MDL Assignment 2
Report
Name: Pranav Gupta
Roll No.- 2021101095

Task 1:

LinearRegression().fit() is a Function which is defined in Scikit Learn Library which was developed by Google. This function is defined in LinearRegression Class of the Library. It takes 2 Parameters, range of values of x and range of values of y. It looks at values of x and y and corresponding to the Degree of the Polynomial finds the coefficients of all the terms of the Polynomial. It returns a Linear Model that minimises the Residual Sum of Squares between the range of values of y given as Testing Set and the Values predicted using the Values of x given as the Testing Set using the Model.

Task 2:

Gradient Descent is an Algorithm in Machine Learning which is used to find Optimal Coefficients satisfying a given condition. It minimises the MSE and calculates the Gradient(Slope) of the Given Cost Function. The Coefficients is Updated by moving in the Direction of Negative Gradient in an Attempt to Minimise the Total Error since the Gradient points up the Slope and not Down to it.

Task 3:

Degree of Polynomial	Bias	Variance	Mean Squared Error	Irreducible Error
1	0.26871836729048143	0.008497813770900723	0.1232071917341779	-6.886852199627925e-17
2	0.08605576887140859	0.0012093414814370307	0.013400382579330157	5.778363898478745e-17
3	0.033371663718274325	0.0003368845057376074	0.00504092852545256	-1.9232486154867334e-16
4	0.025076454411252443	0.0004251158604658545	0.004681399713658434	3.6301799240245904e-18
5	0.025236787898970884	0.0004877320093486093	0.00473197796425213	1.507661725498599e-16
6	0.024588723366002378	0.0010112663127589145	0.005249558685586174	3.093342639329255e-16
7	0.02612092268152255	0.0016693587509950753	0.006043996443317383	3.377021427254656e-16
8	0.027803234494635988	0.0037165422378080677	0.008289759910284224	-7.370395526534734e-16
9	0.02510707358089367	0.008089235188090042	0.012364419429823761	3.717366041725012e-16
10	0.02570044828968239	0.011044891996993012	0.015303508699854185	8.630857801453923e-16
11	0.028692068236282908	0.019945887304916043	0.024469986916644926	8.968612527984753e-16
12	0.05930013173658381	0.16802006436697625	0.19487262985054288	9.009138649939547e-16
13	0.07534749384257973	0.5438295847178355	0.6034181871502564	8.20781431102452e-16
14	0.11685165066241489	2.2274558842517673	2.399618168717727	7.419051809530453e-16
15	0.11874364204699965	2.303234625127227	2.4471932690332627	9.046702189458023e-16

By running the Python Program attached Multiple Times, we observe that, in general, Bias first decreases slowly with the Increase in Degree of Polynomial

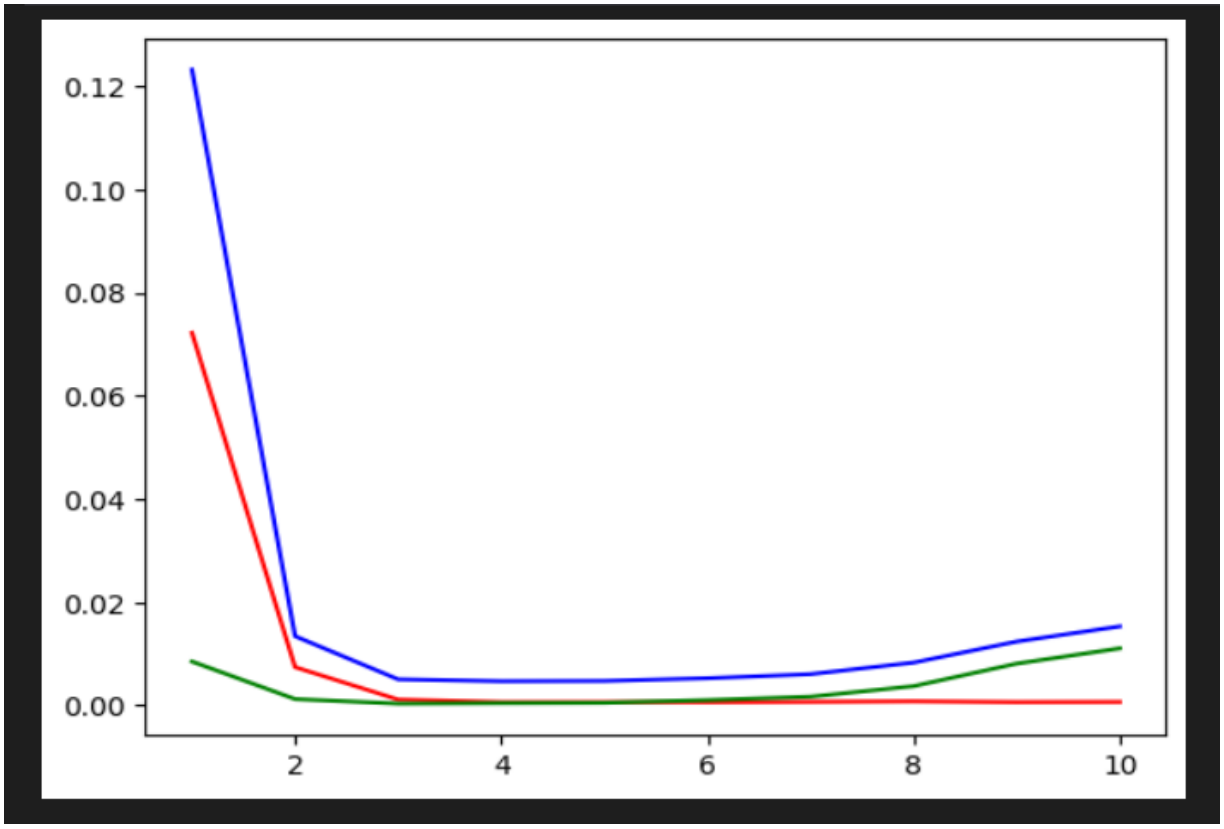
then Increases while Variance first decreases, reaches a Minima and then Increases Sharply. As Degree increases from 1 to 15, Bias Dips Slightly Overall, and Variance rises Rapidly. Since Bias measures the Error Induced by making Incorrect Assumptions and Variance measures Error Induced by Small Fluctuations in Training Set, as degree of Polynomial Increases, Model becomes more Complex and fits the Data Points accurately and closely, reducing Bias while at the same time Increasing Scope of Sensitivity to Noise present in the Data, thereby Increasing Variance. But because we are making all the Measurements in the Test Set, we are getting this Behaviour (as mentioned above). Also, the Behaviour of Bias and Variance is largely dependent on the given Datasets. If the Values are Randomly selected, then we get to see large Values of Bias and Variance while if values in the Dataset are close to each other, then we see relatively Smaller Values of Bias and Variance in conformation to Basic Principles of Probability nd Statistics.

