

# CSE 541: Interactive Learning

## Lecture 1: Logistics, generalization, and online learning

Kevin Jamieson

Paul G. Allen School of Computer Science & Engineering  
University of Washington

Wednesday, September 30, 2026



Everything in these slides is on the course website:

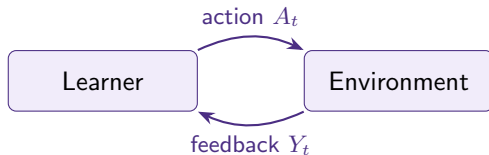
<https://courses.cs.washington.edu/courses/cse541/26au/>



# What is interactive learning?

Traditional ML: collect a labeled dataset, *then* learn.

Interactive learning: use the data collected so far to decide **which measurement to take next**, in a closed loop.



## Pure exploration

Identify a good-enough answer using as few measurements as possible. E.g., find  $\hat{i}$  with  $\mathbb{P}(\hat{i} \neq i^*) \leq \delta$ .

Classification, drug discovery, science.

## Regret minimization

Balance learning (explore) against earning (exploit):

$$R_T = T\mu^* - \mathbb{E}\left[\sum_{t=1}^T \mu_{A_t}\right].$$

Recommendation, treatment design, ad serving.

# Course roadmap

We classify problems by hypothesis space, available actions, measurement model, and side information, and focus on **general algorithmic strategies** and **common proof techniques**.

1. **(Non-)stochastic online learning**: exponential weights, online gradient descent
2. **(Non-)stochastic multi-armed bandits**: concentration, UCB, elimination, EXP3, lower bounds
3. **(Non-)stochastic linear bandits and experimental design**: optimal design, LinUCB, martingale confidence bounds
4. **(Non-)stochastic contextual bandits**: EXP4, off-policy evaluation, model-free and model-based methods

By the end, you should be ready to start research in this area, or to lead an interactive-learning implementation in industry.

|                             | Office hours     | Where     |
|-----------------------------|------------------|-----------|
| Kevin Jamieson (instructor) | Tue 3:00–4:00 PM | CSE2 340  |
| Junyan Liu (TA)             | Wed 2:00–3:00 PM | Zoom      |
| Rohan Baijal (TA)           | Fri 2:30–3:30 PM | Gates 152 |

Junyan's Zoom link is on Canvas.

**Lectures:** Mon & Wed, 10:00–11:20 AM, ECE 037

Slides are posted on the website before each class.

## Where things live

- **Website:** schedule, slides, handouts, policies
- **Ed:** questions and discussion (first stop!)
- **Gradescope:** homework and writeups
- **Email:** cse541-staff@cs, for private matters

# Grading and deliverables

| Component         | Weight   |
|-------------------|----------|
| Homeworks 1, 2, 3 | 10% each |
| Midterm project   | 35%      |
| Final project     | 35%      |

## Tentative due dates (11:59 PM)

|                 |            |
|-----------------|------------|
| HW1             | Tue Oct 20 |
| Midterm project | Sun Nov 8  |
| HW2             | Sun Nov 22 |
| HW3             | Sun Dec 13 |
| Final project   | Tue Dec 15 |

## Each project

- Pick a topic that extends a tool from the course
- **5-minute YouTube video** ( $\pm 30$  s, may be unlisted, captioned). Clarity and accuracy count; production quality doesn't.
- **2–3 page LaTeX writeup** on Gradescope
- Share both in an Ed post
- Individual work. **No late submissions.**

## Homework policies

- **Submission:** one typeset PDF per homework on Gradescope (LaTeX strongly recommended; no photos or scans). Code goes at the end of each problem.
- **Collaboration:** discuss freely with classmates, but write up your own answers and code, and *list who you worked with*. Copying code or  $\text{\LaTeX}$  from another student crosses the line, and so does using existing solutions from past offerings or the web.
- **Late policy:** an extra 24 hours on any homework, no questions asked. Need more? Email me *before* the deadline. (Projects have firm deadlines.)
- **Regrades:** through Gradescope within 7 days of grades being released. The whole assignment may be regraded.

## AI policy: AI for Collaboration

CSE 541 adopts the Allen School's “**AI for Collaboration**” policy, plus two course-specific rules on the next slide.

### You may

use AI on any course work, including homeworks and projects, unless a specific activity says otherwise: explanations, brainstorming, coding, debugging, reviewing your work.

### You must

- **Attribute:** name the tool and how you used it
- **Own it:** be able to explain and defend everything you submit
- **Protect data:** no restricted course materials or personal info; scope what coding agents can read

Please read the [full policy](#) on the course website carefully.

# AI policy: you supply the algorithm and the proof

## Not acceptable

“Implement UCB.”

The AI decides what the algorithm is.

## Acceptable

“At each round  $t$ , compute for each arm  $i$

$$\text{UCB}(i) = \hat{\mu}(i) + \sqrt{\frac{2 \log t}{N(i)}},$$

pull  $\arg \max_i \text{UCB}(i)$ , observe  $X_{i,t} = \mu(i) + \eta_t$  with  $\eta_t \sim \mathcal{N}(0, 1)$ , update  $\hat{\mu}$  and  $N$ , and repeat for  $T$  rounds. Write this loop in Python.”

You specified every step; the AI just translates it into code.

**Proofs:** use AI minimally, to clarify a concept, brainstorm approaches, or fix  $\text{\LaTeX}$  formatting. Never ask it to write, complete, or fix your proof.

## Before next class

**Prerequisites:** fluency in linear algebra, probability and statistics, and calculus. The course is analysis-heavy. Intro ML (CSE 446/546) helps but is not required.

### To do this week:

- Work through the [HW0 self-test](#) (not collected)
- Join Ed and Gradescope
- Read the course policies, including the AI policy

### Main references

[Lattimore–Szepesvári] *Bandit Algorithms*  
[Jamieson] course lecture notes

### Accommodations

Have an Access Plan with DRS? Activate it for CSE 541 early and let's talk.  
Otherwise, contact DRS to set one up.

# Supervised learning and generalization

What the IID assumption buys you

# The standard machine learning paradigm

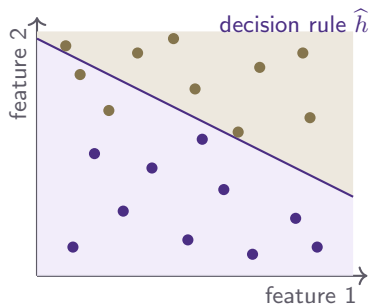
- **Data:** past observations, drawn IID from some fixed  $\nu$
- **Hypotheses:** a class  $\mathcal{H}$  devised to capture patterns in the data
- **Prediction:** fit  $h \in \mathcal{H}$  to the data, apply it to future draws from the *same*  $\nu$

# The standard machine learning paradigm

- **Data:** past observations, drawn IID from some fixed  $\nu$
- **Hypotheses:** a class  $\mathcal{H}$  devised to capture patterns in the data
- **Prediction:** fit  $h \in \mathcal{H}$  to the data, apply it to future draws from the *same*  $\nu$

## Do these fit the paradigm?

Movie recommendations • ad serving • spam filtering • clinical trials • self-driving cars • credit scoring



Learn  $\hat{h}$  from labeled points, apply to new points from the same distribution.

