CS272 - Theoretical Foundations of Learning, Decisions, and Games

Lecture 8: Online Learning with Smoothed Adversaries

September 23, 2025

Lecturer: Nika Haghtalab *Readings: Haghtalab et al.* JACM 2024
Scribe: Isabel Agostino, Ryan Cheng

1 Lecture Overview

From our previous lectures, we covered both Offline and Online learning, with the results of bounds of the Regret when the consistency assumption doesn't hold and does hold plotted in the table below. Relevant papers where these bounds are derived are cited in the table. As a reminder, $\tilde{\Theta}$ indicates that that particular bound is tight to some polylogarithmic function of T.

Our goal for this lecture is to explore a middle ground between Offline and Online learning, since we observed that the Offline case is rather easy while the Online case is exceedingly difficult. This was seen in the analysis of the Littlestone dimension (LDim) of 1-dimensional thresholds—an example that we will return to in motivating the case for a middle ground. Indeed, the Littlestone dimension of most natural classes in machine learning is infinite, as whenever a class embeds a *thresholding* behavior (as we often strive for in classification), the class is known to suffer from infinite Littlestone behavior.

We will see that a middle ground can be constructed by constraining the adversary to have a "shaking hand" — in the sense that they are still able to adversarially construct the data and labels, as in the online setting, but there is an added random perturbation to the data outside the adversary's control.

	Regret Agnostic	Regret/Mistake Bound with "consistency" assumption
Offline	$\tilde{\Theta}(\sqrt{T \cdot \text{VCD}})$ Haussler [1992]	$\tilde{\Theta}(\text{VCD} \cdot \ln T)$ Valiant [1984]
σ-smoothed Adversary	$\mathcal{O}\left(\sqrt{T \cdot \text{VCD} \cdot \ln(\frac{1}{\sigma})}\right)$ Haghtalab et al. [2024]	$\mathcal{O}(\text{VCD} \cdot \text{poly}\log(\frac{T}{\sigma}))$
Online	$\tilde{\Theta}(\sqrt{T \cdot \text{LDim}})$ Ben-David et al. [2009] Alon et al. [2021]	$\Theta(\text{LDim})$ Littlestone [1988] Littlestone and Warmuth [1994]

As a reminder, the main difference between the Offline and Online paradigms is that the data in the Offline setting are drawn independently and identically from the same distribution, while Online data can be arbitrarily dependent on past data. In practice, allowing dependence on the past is important for prediction and generation. Examples include stock forecasting, where future performance can depend on historical performance and predictions, and text generation, where the next token generated is conditioned on the history of the text up to that point.

We will see that the limits of our model for a middle ground—the σ-smoothed Adversary—reduce to the Offline and Online cases.

- Offline: x_t is i.i.d. from $x_1, \ldots, x_{t-1}$
- Online: x_t could be dependent on $x_1, \ldots, x_{t-1}$

2 Middle Ground in 1D Thresholding

We would like to construct a paradigm that is weaker than online learning, but still one that allows future instances to depend on earlier instances. We will focus on the one-dimensional threshold as a canonical example of a hard class to online learn (Littlestone dimension of infinity) in our initial investigation.

Example: Shaky Hand Adversary

Let's consider a situation where x_{t+1} can depend on $x_1, \ldots, x_t$, but there is still randomness. Suppose our adversary thinks of $\hat{x}_{t+1}$ based on $x_1, \ldots, x_t$, but then their hand shakes when they write it down, so there is error. Mathematically, we can think of this as:

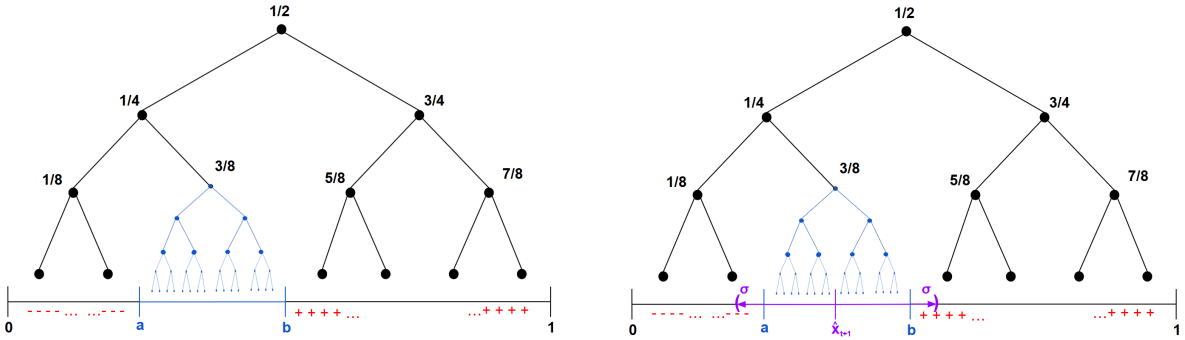
$$x_{t+1} \sim \text{Uniform}[\hat{x}_{t+1} - \sigma, \hat{x}_{t+1} + \sigma]$$

for a parameter σ. This is a natural assumption to make since in real life, all the measurements we take are approximations.

Informal claim: The adversary cannot use the infinite-depth shattered tree strategy we saw in the previous lecture to inflict an infinite number of mistakes in the consistency model.

We analyze the example of the one-dimensional threshold game studied in the last lecture, where the adversary queries the learner on the labels of adaptive instances that form a binary search tree. The shattered tree in this sense is a tree like the one shown in Figure 1(a). Note that the adversary traverses the tree such that x_{t+1} is the right child of x_t if $y_t = -1$, and the left child of x_t if $y_t = +1$. Moreover, y_t is chosen at every round as the opposite of the learner's prediction $\hat{y}_t$.

Let $[a_t, b_t]$ be the interval between the rightmost negatively labeled point and the leftmost positively labeled point—that is, the interval outside of which the learner knows the labels of future



(a) Worst-Case Adversary

(b) Shaky Hand Adversary

points with certainty. No matter how small $[a_t, b_t]$ is, a worst-case adversary can recursively construct a shattered tree within such an interval. Therefore, this worst-case adversary can force the learner to make an arbitrarily large number of mistakes.

Now consider the adversary with shaky hands, as in Figure 1b. Note that when $|a - b| \ll \sigma$, no matter what $\hat{x}_t$ the adversary chooses, it is very likely that $x_t \notin [a, b]$. Since, outside of $[a, b]$, the learner knows the label of the queried point with certainty (due to the promise of consistency with some threshold function), such a query can be answered with no mistake on the part of the learner. This limits the adversary's power for recursively going down the shattered tree, and in some way, σ defines an interval or resolution within which the adversary does not have control.

3 Smoothed Adaptive Adversary

We now introduce a more general model of adaptive adversaries inspired by the above example. Note that the main limiting factor on the adversary was that they could not focus on arbitrarily small areas of the domain. In other words, they could not concentrate their queries too much in a way that was unpredictable to the learner.

Let μ be some base measure (e.g., a uniform distribution or Gaussian). A smoothed adversary operates as follows:

1. At time t, the adversary picks a distribution D_t on $\mathcal{X}$ (depending on the history thus far), with the only constraint being that D_t is σ-smooth with respect to μ.

2. $x_t \sim D_t$ and is revealed to the learner.

3. The learner predicts $\hat{y}_t$ and later observes y_t.

Definition 3.1 (σ-smoothness)**.** A distribution D is **σ-smooth** with respect to μ if $\forall A \subseteq \mathcal{X}$, $D(A) \leq \frac{\mu(A)}{\sigma}$ where $0 < \sigma \leq 1$.

Remark 3.2. Here μ is known in advance to the learner. The definition of σ-smooth can be generalized to requiring $\frac{d\vec{p}}{d\mu} \leq \frac{1}{\sigma}$ where $\frac{d\vec{p}}{d\mu}$ is the Radon-Nikodym derivative.

