Assignment 1

Advanced Machine Learning

Chirobocea Mihail-Bogdan

Exercise 1.

a) Finite example

$Let\ A = \{0,1\}^{2024}, and\ B \subset A\ as\ so\ |B| = 2024, C = A \setminus B.$

$Let\ \mathcal{H} = \{h: A \to \{0,1\} |\ h(x) = 0, x \in C\}$

$We\ can\ write\ \mathcal{H}\ as\ \mathcal{H} = \left\{h: B \to \{0,1\}\right\} \cup \{h: C \to \{0,1\}, h(x) = 0\}$

$\left|\left\{h: B \to \{0,1\}\right\}\right| = 2^{2024}$

$|\{h: C \to \{0,1\}, h(x) = 0\}| = 2^{2024} - 2024$

$Note\ that\ \left\{h: B \to \{0,1\}\right\} \cap \{h: C \to \{0,1\}, h(x) = 0\} = \emptyset, because\ \ C = A \setminus B$

$We\ have\ |\mathcal{H}| = 2^{2024} + 2^{2024} - 2024 = 2^{2025} - 2024 <\ 2^{2025}$

$Recall\ that\ VCdim(\mathcal{H}) \leq log_2|\mathcal{H}|\ from\ lecture\ 6,$

$We\ have\ VCdim(\mathcal{H}) \leq log_2|\mathcal{H}| < log_2 2^{2025} = 2025$

$It\ is\ obvious\ that\ |\mathcal{H}_B| = 2^{2024}, as\ \mathcal{H}_B = \left\{h: B \to \{0,1\}\right\} \subset \mathcal{H}\ \Rightarrow\ VCdim(\mathcal{H}) \geq 2024$

$We\ end\ up\ with\ VCdim(\mathcal{H}) = 2024$

b) Infinite example

$$Let\ \mathcal{HS}_0^{2024} = \left\{h_{w,0}: \mathbb{R}^{2024} \to \{-1,1\}, h_{w,0}(x) = sign\left(\sum_{i=1}^{2024} w_i x_i\right) \middle|\ w \in \mathbb{R}^{2024}\right\}$$

$Because\ h\ is\ defined\ on\ \mathbb{R}^{2024}\ we\ have\ that\ \left|\mathcal{HS}_0^{2024}\right| = \infty$

$Recalling\ lecture\ 7\ we\ have\ that\ VCdim(\mathcal{HS}_0^n) = n, so\ VCdim(\mathcal{HS}_0^{2024}) = 2024$

$Another\ example\ is\ the\ following$

$\mathcal{H}_{hyperplane_{2023}} = \left\{h_a: \mathbb{R}^{2023} \to \{0,1\}, h_a(x) = 1_{[a_1x_1 + \cdots + a_{2023}x_{2023} + a_{2024} > 0]} \middle| a \in \mathbb{R}^{2024}\right\}$

$Note\ that\ \ \mathcal{H}_{hyperplane_n}\ is\ a\ generalisation\ of\ \ \mathcal{H}_{lines}\ from\ \mathbb{R}^2\ to\ \mathbb{R}^n$

$Let\ consider\ A\ a\ set\ of\ 2024\ points\ in\ \mathbb{R}^{2023}\ that\ are\ not\ on\ the\ same\ hyperplane,$

$\mathcal{H}$ *shatters* A. *Even more, for any set* B *of* 2025 *points in* $\mathbb{R}^{2023}$ $\mathcal{H}$ *does not shatters* B.

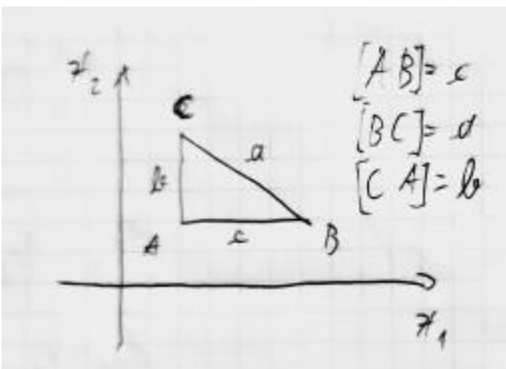
This is because of Cover's Function Counting Theorem.