## Setup: classification with a finite hypothesis class

- Fix a **finite** class  $\mathcal{H} = \{h_1, h_2, \dots\}$  of classifiers  $h : \mathcal{X} \rightarrow \{-1, +1\}$ .
- Observe  $(x_1, y_1), \dots, (x_n, y_n) \stackrel{iid}{\sim} \nu$  with  $y_i \in \{-1, +1\}$ .
- For each  $h \in \mathcal{H}$  define the **true risk** and the **empirical risk**

$$R(h) = \mathbb{P}_{(X,Y) \sim \nu}(h(X) \neq Y), \quad \hat{R}_n(h) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}\{h(x_i) \neq y_i\},$$

so that  $\mathbb{E}[\hat{R}_n(h)] = R(h)$  for each fixed  $h$ .

- **Empirical risk minimization (ERM):**  $\hat{h} \in \arg \min_{h \in \mathcal{H}} \hat{R}_n(h)$ .

## Setup: classification with a finite hypothesis class

- Fix a **finite** class  $\mathcal{H} = \{h_1, h_2, \dots\}$  of classifiers  $h : \mathcal{X} \rightarrow \{-1, +1\}$ .
- Observe  $(x_1, y_1), \dots, (x_n, y_n) \stackrel{iid}{\sim} \nu$  with  $y_i \in \{-1, +1\}$ .
- For each  $h \in \mathcal{H}$  define the **true risk** and the **empirical risk**

$$R(h) = \mathbb{P}_{(X,Y) \sim \nu}(h(X) \neq Y), \quad \hat{R}_n(h) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}\{h(x_i) \neq y_i\},$$

so that  $\mathbb{E}[\hat{R}_n(h)] = R(h)$  for each fixed  $h$ .

- **Empirical risk minimization (ERM):**  $\hat{h} \in \arg \min_{h \in \mathcal{H}} \hat{R}_n(h)$ .

### Question

How “good” is  $\hat{h}$ ? That is, how large can  $R(\hat{h})$  be, and how does it depend on  $n$  and  $|\mathcal{H}|$ ?

Why finite  $\mathcal{H}$ ? Clean proofs, all the ideas. Infinite classes need VC dimension, covering numbers, etc.

### Theorem (realizable ERM)

Let  $|\mathcal{H}| < \infty$  and  $(x_i, y_i) \stackrel{iid}{\sim} \nu$ . Assume there exists  $h_* \in \mathcal{H}$  with  $R(h_*) = 0$ . If  $\hat{h} \in \arg \min_{h \in \mathcal{H}} \hat{R}_n(h)$ , then with probability at least  $1 - \delta$ ,

$$R(\hat{h}) \leq \frac{\log(|\mathcal{H}|/\delta)}{n}.$$

### Theorem (realizable ERM)

Let  $|\mathcal{H}| < \infty$  and  $(x_i, y_i) \stackrel{iid}{\sim} \nu$ . Assume there exists  $h_* \in \mathcal{H}$  with  $R(h_*) = 0$ . If  $\hat{h} \in \arg \min_{h \in \mathcal{H}} \hat{R}_n(h)$ , then with probability at least  $1 - \delta$ ,

$$R(\hat{h}) \leq \frac{\log(|\mathcal{H}|/\delta)}{n}.$$

Equivalently,  $\mathbb{P}(R(\hat{h}) > \epsilon) \leq |\mathcal{H}|e^{-n\epsilon}$ : with  $\epsilon$  = “approximately” and  $\delta$  = “probably”, this is a **Probably Approximately Correct** (PAC) guarantee.

### Theorem (realizable ERM)

Let  $|\mathcal{H}| < \infty$  and  $(x_i, y_i) \stackrel{iid}{\sim} \nu$ . Assume there exists  $h_* \in \mathcal{H}$  with  $R(h_*) = 0$ . If  $\hat{h} \in \arg \min_{h \in \mathcal{H}} \hat{R}_n(h)$ , then with probability at least  $1 - \delta$ ,

$$R(\hat{h}) \leq \frac{\log(|\mathcal{H}|/\delta)}{n}.$$

Equivalently,  $\mathbb{P}(R(\hat{h}) > \epsilon) \leq |\mathcal{H}|e^{-n\epsilon}$ : with  $\epsilon =$  “approximately” and  $\delta =$  “probably”, this is a **Probably Approximately Correct** (PAC) guarantee.

**In expectation.** Integrate the tail, splitting at  $\epsilon_0 = \log(|\mathcal{H}|)/n$ :

$$\mathbb{E}[R(\hat{h})] = \int_0^1 \mathbb{P}(R(\hat{h}) > \epsilon) d\epsilon \leq \epsilon_0 + \int_{\epsilon_0}^{\infty} |\mathcal{H}|e^{-n\epsilon} d\epsilon = \frac{1 + \log |\mathcal{H}|}{n}.$$

### Theorem (realizable ERM)

Let  $|\mathcal{H}| < \infty$  and  $(x_i, y_i) \stackrel{iid}{\sim} \nu$ . Assume there exists  $h_* \in \mathcal{H}$  with  $R(h_*) = 0$ . If  $\hat{h} \in \arg \min_{h \in \mathcal{H}} \hat{R}_n(h)$ , then with probability at least  $1 - \delta$ ,

$$R(\hat{h}) \leq \frac{\log(|\mathcal{H}|/\delta)}{n}.$$

Equivalently,  $\mathbb{P}(R(\hat{h}) > \epsilon) \leq |\mathcal{H}|e^{-n\epsilon}$ : with  $\epsilon =$  “approximately” and  $\delta =$  “probably”, this is a **Probably Approximately Correct** (PAC) guarantee.

**In expectation.** Integrate the tail, splitting at  $\epsilon_0 = \log(|\mathcal{H}|)/n$ :

$$\mathbb{E}[R(\hat{h})] = \int_0^1 \mathbb{P}(R(\hat{h}) > \epsilon) d\epsilon \leq \epsilon_0 + \int_{\epsilon_0}^{\infty} |\mathcal{H}|e^{-n\epsilon} d\epsilon = \frac{1 + \log |\mathcal{H}|}{n}.$$

**Takeaway:** risk decays like  $1/n$ , and the class size enters only through  $\log |\mathcal{H}|$ ; a million hypotheses cost about 14 extra units of  $n\epsilon$ .

## Realizable case: proof

**Two ingredients.** (i) **Union bound:**  $\mathbb{P}(A \cup B) \leq \mathbb{P}(A) + \mathbb{P}(B)$ . (ii)  $1 - x \leq e^{-x}$  for all  $x$ .

Since  $\hat{R}_n(h_*) = 0$ , ERM returns some  $\hat{h}$  with  $\hat{R}_n(\hat{h}) = 0$ . So a “bad” event needs a hypothesis that is bad *and* looks perfect on the data:

$$\mathbb{P}(R(\hat{h}) > \epsilon) =$$

## Realizable case: proof

**Two ingredients.** (i) **Union bound:**  $\mathbb{P}(A \cup B) \leq \mathbb{P}(A) + \mathbb{P}(B)$ . (ii)  $1 - x \leq e^{-x}$  for all  $x$ .

