AdaBoost: an Introduction

Carlo Ciliberto

SINA – Genova – 28/11/2011
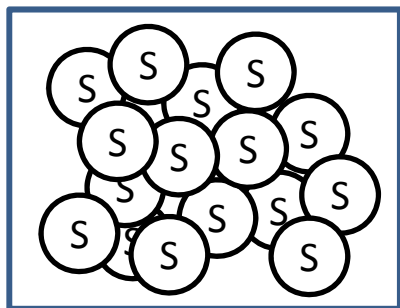
AdaBoost

AdaBoost

Inputs:

- Training set $\{(x_i, y_i)\}_{i=1}^{n}$ $\qquad x \in \mathcal{X} \quad y_i \in [-1, 1]$
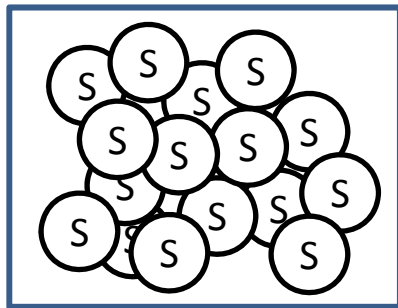
Training set
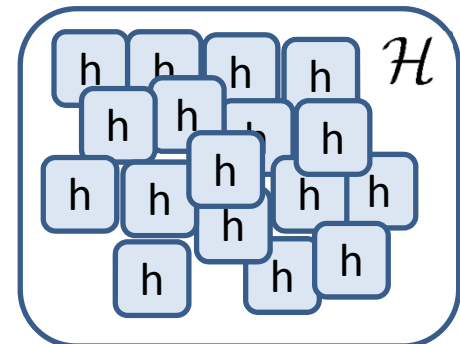
AdaBoost

Inputs:

- Training set $\{(x_i, y_i)\}_{i=1}^{n}$ $x \in \mathcal{X}$ $y_i \in [-1, 1]$

- Weak Learners $\mathcal{H}$
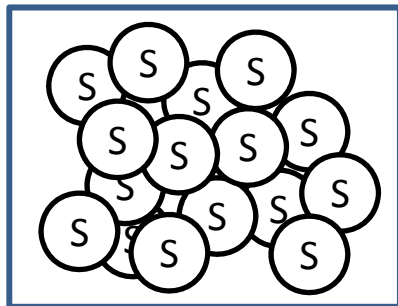
Training set

Weak learners

AdaBoost

Inputs:

- Training set $\{(x_i, y_i)\}_{i=1}^n$ $x \in \mathcal{X}$ $y_i \in [-1, 1]$
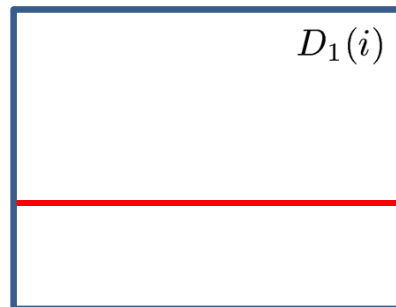
- Weak Learners $\mathcal{H}$

Initialization:
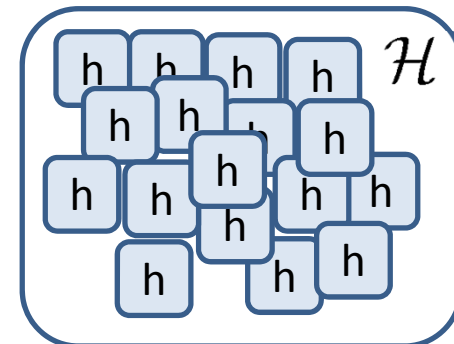
- Uniform initial distribution $D_1(i) = 1/n$
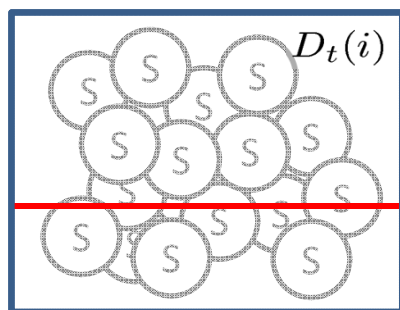
Training set

Weight distribution

$D_1(i)$

Weak learners

$\mathcal{H}$

AdaBoost

Training:

For $t = 1, ..., T$:

Training set
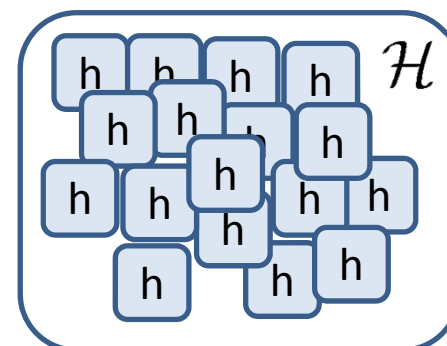
$D_t(i)$

Weak learners

$\mathcal{H}$

h h h h
h h
h
h h h h
h
h h h

AdaBoost

Training:

For $t = 1,...,T$:

- Find the current best weak learner: $h_t = arg\underset{h}{min} \sum_{i/y_i \neq h(x_i)} D_t(i)$

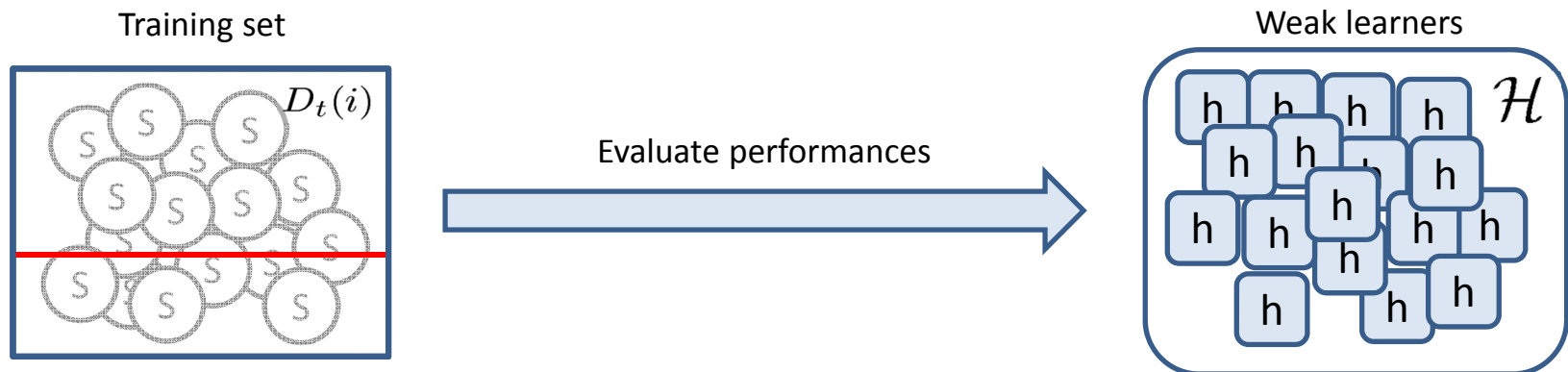Training set

Weak learners

Evaluate performances

$D_t(i)$

$\mathcal{H}$

AdaBoost

Training:

For $t = 1,...,T$:

- Find the current best weak learner: $h_t = arg\underset{h}{min} \sum_{i/y_i \neq h(x_i)} D_t(i)$

Training set

$D_t(i)$

Best weak learner

h_t

Weak learners

$\mathcal{H}$

AdaBoost

Training:

For $t = 1,...,T$:

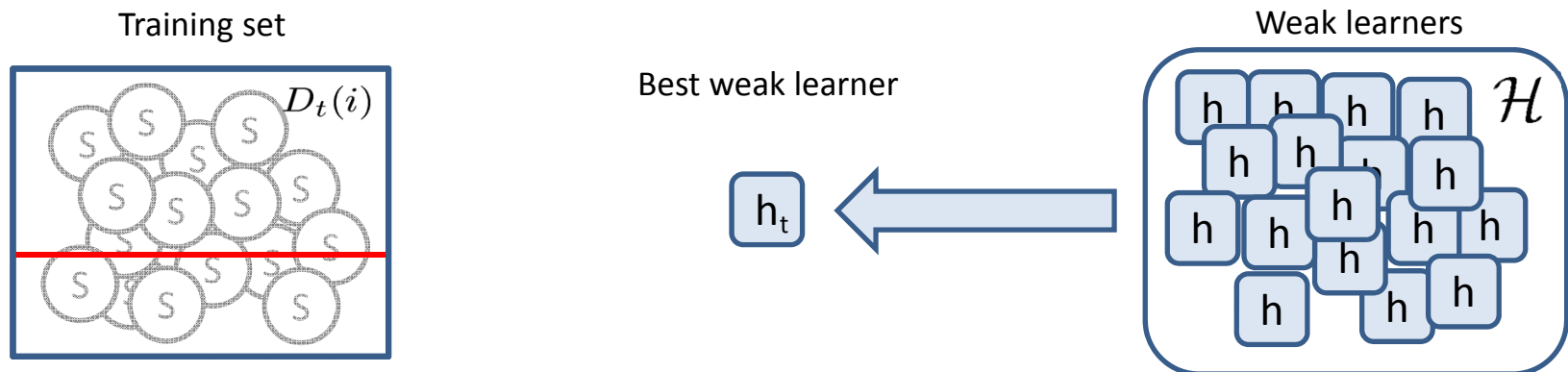- Find the current best weak learner: $h_t = \underset{h}{argmin} \sum_{i/y_i \neq h(x_i)} D_t(i)$

- Set $\alpha_t = \frac{1}{2} \log \frac{1 - \epsilon_t}{\epsilon_t}$, $\epsilon_t = \sum_{i/y_i \neq h(x_i)} D_t(i)$

Training set

$D_t(i)$

Best weak learner

h_t

Weak learners

$\mathcal{H}$

h h h h
h h h
h
h h h h
h
h h h

AdaBoost

Training:

For $t = 1,...,T$:

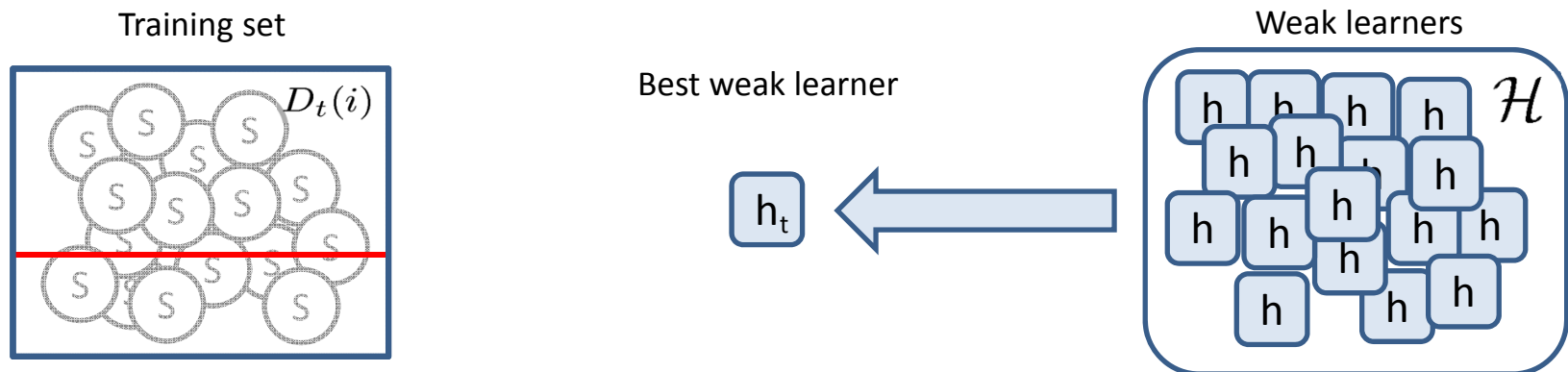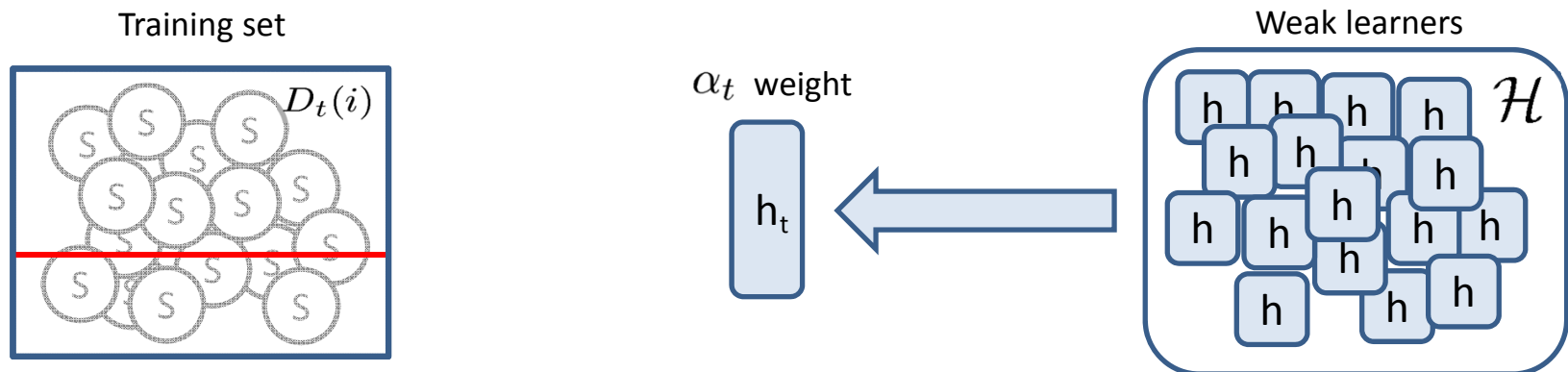- Find the current best weak learner: $h_t = \underset{h}{argmin} \sum_{i/y_i \neq h(x_i)} D_t(i)$

- Set $\alpha_t = \frac{1}{2} \, log \, \frac{1 - \epsilon_t}{\epsilon_t} \; , \; \epsilon_t = \sum_{i/y_i \neq h(x_i)} D_t(i)$

Training set

$D_t(i)$

α_t weight

h_t

Weak learners

$\mathcal{H}$

AdaBoost

Training:

For $t = 1,...,T$:

- Find the current best weak learner: $h_t = \underset{h}{argmin} \sum_{i/y_i \neq h(x_i)} D_t(i)$

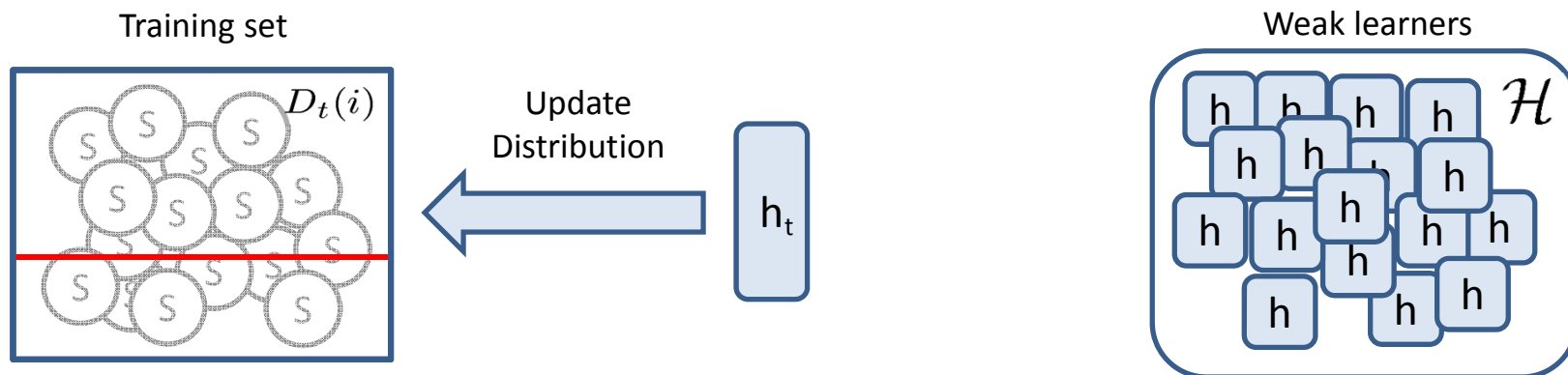- Set $\alpha_t = \frac{1}{2} log \frac{1 - \epsilon_t}{\epsilon_t}$, $\epsilon_t = \sum_{i/y_i \neq h(x_i)} D_t(i)$

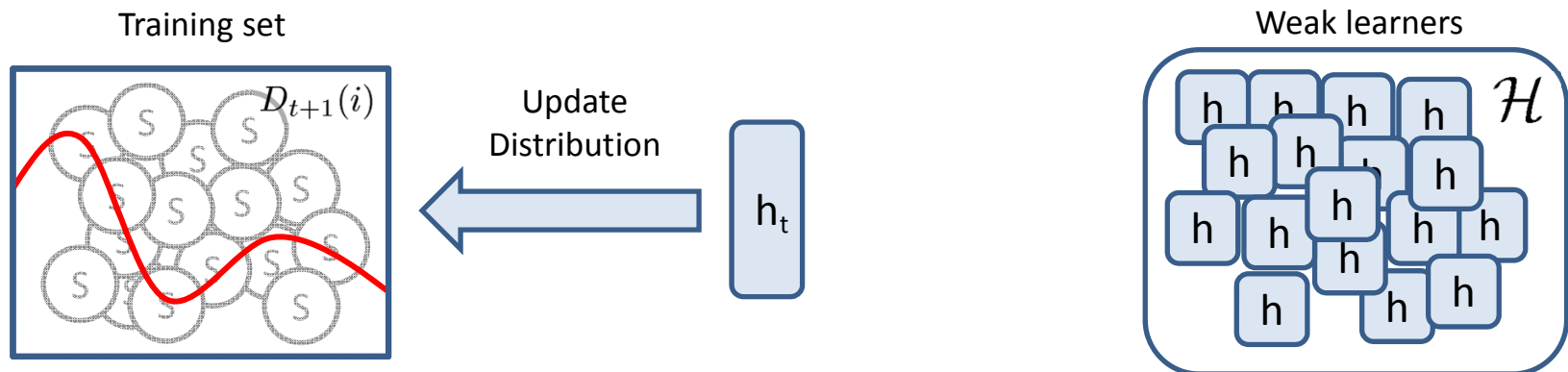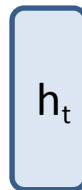- Update the weights $D_{t+1}(i) = \frac{D_t(i)exp(-\alpha_t y_i h_t(x_i))}{Z_t}$

Training set

Update Distribution

h_t

Weak learners

AdaBoost

Training:

For $t = 1,...,T$:

- Find the current best weak learner: $h_t = \underset{h}{argmin} \sum_{i/y_i \neq h(x_i)} D_t(i)$

- Set $\alpha_t = \frac{1}{2} log \frac{1 - \epsilon_t}{\epsilon_t}$, $\epsilon_t = \sum_{i/y_i \neq h(x_i)} D_t(i)$

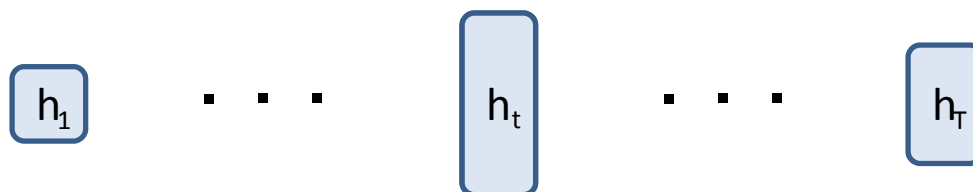- Update the weights $D_{t+1}(i) = \frac{D_t(i) exp(-\alpha_t y_i h_t(x_i))}{Z_t}$

Training set

$D_{t+1}(i)$

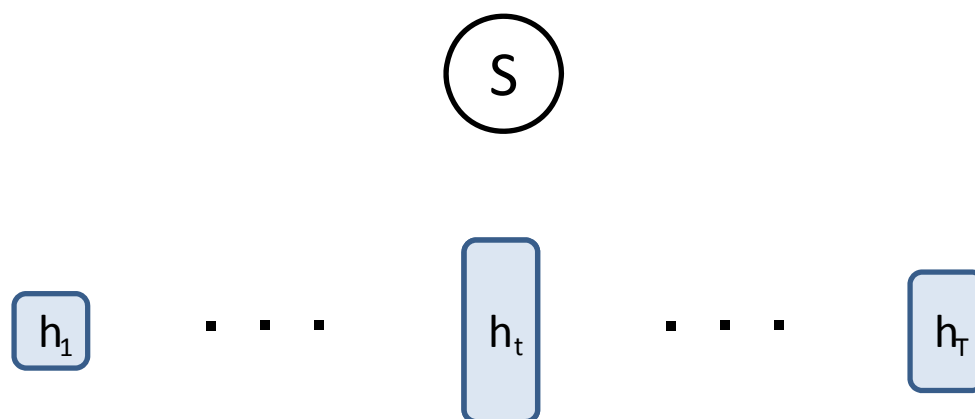Update Distribution

h_t

Weak learners

$\mathcal{H}$

h h h h h h h h h h h h h h h h h h

AdaBoost

Test:

h_t

AdaBoost

Test:

$$h_1 \quad \cdots \quad h_t \quad \cdots \quad h_T$$

AdaBoost

Test:

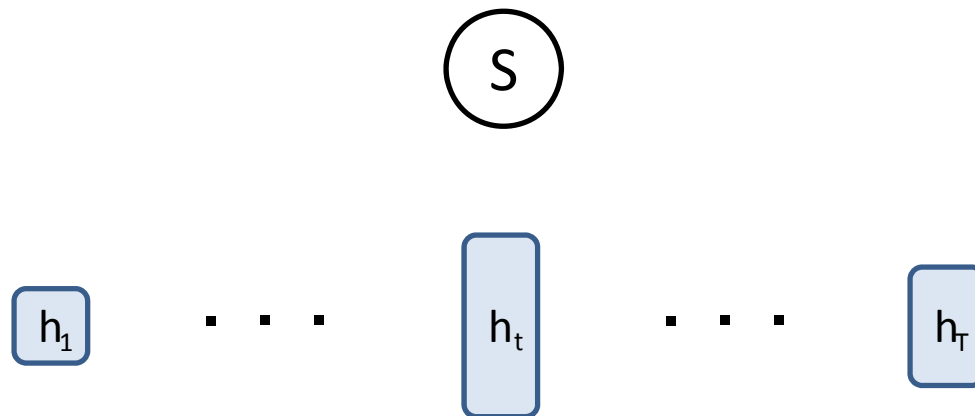For any novel example $x \in \mathcal{X}$:

S

h_1 · · · h_t · · · h_T

AdaBoost

Test:

For any novel example $x \in \mathcal{X}$:

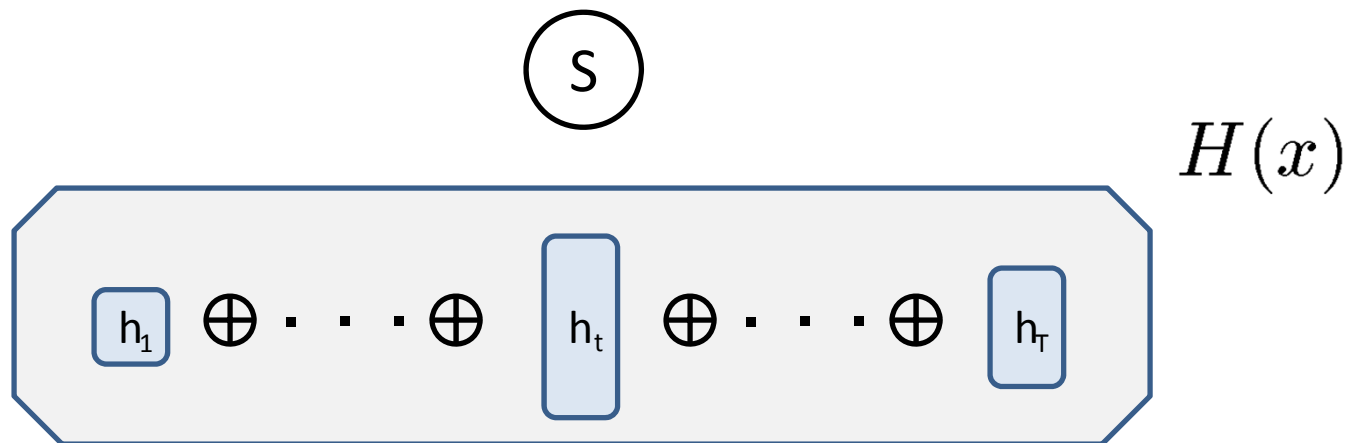- The strong classifier returns $H(x) = sign\left(\sum_{t=1}^{T} \alpha_t h_t(x)\right)$

S

h₁ · · · hₜ · · · h_T

AdaBoost

Test:

For any novel example $x \in \mathcal{X}$:

- The strong classifier returns $H(x) = sign\left(\sum\limits_{t=1}^{T} \alpha_t h_t(x) \right)$
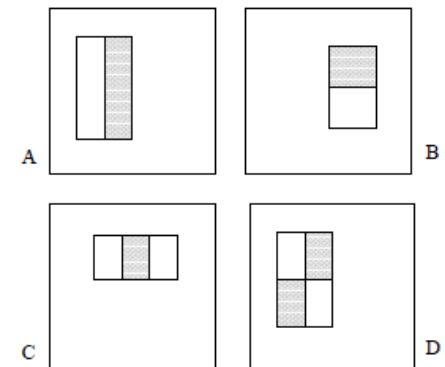
Face Detection

Viola and Jones 2001

Face Detection

- Training set $\{(x_i, y_i)\}_{i=1}^{n}$

- Weak Learners $\mathcal{H}$

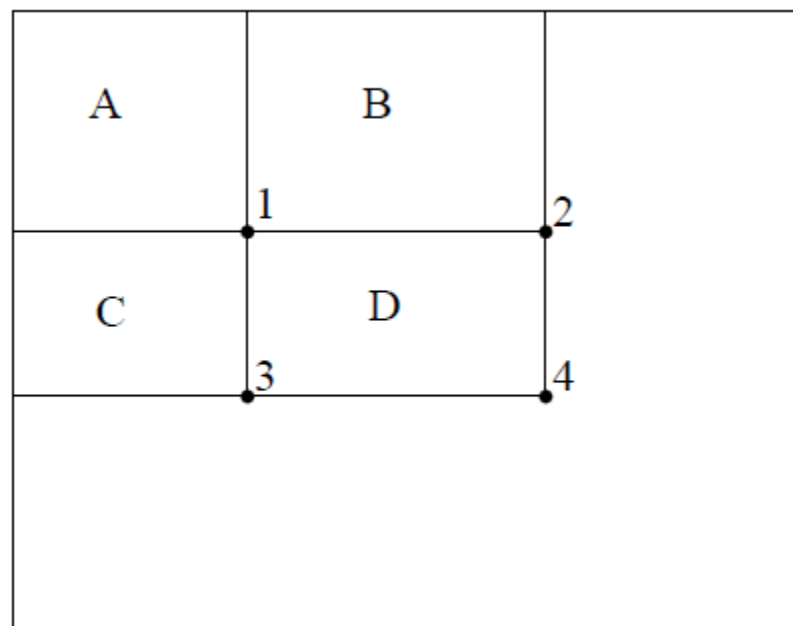$$h_j(x) = \begin{cases} 1 & \text{if } p_j f_j(x) < p_j \theta_j \\ -1 & \text{otherwise} \end{cases}$$