So we have $VCdim(\mathcal{H}_{hyperplane_{2023}}) = 2024$, *obsiously* $\mathcal{H}_{hyperplane_{2023}}$ *is infinite.*

Exercise 2.

$$\mathcal{H}_\alpha = \left\{ h_{x_{1_A},x_{2_A},x_{1_B},x_{2B},x_{1_C},x_{2_C}}: \mathbb{R}^2 \to \{0,1\}, x_{1_A} = x_{1_B} \leq x_{1_C}, x_{2_A} = x_{2_C} \leq x_{2_B}, \right.$$

$$\left. h_{x_{1_A},x_{2_A},x_{1_B},x_{2B},x_{1_C},x_{2_C}}(x_1,x_2) = \begin{cases} 1, & \begin{matrix} x_{1_A} \leq x_1 \leq x_{1_B} \\ x_{2_A} \leq x_2 \leq x_{2_C} \\ \alpha(x_2 - x_{2_C}) + (x_1 - x_{1_C}) \leq 0 \end{matrix} \\ & 0, otherwise \end{cases} \right\}$$



The equation of line BC is $\dfrac{x_2 - x_{2_C}}{x_{2_b} - x_{2_C}} = \dfrac{x_1 - x_{1_C}}{x_{1_b} - x_{1_C}} \Rightarrow$

$\Rightarrow (x_2 - x_{2_C})(x_{1_B} - x_{1_C}) - (x_1 - x_{1_C})(x_{2_B} - x_{2_C}) = 0$

$\Rightarrow (x_2 - x_{2_C})c - (x_1 - x_{1_C})b = 0$

$\Rightarrow b\left(\alpha(x_2 - x_{2_C}) - (x_1 - x_{1_C})\right) = 0$

$\Rightarrow \alpha(x_2 - x_{2_C}) - (x_1 - x_{1_C}) = 0$

We are under the realizability assumption, so there exist a labeling function $f \in \mathcal{H} = \mathcal{H}_\alpha$,

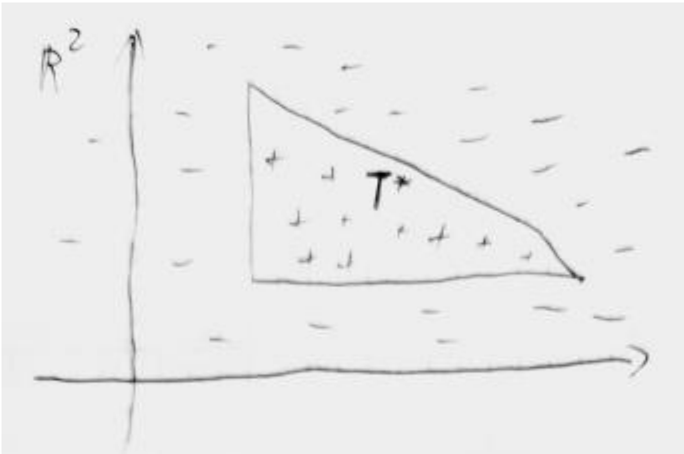
$f = h^*_{x_{1_A},x_{2_A},x_{1_B},x_{2B},x_{1_C},x_{2_C}}$ *that labels the training data*

Consider the training set $S = \left\{(x_1,y_1),\ldots,(x_m,y_m) | y_I = h^*_{x_{1_A},x_{2_A},x_{1_B},x_{2B},x_{1_C},x_{2_C}}(x_i), x_i \in \mathbb{R}^2\right\}$

h^* *labels each point drawn from the triangle* T^* *with label* 1 *and all other points with label* 0

$T^* = \{(x_1, x_2) \in \mathbb{R}^2 | x_{1_A} \le x_1 \le x_{1_B}, x_{2_A} \le x_2 \le x_{2_C}, \alpha(x_2 - x_{2_C}) + (x_1 - x_{1_C}) \le 0\}$

$So\ we\ have\ h^*_{x_{1_A}, x_{2_A}, x_{1_B}, x_{2B}, x_{1_C}, x_{2_C}} = 1_{T^*}$



$Consider\ the\ following\ algorithm\ A, that\ takes\ as\ input\ the\ training\ set\ S\ and\ outputs\ h_s.$

$h_s = h_{x_{1_{A_S}}, x_{2_{A_S}}, x_{1_{B_S}}, x_{2B_S}, x_{1_{C_S}}, x_{2_{C_S}}}\ where$

$$x_{1_{A_S}} = x_{1_{C_S}} = \min_{\substack{i=1,m \\ y_i=1}} x_{i_1}$$

$$x_{2_{A_S}} = x_{2_{B_S}} = \min_{\substack{i=1,m \\ y_i=1}} x_{i_2}$$

$For\ the\ rest\ of\ the\ points\ we\ have\ to\ consider\ the\ equation\ of\ line\ BC.$

$For\ simplicity\ we\ will\ consider\ the\ equation\ of\ the\ parallel\ line\ with\ BC\ that\ passes\ through\ origin$

