

Report

MDL

Assignment 1

pAIn

Aparna Agrawal + Soveet Nayak



Write a brief about what function the method `LinearRegression().fit()` performs.

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

`model = LinearRegression()` creates the variable `model` as the instance of `LinearRegression`. You can provide several optional parameters to `LinearRegression`:

`model.fit(x, y)` calculates the optimal values of the weights b_0 and b_1 , using the existing input and output (x and y) as the arguments. In other words, `.fit()` fits the model. It returns itself, which is the variable `model` itself.



Tabulate the values of bias and variance and also write a detailed report explaining how bias and variance change as you vary your function classes.

```
Degree: 1
MSE: 533058.8281885165
VAR: 30819.371079803183
BIAS: 181.00131627457168
IRE: 0.0
```

Degree: 2
MSE: 526616.3391609614
VAR: 53993.29898506112
BIAS: 179.99120337406302
IRE: -1.1641532182693481e-10

Degree: 3
MSE: 60320.35452107467
VAR: 52836.01732782172
BIAS: 13.426254313186728
IRE: -5.4569682106375694e-12

Degree: 4
MSE: 82766.66746911078
VAR: 75710.59543364072
BIAS: 18.01138423770967
IRE: -1.9099388737231493e-11

Degree: 5
MSE: 116718.58564080132
VAR: 109867.33015604946
BIAS: 21.130834053162097
IRE: -1.7280399333685637e-11

Degree: 6
MSE: 123176.21668352601
VAR: 111273.17841677714
BIAS: 21.8686472512641
IRE: 1.0913936421275139e-11

Degree: 7
MSE: 146643.08205046545
VAR: 133608.62275827563
BIAS: 28.1734637234538
IRE: -1.0913936421275139e-11

Degree: 8
MSE: 153265.84724898366
VAR: 139374.0347671422
BIAS: 30.410382357866332
IRE: -1.8189894035458565e-11

Degree: 9
MSE: 180441.16946148925
VAR: 167409.13758883774
BIAS: 31.68188293144221
IRE: -3.2741809263825417e-11

```
Degree: 10
MSE: 206338.73374764994
VAR: 192579.10130006564
BIAS: 28.881972603214166
IRE: 0.0

Degree: 11
MSE: 252054.83308781227
VAR: 227795.26497119124
BIAS: 28.356073433559537
IRE: 3.637978807091713e-11

Degree: 12
MSE: 268332.28962305444
VAR: 228395.0312055635
BIAS: 33.267425398402466
IRE: -1.4551915228366852e-11

Degree: 13
MSE: 257744.47710103323
VAR: 239416.48929436025
BIAS: 15.909103321641236
IRE: 1.0913936421275139e-11

Degree: 14
MSE: 294885.8371829187
VAR: 223979.5410808011
BIAS: 22.472020850466528
IRE: 0.0

Degree: 15
MSE: 248090.55687284423
VAR: 227059.01685809018
BIAS: 14.888105551529183
IRE: 1.4551915228366852e-11
```

The bias and variance change when we vary the function classes. We observe bias is very high at the beginning but tends to decrease and become somewhat constant. Variance seems to increase with the increase in the degree of the function classes.



Tabulate the values of irreducible error for the models in Task 2 and also write a detailed report explaining why or why not the value of irreducible error changes as you vary your class function.

The values of irreducible error can be found on the .ipynb. Here is a particular run from the notebook. We observe that the irreducible error has the least value generally around Degree 3 and 4.



Plot the Bias² – Variance trade-off graph and write your observations in the report with respect to underfitting, overfitting and also comment on the type of data just by analyzing the Bias² – Variance plot.

Overfitting – High variance and low bias

When a model gets trained with so much data, it starts learning from the noise and inaccurate data entries in our data set. Then the model does not categorize the data correctly, because of too many details and noise.