$$E[\text{test}] = E[(y-\hat{y})^2]$$

$$= E[(f(x) + \varepsilon - \hat{f}(x))^2]$$

$f(x) \rightarrow$ best model

$\hat{f}(x) \Rightarrow$ estimated model

$\varepsilon \rightarrow$ random noise

$$\underset{(MSE)}{E[\text{test}]} = E[(f - \hat{f} + \epsilon)^2]$$

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

$$= E[(f-\hat{f})^2] + E[\epsilon^2] + 2E[(f-\hat{f})]\cancelto{0}{E[\epsilon]}$$

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

$$= E[(f - E[\hat{f}] + E[\hat{f}] - \hat{f})^2] + \text{var}(\epsilon)$$

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

$$+ 2E[(f-E[\hat{f}])(E[\hat{f}]-\hat{f})]$$

$$+ \text{Var}(\epsilon)$$

$$= \text{Bias}[\hat{f}]^2 + \text{Var}[\hat{f}] + \text{Var}(\epsilon)$$

$$+ 2\left(E[f]E[\hat{f}] - \cancel{E[\hat{f}]^2} - E[f\hat{f}] + \cancel{E[\hat{f}]^2}\right)$$

$$E[\text{test}] = \text{Bias}[\hat{f}]^2 + \text{Var}[\hat{f}] + \text{Var}(\epsilon)$$

$$+ 2\left(E[f]E[\hat{f}] - E[f\hat{f}]\right)$$

f & $\hat{f}$ are independent
therefore it will cancel out

$$\Rightarrow \boxed{E[\text{test}] = \text{Bias}[\hat{f}]^2 + \text{Var}[\hat{f}] + \text{Var}(\epsilon)}$$