Online Set Learning from Precision and Recall Feedback

Lee Cohen[1] Yishay Mansour[2,3] Shay Moran[4,3] Han Shao[5]
[1]Stanford University [2]Tel Aviv University [3]Google Research [4]Technion [5]University of Maryland

arXiv:2605.09565
NeurIPS 2026

TL;DR

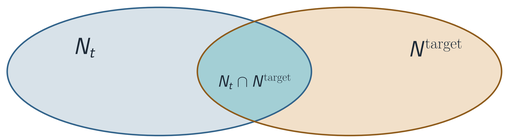
Finite VC dimension characterizes learnability, but precision feedback fundamentally changes the algorithms: ordinary ERM can fail, and proper learning can be impossible.

Motivation

Many applications require predicting a **set of relevant items**: **content recommendation**, **information retrieval**, **hiring**, and **cybersecurity**. We want to **cover** the target while **avoiding irrelevant items**—yet receive feedback on only one sampled item per round.

Online Set Learning Model

- Universe $\mathcal{X}$, known available set $X \subseteq \mathcal{X}$; hidden target $N^{\text{target}} \subseteq X$.
- Hypotheses $\mathcal{H} \subseteq 2^{\mathcal{X}}$; at round t, predict a set $N_t \subseteq X$.
- One of two feedback types is drawn uniformly at random:



Precision $p(N_t) = \dfrac{|N_t \cap N^{\text{target}}|}{|N_t|}$ **Recall** $r(N_t) = \dfrac{|N_t \cap N^{\text{target}}|}{|N^{\text{target}}|}$

Expected reward: $g(N_t) = \frac{1}{2}(p(N_t) + r(N_t))$

Precision feedback: $v_t \sim \text{Unif}(N_t)$, observe $\mathbf{1}\{v_t \in N^{\text{target}}\}$.
Recall feedback: $u_t \sim \text{Unif}(N^{\text{target}})$, observe $\mathbf{1}\{u_t \in N_t\}$.

$$\text{Reg}_T = T \max_{N \in \mathcal{H}} g(N) - \Sigma_{t=1}^{T} g(N_t).$$

The Fundamental Challenge

- **Recall is full-information-like:** a positive item u_t can be checked against *every* candidate $N \in \mathcal{H}$.
- **Precision is bandit-like:** the sampled item v_t depends on *the set we chose*; unplayed hypotheses cannot be evaluated directly.

Main Results

Question	Answer
Which classes are learnable?	$\text{VCdim}(\mathcal{H}) < \infty$ *iff* learnable
Realizable regret	$O(d \log^2 T)$; $\Omega(d)$ lower bound
Agnostic regret	$O(d^{1/4} T^{3/4})$
Can the algorithm be proper?	**No:** proper regret can be $\Omega(T)$, even for $d = 1$

$d = \text{VCdim}(\mathcal{H})$. Bounds suppress confidence / logarithmic factors as appropriate.

Realizable Setting: Why ERM Fails

Let $N^{\text{target}} = N^\star \in \mathcal{H}$.

Issue: Consistency with recall samples does not guarantee high precision.

$$\mathcal{H} = \big|\{0\}, \{0,1\}, \ldots, \{0,n\}\big|, \qquad N^{\text{target}} = \{0\}.$$

Every recall sample equals 0, but $\{0, i\}$ has precision $1/2$.

Lesson: Favor smaller consistent sets to control false positives.

Algorithm: Minimum-Size Consistent Hypothesis

Choose the **smallest** set consistent with all observed recall items:

$$N_t \in \arg \min_{\substack{N \in \mathcal{H} \\ u_1, \ldots, u_{t-1} \in N}} |N \cap X|.$$

This is a **size-regularized ERM / maximum-likelihood rule**.

Importantly, it uses only recall observations.

Why Does This Work?

1. **Uniform convergence for recall.** Recall samples are i.i.d. from $\text{Unif}(N^{\text{target}})$. With $d = \text{VCdim}(\mathcal{H})$, a consistent hypothesis has
$$1 - r(N_t) = O(d/t).$$
2. **Minimum cardinality controls precision.** Because the realizable target is consistent, $|N_t| \le |N^{\text{target}}|$. Thus
$$p(N_t) = r(N_t) \frac{|N^{\text{target}}|}{|N_t|} \;\ge\; r(N_t).$$
3. **Both errors become small.** A high-recall, minimum-size hypothesis automatically has high precision, *without* precision feedback.

Realizable Regret Guarantee

$$\text{Reg}_T = O\big(d \log^2 T\big)$$

Any learner incurs $\Omega(d)$ expected regret on some realizable instance.
The logarithmic gap remains open.

A Connection to Online Classification

A binary online classifier with mistake bound M can be converted into a set learner with regret at most M; in particular,

$$\text{Reg}_T \le \text{Ldim}(\mathcal{H}).$$

But **Littlestone dimension is not the right characterization**: it may be infinite even when VC dimension is finite and our set-learning problem is learnable.

Open Questions

- Can agnostic regret reach $O(\sqrt{dT})$ in general?
- Can the realizable $\log^2 T$ factor be removed?

Agnostic Setting: Proper Learning Can Fail

Now N^{target} need not correspond to any member of $\mathcal{H}$. A striking obstacle is that **every proper algorithm** can have **linear regret**, even when $\text{VCdim}(\mathcal{H}) = 1$.

- Construct many large, disjoint hypotheses: finding the best one's precision becomes a many-armed exploration problem.
- Yet trivial *improper* predictions $\varnothing$ or X obtain reward at least $1/2$.

Lesson: competing with $\mathcal{H}$ may require playing **outside** $\mathcal{H}$.

Key Insight: Recover Precision from Recall

Recall is estimable *simultaneously for all candidate sets*, but precision is observed only on the played set. The missing bridge is the **size of the hidden target**:

$$p(N) = r(N) \frac{|N^{\text{target}}|}{|N|}.$$

For a played N_t, precision feedback provides $y_t \in \{0,1\}$ with $\mathbb{E}[y_t \mid N_t] = p(N_t)$; hence,

$$\mathbb{E}\left[\frac{y_t |N_t|}{r(N_t)} \,\middle|\, N_t\right] = |N^{\text{target}}|, \qquad r(N_t) > 0.$$

Replacing r by its empirical estimate gives an estimator $n \approx |N^{\text{target}}|$, and therefore

$$p(N) = r(N) \frac{n}{|N|} \qquad \text{for all nonempty candidate sets } N.$$

The Estimation Challenge

If **recall or precision is tiny**, the target-size estimator is unstable: inverse-recall weights and sampling noise become too large.
An algorithm must first **find informative sets** before comparing sets.

APRIL: High-Level Algorithmic Ideas

1. Recall screening. Estimate recall and eliminate low-recall sets.
Select a candidate with a favorable estimated recall-to-size ratio.

↓

2. Precision calibration. Play the candidate to estimate $|N^{\text{target}}|$.
Infer precision and eliminate low-precision sets.

↓

3. Learn or fall back. Select the set maximizing estimated reward; if no set survives, predict $\varnothing$.

Principle: combine uniform recall estimates with safe precision estimates; use improper predictions when needed.

Agnostic Regret Guarantee

$$\text{Reg}_T = O\big(d^{1/4} T^{3/4}\big)$$

With an optimality margin bounded below by a constant, $\alpha_{\mathcal{H}} = 2 \max_{N \in \mathcal{H}} g(N) - 1 \ge c > 0$, the bound improves to $O(\sqrt{dT})$.