Data → X, y → LR Algo → LR Model

LR Model → c, m, eq.

X_{train}, X_{test} → LR model → $\hat{y}$

$\hat{y} = m\underline{x} + c$

Error

y_{train}	$\hat{y}_{train}$
y_{test}	$\hat{y}_{test}$

$\boxed{MSE} = \frac{1}{N} \Sigma (y-\hat{y})^2$

$\cancel{MAE} = \frac{1}{N} \Sigma |y-\hat{y}|$

$\checkmark\ \boxed{RSME} = \frac{1}{N} \sqrt{\Sigma (y-\hat{y})^2}$

$\hat{y} = mx + c \pm \underline{RSME}$

SD → RMR

15.95 - 16.00

6

15.99

4

14

13.99

$\hat{y} = x + 9.99$

0.01

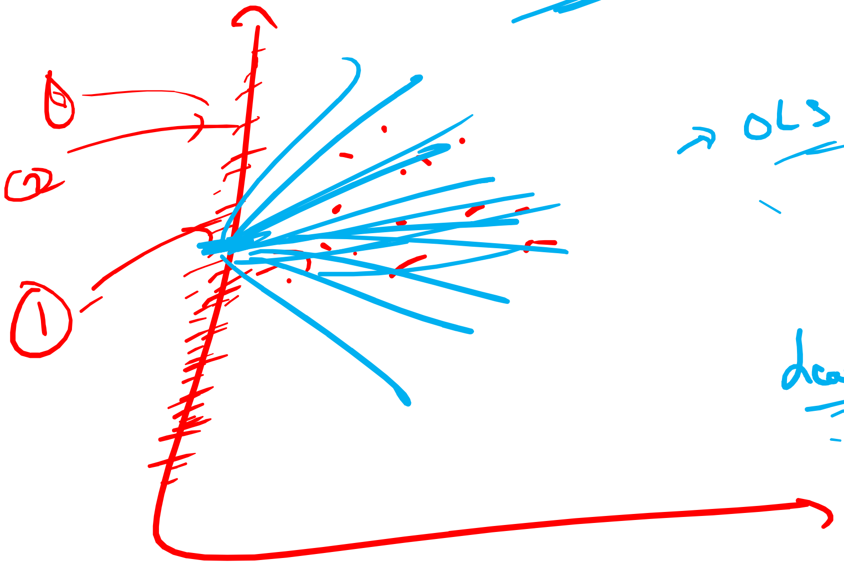
15.99

16

± 0.01

LR Algo → Gradient Descent

↳ OLS → Ordinary Least Sq



0 - 100

→ OLS

In life

hit & trial

you learn by making mistake

OLS

Stop

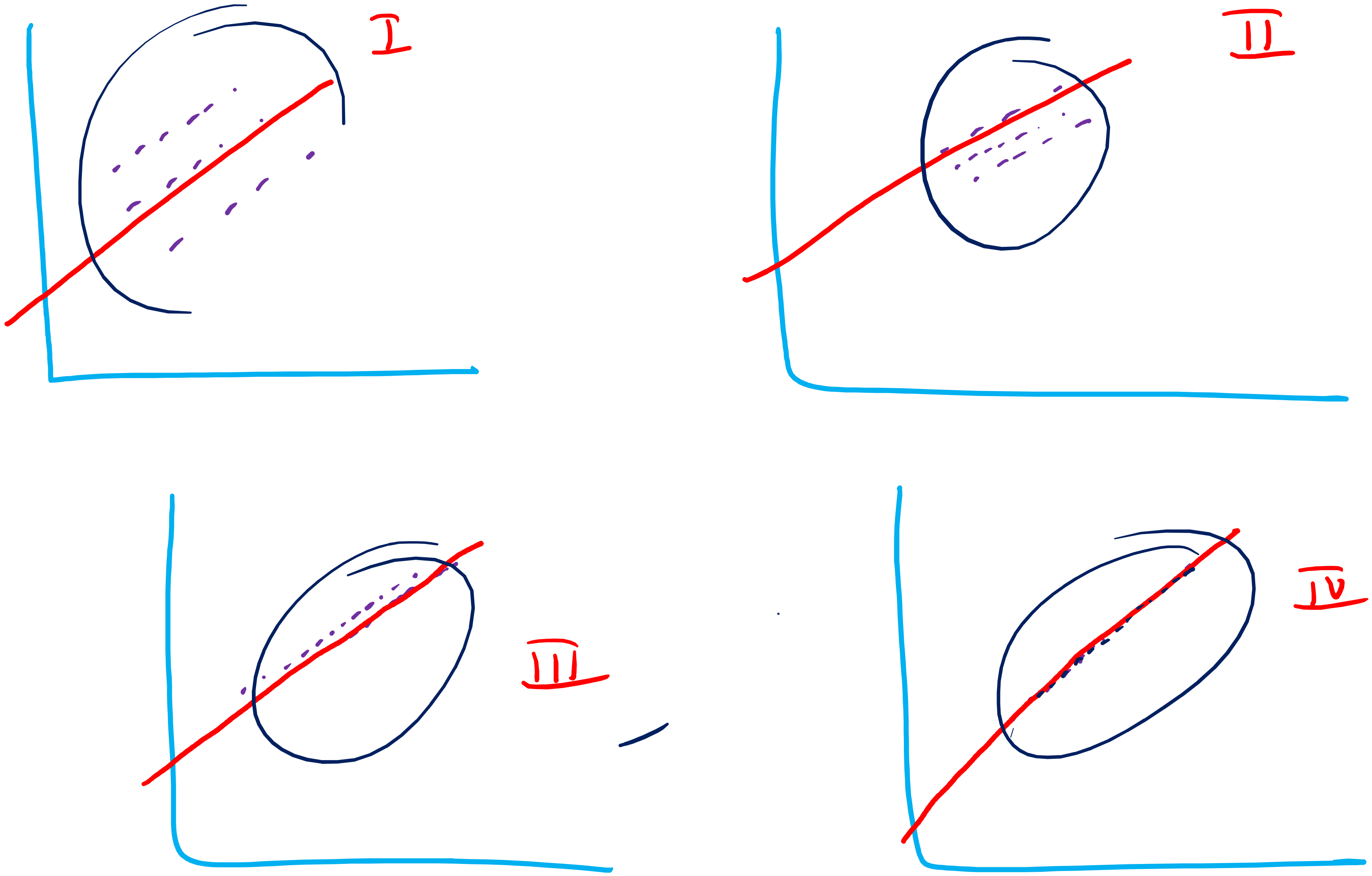
you learn from others & reduce the chance of mistake