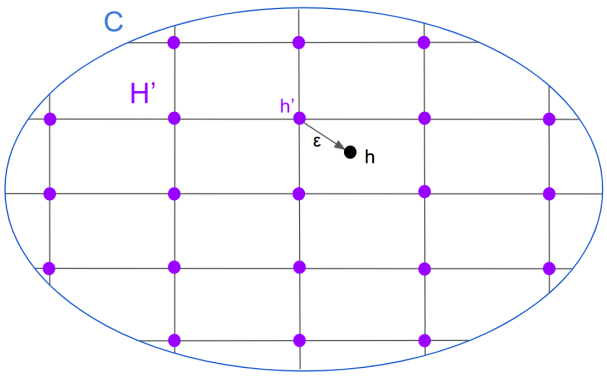


Figure 2: A representation of the finite set H', which is chosen as an ϵ-net of the concept space C with respect to μ for some ϵ. For all $h \in C, \exists h' \in H'$, which is at most ϵ distance away from h, as measured by their agreement on μ (i.e., $\mathbb{P}_{x\sim\mu}(h(x) \neq h'(x))$).

Theorem 3.3. *Let $\mathcal{D}$ denote the joint distribution on $x_1, \ldots, x_T$ where T is the final time. There is an algorithm in online learning with smooth adaptive adversaries with expected regret*

$$\mathbb{E}_{\mathcal{D}}\left[\sum_{t=1}^{T} \mathbb{1}(\hat{y}_t \neq y_t) - \inf_{h\in C} \sum_{t=1}^{T} \mathbb{1}(h(x_t) \neq y_t)\right] \leq \mathcal{O}\left(\sqrt{VCD \cdot T \cdot \ln(\tfrac{1}{\sigma})}\right).$$

Next, we will sketch a proof for this theorem with a more relaxed bound $\tilde{\mathcal{O}}(\sqrt{\text{VCD} \cdot T/\sigma})$.

Proof Sketch of Theorem 3.3: Let $H' \subseteq C$ be some finite set. Ignore $C \setminus H'$ and play randomized weighted majority (RWM)/multiplicative weight update (MWU) algorithm on H'. Then the regret of this algorithm is

$$\text{regret} \leq \underbrace{\mathop{\mathbb{E}}_{\mathcal{D}}\left[\sum_{t=1}^{T} \mathbb{1}(\hat{y}_t \neq y_t) - \inf_{h'\in H'} \sum_{t=1}^{T} \mathbb{1}(h'(x_t) \neq y_t)\right]}_{\text{expected regret in mistake-bound setting} \leq \sqrt{T\log(H')}}$$
$$+ \underbrace{\mathop{\mathbb{E}}_{\mathcal{D}}\left[\inf_{h'\in H'} \sum_{t=1}^{T} \mathbb{1}(h'(x_t) \neq y_t) - \inf_{h\in C} \sum_{t=1}^{T} \mathbb{1}(h(x_t) \neq y_t)\right]}_{A_T = \text{approximation error of } H'}$$

What is a good choice of H'? $|H'| \leq (\frac{1}{\epsilon})^{\text{VCD}(C)}, (\frac{\text{VCD}(C)}{\epsilon})^{\text{VCD}(C)}, 2^{\text{VCD}(C)}$ are all acceptable because we want $\log(H') < \tilde{\mathcal{O}}(\text{VCD}(C))$. What we want is for H' to have small approximation error (i.e., $\sqrt{\text{VCD}(C) \cdot T \cdot \textit{something}}$ with $\textit{something} \propto \frac{1}{\sigma}$ or $\log(\frac{1}{\sigma})$.

It is important for the approximation error to be small when measured over instances drawn from $\mathcal{D}$. But this is difficult to discuss due to the dependencies between the x_ts in this distribution. So, to get some more intuition we start with a scenario where $\mathcal{D}$ is the distribution of i.i.d samples

from μ? How could we ensure that

$$\mathbb{E}\left[\inf_{h'\in H'}\sum_{t=1}^{T}\mathbb{1}(h'(x_t)\neq y_t)-\inf_{h\in C}\sum_{t=1}^{T}\mathbb{1}(h(x_t)\neq y_t)\right]\leq\mathbb{E}\left[\sup_{h\in C}\inf_{h'\in H'}\sum_{t=1}^{T}\mathbb{1}(h(x_t)\neq h'(x_t))]\right]$$

is small? We introduce the following lemma, which considers this case.

Lemma 3.4 (Haussler)**.** *For any distribution μ, there is $H'\subseteq C$ of size $|H'|\leq(\frac{1}{\epsilon})^{41\cdot VCD(C)}$ such that*

$$\forall h\in C,\exists h'\in H' \text{ s.t. } \mathbb{P}_\mu(h(x)\neq h'(x))\leq\epsilon. \tag{1}$$

Remark 3.5. Statement 1 is that H' is an ϵ-cover (or ϵ-net) of C with respect to μ.

