



A handwritten yellow diagram showing a magnifying glass with a single line for a handle, focusing on a circled weight  $w$ . Above the magnifying glass is a handwritten polynomial equation:  $w_0 + w_1x + w_2x^2 + w_3x^3$ . The  $w_1$  term in the equation is also circled in yellow.



# Overfitting and Underfitting

Mohd Ayyoob  
(M.Tech CSA, IISc Bangalore)

2024

# AGENDA

- ① Overfitting
  - ② Underfitting
  - ③ Bias
  - ④ Variance
  - ⑤ Code-snippets and Visualizations.
- } why do we even need it?
- } 'hand-wavy introduction'

## GOAL in Machine learning:

- ① We want a model to 'generalize' well to unseen data data. <sup>TEST.</sup>
- ② We want our model to have high/low generaliz<sup>n</sup> accuracy.
- ③ We want our model to have high/low generaliz<sup>n</sup> error

$$\theta^+ = \arg \min_{\theta} \underbrace{\frac{1}{2} (\hat{y} - y)^2}_{\text{MSE}} = \arg \min_{\theta} \frac{1}{2} (\underbrace{h_{\theta}(x)}_{\text{TEST-ERROR}} - \underbrace{y}_{\text{TRAIN-ERROR}})^2$$

TRAIN-ERROR = 0

TEST-ERROR ↓↓↓

M1  
 $E_{\text{train}} \uparrow$

M2  
 $E_{\text{test}} \uparrow \uparrow$

M3  
 $E_{\text{test}} \downarrow$   
 $E_{\text{train}} \uparrow$

M4  
 $A_{\text{train}} \uparrow$   
 $A_{\text{test}} \downarrow$

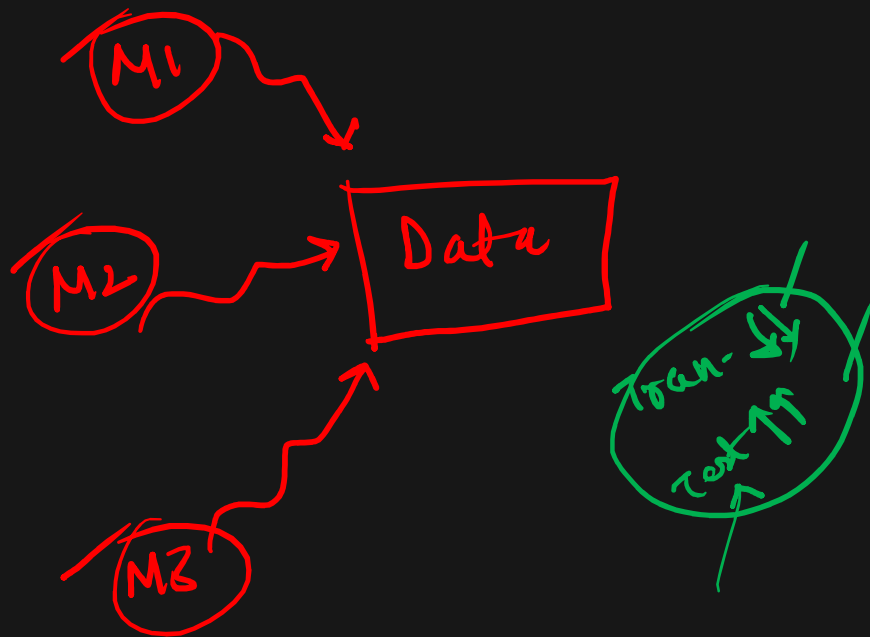
— overfit  
— underfit

FRAMEWORK

$w_0 + w_1 x'$

EVALUATE A PARTICULAR MODEL

Linear model



— TEST-error ↓↓ ✓

### FRAMEWORK

✕k compare different models on same D.

