2025-06-23

23 June 2025 10:17

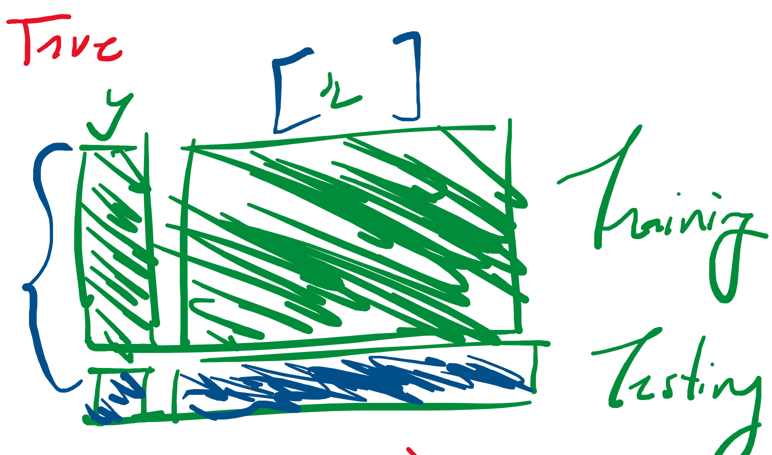
-> Linear Reg Model :

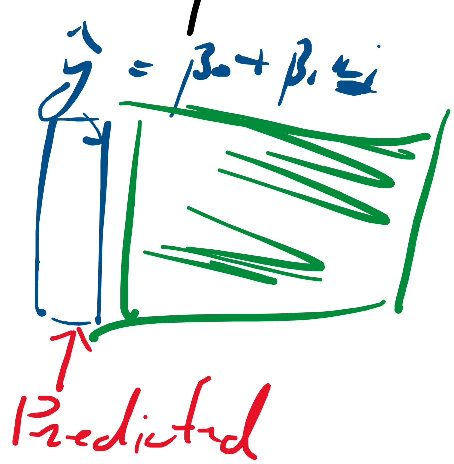
$$\hat{y} = \beta_0 + \beta_1 \cdot x_i$$

$$\hat{y} = 41.63 + 0.2076 \times x_i$$

=> Validation in Learning :

Loss Functions : Used to optimize the parameters.





(Loss)
Error = True Values | Predicted Values

=> Loss Metric in Regression

[Mean Squared Error]

$$MSE = \frac{1}{N} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

$$\text{Error} = (\underbrace{y_i}_{\text{True label}} - \underbrace{\hat{y}_i}_{\text{Predicted label}})$$

$$\text{Squared Error} = (y_i - \hat{y}_i)^2$$

$$\text{Sum of Squared Errors} = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

$$\text{Mean}(SSE) = \left[\frac{SSE}{N}\right]$$

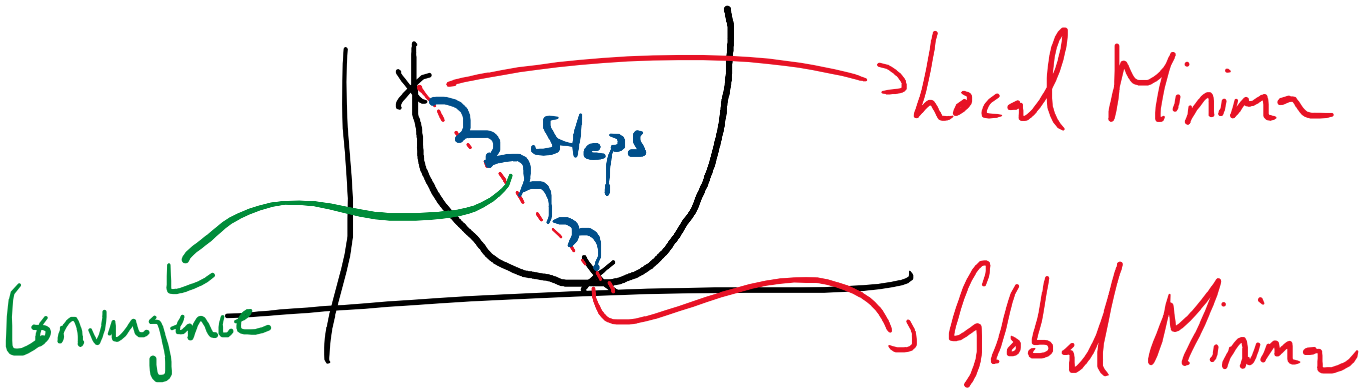
Note: This loss should be close to zero as much as possible.

=> Is there a way to reduce it.

=> Loss Optimization: Process of

=> Loss Optimization: Process of converging the loss.

=> Gradient Descent (Loss Optimizer)



$$\text{Loss} = \frac{1}{N} \sum_{i=1}^{n} \left(y_i - \hat{y}_i\right)^2$$

$$= \frac{1}{N} \sum_{i=1}^{n} \left(y_i - \{\beta_0 + \beta_1(x_i)\}\right)^2$$

Process:

S1 : Perform Convergence (Derivative)

S2 : Take learning rate (α) as an arbitrary value (Steps taken for convergence)

Eg: 0.0001

Eg: 0.0001 for convergence

S3: Multiply your derivative with the α & subtract it from original parameter.

