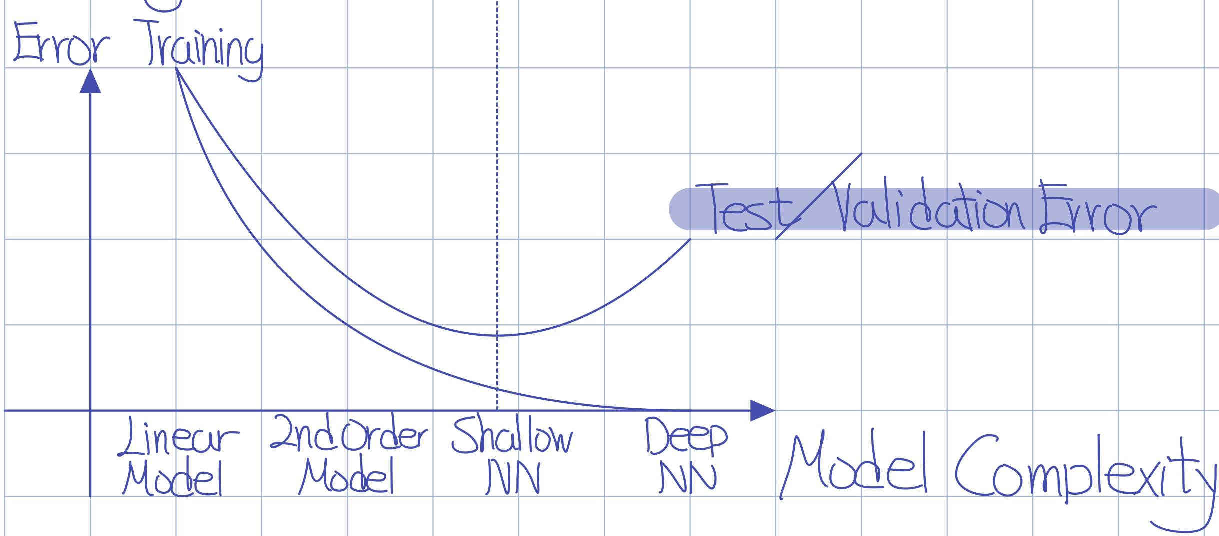
ECE 324 → Machine Learning Courses
APS 360 → Machine Learning Courses

Probability Approximately Correct (PAC)

Given a training dataset which model should be select for training ?



Mathematical Framework

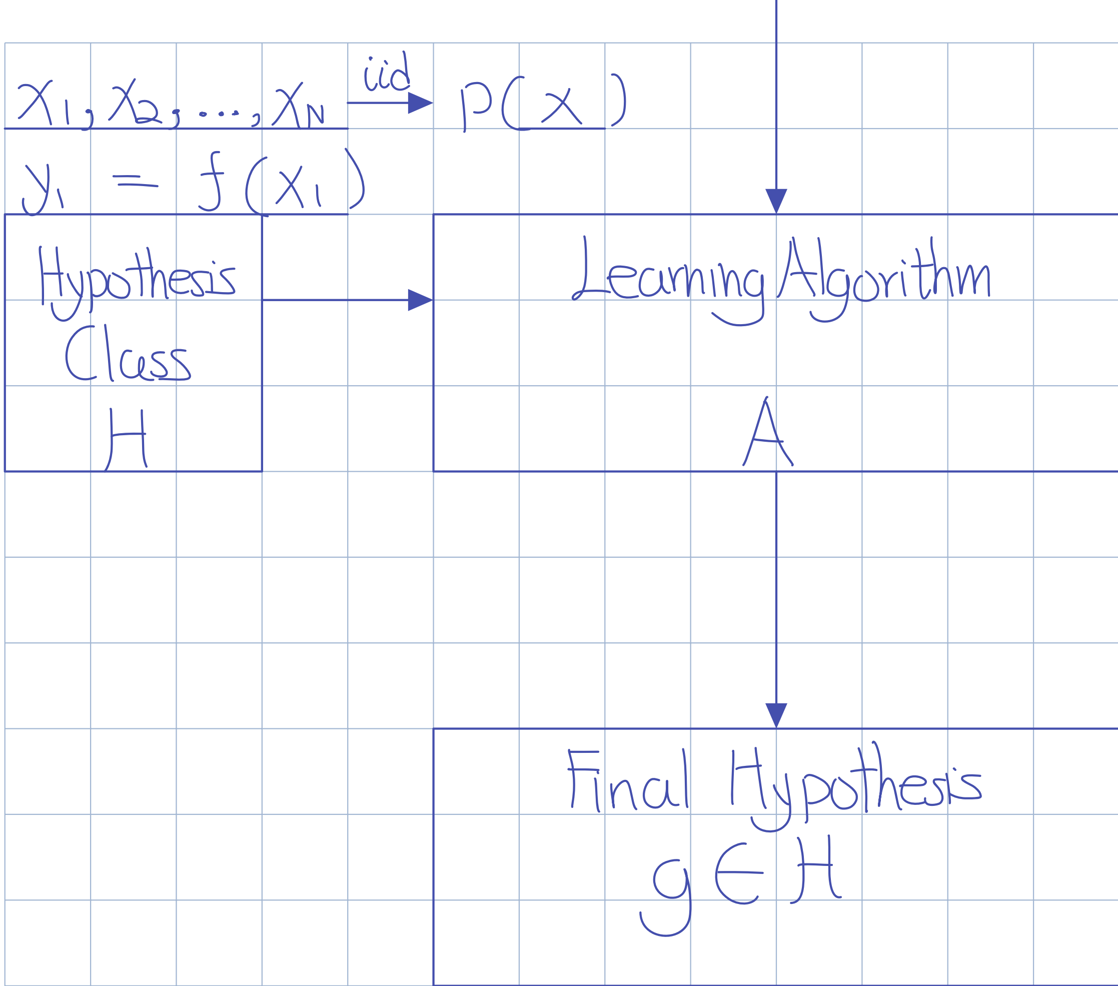
Binary Classification $\underline{x} \in \mathbb{R}^d$, $y \in \{-1, 1\}$

