

Fact:

$$\mathbb{E}[(Y - \hat{f}(x))^2 \mid X=x] \stackrel{?}{=} \underbrace{(f(x) - \hat{f}(x))^2}_{\text{Reducible}} + \underbrace{\text{Var}[\varepsilon]}_{\text{Irreducible}}$$

→:

By our model:

$Y = f(X) + \varepsilon$  w/  $\varepsilon$  independent from  $X$   
and  $\mathbb{E}[\varepsilon] = 0$

$$\begin{aligned} \mathbb{E}[(f(x) + \varepsilon - \hat{f}(x))^2 \mid X=x] &= (\text{linearity of expectation}) \\ &= \mathbb{E}[(f(x) - \hat{f}(x))^2 \mid X=x] \\ &\quad + 2 \mathbb{E}[(f(x) - \hat{f}(x)) \cdot \varepsilon \mid X=x] \quad (\varepsilon \text{ independent from } X) \\ &\quad + \mathbb{E}[\varepsilon^2 \mid X=x] \\ &= (f(x) - \hat{f}(x))^2 \\ &\quad + \cancel{2 \mathbb{E}[(f(x) - \hat{f}(x)) \mid X=x] \cdot \mathbb{E}[\varepsilon]} \\ &\quad + \text{Var}[\varepsilon] \quad \square \end{aligned}$$

In general, for any r.v.  $W$

$$\text{Var}[W] = \mathbb{E}[W^2] - (\mathbb{E}[W])^2$$

$$\mathbb{E}[W^2] = \text{Var}[W] + (\mathbb{E}[W])^2$$

If we take  $\varepsilon = W$  and note  $\mathbb{E}[\varepsilon] = 0$

$$\Rightarrow \mathbb{E}[\varepsilon^2 \mid X=x] \underset{\substack{\uparrow \\ \text{independence}}}{=} \mathbb{E}[\varepsilon^2] = \text{Var}[\varepsilon] + (\mathbb{E}[\varepsilon])^2 = \text{Var}[\varepsilon]$$