A   B

C   D

Viola and Jones 2001

Integral Image

$$ii(x,y) = \sum_{x' \leq x, y' \leq y} i(x',y')$$

(x,y)
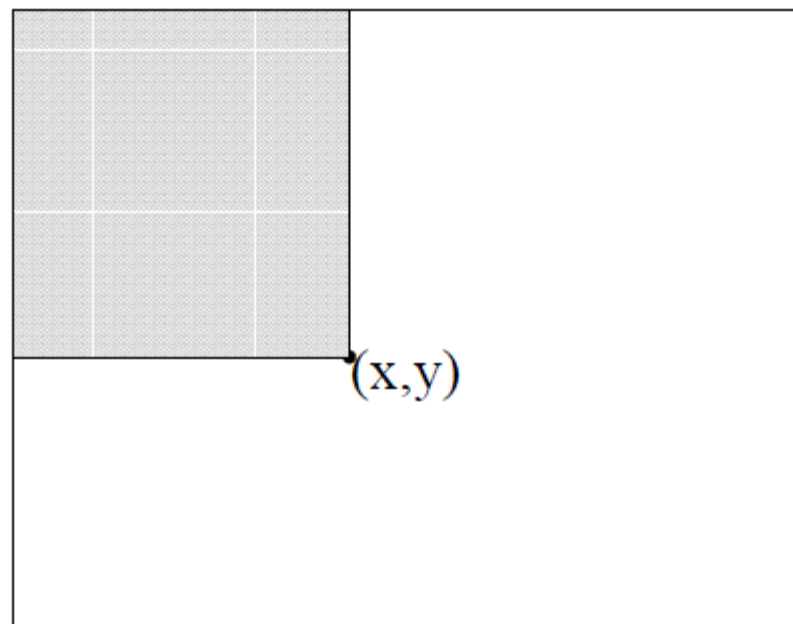
Integral Image

Integral Image

$D = ii(x_4, y_4)$

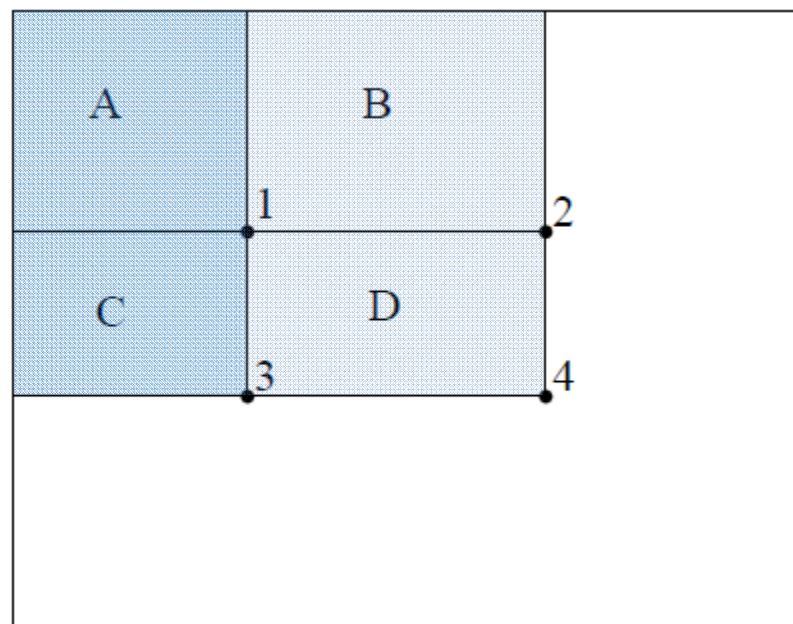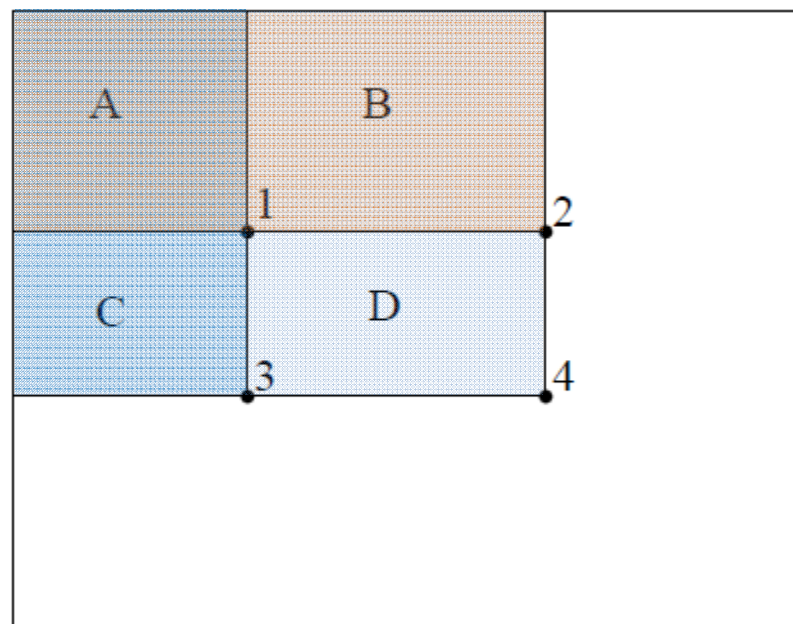*Viola and Jones 2001*
