Admin:  PSS due Thurs
        Rocco OH 11-1 TODAY  (no Rocco OH
        OH 9-11 Fri.)

Last time:
        review; new SQ model:
        $\text{pred} \, x(\cdot, \cdot), \quad \text{STAT}(c, \theta) \rightarrow \text{est} \, \hat{P}_x$
        $\forall \alpha, \beta \frac{1}{2},$ \text{tol.}

$P_x = \frac{P_r \left[ x(x, c(x)) = 1 \right]}{(x, c(x)) \sim \text{EX}(c, \theta)}$

$P_x - \alpha \leq \hat{P}_x \leq P_x + \beta,$ \text{(easy)}

E eff SQ learnable $\Rightarrow$ E eff PAC learnable.

Today: interesting variant:

Thm: If E eff. SQ learnable, then
      E is eff. PAC learnable even in presence
      of RCN \, $0 < m < \frac{1}{2}$: \, time \ldots $\frac{1}{1-2m}$ \ldots

Rel. betw. SQ learning & PAC learning.

$x = 1001001101; \quad c(x) = 0 \text{ or } 1 \text{? PAC }$

"$x_1 = 1": \quad 0.734$\n
"$x_1 = 1, x_2 = 0": \quad 0.216$

To establish this:

show how to simulate a STAT(\(c, \theta\)) oracle
when all we have is EX(\(c, \theta\)).
We're given a pair \((X, \mathcal{E})\).

We get: access to \(\mathbb{E}X \sim (\mathcal{E}, \theta)\).

We need to: estimate \(P_X\).

Main idea: break up \(X\) into 2 pieces, where we completely understand effect of noise on each piece.

\[ X = \{ x \in X: x(x, 0) \neq x(x, 1) \} \]

\[ X_2 = \{ x \in X: x(x, 0) = x(x, 1) \} \]

Key obs: Given \(x\), can eval. \(x(x, 0), x(x, 1)\) & determine whether \(x \in X_1\) or \(x \in X_2\).

Consider \(x(x, b) = \text{"b = 1 \& x = 0"}\)

\[ X_1 = \{ x: x_1 = 0 \} \quad X_2 = \{ x: x_1 = 1 \} \]

Let \(\mathbb{P}_1 = \mathbb{P}_{\mathcal{E}}[x \in X_1]. \quad \forall \mathcal{E} \subseteq \mathcal{X}_1\)

Let \(\mathbb{P}_1 = \mathbb{P}_{\mathcal{E}}[x \in \mathcal{E}] = \mathbb{P}_{\mathcal{E}}[x \in \mathcal{E} | x \in \mathcal{E}_1] \frac{x \in \mathcal{E}_1}{x \in \mathcal{E}}\)

Goal follows immed. from:
\[ L1:\text{ Have } P_x = \frac{Pr_{(x,b) \sim \text{EX}^\sim(c,\theta)}}{1-2-m} \]

\[ + Pr_{(x,b) \sim \text{EX}^\sim(c,\theta)} [x(x,b) = 1 + x \in X_2] \]

\[ L2: \text{ Poss. to eff. estimate } P_x \text{ given } x, \text{ EX}^\sim(c,\theta). \]

\[ \text{ Pf of L1: } P_x = Pr_{x \sim \theta} [x(x,c(x)) = 1] \]

\[ = Pr_{x \sim \theta} [x(x,c(x)) = 1 + x \in X_2] + Pr_{x \sim \theta} [x(x,c(x)) = 1 + x \in X_2] \]

\[ = Pr_{x \sim \theta} [x(x,c(x)) = 1] \quad \text{B/C } D_1 \text{ supp. on } X, \text{ where "noise always matters"} \]

\[ = Pr_{x \sim \theta_1} [x(x,c(x)) = 1] \quad \text{Consider noisy prob:} \]

\[ Pr_{(x,b) \sim \text{EX}^\sim(c,\theta)} [x(x,b) = 1] = (1-m) Pr_{x \sim \theta_1} [x(x,c(x)) = 1] \]

\[ + m \cdot Pr_{x \sim \theta_1} [x(x,c(x)) = 0] \]

\[ 1 - Pr_{x \sim \theta_1} [x(x,c(x)) = 1] \]
\[ y = \eta + (1 - 2\eta) \cdot \Pr_{x \sim d_1} [x(x, c(x)) = 1]. \]

So \( \Pr_{x \sim d_1} [x(x, c(x)) = 1] = \frac{\Pr_{(x, b) \sim \mathcal{E}(c, \theta_1)} [x(x, b) = 1] - m}{1 - 2\eta}. \)

**L2:** Can est. \( p_x \) given \( x, \mathcal{E}(c, \theta_1) \).
- Can est. \( p_i \): \( p_i = \Pr_{x \sim d_1} [x \in X], \)
  \[ = \Pr_{x \sim d_1} [x(x, 0) \neq x(x, 1)]. \]
  Draw \( (x, b) \sim \mathcal{E}(c, \theta_1) \), compute \( x(x, 0), x(x, 1) \).
- Want \( p_i \): \( \Pr [ \_ ] - m \) could be as big as \( \frac{1 - m}{1 - 2\eta}, \) large...
  
  If \( p_i \) very small, so is \( \eta, \) \( \frac{1 - m}{1 - 2\eta} \) large...
  
  (depends on \( 1 - 2\eta \ldots \)) \( 0 \) is good est. of product.

If \( p_i \) not so small: \( \Pr_{x \sim d_1} [x \in X] \) not so small, so can draw \( x \sim d_1 \) at a \( 1/p_i \) factor slowdown, which is affordable. In this case, we'll est. the frac.

To est. the frac:
\[ \Pr \left( x(x,b) = 1 \right) - m \] have
\[ (x,b) \sim \text{EX}(c, \theta) \]
\[ 1 - 2m \]
\[ \rightarrow \text{reject draws where } x \in \mathbb{X}_2. \text{ Gives access} \]
\[ \text{to } \text{EX}^{-1}(c, \theta), \text{ etc.} \]

Last piece: given \( \text{EX}^{-1}(c, \theta) \), easy to est
\[ \Pr \left( x(x,b) = 1 \land x \in \mathbb{X}_2 \right) \]
\[ (x,b) \sim \text{EX}^{-1}(c, \theta) \]

Fact: Many PAC algs can be easily translated
to SQ algs. All such algs yield RCN-tolerant
PAC algs.

Recap:
\( C \text{ eff SQ learnable } \Rightarrow C \text{ eff PAC learnable.} \)

Q:
\( \Rightarrow \) ?

\( \boxed{\text{NO}} \)

\( C = \text{all parity functions over } \{0,1\}^n \).

\( S \subseteq \{1,2,\ldots,n\} \)
\( S = \{3, 4, 7, 8\} \)
\[ \text{PAR}_5(x_1, \ldots, x_n) = x_3 \oplus x_4 \oplus x_7 \oplus x_8 \]
\[ = x_3 + x_4 + x_7 + x_8 \mod 2 \]

\[ \text{FACT} : \text{There is an eff. PAC learning alg.} \]
\[ \text{for } C = \text{all } 2^n \text{ PAR}_5 \text{ functions.} \]

\[ \text{Idea: solve syst. of lin. eq. to find} \]
\[ \text{consistent parity fn., given a labeled data} \]
\[ \text{set.} \]
\[ a_1, \ldots, a_n \text{ are 0/1 variables} \]
\[ a_i = 1 \iff x_i \text{ is in target PAR}_5. \]
\[ (1, 1, 0, 0, 1; 1) \]
\[ 0, 1, 0, 1, 0; 0 \]
\[ \rightarrow \]
\[ a_1 + a_2 + a_5 = 1 \]
\[ a_2 + a_4 = 0 \]
\[ \text{combine} \]
\[ a_1 + a_4 + a_5 = 1. \]

Gaussian elim. provides an CHF.

\[ |E| = 2^n \implies m = O \left( \frac{n}{\epsilon} + \frac{1}{\epsilon} \log \frac{1}{\delta} \right) \text{ many ex.} \]

Unclear how to have an SQ variant of this (seems to use indiv. examples...)

Turns out: impossible.
Next time: define a criterion (having many uncorrelated concepts).

- \( C = \text{PAR Functions} \) meets this criterion.

---

OH 11-1 today