Siddhardhan

Linear Regression - intuition

Linear Regression

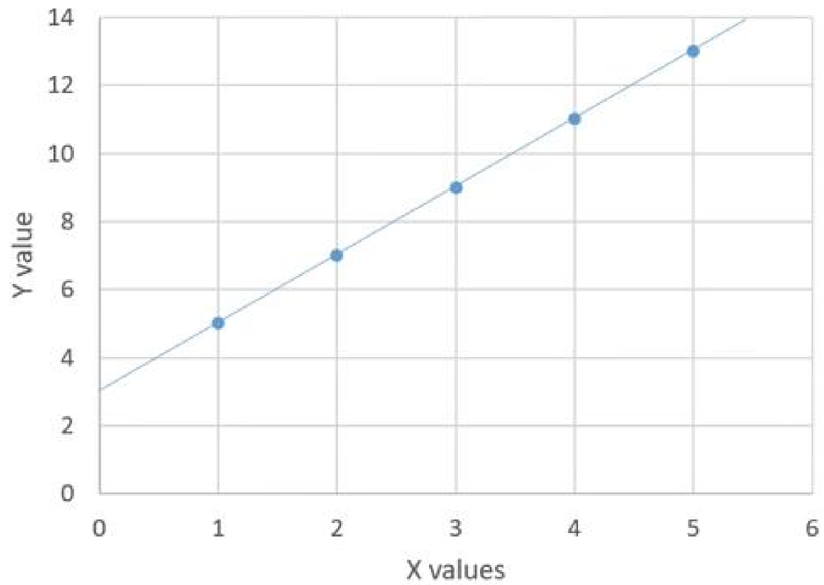
Experience in Years	0	2	4	5	6
Salary	2,00,000	4,00,000	8,00,000	10,00,000	12,00,000

What would be the **salary** of a person with **3 years of Experience?**

~ ₹ 650000 per Year

Linear Regression

X	1	2	3	4	5
Y	5	7	9	11	13

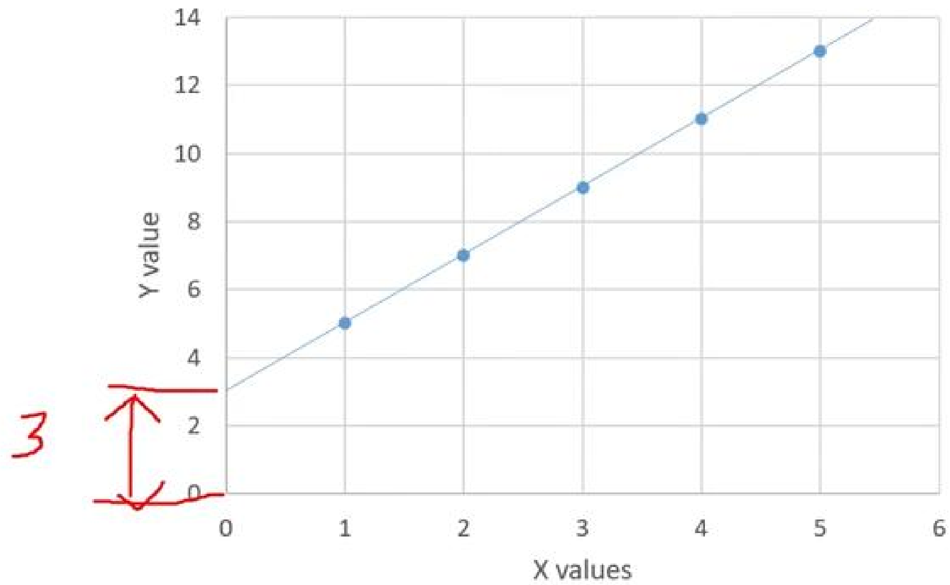


Y = mX + c

X --> X value
Y --> Y value
m --> Slope
c --> Intercept

Linear Regression

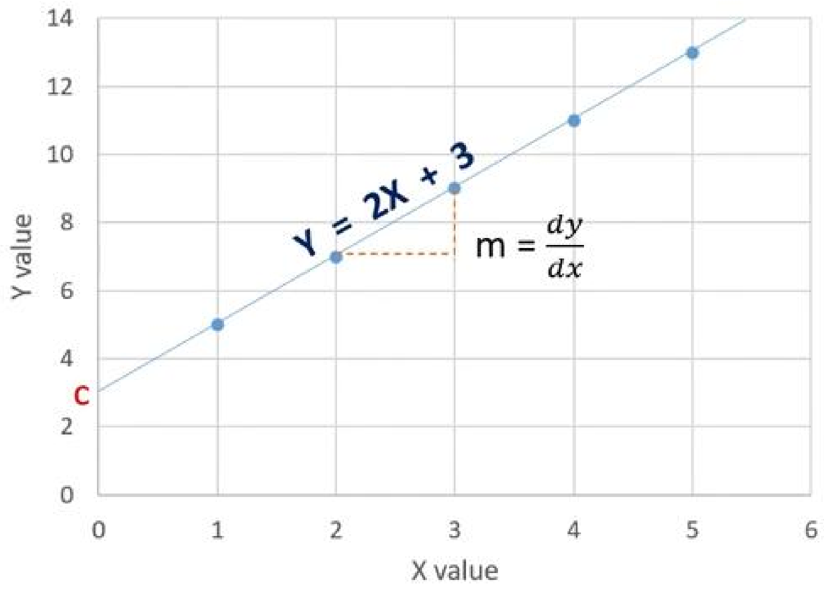
X	1	2	3	4	5
Y	5	7	9	11	13



Y = mX + c

X --> X value
Y --> Y value
m --> Slope
c --> Intercept

Linear Regression



Inference: The above Line equation is a function that relates X and Y. For a given value of X, we can find the corresponding value of Y

Equation of a Straight Line : **Y = mX + c**

Find the values of m and c:

Point P1 (2,7)
Point P2 (3,9)

Slope, m = $\frac{y_2 - y_1}{x_2 - x_1} = \frac{9 - 7}{3 - 2} = 2$

m = 2

Intercept, c:

Point (4,11)

Y = 2X + c

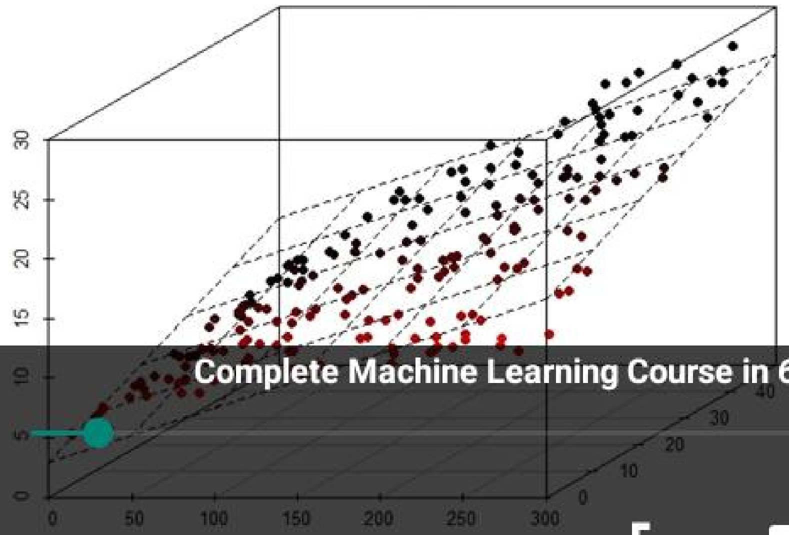
11 = 2(4) + c

c = 3

What if there are more than 2 Variables?

Multiple Linear Regression

Multiple linear regression is a model for predicting the value of one dependent variable based on two or more independent variables.



