Computational Learning Theory

2 Frameworks

- PAC :- Probably Approximately Correct
- Mistake Bound

① Sample Complexity :- How many training data are needed for learner to converge

② Computational Complexity :- How much Computation per data point

(3) Mistake bound :- How mistakes the learner makes before converging

$\mathcal{E}$ - Exhaustive Version Space

h_i = true erro $(h_i) = \epsilon$

probability of getting one random instance correct by h_i = $(1 - \epsilon)$

training data contain m random samples

probability true error (h_i) is ϵ and getting m random instance correct

$= (1 - \epsilon)^m \leq e^{-\epsilon m}$

$$(1-\varepsilon)^m = e^{-\varepsilon m}$$

$$k e^{-\varepsilon m}$$

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

$$m \geq \frac{1}{\varepsilon}\left[\ln(|H|) + \ln\left(\tfrac{1}{\delta}\right)\right]$$

$$\Pr\Big[\text{true probability of head} > \text{Calculate probability of head}\Big] \leq e^{-2m\varepsilon^2}$$

$$\Pr\Big[\text{true error} > \text{Observed / Training Error}\Big] \leq e^{-2m\varepsilon^2}$$

$$\delta \leq |H| e^{-2m\varepsilon^2}$$

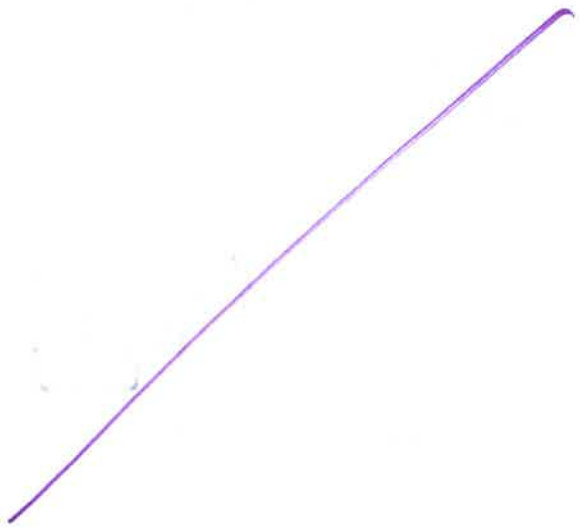
$$m \geq \frac{1}{2\varepsilon^2}\left[\ln|H| + \ln\left(\frac{1}{\delta}\right)\right]$$

