Task 1

LinearRegression

1.Linear regression is a statistical regression method which is used for predictive analysis.

2.It is one of the very simple and easy algorithms which works on regression and shows the relationship between the continuous variables.

3.It is used for solving the regression problem in machine learning.

4.Linear regression shows the linear relationship between the independent variable (X-axis) and the dependent variable (Y-axis), hence called linear regression.

5.If there is only one input variable (x), then such linear regression is called simple linear regression. And if there is more than one input variable, then such linear regression is called multiple linear regression.

LinearRegression().fit()

LinearRegression fits a linear model with coefficients a = (a_1,a_2,a_3..................) to minimize the residual sum of squares between the observed targets in the dataset, and the targets predicted by linear approximation.

The class `sklearn.linear_model.LinearRegression` will be used to perform linear and polynomial regression and make predictions accordingly.

```
model = LinearRegression()
```

This statement creates the baraiable model as the instance of LinearRegression .

```
model.fit(x, y)
```

Here x is two dimensional array and y is one dimensional array.

With fit() you calculate the optimal values of the weights (intercept and coefficients). Once you have your model fitted, you can get the results to check whether the model works satisfactorily and interpret it.

The attributes of model are `.intercept_` , which represents the coefficient, and `.coef_` . `.intercept_` is scalar and `.coef_` is an array.

When applying `.predict()` , you pass the regressor as the argument and get the corresponding predicted response.

We can perform single variable linear regression as well as multiple variable regression. Example of single variable is plot for area and prices and multiple variable example is plot of [area,bedroom,age] as x and prices as y.

Task 2

Calculating Bias and Variance

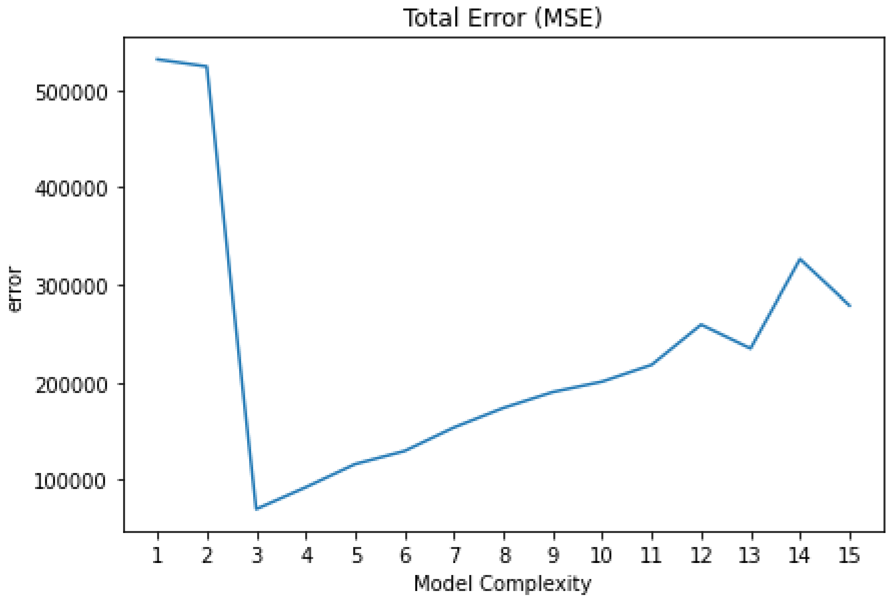
BIAS

The bias is a measure of how close the model can capture the mapping function between inputs and outputs. It captures the rigidity of the model, the strength of the assumption the model has about the funtional form of the mapping between inputs and outputs. Technically, we can define bias as the error between average model prediction and the ground truth. Moreover, it describes how well the model matches the training data set: 1. A model with a higher bias would not match the data set closely. 2. A low bias model will closely match the training data set. Signs of High Bias : 1. Failure to cature data trends 2. Underfitting 3. Overly simplified 4. High error rate The bias is always positive.