Since  $\hat{R}_n(h_*) = 0$ , ERM returns some  $\hat{h}$  with  $\hat{R}_n(\hat{h}) = 0$ . So a “bad” event needs a hypothesis that is bad *and* looks perfect on the data:

$$\begin{aligned}\mathbb{P}(R(\hat{h}) > \epsilon) &= \mathbb{P}(R(\hat{h}) > \epsilon \text{ and } \hat{R}_n(\hat{h}) = 0) \leq \mathbb{P}\left(\bigcup_{h \in \mathcal{H}} \{R(h) > \epsilon \text{ and } \hat{R}_n(h) = 0\}\right) \quad (\hat{h} \in \mathcal{H}) \\ &\leq \sum_{h \in \mathcal{H}: R(h) > \epsilon} \mathbb{P}(\hat{R}_n(h) = 0) \quad \text{(union bound)} \\ &= \sum_{h \in \mathcal{H}: R(h) > \epsilon} \prod_{i=1}^n \mathbb{P}(h(x_i) = y_i) \leq |\mathcal{H}|(1 - \epsilon)^n \leq |\mathcal{H}|e^{-n\epsilon}. \quad \text{(IID)}\end{aligned}$$

Set  $|\mathcal{H}|e^{-n\epsilon} = \delta$  and solve for  $\epsilon$ . ■

## Realizable case: proof

**Two ingredients.** (i) **Union bound:**  $\mathbb{P}(A \cup B) \leq \mathbb{P}(A) + \mathbb{P}(B)$ . (ii)  $1 - x \leq e^{-x}$  for all  $x$ .

Since  $\hat{R}_n(h_*) = 0$ , ERM returns some  $\hat{h}$  with  $\hat{R}_n(\hat{h}) = 0$ . So a “bad” event needs a hypothesis that is bad *and* looks perfect on the data:

$$\begin{aligned}\mathbb{P}(R(\hat{h}) > \epsilon) &= \mathbb{P}(R(\hat{h}) > \epsilon \text{ and } \hat{R}_n(\hat{h}) = 0) \leq \mathbb{P}\left(\bigcup_{h \in \mathcal{H}} \{R(h) > \epsilon \text{ and } \hat{R}_n(h) = 0\}\right) \quad (\hat{h} \in \mathcal{H}) \\ &\leq \sum_{h \in \mathcal{H}: R(h) > \epsilon} \mathbb{P}(\hat{R}_n(h) = 0) \quad \text{(union bound)} \\ &= \sum_{h \in \mathcal{H}: R(h) > \epsilon} \prod_{i=1}^n \mathbb{P}(h(x_i) = y_i) \leq |\mathcal{H}|(1 - \epsilon)^n \leq |\mathcal{H}|e^{-n\epsilon}. \quad \text{(IID)}\end{aligned}$$

Set  $|\mathcal{H}|e^{-n\epsilon} = \delta$  and solve for  $\epsilon$ . ■

### Where did we use IID?

Exactly once, for a *fixed*  $h$ :  $\mathbb{P}(\hat{R}_n(h) = 0) = \prod_i \mathbb{P}(h(x_i) = y_i)$ . The union bound handles the fact that  $\hat{h}$  depends on the data.

## Agnostic (non-realizable) case: the tool

Usually no  $h \in \mathcal{H}$  is perfect. Compare instead to the best in class,  $h_* \in \arg \min_{h \in \mathcal{H}} R(h)$ . We now need  $\hat{R}_n(h) \approx R(h)$ , i.e., an average of bounded IID variables concentrates around its mean.

### Lemma (Hoeffding's inequality)

Let  $Z_1, \dots, Z_n$  be independent with  $Z_i \in [a, b]$  a.s. and  $\mathbb{E}[Z_i] = \mu$ . For any  $\epsilon > 0$ ,

$$\mathbb{P}\left(\frac{1}{n} \sum_{i=1}^n Z_i \geq \mu + \epsilon\right) \leq \exp\left(\frac{-2n\epsilon^2}{(b-a)^2}\right).$$

## Agnostic (non-realizable) case: the tool

Usually no  $h \in \mathcal{H}$  is perfect. Compare instead to the best in class,  $h_* \in \arg \min_{h \in \mathcal{H}} R(h)$ . We now need  $\hat{R}_n(h) \approx R(h)$ , i.e., an average of bounded IID variables concentrates around its mean.

### Lemma (Hoeffding's inequality)

Let  $Z_1, \dots, Z_n$  be independent with  $Z_i \in [a, b]$  a.s. and  $\mathbb{E}[Z_i] = \mu$ . For any  $\epsilon > 0$ ,

$$\mathbb{P}\left(\frac{1}{n} \sum_{i=1}^n Z_i \geq \mu + \epsilon\right) \leq \exp\left(\frac{-2n\epsilon^2}{(b-a)^2}\right).$$

**Sanity check via the CLT.**  $\frac{1}{n} \sum_i Z_i - \mu \approx \mathcal{N}(0, \sigma^2/n)$  with  $\sigma^2 \leq (b-a)^2/4$ , so fluctuations are of size  $(b-a)/(2\sqrt{n})$  and Gaussian tails look like  $e^{-n\epsilon^2/(2\sigma^2)}$ . Hoeffding says the same thing, **non-asymptotically** and for every  $n$ .

**For us:**  $Z_i = \mathbf{1}\{h(x_i) \neq y_i\} \in [0, 1]$  with mean  $R(h)$ , so for a fixed  $h$ ,  $\mathbb{P}(|\hat{R}_n(h) - R(h)| \geq \epsilon) \leq 2e^{-2n\epsilon^2}$ .

# Proof of Hoeffding: the Chernoff method

## Hoeffding's lemma

If  $X \in [a, b]$  a.s. and  $\mathbb{E}[X] = 0$ , then  $\log \mathbb{E}[e^{\lambda X}] \leq \lambda^2(b-a)^2/8$  for all  $\lambda \in \mathbb{R}$ .

(Proof idea:  $\psi(\lambda) = \log \mathbb{E}[e^{\lambda X}]$  has  $\psi(0) = \psi'(0) = 0$  and  $\psi''(\lambda)$  is the variance of a  $[a, b]$ -valued tilted variable, hence  $\leq (b-a)^2/4$ . Taylor.)

# Proof of Hoeffding: the Chernoff method

## Hoeffding's lemma

If  $X \in [a, b]$  a.s. and  $\mathbb{E}[X] = 0$ , then  $\log \mathbb{E}[e^{\lambda X}] \leq \lambda^2(b-a)^2/8$  for all  $\lambda \in \mathbb{R}$ .

(Proof idea:  $\psi(\lambda) = \log \mathbb{E}[e^{\lambda X}]$  has  $\psi(0) = \psi'(0) = 0$  and  $\psi''(\lambda)$  is the variance of a  $[a, b]$ -valued tilted variable, hence  $\leq (b-a)^2/4$ . Taylor.)

**Chernoff method.** For any  $\lambda > 0$ , exponentiate, apply Markov's inequality, then use independence:

$$\mathbb{P}\left(\sum_{i=1}^n (Z_i - \mu) \geq n\epsilon\right) =$$

# Proof of Hoeffding: the Chernoff method

## Hoeffding's lemma

If  $X \in [a, b]$  a.s. and  $\mathbb{E}[X] = 0$ , then  $\log \mathbb{E}[e^{\lambda X}] \leq \lambda^2(b-a)^2/8$  for all  $\lambda \in \mathbb{R}$ .

(Proof idea:  $\psi(\lambda) = \log \mathbb{E}[e^{\lambda X}]$  has  $\psi(0) = \psi'(0) = 0$  and  $\psi''(\lambda)$  is the variance of a  $[a, b]$ -valued tilted variable, hence  $\leq (b-a)^2/4$ . Taylor.)

