression:
[[0.24189709]
[10.06499425]
[-31.47014555]
[21.07750372]]

Learned 3rd order polynomial function:
$0.24189709 + 10.06499425 * x - 31.47014555 * x^2 + 21.07750372 * x^3$

3. Predictions made using the learned model:

Prediction made on x=0.0:	0.24189709106640578
Prediction made on x=0.25:	1.1205975526686522
Prediction made on x=0.5:	0.04154579476072229
Prediction made on x=0.75:	-1.0192422090071567
Prediction made on x=1.0:	-0.08575048498475724

4.



P2)

For each fold in the 5-fold cross-validation

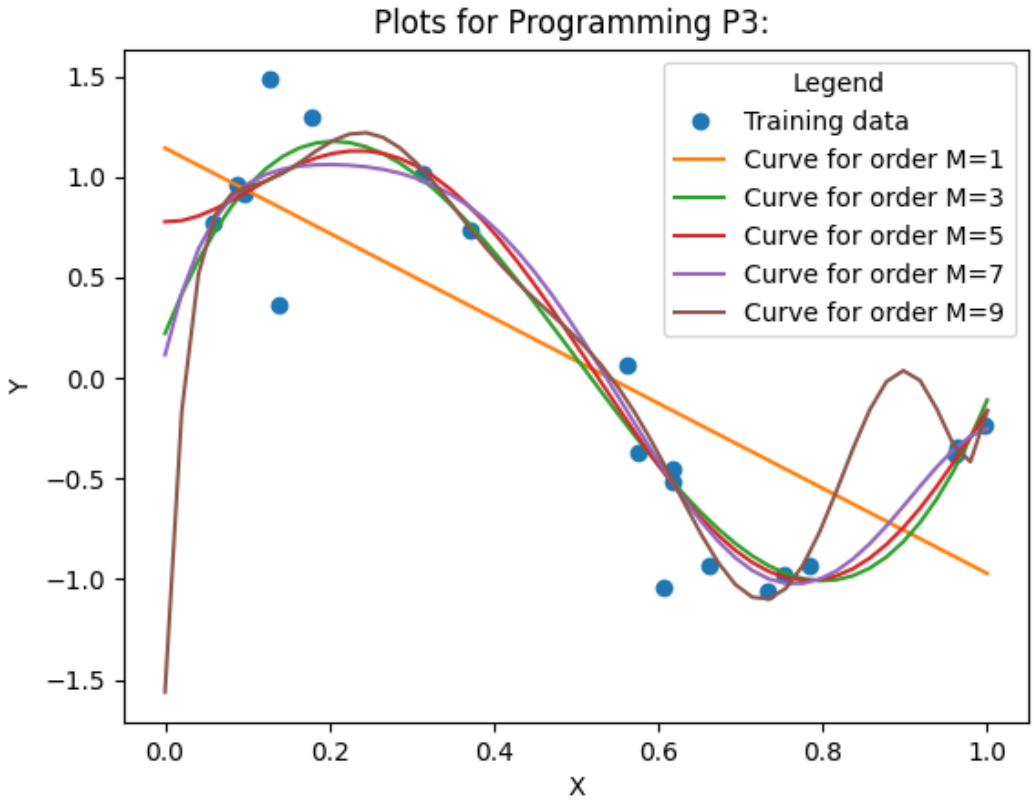
- For fold = 1, learned parameters = [[0.22226321], [10.19291896], [-31.22146407], [20.69716381]], Prediction error = 0.38017503702346417,
- For fold = 2, learned parameters = [[0.21045248], [10.46563737], [-32.56251788], [21.8943883]], Prediction error = 0.22977697921680748
- For fold = 3, learned parameters = [[0.25384085], [9.89265832], [-30.9305621], [20.69091463]], Prediction error = 0.21105793413409554
- For fold = 4, learned parameters = [[0.24317812], [10.04152729], [-31.36993417], [20.97340248]], Prediction error = 0.13872157665602158
- For fold = 5, learned parameters = [[0.2438996], [10.0669402], [-31.60418117], [21.21347486]], Prediction error = 0.18786063458947