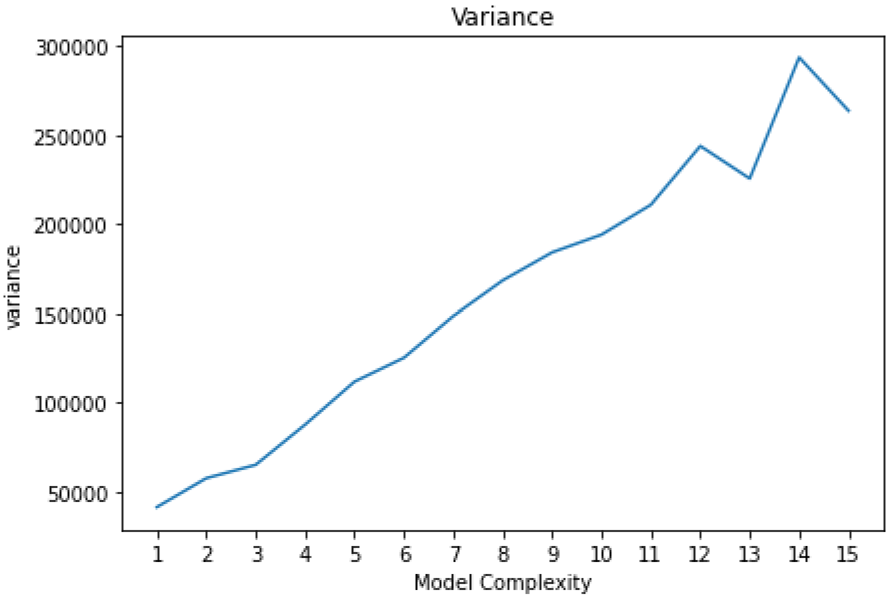
VARIANCE

The variance of the model is the amount the performance of the model changes when it is fit on different training data. It captures the impact of the speciies the data has on the model. A model with high variance will change a lot with small changes to the training dataset. Conversely, a model with low variance will change little with small or even large changes to the training dataset. Low Variance: Small changes to the model with changes to the training dataset. High Variance: Large changes to the model with changes to the training dataset. Signs of High variance: 1. Noise in the data set 2. Potential towerds overfitting 3. Complex models 4. Trying to put all data points as close as possible. The variance is always positive.