$We\ have\ the\ equation: \alpha x_1 + x_2 = 0$

$Let\ f: \mathbb{R}^2 \to \mathbb{R}, f(x_1, x_2) = \alpha x_1 + x_2$

$$x_{1_{max}}, x_{2_{max}} = \max_{\substack{i=1,m \\ y_i=1}} f(x_{i_1}, x_{i_2})$$

$Having\ this\ point\ we\ can\ now\ define\ the\ parallel\ line\ with\ BC\ that\ passes\ through\ the\ farest\ point:$

$We\ obtain\ \alpha(x_2 - x_{2_{max}}) + (x_1 - x_{1_{max}}) = 0$

$To\ extract\ C_S\ and\ B_S\ we\ have\ to\ intersect\ this\ line\ with\ the\ lines\ A_SC_S, respectively\ A_SB_S.$

$$\begin{cases} \alpha(x_2 - x_{2_{max}}) + (x_1 - x_{1_{max}}) = 0 \\ x_1 = x_{1_{C_S}} \end{cases} \Rightarrow x_{2_{C_S}} = \frac{x_{1_{max}} - x_{1_{C_S}}}{\alpha} + x_{2_{max}}$$

$$\begin{cases} \alpha(x_2 - x_{2_{max}}) + (x_1 - x_{1_{max}}) = 0 \\ x_2 = x_{2_{B_S}} \end{cases} \Rightarrow x_{1_{B_S}} = \alpha\left(x_{2_{max}} - x_{2_{B_S}}\right) + x_{1_{max}}$$

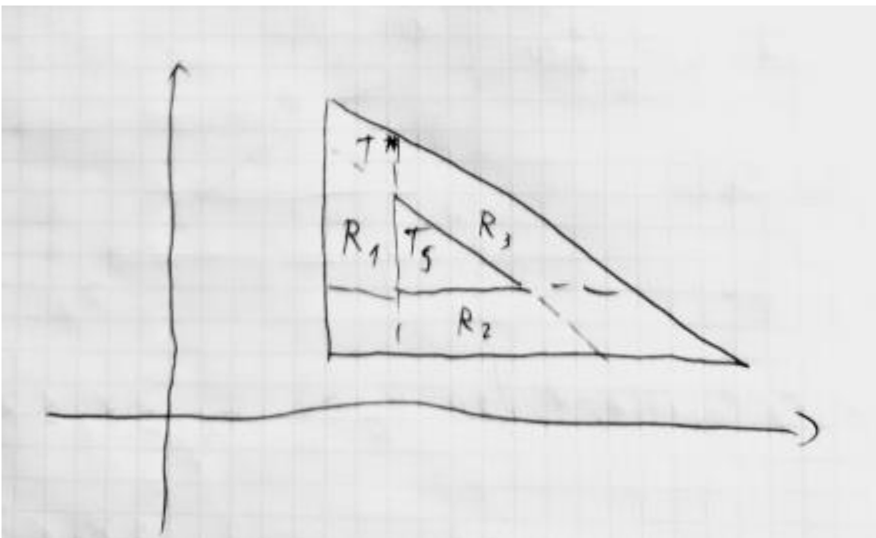
$If\ all\ y_i = 0, then\ all\ points\ x_i\ have\ label\ 0, so\ there\ is\ no\ positive\ example.$

$In\ this\ case, chose\ z_A = (z_1, z_2), z_B = (z_3, z_2), z_C = (z_1, z_4)\ that\ are\ not\ in\ the\ training\ set\ S$

$and\ take\ x_{1_{A_S}} = x_{1_{C_S}} = z_1, x_{2_{A_S}} = x_{2B_S} = z_2, x_{1_{B_S}} = z_3\ , x_{2_{C_S}} = z_4$

$h_{x_{1_{A_S}}, x_{2_{A_S}}, x_{1_{B_S}}, x_{2B_S}, x_{1_{C_S}}, x_{2_{C_S}}}\ is\ the\ indicator\ function\ of\ the\ tighest\ triangle\ encolsing\ all\ positive\ examples.$

$By\ construction, A\ is\ an\ ERM, meaning\ that\ h_S\ dosen't\ make\ any\ trianing\ errors\ on\ training\ set\ S.$



$We\ make\ the\ observation\ that\ h_S\ makes\ errors\ in\ region\ T^*\backslash T_S, assigning\ the\ label\ 0\ to\ points\ that\ should\ get\ lable\ 1. All\ points\ \in T_S\ wiil\ be\ labeled\ correctly\ (label\ 1), all\ points\ outside\ T^*\ will\ be\ labeled\ correctly\ (label\ 0).$

