$\boxed{\text{4 Sep 2024}}$ The Multiplicative Weights Algorithm

Binary Labeling With Expert Advice

Time proceeds in rounds $1, 2, \ldots, T$.
There are K experts, $1, \ldots, K$.
There are 2 labels, $0, 1$

Each round:

- Each expert recommends a label.
- Algorithm guesses label after seeing recommendations.
- Correct label revealed.

Goal: Minimize alg # of mistakes.

(I) Suppose $\exists$ a perfectly reliable expert who makes no mistakes.

The MAJORITY algorithm works as follows.

- Initialize $w_i = 1$ for all $i \in [K]$.
- Set an expert's weight to $\emptyset$ whenever it makes a mistake.
- Predict label $\emptyset$ or 1 according to a weighted majority vote of experts.

This makes $\leq \lfloor \log_2(K) \rfloor$ mistakes.

(II) No guarantee that $\exists$ perfect expert.
We want to make
not-too-many-more mistakes
than best expert.

The WEIGHTED MAJORITY algorithm

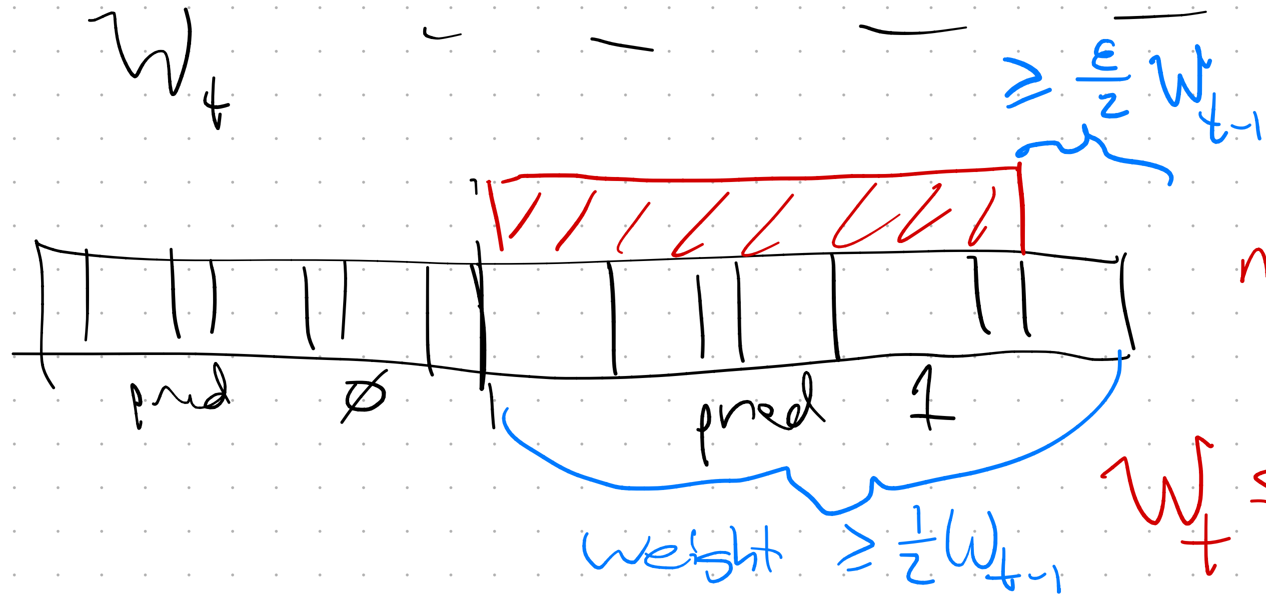
- Initialize $w_i = 1 \quad \forall\, i \in [K]$
- Whenever an expert makes a mistake, multiply w_i by $1-\epsilon$.
- Predict labels by taking weighted majority vote.

  predict $\emptyset$ if

  $$\sum_{i \text{ predicting } \emptyset} w_i > \sum_{i \text{ predicting } 1} w_i$$

  else predict 1.

If W_{t-1} is combined weight after $t-1$ rounds
W_t — — — — — t rounds



mistake

$\Downarrow$

$$W_t \leq \left(1 - \frac{\epsilon}{2}\right) W_{t-1}$$

If $\exists$ an expert that makes only m mistakes, and WM alg. makes m_{ALG} mistakes

$$(1-\varepsilon)^m < W_T \leq W_0 \cdot \left(1-\frac{\varepsilon}{2}\right)^{m_{ALG}}$$

$$W_0 = K$$

$$m \ln(1-\varepsilon) < \ln K + m_{ALG} \ln\left(1-\frac{\varepsilon}{2}\right)$$

$$-m_{ALG} \ln\left(1-\frac{\varepsilon}{2}\right) < \ln K - m \ln(1-\varepsilon)$$

$$m_{ALG} < -\frac{\ln K}{\ln(1-\varepsilon/2)} + m \frac{\ln(1-\varepsilon)}{\ln\left(1-\frac{\varepsilon}{2}\right)}$$

For $0 < \varepsilon < 1/2$

$$\varepsilon < -\ln(1-\varepsilon) < \varepsilon + \varepsilon^2$$

$$< \frac{2 \ln K}{\varepsilon} + 2(1+\varepsilon) m$$