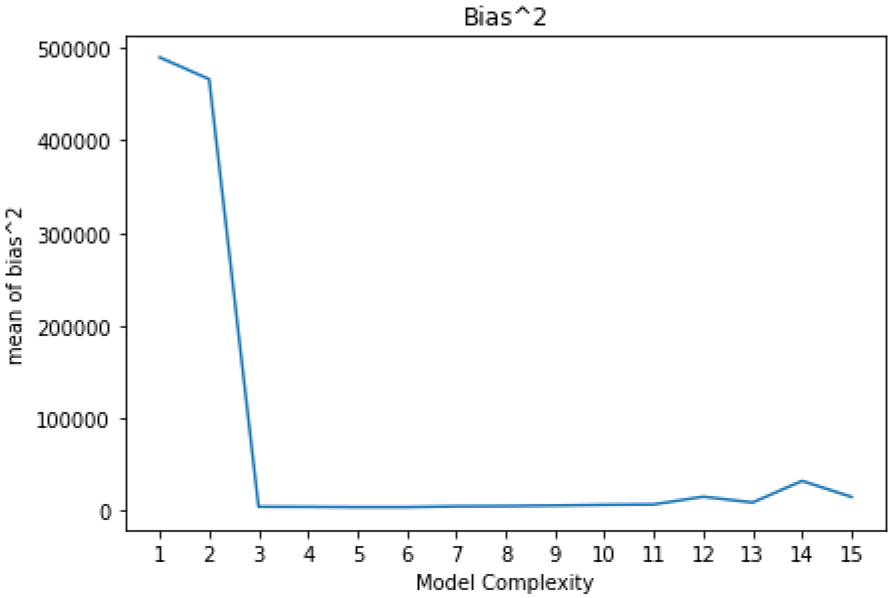
```
+----------+----------+-----------+------------+
|   Degree |      MSE |    Bias^2 |   Variance |
+==========+==========+===========+============+
|        1 | 531098   | 489775    |    41323   |
+----------+----------+-----------+------------+
|        2 | 523819   | 466255    |    57564   |
+----------+----------+-----------+------------+
|        3 |  69395.1 |   4323.2  |    65071.9 |
+----------+----------+-----------+------------+
|        4 |  91813.4 |   4176.54 |    87636.9 |
+----------+----------+-----------+------------+
|        5 | 115657   |   3876.96 |   111780   |
+----------+----------+-----------+------------+
|        6 | 129094   |   3902.1  |   125192   |
+----------+----------+-----------+------------+
|        7 | 153333   |   4759.1  |   148574   |
+----------+----------+-----------+------------+
|        8 | 173393   |   4994.45 |   168399   |
+----------+----------+-----------+------------+
|        9 | 189731   |   5605.39 |   184125   |
+----------+----------+-----------+------------+
|       10 | 200474   |   6392.54 |   194081   |
+----------+----------+-----------+------------+
|       11 | 217713   |   6841.57 |   210872   |
+----------+----------+-----------+------------+
|       12 | 259012   |  15151.3  |   243860   |
+----------+----------+-----------+------------+
|       13 | 234472   |   8888.79 |   225583   |
+----------+----------+-----------+------------+
|       14 | 325870   |  32238.6  |   293632   |
+----------+----------+-----------+------------+
|       15 | 278611   |  14946.5  |   263665   |
+----------+----------+-----------+------------+
```