**Chernoff method.** For any  $\lambda > 0$ , exponentiate, apply Markov's inequality, then use independence:

$$\begin{aligned} \mathbb{P}\left(\sum_{i=1}^n (Z_i - \mu) \geq n\epsilon\right) &= \mathbb{P}\left(e^{\lambda \sum_i (Z_i - \mu)} \geq e^{\lambda n\epsilon}\right) \leq e^{-\lambda n\epsilon} \mathbb{E}\left[e^{\lambda \sum_i (Z_i - \mu)}\right] = e^{-\lambda n\epsilon} \prod_{i=1}^n \mathbb{E}\left[e^{\lambda (Z_i - \mu)}\right] \\ &\leq \exp\left(-\lambda n\epsilon + \frac{n\lambda^2(b-a)^2}{8}\right) \stackrel{\lambda=4\epsilon/(b-a)^2}{=} \exp\left(\frac{-2n\epsilon^2}{(b-a)^2}\right). \quad \blacksquare \end{aligned}$$

# Proof of Hoeffding: the Chernoff method

## Hoeffding's lemma

If  $X \in [a, b]$  a.s. and  $\mathbb{E}[X] = 0$ , then  $\log \mathbb{E}[e^{\lambda X}] \leq \lambda^2(b-a)^2/8$  for all  $\lambda \in \mathbb{R}$ .

(Proof idea:  $\psi(\lambda) = \log \mathbb{E}[e^{\lambda X}]$  has  $\psi(0) = \psi'(0) = 0$  and  $\psi''(\lambda)$  is the variance of a  $[a, b]$ -valued tilted variable, hence  $\leq (b-a)^2/4$ . Taylor.)

**Chernoff method.** For any  $\lambda > 0$ , exponentiate, apply Markov's inequality, then use independence:

$$\begin{aligned} \mathbb{P}\left(\sum_{i=1}^n (Z_i - \mu) \geq n\epsilon\right) &= \mathbb{P}\left(e^{\lambda \sum_i (Z_i - \mu)} \geq e^{\lambda n\epsilon}\right) \leq e^{-\lambda n\epsilon} \mathbb{E}\left[e^{\lambda \sum_i (Z_i - \mu)}\right] = e^{-\lambda n\epsilon} \prod_{i=1}^n \mathbb{E}\left[e^{\lambda (Z_i - \mu)}\right] \\ &\leq \exp\left(-\lambda n\epsilon + \frac{n\lambda^2(b-a)^2}{8}\right) \stackrel{\lambda=4\epsilon/(b-a)^2}{=} \exp\left(\frac{-2n\epsilon^2}{(b-a)^2}\right). \quad \blacksquare \end{aligned}$$

## Remember the shape of this argument

Bound a tail by bounding a moment generating function, then optimize  $\lambda$ . It reappears with martingales, sub-Gaussian noise, and confidence bounds throughout the course.

## Agnostic (non-realizable) case

### Theorem (agnostic ERM)

Let  $|\mathcal{H}| < \infty$ ,  $(x_i, y_i) \stackrel{iid}{\sim} \nu$ ,  $h_* \in \arg \min_{h \in \mathcal{H}} R(h)$  and  $\hat{h} \in \arg \min_{h \in \mathcal{H}} \hat{R}_n(h)$ . With probability at least  $1 - \delta$ ,

$$R(\hat{h}) - R(h_*) \leq \sqrt{\frac{2 \log(2|\mathcal{H}|/\delta)}{n}}.$$

### Theorem (agnostic ERM)

Let  $|\mathcal{H}| < \infty$ ,  $(x_i, y_i) \stackrel{iid}{\sim} \nu$ ,  $h_* \in \arg \min_{h \in \mathcal{H}} R(h)$  and  $\hat{h} \in \arg \min_{h \in \mathcal{H}} \hat{R}_n(h)$ . With probability at least  $1 - \delta$ ,

$$R(\hat{h}) - R(h_*) \leq \sqrt{\frac{2 \log(2|\mathcal{H}|/\delta)}{n}}.$$

**Proof.** Apply Hoeffding to  $Z_i = \mathbf{1}\{h(x_i) \neq y_i\} \in [0, 1]$  for a fixed  $h$ , to both tails, then union bound over  $h \in \mathcal{H}$ : with probability  $\geq 1 - \delta$ , simultaneously for all  $h \in \mathcal{H}$ ,

$$|\hat{R}_n(h) - R(h)| \leq \epsilon_n := \sqrt{\frac{\log(2|\mathcal{H}|/\delta)}{2n}}.$$

On this event, since  $\hat{h}$  minimizes  $\hat{R}_n$ ,

$$R(\hat{h}) \leq$$

## Theorem (agnostic ERM)

Let  $|\mathcal{H}| < \infty$ ,  $(x_i, y_i) \stackrel{iid}{\sim} \nu$ ,  $h_* \in \arg \min_{h \in \mathcal{H}} R(h)$  and  $\hat{h} \in \arg \min_{h \in \mathcal{H}} \hat{R}_n(h)$ . With probability at least  $1 - \delta$ ,

$$R(\hat{h}) - R(h_*) \leq \sqrt{\frac{2 \log(2|\mathcal{H}|/\delta)}{n}}.$$

**Proof.** Apply Hoeffding to  $Z_i = \mathbf{1}\{h(x_i) \neq y_i\} \in [0, 1]$  for a fixed  $h$ , to both tails, then union bound over  $h \in \mathcal{H}$ : with probability  $\geq 1 - \delta$ , simultaneously for all  $h \in \mathcal{H}$ ,

$$|\hat{R}_n(h) - R(h)| \leq \epsilon_n := \sqrt{\frac{\log(2|\mathcal{H}|/\delta)}{2n}}.$$

On this event, since  $\hat{h}$  minimizes  $\hat{R}_n$ ,

$$R(\hat{h}) \leq \hat{R}_n(\hat{h}) + \epsilon_n \leq \hat{R}_n(h_*) + \epsilon_n \leq R(h_*) + 2\epsilon_n. \quad \blacksquare$$

## Agnostic (non-realizable) case

### Theorem (agnostic ERM)

Let  $|\mathcal{H}| < \infty$ ,  $(x_i, y_i) \stackrel{iid}{\sim} \nu$ ,  $h_* \in \arg \min_{h \in \mathcal{H}} R(h)$  and  $\hat{h} \in \arg \min_{h \in \mathcal{H}} \hat{R}_n(h)$ . With probability at least  $1 - \delta$ ,

$$R(\hat{h}) - R(h_*) \leq \sqrt{\frac{2 \log(2|\mathcal{H}|/\delta)}{n}}.$$

**Proof.** Apply Hoeffding to  $Z_i = \mathbf{1}\{h(x_i) \neq y_i\} \in [0, 1]$  for a fixed  $h$ , to both tails, then union bound over  $h \in \mathcal{H}$ : with probability  $\geq 1 - \delta$ , simultaneously for all  $h \in \mathcal{H}$ ,

$$|\hat{R}_n(h) - R(h)| \leq \epsilon_n := \sqrt{\frac{\log(2|\mathcal{H}|/\delta)}{2n}}.$$

On this event, since  $\hat{h}$  minimizes  $\hat{R}_n$ ,

$$R(\hat{h}) \leq \hat{R}_n(\hat{h}) + \epsilon_n \leq \hat{R}_n(h_*) + \epsilon_n \leq R(h_*) + 2\epsilon_n. \quad \blacksquare$$

**In expectation** (integrate the tail as before):  $\mathbb{E}[R(\hat{h})] - R(h_*) \leq \sqrt{8 \log(2|\mathcal{H}|)/n}$ .

# What we learned from the IID world

|            | Excess risk of ERM                               | Tool              |
|------------|--------------------------------------------------|-------------------|
| Realizable | $\frac{\log( \mathcal{H} /\delta)}{n}$           | union bound       |
| Agnostic   | $\sqrt{\frac{2 \log(2 \mathcal{H} /\delta)}{n}}$ | Hoeffding + union |

- **Fast rate**  $1/n$  vs. **slow rate**  $1/\sqrt{n}$ ; Bernstein's inequality interpolates:  
 $\sqrt{R(h_*) \log |\mathcal{H}|/n} + \log |\mathcal{H}|/n.$
- Infinite  $\mathcal{H}$  (e.g., all hyperplanes): replace  $\log |\mathcal{H}|$  by VC dimension, Rademacher complexity, covering numbers.

