Assignment 1
Machine Learning
(CS60050)

Name: Prerit Jain
Roll No.: 16IM10035

Ans. 1

Polynomial Model with degree 1
Final training set loss value is 0.14455826215760814
These are the learned coefficient matrix

[[0.65829118]
 [-2.14654108]]
The loss function value on test set is 0.057458752833825115

Polynomial Model with degree 2
Final training set loss value is 0.1273280881054223
These are the learned coefficient matrix

[[1.38694231]
 [-1.56898195]
 [-0.65174781]]
The loss function value on test set is 0.0667040824372555

Polynomial Model with degree 3
Final training set loss value is 0.017699799248041048
These are the learned coefficient matrix

[[1.53862544]
 [-0.82397054]
 [-1.03325378]
 [-0.47228282]]
The loss function value on test set is 0.3552090393131233

Polynomial Model with degree 4
Final training set loss value is 0.11099612932135834
These are the learned coefficient matrix

[[0.75986593]
 [-2.3355439]
 [-0.97013078]
 [0.40567128]
 [1.42790743]]

The loss function value on test set is 0.22247355954831158

Polynomial Model with degree 5
Final training set loss value is 0.16254065359261416
These are the learned coefficient matrix

[[0.22058382]
 [-0.33262796]
 [-1.51674352]
 [-0.12362327]
 [-0.05404843]
 [0.95890957]]
The loss function value on test set is 0.0023083527822429944

Polynomial Model with degree 6
Final training set loss value is 0.12266191663882434
These are the learned coefficient matrix

[[0.78375785]
 [-1.35684264]
 [-1.13249384]
 [-0.35181386]
 [0.15543997]
 [0.64689522]
 [0.90997855]]
The loss function value on test set is 0.32569156157736123

Polynomial Model with degree 7
Final training set loss value is 0.011481106475400963
These are the learned coefficient matrix

[[0.64943308]
 [-1.73169876]
 [-1.52196669]
 [-0.23068049]
 [-0.05069072]
 [0.98919986]
 [0.74843291]
 [1.02672488]]
The loss function value on test set is 0.15816034316178057

Polynomial Model with degree 8
Final training set loss value is 0.11650571478022245
These are the learned coefficient matrix

[[0.56760918]
 [-0.76519348]
 [-1.6391063]
 [-1.67735992]
 [-0.07533587]
 [0.5512103]
 [0.99914221]
 [0.72542669]
 [0.84448736]]
The loss function value on test set is 0.003943217588839219

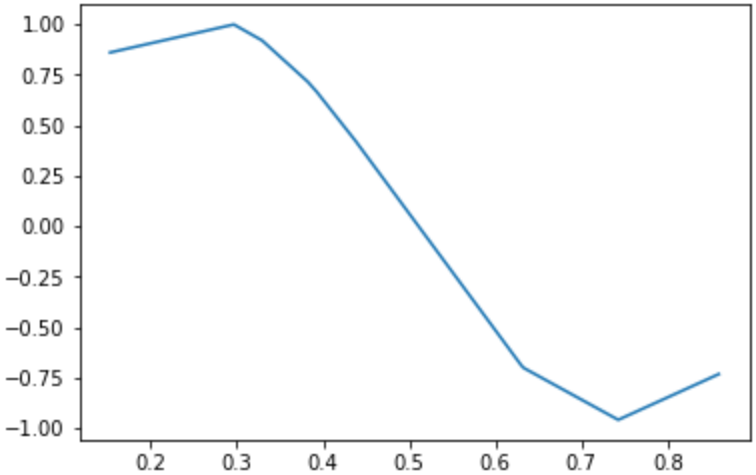
Polynomial Model with degree 9
Final training set loss value is 0.09954738842253541
These are the learned coefficient matrix

[[1.02915342]
 [0.39695657]
 [-1.64688704]
 [-1.01334774]
 [-0.42339676]
 [-0.43007748]
 [-0.62438718]
 [-0.11161208]
 [0.47262968]
 [0.176923]]
The loss function value on test set is 0.07494342333996587