MSE

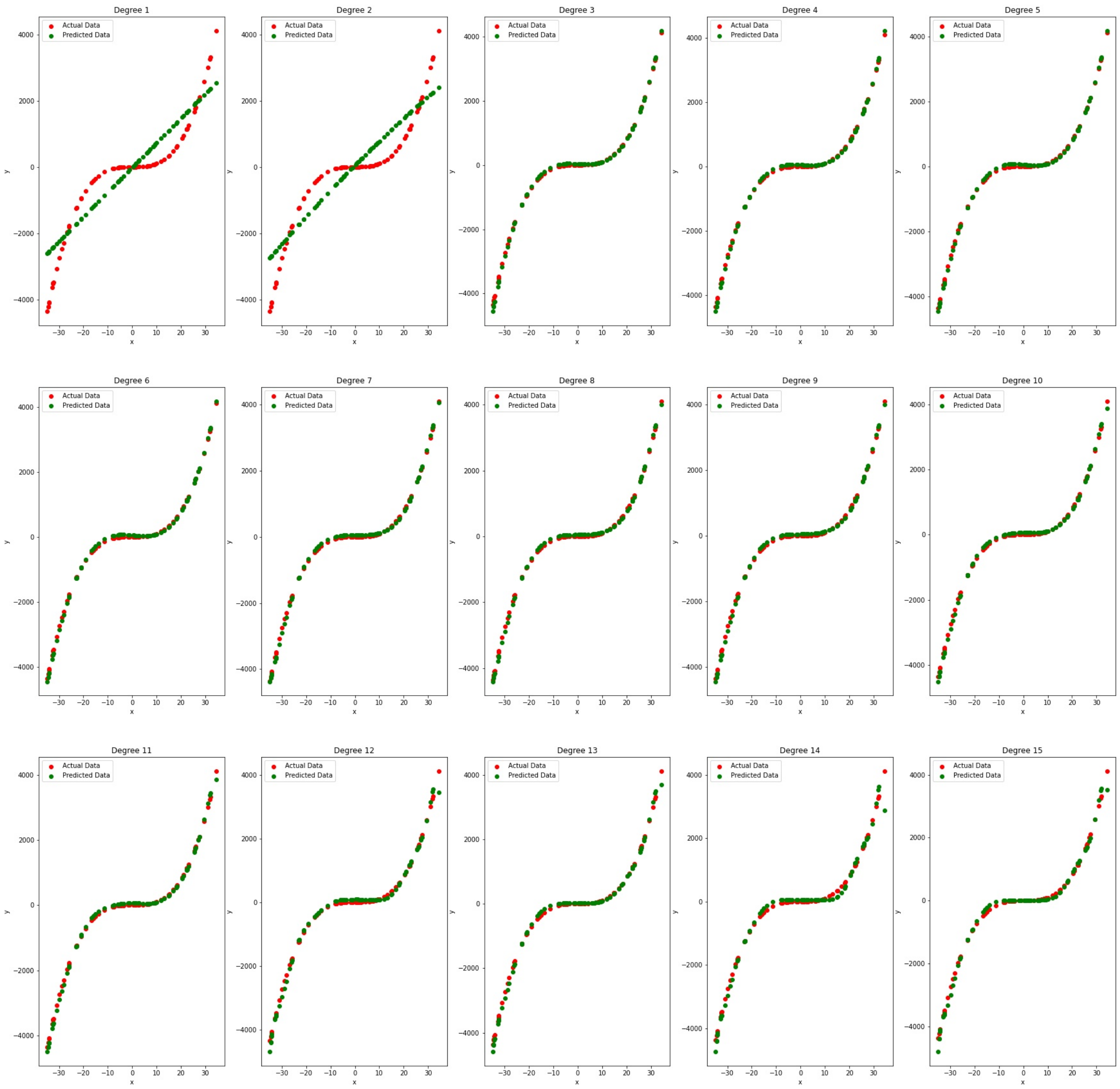


Variance



Squared Bias

These plots show how the model trained over the 16 partitions , and from these plots we can see the difference between the actual and predicted data upto 15 degree polynomial.



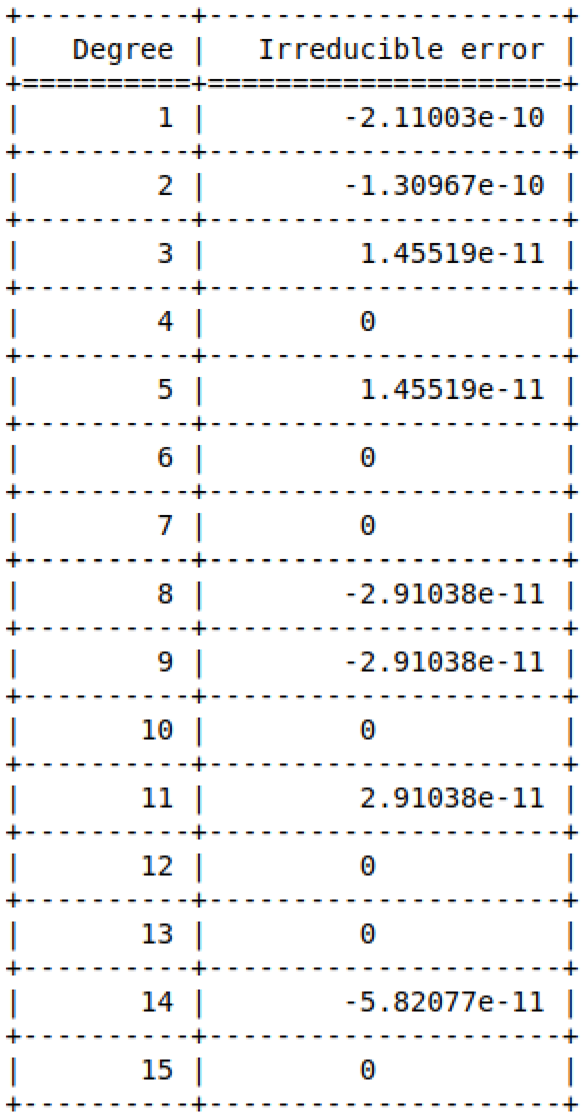
We can observe from above plots for degree of 1 or 2 , the model shows underfitting , for degree 3-11 the model is in best fit case, and degrees above it are case of overfitting.