## Everything relied on IID

Concentration of  $\widehat{R}_n(h)$  around  $R(h)$ , the very definition of “risk,” and the promise that the future looks like the past.

# What we learned from the IID world

|            | Excess risk of ERM                               | Tool              |
|------------|--------------------------------------------------|-------------------|
| Realizable | $\frac{\log( \mathcal{H} /\delta)}{n}$           | union bound       |
| Agnostic   | $\sqrt{\frac{2 \log(2 \mathcal{H} /\delta)}{n}}$ | Hoeffding + union |

- **Fast rate**  $1/n$  vs. **slow rate**  $1/\sqrt{n}$ ; Bernstein's inequality interpolates:  
 $\sqrt{R(h_*) \log |\mathcal{H}|/n} + \log |\mathcal{H}|/n.$
- Infinite  $\mathcal{H}$  (e.g., all hyperplanes): replace  $\log |\mathcal{H}|$  by VC dimension, Rademacher complexity, covering numbers.

## Everything relied on IID

Concentration of  $\widehat{R}_n(h)$  around  $R(h)$ , the very definition of “risk,” and the promise that the future looks like the past.

## The question for the rest of today

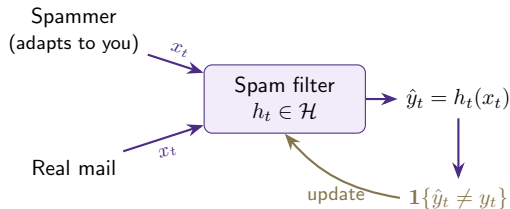
Spammers adapt to the filter.  
Markets react to trades. Opponents play to win.

Can we say *anything* when  $(x_t, y_t)$  are chosen **adversarially**?

# Online learning

Dropping the IID assumption

# Online learning: the spam filter



## Goal

Minimize the number of mistakes

$$\sum_{t=1}^T \mathbf{1}\{h_t(x_t) \neq y_t\}.$$

## Online classification

**Input:**  $\mathcal{H}$  with  $|\mathcal{H}| < \infty$

**for**  $t = 1, 2, \dots, T$ :

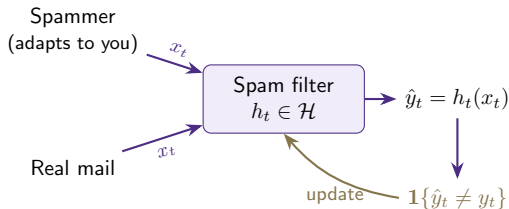
$x_t$  arrives

Player picks  $h_t \in \mathcal{H}$  (possibly at random)

$y_t \in \{-1, +1\}$  is revealed

Player suffers loss  $\mathbf{1}\{h_t(x_t) \neq y_t\}$

# Online learning: the spam filter



## Online classification

**Input:**  $\mathcal{H}$  with  $|\mathcal{H}| < \infty$

**for**  $t = 1, 2, \dots, T$ :

$x_t$  arrives

Player picks  $h_t \in \mathcal{H}$  (possibly at random)

$y_t \in \{-1, +1\}$  is revealed

Player suffers loss  $\mathbf{1}\{h_t(x_t) \neq y_t\}$

## Goal

Minimize the number of mistakes

$$\sum_{t=1}^T \mathbf{1}\{h_t(x_t) \neq y_t\}.$$

**Two settings:** **IID:**  $(x_t, y_t) \stackrel{iid}{\sim} \nu$ .

**Adversarial:**  $(x_t, y_t)$  arbitrary; the adversary may know your algorithm, but not your future coin flips. Strictly harder, so adversarial guarantees are also IID guarantees.

# Realizable case, IID: we already know what to do

## Online classification

**Input:**  $\mathcal{H}$  with  $|\mathcal{H}| < \infty$

**for**  $t = 1, 2, \dots, T$ :

$x_t$  arrives

Player picks  $h_t \in \mathcal{H}$  (possibly at random)

$y_t \in \{-1, +1\}$  is revealed

Player suffers loss  $\mathbf{1}\{h_t(x_t) \neq y_t\}$

**Assume:**  $y_t = h_*(x_t)$  for some  $h_* \in \mathcal{H}$ , and  $(x_t, y_t) \stackrel{iid}{\sim} \nu$ .

**Strategy:** run ERM on everything seen so far,

$$h_t \in \arg \min_{h \in \mathcal{H}} \sum_{s=1}^{t-1} \mathbf{1}\{h(x_s) \neq y_s\}.$$

# Realizable case, IID: we already know what to do

## Online classification

**Input:**  $\mathcal{H}$  with  $|\mathcal{H}| < \infty$

**for**  $t = 1, 2, \dots, T$ :

$x_t$  arrives

Player picks  $h_t \in \mathcal{H}$  (possibly at random)

$y_t \in \{-1, +1\}$  is revealed

Player suffers loss  $\mathbf{1}\{h_t(x_t) \neq y_t\}$

**Assume:**  $y_t = h_*(x_t)$  for some  $h_* \in \mathcal{H}$ , and  $(x_t, y_t) \stackrel{iid}{\sim} \nu$ .

**Strategy:** run ERM on everything seen so far,

$$h_t \in \arg \min_{h \in \mathcal{H}} \sum_{s=1}^{t-1} \mathbf{1}\{h(x_s) \neq y_s\}.$$

$h_t$  depends only on rounds  $< t$ , and  $(x_t, y_t)$  is a fresh draw, so

$\mathbb{E}[\mathbf{1}\{h_t(x_t) \neq y_t\}] = \mathbb{E}[R(h_t)]$  and the realizable bound in expectation gives

$$\mathbb{E}\left[\sum_{t=1}^T \mathbf{1}\{h_t(x_t) \neq y_t\}\right] \leq 1 + \sum_{t=2}^T \mathbb{E}[R(h_t)] \leq 1 + \sum_{t=2}^T \frac{1 + \log |\mathcal{H}|}{t-1} \leq 1 + (1 + \log |\mathcal{H}|)(1 + \log T).$$

The number of mistakes grows only **logarithmically** in  $T$ .

# Realizable case, adversarial: ERM breaks

## Online classification

**Input:**  $\mathcal{H}$  with  $|\mathcal{H}| < \infty$

**for**  $t = 1, 2, \dots, T$ :

$x_t$  arrives

Player picks  $h_t \in \mathcal{H}$  (possibly at random)

$y_t \in \{-1, +1\}$  is revealed

Player suffers loss  $\mathbf{1}\{h_t(x_t) \neq y_t\}$

**Assume:**  $y_t = h_*(x_t)$  for some  $h_* \in \mathcal{H}$ , but  $(x_t, y_t)$  are chosen by an adversary.

**Same strategy?**

$$h_t \in \arg \min_{h \in \mathcal{H}} \sum_{s < t} \mathbf{1}\{h(x_s) \neq y_s\}.$$

# Realizable case, adversarial: ERM breaks

## Online classification

**Input:**  $\mathcal{H}$  with  $|\mathcal{H}| < \infty$

**for**  $t = 1, 2, \dots, T$ :

$x_t$  arrives

Player picks  $h_t \in \mathcal{H}$  (possibly at random)

$y_t \in \{-1, +1\}$  is revealed

Player suffers loss  $\mathbf{1}\{h_t(x_t) \neq y_t\}$

**Assume:**  $y_t = h_*(x_t)$  for some  $h_* \in \mathcal{H}$ , but  $(x_t, y_t)$  are chosen by an adversary.

**Same strategy?**

$$h_t \in \arg \min_{h \in \mathcal{H}} \sum_{s < t} \mathbf{1}\{h(x_s) \neq y_s\}.$$

## Claim

There is a sequence  $\{(x_t, y_t)\}$  consistent with some  $h_* \in \mathcal{H}$  on which this strategy makes  $\min\{|\mathcal{H}| - 1, T\}$  mistakes.