G.D.

$$\text{Least Sq.} = \frac{Cov(x,y)}{Var(x)}$$

$$\Rightarrow \frac{\frac{\sum (x-\bar{x})(y-\bar{y})}{\cancel{n-1}}}{\frac{\sum (x-\bar{x})^2}{\cancel{n-1}}} = \frac{\sum (x-\bar{x})(y-\bar{y})}{\sum (x-\bar{x})^2}$$

Not that much Imp



Straight

$r, \rho \rightarrow$ corr

LR → Straight → coeff. of Regression | (R^2) | coeff. of Determination

$-1 \le r \le 1$

$\frac{SS}{N} = \sigma^2$

$\bar{x}$ $\bar{y}$ σ^2

$0 \le R^2 \le 1$

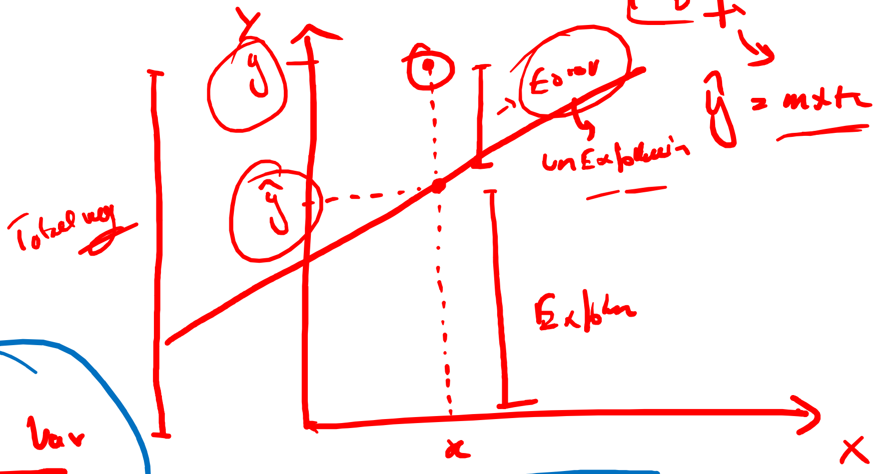
↓ Worst ↓ Best

$R^2 \boxed{\geq 0.7}$ → Then only, the model is of good quality model

$\bar{y}$ → Avg.