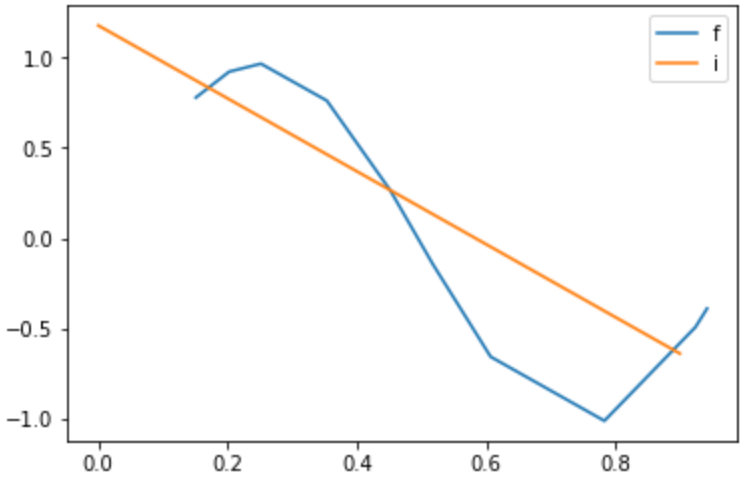
Ans. 2(a)

Plot for Dataset Size: 10

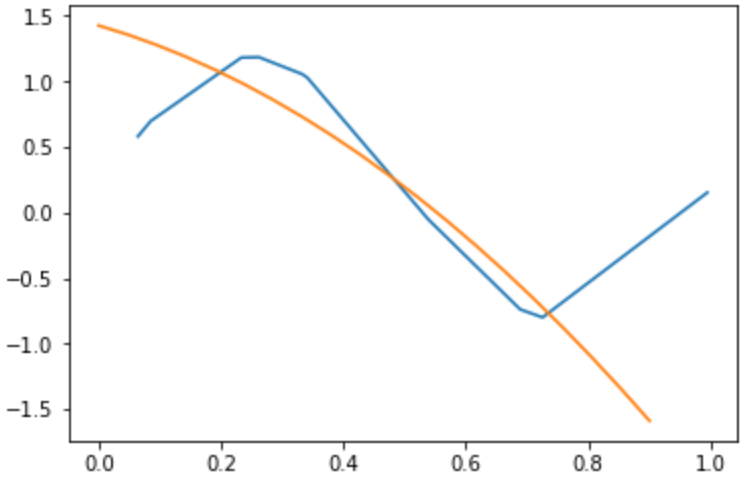
a) **Synthetic Dataset**



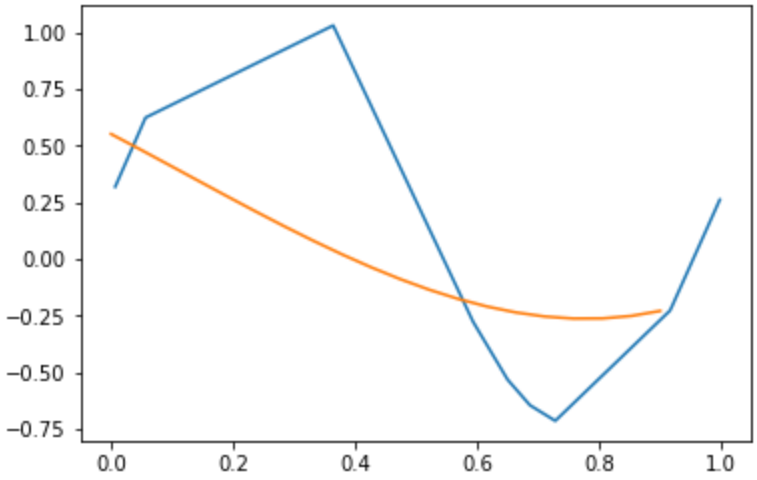
b) Polynomial degree 1



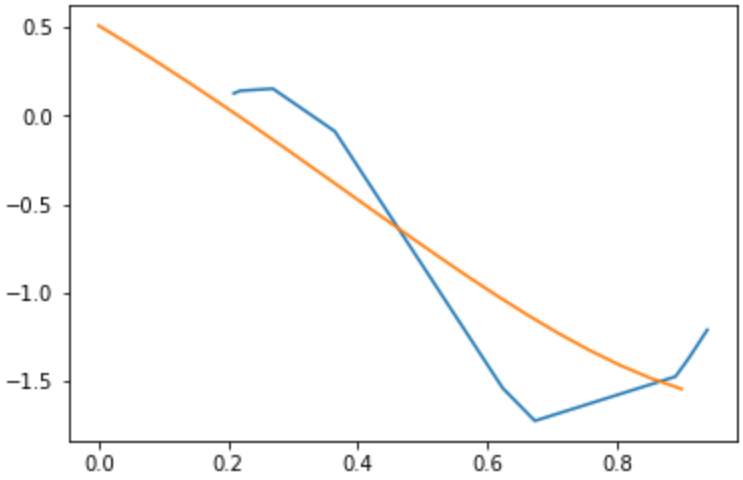
c) Polynomial degree 2



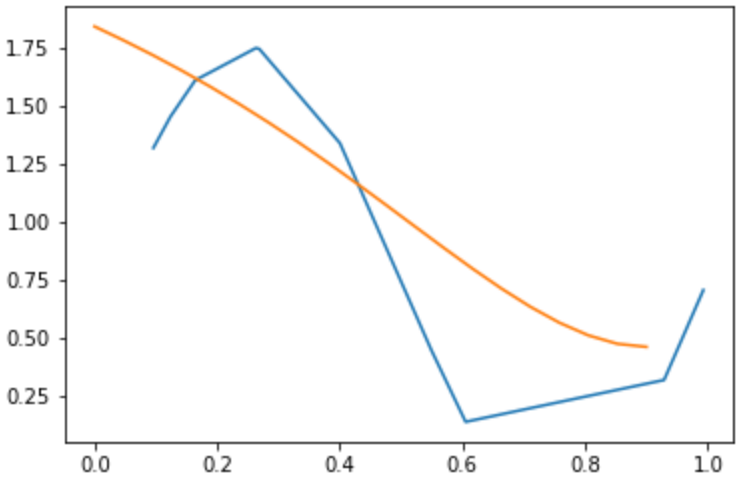
d) Polynomial degree 3



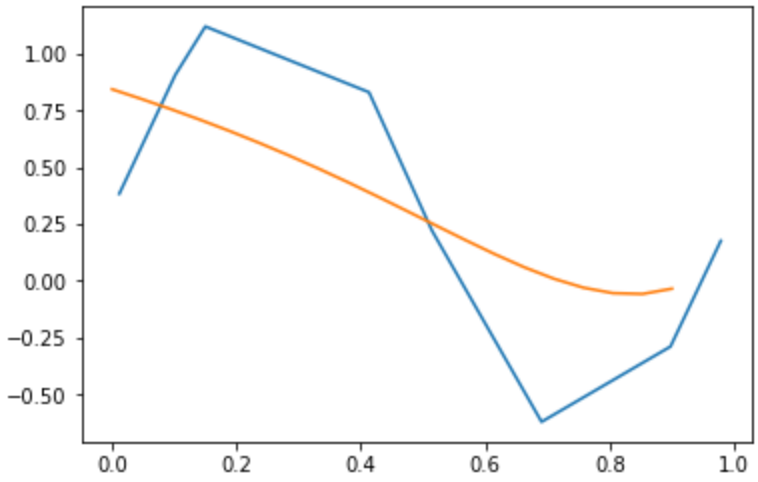
e) Polynomial degree 4



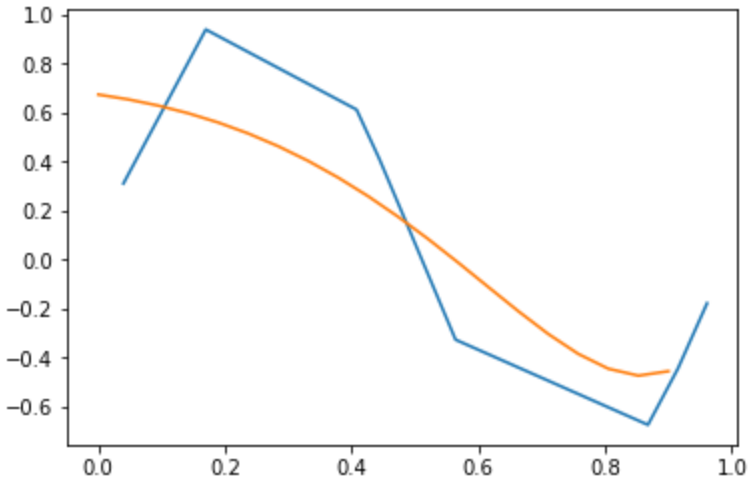
