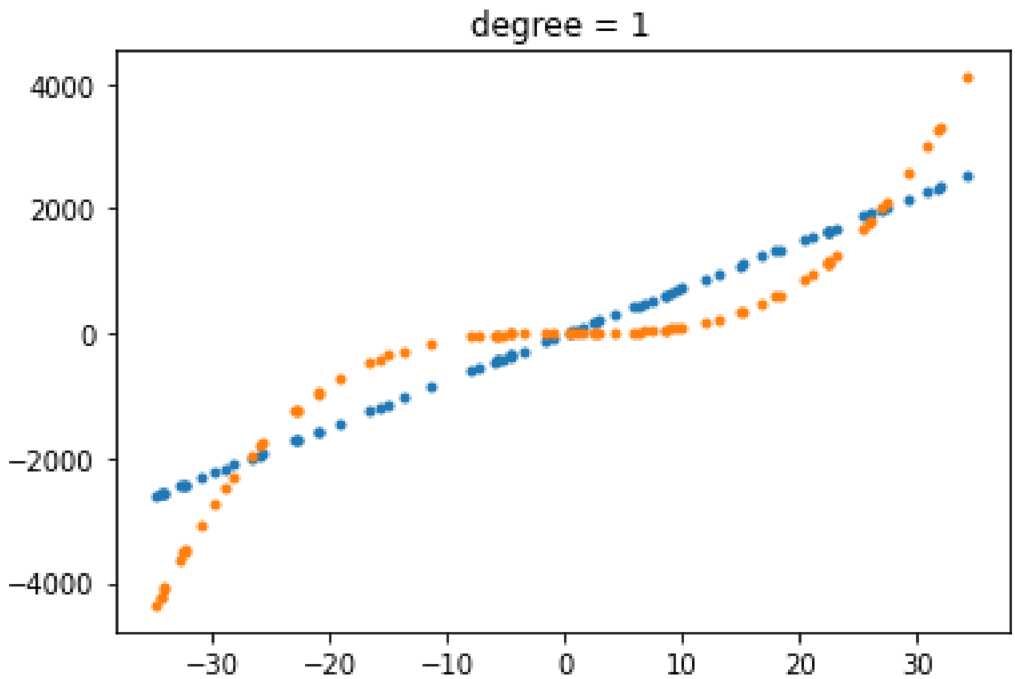
MDL Assignment 1

Team Name: RRK
Siddharth Jain: 2020113014
Tanmay Goyal: 2020111025

Task 1: Linear Regression

Linear Regression is a Machine Learning algorithm that allows us to map numeric inputs to numeric outputs, fitting a line into the data points.

LinearRegression().fit() fits a linear model with coefficients to minimize the residual sum of squares between the observed targets in the dataset, and the targets predicted by the linear approximation.