### ① No of param

M1: degree 1: LR —

M2: degree 5: poly LR —

$M2 > M1$

$M2$  is complex  $> M1$

### ② underlying algo:

M1: LR

M2: SVM

M3: NN

No of parameter

$M1 > M2 > M3$

$$\left[ \text{TEST-ERROR} = \frac{\text{BIAS}^2}{\text{}} + \frac{\text{VARIANCE}}{\text{}} + \sigma^2 \right] \text{irreducible error}$$

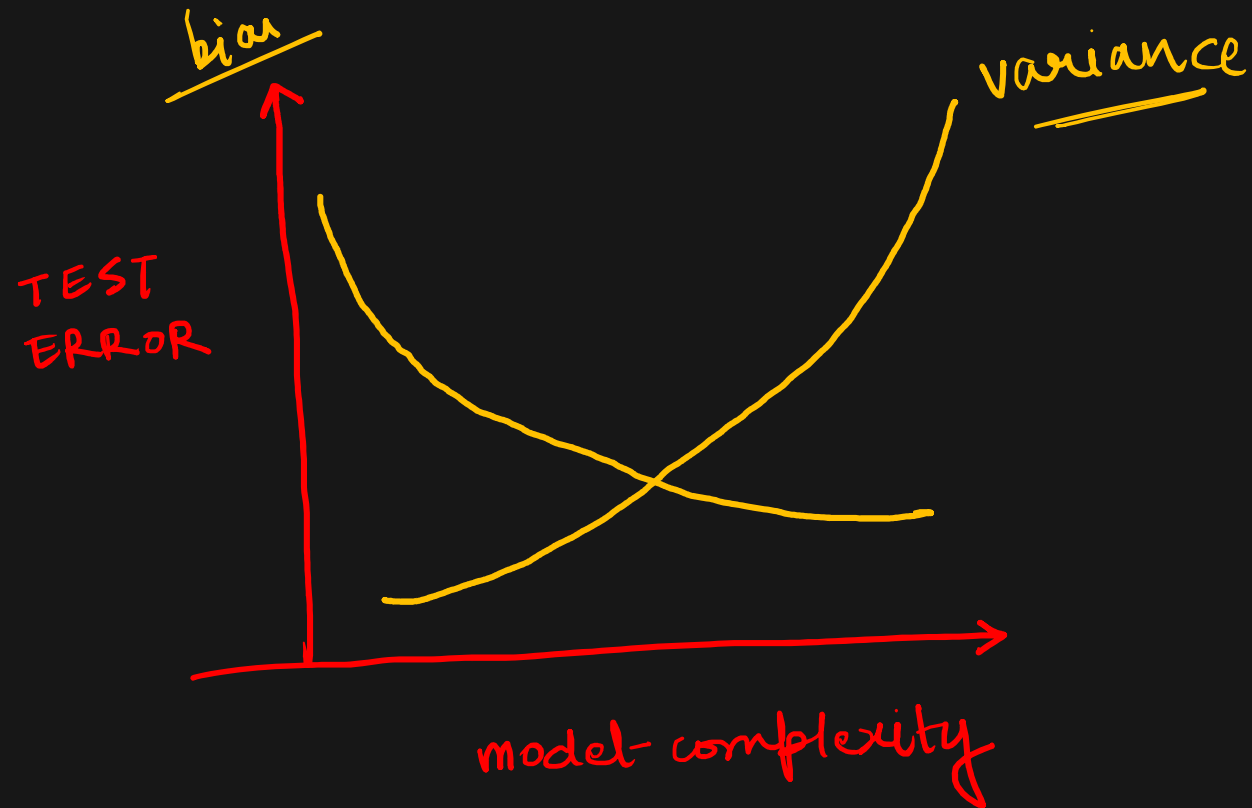
How good my assumed structure actually representing the truth

