Dana Van Aken

Problem 2: Learning Theory

Problem 2.1: Learning Decision Lists

1. Upper bound for number of all possible DLs for an n variable input.

   We have:

   - 2^n possible inputs (n variables, inputs $\in \{1,0\}$)
   - 2^n possible outputs (n variables, outputs $\in \{1,0\}$)
   - $n!$ possible ways to arrange the order of the decision lists.

   upper bound on size of hypothesis space,

   $$\boxed{|H| = 2^n \cdot 2^n \cdot n! = 4^n n!}$$

2. Provide an upper bound on the probability that h is consistent with the provided training examples

   - h: error $(h) > \epsilon$

   $P(h \text{ is consistent with 1 example}) \leq 1-\epsilon$

   $P(h \text{ is consistent with all m training examples}) \leq (1-\epsilon)^m$

3. Provide an upper bound on the probability that our algorithm learns a DL with error$(h) > \epsilon$

   - If there are K hypotheses (DLs):

     $h_1 \ldots h_K$ s.t. $\{\text{error}(h_i) > \epsilon\}_{i=1}^{K}$

   - then $P(\text{algorithm learns a DL}) = K(1-\epsilon)^m \leq |H|(1-\epsilon)^m \leq \underbrace{|H|e^{-\epsilon m}}_{\text{convex upper bound}}$

$$\boxed{P(\text{error}(h) \geq \epsilon) \leq |H|e^{-\epsilon m} = 4^n n! \, e^{-\epsilon m} \leq \delta}$$

4. Provide a lower bound on the number of training examples needed to guarantee that our algorithm learns a DL with error $(h) > \epsilon$ is less than δ

$$|H|e^{-\epsilon m} \leq \delta$$

$$= \ln\left[|H|e^{-\epsilon m} \leq \delta\right]$$

$$= \ln|H| + \ln(e^{-\epsilon m}) \leq \ln\delta$$

$$= \ln|H| - \epsilon m \leq \ln\delta$$

$$= \frac{-\cancel{\epsilon} m}{-\cancel{\epsilon}} \leq \frac{\ln\delta - \ln|H|}{-\epsilon}$$

$$= m \geq \frac{-\ln\delta + \ln(4^n n!)}{\epsilon}$$

$$\boxed{m \geq \frac{\ln\frac{1}{\delta} + n\cdot\ln 4 + \ln n!}{\epsilon}}$$

Problem 2.2: VC Dimension

1. Intervals in $\mathbb{R}$:

- 1 point: [+]   − [ ]
- 2 points: [+] −   [ ] − −   − [+]   [+ +]
- 3 points: [+ +] −   − [+] −   + − + ← cannot be shattered by H

$\boxed{\text{VC-dimension} = 2}$

2. Circles in $\mathbb{R}^2$

- 1 point:

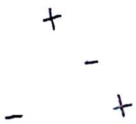
- 2 points:

- 3 points:

- 4 points:

← cannot be shattered

← cannot be shattered



← cannot be shattered

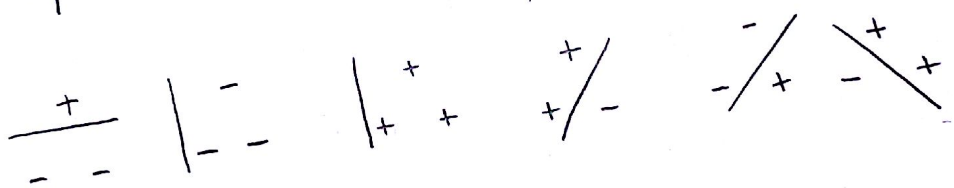
$\boxed{\text{VC-dimension} = 3}$

3. A Line in $\mathbb{R}^2$:

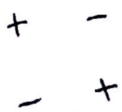
- 1 point:

- 2 points:

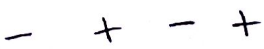
- 3 points



- 4 points:

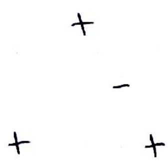


← cannot be shattered



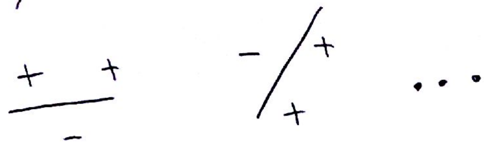
← cannot be shattered

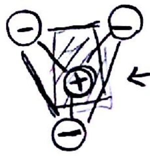
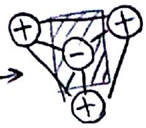
$\boxed{\text{VC-dimension} = 3}$

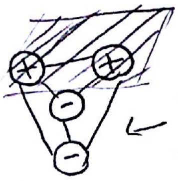


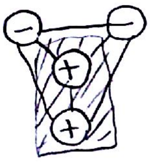
← cannot be shattered

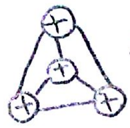
4. A plane in $\mathbb{R}^3$

- 1 point 
- 2 points 
- 3 points 
- 4 points (examine tetrahedrals)

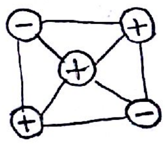
 ← Plane slices through top → 

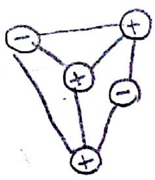
 ← plane slices diagonally

 ← plane slices diagonaly

 ← plane above or below (same case for ⊖)

- 5 points

 ← square pyramid cannot be shattered

 ← cannot be shattored → 

$\boxed{\text{VC-dimension} = 4}$

5. Problems 1, 3, and 4 represent half-planes in $\mathbb{R}^1$, $\mathbb{R}^2$, and $\mathbb{R}^3$, respectively. In each case, the VC-dimension is equal to the dimension + 1. Therefore, I would expect a half-space in $\mathbb{R}^n$ to have a VC-dimension of $n+1$.

Scanned by CamScanner