

$$\Delta P[|E_{tr}(h) - E(h)| > \epsilon] \leq 2e^{-2\epsilon^2 N}$$

$\underbrace{\quad}_{\text{training error}}$
 $\underbrace{\quad}_{\text{transition error?}}$

Δ also, when we hv multiple hypothesis:
 we want an upper bound

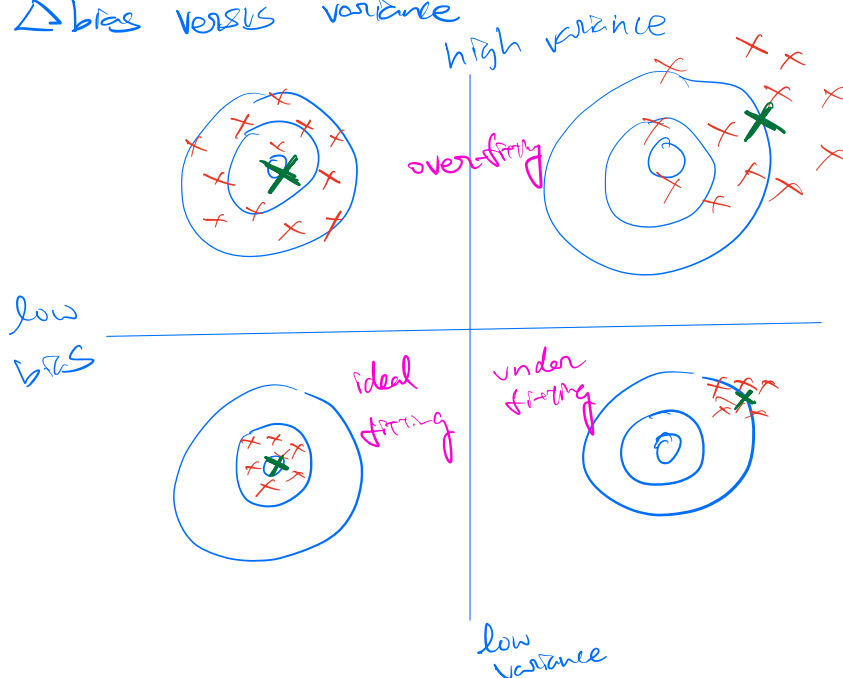
$$P[\sup_{h \in \mathcal{H}} |E_{tr}(h) - E(h)| > \epsilon]$$

Δ How many training example do you need.

$$N \geq \frac{1}{2\epsilon^2} \log \frac{2/|\mathcal{H}|}{\delta} \quad \text{where } \delta = 2/|\mathcal{H}| e^{-2\epsilon^2 N}$$

$$E_{tr}(h^*) - r \leq E(h^*)$$

Δ bias versus variance



$$E_{tr}(h^*) - r$$

large E_{Train}
 large E_{Test}

overfitting. \Rightarrow high variance
 Sensitive to noise

$$\Delta H = w^T x + b$$

$1/|\mathcal{H}| \rightarrow$ is actually infinity
 (non-trivial)

How do we shrink it?
 (improve)

Δ on P17 - LG.

h_1 & h_2 although
 mathematical
 different

It is empirically same.

