# BOOSTING APPROACH (ADABOOST ALGORITHM)

# Motivation & Procedure

- Is it possible to boost a weak learning algorithm (which performs slightly better than a random guessing) into a strong algorithm?

- To achieve it
  - Adopt a weak classifier known as the base classifier.
  - By employing the base classifier, design a series of classifiers one by one
    - Each time, use a different weighting of training samples
    - Emphasis is given on hard samples
  - Combine the classifiers by a weighted average procedure.

# AdaBoost

- Construct an optimal classifier of the form

$$f(\mathbf{x}) = \text{sign}\{F(\mathbf{x})\}$$

where

$$F(\mathbf{x}) = \sum_{k=1}^{K} \alpha_k \phi(\mathbf{x}; \boldsymbol{\theta}_k)$$

- Optimize $\alpha_k$ and $\boldsymbol{\theta}_k$ from $k = 1$ to $K$ sequentially to minimize

$$\sum_{i=1}^{N} \exp\left(-y_i F(\mathbf{x}_i)\right)$$

# AdaBoost

- Notations
  - Partial sum

$$F_m(\mathbf{x}) = \sum_{k=1}^{m} \alpha_k \phi(\mathbf{x}; \boldsymbol{\theta_k}) = F_{m-1}(x) + \alpha_m \phi(\mathbf{x}; \boldsymbol{\theta}_m)$$

  - Weight

$$w_i^{(m)} = \frac{1}{\beta} \exp(-y_i F_{m-1}(\mathbf{x}_i))$$

  - Weighted empirical error

$$P_m = \sum_{i}^{N} w_i^{(m)} I(y_i \neq \phi(\mathbf{x}; \boldsymbol{\theta}_m))$$

# AdaBoost

- Optimal $\boldsymbol{\theta}_m$

$$\boldsymbol{\theta}_m = \arg\min_{\boldsymbol{\theta}} \sum_{i}^{N} w_i^{(m)} I(y_i \neq \phi(\mathbf{x}; \theta))$$

- Optimal $\alpha_m$

$$\alpha_m = \frac{1}{2} \ln \frac{1 - P_m}{P_m}$$

# AdaBoost – Pseudocode

- Initialize: $w_i^{(1)} = \frac{1}{N},\ i = 1, 2 \ldots, N$
- Initialize: $m = 1$
- Repeat
  - Compute optimum $\boldsymbol{\theta}_m$ in $\phi(\cdot; \boldsymbol{\theta}_m)$ by minimizing $P_m$; (4.135)
  - Compute the optimum $P_m$; (4.135)
  - $\alpha_m = \frac{1}{2} \ln \frac{1-P_m}{P_m}$
  - $Z_m = 0.0$
  - For $i = 1$ to $N$
    - $w_i^{(m+1)} = w_i^{(m)} \exp\left(-y_i \alpha_m \phi(\boldsymbol{x}_i; \boldsymbol{\theta}_m)\right)$
    - $Z_m = Z_m + w_i^{(m+1)}$
  - End{For}
  - For $i = 1$ to $N$
    - $w_i^{(m+1)} = w_i^{(m+1)} / Z_m$
  - End {For}
  - $K = m$
  - $m = m + 1$
- Until a termination criterion is met.
- $f(\cdot) = \text{sign}(\sum_{k=1}^{K} \alpha_k \phi(\cdot, \boldsymbol{\theta}_k))$