Task 3

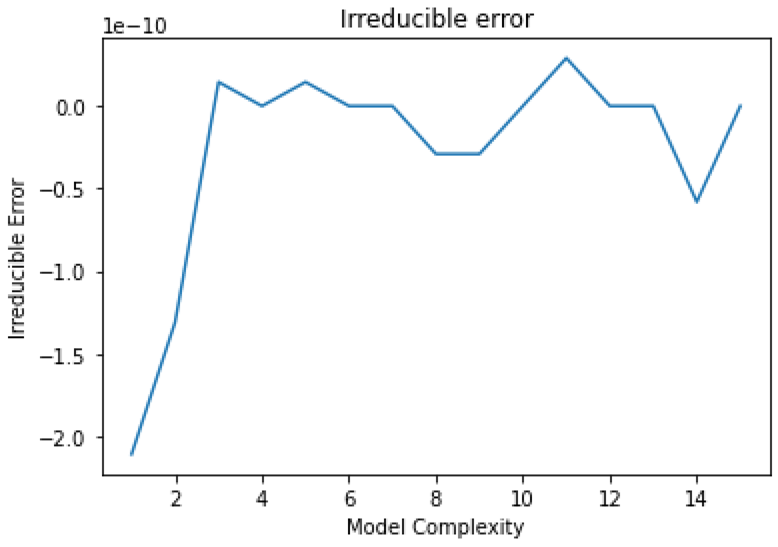
Calculating Irreducibile Error

IRREDUCIBLE ERROR

On the whole, the error of a model consists of reducible error and irreducible error. Model Error = Reducible Error + Irreducible Error The reducible error is the element that we can improve. It is the quantity that we reduce when the model is learning on a training dataset and we try to get this number as close to zero as possible. The irreducible error is the error that we can not remove with our model, or with any model. The error is caused by elements outside our control, such as statistical noise in the observations. It is reminder that no model is perfect.



```
+----------+--------------------+
|   Degree |   Irreducible error |
+==========+====================+
|        1 |       -2.11003e-10 |
+----------+--------------------+
|        2 |       -1.30967e-10 |
+----------+--------------------+
|        3 |        1.45519e-11 |
+----------+--------------------+
|        4 |        0           |
+----------+--------------------+
|        5 |        1.45519e-11 |
+----------+--------------------+
|        6 |        0           |
+----------+--------------------+
|        7 |        0           |
+----------+--------------------+
|        8 |       -2.91038e-11 |
+----------+--------------------+
|        9 |       -2.91038e-11 |
+----------+--------------------+
|       10 |        0           |
+----------+--------------------+
|       11 |        2.91038e-11 |
+----------+--------------------+
|       12 |        0           |
+----------+--------------------+
|       13 |        0           |
+----------+--------------------+
|       14 |       -5.82077e-11 |
+----------+--------------------+
|       15 |        0           |
+----------+--------------------+
```



Irr Error = MSE - bias^2 - variance

The value of the irreducible error is very close to zero. The negative values are of the order 1e-11 becuase of floating point precision error. This error is not reducible when we train any model and test it on furtur cases. The noise is nearly zero in the dataset given of order 1e-11. It was expected to stay the same as the data is very structured , there is almost no irreducible error.