$Let'sfix\ \epsilon > 0, \delta > 0\ and\ consider\ a\ distribution\ D\ over\ \mathbb{R}^2.$

$Case\ 1)$

$If\ D(T^*) = \underset{x\sim D}{P}(x \in \mathbb{R}^2) \leq \epsilon\ then\ in\ this\ case$

$L_{h^*,D}(h_S) = \underset{x\sim D}{P}\big(h_S(x) \neq h^*(x)\big) = \underset{x\sim D}{P}(x \in T^*\backslash T_S) \leq \underset{x\sim D}{P}(x \in T^*) \leq \epsilon\ so\ we\ have\ that$

$\underset{S\sim D^m}{P}\big(L_{h^*,D}(h_S) \leq \epsilon\big) = 1\ (this\ happens\ all\ the\ time)$

$Case\ 2)$

$If\ D(T^*) = \underset{x\sim D}{P}(x \in \mathbb{R}^2) > \epsilon$

$We\ define\ three\ trapezes\ R_1, R_2, R_3\ with\ D(R_i) = \frac{\epsilon}{3}$

$If\ R_1, R_2, R_3\ intersetcs\ T_S\ than:$

$$L_{h^*,D}(h_S) = \underset{x\sim D}{P}(x \in T^* \backslash T_S) \le \underset{x\sim D}{P}(x \in R_1 \cup R_2 \cup R_3) \le \underset{x\sim D}{P}(x \in R_1) + \underset{x\sim D}{P}(x \in R_2) + \underset{x\sim D}{P}(x \in R_3)$$

$$\underset{x\sim D}{P}(x \in R_1) + \underset{x\sim D}{P}(x \in R_2) + \underset{x\sim D}{P}(x \in R_3) = \epsilon \ \Rightarrow\ L_{h^*,D}(h_S) < \epsilon \ \Rightarrow\ \underset{S\sim D^m}{P}\left(L_{h^*,D}(h_S) \le \epsilon\right) = 1$$

$If\ none\ of\ R_1, R_2, R_3 intersects\ with\ T_S\ than:$

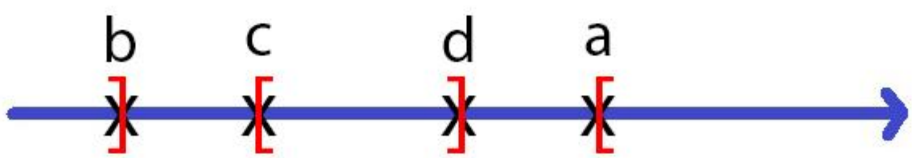
$$\underset{S\sim D^m}{P}\left(L_{h^*,D}(h_S) > \epsilon\right) \le 3\left(1 - \frac{\epsilon}{3}\right)^m \le 3\epsilon^{\frac{\epsilon}{3}} < \delta$$

$So\ if\ we\ take\ m \ge \boldsymbol{m_{\mathcal{H}_\alpha}(\epsilon, \delta) = \frac{3}{\epsilon} log\left(\frac{3}{\delta}\right)}\ we\ obtain\ the\ desired\ result: \mathcal{H}_\alpha\ PAC - learnable.$

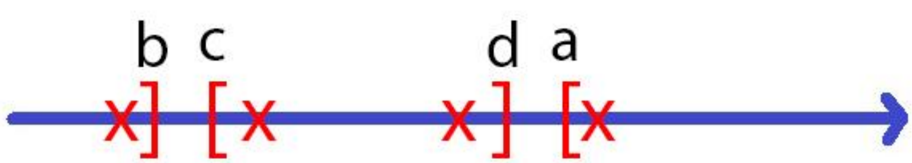
Exercise 3.

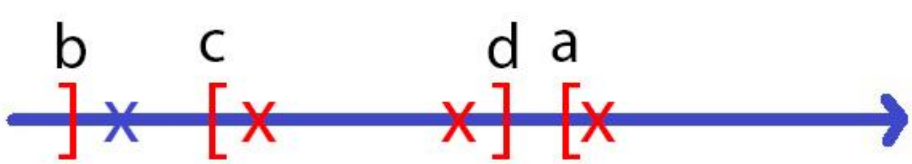
$Let\ C = \{x_1, x_2, x_3, x_4 \in \mathbb{R} | x_1 < x_2 < x_3 < x_4\}$