model is too sensitive to data  
→

M1: High Bias

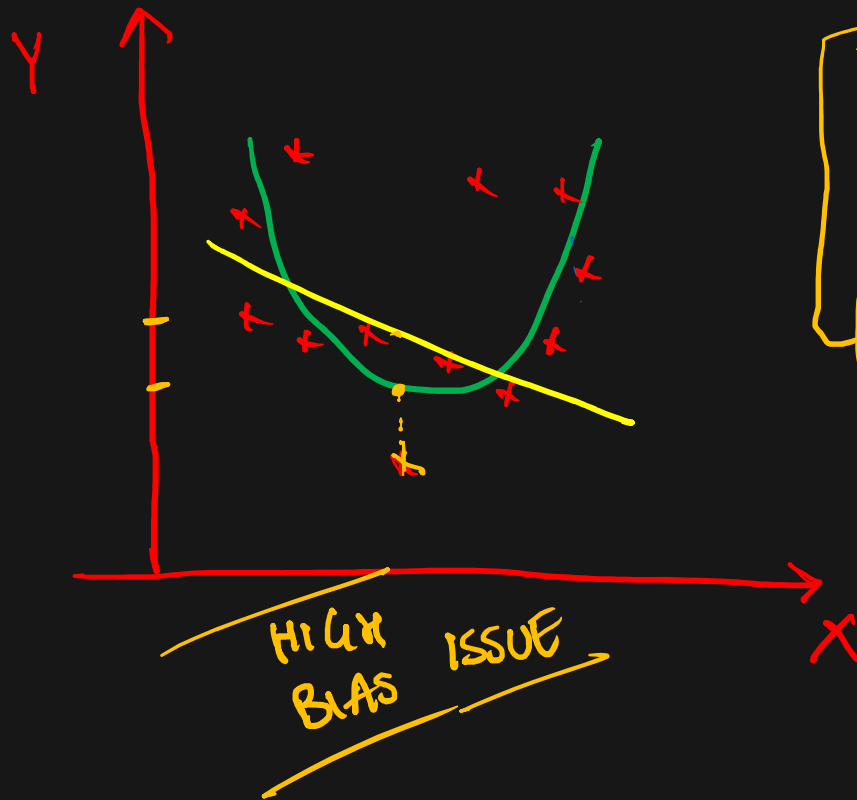
M2: High Variance

# BIAS-VARIANCE TRADEOFF



# UNDER-FITTING

- : true func<sup>n</sup> → train-error: ↑↑  
- : MI                      test-error: ↑↑



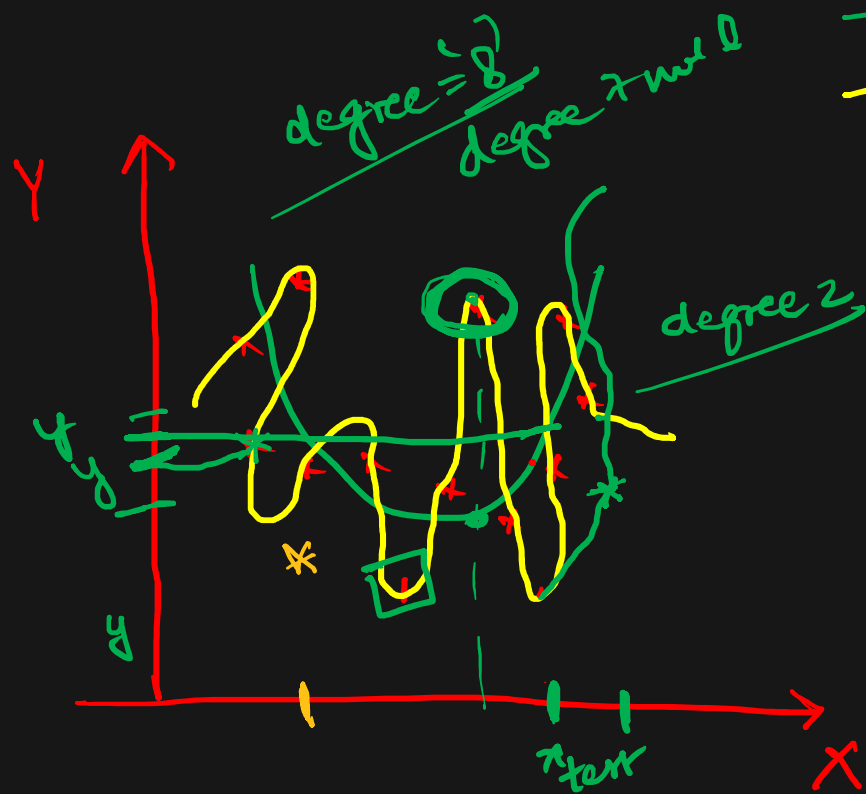
Assumed: model that we are fitting  
is from linear family WTF

→ model is too simple  
→ make model more complex

BIAS



# OVERFITTING



variance of model's prediction

—: true func<sup>n</sup>.

—: M<sub>2</sub>

- train-error: 0 ↓↓

- test-error: ↑↑

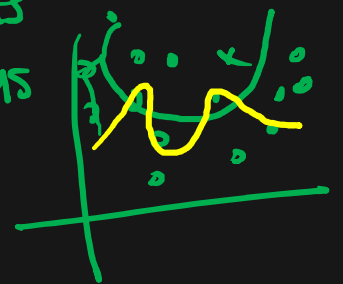
model is too complex —

- model is fitting noises

- model is too sensitive to changing training data. ↗

HIGH VARIANCE

→ Adding more points would help in this



## Good - FIT

-: true func<sup>n</sup>.

-: M3

train-error:

test-error:

