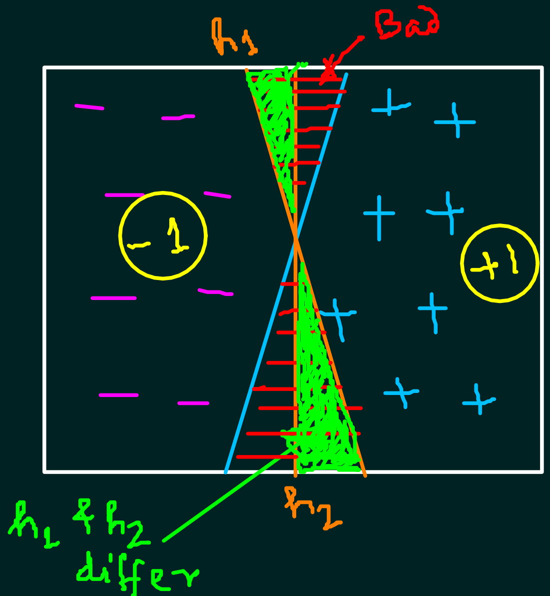
▶ Generalization in Learning: $\boxed{E_{out}(g) \approx 0}$ $\Rightarrow$ $E_{in}(g) \approx 0$ and $E_{in}(g) \approx E_{out}(g)$

↳ Can $E_{out}(g) \approx E_{in}(g)$ ? $\rightarrow$ $P\left[\,|E_{in}(g) - E_{out}(g)| > \epsilon\,\right] \leq 2\boxed{M}e^{-2\epsilon^2 N}$

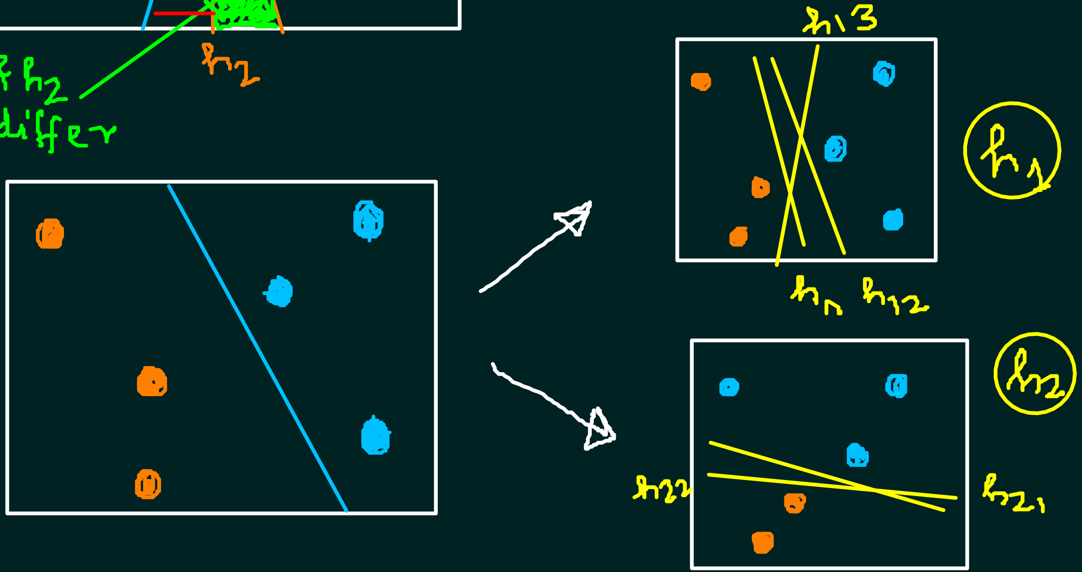
How to reduce?




Two hypothesis $h_1$ and $h_2$

overlap $\Rightarrow$ $|E_{in}(h_1) - E_{out}(h_1)| \approx |E_{in}(h_2) - E_{out}(h_2)|$

Question: What can replace M ?