f) Polynomial degree 5



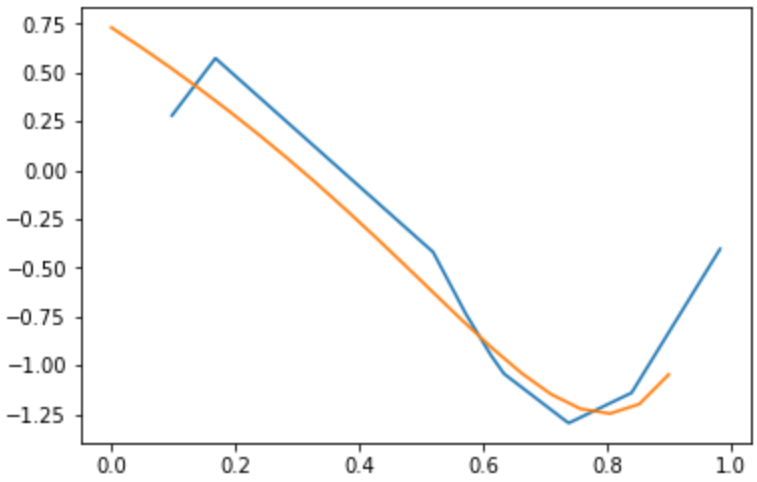
g) Polynomial Degree 6



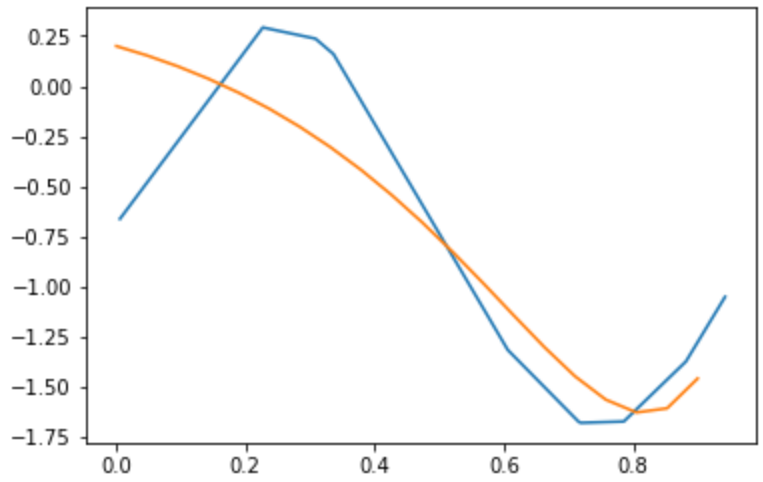
h) Polynomial degree 7



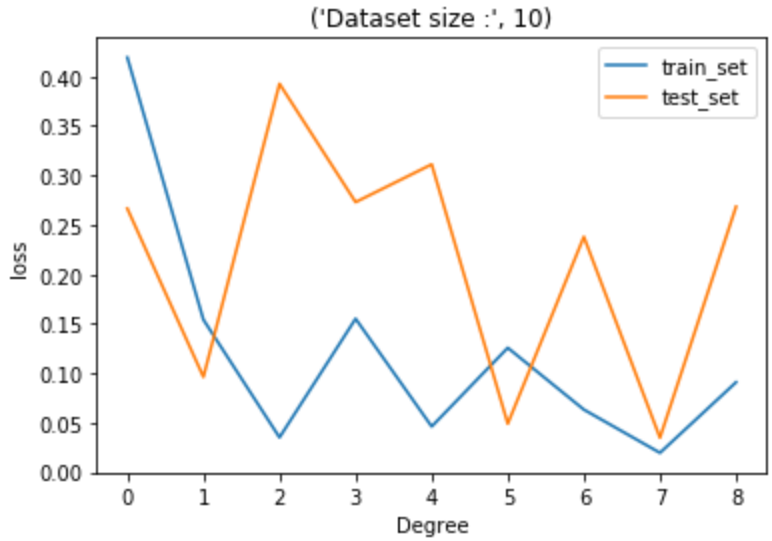
i) Polynomial degree 8



j) Polynomial degree 9



k) Loss vs Degree



All other plots are given in the Respective folders of the Dataset as submitted in the respective folder names. I am not adding it hear in order to avoid making a lengthy report.

Ans 2(b) The higher value of n are more suitable for applying the polynomial regression. It is not clear from a small dataset of size 10 but as we proceed further and plot the learning curves for larger datasets, it is quite evident that using higher degree polynomials help in better prediction.

The following observation can be justified as the expansion of sin curve is a polynomial of degree infinite, so as we add higher degree terms it helps in the generating better predictions.