Machine Learning Homework 6

He Li, 2014012684

2017/11/17

Problem 1. *Prove that* $yf(x) = y\sum_t \alpha_t h_t(x)$ *is a distance of* (x, y).

Proof. We can rewrite the expression as:

$$\frac{\sum_{t=1}^{T} \alpha_t h_t(x)}{\sum_{t=1}^{T} \alpha_t} = \frac{< \alpha, h(x) >}{|\alpha|_1}$$

where $\alpha = (\alpha_1, ..., \alpha_T)$ and $h(x) = (h_1(x), ..., h_T(x))$. Therefore, it can be seen as the distance from $h(x)$ to the line α.

Here margin can be seen as the distance between the predict vector to the linear combination of classifiers. □

Problem 2. *Let* Q *be a distribution of classifiers. For stochastic classifier, stochastic error is* $error_D(f_Q) = \mathbb{P}_{Q,(x,y)}(f(x) \neq y)$. *For voting classifier* $f_v(x) = y \iff \mathbb{P}_Q(f(x) = y) > 1/2$, *the voting error is defined as* $error_D(v_Q) = \mathbb{P}_{(x,y)}(f_v(x) \neq y)$. *Consider the relationship between voting error and stochastic error.*

Proof. We know that if $f_v(x) \neq y$, then

$$\mathbb{P}_Q(f(x) \neq y) > 1/2$$

Let $f(x, y) = \mathbb{P}_Q(f(x) \neq y)$,

$$\text{error}_D(v_Q) = \mathbb{P}_{(x,y)}(f_v(x) \neq y) = \mathbb{E}_{(x,y)}(I[f(x, y) > \frac{1}{2}])$$

also,

$$\begin{aligned}
\text{error}_D(f_Q) &= \mathbb{P}_{Q,(x,y)}(f(x) \neq y) \\
&= \mathbb{E}_{(x,y)}(f(x, y)) \\
&= \frac{1}{2}\mathbb{E}_{(x,y)}(2f(x, y)) \\
&\geq \frac{1}{2}\mathbb{E}_{(x,y)}(I[f(x, y) > 0.5]) \\
&= \frac{1}{2}\text{error}_D(v_Q)
\end{aligned}$$

Therefore,

$$\text{error}_D(f_Q) \geq \frac{1}{2}\text{error}_D(v_Q)$$

□