# Realizable case, adversarial: ERM breaks

## Online classification

**Input:**  $\mathcal{H}$  with  $|\mathcal{H}| < \infty$

**for**  $t = 1, 2, \dots, T$ :

$x_t$  arrives

Player picks  $h_t \in \mathcal{H}$  (possibly at random)

$y_t \in \{-1, +1\}$  is revealed

Player suffers loss  $\mathbf{1}\{h_t(x_t) \neq y_t\}$

**Assume:**  $y_t = h_*(x_t)$  for some  $h_* \in \mathcal{H}$ , but  $(x_t, y_t)$  are chosen by an adversary.

**Same strategy?**

$$h_t \in \arg \min_{h \in \mathcal{H}} \sum_{s < t} \mathbf{1}\{h(x_s) \neq y_s\}.$$

## Claim

There is a sequence  $\{(x_t, y_t)\}$  consistent with some  $h_* \in \mathcal{H}$  on which this strategy makes  $\min\{|\mathcal{H}| - 1, T\}$  mistakes.

**Why.** Many hypotheses are consistent with the past. The adversary picks an  $x_t$  on which they disagree, waits for your  $h_t$  (it knows your tie-breaking rule), and sets  $y_t = -h_t(x_t)$ . You eliminate one hypothesis per round.

# Realizable case, adversarial: ERM breaks

## Online classification

**Input:**  $\mathcal{H}$  with  $|\mathcal{H}| < \infty$

**for**  $t = 1, 2, \dots, T$ :

$x_t$  arrives

Player picks  $h_t \in \mathcal{H}$  (possibly at random)

$y_t \in \{-1, +1\}$  is revealed

Player suffers loss  $\mathbf{1}\{h_t(x_t) \neq y_t\}$

**Assume:**  $y_t = h_*(x_t)$  for some  $h_* \in \mathcal{H}$ , but  $(x_t, y_t)$  are chosen by an adversary.

**Same strategy?**

$$h_t \in \arg \min_{h \in \mathcal{H}} \sum_{s < t} \mathbf{1}\{h(x_s) \neq y_s\}.$$

## Claim

There is a sequence  $\{(x_t, y_t)\}$  consistent with some  $h_* \in \mathcal{H}$  on which this strategy makes  $\min\{|\mathcal{H}| - 1, T\}$  mistakes.

**Why.** Many hypotheses are consistent with the past. The adversary picks an  $x_t$  on which they disagree, waits for your  $h_t$  (it knows your tie-breaking rule), and sets  $y_t = -h_t(x_t)$ . You eliminate one hypothesis per round.

**Fix:** eliminate *many* hypotheses per mistake, not one.

# The Halving algorithm

## Halving algorithm

**Input:**  $\mathcal{H}$  with  $|\mathcal{H}| < \infty$

**Initialize:** version space  $V_1 = \mathcal{H}$

**for**  $t = 1, 2, \dots$ :

$x_t$  arrives

    Predict the majority vote

$\hat{y}_t = \text{sign}\left(\sum_{h \in V_t} h(x_t)\right)$

$y_t$  is revealed

$V_{t+1} = \{h \in V_t : h(x_t) = y_t\}$

## Theorem (Halving)

If  $y_t = h_*(x_t)$  for all  $t$  for some  $h_* \in \mathcal{H}$ , then for *any* sequence  $x_1, x_2, \dots$  the Halving algorithm makes at most  $\lfloor \log_2 |\mathcal{H}| \rfloor$  mistakes.

# The Halving algorithm

## Halving algorithm

**Input:**  $\mathcal{H}$  with  $|\mathcal{H}| < \infty$

**Initialize:** version space  $V_1 = \mathcal{H}$

**for**  $t = 1, 2, \dots$ :

$x_t$  arrives

    Predict the majority vote

$\hat{y}_t = \text{sign}\left(\sum_{h \in V_t} h(x_t)\right)$

$y_t$  is revealed

$V_{t+1} = \{h \in V_t : h(x_t) = y_t\}$

## Theorem (Halving)

If  $y_t = h_*(x_t)$  for all  $t$  for some  $h_* \in \mathcal{H}$ , then for *any* sequence  $x_1, x_2, \dots$  the Halving algorithm makes at most  $\lfloor \log_2 |\mathcal{H}| \rfloor$  mistakes.

**Proof.**  $h_* \in V_t$  for all  $t$ , so  $|V_t| \geq 1$ . On a mistake, more than half of  $V_t$  predicted  $\hat{y}_t \neq y_t$  and is removed, so  $|V_{t+1}| \leq |V_t|/2$ . After  $m$  mistakes,  $1 \leq |V_t| \leq |\mathcal{H}|/2^m$ . ■

# The Halving algorithm

## Halving algorithm

**Input:**  $\mathcal{H}$  with  $|\mathcal{H}| < \infty$

**Initialize:** version space  $V_1 = \mathcal{H}$

**for**  $t = 1, 2, \dots$ :

$x_t$  arrives

Predict the majority vote

$\hat{y}_t = \text{sign}\left(\sum_{h \in V_t} h(x_t)\right)$

$y_t$  is revealed

$V_{t+1} = \{h \in V_t : h(x_t) = y_t\}$

## Theorem (Halving)

If  $y_t = h_*(x_t)$  for all  $t$  for some  $h_* \in \mathcal{H}$ , then for *any* sequence  $x_1, x_2, \dots$  the Halving algorithm makes at most  $\lfloor \log_2 |\mathcal{H}| \rfloor$  mistakes.

**Proof.**  $h_* \in V_t$  for all  $t$ , so  $|V_t| \geq 1$ . On a mistake, more than half of  $V_t$  predicted  $\hat{y}_t \neq y_t$  and is removed, so  $|V_{t+1}| \leq |V_t|/2$ . After  $m$  mistakes,  $1 \leq |V_t| \leq |\mathcal{H}|/2^m$ . ■

## Compare

IID + ERM:  $O(\log |\mathcal{H}| \log T)$  expected mistakes. Adversarial + Halving:  $\log_2 |\mathcal{H}|$ , deterministically, for any  $T$ .

We dropped the IID assumption and lost nothing. (Halving is also minimax optimal.)

# The non-separable case

Regret and exponential weights

## When no hypothesis is perfect: regret

Without realizability, an adversary can force a mistake every round (set  $y_t$  against your prediction), so “number of mistakes” is meaningless. Ask instead to be competitive with the best fixed hypothesis in hindsight.

**Goal: minimize regret**

$$\text{Regret}_T = \mathbb{E} \left[ \sum_{t=1}^T \mathbf{1}\{h_t(x_t) \neq y_t\} \right] - \min_{h \in \mathcal{H}} \sum_{t=1}^T \mathbf{1}\{h(x_t) \neq y_t\},$$

where the expectation is over the player's internal randomness. We want  $\text{Regret}_T = o(T)$ .

## When no hypothesis is perfect: regret

Without realizability, an adversary can force a mistake every round (set  $y_t$  against your prediction), so “number of mistakes” is meaningless. Ask instead to be competitive with the **best fixed hypothesis in hindsight**.

