

# 机器学习第四次作业

2020年11月

## 1 第一题

请推导Adaboost算法得到的分类器 $h$ 的经验误差界，证明：

$$\hat{\epsilon}(h) \leq \exp \left[ -2 \sum_{t=1}^T \left( \frac{1}{2} - \epsilon_t \right)^2 \right]$$

其中 $\epsilon_t$ 是基分类器 $h_t$ 上的误差， $h(x) = \text{sgn}(f(x))$ ， $f(x) = \sum_{t=1}^T \alpha_t h_t(x)$ 。

## 2 第二题

假设训练集样本为 $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ ，我们希望差异化惩罚不同样本上的误差，为此我们给每一个样本 $(x_i, y_i)$  赋予非负重要性权重 $w_i$ ，并定义目标函数： $F(\alpha) = \sum_{i=1}^n w_i e^{-y_i f(x_i)}$ ，其中 $f(x) = \sum_{j=1}^N \alpha_j h_j(x)$ 。请用该目标函数推导出一种boosting算法。