▷ Hypothesis, $h : \mathcal{X} \rightarrow \{+1, -1\}$

▷ Dichotomies, $h : \{x_1 \dots x_N\} \rightarrow \{+1, -1\}$

$|\mathcal{H}| \rightarrow$ infinite

$|\mathcal{H}(x_1 \dots x_N)| \leq 2^N$ ← at most

$\Downarrow$

growth function $\boxed{m_{\mathcal{H}}(N)}$ ✓

$$m_{\mathcal{H}}(N) = \max_{x_1 \dots x_N \in \mathcal{X}} |\mathcal{H}(x_1, \dots, x_N)|$$

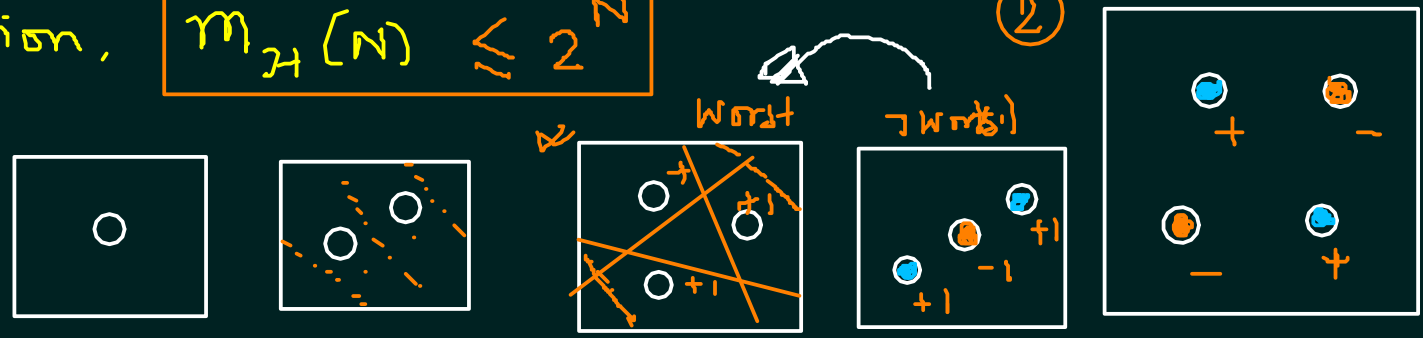
# Solution: Count the number of $\boxed{\text{DICHOTOMIES}}$ ✓ with N training Points

▷ The Growth Function, $\boxed{m_{\mathcal{H}}(N) \leq 2^N}$ 2

→ 2D - Perceptron:

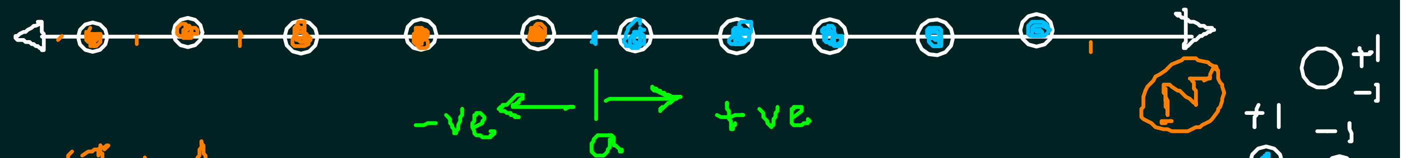
$m_{\mathcal{H}}(N) < 2^N$

Break Point $(K = 4)$



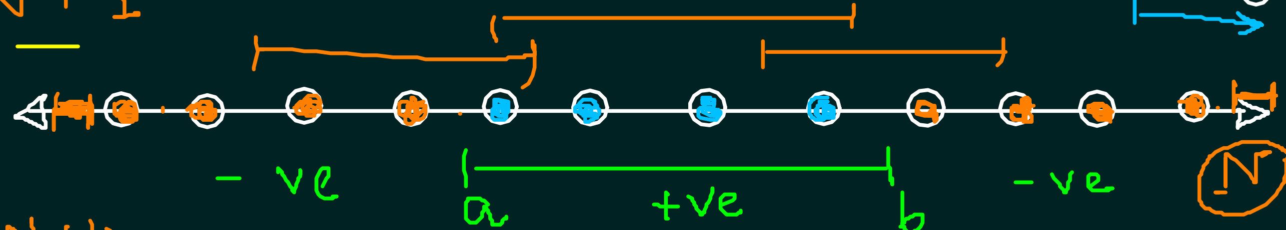

$m_{\mathcal{H}}(1) = 2$ $m_{\mathcal{H}}(2) = 4$ $m_{\mathcal{H}}(3) = 8$ $\not\Rightarrow$ $m_{\mathcal{H}}(4) = 14$

→ Positive Rays:



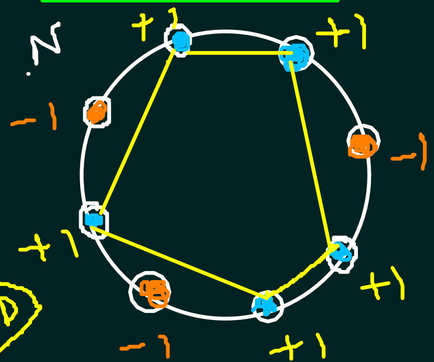

(poly) B.P. = 2 $m_{\mathcal{H}}(N) = N + 1$

→ Positive Interval:




(poly) B.P. = 3 $m_{\mathcal{H}}(N) = \binom{N+1}{2} + 1$

→ Convex Set:

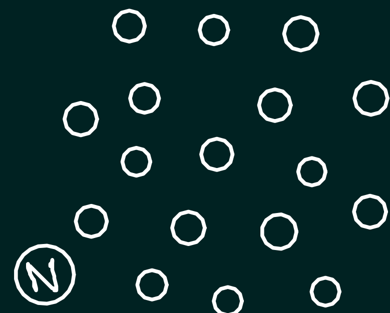


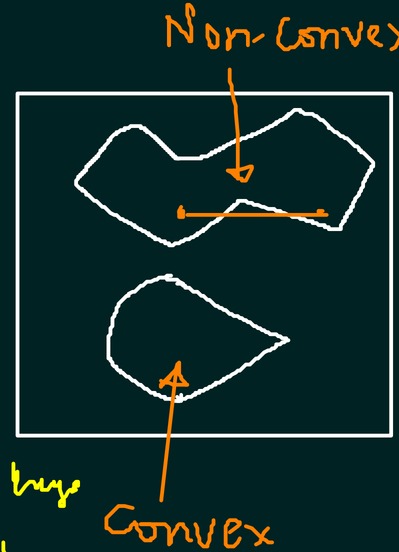

B.P. = ∞

SHATTERED

$m_{\mathcal{H}}(N) = 2^N$

(exp)




▷ $P\left[\,|E_{in} - E_{out}| > \epsilon\,\right] \leq 2 M e^{-2\epsilon^2 N}$ → huge

poly 0.01

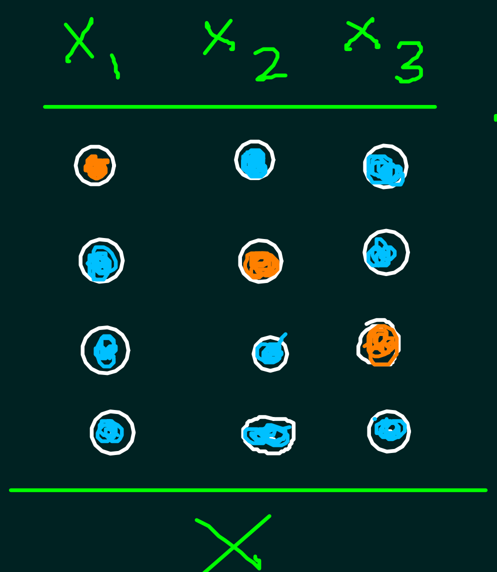
- $m_{\mathcal{H}}(N) \leq 2^N \longrightarrow$ potential candidate for M

- Break Point, K when $m_{\mathcal{H}}(K) < 2^K$

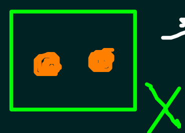
  (K-1) can be SHATTERED!!

- Prove that:

  1. if B.P. exists $\Rightarrow m_{\mathcal{H}}(N)$ is Poly.
     B.P. does not exist $(\infty) \Rightarrow 2^N$
  2. Indeed $m_{\mathcal{H}}(N)$ can replace M.



$$2^3 = 8 \rightarrow 5$$

$\boxed{B.P. = 2} \rightarrow$  $\rightarrow$ Prohibited

$\rightarrow$ Reduces the $2^N$ space

$\rightarrow$ K be a break point so does (K+1) too.