**Goal: minimize regret**

$$\text{Regret}_T = \mathbb{E} \left[ \sum_{t=1}^T \mathbf{1}\{h_t(x_t) \neq y_t\} \right] - \min_{h \in \mathcal{H}} \sum_{t=1}^T \mathbf{1}\{h(x_t) \neq y_t\},$$

where the expectation is over the player's internal randomness. We want  $\text{Regret}_T = o(T)$ .

**Randomization is essential.** If  $h_t$  is a deterministic function of the past, the adversary knows it and sets  $y_t = -h_t(x_t)$ : you make  $T$  mistakes. If  $h, -h \in \mathcal{H}$ , one of them makes  $\leq T/2$ , so  $\text{Regret}_T \geq T/2$ .

**Rules of the game.** The adversary may know your algorithm and the distribution  $p_t$  you sample  $h_t$  from, but must commit to  $(x_t, y_t)$  before seeing the realized  $h_t$ .

## Non-separable case, IID: ERM again

**Strategy:**  $h_t \in \arg \min_{h \in \mathcal{H}} \sum_{s < t} \mathbf{1}\{h(x_s) \neq y_s\}$  (also called *follow the leader*).

Let  $h_* \in \arg \min_{h \in \mathcal{H}} R(h)$ . Since

$$\mathbb{E}[\min_h \sum_t \mathbf{1}\{h(x_t) \neq y_t\}] \leq \min_h \mathbb{E}[\sum_t \mathbf{1}\{h(x_t) \neq y_t\}] = TR(h_*),$$

## Non-separable case, IID: ERM again

**Strategy:**  $h_t \in \arg \min_{h \in \mathcal{H}} \sum_{s < t} \mathbf{1}\{h(x_s) \neq y_s\}$  (also called *follow the leader*).

Let  $h_* \in \arg \min_{h \in \mathcal{H}} R(h)$ . Since

$$\mathbb{E}[\min_h \sum_t \mathbf{1}\{h(x_t) \neq y_t\}] \leq \min_h \mathbb{E}[\sum_t \mathbf{1}\{h(x_t) \neq y_t\}] = TR(h_*),$$

$$\text{Regret}_T \leq \sum_{t=1}^T (\mathbb{E}[R(h_t)] - R(h_*)) \leq 1 + \sum_{t=2}^T \sqrt{\frac{8 \log(2|\mathcal{H}|)}{t-1}} \leq 1 + \sqrt{32 T \log(2|\mathcal{H}|)},$$

using the agnostic bound in expectation and  $\sum_{s=1}^T s^{-1/2} \leq 2\sqrt{T}$ .

## Non-separable case, IID: ERM again

**Strategy:**  $h_t \in \arg \min_{h \in \mathcal{H}} \sum_{s < t} \mathbf{1}\{h(x_s) \neq y_s\}$  (also called *follow the leader*).

Let  $h_* \in \arg \min_{h \in \mathcal{H}} R(h)$ . Since

$$\mathbb{E}[\min_h \sum_t \mathbf{1}\{h(x_t) \neq y_t\}] \leq \min_h \mathbb{E}[\sum_t \mathbf{1}\{h(x_t) \neq y_t\}] = TR(h_*),$$

$$\text{Regret}_T \leq \sum_{t=1}^T (\mathbb{E}[R(h_t)] - R(h_*)) \leq 1 + \sum_{t=2}^T \sqrt{\frac{8 \log(2|\mathcal{H}|)}{t-1}} \leq 1 + \sqrt{32 T \log(2|\mathcal{H}|)},$$

using the agnostic bound in expectation and  $\sum_{s=1}^T s^{-1/2} \leq 2\sqrt{T}$ .

### So in the IID world

$\text{Regret} = O(\sqrt{T \log |\mathcal{H}|})$ , i.e., average regret per round  $\rightarrow 0$  like  $1/\sqrt{T}$ , matching the  $1/\sqrt{n}$  agnostic rate.

## Non-separable case, IID: ERM again

**Strategy:**  $h_t \in \arg \min_{h \in \mathcal{H}} \sum_{s < t} \mathbf{1}\{h(x_s) \neq y_s\}$  (also called *follow the leader*).

Let  $h_* \in \arg \min_{h \in \mathcal{H}} R(h)$ . Since

$$\mathbb{E}[\min_h \sum_t \mathbf{1}\{h(x_t) \neq y_t\}] \leq \min_h \mathbb{E}[\sum_t \mathbf{1}\{h(x_t) \neq y_t\}] = TR(h_*),$$

$$\text{Regret}_T \leq \sum_{t=1}^T (\mathbb{E}[R(h_t)] - R(h_*)) \leq 1 + \sum_{t=2}^T \sqrt{\frac{8 \log(2|\mathcal{H}|)}{t-1}} \leq 1 + \sqrt{32 T \log(2|\mathcal{H}|)},$$

using the agnostic bound in expectation and  $\sum_{s=1}^T s^{-1/2} \leq 2\sqrt{T}$ .

### So in the IID world

$\text{Regret} = O(\sqrt{T \log |\mathcal{H}|})$ , i.e., average regret per round  $\rightarrow 0$  like  $1/\sqrt{T}$ , matching the  $1/\sqrt{n}$  agnostic rate.

**Adversarial?** Follow the leader is deterministic, so it suffers  $\Omega(T)$  regret. We need to **hedge**: keep playing hypotheses that have done well, but never commit entirely.

## Exponential weights (Hedge, multiplicative weights)

**Input:**  $d$  actions (“experts”), learning rate  $\eta > 0$

**Initialize:**  $w_1(i) = 1$  for all  $i \in [d]$

**for**  $t = 1, 2, \dots, T$ :

$W_t = \sum_{i=1}^d w_t(i)$  and  $p_t(i) = w_t(i)/W_t$

Play  $I_t \sim p_t$

Adversary reveals losses  $\ell_t \in [0, 1]^d$ ; pay  $\ell_t(I_t)$

$w_{t+1}(i) = w_t(i) e^{-\eta \ell_t(i)}$  for all  $i \in [d]$

**In words.** Unrolling the update,

$$p_t(i) \propto \exp\left(-\eta \sum_{s < t} \ell_s(i)\right),$$

a *softmax* over cumulative losses.

- $\eta \rightarrow \infty$ : all mass on the leader (follow the leader; fails)
- $\eta \rightarrow 0$ : uniform forever (ignores data; fails)
- In between: hedge, with  $\eta$  trading off the two

# Exponential weights

## Exponential weights (Hedge, multiplicative weights)

**Input:**  $d$  actions (“experts”), learning rate  $\eta > 0$

**Initialize:**  $w_1(i) = 1$  for all  $i \in [d]$

**for**  $t = 1, 2, \dots, T$ :

$W_t = \sum_{i=1}^d w_t(i)$  and  $p_t(i) = w_t(i)/W_t$

Play  $I_t \sim p_t$

Adversary reveals losses  $\ell_t \in [0, 1]^d$ ; pay  $\ell_t(I_t)$

$w_{t+1}(i) = w_t(i) e^{-\eta \ell_t(i)}$  for all  $i \in [d]$

**In words.** Unrolling the update,

$$p_t(i) \propto \exp\left(-\eta \sum_{s < t} \ell_s(i)\right),$$

a *softmax* over cumulative losses.