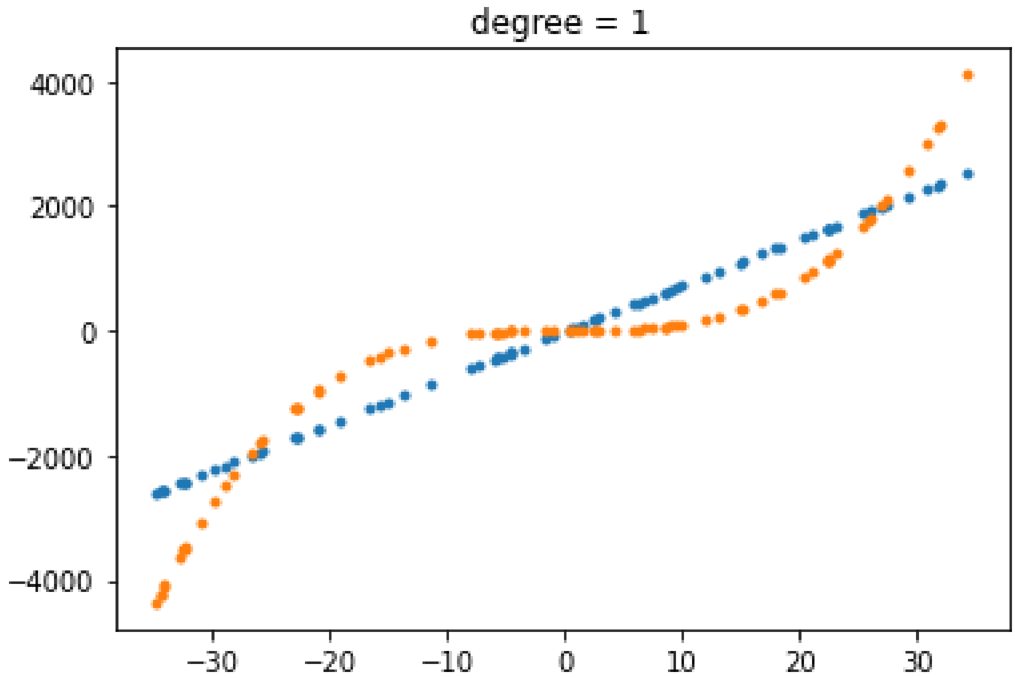


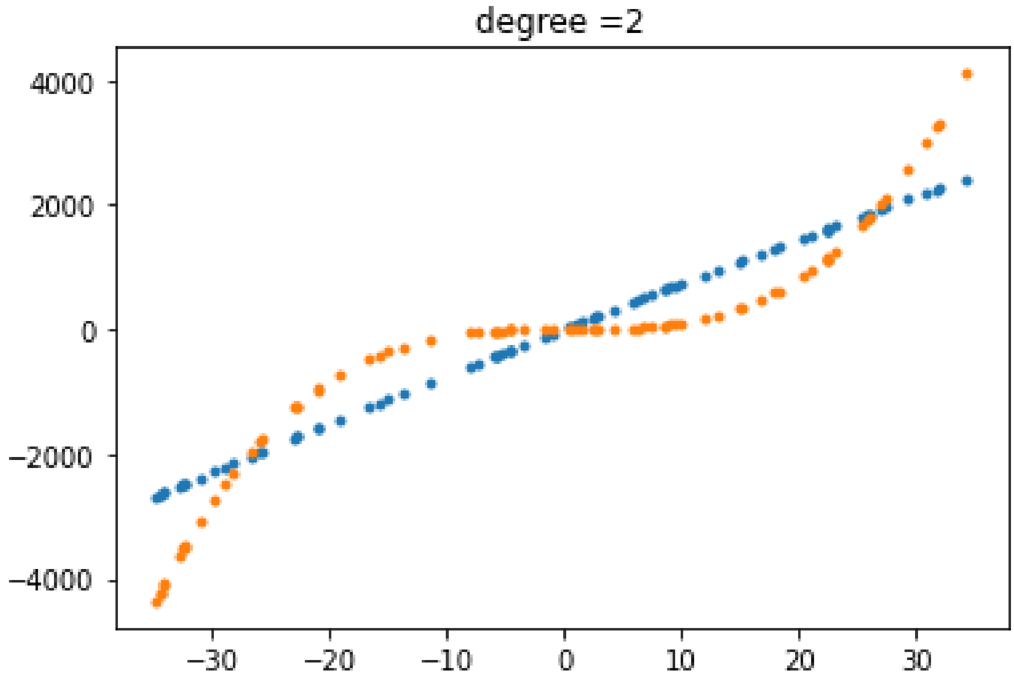
The. *fit()* function finds the optimal values of the intercepts where the arguments are the existing inputs and outputs,. To anticipate the outputs of the unseen inputs, an estimator is fitted to the model. A classifier is a type of estimator that is fitted to a model and whose goal is to learn from it. The .*fit()* function fits any Linear Regression instance.

Task 2: Calculating Bias and Variance

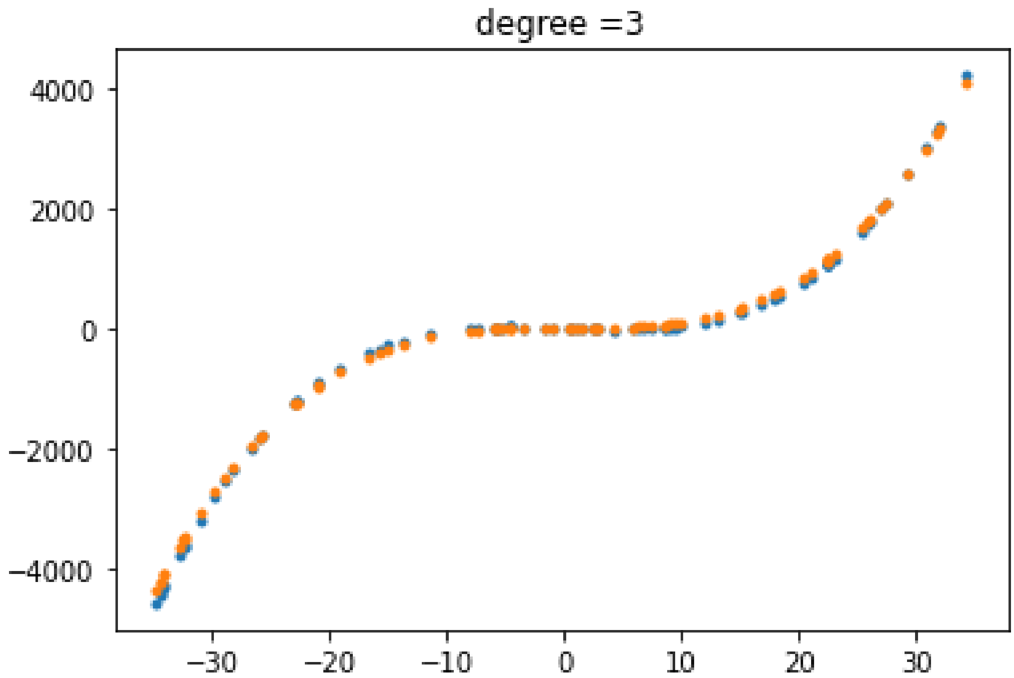
Below are the graphs depicting the predicted values against the original testing values, that is, the relationship between predictor and response.

The functions with degrees 1 and 2 do not have overlapping values for anticipated and original values, as seen in the graphs below, whereas functions with degrees 3 to 5 have the best overlap. As a result, we assume that our data represents a function of degree 3,4 or 5.