β_1 & β_0 as parameters.

$$\partial/\partial\beta_1 = \frac{1}{N}\sum(y_i - \hat{y}_i)^2$$

$$= \frac{1}{N}\sum(y_i - \{\beta_0 + \beta_1 \cdot (x_i)\})^2$$

$$\partial/\partial\beta_1 = \frac{1}{N}\sum 2(y_i - \{\beta_0 + \beta_1 \cdot x_i\}) \times \partial/\partial\beta_1 (y_i - \{\beta_0 + \beta_1 \cdot x_i\})$$

$$\partial/\partial\beta_1 = \frac{2}{N}\sum(y_i - \{\beta_0 + \beta_1 \cdot x_i\}) \times -(x_i)$$

$$\partial/\partial\beta_1 = \left[-\frac{2}{N}\sum(\underline{y_i - \{\beta_0 + \beta_1 \cdot x_i\}}) \times (x_i)\right]$$

$$\partial/\partial\beta_0 = \frac{1}{N}\underline{\sum(y_i - \hat{y}_i)}^2$$

$$\partial/\partial\beta_0 = \frac{-2}{N}\sum(y_i - \hat{y}_i)$$

$$\partial/\partial\beta_0 = \left[\left(\frac{-2}{N}\sum\left(y_i - \{\beta_0 + \beta_1 \cdot x_i\}\right)\right)\right]$$

S3: $\{$

$$\beta_1 = \beta_1 - \left(\alpha \times \partial/\partial\beta_1 \text{ Loss}\right)$$

$$\beta_0 = \beta_0 - \left(\alpha \times \partial/\partial\beta_0 \text{ Loss}\right)$$

$\}$

=> Gradient Descent Implementation:

Epoch = n

$\{$

$$\partial/\partial\beta_1 = \frac{-2}{N}\sum(x_i) \times \left(y_i - \{\beta_0 + \beta_1 \cdot x_i\}\right)$$

$$\partial/\partial\beta_0 = \frac{-2}{N}\sum\left(y_i - \{\beta_0 + \beta_1 \cdot x_i\}\right)$$

$$\beta_1 = (\beta_1) - \left(\alpha \times \partial/\partial\beta_1 \cdot \text{Loss}\right)$$

$$\beta_0 = (\beta_0) - \left(\alpha \times \partial/\partial\beta_0 \cdot \text{Loss}\right)$$

$$\beta_0 = (\beta_0) - (\alpha \times \partial/\partial\beta_0 \cdot Loss)$$

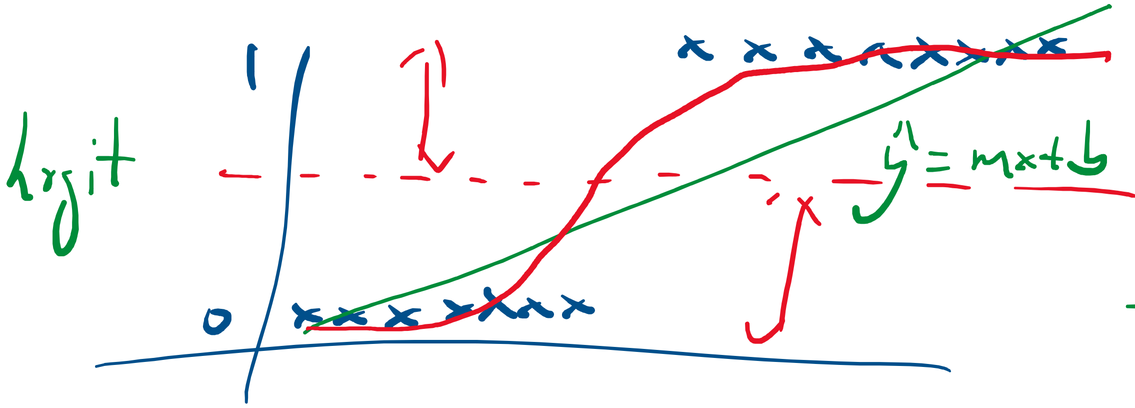
}

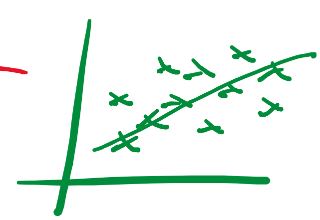
CLASSIFICATION: Discrete Outcomes always

⇒ No. of hours a student studies	Outcome in exams	
4.3	P	1
2.0	P	1
0.8	F	0

⇒ Logistic Regression:

Binary Classification / Supervised / Parametric





Logit / Sigmoid

$$\sigma(z) = \frac{1}{1+e^{-z}}$$

$(\hat{y})$

$$z = \beta_0 + \beta_1 \cdot x_i$$

$$\sigma(z) = \frac{1}{1+e^{-\{\beta_0 + \beta_1 \cdot x_i\}}}$$

Pred $\hat{y} = \sigma(z)$