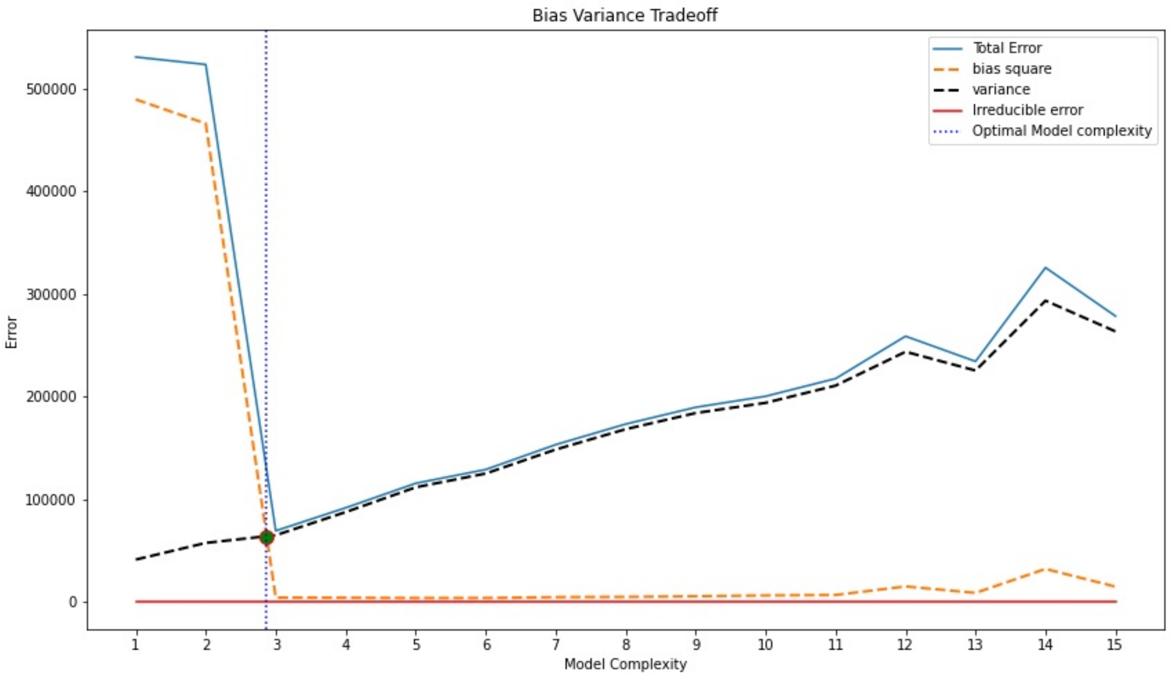
Task 4

Plotting Bias^2-Variance Graph

BIAS-VARIANCE TRADEOFF

The bias and the variance of a model's performance are connected. Ideally, we would prefer a model with low bias and low variance, although in practice, this is very challenging. In fact, this could be described as the goal of applied machine learning for a given predictive modeling problem, Reducing the bias can easily be achieved by increasing the variance. Conversely, reducing the variance can easily be achieved by increasing the bias. Bias and variance are inversely related to each other. We can choose a model based on its bias or variance. Simple models, such as linear regression and logistic regression, generally have a high bias and a low variance. Complex models, such as random forest, generally have a low bias but a high variance. High bias is not always bad, nor is high variance, but they can lead to poor results. We often must test a suite of different models and model configurations in order to discover what works best for a given dataset. A model with a large bias may be too rigid and underfit the problem. Conversely, a large variance may overfit the problem. We may decide to increase the bias or the variance as long as it decreases the overall estimate of model error.

Bias variance tradeoff for the task 2.



The point of intersection of the two curves or optimal model complexity is around 3 which is visible in the above graph.

From the above plot, we can see that below 3 , the bias is very high and variance is low, this makes it a case of underfitting. This shows that the model is not able to predict the values at all.

On the other hand , after degree 3 the bias is constantly low and variance is increasing which makes total error also increasing. This shows the model is picking even the small irregularities in the data and this is the case of overfitting.