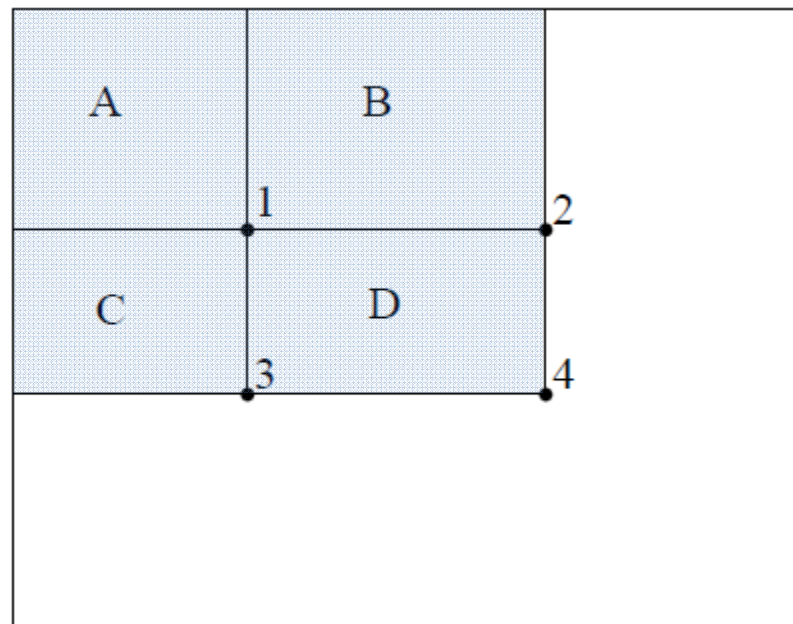
Integral Image

$$D = ii(x_4, y_4) - ii(x_3, y_3)$$
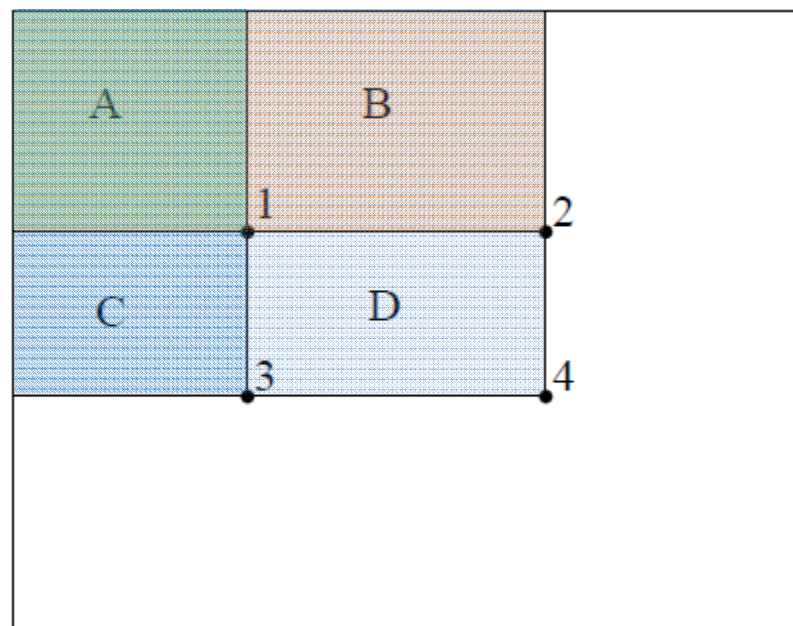
Integral Image

$$D = ii(x_4, y_4) - ii(x_3, y_3) - ii(x_2, y_2)$$

Integral Image

$$D = ii(x_4, y_4) - ii(x_3, y_3) - ii(x_2, y_2) + ii(x_1, y_1)$$

Integral Image

Viola and Jones 2001