Task 4:

Degree of Polynomial	Bias	Variance	Mean Squared Error	Irreducible Error
1	0.26871836729048143	0.008497813770900723	0.1232071917341779	-6.886852199627925e-17
2	0.08605576887140859	0.0012093414814370307	0.013400382579330157	5.778363898478745e-17
3	0.033371663718274325	0.0003368845057376074	0.00504092852545256	-1.9232486154867334e-16
4	0.025076454411252443	0.0004251158604658545	0.004681399713658434	3.6301799240245904e-18
5	0.025236787898970884	0.0004877320093486093	0.00473197796425213	1.507661725498599e-16
6	0.024588723366002378	0.0010112663127589145	0.005249558685586174	3.093342639329255e-16
7	0.02612092268152255	0.0016693587509950753	0.006043996443317383	3.377021427254656e-16
8	0.027803234494635988	0.0037165422378080677	0.008289759910284224	-7.370395526534734e-16
9	0.02510707358089367	0.00808923518809042	0.012364419429823761	3.717366041725012e-16
10	0.02570044828968239	0.011044891996993012	0.015303508699854185	8.630857801453923e-16
11	0.028692068236282908	0.019945887304916043	0.024469986916644926	8.968612527984753e-16
12	0.05930013173658381	0.16802006436697625	0.19487262985054288	9.009138649939547e-16
13	0.07534749384257973	0.5438295847178355	0.6034181871502564	8.20781431102452e-16
14	0.11685165066241489	2.2274558842517673	2.399618168717727	7.419051809530453e-16
15	0.11874364204699965	2.303234625127227	2.4471932690332627	9.046702189458023e-16

The irreducible error in a machine learning model is the error that cannot be reduced by improving the model, as it is inherent to the data itself. By running the Python Program attached Multiple Times, we observe that, in general, as the Degree of the Polynomial increases the Irreducible Error stays more or less the same. Sometimes Irreducible Error turns out to be negative as well but that too is order of 10^(-17). So, because the Absolute Values of Irreducible Error is very very Small, because of Minor Fluctuations it is sometimes becoming positive while sometimes it is becoming negative. It does not directly depend on Degree of Polynomial but in Case of Overfitting, it may Increase with the Degree of the Polynomial. Since, in our Program, Overfitting does not occur, Irreducible Error Stays almost the same as Degree of Polynomial Increases and is therefore almost Independent.

Task 5:



This is the Graph obtained for Task 5 of the Assignment. Blue Line signifies Mean Squared Error, Red Line signifies Variance and Green Line signifies Bias^2 Values. Detailed Explanation about the Observed Values of Bias and Variance is provided in the Preceding Parts.
We observe that The Optimum Degree is around 4. Also, we note that the value of Bias^2 and Variance intersect each other at Degree=4, after which the Value of Variance overtakes the Value of Bias^2. Also, we note that there is some Noise which is present in the Data, because of which Minor Fluctuations lead to some Disruption in Observed Values of Bias^2 and Variance as explained above.