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

training error — transition error?

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

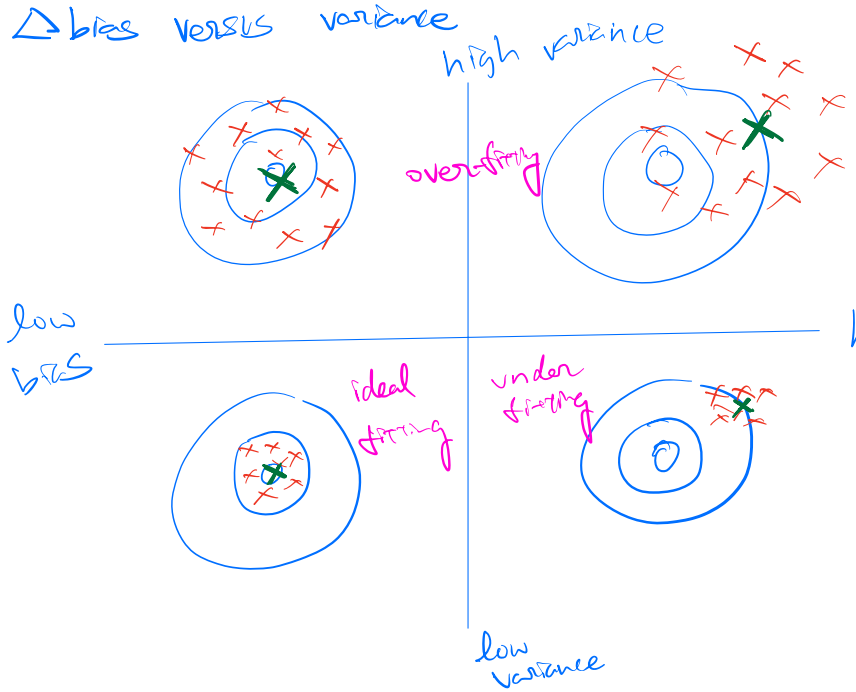
$P[\sup_{h \in \mathcal{H}} |E_{tr}(h) - E(h)| > \epsilon]$

△ How many training training example do you need.

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

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

△ bias versus variance



~~$E_{tr}(h^*) - \gamma$~~

large E_{Train}
large E_{test}

overfitting = high variance
Sensitive to noise

△ $H = w^T x + b$

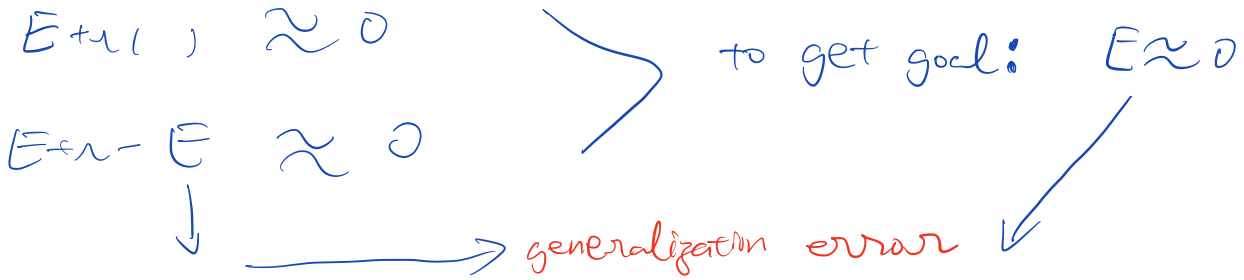
$|\mathcal{H}| \rightarrow$ is actually infinity (non-trivial)

How do we shrink it? (improve)

△ on P17_L8.

h_1 & h_2 although mathematical different

It is empirically same.


$E_{train} \approx 0$
$E_{val} - E \approx 0$
to get good: $E \approx 0$
generalization error