The average prediction error on all validation folds = 0.22951843232397176

P3)

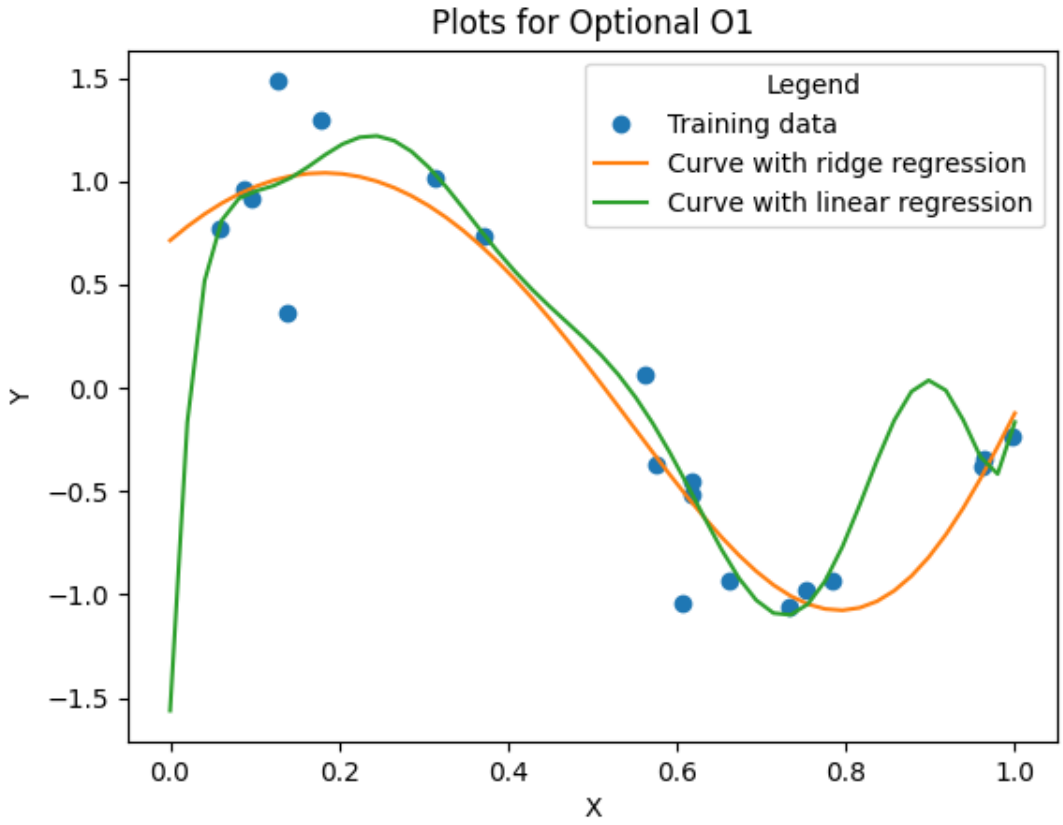
1. For polynomial function of order 1, average cross-validation error = 0.40816434108178684
For polynomial function of order 3, average cross-validation error = 0.22951843232397176
For polynomial function of order 5, average cross-validation error = 0.18802398695133443
For polynomial function of order 7, average cross-validation error = 0.30459785283210544
For polynomial function of order 9, average cross-validation error = 0.42448085486238324
The Optimal order for polynomial function is 5th order.

2.



3) Optional Programming

O1)



O2)
Lambda values taken for testing: [0.00001, 0.0001, 0.001, 0.05, 0.1, 1]

Average prediction error of lambda=1e-05: 0.24107344514577012
Average prediction error of lambda=0.0001: 0.20142770637123766
Average prediction error of lambda=0.001: 0.23376189476378412
Average prediction error of lambda=0.05: 0.2547000178679089
Average prediction error of lambda=0.1: 0.24018854286865826
Average prediction error of lambda=1: 0.332057262713207

Optimal lambda over [1e-05, 0.0001, 0.001, 0.05, 0.1, 1]: 0.0001