

Regulation.....

Q What do you mean by Regularization?

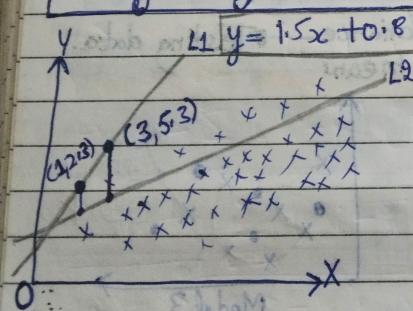
Ans Regularization is used to reduce the Overfitting problem you induce some added information in the machine learning model.

• Basically it is used to reduce the problem of Overfitting...

Types of Regularization techniques

- Ridge Regularization or L2 Regularization
- Lasso Regularization or L1 Regularization
- Elastic Net Regularization

Ridge Regularization



• As per the Graph, you can see that the Overfitting is there because, they are relying completely on the training data [L1].

• To select the [L2] fit line. We use the Loss Function.

$$\text{Loss} = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

In order to select the best fit line which is [L2] we have to add something to our Loss Function, which

$$\text{is } \lambda(m^2)$$

$$\text{where } \lambda = [0, \infty]$$

Therefore the Formula for Ridge Regularization will be:-

$$\text{Loss} = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 + \lambda(m^2)$$

Date _____

$$\text{Equation of L1} = [1.5x + 0.8], \text{ Equation of L2} = [0.9x + 1.5]$$

$$\begin{aligned} \text{Loss}_{L1} &= \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 + \lambda(m^2) \\ &\quad \boxed{\lambda=1} \text{ Assumption} \\ \Rightarrow & 0 + \lambda(1.5)^2 \\ \Rightarrow & \boxed{2.25} - \textcircled{I} \end{aligned}$$

$$\begin{aligned} \text{Loss}_{L2} &= \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 + \lambda(m^2) \\ &\quad \boxed{\lambda=1} \text{ Assumption} \\ \Rightarrow & (2.3 - 0.9(1) + 1.5)^2 + (5.3 - 0.9(3) - 1.5)^2 \\ \Rightarrow & (2.3 - 0.9 - 1.5)^2 + (5.3 - 2.7 - 1.5)^2 \\ \Rightarrow & (-0.100)^2 + (1.100)^2 \\ \Rightarrow & 1.00 + 1.210 \\ \Rightarrow & \boxed{2.21} \rightarrow \textcircled{II} \end{aligned}$$

So, from \textcircled{I} and \textcircled{II} we can say that Loss of \textcircled{II} is less, therefore L2 is more suitable fit line for the testing data.

Q. Why Ridge Regularization is also Known as L^2 Regularization?

As per the Ridge Regularization Formula, we add λm^2 with the Loss Function. When there is One Input Column. But, if we have 2 or 3 Input Columns, then, how would we do.

Then for every Input Column added, we generally add the coef's square value with the λ .

For example:- Let say we have 2 Input Columns.. For that formula will be-

$$\begin{array}{|c|c|c|} \hline X_1 & X_2 & Y \\ \hline \end{array}$$

→ $Y = m_1 X_1 + b_1 \uparrow \textcircled{I} + \lambda(X - \bar{X})^2$

→ $Y = m_2 X_2 + b_2$

$$\text{Loss Function} = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 + \lambda(m_1^2 + m_2^2)$$

• Since we are adding with the Square value of "m". therefore it is termed as L^2 Regularization..