- $\eta \rightarrow \infty$ : all mass on the leader (follow the leader; fails)
- $\eta \rightarrow 0$ : uniform forever (ignores data; fails)
- In between: hedge, with  $\eta$  trading off the two

**Classification** is the case  $d = |\mathcal{H}|$ : action  $i$  is hypothesis  $h_i$  and, once  $y_t$  is revealed,  $\ell_t(i) = \mathbf{1}\{h_i(x_t) \neq y_t\}$ ; the expected loss is  $\mathbb{E}[\ell_t(I_t)] = \langle p_t, \ell_t \rangle$ . But nothing is specific to it: the adversary may pick *any*  $\ell_t \in [0, 1]^d$  after seeing  $p_t$  (not  $I_t$ ).

# Exponential weights: regret bound

## Theorem

For *any* sequence of loss vectors  $\ell_1, \dots, \ell_T \in [0, 1]^d$ , exponential weights satisfies

$$\text{Regret}_T = \sum_{t=1}^T \langle p_t, \ell_t \rangle - \min_{i \in [d]} \sum_{t=1}^T \ell_t(i) \leq \frac{\log d}{\eta} + \frac{\eta T}{8}.$$

With  $\eta = \sqrt{8 \log(d)/T}$  this is  $\text{Regret}_T \leq \sqrt{T \log(d)/2}$ .

# Exponential weights: regret bound

## Theorem

For any sequence of loss vectors  $\ell_1, \dots, \ell_T \in [0, 1]^d$ , exponential weights satisfies

$$\text{Regret}_T = \sum_{t=1}^T \langle p_t, \ell_t \rangle - \min_{i \in [d]} \sum_{t=1}^T \ell_t(i) \leq \frac{\log d}{\eta} + \frac{\eta T}{8}.$$

With  $\eta = \sqrt{8 \log(d)/T}$  this is  $\text{Regret}_T \leq \sqrt{T \log(d)/2}$ .

## Compare to IID + ERM (classification, $d = |\mathcal{H}|$ )

IID:  $O(\sqrt{T \log |\mathcal{H}|})$ , using IID in an essential way. Adversarial:  $\sqrt{T \log |\mathcal{H}|/2}$  with **no assumptions at all**. Same rate.

- Optimal up to constants: no algorithm has  $o(\sqrt{T \log d})$  worst-case regret.
- $\text{Regret}_T/T \rightarrow 0$ : as good per round as the best fixed action, eventually.
- Cost per round:  $O(d)$ . Hold that thought.

## Exponential weights: proof

**Potential function:**  $W_t = \sum_{i=1}^d w_t(i)$ . We bound  $\log(W_{T+1}/W_1)$  from above and below.

## Exponential weights: proof

**Potential function:**  $W_t = \sum_{i=1}^d w_t(i)$ . We bound  $\log(W_{T+1}/W_1)$  from above and below.

**Upper bound.** One step of the update is an expectation under  $p_t$ :

$$\frac{W_{t+1}}{W_t} =$$

## Exponential weights: proof

**Potential function:**  $W_t = \sum_{i=1}^d w_t(i)$ . We bound  $\log(W_{T+1}/W_1)$  from above and below.

**Upper bound.** One step of the update is an expectation under  $p_t$ :

$$\frac{W_{t+1}}{W_t} = \sum_{i=1}^d \frac{w_t(i)}{W_t} e^{-\eta \ell_t(i)} = \mathbb{E}_{I \sim p_t} [e^{-\eta \ell_t(I)}].$$

Since  $\ell_t(i) \in [0, 1]$ , **Hoeffding's lemma** (same tool as before, now applied to  $I \sim p_t$ ) gives

$$\log \mathbb{E}_{I \sim p_t} [e^{-\eta \ell_t(I)}] \leq -\eta \langle p_t, \ell_t \rangle + \frac{\eta^2}{8} \quad \implies \quad \log W_{T+1} \leq \log d - \eta \sum_{t=1}^T \langle p_t, \ell_t \rangle + \frac{\eta^2 T}{8}$$

by telescoping over  $t$  with  $W_1 = d$ .

## Exponential weights: proof

**Potential function:**  $W_t = \sum_{i=1}^d w_t(i)$ . We bound  $\log(W_{T+1}/W_1)$  from above and below.

**Upper bound.** One step of the update is an expectation under  $p_t$ :

$$\frac{W_{t+1}}{W_t} = \sum_{i=1}^d \frac{w_t(i)}{W_t} e^{-\eta \ell_t(i)} = \mathbb{E}_{I \sim p_t} [e^{-\eta \ell_t(I)}].$$

Since  $\ell_t(i) \in [0, 1]$ , **Hoeffding's lemma** (same tool as before, now applied to  $I \sim p_t$ ) gives

$$\log \mathbb{E}_{I \sim p_t} [e^{-\eta \ell_t(I)}] \leq -\eta \langle p_t, \ell_t \rangle + \frac{\eta^2}{8} \implies \log W_{T+1} \leq \log d - \eta \sum_{t=1}^T \langle p_t, \ell_t \rangle + \frac{\eta^2 T}{8}$$

by telescoping over  $t$  with  $W_1 = d$ .

**Lower bound.** For any fixed  $i \in [d]$ :  $\log W_{T+1} \geq \log w_{T+1}(i) = -\eta \sum_{t=1}^T \ell_t(i)$ .

**Combine** and divide by  $\eta$ :  $\sum_{t=1}^T \langle p_t, \ell_t \rangle - \sum_{t=1}^T \ell_t(i) \leq \frac{\log d}{\eta} + \frac{\eta T}{8}$  for every  $i$ . ■

# Today in one table

| Setting            | Assumption | Algorithm           | Guarantee                                                         |
|--------------------|------------|---------------------|-------------------------------------------------------------------|
| Batch, realizable  | IID        | ERM                 | $R(\hat{h}) \leq \log( \mathcal{H} /\delta)/n$                    |
| Batch, agnostic    | IID        | ERM                 | $R(\hat{h}) - R(h_*) \leq \sqrt{2 \log(2 \mathcal{H} /\delta)/n}$ |
| Online, realizable | IID        | ERM each round      | $O(\log  \mathcal{H}  \log T)$ mistakes                           |
| Online, realizable | none       | Halving             | $\log_2  \mathcal{H} $ mistakes                                   |
| Online, agnostic   | IID        | ERM each round      | $O(\sqrt{T \log  \mathcal{H} })$ regret                           |
| Online, agnostic   | none       | Exponential weights | $\sqrt{T \log( \mathcal{H} )/2}$ regret                           |

# Today in one table

| Setting            | Assumption | Algorithm           | Guarantee                                                         |
|--------------------|------------|---------------------|-------------------------------------------------------------------|
| Batch, realizable  | IID        | ERM                 | $R(\hat{h}) \leq \log( \mathcal{H} /\delta)/n$                    |
| Batch, agnostic    | IID        | ERM                 | $R(\hat{h}) - R(h_*) \leq \sqrt{2 \log(2 \mathcal{H} /\delta)/n}$ |
| Online, realizable | IID        | ERM each round      | $O(\log  \mathcal{H}  \log T)$ mistakes                           |
| Online, realizable | none       | Halving             | $\log_2  \mathcal{H} $ mistakes                                   |
| Online, agnostic   | IID        | ERM each round      | $O(\sqrt{T \log  \mathcal{H} })$ regret                           |
| Online, agnostic   | none       | Exponential weights | $\sqrt{T \log( \mathcal{H} )/2}$ regret                           |

## The message

Every IID guarantee has an adversarial twin with the same rate: the same  $\log |\mathcal{H}|$ , the same  $\sqrt{T}$ .  
Dropping the IID assumption cost nothing.

Not yet addressed: losses with **curvature** (squared loss, log-wealth), and **computation** when  $|\mathcal{H}|$  is huge.