Weighted Majority and Winnow

Wednesday, October 13, 2021 5:52 AM

Yufan

Predicting from Expert advice

n experts, all make 0/1 predictions

Algo makes predictions based on experts' advice.

M = # mistakes by Algo

m = # mistakes by best expert.

- Predict Majority
- On each mistake remove all experts that got it wrong.
- If $\exists$ perfect expert, $M \le \log n$.

If no perfect expert, after all experts eliminated, restart.

- in each round $M \le \log n$ and best expert makes ≥ 1 mistake

$$\Rightarrow \quad M \le m \log_2 n .$$

(Multiplicative weights)

Weighted Majority (Multiplicative weights)

Start with $w_i = 1$. $\quad (W = \sum_i w_i)$

Predict according to weighted Majority.

On a mistake $w_i \leftarrow \frac{w_i}{2}$ for all experts that erred.

Total weight decreases by $\frac{3}{4}$.

So after M mistakes, $W \leq n\left(\frac{3}{4}\right)^M$

but best expert has $w_i \geq \left(\frac{1}{2}\right)^m$.

$$\Rightarrow \quad n\left(\frac{3}{4}\right)^M \geq \left(\frac{1}{2}\right)^m$$

$$\log_2\left(\frac{4}{3}\right) \cdot M \leq m + \log_2 n \quad \text{i.e. } M \leq \frac{5}{2}(m + \log_2 n).$$

Randomized weighted Majority.

Predict according to expert i with prob. $\frac{w_i}{W}$.

for each expert that makes an error,

W

– for each expert that makes an error,

$$w_i \leftarrow (1-\varepsilon) w_i$$

– Let f_t: fraction of experts (weighted) that make a mistake at time t.

– Total weight at time T, $W_T = n \prod_{t=1}^{T} (1-\varepsilon f_t)$

$$\ln W = \ln n + \sum_t \ln(1-\varepsilon f_t)$$

$$\leq \ln n - \varepsilon \sum_t f_t$$

$$= \ln n - \varepsilon \, \mathbb{E}(M)$$

– Weight of best expert $\geq (1-\varepsilon)^m$.

$$m \ln(1-\varepsilon) \leq \ln n - \varepsilon \, \mathbb{E}(M)$$

$$\mathbb{E}(M) \leq (1+\varepsilon) m + \frac{\ln n}{\varepsilon}$$

– Set $\varepsilon = \sqrt{\frac{\ln n}{m}}$. $\leq m + 2\sqrt{m \ln n}$

$$\frac{\mathbb{E}(M)}{T} \leq \frac{m}{T} + 2 \frac{\sqrt{m \ln n}}{T}$$

$$\leq \frac{m}{T} + 2 \sqrt{\frac{\ln n}{T}} \quad \Big| \, m \leq T$$

$$\leq \frac{m}{T} + 2\underbrace{\sqrt{\frac{\ln n}{T}}}_{\xrightarrow[T\to\infty]{} 0} . \quad \Big| m \leq T$$

WINNOW

Learn an OR of r out of n variables.

- start with $w_i = 1$.
- Predict $+$ if $w_i \cdot X \geq n$
- $-$ otherwise.
- mistake on +ve X, $w_i \leftarrow 2w_i$ if $X_i = 1$
- ——— −ve X, $w_i \leftarrow \frac{1}{2} w_i$ if $X_i = 1$.

positive Mistakes, M_+, $\leq r \log_2 n$ (at least one of the r has $X_i = 1$)

On each +ve mistake, weight goes up by $\leq n$

——— −ve ——— ——— down — $\geq \frac{n}{2}$

$\Rightarrow M_- \leq 2M_+$.

So $M \leq 3r \log_2 n$.

So $M \le 3r\log_2 n$.

Learn a k-out-of-r majority function.

— start $w_i = 1$, predict + if $w \cdot x \ge n$
— otherwise.

— Mistake on +ve example, $w_i \leftarrow (1+\varepsilon) w_i$ if $x_i = 1$

— -ve —, $w_i \leftarrow \dfrac{w_i}{(1+\varepsilon)}$ if $x_i = 1$

Then $kM_+ - (k-1)M_- \le r\log_{1+\varepsilon} n$

On each of the r variables, # $(1+\varepsilon)$ factors is at most $\log_{1+\varepsilon} n$.

And $$n + \underbrace{(\varepsilon n)}_{\text{increase}} M_+ \ge \frac{\varepsilon n M_-}{1+\varepsilon}$$

$$\Rightarrow \quad M_- \le (1+\varepsilon) M_+ + \frac{1+\varepsilon}{\varepsilon}.$$

Using this above,

$$(k - (k-1)(1+\varepsilon))\, M_+ \le r\log n + (k-1)\frac{(1+\varepsilon)}{\varepsilon}$$

$\varepsilon = 1/$[illegible] $\Rightarrow$ $M = O(kr\log n)$

$$\varepsilon = \frac{1}{2(k-1)} \Rightarrow M_+ = O(kr\log n)$$

$$\text{and} \quad M = O(kr\log n).$$

Halfspaces. $w_1^* x_1 + \dots \; w_n^* x_n \geq w_0^*$.

Assume $w_i^* \geq 0$, else set $y_i = 1 - x_i$

Assume w_i^* integer by scaling up.

Duplicate each variable x_i $\sum_i w_i^*$ times.

So now we have the w_0^* out of $W = \sum_{i=1}^{n} w_i^*$ problem.

$$\#\text{mistakes} = O\left(w_0^* \cdot W \log(nW)\right)$$
$$= O\left(W^2 \log(nW)\right).$$

Generalizing to arbitrary w, x, with $\gamma = \min_x |w^* \cdot x|$

$$\#\text{mistakes} = O\left(\frac{\|w^*\|_1^2 \|x\|_\infty^2 \log(\|w^*\|_1 n)}{\gamma^2}\right).$$

Perceptron vs Winnow

$x \in \{0,1\}^n$

1. $w^* = (\underbrace{1, 1, \ldots, 1}_{k}, 0 \ldots 0)$ $\quad x \in \{0,1\}^n$

WINNOW: $O\left(\frac{k^2 \log n}{\gamma^2}\right)$ $\quad$ PERCEPTRON: $O\left(\frac{k \cdot n}{\gamma^2}\right)$

2. $w^* = \left(\frac{1}{\sqrt{n}}, \ldots \frac{1}{\sqrt{n}}\right)$ $\quad \|x\|_2 = 1$

WINNOW: $O\left(\frac{n \log n}{\gamma^2}\right)$ $\quad$ PERCEPTRON: $O\left(\frac{1}{\gamma^2}\right)$