Unknown Target Function $y = f(\underline{x})$

Unknown Data Distribution $F(\underline{x})$

↓

Training Set
$D = \{(\underline{x}_1, y_1), \ldots (x_N, y_N)\}$



Summary

1) Unknown Function $f : \mathbb{R}^d \longrightarrow \{-1, +1\}$

2) Unknown Distribution $p(\underline{x})$

$$\underline{x_1, x_2, \ldots, x_N} \xrightarrow{iid} p(\underline{x})$$

$$y_1 = f(\underline{x_1})$$

3) Hypothesis Class $H = \{h_1, h_2, \ldots, h_M\}$

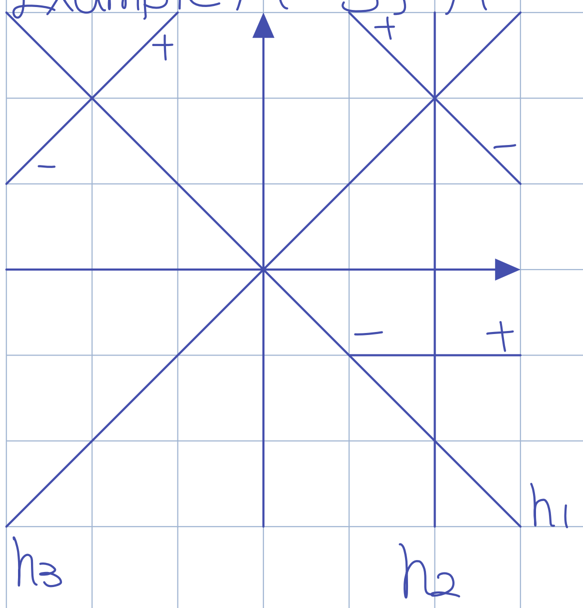
$h_i : \mathbb{R}^d \longrightarrow \{-1, 1\}$

Output hypothesis $g \in H$

Chapter 1: M is finite

Chapter 2: M is infinite

Example $M = 3$, $H = \{ h_1, h_2, h_3 \}$



A will output either

$g = h_1$

or $g = h_2$

or $g = h_3$

Performance Metrics

In Sample Training Errors

$$\boxed{E_{in}(g) = \frac{1}{N}\sum_{i=1}^{N} e(y_i, g(\underline{x_i}))}$$

Binary Classification

$$e(y_i, g(\underline{x_i})) = \begin{cases} 1, & y \neq g(\underline{x_i}) \\ 0, & y = g(\underline{x_i}) \end{cases}$$

$$= \mathbb{1}\{ y \neq g(\underline{x_i}) \}$$

Test Error

$$E_{out}(g) = \mathbb{E}_x[e(y, g(\underline{x}))]$$

Integral In D-Dim

$$= \oint_{x \in \mathbb{R}^d} e(y, g(\underline{x}))\, p(\underline{x})\, dx$$

$$= \oint_{x\in\mathbb{R}^d} \underline{1}\left\{ y \neq g(\underline{x_i}) \right\} p(\underline{x})\,dx$$

$$= \cancelto{\text{Zero}}{\oint_{x\in\mathbb{R}^d, y=g(x)} \underline{1}\left\{ y \neq g(\underline{x_i}) \right\} p(\underline{x})\,dx} +$$

$$\oint_{x\in\mathbb{R}^d, y\neq g(x)} \underline{1}\left\{ y \neq g(\underline{x_i}) \right\} p(\underline{x})\,dx$$

Splitting into two integrals

$$= \oint_{x\in\mathbb{R}^d, y\neq g(x)} \underline{1}\left\{ y \neq g(\underline{x_i}) \right\} p(\underline{x})\,dx$$

$$= \oint_{x\in\mathbb{R}^d, y\neq g(x)} p(\underline{x})\,dx$$

$$= Pr(y \neq g(\underline{x}))$$

$\boxed{E_{out}(g) = P_r(y \neq g(x))}$ (What We Care About)

$\underline{X}$ = test data point (no access)

$\underline{X} \sim P(\underline{x})$

also $\underline{X}$ is independent of $\underline{X_1, X_2, \ldots, X_N}$

PAC Learning Gaurantee Condition

2 Performance Metric

Training Error	Computable
Testing Error	Not Computable