Bias Variance Tradeoff

derivation of loss f^n = bias^2 + variance + U(RE)

irreducible(E)

Setup

cgpa	iq	lpa	prediction	Error
		8	9	1
		8	8.1	.1
		7	6.9	.1
		9	10.1	1.1

Error → irreducible + reducible error

$$\boxed{\text{loss } f^n = \sum_{i=1}^{n} (y_i - \hat{y})^2}$$

- Our goal is to minimize this loss function
- Your model → lr → $\hat{y} = \beta_0 + \beta_1 \, cgpa + \beta_2 \, iq$
- There will be some error and error will be combination of reducible error and irreducible error

Assumption about irreducible error

The irreducible error :-

mean $= 0$

var $= \sigma^2$ (constant)

Derivation

$$m.se = \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n}$$

m.se can be rewrite as :-

$$E[(y-\hat{y})^2] \longrightarrow (1) \text{ (Expected)}$$

True function vs our function

(True) $y = f(x) + E = \varphi + E \quad (\because \varphi = f(x))$

(Our) $\hat{y} = f'(x) = \hat{\varphi}$

E is irreducible error.

putting y & $\hat{y}$ in (1)

$$= E[(\varphi + E - \hat{\varphi})^2]$$

$$= E[((\varphi - \hat{\varphi}) + E)^2] \quad (\because (a+b)^2 = a^2 + b^2 + 2ab)$$

$$= E[(\varphi - \hat{\varphi})^2 + E^2 + 2E(\varphi - \hat{\varphi})]$$

$$\because E[x+y] = E[x] + E[y]$$

$$= E[(\varphi - \hat{\varphi})^2] + E[E^2] + E[2E(\varphi - \hat{\varphi})]$$

$$= E[(\varphi - \hat{\varphi})^2] + E[E^2] + E[2]\, E[E]\, E[\varphi - \hat{\varphi}]$$

$\because$ ϵ is irreducible error

$\therefore$ Expected value of ϵ (irreducible error) will be zero (check assumption)

$\therefore$ $E(\epsilon) = 0$ putting this in eq^n

continue

$$\Rightarrow E[(\theta - \hat{\theta})^2] + E[\epsilon^2] + \underbrace{E[2]\,E(\epsilon)\,E[\theta - \hat{\theta}]}_{0}$$

$$\Rightarrow E[(\theta - \hat{\theta})^2] + E[\epsilon^2]$$

We know that $\text{var}(\epsilon) = E[(\epsilon - E[\epsilon])^2]$

since this = 0

$$\text{var}(\epsilon) = E[(\epsilon)^2]$$

$$\Rightarrow mse = E[(\theta - \hat{\theta})^2] + \text{Var}(\epsilon)$$

$\rightarrow$ This is irreducible error ($\because$ variance of population irreducible error)

Detour

We needed to prove: Loss fn = Bias^2 + variance + Var(ϵ)

where var (ϵ) is irreducible error

We proved already this part

Now we have to work on Bias^2 + variance

Continue

$$mse = \underbrace{E[(\theta - \hat{\theta})^2]} + \text{Var}(\epsilon)$$

Adding & substracting by $E[\hat{\varphi}]$

$= \{$ We have $E[(\varphi - \hat{\varphi})^2]$

$= E[(\underbrace{\varphi - E[\hat{\varphi}]}_{a}) + (\underbrace{E[\hat{\varphi}] - \hat{\varphi}}_{b})]^2$

$= E[(\varphi - E[\hat{\varphi}])^2 + (E[\hat{\varphi}] - \hat{\varphi})^2 + 2(\varphi - E[\hat{\varphi}])(E[\hat{\varphi}] - \hat{\varphi})]$

$= E[(\varphi - E(\hat{\varphi}])^2] + E[(E[\hat{\varphi}] - \hat{\varphi})^2]$ $(\because E[x+y] = E[x] + E[y])$

$+ E[2(\varphi - E[\hat{\varphi}])(E[\hat{\varphi}] - \hat{\varphi})] \rightarrow 0$

$= \varphi[$

because

$E[2] \; E[(\varphi - E(\hat{\varphi}])] \; E[(E[\hat{\varphi}] - \hat{\varphi})]$

$2 \quad (\varphi - E[\hat{\varphi}]) \{E[E[\hat{\varphi}]] - E[\hat{\varphi}]\}$

$E[\hat{\varphi}] - E[\hat{\varphi}]$

$\leftarrow \times \; 0 = 0$

$(\because E[1]$ constant $= 1$

$= E[\varphi - E[\hat{\varphi}])^2] + E[(E[\hat{\varphi}] - \hat{\varphi})^2]$

$= \underbrace{(\varphi - E[\hat{\varphi}])^2}_{\text{Bias}} + \underbrace{(E[\hat{\varphi}] - \hat{\varphi})^2}_{\text{Variance}}$

$MSE = \underbrace{Bias^2 + Variance}_{\text{reducible error}} + \underbrace{Var(\epsilon)}_{\text{Irreducible error}}$ Hence proved

mse = reducible + irreducible error

$\downarrow$ $\qquad\qquad$ $\downarrow$

$(Bias^2 + \text{variance})$ $\qquad$ $var(\epsilon)$

$\downarrow$ $\qquad\qquad$ $\downarrow$ noise variance

ml model variance