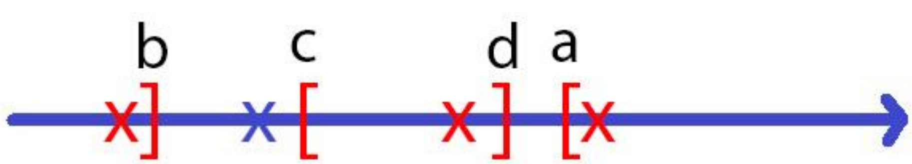
$In\ the\ following\ representation, red\ X\ means\ label\ 1\ and\ blue\ X\ means\ label\ 0.$

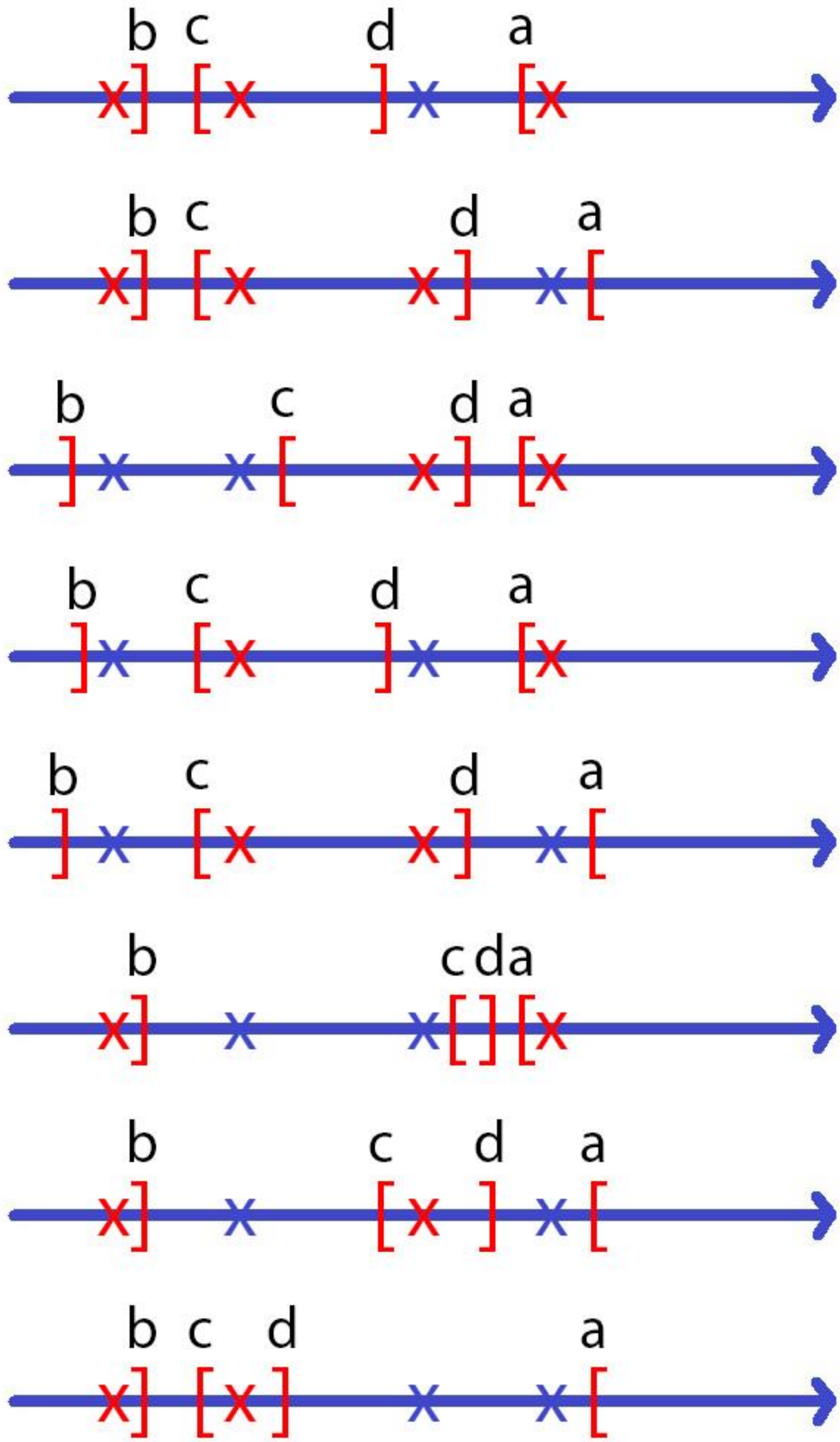


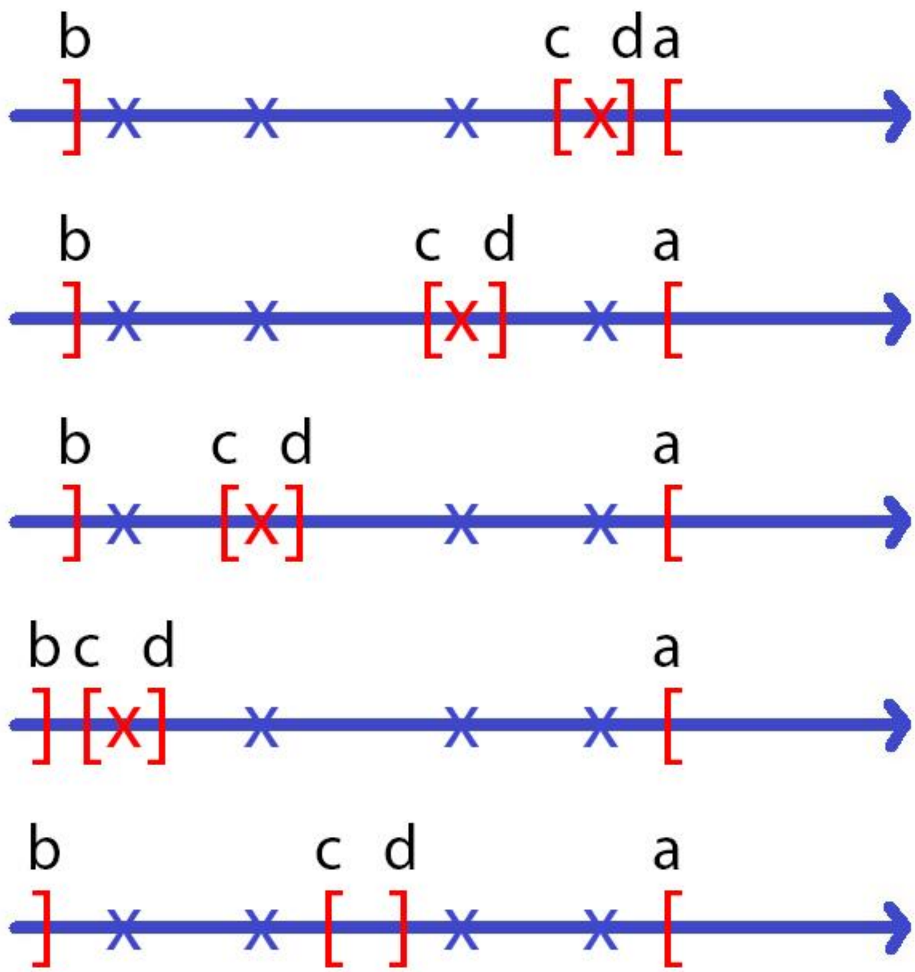
$The\ followings\ are\ the\ 16\ representations\ of\ \mathcal{H}_C\ proving\ that\ |\mathcal{H}_C| = 2^4, where\ |C| = 4$ [1]








b c d a
b c d a
b c d a
b c d a
b c d a
b cda
b c d a
b c d a



$We\ have\ VCdim(\mathcal{H}) \geq 4$ [2]

$Let\ D = \{x_1, x_2, x_3, x_4, x_5 \in \mathbb{R} | x_1 \leq x_2 \leq x_3 \leq x_4 \leq\ x_5\}\ with\ the\ following\ labels$

$x_1 \rightarrow 0, \quad x_2 \rightarrow\ 1, \quad x_3 \rightarrow\ 0, \quad x_4 \rightarrow\ 1, \quad x_5 \rightarrow\ 0$

$If\ b \geq x_1\ than\ x_1\ will\ have\ label\ 1\ \Rightarrow b < x_1$

$If\ c > x_2\ and\ a > x_2\ than\ x_2\ will\ have\ label\ 0\ \Rightarrow either\ c \leq x_2\ or\ a \leq x_2$

$If\ a \leq x_2\ than\ x_3\ will\ be\ 1 \Rightarrow a > x_2\ and\ c \leq x_2$

$If\ d < x_2\ than\ x_2\ will\ be\ 0\ \Rightarrow d \geq x_2$

$If\ d \geq x_3\ than\ x_3\ will\ be\ 1\ \Rightarrow\ d < x_3$

$If\ a > x_4\ than\ x_4\ will\ be\ 0, else\ x_5\ will\ be\ 1.$

$As\ we\ covered\ all\ cases\ and\ we\ still\ ended\ up\ with\ a\ contradiction\ we\ have\ that:$

$!\exists\ D \subset \mathbb{R}\ with\ |D| = 5\ as\ that\ |\mathcal{H}_D| = 2^5$

$[1], [2]\ \Rightarrow\ VCdim(\mathcal{H}) = 4$

Exercise 4.

