Outline for today

- Finish VC Dimension discussion
- Agnostic learning
- Mistake bounded learning

PAC Learning: Summary

- Probability[some hypothesis $h \in H$ consistent with m training examples will have $error(h) > \epsilon] \leq \delta$, provided

$$m \geq \frac{1}{\epsilon}(\ln |H| + \ln(1/\delta))$$

or

$$m \geq \frac{1}{\epsilon}(8VC(H) \log(13/\epsilon) + 4 \log(2/\delta))$$

- Does this apply even if $c \notin H$?

- What if learner avoids the assumption that $C \subseteq H$?
  - This setting called *agnostic learning*
  - Assume learner outputs $h \in H$ with minimum training error...

Agnostic learning

- Output h with lowest training error $r(h)$
- What is probability that true $error(h) > r(h) + \epsilon$?
- Can use Hoeffding bounds: bounds deviation between probability of heads (in flipping a coin), and the fraction of heads observed in m trials

Agnostic learning cont'd.

- Hoeffding bound: For a *single* hypothesis h_0

$$Pr[error(h_0) > r(h_0) + \epsilon] \leq e^{-2m\epsilon^2}$$

- As before, must consider all $|H|$ hypotheses:

$$Pr[(\exists h_i \in H) error(h_i) > r(h_i) + \epsilon] \leq |H| e^{-2m\epsilon^2}$$

- So, can reduce this probability to $\leq \delta$ by using m examples

$$m \geq \frac{1}{2\epsilon^2}(\ln |H| + \ln(1/\delta))$$

Compare with earlier result, which assumed $r(h) = 0$ (i.e., $C \subseteq H$):

$$m \geq \frac{1}{\epsilon}(\ln |H| + \ln(1/\delta))$$

Mistake Bounded Learning

New question:

- Suppose learner must predict the target value for each new instance x, before teacher provides the correct target value $c(x)$

- Can we bound the *total number of mistakes* learner will make before *exactly* learning c (i.e., before outputing an h such that $(\forall x)(h(x) = c(x))$?

Mistake Bound for FindS

Consider $H = CNF$, where

$h =$ conjunction of up to n boolean literals and negations

(e.g., $h = Tall \wedge \neg Young$)

Consider the algorithm $FindS$:

- Initialize S to $l_1 \wedge \neg l_1 \wedge l_2 \wedge \neg l_2 \ldots$
- For each positive training instance x
  - Remove any literal inconsistent with x

How many mistakes will $FindS$ make before exactly learning $c(x)$?