639 Problem Set One

Yanpeng Zhao

October 2020

1 The First Problem

a) Disjunctions and Conjunctions are special cases because any form of conjunctions and disjunctions can be represented in decision list. And if any element in disjunctions is 1, then out put 1, else output 0:

$$[(x_0, 1), (x_1, 1), (x_2, 1)...(x_k, 1), 0]$$

If any element in conjunctions is 0, then output 0, else output 1:

$$[(\overline{x_0}, 0), (\overline{x_1}, 0), (\overline{x_2}, 0)...(\overline{x_k}, 0), 1]$$

b) Write decision list into linear thresh hold function:

$$[(w_1, x_1, b_1), (w_2, x_2, b_2), ...(w_k, x_k, b_k), b_k + 1]$$

$$f(x) = +iff w_1x_1 + ... + w_nx_n >= 0$$

but for decision list writing into linear thresh hold function, we need to make sure w1 is greater than the sum of other weights that

$$|w_1| > |w_2| + |w_3| + ... + |w_k|$$

2 The Second Problem

suppose we now have a algorithm A having $(1 + \epsilon)M_c + O(size(c))$ Mistake bound over sample $x_1, x_2, x_3, ..., x_n$ And $A^{'}$ is similar to A but $A^{'}$ will check $C(x_i) = h_{i+1}(x_i)$ when $A^{'}$ gets to i^{th} term. So we have the following algorithm:
1.Run $A^{'}$ on sample $x_1, x_2, ...x_n$.
2.Each time reconstructing a hypothesis h_i run the hypothesis on x_{i+1}
3.If $A^{'}$doesn't make any mistake, keep current hypothesis h_i, and terminate. Else, continue running $A^{'}$
4.After all x, output current hypothesis of $A^{'}$
In this case, because we have exactly same hypothesises, we are having the same mistake bound $(1 + \epsilon)M_c + O(size(c))$

3 The Third Problem

a)For computing mistake bound of winnow algorithm, we need to add mistake bound of positive, negative and correct cases together.
Positive case: let $l$be number of times i being doubled. Since weight can never exceed n. We have:

$$2^{l-1} < n$$

$$l < log_n + 1$$

The total will be $k(log_n + 1)$ In demotion process, each promotion increase the weight by $(\alpha - 1)\theta$ Each elimination decrease $\sum_{i=1}^{n} w_i$ at level θ j is the number of promotion in correct case.

$$0 \leq \sum_{i=1}^{n} w_i \leq n + \theta(\alpha - 1)u + \theta(\alpha - 1)j - \theta v$$

$$v \leq \frac{n}{\theta} + (\alpha - 1)(u + j)$$

$$v \leq \frac{n}{\theta} + (\alpha - 1)2\alpha k(log\theta + 1)$$

total:

$$k(\log n + 1) + k(\log n + 1) + (2\alpha k - 2k)(\log \theta + 1) + \frac{n}{\theta}$$

$$= 2\alpha k(\log \theta + 1) + \frac{n}{\theta}$$

in this case α is 2, θ is 1, we get final mistake bound $= 4k(\log n + 1) + 1$
b) the upper bound is $2k \log(n + 1) + 1 \geq k \log n$
lower bound is $k \log \frac{n}{k} \leq k \log n$ so it is possible to make $\Omega(k \log n)$

4 The Forth Problem

we know $\omega \leftarrow \omega - \omega(\omega \cdot x)x$ so, we can have $||\omega^{'}||^2 = ||\omega||^2 - 2||\omega(\omega x)x|| + ||(\omega x)x||^2$ because $|x|_2 = 1$ we can eliminate x term and add x as we want. Finally, $||\omega^{'}|| = ||\omega||^2 - ||\omega \cdot x||^2$ Divide this for formula with $||\omega||^2$: $\frac{||\omega^{'}||}{||\omega||} = 1 - \frac{||\omega \cdot x||^2}{||\omega||^2} \leq 1 - \delta^2$ and after m mistakes $||\omega||^2 \leq (1 - \delta^2)^m$, $||\omega|| \leq \sqrt{(1 - \delta^2)^m}$ Because, $\omega - \omega_0 = -(\omega \cdot)x$ and sum each difference after making m mistakes: $\omega - \omega_0 = \sum_{i=0}^{m} -(\omega \cdot x)x$,$\omega v = -\sum_{i=0}^{m}(w \cdot x)xv + \omega_0 v$ because $\omega_0 \cdot v \geq \gamma$ and $||v||_2 = 1$, $||\omega|| \geq \gamma$ bringing $||\omega|| \leq \sqrt{(1 - \delta^2)^m}$ into this formula, we get $\sqrt{(1 - \delta^2)^m} \geq \gamma$, ln()both side:$\frac{m}{2} \ln(1 - \delta^2) \geq \ln(\gamma)$,because $\ln(1 - \delta^2) \leq \delta^2$, we finally get $\frac{m}{2}\delta^2 \geq \ln(\gamma)$, $\delta^2 \geq \frac{2\ln(\gamma)}{m}$, $m \leq \frac{2\ln(\gamma)}{\delta^2}$

5 The Fifth Problem

when each example have super small angle changes between each other. The mistake bound could be higher. $[(<1,0,0,...>,+),(<\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}},0,0,...>,+),(<\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}}0,0,0,...>,+)...]$ and $C=[\frac{n}{\sqrt{\frac{n^2+n}{2}}},\frac{n-1}{\sqrt{\frac{n^2+n}{2}}},\frac{n-2}{\sqrt{\frac{n^2+n}{2}}}...]$ In this case, it could make $\Omega(1/\gamma^2)$

6 The sixth Problem

Algorithm:

$$\omega_0 = P_1, P_2, P_3...P_n$$

and the algorithm predict based on majority that output$=1$ if $\sum h_i(x)=1 > \sum h_i(x)=0$, else output$=0$. When it doesn't make mistake, do nothing. When it makes mistake, eliminate all P predict wrong.
Because we eliminate a half of programs each time the algorithm makes mistake: $||\omega||_2 \leq 2^{-m}(n)$ which $m \leq \log(n)$and before finding f_t, we make $O(\log(t))$ mistake