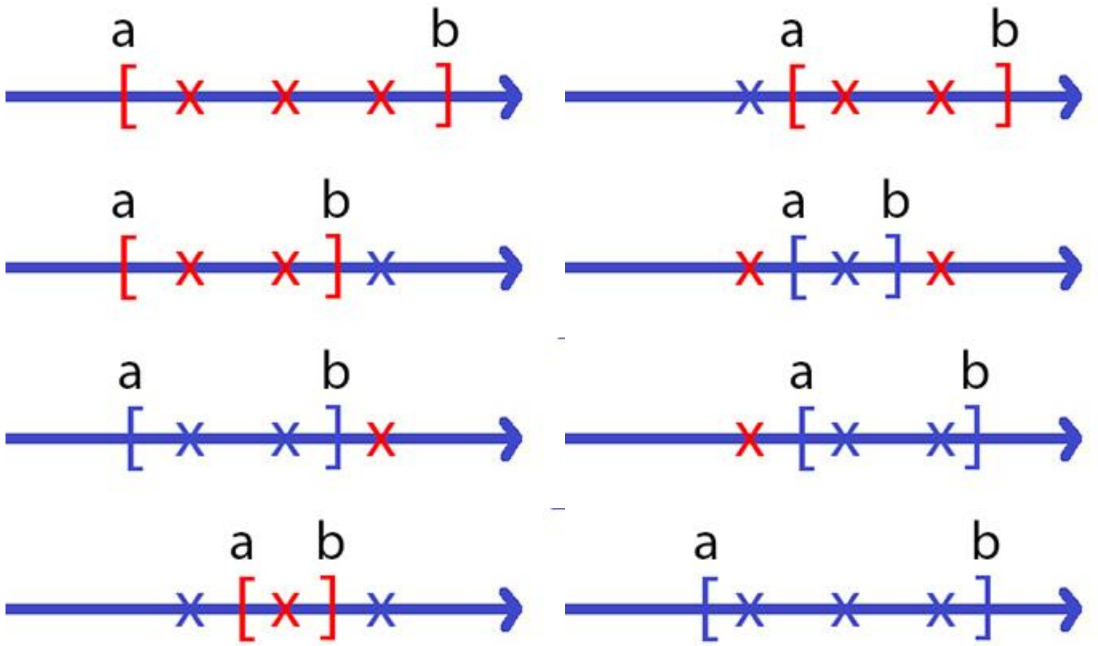
$Let\ C = \{x_1, x_2, x_3, \in \mathbb{R} | x_1 < x_2 < x_3\}$

$In\ the\ following\ representation, red\ X\ means\ label\ 1\ and\ blue\ X\ means\ label - 1.$

$The\ red\ squared\ brackets\ represents\ the\ function\ h_{a,b,1}(x) = \begin{cases} 1, & x \in [a,b] \\ -1, & else \end{cases}$

$The\ blue\ squared\ brackets\ represents\ the\ function\ h_{a,b,-1}(x) = \begin{cases} -1, & x \in [a,b] \\ 1, & else \end{cases}$

$The\ followings\ are\ the\ 8\ representations\ of\ \mathcal{H}_C\ proving\ that\ |\mathcal{H}_C| = 2^3, where\ |C| = 3$ [1]



$Let\ consider\ l_1, \dots, l_n \in \{-1, 1\}\ with\ n \in \mathbb{N}.$

$We\ define\ a\ change\ as\ the\ event\ when\ l_i \neq l_{i+1}, i \in \overline{1, n-1}$

$We\ want\ to\ prove\ that:$

$\forall\ l_1, \dots, l_n \in \{-1, 1\}, n \in \mathbb{N}, n \geq 4\ than\ it\ is\ sufficient\ and\ necesary\ to\ have\ 3\ changes$

$as\ so\ l_1, \dots, l_n\ could\ not\ be\ represented\ using\ h_{a,b,s}\ with\ a, b \in \mathbb{R}, a \leq b, s \in \{-1, 1\}$

$Let\ consider\ x_1, \dots, x_n \in \mathbb{R}\ with\ corresponding\ labels\ l_1, \dots, l_n \in \{-1, 1\}, n \in \mathbb{N}, n \geq 4$

$and\ three\ changes\ on\ l_{k_i} - l_{k_i+1}, l_{k_j} - l_{k_j+1}, l_{k_h} - l_{k_h+1}\ with\ k_i < k_j < k_h$

$We\ can\ consider\ that\ the\ first\ change\ starts\ with\ 1.$

$So\ we\ have\ l_{k_i} = 1, l_{k_i+1} = -1, l_{k_j} = -1, l_{k_j+1} = 1, l_{k_h} = 1, l_{k_h+1} = -1$