Proof Sketch of Theorem 3.3 (cont.): From now on, take H' to be an ϵ-net of C with respect to μ for some ϵ.

Claim: When $H'\subseteq C$ is an ϵ-net for C with respect to μ, its expected approximation error (A_T) with respect to C and any $\mathcal{D}$ that is σ-smooth is also small:

$$A_T\leq\mathbb{E}_{\mathcal{D}}\left[\sup_{h\in C}\inf_{h'\in H'}\sum_{t=1}^{T}\mathbb{1}(h(x_t)\neq h'(x_t))\right]\leq\mathcal{O}\Big(T\cdot K\cdot\epsilon+\sqrt{T\cdot K\cdot\text{VCD}(C)}+T(1-\sigma)^K\Big).$$

Proof of claim: Define $g_h(x)=\mathbb{1}(h(x)\neq h'_h(x))$ where h'_h is the closest neighbor in H' to h. Then, we can write:

$$\mathbb{E}_{\mathcal{D}}\left[\sup_{h\in C}\inf_{h'\in H'}\sum_{t=1}^{T}\mathbb{1}(h(x_t)\neq h'(x_t))\right]=\mathbb{E}_{\mathcal{D}}\left[\sup_{g\in G}\sum_{t=1}^{T}g(x_t)\right].$$

A fact from homework 2 tells us VCDim(G) $\leq$ 2VCDim(C). So, this fact and the union bound allow us to write:

$$\mathbb{E}_{\mathcal{D}}\left[\sup_{g\in G}\sum_{t=1}^{T}g(x_t)\right]\leq\sup_{g\in G}\mathbb{E}\left[\sum_{t=1}^{T}g(x_t)\right]+\tilde{\mathcal{O}}\Big(\sqrt{T\cdot\text{VCD}(C)}\Big).$$

The next lemma is arguably the most magical property of smoothed adversaries. Up to now, we have built some intuition regarding instances where $\mathcal{D}$ is an i.i.d. process. Of course, to address smoothed adversaries, we need to go beyond the independence assumption. The following lemma states, in general, how adaptive smoothed adversaries over a time horizon of length T can be viewed as adversaries that generate approximately T/σ samples and then select a subset of them. More formally, we have:

Lemma 3.6 (Coupling between smoothed adaptive and i.i.d processes)**.** *Suppose $\mathcal{D}$ is a σ-smooth joint distribution with respect to μ. There is a coupling Π over $(x_t,z_t^j)_{1\leq t\leq T;1\leq j\leq K}$ for any T,K such that:*

1. *x_t's are distributed with respect to $\mathcal{D}_t(\cdot|x_1, \ldots, x_{t-1})$*
2. *$\{z_t^j\}_{t=1,\ldots,T;j=1,\ldots,K}$ i.i.d from μ*
3. *For each t, $\mathbb{P}_\Pi(x_t \notin \{z_t^j\}_{j=1,\ldots,K}) \leq (1-\sigma)^K$*

Proof of 3.3 (cont.): *Proof of claim (cont.):* Applying the coupling lemma, we see

$$\mathbb{E}_{\mathcal{D}}\left[\sup_{g\in G}\sum_{t=1}^{T} g(x_t)\right] \leq \underbrace{\mathbb{E}_{\Pi}\left[\sup_{g\in G}\sum_{t=1}^{T}\sum_{j=1}^{K} g(z_t^j)\right]}_{B} + T(1-\sigma)^K$$

and

$$B \leq \underbrace{\sup_{g\in G}\mathbb{E}_{\mu}\left[\sum_{t=1}^{T}\sum_{j=1}^{K} g(z_t^j)\right]}_{T\cdot K\cdot\epsilon} + \sqrt{T\cdot K\cdot \text{VCD}(C)}$$

We can see the supremum above is upper bounded by $TK\epsilon$ using the definition of $g(z_t^j)$ as an indicator function and the fact that h' is in the ϵ-cover.

