

Foundations of Machine Learning — Homework Assignment 1

Anirudhan J Rajagopalan
N18824115
ajr619

October 11, 2015

C. Randomized Halving

1. Psuedo code

Algorithm 1 Randomized Halving

```

1:  $H_1 \leftarrow H$ 
2: for  $t \leftarrow 1$  to  $T$  do
3:    $RECEIVE(x_t)$ 
4:    $r_t \leftarrow \frac{\sum_{i: y_{t,i}=1} 1}{|H_t|}$ 
5:    $p_t \leftarrow 1$ 
6:   if  $r_t \leq \frac{3}{4}$  then
7:      $p_t \leftarrow \lceil \frac{1}{2} \log_2 \frac{1}{1-r_t} \rceil$ 
8:    $\hat{y}_t \leftarrow GetRandomNumberWithProbability([1, 0], [p_t, 1 - p_t])$ 
9:    $RECEIVE(y_t)$ 
10:  if  $\hat{y}_t \neq y_t$  then
11:     $H_{t+1} \leftarrow \{C \in H_t : C(x_t) = y_t\}$ 
return  $H_{T+1}$ 

```

2. Prove $\forall t \geq 1, E[\mu] \leq \frac{\phi_t - \phi_{t+1}}{2}$

Given: Potential function: $\phi_t = \log_2 |H_t|$ and $\mu_t = 1_{y_t \neq \hat{y}_t}$

Proof:

We are only considering the case when the predicted value \hat{y}_t is not equal to the received value y_t . The value of expectation can be written as

$$\begin{aligned} E[\mu_t] &= p_t * 1 + (1 - p_t) * 0 \\ &= p_t * 1 \end{aligned}$$

The probability of predicting 1 by the randomized algorithm is the probability of making a mistake since we are only considering the cases in which we make mistakes ($\mu_t = 1_{y_t \neq \hat{y}_t}$)

Therefore,

$$\begin{aligned} E[\mu_t] &= p_t \\ &= \lceil \frac{1}{2} \log_2 \frac{1}{1-r_t} \rceil 1_{r_t \leq \frac{3}{4}} + 1_{r_t > \frac{3}{4}} \\ &\leq \lceil \frac{1}{2} \log_2 \frac{1}{1-r_t} \rceil \end{aligned}$$

Since, the Expectation will be 1 when $r_t > \frac{3}{4}$ which corresponds to the maximum expectation here, we can upper bound the expectation by using $r_t \leq \frac{3}{4}$ as $\lceil \frac{1}{2} \log_2 \frac{1}{1-r_t} \rceil$ equals 1 when $r_t = \frac{3}{4}$

$$\begin{aligned}
 E[\mu_t] &\leq \left\lfloor \frac{1}{2} \log_2 \frac{1}{1-r_t} \right\rfloor \\
 &= \left\lfloor \frac{\log_2 \frac{1}{1-r_t}}{2} \right\rfloor \\
 &= \left\lfloor \frac{\log_2 \frac{1}{1-r_t}}{2} \right\rfloor \\
 &= \left\lfloor \frac{\log_2 \frac{1}{1-\frac{|H_1|}{|H_t|}}}{2} \right\rfloor \\
 &= \left\lfloor \frac{\log_2 \frac{|H_t|}{|H_t|-|H_1|}}{2} \right\rfloor \\
 &= \left\lfloor \frac{\log_2 \frac{|H_t|}{|H_0|}}{2} \right\rfloor \\
 &= \left\lfloor \frac{\log_2 |H_t| - \log_2 |H_0|}{2} \right\rfloor \\
 &= \left\lfloor \frac{\log_2 |H_t| - \log_2 |H_{t+1}|}{2} \right\rfloor \\
 &= \left\lfloor \frac{\phi_t - \phi_{t+1}}{2} \right\rfloor \\
 \therefore E[\mu_t] &\leq \frac{\phi_t - \phi_{t+1}}{2}
 \end{aligned}$$

3. Expected number of mistakes.

Given: N be the total number of experts at the beginnning of the iterations (denoted by H_1). Since we are considering a relizable scenario at the end of the algorithm the number of experts should be atleast one. So $H_T = 1$.

To Prove: The expected number of mistakes made by Randomized Halving is at most $\frac{1}{2} \log_2 N$

Proof:

Lets consider the total expectation of mistakes of the Randomized Halving

algorithm over T iterations.

$$\begin{aligned}
E[\mu_T] &\leq \sum_{t=1}^T \frac{\phi_t - \phi_{t+1}}{2} \\
&\leq \frac{(\phi_{H_1} - \phi_{H_2}) + (\phi_{H_2} - \phi_{H_3}) \cdots (\phi_{H_{T-2}} - \phi_{H_{T-1}}) + (\phi_{H_{T-1}} - \phi_{H_T})}{2} \\
&\leq \frac{\phi_{H_1} - \phi_{H_T}}{2} \\
&\leq \frac{\phi_N - \phi_T}{2} \\
&\leq \frac{\phi_N}{2} \\
&\leq \frac{1}{2} \log_2 N
\end{aligned}$$

Hence proved.

Here $\phi_T = 0$ because the number of experts at line T is 1. Therefore, log of T will be zero.

4. [Bonus Question]

As we have seen in the previous answer, the mistakes made by the randomized algorithm is bounded by $\frac{1}{2} \log_2 N$. This upper bound is dependent only on the number of initial experts, N . Therefore any randomized algorithm that is dependent on the opinion of the experts to generate its predictions will have similar upper bound of $\lfloor \frac{1}{2} \log_2 N \rfloor$. The floor function is used as mistakes are natural numbers.