↓ LR model

$\hat{y} = mx + c$



$$R^2 = \frac{\text{Explained Var}}{\text{Total variation}}$$

$$\Rightarrow \frac{\text{Total var} - \text{UnExplained var}}{\text{Total variation}} \Rightarrow 1 - \frac{\text{Unexplained Var}}{\text{Total variance}}$$

$$\Rightarrow \boxed{1 - \frac{\sum (y-\hat{y})^2}{\sum (y-\bar{y})^2} = R^2}$$

→ R^2 → quality of model ($\hat{y} = mx + c$)

↳ strength of Rel^n b/w (x, y)

10hrs → 2 hrs

c.D. →

→ Adj R^2

$d x_n$ → slope (+) (−ve)

data + LR

$$R^2 = 1 - \frac{SSE/N}{SST/N}$$

$$df_T = N - 1$$

$$df_E \Rightarrow N_{ex} - 1$$

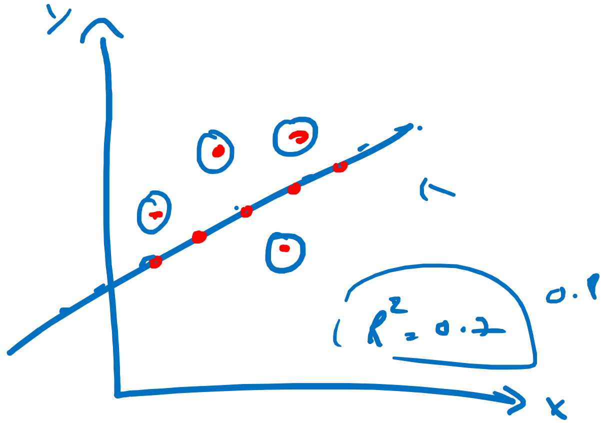
Penalty system

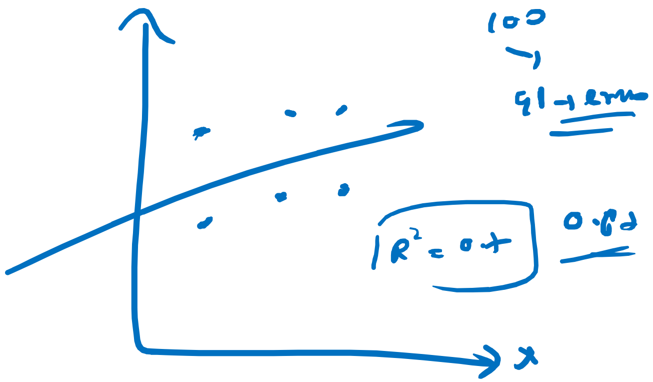
$$df_E \leq df_T$$

$$\text{Adj } R^2 = 1 - \frac{SSE/df_E}{SST/df_T}$$

$\bar{R}^2 \Rightarrow$ −ve

1





→ Use R^2 & Adj R^2 for feature Engg.

		R^2	Adj R^2
y →	x_1	0.8	0.74
+ →	x_2	0.82	0.8
+ →	x_3	0.67	0.6

→ good features

bad features

$$1 - \frac{10/5}{10/5} = 1 - 0$$

SSE = 10

SST = 10

N = 5

5 → 1

3 Predictor

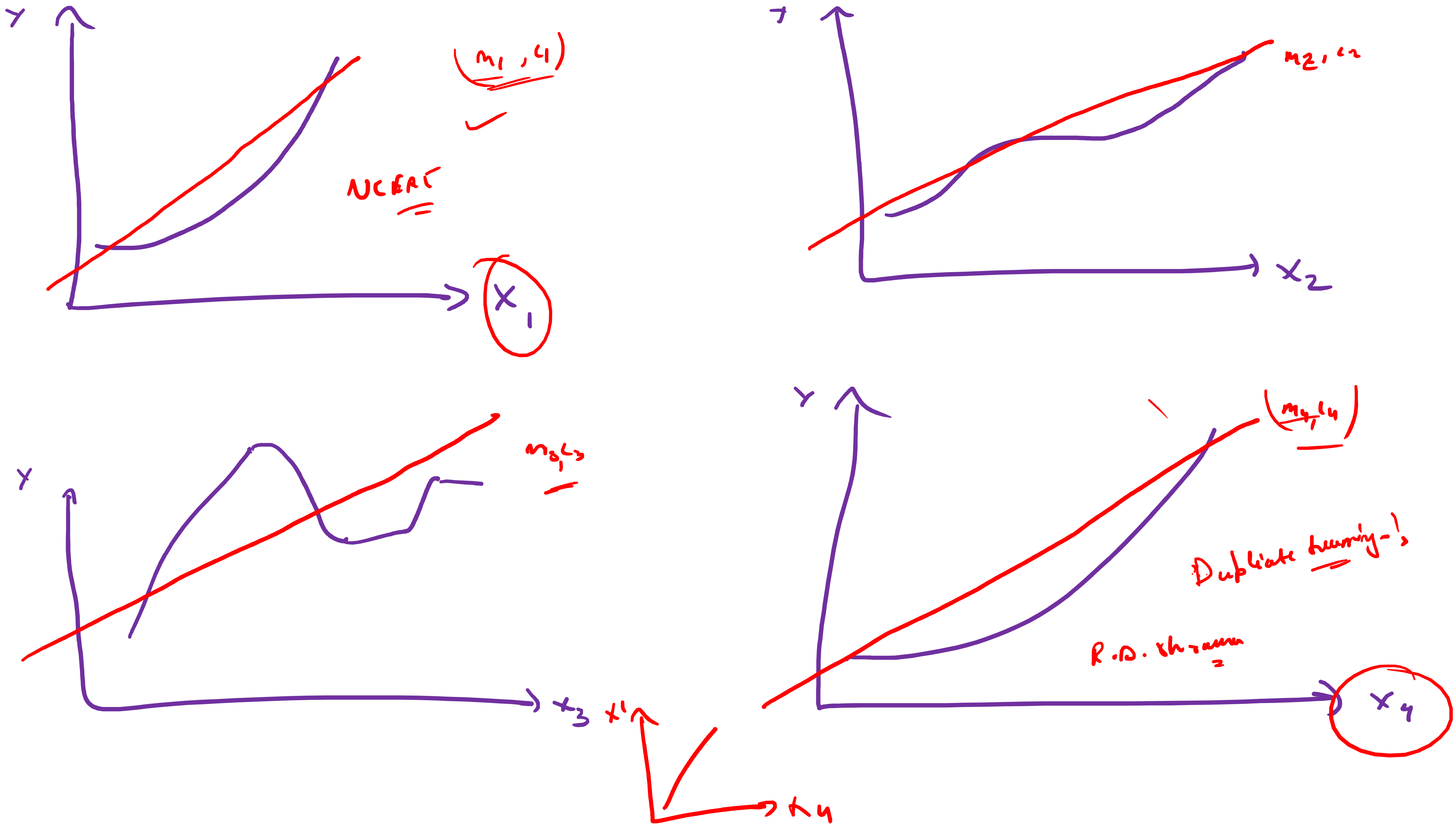
$$1 - \frac{SSE / df.error}{SST / d.f.T}$$

