

$$\boxed{y \leftarrow x_1 + x_2 + x_3 + x_4 + x_5} \rightarrow R^2, \text{ Adj. } R^2$$

~~Stepwise Regression~~

$$y \leftarrow \underline{x_1 + x_2 + x_3 + x_4} - \rightarrow R^2, \text{ Adj. } R^2$$

$$y \leftarrow x_1 + x_2 + x_4 + x_5 \rightarrow R^2, \text{ Adj. } R^2$$

$$y \leftarrow x_1 + x_3 + x_4 + x_5 \rightarrow R^2, \text{ Adj. } R^2$$

$$y \leftarrow x_2 + x_3 + x_4 + x_5 \rightarrow R^2, \text{ Adj. } R^2$$

$$y \leftarrow x_1 + x_2 + x_3 \rightarrow R^2, \text{ Adj. } R^2$$

⋮

$$y \leftarrow x_1 \rightarrow R^2, \text{ Adj. } R^2$$

Backward elimination

Stepwise regression

backward elimination

forward selection

$$T \begin{pmatrix} x_1, x_2, x_3, x_4, x_5 \end{pmatrix}$$

$$y \leftarrow x_1 \rightarrow \underline{\mathbb{R}^2}$$

$$y \leftarrow x_2 \rightarrow \underline{\mathbb{R}^2}$$

$$y \leftarrow x_3 \rightarrow \underline{\mathbb{L}}$$

$$y \leftarrow x_4 \rightarrow \underline{\mathbb{R}^2}$$

$$y \leftarrow x_1 + x_2 + x_3 + x_4 + x_5 \rightarrow \underline{\mathbb{R}^2}$$

①

$$y \leftarrow x_1 + 2x_2 + 3x_3 + 4x_4 + 5x_5 \rightarrow 0.9225 \quad \text{HhVaria}$$