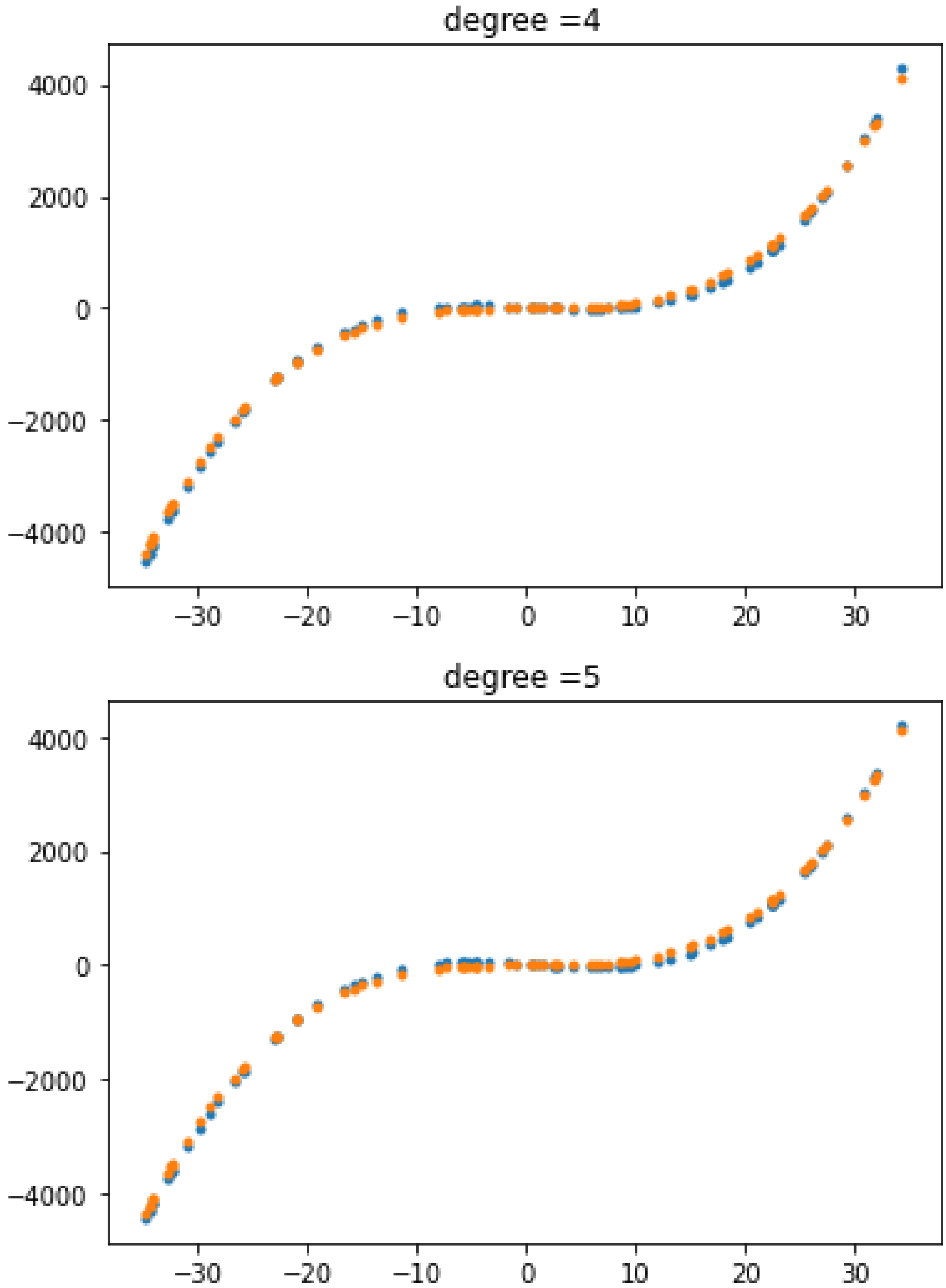


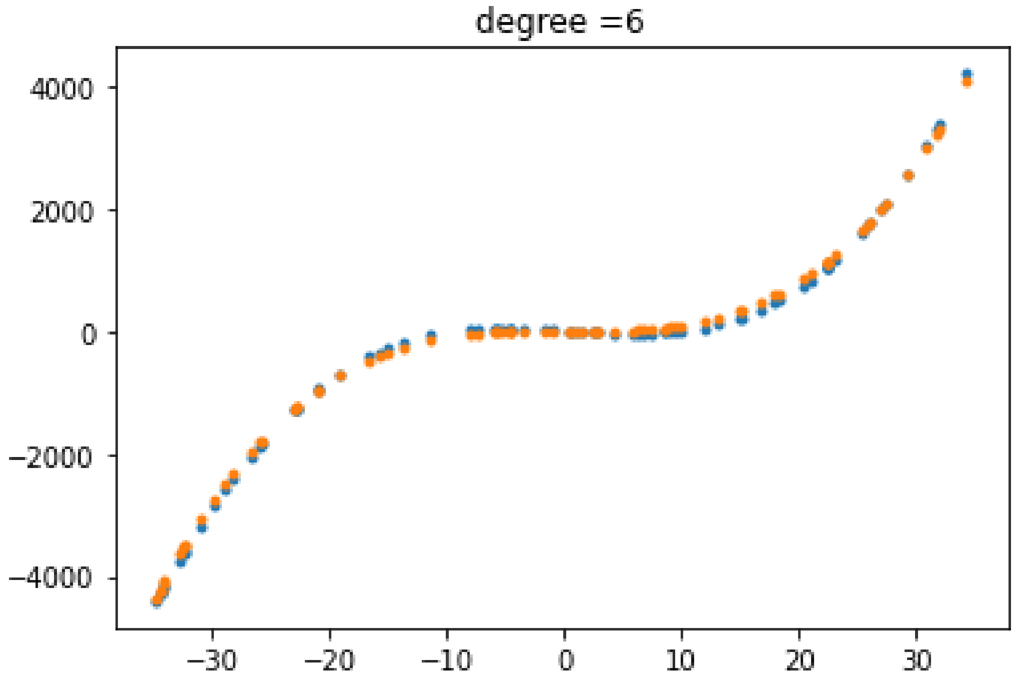