$$\begin{aligned}
\mathbb{E}_{\mu}\left[\sum_{t=1}^{T}\sum_{j=1}^{K} g(z_t^j)\right] &= \sum_{t=1}^{T}\sum_{j=1}^{K}\mathbb{E}_{\mu}\left[\mathbb{1}(h(z_t^j) \neq h'_h(z_t^j))\right] \\
&= \sum_{t=1}^{T}\sum_{j=1}^{K}\mathbb{P}_{\mu}(h(z_t^j) \neq h'(z_t^j)) \\
&\leq TK\epsilon
\end{aligned}$$

Combining the above results, we see that

$$A_T \leq \mathbb{E}_{\mathcal{D}}\left[\sup_{g\in G}\sum_{t=1}^{T} g(x_t)\right] \leq T\cdot K\cdot\epsilon + \sqrt{T\cdot K\cdot \text{VCD}(C)} + T(1-\sigma)^K.$$

If $\epsilon = \frac{\sigma^2}{T^2}$ and $K = \frac{1}{\sigma}\log T$,

$$A_T \leq \underbrace{T\exp(-K\sigma)}_{\mathcal{O}(1)} + \underbrace{\tfrac{TK\epsilon}{\sigma}}_{\mathcal{O}(1)} + \sqrt{\tfrac{T}{\sigma}\ln(T)\text{VCD}(C)}.$$

Then

$$\begin{aligned}
\text{regret} &\leq \sqrt{T\log(H')} + \mathcal{O}\left(\sqrt{\tfrac{T}{\sigma}\ln(T)\text{VCD}(C)}\right) \\
&= \sqrt{\tfrac{T}{\sigma}\text{VCD}(C)}\cdot\left(\sqrt{\tfrac{\sigma\log H'}{\text{VCD}(C)}} + \mathcal{O}\left(\sqrt{\ln T}\right)\right) \\
&= \tilde{\mathcal{O}}\left(\sqrt{\tfrac{T}{\sigma}\text{VCD}(C)}\right). \quad \square
\end{aligned}$$

Remark 3.7. This proof sketch achieves a polynomial dependence on $\frac{1}{\sigma}$, compared to the stated dependence of $\log(1/\sigma)$. To obtain the tight bound stated in Theorem 3.3, the main idea is to use a stronger concentration inequality (Bernstein rather than Hoeffding) that leverages the fact that $\mathbb{E}[g(z_t^j)] \leq \epsilon$ to achieve a tighter bound on the approximation error.

References

Noga Alon, Omri Ben-Eliezer, Yuval Dagan, Shay Moran, Moni Naor, and Eylon Yogev. Adversarial laws of large numbers and optimal regret in online classification. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2021, page 447–455, New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450380539. doi: 10.1145/3406325.3451041. URL https://doi.org/10.1145/3406325.3451041.

Shai Ben-David, Dávid Pál, and Shai Shalev-Shwartz. Agnostic online learning. In *Annual Conference Computational Learning Theory*, 2009. URL https://api.semanticscholar.org/CorpusID:9403043.

Nika Haghtalab, Tim Roughgarden, and Abhishek Shetty. Smoothed analysis with adaptive adversaries. *J. ACM*, 71(3), June 2024. ISSN 0004-5411. doi: 10.1145/3656638. URL https://doi.org/10.1145/3656638.

David Haussler. Decision theoretic generalizations of the pac model for neural net and other learning applications. *Information and Computation*, 100(1):78–150, 1992. ISSN 0890-5401. doi: https://doi.org/10.1016/0890-5401(92)90010-D. URL https://www.sciencedirect.com/science/article/pii/089054019290010D.

N. Littlestone and M.K. Warmuth. The weighted majority algorithm. *Information and Computation*, 108(2):212–261, 1994. ISSN 0890-5401. doi: https://doi.org/10.1006/inco.1994.1009. URL https://www.sciencedirect.com/science/article/pii/S0890540184710091.

Nick Littlestone. Learning quickly when irrelevant attributes abound: A new linear-threshold algorithm. *Mach. Learn.*, 2(4):285–318, April 1988. ISSN 0885-6125. doi: 10.1023/A:1022869011914. URL https://doi.org/10.1023/A:1022869011914.

Leslie G. Valiant. A theory of the learnable. *Commun. ACM*, 27:1134–1142, 1984. URL https://api.semanticscholar.org/CorpusID:59712.