Ridge and Lasso Regression

No
Date

→ Overfitting

Train Accuracy = 90%
Test Accuracy = 70%
example

→ model is biased in training dataset due which our test predicted accuracy is ~~reduced~~ decreasing

→ Underfitting

Train Accuracy = 60%
Test Accuracy = 62%

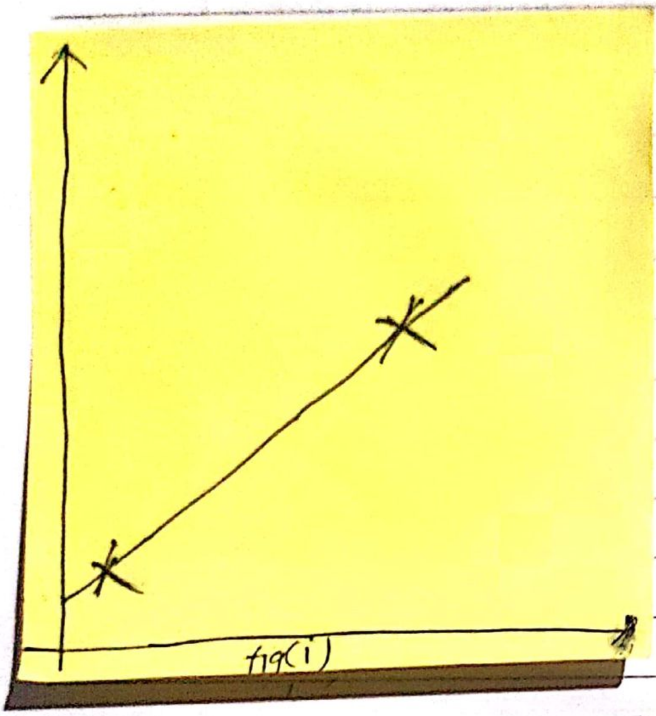
Overfitting	Underfitting
→ Low Bias (Train)	→ High bias (Train)
→ High variance (Test)	→ High Variance (Test)

Our main is to build Generalized model.
→ Low Bias (Train) eg: 90%
→ Low Variance (Test) Eg: 89%

> To prevent Overfitting, we use Ridge regression,
→ Ridge ~~lassi~~ regression is also called (L_2 regularization).

NISSAN

Ridge Regression



fig(i)

fig of

→ This Graph ~~has~~ of linear regression ~~(which is)~~ has ~~over~~ best fitted line which is overfitted with actual values.

→ for the given regression

Cost function:

$$= \frac{1}{2m} \sum_{i=1}^{m} (h_\theta(x)^{(i)} - (y)^{(i)})^2$$

↓

will be zero which indicates highly overfitted

Note: training data should not be overfitted.

Working of Ridge regression

→ When the model is overfitted with best fit line, Ridge regression adds two variables to the cost function which will be prevent overfitting condition.

$$\boxed{\frac{1}{2m} \sum_{i=1}^{m} (h_\theta(x)^{(i)} - y^{(i)})^2 + \lambda(\text{slope})^2}$$

for example

from the graph, fig(i), the value of cost function is zero because the residual error i.e difference between predicted value and actual value is zero. The model will be overfitted.

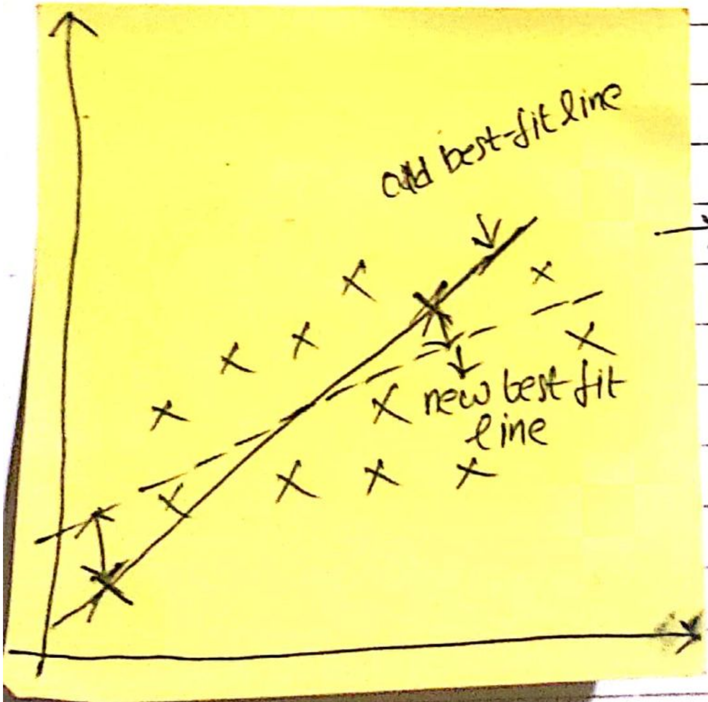
Now, we can use Ridge Regression i.e

$$\underbrace{C.F = \frac{1}{2m}\sum_{i=1}^{m}\left(h_\theta(x)^{(i)} - y^{(i)}\right)^2}_{0} + \lambda(\text{slope})^2$$

$$+ 1 \cdot (2)^2$$

$$= 4 \quad (\text{Assuming } \lambda = 1,\ \text{slope} = 2)$$

Now, we have to reduce C.F



Now,

$$= \text{C.F} + \lambda(\text{slope})^2$$

$$= \{\text{small value}\} + 1 \cdot (1.3)^2$$

$$\approx 2.05 \downarrow\downarrow$$

→ Which will cover most of the data points resulting the removal of overfitting.