$l_1 \quad l_2 \quad \ldots \quad l_n$

$l_1 \wedge l_2$

$l_3 \wedge \neg l_5$

$|H| = 3^n$



VC dimension of linear decision surface in an n-dimensional space is n+1

Let X be 3 boolean literals

X		
0	0	0
1	0	1

H → Conjunction of 3 boolean literals

VC(H) ? VC(H) = 3 ⇒ n

instance 1 :- 100 } $\neg l_1 \wedge \neg l_2 \wedge \neg l_3$

instance 2 :- 010

instance 3 :- 001 } $\neg l_1 \wedge l_3$

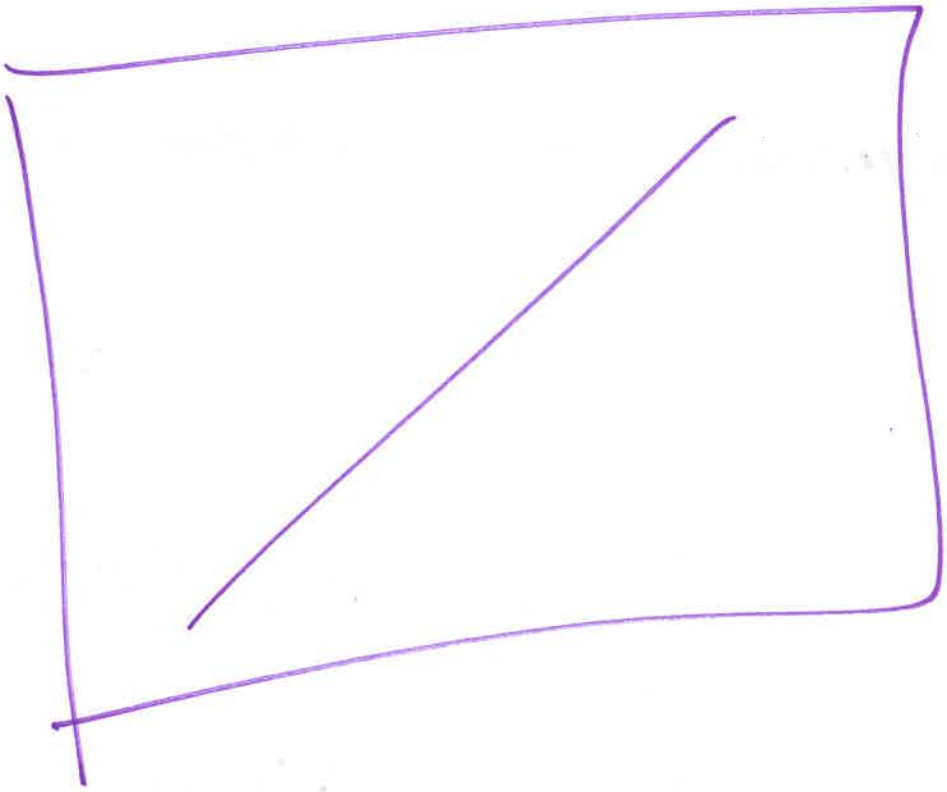
$VC(H) = 2$

$VC(H) \neq 3$

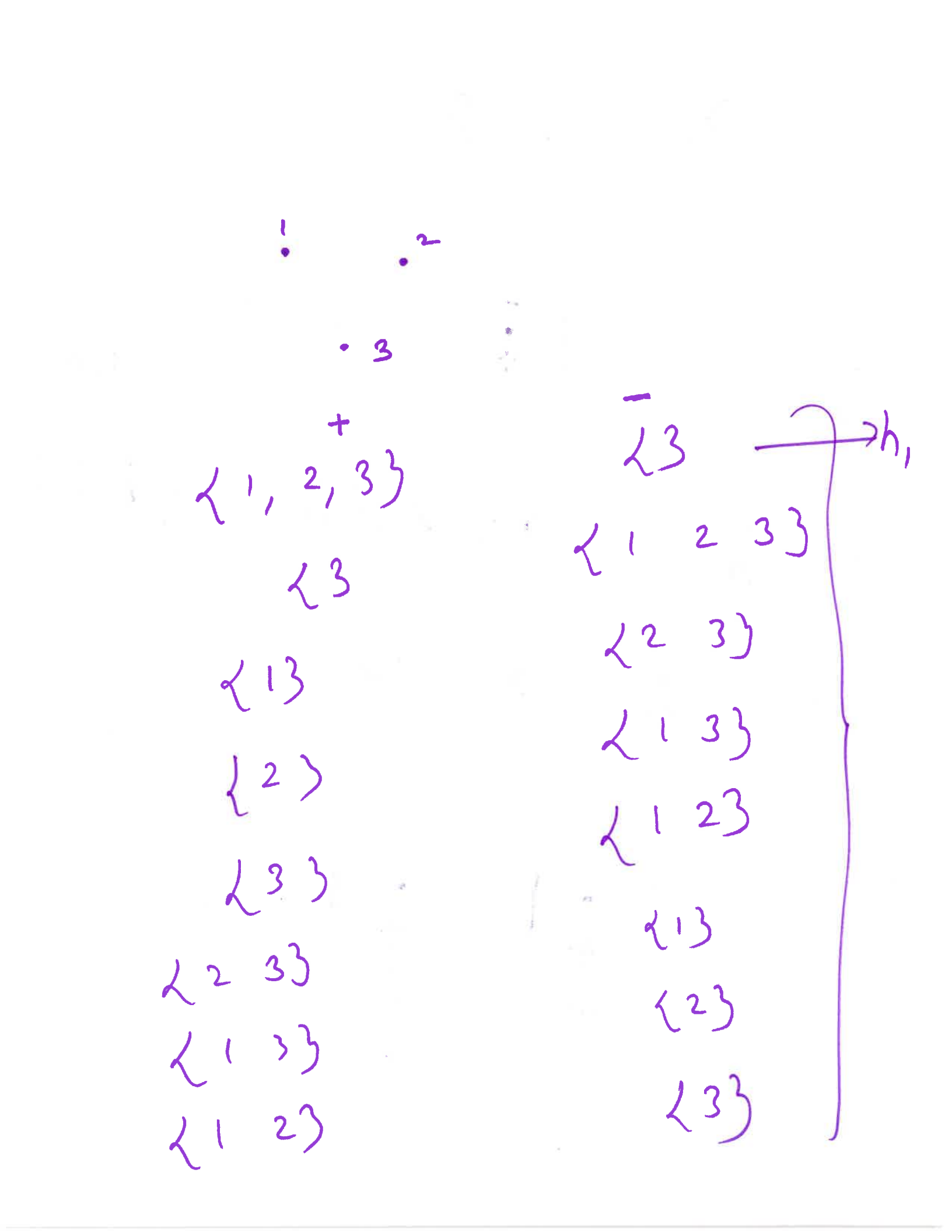
$|H| = \infty$

$X = \{x_1, x_2\}$

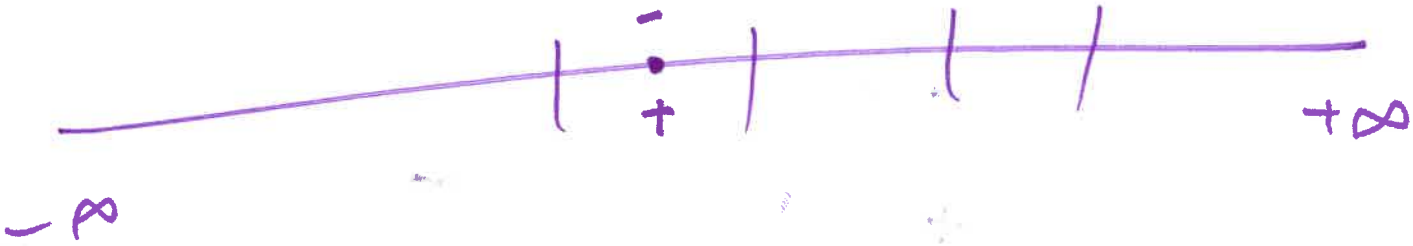
$x_1, x_2 = \{\mathbb{R}\}$



$H =$ Line


1
2
3
+
{1, 2, 3}
{3
{1 3
{2
{3}
{2 3}
{1 3}
{1 2}
-
{3
{1 2 3}
{2 3}
{1 3}
{1 23
{1
{2}
{3}
h1

$X = \{\mathbb{R}\}$



H $a \leq x \leq b \quad a, b \in \mathbb{R}$

$2 \leq x \leq 10$