②

$$\boxed{y \leftarrow x} \quad \leftarrow 0.9223 \quad \text{bw varia}$$

0.003 → ↑

Saw 10<sup>3</sup> ← ↑

$$\text{VIF} = \frac{1}{1-R^2} \rightarrow \underline{\lambda_m}$$

$x_1 \ x_2 \ x_3 \ x_4$

$$x_1 + x_2 + x_3 \xrightarrow{LR} \underline{x_4} \rightarrow R^2$$

$$x_1 + x_2 + \underline{x_3} \xrightarrow{LR} x_3 \rightarrow R^2$$

$$x_1 + x_3 + \underline{x_4} \xrightarrow{LR} x_2 \rightarrow R^2$$

$$x_2 + x_3 + \underline{x_4} \Rightarrow x_1 \rightarrow R^2$$

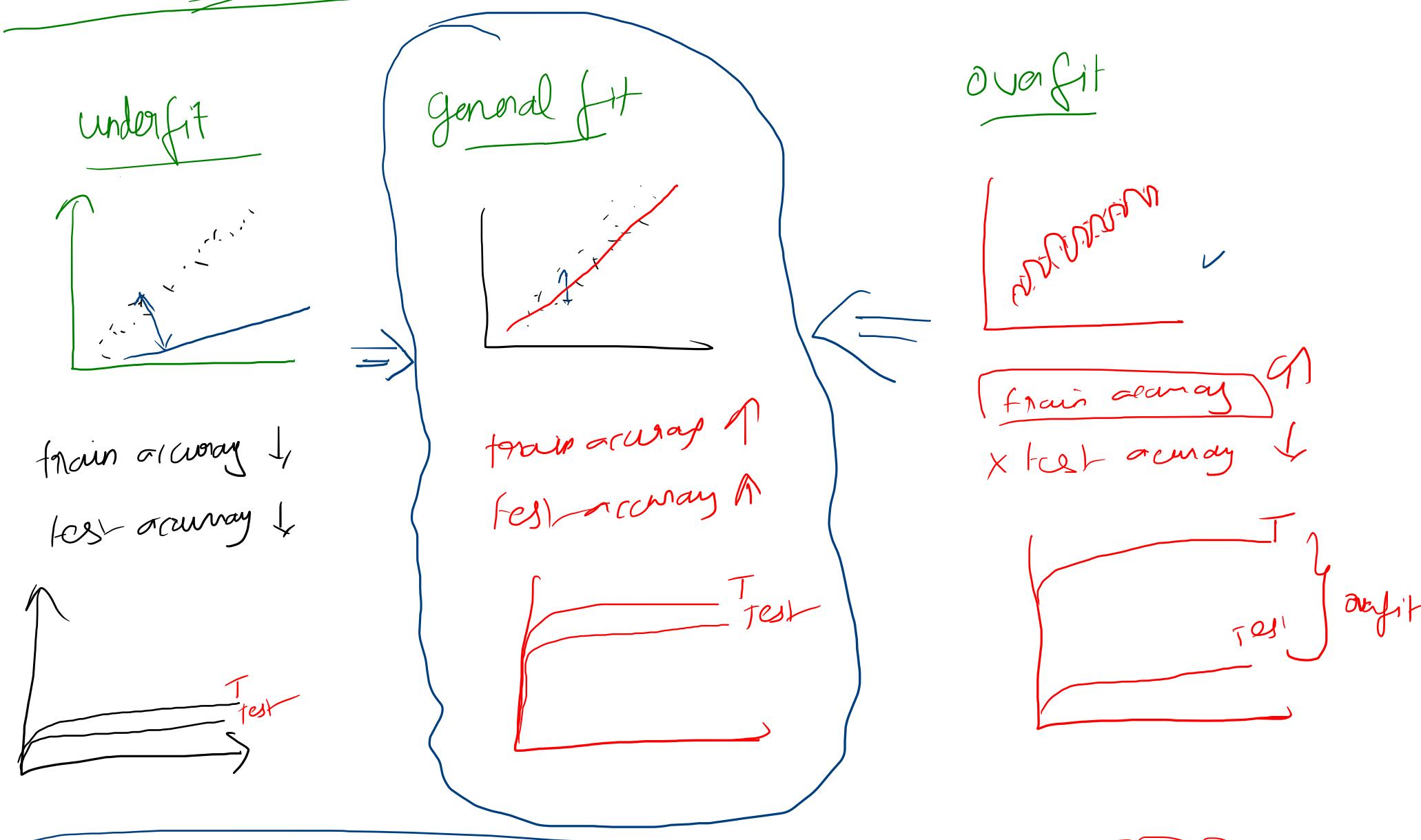
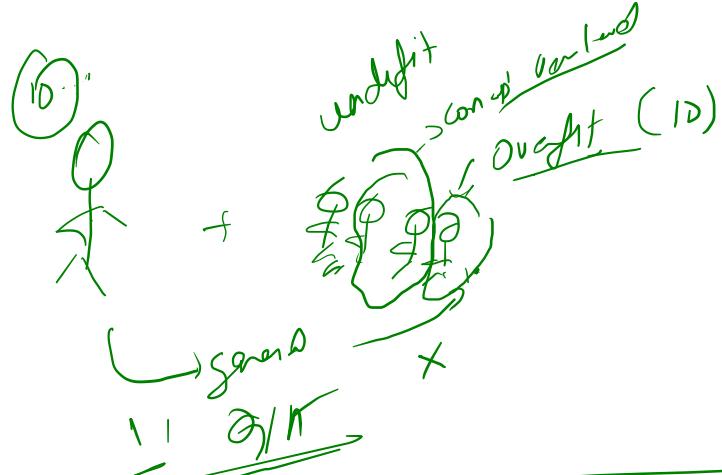
VIF > 5 highly affect by  
multi collinearity

$$\rightarrow \text{VIF} = \frac{1}{1-R^2} = \left\{ \begin{array}{l} x \rightarrow y \\ x \rightarrow p \\ x \end{array} \right.$$