▷ If $k$ is a Break Point, then $m_{\mathcal{H}}(N)$ is Poly. $O(N^{k-1})$

→ $m_{\mathcal{H}}(N) \le \underset{①}{- - - -} \le \underset{②}{- - -} \le$ A poly. $\le \sum_{i=0}^{k-1} \binom{N}{i}$

$\boxed{B(N,k)}$ ← max no. of dichotomies with $N$ points having B.P. $= k$ ×

?

$$B(N,k) = \alpha + 2\beta$$

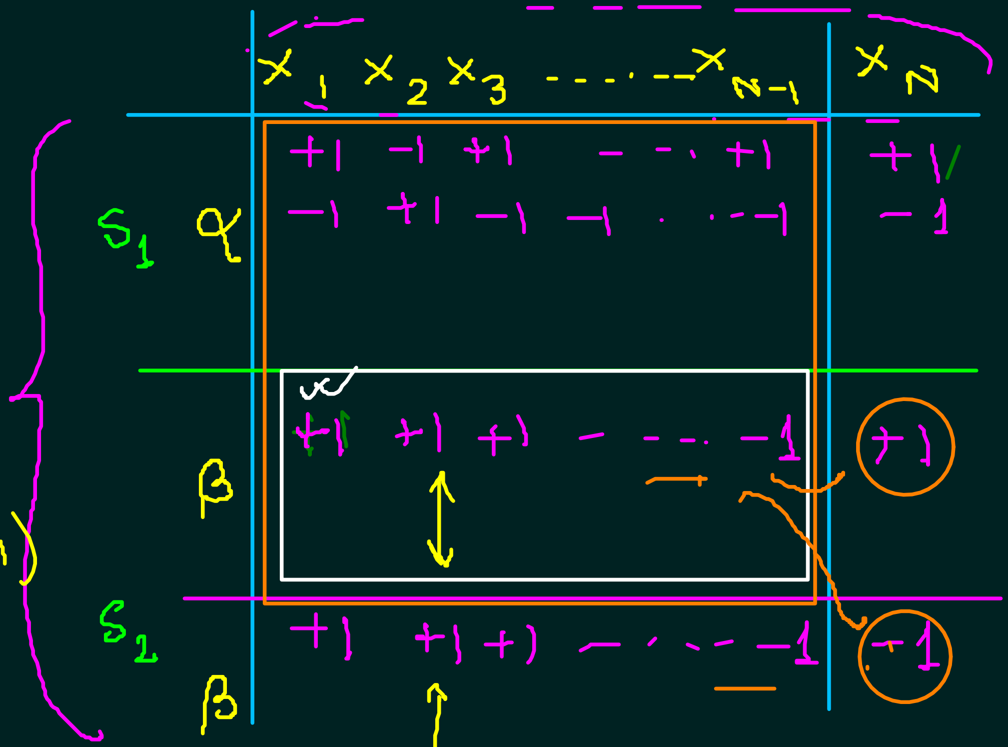
$$\alpha + \beta \le B(N-1,k)$$

$$\beta \le B(N-1, \underline{k-1})$$

(*)

$$\Rightarrow B(N,k) \le B(N-1,k) + B(N-1,k-1)$$

①




| N \ K | 1 | 2 | 3 | 4 | 5 | 6 ... |
|---|---|---|---|---|---|---|
| 1 | 1 | 2 | 2 | 2 | 2 | 2 |
| 2 | 1 | 3 | 4 | 4 | 4 | 4 |
| 3 | 1 | 4 | 7 | 8 | 8 | 8 |
| 4 | 1 | 5 | 11 | 15 | | |
| 5 | 1 | | | | | |

$$B(N,k) \le \boxed{\sum_{i=0}^{k-1} \binom{N}{i}}$$

②

M.I. Induction

Positive Rays: $m_{\mathcal{H}}(N) = N+1 \leq \sum_{i=0}^{2-1} \binom{N}{i} = N+1$ (B.P. = 2)

Positive Interval: $m_{\mathcal{H}}(N) = \frac{1}{2}N^2 + \frac{1}{2}N + 1 \leq \sum_{i=0}^{3-1} \binom{N}{i}$ (BP = 3)

$= \frac{1}{2}N^2 + \frac{1}{2}N + 1$

2D-Perceptron: $m_{\mathcal{H}}(N) \leq \sum_{i=0}^{4-1} \binom{N}{i} = \frac{1}{6}N^3 + \frac{5}{6}N + 1$

(BP = 4)

very simple poly.

$$P\left[\,|E_{in} - E_{out}| > \epsilon\,\right] \leq 2\left(\boxed{m_{\mathcal{H}}(N)}\right) e^{-2\epsilon^2 N}$$

#TE

Prove $\longrightarrow$ Can $m_{\mathcal{H}}(N)$ replace M ?