Simple Linear Regression $y = b_0 + b_1{*}x_1$

Multiple Linear Regression $y = b_0 + b_1{*}x_1 + b_2{*}x_2 + \dots + b_n{*}x_n$

Complete Machine Learning Course in 60 Hours - Part 3 | Full Machine Learning Course for Beginners

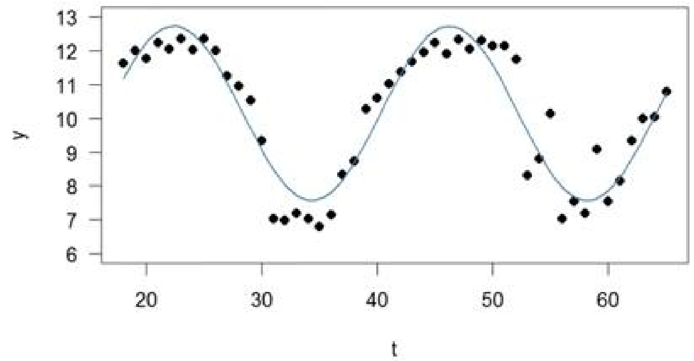
26:25 -11:25:33

HW 1.50X

Linear Regression

Advantages:

1. Very simple to implement
2. Performs well on data with linear relationship



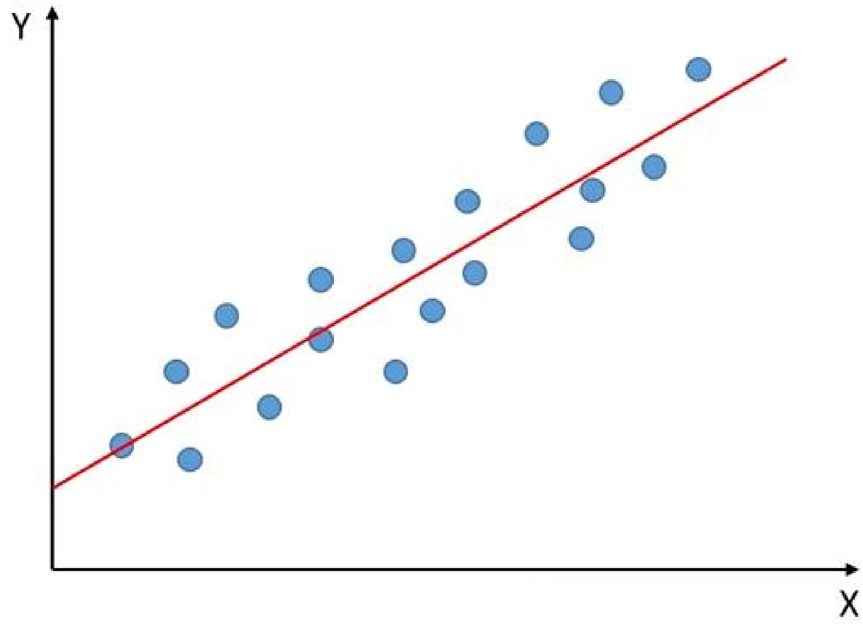
Disadvantages:

1. Not suitable for data having non-linear relationship
2. Underfitting issue

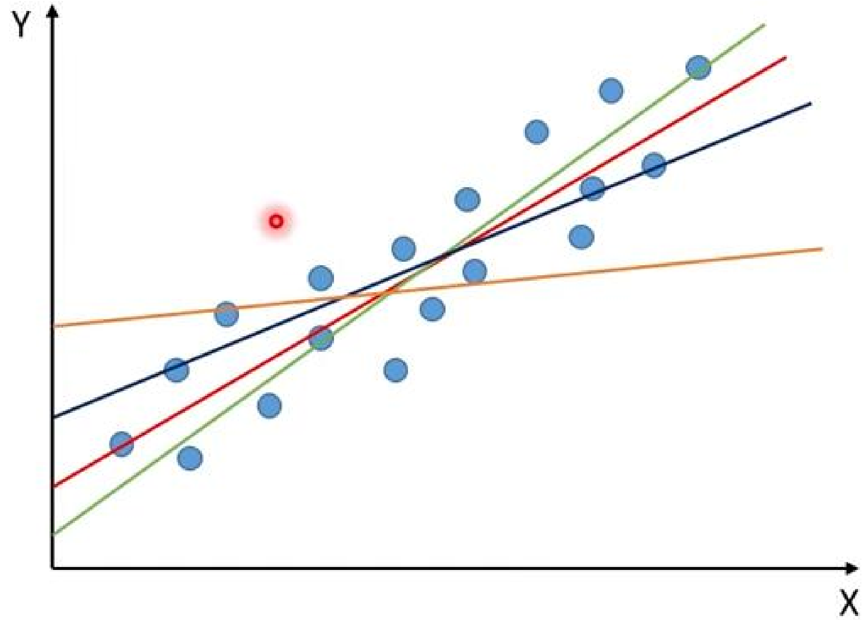
Siddhardhan

Linear Regression - Mathematical Understanding

Linear Regression



Linear Regression



Loss Function

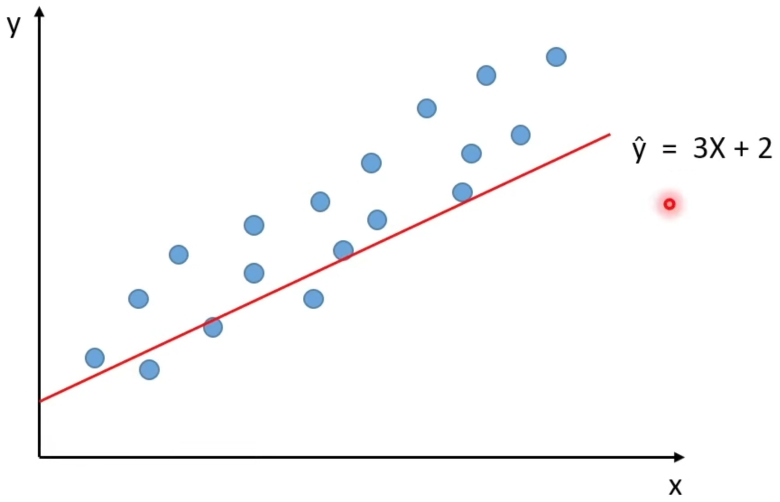
Loss function measures how far an estimated value is from its true value.

It is helpful to determine which model performs better & which parameters are better.

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

Loss Function

Randomly assigned Parameters: m = 3; c = 2



x	y	ŷ
2	10	8
3	14	11
4	18	14
5	22	17
6	26	20

Loss Function

x	y	ŷ
2	10	8
3	14	11
4	18	14
5	22	17
6	26	20

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

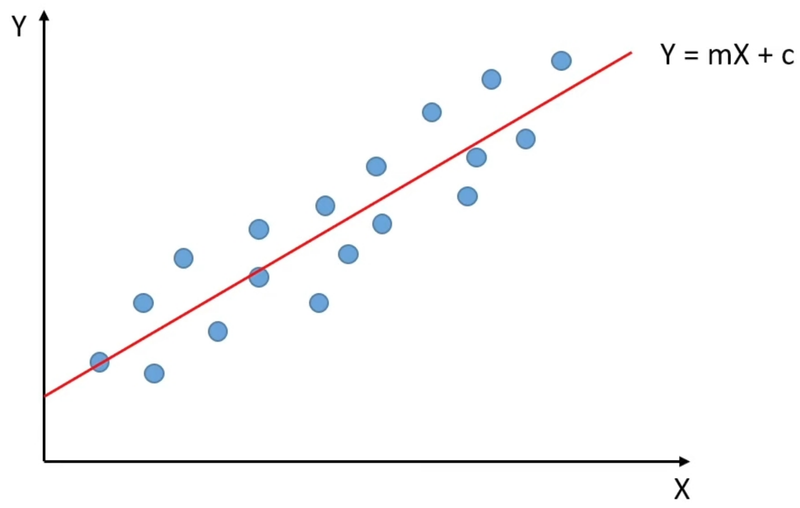
Loss = [$(10 - 8)^2 + (14 - 11)^2 + (18 - 14)^2 + (22 - 17)^2 + (26 - 20)^2$] / 5

Loss = [4 + 9 + 16 + 25 + 36] / 5

Loss = 18

Low Loss value → High Accuracy

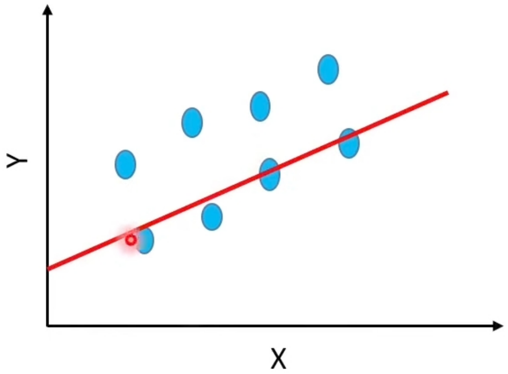
Linear Regression



Best Fit

Model Optimization

Optimization refers to determining best parameters for a model, such that the loss function of the model decreases, as a result of which the model can predict more accurately.