$$\rightarrow \text{VIF}$$

$$\rightarrow \text{UIF}$$

$$\rightarrow \text{UIF^-}$$



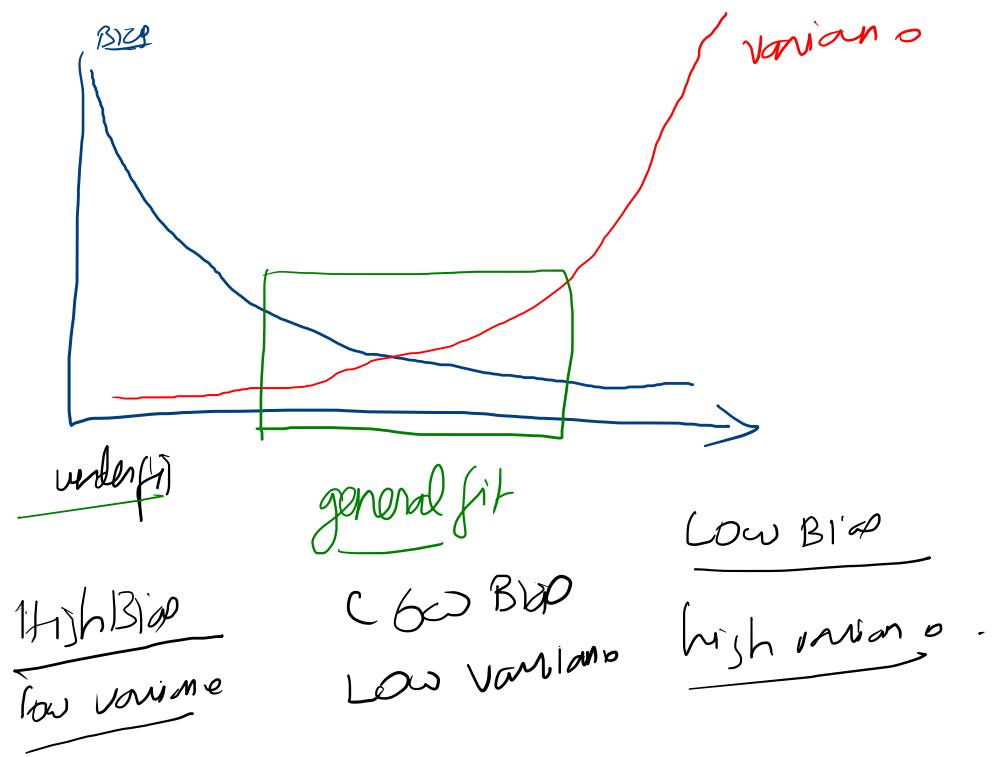
Bias  $\rightarrow$   $y = \alpha + \beta x_1 + \epsilon$

$y \in \mathbb{R}$

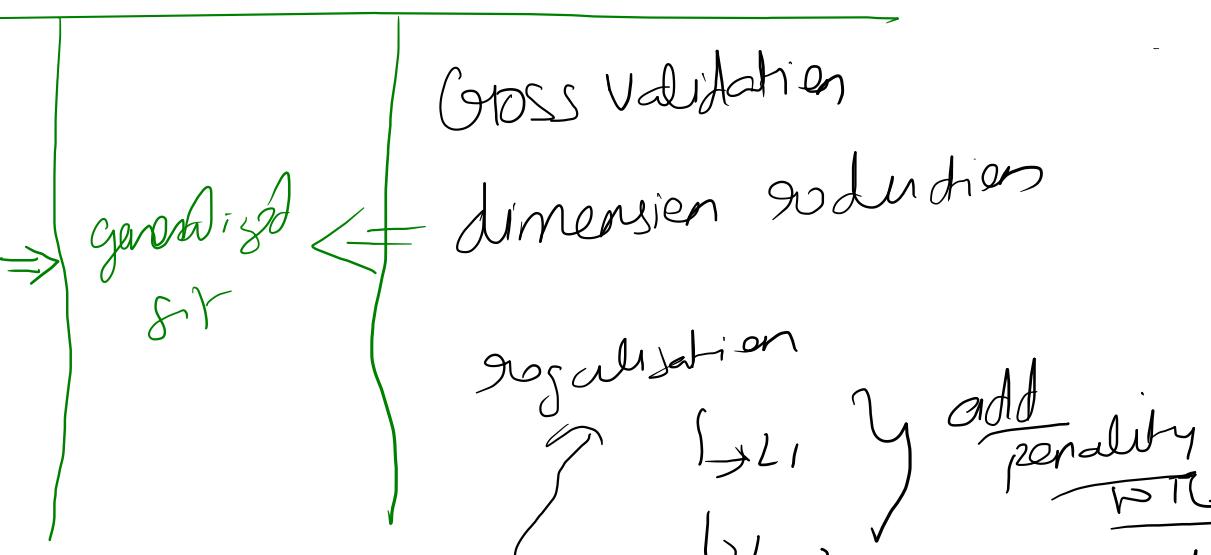
$y = f(x_1, x_2, \dots, x_n) + \epsilon$

Variance

# Bias (vs) Variance tradeoff



- ① train with different algo
- ② Add more variables
- ③ grid search Hyper parameters tuning



$$y \leftarrow 1x_1 + 2x_2 + 3x_3$$

Transo pending to copy

$$y \leftarrow 0.5x_1 + 1.2x_2 + 2x_3$$