

Let  $m(X) = E[Y|X]$  (regression function)

Then for any  $f: X \rightarrow \mathbb{R}$ ,

$$\begin{aligned} E(f(X) - Y)^2 &= E(f(X) - m(X) + (m(X) - Y))^2 \\ &= E(f(X) - m(X))^2 + E(m(X) - Y)^2 \\ &\quad + 2E[(f(X) - m(X))(m(X) - Y)] \\ &= E E[(f(X) - m(X))(m(X) - Y) | X] \\ &= E(f(X) - m(X)) \underbrace{E[m(X) - Y | X]}_{=0} = 0 \end{aligned}$$

$$= \underbrace{E(f(X) - m(X))^2}_{\geq 0} - E(m(X) - Y)^2$$

$$\geq E(m(X) - Y)^2,$$

and therefore  $m$  is the optimal predictor under the squared loss.