Estimate $\hat{f}$ = Learn $\hat{f}$

sources of error:

irreducible error ϵ

reducible error $\hat{f}$

the squared error for a given estimate $\hat{f}$ is

$$E(\text{actual} - \text{predicted})^2 = E(Y - \hat{Y})^2$$

which factors as

$$E[f(\boldsymbol{X}) + \epsilon - \hat{f}(\boldsymbol{X})]^2$$

$$\underbrace{[f(\boldsymbol{X}) - \hat{f}(\boldsymbol{X})^2]}_{\text{reducible}} + \underbrace{\text{Var}(\epsilon)}_{\text{irreducible}}$$

until now, training data was the only data we considered

we compute reducible error (or MSE) on the same data used to learn $\hat{f}$

let's change that!

$$Y = f(\boldsymbol{X}) + \epsilon$$

Estimate $\hat{f}$ = Learn $\hat{f}$

$$Y = f(\boldsymbol{X}) + \epsilon$$

sources of error:

irreducible error ϵ

reducible error $\hat{f}$

the squared error for a given estimate $\hat{f}$ is

$$E(\text{actual} - \text{predicted})^2 = E(Y - \hat{Y})^2$$

which factors as

$$E[f(\boldsymbol{X}) + \epsilon - \hat{f}(\boldsymbol{X})]^2$$

$$\underbrace{[f(\boldsymbol{X}) - \hat{f}(\boldsymbol{X})^2]}_{\text{reducible}} + \underbrace{\text{Var}(\epsilon)}_{\text{irreducible}}$$

until now, training data was the only data we considered
we compute reducible error (or MSE) on the same data used to learn $\hat{f}$

let's change that!