

Probability distribution

Don't see $X \neq Y$

data

$\rightarrow (x_i, y_i), (x_1, y_1), \dots$

Training Data

$(\vec{x}_1, y_1), (\vec{x}_2, y_2), \dots$

Test data

$(\vec{x}_{t_1}, y_{t_1}), \dots, (\vec{x}_{t_n}, y_{t_n})$

Hypothesis / Model set H

$h_{\vec{w}}(\vec{x}) = \text{sign}(\langle \vec{w}, \vec{x} \rangle)$

Algorithm A

Final hypothesis

$h_{\vec{w}^*}$ (fixed \vec{w}^*)

Estimate the

generalization error

Loss function $\ell()$

$\ell(h_{\vec{w}}, (\vec{x}, y)) = \max\{0, -y \langle \vec{w}, \vec{x} \rangle\}$

ERM: pick \vec{w}^* that gives the minimum training error