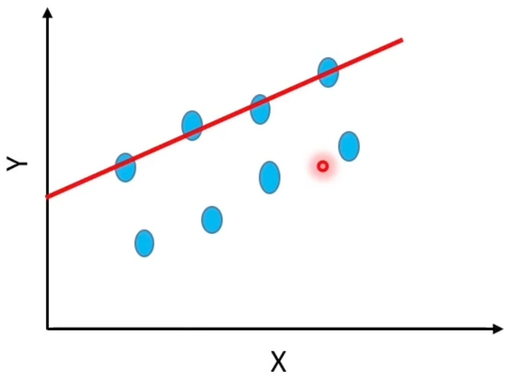


$Y = m_1X + C_1$

(m_1 & C_1 are the parameters of the line)

Model Optimization

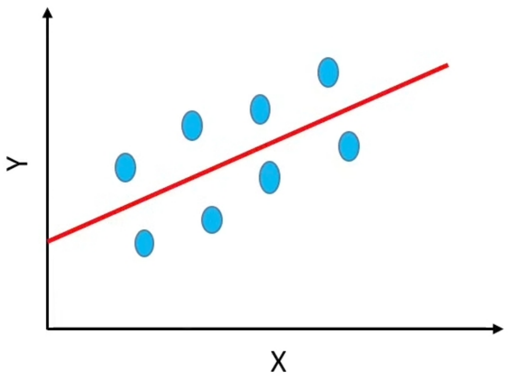
Optimization refers to determining best parameters for a model, such that the loss function of the model decreases, as a result of which the model can predict more accurately.



$$Y = m_2X + C_2$$

Model Optimization

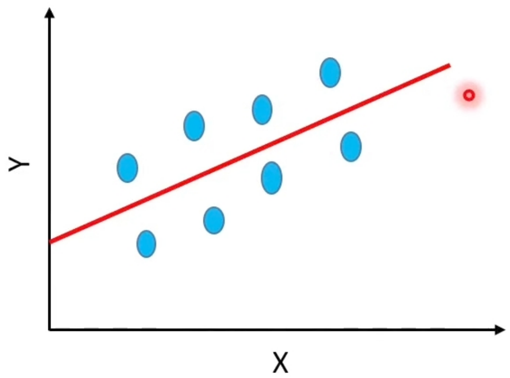
Optimization refers to determining best parameters for a model, such that the loss function of the model decreases, as a result of which the model can predict more accurately.



$Y = m_3X + C_3$

Model Optimization

Optimization refers to determining best parameters for a model, such that the loss function of the model decreases, as a result of which the model can predict more accurately.

