$$Y = f(X) + \varepsilon$$

$$E\left[(y_0 - \hat{f}(x_0))^2\right]$$

$$= E\left[(f(x_0) - \hat{f}(x_0) + \varepsilon)^2\right]$$

$$= E\left[(f(x_0) - \hat{f}(x_0))^2 + \varepsilon^2 + 2(f(x_0) - \hat{f}(x_0))\varepsilon\right]$$

$$= E\left[(f(x_0) - \hat{f}(x_0))^2\right] + E\varepsilon^2 + \boxed{2E\left[(f(x_0) - \hat{f}(x_0))\varepsilon\right]}$$

$$\boxed{2E\left[(f(x_0) - \hat{f}(x_0))\varepsilon\right]} = 2E\left[f(x_0) - \hat{f}(x_0)\right]E\varepsilon = 0$$

$$= E\left[(f(x_0) \underline{- E\hat{f}(x_0)} \underline{+ E\hat{f}(x_0)} - \hat{f}(x_0))^2\right] + Var\,\varepsilon$$

$$= E\left[(f(x_0) - E\hat{f}(x_0))^2 + (E\hat{f}(x_0) - \hat{f}(x_0))^2 + 2(f(x_0) - E\hat{f}(x_0))(E\hat{f}(x_0) - \hat{f}(x_0))\right] + Var\,\varepsilon$$

$$E\left[\underbrace{(f(x_0) - E\hat{f}(x_0))}_{\text{not random}}(E\hat{f}(x_0) - \hat{f}(x_0))\right]$$

$$= \underline{(f(x_0) - E\hat{f}(x_0))}\, E\left[E\hat{f}(x_0) - \hat{f}(x_0)\right]$$

$$= (f(x_0) - E\hat{f}(x_0))\left[E\hat{f}(x_0) - E\hat{f}(x_0)\right] = 0$$

$$= (f(x_0) - E\hat{f}(x_0))^2 + E\left[(\hat{f}(x_0) - E\hat{f}(x_0))^2\right] + Var\,\varepsilon$$

$$= (Bias\,\hat{f}(x_0))^2 + Var\,\hat{f}(x_0) + Var\,\varepsilon$$

Flexibility ↗ $\Longrightarrow$ Variance ↗

Bias ↘

$$\boxed{f(x_0)} - \boxed{E\hat{f}(x_0)}$$

not random   not random

$$\begin{matrix}(X_1, Y_1) \\ \vdots \\ (X_n, Y_n)\end{matrix} \Longrightarrow \hat{f}$$

$$Y_i = \boxed{f(X_i)} + \varepsilon_i$$

not random   random