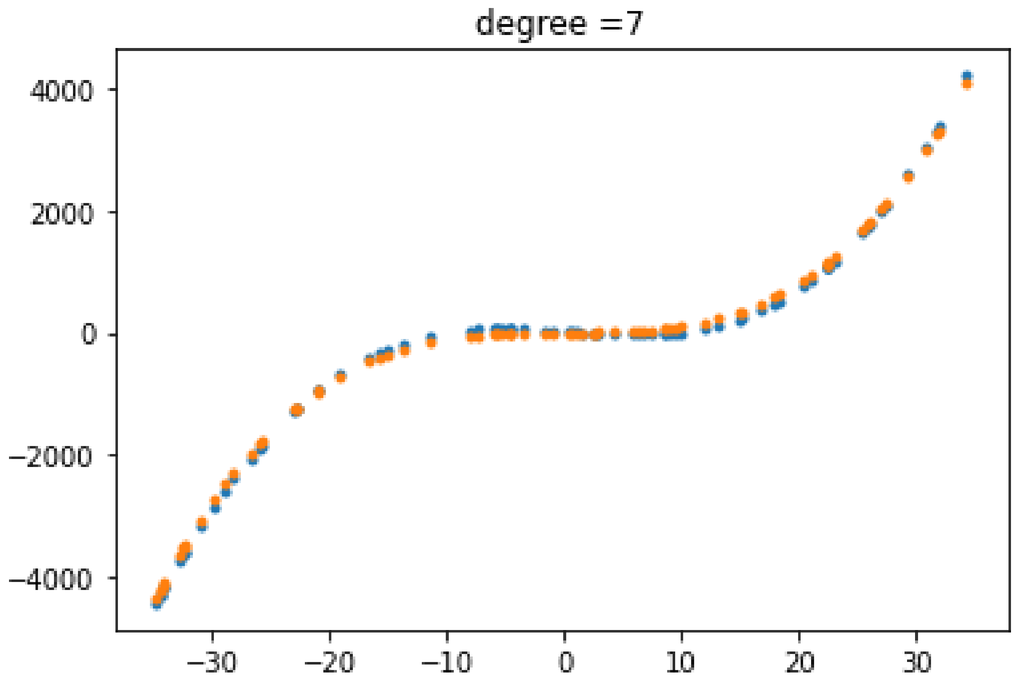
degree =2
4000
2000
0
−2000
−4000
−30
−20
−10
0
10
20
30