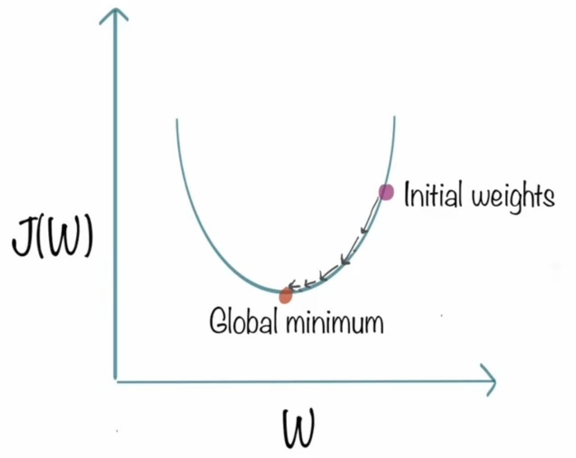


$Y = m_3X + C_3$

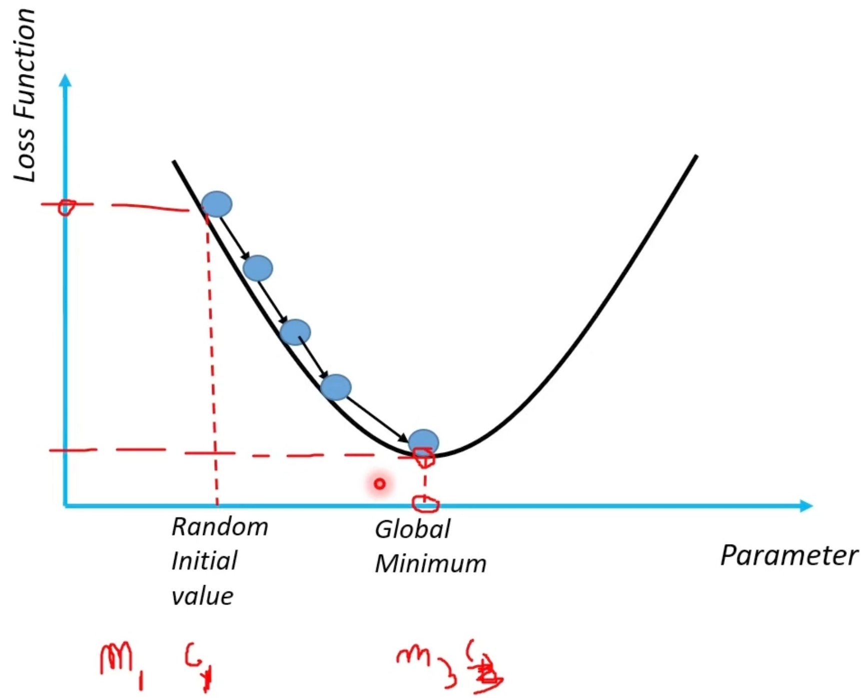
Hence, m_3 & C_3 are the best parameters

Siddhardhan

Gradient Descent for Linear Regression



Gradient Descent



Gradient Descent

Gradient Descent is an optimization algorithm used for minimizing the loss function in various machine learning algorithms. It is used for updating the parameters of the learning model.

$$m_2 = m_1 - LD_m$$

$$c_2 = c_1 - LDc$$

m --> slope
c --> intercept
L --> Learning Rate

D_m --> Partial Derivative of loss function with respect to m

D_c --> Partial Derivative of loss function with respect to c

Gradient Descent

$$D_m = \frac{\partial(Cost\ Function)}{\partial m} = \frac{\partial}{\partial m}\left(\frac{1}{n}\sum_{i=0}^{n}(y_i - y_{i\,pred})^2\right)$$

$$= \frac{1}{n}\frac{\partial}{\partial m}\left(\sum_{i=0}^{n}(y_i - (mx_i + c))^2\right)$$

$$= \frac{1}{n}\frac{\partial}{\partial m}\left(\sum_{i=0}^{n}(y_i^2 + m^2x_i^2 + c^2 + 2mx_ic - 2y_imx_i - 2y_ic)\right)$$

$$= \frac{-2}{n}\sum_{i=0}^{n} x_i(y_i - (mx_i + c))$$

$$= \frac{-2}{n}\sum_{i=0}^{n} x_i(y_i - y_{i\,pred})$$

$$D_c = \frac{\partial(Cost\ Function)}{\partial c} = \frac{\partial}{\partial c}\left(\frac{1}{n}\sum_{i=0}^{n}(y_i - y_{i\,pred})^2\right)$$

$$= \frac{1}{n}\frac{\partial}{\partial c}\left(\sum_{i=0}^{n}(y_i - (mx_i + c))^2\right)$$

$$= \frac{1}{n}\frac{\partial}{\partial c}\left(\sum_{i=0}^{n}(y_i^2 + m^2x_i^2 + c^2 + 2mx_ic - 2y_imx_i - 2y_ic)\right)$$

$$= \frac{-2}{n}\sum_{i=0}^{n}(y_i - (mx_i + c))$$

$$\frac{-2}{n}\sum_{i=0}^{n}(y_i - y_{i\,pred})$$