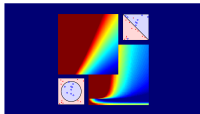


Machine Learning Foundations

(機器學習基石)



Lecture 5: Training versus Testing, Extended

Hsuan-Tien Lin (林軒田)

htlin@csie.ntu.edu.tw

Department of Computer Science
& Information Engineering

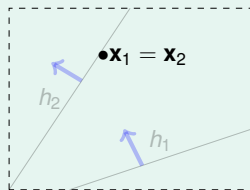
National Taiwan University
(國立台灣大學資訊工程系)



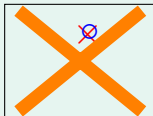
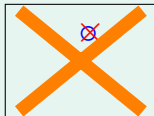
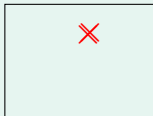
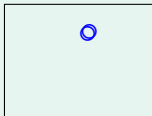
How Many Lines Are There for Overlapping Inputs?

$$\mathcal{H} = \left\{ \text{all lines in } \mathbb{R}^2 \right\}$$

- how many **kinds of** lines if viewed from two inputs $\mathbf{x}_1, \mathbf{x}_2$, but $\mathbf{x}_1 = \mathbf{x}_2$?



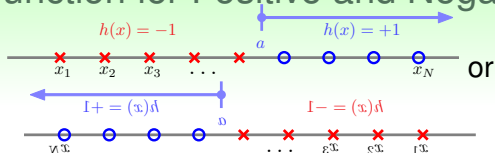
2:



overlapping generally 'decreases'

$$\mathcal{H}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$$

Growth Function for Positive and Negative Rays



- $\mathcal{X} = \mathbb{R}$ (one dimensional)
- $\mathcal{H}$ contains h , where **each** $h(x) = \text{sign}(x - a)$ for threshold a or $h(x) = -\text{sign}(x - a)$
- 1D perceptrons, called decision stump

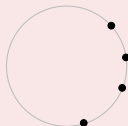
$$m_{\mathcal{H}}(N) = (N + 1) + (N + 1) - 2 = 2N$$

- $N + 1$ dichotomies from positive rays, **same from negative rays by symmetry**
- 2 overlapping, $(\circ, \circ, \dots, \circ)$ and $(\times, \times, \dots, \times)$

can be extended to compute
growth function of **hypothesis-set union**

Growth Function for Origin-Passing Perceptrons

- one possible set of N inputs:
 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$ on a unit half-circle



- **angle-sweeping** reduces the problem to **decision stump**: $2N$ dichotomies
- other cases
 - not on unit circle: **normalize**
 - not on half-circle: **flip as equivalent dichotomy**
 - origin: fewer dichotomies anyway

$m_{\mathcal{H}}(N) = 2N$ for origin-passing perceptrons
by connecting (reducing) to 1D perceptrons