02-750 Week 2

Automation of Scientific Research

Aidan Jan

January 22, 2026

Toy Online Learning Algorithm, Halving (Continued)

Regret Analysis

- Let $W_t = \sum_{h_i \in \mathcal{H}} w_i(t)$ be the total weight at step t (after re-weighting)
- Case analysis summary:
  - If $\hat{y} = y_t$, then $W_t \leq W_{t-1}$ and $R^t = R^{t-1}$. i.e., total weight *may* decrease, but the regret stays the same
  - If $\hat{y} \neq y_t$, then $W_t \leq \frac{W_{t-1}}{2}$ and $R^t = 1 + R^{t-1}$. i.e., the total weight *must* decrease by a factor of at least 2, because we took the **majority**, and they were wrong, and the total regret increases.

Best case analysis at each iteration:

- $t = 0$, $W_0 = |\mathcal{H}|$ and $R^0 = 0$
- $t = 1$, (Best case) if $\hat{y} = y_t$, then $W_1 \leq W_0$ and $R^1 = R^0 + 0 = 0$
- $t = 2$, (Best case) if $\hat{y} = y_t$, then $W_2 \leq W_1$ and $R^2 = R^1 + 0 = 0$
- $t = i$, (Best case) if $\hat{y} = y_t$, then $W_i \leq W_{i-1}$ and $R^i = R^{i-1} + 0 = 0$

Worst case analysis at each iteration:

- $t = 0$, $W_0 = |\mathcal{H}|$ and $R^0 = 0$
- $t = 1$, (Worst case) if $\hat{y} \neq y_t$, then $1 \leq W_1 \leq \frac{W_0 = |\mathcal{H}|}{2}$ and $R^1 = R^0 + 1 = 1$
- $t = 2$, (Worst case) if $\hat{y} \neq y_t$, then $1 \leq W_2 \leq \frac{W_0 = |\mathcal{H}|}{2^2}$ and $R^2 = R^1 + 1 = 2$
- $t = i$, (Worst case) if $\hat{y} \neq y_t$, then $1 \leq W_i \leq \frac{W_0 = |\mathcal{H}|}{2^i}$ and $R^i = R^{i-1} + 1 = i$

We know that each time the majority is wrong, $W_t \leq \frac{W_{t-1}}{2}$. Thus, $W_t \leq \frac{|\mathcal{H}|}{2^{M_t}}$, where M_t is the number of mistakes made by the majority up to time t (proof by induction)

We also know that $W_t \geq 1$ because there *is* a perfect model, and so at least one model always has weight 1. Thus, $1 \leq W_t \leq \frac{|\mathcal{H}|}{2^{M_t}}$ which implies that $M_t \leq \log_2 |\mathcal{H}| \in o(T)$. In summary, the number of mistakes is $o(T)$. Under our assumptions, the number of mistakes is the same as the regret, so the toy algorithm is no regret.

Weighted Majority Algorithm

Of course, the toy algorithm is not very useful in practice, because it assumes that $\mathcal{H}$ contains a perfect model. The Weighted Majority Algorithm (WMA) does not make that assumption.

- The two key differences between the toy algorithm and WMA are:
  - The weights are real-valued in the semi-closed interval, $(0, 1]$
  - The weights are updated using a different rule:

$$w_i(t) = \begin{cases} w_i(t-1) & \text{if } h_i(x_t) = y_t \\ \frac{w_i(t-1)}{2} & \text{if } h_i(x_t) \neq y_t \end{cases}$$

  - That is, half the model's weight, each time it makes a mistake.

Like the toy algorithm, predictions are made via **weighted majority**, but now these weights are real-valued. That is, for each new training instance, (x_t, y_t)

1. Output the majority label from model $\hat{h}_t(x_t)$

$$\hat{y} = \arg\max_{c \in \mathfrak{C}} \sum_{h_i \in \mathcal{H}} w_i \cdot \mathbb{1}(h_i(x_t) = c)$$

2. Re-weight each model in $\mathcal{H}$ using the following rule:

$$w_i(t) = \begin{cases} w_i(t-1) & \text{if } h_i(x_t) = y_t \\ \frac{w_i(t-1)}{2} & \text{if } h_i(x_t) \neq y_t \end{cases}$$

Example

Suppose we have a collection of models for predicting whether two molecules bind.

h_1	h_2	h_3	w_1	w_2	w_3	Weighted vote for $y_t = 0$	Weighted vote for $y_t = 1$	$\hat{h}_i(x_t) = \hat{y}$	y_t
1	0	1	1	1	1	1	2	1	1
0	1	0	1	0.5	1	2	0.5	0	1
1	0	1	0.5	0.5	0.5	0.5	1	1	1
0	1	1	0.5	0.25	0.5	0.5	0.75	1	1
1	0	1	0.25	0.25	0.5	0.25	0.75	1	1

Regret Analysis (after T steps)

- Let m be the number of mistakes by the best model $h^* \in \mathcal{H}$. (offline learning)
- Let M be the number of mistakes by the WMA weighted majority. (online learning)
- Let $W_t = \sum_{h_i \in \mathcal{H}} w_i(t)$ be the total weight after t steps, where
  - $w_i(t) = \left(\frac{1}{2}\right)^{m_i(t)}$ is the weight of the i-th model after the first t steps, and $m_i(t)$ is the number of mistakes made by the i-th model during the first t steps.
- $W_0 = |\mathcal{H}|$ because we initialized each weight to 1.

