

输入：训练数据集(包含 N 个样本)，弱学习算法及对应的弱分类器一般选择决策树桩、迭代次数 M ：

过程：

1. 初始化训练数据的权值分布

$$W_0 = (w_{01}, \dots, w_{0N})$$

批注 [A1]: $1/N$

2. 对 $k = 0, 1, \dots, M - 1$:

1. 使用权值分布为 W_k 的训练数据集训练弱分类器

批注 [A2]: 弱分类器为单层的二叉决策树，划分标准是加权错误率

$$g_{k+1}(x): X \rightarrow \{-1, +1\}$$

2. 计算 $g_{k+1}(x)$ 在训练数据集上的加权错误率

$$e_{k+1} = \sum_{i=1}^M w_{ki} I(g_{k+1}(x_i) \neq y_i)$$

批注 [A3]: I 是示性函数
这就是本步的加权错误率

以上是单层决策树

批注 [A4]: 上面是弱分类器的实现过程

3. 根据加权错误率 $e_{k+1}(x)$ 的“话语权”

批注 [A5]: 弱分类器的总体话语权

$$\alpha_{k+1} = \frac{1}{2} \ln \frac{1 - e_{k+1}}{e_{k+1}}$$

4. 根据 $g_{k+1}(x)$ 的表现更新训练数据集的权值分布：被 $g_{k+1}(x)$ 误分的样本（ $y_i g_{k+1}(x_i) < 0$ 的样本）要相对地（以 $e^{\alpha_{k+1}}$ 为比例地）增大其权重，反之则要（以 $e^{-\alpha_{k+1}}$ 为比例地）减少其权重

$$w_{k+1,i} = \frac{w_{ki}}{Z_k} \cdot \exp(-\alpha_{k+1} y_i g_{k+1}(x_i))$$
$$W_{k+1} = (w_{k+1,1}, \dots, w_{k+1,N})$$

批注 [A6]: 弱分类器对每个样本权重的调整，既对该样本的识别效果有关，也与整体识别效果有关

5. 这里的 Z_k 是规范化因子，如下所示。它的作用是将 W_{k+1} 归一化为一个概率分布

$$Z_k = \sum_{i=1}^N w_{ki} \cdot \exp(-\alpha_{k+1} y_i g_{k+1}(x_i))$$

批注 [A7]: 向量乘法，矩阵乘法

3. 加权集成弱分类器

$$f(x) = \sum_{i=1}^M \alpha_k g_k$$

批注 [A8]: 根据每个弱分类器的效果加权预测

输出：最终分类器 $g(x)$

$$g(x) = \text{sign}(f(x)) = \text{sign}\left(\sum_{k=1}^M \alpha_k g_k(x)\right)$$

批注 [A9]: sign 是符号函数
使用加权后的符号作为预测结果