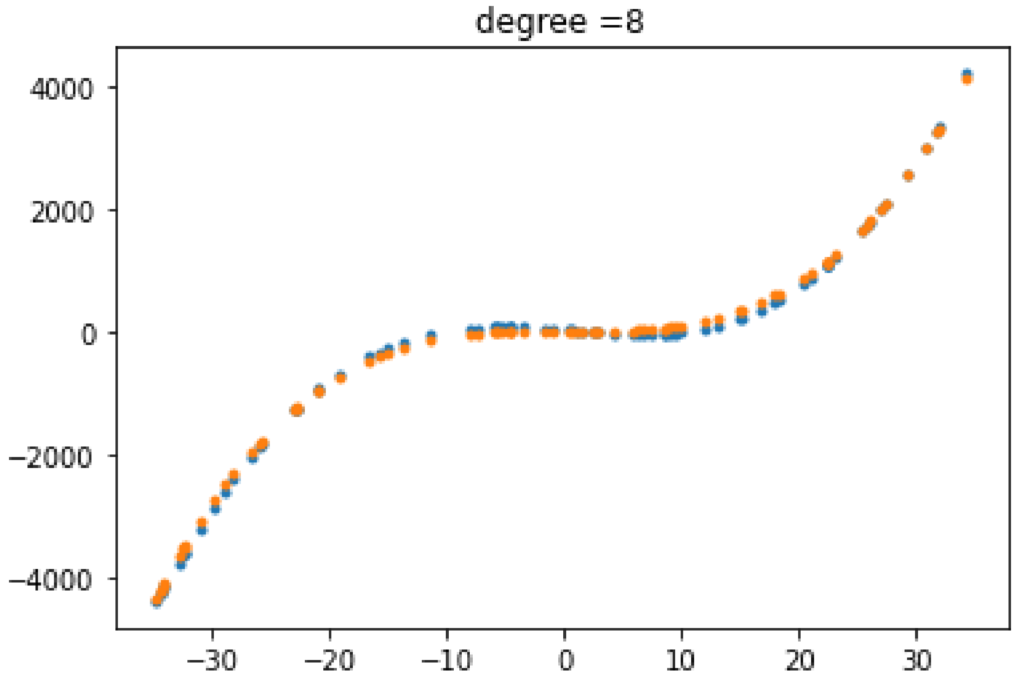


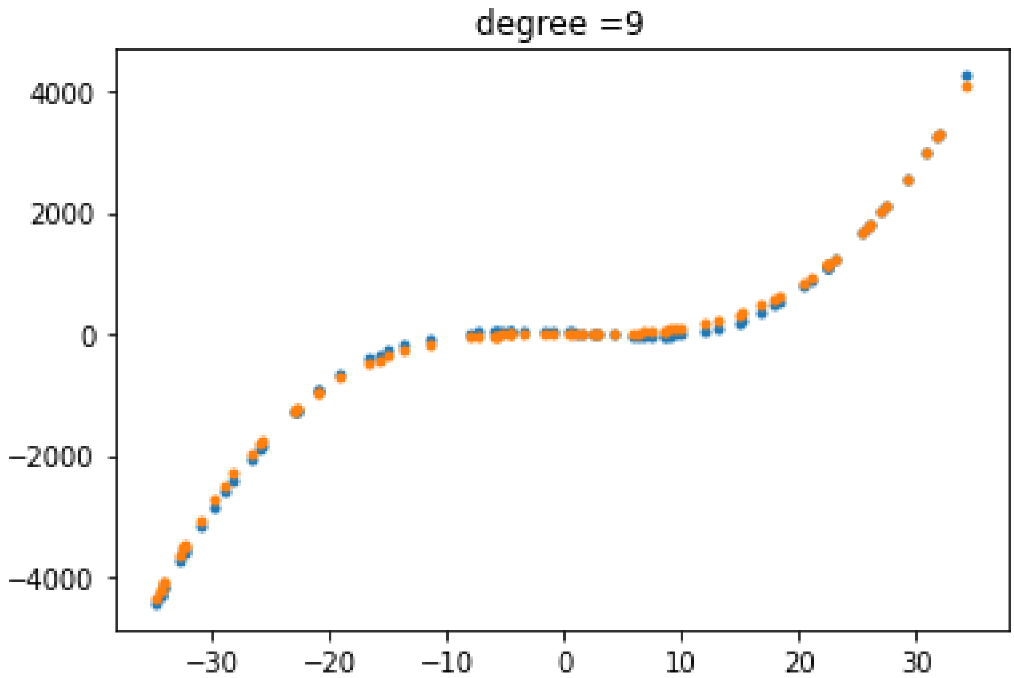
degree =3
4000
2000
0
−2000
−4000
−30
−20
−10
0
10
20
30

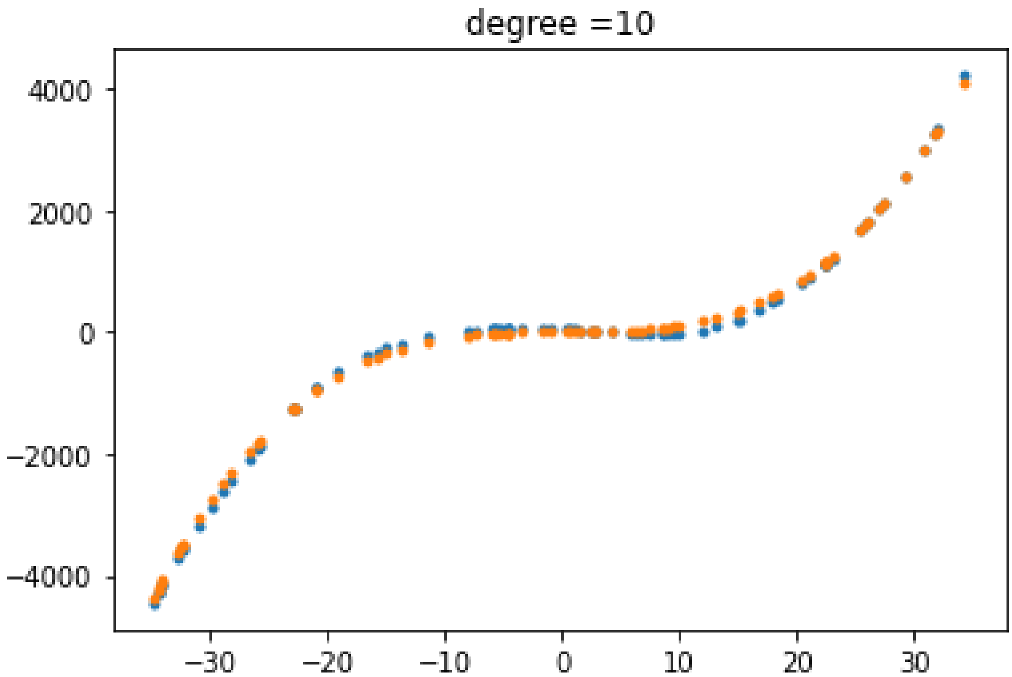

degree =4
4000
2000
0
−2000
−4000
−30
−20
−10
0
10
20
30
degree =5
4000
2000
0
−2000
−4000
−30
−20
−10
0
10
20
30