Claim: A mistake on step t implies that $W_t \leq \left(\frac{3}{4}\right) W_{t-1}$

Proof:

- The mass associated with the majority is **at least** $\left(\frac{1}{2}\right) W_{t-1}$
- The majority was wrong, so we reduce the combined weight of the majority by half (thus, the reduction is *at least* $\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) W_{t-1} = \left(\frac{1}{4}\right) W_{t-1}$)
- Therefore, $W_t \leq \left(\frac{3}{4}\right) W_{t-1}$

Therefore,

- After T steps, the total weight is *at most* $W_t \leq |\mathcal{H}| \left(\frac{3}{4}\right)^M$
- After T steps, the best model will have weight $w^* = \left(\frac{1}{2}\right)^m$
- Thus, $\left(\frac{1}{2}\right)^m \leq W_t \leq |\mathcal{H}| \left(\frac{3}{4}\right)^M$
- Solve for M, we get

$$m \log_2 0.5 \leq \log_2 |\mathcal{H}| + M \log_2 (0.75) M \leq 2.38 \log_2 |\mathcal{H}| + 2.38m$$

- Since M is bounded by the above equation, the worst case= regret for WMA is

$$R^T = (2.38 \log_2 |\mathcal{H}| + 2.38m) - m = 2.38 \log_2 |\mathcal{H}| + 1.38m$$

- This presents a theoretical problem, because m **could be** $O(T)$, thus

$$\lim_{T \to \infty} \frac{R^T}{T} = \frac{O(T)}{T} > 0$$

- WMA is **not** a no-regret algorithm.

Interesting Theoretical Result

If $\mathcal{H}$ does **not** contain a perfect model, then one can show that **no deterministic algorithm can achieve** $R^T \in o(T)$ using a 0-1 loss function

- That is, there are no deterministic online algorithms that achieve no-regret, unless $\mathcal{H}$ contains a perfect model (like our *Toy algorithm*)
- Proof sketch for any deterministic learning algorithm
  - (Online Learning) Let $L^T = \sum_{t=1}^{T} \mathcal{L}(h_t, x_t, y_t)$ be the loss for the online algorithm over T instances
  - An adversary can always select instances such that $L^T = T$
  - (Offline Learning) Over the same set of T instances, the best model $h^* \in \mathcal{H}$ (offline) can be wrong no more than 50% of the time, otherwise we could create a better model by simply outputting the opposite prediction. Thus, $L^* = \sum_{t=1}^{T} \mathcal{L}(h^*, x_t, y_t) \leq \frac{T}{2}$
  - Hence, $R^T = L^- L^* \geq \frac{T}{2} = O(T) \to \lim_{t \to \infty} \frac{R^T}{T} = \frac{O(T)}{T} > 0$

The next algorithm addresses this challenge by using randomization.

Hedge Algorithm (aka. Randomized WMA)

- The Hedge algorithm is basically a **randomized** version of WMA
  - Rather than outputting the weighted majority class, Hedge selects a model at random
  - The probability that any particular model is selected is proportional to its relative weight
- This seemingly simple change has a profound effect, in terms of regret
  - Specifically, Hedge is a no regret algorithm

The Algorithm

- Let $(1-\epsilon)$ be the **learning rate**. Note: the effective learning rate of WMA was 0.5.
- Initialize all weights to 1
- For each new training instance, (x_t, y_t)
  1. Select model **at random** from the multinomial distribution
  $$h_t \sim Multinomial_t(p_1, \cdot, p_n)$$
  where $p_i = \frac{w_i(t)}{W_t}$ and $n = |\mathcal{H}|$
  2. Output h_t's prediction: $\hat{y}_t = h_t(x_t)$
  3. Reweight *all* models $h_t \in \mathcal{H}$ as follows:
  $$w_i(t) = w_i(t-1)(1-\epsilon)^{l(h_i, x_t, y_t)}$$
     - Assume 0-1 loss.

Example

Suppose we have a collection of models for predicting whether two molecules bind. $(\epsilon = 0.5)$

h_1	h_2	h_3	w_1	w_2	w_3	α = Random selection at step t	$\hat{h}_\alpha(x_t) = \hat{y}$	y_t
1	0	1	1	1	1	2	0	1
0	1	0	1	0.5	1	3	0	1
1	0	1	0.5	0.5	0.5	1	1	1
0	1	1	0.5	0.25	0.5	3	1	1
1	0	1	0.25	0.25	0.5	3	1	1

Green identifies the model randomly selected in round t

Regret Analysis (after T steps)

- Let m be the number of mistakes by the best model $h^* \in \mathcal{H}$ (Offline learning)
- Let M be the number of mistakes by the Hedge algorithm (Online learning)
- Let $W_t = \sum_{h_i \in \mathcal{H}} w_i(t)$ be the total weight after t steps, where
  - $w_i(t) = (1-\epsilon)^{m_i(t)}$ is the weight of the i-th model after the first t steps, and $m_i(t)$ is the number of mistakes made by the i-th model during the first t steps.
- $W_0 = |\mathcal{H}|$ because we initialized each weight to 1
- Let $l_i(t) = \mathcal{L}(h_i, x_t, y_t)$ (i.e., the 0-1 loss of the i-th model on sample t)

Summary

- Online learning gives rise to the concept of **regret**
- Some online learning algorithms can achieve no regret
  - That is, there is no significant penalty associated with such algorithms, relative to offline learning
- Examples:

  - The toy algorithm was no regret, but assumes that there is a perfect model
  - WMA **is not** a no regret algorithm. Indeed, no deterministic algorithm can achieve no regret for the 0-1 loss, if the hypothesis space doesn't have a perfect model

- The Hedge algorithm is a relatively minor modification to WMA that employs randomization and restores the no regret guarantee.