PROOF THAT $MSE = BIAS^2 + VARIANCE$

θ_m = Model Parameter

with θ being its true value

θ_m is inferred from training data ; so θ_m is a random Variable with mean $\bar{\theta}_m$

Mean Square Error = MSE = $E[(\theta_m - \theta)^2]$

Bias = $E[\theta_m - \theta] = E[\theta_m] - \theta = \bar{\theta}_m - \theta$

Variance of θ_m = $E[(\theta_m - \bar{\theta}_m)^2]$

We wish to prove that $MSE = Bias^2 + Variance$

or $E[(\theta_m - \theta)^2] = (\bar{\theta}_m - \theta)^2 + E[(\theta_m - \bar{\theta}_m)^2]$ ← Plugging in definitions above

Lets take RHS : $(\bar{\theta}_m - \theta)^2 + E[(\theta_m - \bar{\theta}_m)^2]$

$$= \bar{\theta}_m^2 + \theta^2 - 2\theta\bar{\theta}_m + E[\theta_m^2 + \bar{\theta}_m^2 - 2\theta_m\bar{\theta}_m]$$

$$= \bar{\theta}_m^2 + \theta^2 - 2\theta\bar{\theta}_m + E[\theta_m^2] + \bar{\theta}_m^2 - 2E[\theta_m]\bar{\theta}_m$$

$$= \bar{\theta}_m^2 + \theta^2 - 2\theta\bar{\theta}_m + E[\theta_m^2] + \bar{\theta}_m^2 - 2\bar{\theta}_m^2$$

$$= \theta^2 + E[\theta_m^2] - 2\theta\bar{\theta}_m$$

$$= E[\theta^2 + \theta_m^2 - 2\theta\theta_m]$$

$$= E[(\theta - \theta_m)^2]$$

$$= MSE = LHS$$

Scanned by CamScanner