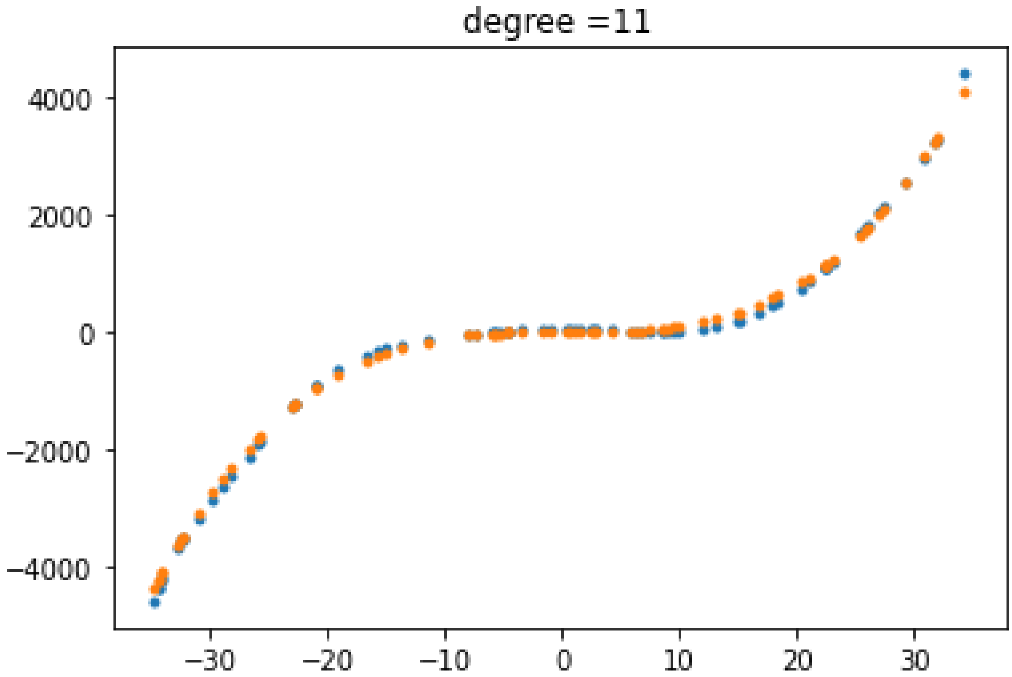

degree =6
4000
2000
0
−2000
−4000
−30
−20
−10
0
10
20
30

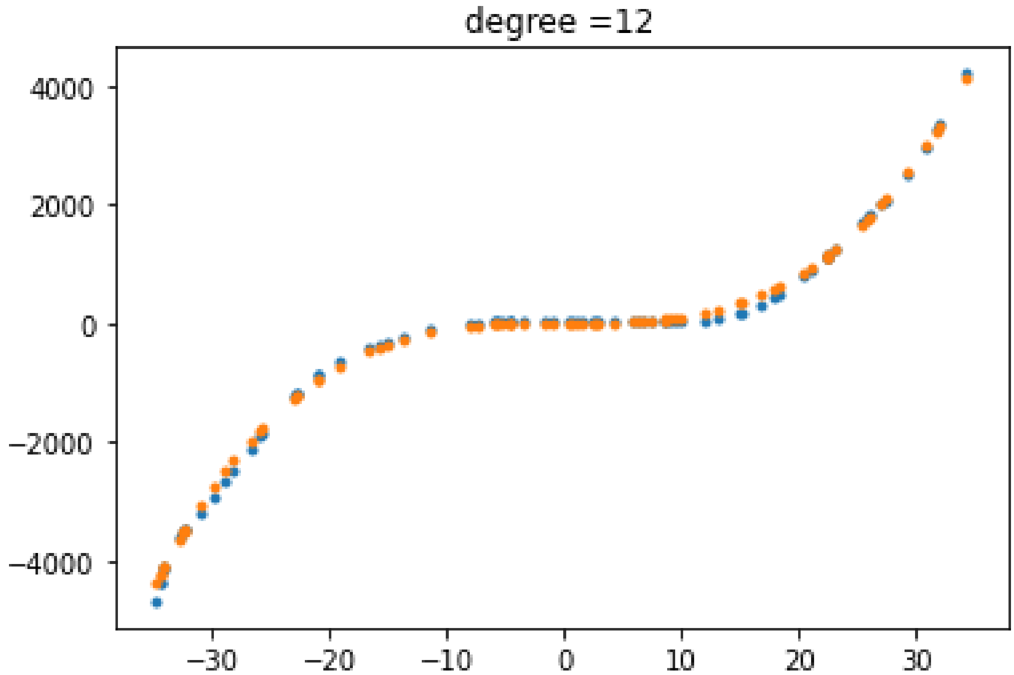

degree =7
4000
2000
0
−2000
−4000
−30
−20
−10
0
10
20
30

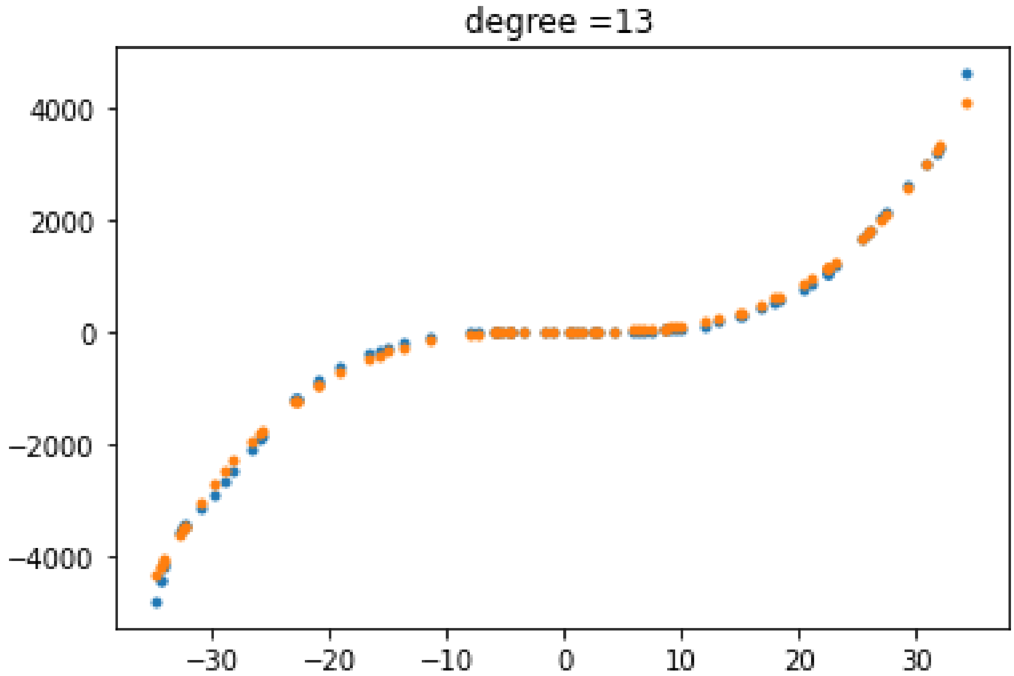

degree =8
4000
2000
0
−2000
−4000
−30
−20
−10
0
10
20
30

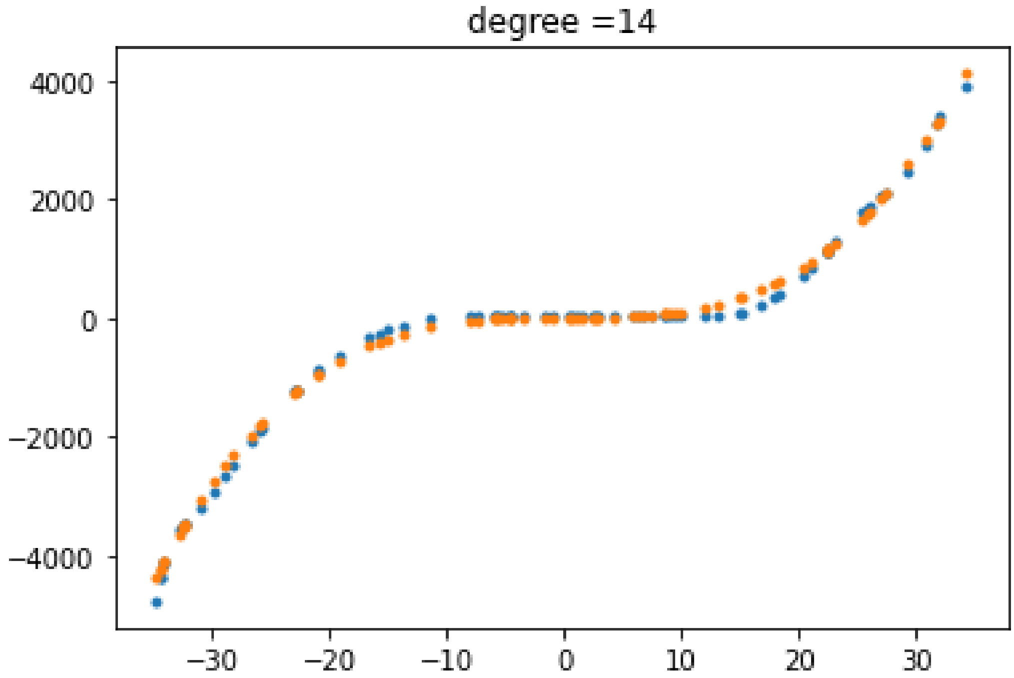

degree =9
4000
2000
0
−2000
−4000
−30
−20
−10
0
10
20
30

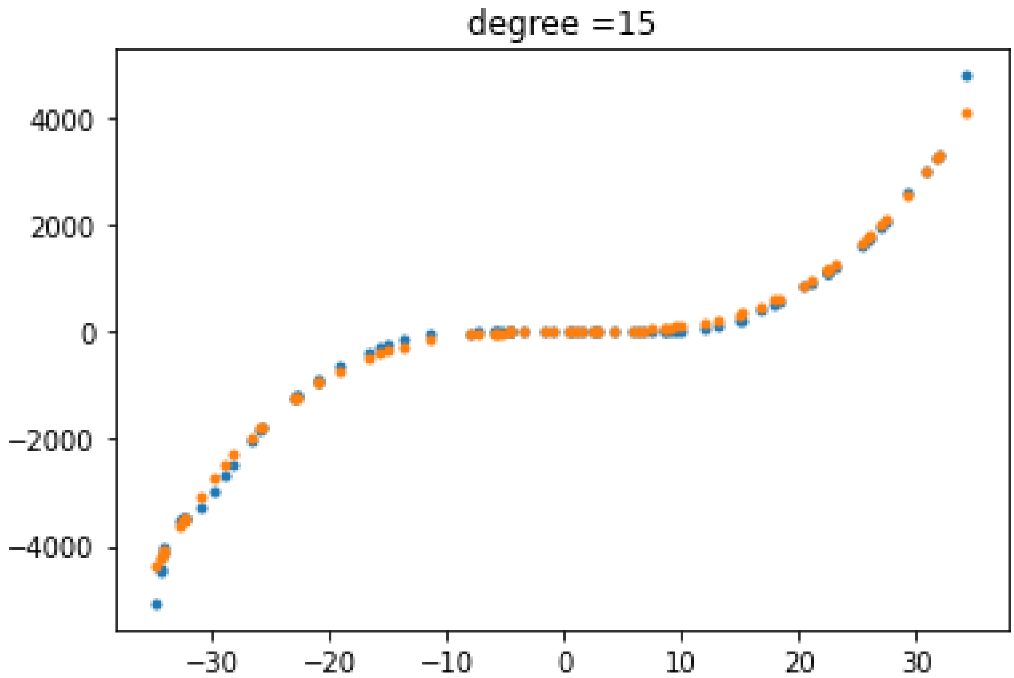

degree =10
4000
2000
0
−2000
−4000
−30
−20
−10
0
10
20
30


degree =11
4000
2000
0
−2000
−4000
−30
−20
−10
0
10
20
30


degree =12
4000
2000
0
−2000
−4000
−30
−20
−10
0
10
20
30


degree =13
4000
2000
0
−2000
−4000
−30
−20
−10
0
10
20
30


degree =14
4000
2000
0
−2000
−4000
−30
−20
−10
0
10
20
30


degree =15
4000
2000
0
−2000
−4000
−30
−20
−10
0
10
20
30

As can be seen from the below table, the value of Bias does not vary a lot across the function classes. The bias first decreases then remains almost same and then starts increasing again. On the other hand the value of variance keeps on increasing from degree 1 to degree 15.

When the model's bias and variance are balanced, the model fits well.

Rather of an underfitting or overfitting graph, we should want a trade-off between bias and variance. As a result of the tabulated bias and variance values, we infer that a function of degree 3 or 4 would be an almost perfect fit. We also determined in the previous part that a good match appears to be occurring for functions with degrees 3, 4, and 5.

	bias	variance
1	700.923	16572.866
2	683.129	26477.312
3	69.070	31577.247
4	67.901	55606.043
5	62.037	84365.951
6	52.335	102745.172
7	56.644	121998.006
8	54.896	139445.342
9	54.745	144142.748
10	59.143	154140.266
11	59.662	153595.054
12	98.002	149151.788
13	67.137	141919.723
14	146.205	149626.771
15	98.180	132682.488

Task 3: Calculating Irreducible Error

Irreducible error is the error that can't be reduced however good the model is implying that irreducible error is independent of the model.

For a good model, there should be a good balance between bias and variance for a minimal total error.

Mean squared error = Bias^2 + Variance + Irreducible error

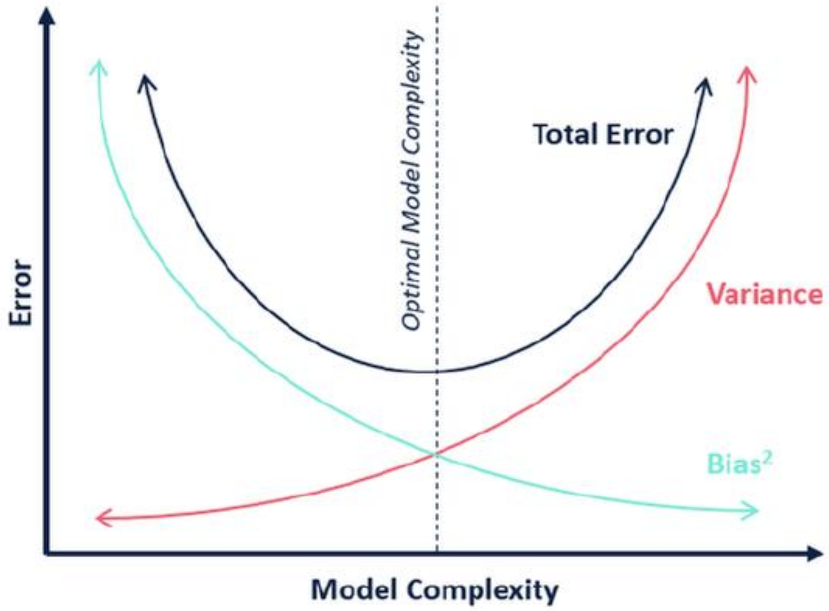
The irreducible error remains more or less the same, that is, close to zero across all the function classes.

	Irreducible error
1	8.003553e-11
2	1.455192e-11
3	7.275958e-12
4	1.455192e-11
5	1.455192e-11
6	1.455192e-11
7	-2.910383e-11
8	0.000000e+00
9	0.000000e+00
10	0.000000e+00
11	-2.910383e-11
12	5.820766e-11
13	-2.910383e-11
14	-2.910383e-11
15	-2.910383e-11

Task 4: Plotting $Bias^2$- Variance graph

The plot of $Bias^2$ and variance is shown below. As can be seen the value of $Bias^2$ decreases and the value of variance increases between the degree 1 and 3. Also from the MSE graph we can observe that the error is minimum at X=3.

A good model is one with a good balance between bias and variance or with a good trade-off between bias and variance that it minimizes the total error.



The value of Bias is very high and that of Variance is significantly small before X=3. After X=3 bias decreases remarkably and variance increases gradually. Hence before X=3, we have high bias and low variance which suggests an **underfitting model**.

After X = 3, variance increases slowly and bias remains almost constant till X = 13 and then increases. Therefore, after X = 3, the bias is low and the variance is comparatively high till X = 15, which points towards an **overfitting graph.**

Since, the error is the lowest at X = 3 it implies a good balance exists between bias and variance at X = 3.

Finally, the data is best suited by a model of degree three. This leads to the conclusion that **Y = f(X) is cubic.**