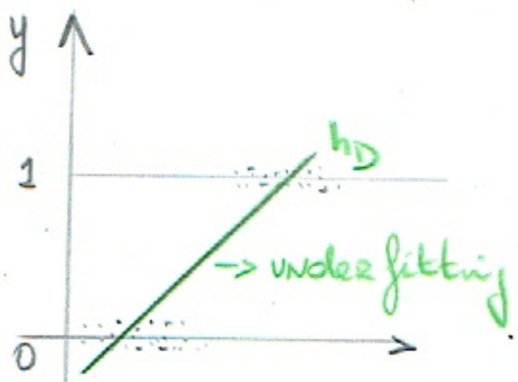
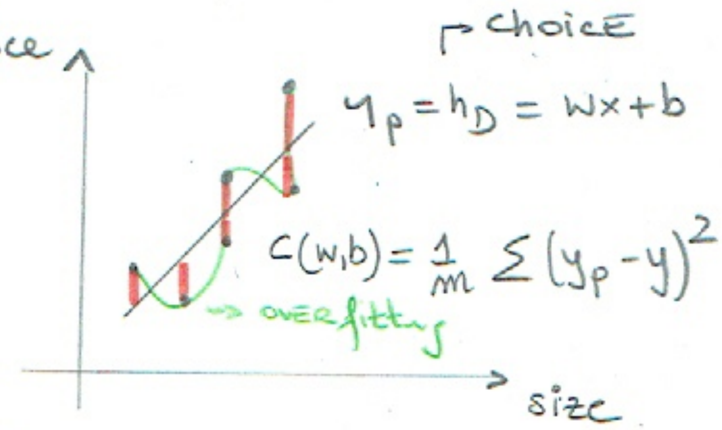
UNderfitting / overfitting

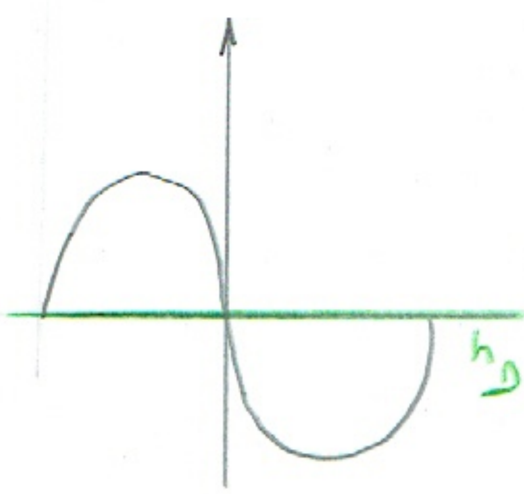
↓ BIAS → GENERALization → simple $h_D$

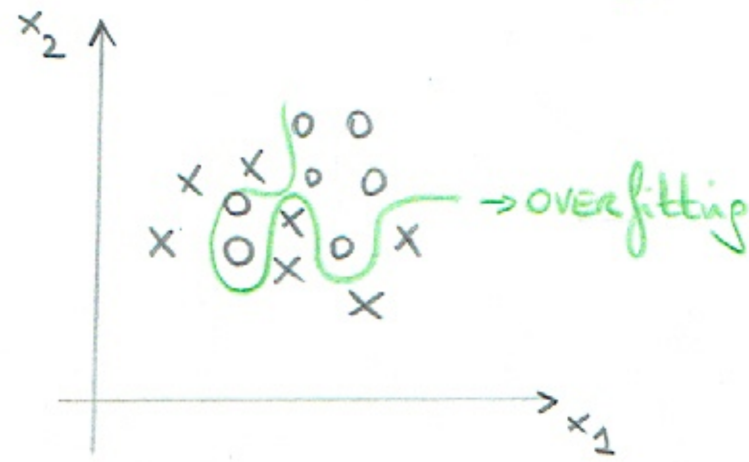
↓ Variance → Approximation ↳ MORE complex $h_D$

∃ trade off!




process

1. DRAW DATAset $D \sim P$
2. TRAIN Algorithm → $h_D$
3. Take Test Point $(x, y)$ → calculate expected ERROR

$$\boxed{\text{Expected ERROR} = \underbrace{\sigma^2}_{\text{noise}} + \underbrace{VAR(h_D)}_{\text{Range of } h_D\text{s}} + \underbrace{BIAS^2(h_D)}_{\text{how far is } h_D \text{ from } E(h_D)?}}$$

→ $h_D$ is RV

Another representation

## BIAS – VARIANCE Decomposition (optional)

$y_p$ : result of training

②

Expected ERROR (regression) $= E\left[(y - h_D)^2\right]$ (real y)

$= E(y^2) + E(h_D^2) - 2E(y h_D)$ (1)

$(a-b)^2 = a^2 + b^2 - 2ab$

SET : $y = h_D + \varepsilon$ and $E[\varepsilon] = 0$ ($\sim N(0, \sigma^2)$)

$var(y) = E[y^2] - E[y]^2$ (statistics refresher) $\Rightarrow E[y^2] = VAR[y] + E[y]^2$ (2)

AND $E[h_D^2] = VAR(h_D) + E[h_D]^2$

NOW $E[y] = E[h_D + \varepsilon] = E[h_D] + \underbrace{E[\varepsilon]}_{0} = h_D$ (3) (deterministic!)

$$(1) \underset{(2)\,\&\,(3)}{=} \underbrace{VAR[y] + E[y]^2}_{(2)} + \underbrace{VAR[h_D] + E[h_D]^2}_{(2)} - 2y\,E[h_D]$$

$E[y]^2 = E[h_D]^2$

$$= VAR[y] + \underbrace{VAR[h_D] + h_D^2 + E[h_D]^2 - \underbrace{2y\,E[h_D]}_{2h_D E[h_D]}}$$

$$VAR[y] = E\left[(y - E[y])^2\right] = E\left[(y - h_D)^2\right] = E\left[(h_D + \varepsilon - h_D)^2\right] = E[\varepsilon^2] = \sigma^2$$

NOISE (irreducible)

$\left[h_D - E[h_D]\right]^2 = \left[BIAS(h_D)\right]^2$

→ how much is classifier biased to an explanation that cannot be found in the data.

$$\rightarrow \boxed{\text{Expected ERROR} = \sigma^2 + VAR[h_D] + BIAS^2(h_D)}$$