d.f. error = 2

d.f. Total = 4

$$1 - \frac{10/2}{10/4}$$

$$1 - \frac{5}{2.5} = 1 - 2 = -1$$

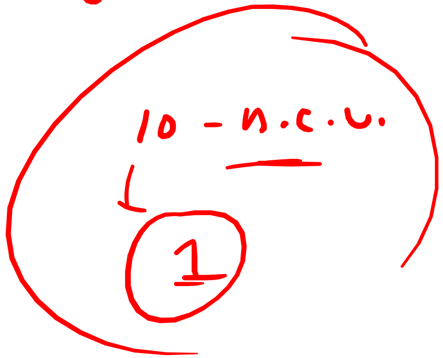

Y
(m_1, c_1)
NCERT
X_1
Y
m_2, c_2
X_2
Y
m_3, c_3
X_3
X'
X_4
Y
(m_4, c_4)
Duplicate learning
R.D. Sharma
X_4

→ Multi colinearity .

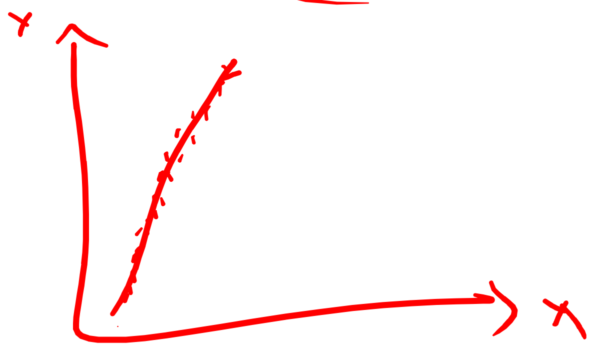
→ The model can learn from any 1 variable among N highly correlated var

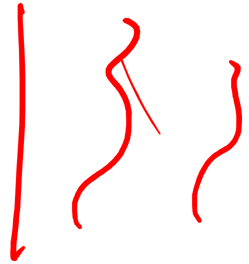
↓

Feature & Feature → No High corr



→ Target & Feature ↳ high corr is req.





→ Corr Matrix ⇒

$R^2 = 0.70$

$Adj R^2 = 0.68$

$>/ 0.7$

→ VIF (Variance Inflation Factor)

$VIF = \frac{1}{1-R^2} \rightarrow [1, \infty]$

$0 \le R^2 \le 1$

If $VIF \le 5$ → M.C. features exist.

Multicolinearity will Exist

If $VIF > 5$ → ✓ (g+g)

10 Features, 1 Target

F_1, F_7, F_8 → r/ρ ↑

↓

7 Feat. + 1 g

↓

8 Features

F_2, F_6, F_{10} → Low/none → Irrelevant

↓

5 features

↳ R^2, $Adj R^2$ → 3 features