Homework 5

Name: 方嘉聪 ID: 2200017849

Problem 1. Consider the following AdaBoost algorithm:

Algorithm 1 AdaBoost Algorithm

Require: Data set $(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)$ and a weak learner $\mathcal{A}$

Ensure: A strong classifier $F(x)$.

Initialize $D_0(i) \leftarrow 1/n, \forall i \in [n]$.

for $i = 1$ to T **do**

  Train a classifier h_t on D_t using $\mathcal{A}$, where $h_t \in \{\pm 1\}$.

  $\varepsilon_t \leftarrow \sum_{i=1}^{n} D_t(i) \mathbb{1}[y_i \neq h_t(x)]$

  $\gamma_t \leftarrow 1 - 2\varepsilon_t$, $\alpha_t \leftarrow \frac{1}{2} \ln \frac{1+\gamma_t}{1-\gamma_t}$

  $Z_t \leftarrow \sum_{i=1}^{n} D_t(i) \exp(-\alpha_t y_i h_t(x_i))$, is the normalization factor.

  $D_{t+1}(i) \leftarrow D_t(i) \exp(-\alpha_t y_i h_t(x_i))/Z_t$.

return $F(x) = \sum_{t=1}^{T} \alpha_t h_t(x)$

Prove the following proposition:

Assume $\gamma_t \geq \gamma > 0$ for all $t = 1, 2, \cdots, T$, the training loss of the strong classifier $F(x)$ has a upper bound:

$$\frac{1}{n} \sum_{i=1}^{n} \mathbb{1}[y_i \cdot F(x_i) \leq 0] \leq (1 - \gamma^2)^{T/2}$$

Hint: Try to prove the exponential loss is upper bounded by $(1 - \gamma^2)^{T/2}$. ◀

Solution. We can prove that the training loss is upper bounded by the exponential loss, i.e.,

$$\frac{1}{n} \sum_{i=1}^{n} \mathbb{1}[y_i \cdot F(x_i) \leq 0] \leq \frac{1}{n} \sum_{i=1}^{n} \exp(-y_i F(x_i)) = \prod_{t=1}^{T} Z_t$$

The last equality is proved in the class. And we know that

$$Z_t = 2\sqrt{\varepsilon_t(1 - \varepsilon_t)} = \sqrt{1 - \gamma_t^2}$$

Since $\gamma_t \geq \gamma > 0$, we have

$$\sqrt{1 - \gamma_t^2} \leq \sqrt{1 - \gamma^2}$$

Thus, we have

$$\frac{1}{n} \sum_{i=1}^{n} \mathbb{1}[y_i \cdot F(x_i) \leq 0] \leq \prod_{t=1}^{T} Z_t \leq \prod_{i=1}^{T} \sqrt{1 - \gamma^2} = (1 - \gamma^2)^{T/2}$$

This completes the proof. ◁

Remark: It implies that when $T = \Omega(\log n)$, $(1 - \gamma^2)^{T/2} < 1/n$, thus the training loss is 0.