$If\ a \leq b \leq x_{k_i}\ \Rightarrow l_{k_h} = l_{k_{h+1}} = s\ wich\ is\ a\ contradiction$

$If\ a \leq x_{k_i}, x_{k_i} < b \leq x_{k_{i+1}}\ \Rightarrow l_{k_h} = l_{k_{h+1}} = s\ wich\ is\ a\ contradiction$

$If\ a \leq x_{k_i}, x_{k_{i+1}} < b\ \Rightarrow l_{k_i}\ = l_{k_{i+1}}\ = s\ wich\ is\ a\ contradiction$

$If\ x_{k_i} < a \leq b \leq x_{k_{i+1}} \Rightarrow l_{k_h} = l_{k_{h+1}} = s\ wich\ is\ a\ contradiction$

$If\ x_{k_i} < a \leq x_{k_{i+1}}, x_{k_{i+1}} < b < x_{k_{j+1}} \Rightarrow l_{k_h} = l_{k_{h+1}} = s\ wich\ is\ a\ contradiction$

$If\ x_{k_i} < a \leq x_{k_{i+1}}, x_{k_{j+1}} \leq b \Rightarrow l_{k_{i+1}} = l_{k_{j+1}} = s\ wich\ is\ a\ contradiction$

$If\ x_{k_{i+1}} < a\ \Rightarrow l_{k_i} = l_{k_{i+1}} = s\ wich\ is\ a\ contradiction$

$As\ we\ cover\ all\ cases\ and\ still\ ended\ up\ with\ a\ contradiction, we\ proved\ that$

$this\ condition\ is\ \boldsymbol{sufficient}$

$To\ prove\ that\ it\ is\ also\ necessary\ we\ need\ to\ check\ the\ case\ with\ only\ less\ changes$

$Let\ consider\ x_1, \dots, x_n \in \mathbb{R}\ with\ corresponding\ labels\ l_1, \dots, l_n \in \{-1, 1\}, n \in \mathbb{N}, n \geq 4$

$and\ two\ changes\ on\ l_{k_i} - l_{k_i+1}, l_{k_j} - l_{k_j+1}\ with\ k_i < k_j$

$We\ can\ consider\ that\ the\ first\ change\ starts\ with\ 1.$

$So\ we\ have\ l_{k_i} = 1, l_{k_i+1} = -1, l_{k_j} = -1, l_{k_j+1} = 1$

$For\ a, b \in \mathbb{R}\ as\ so\ x_{k_i} < a < x_{k_{i+1}}\ and\ x_{k_j} < b < x_{k_{j+1}}\ and\ s = -1\ we\ satisfy\ all\ conditions.$

$Let\ consider\ x_1, \dots, x_n \in \mathbb{R}\ with\ corresponding\ labels\ l_1, \dots, l_n \in \{-1, 1\}, n \in \mathbb{N}, n \geq 4$

$and\ one\ change\ on\ l_{k_i} - l_{k_i+1}$

$We\ can\ consider\ that\ the\ first\ change\ starts\ with\ 1.$

$So\ we\ have\ l_{k_i} = 1, l_{k_i+1} = -1, l_{k_j}$

$For\ a, b \in \mathbb{R}\ as\ so\ a < x_{k_{i1}}\ and\ x_{k_i} < b < x_{k_{i+1}}\ and\ s = 1\ we\ satisfy\ all\ conditions.$

$Let\ consider\ x_1, \dots, x_n \in \mathbb{R}\ with\ corresponding\ labels\ l_1, \dots, l_n \in \{-1, 1\}, n \in \mathbb{N}, n \geq 4$

$and\ no\ change$

$We\ can\ consider\ that\ x_1, \dots, x_n = 1$

$For\ a, b \in \mathbb{R}\ as\ so\ a < x_{k_1}\ and\ x_{k_n} < b\ and\ s = 1\ we\ satisfy\ all\ conditions.$

$All\ the\ above\ demosntration\ work\ symetrically\ when\ assuming\ -1\ instead\ of\ 1$

$We\ proved\ that\ this\ condition\ is\ \boldsymbol{necessary}.$

$Let\ consider\ D \subset \mathbb{R}, with\ |D| = k \geq 4\ than\ |\mathcal{H}_D| = 2^k - 2m,$

$where\ m\ is\ the\ number\ of\ chains\ of\ lenght\ k\ that\ has\ at\ least\ 3\ changes.$

$For\ a\ list\ of\ size\ k\ there\ are\ k-1\ positions\ where\ a\ change\ can\ occur.$

$The\ number\ of\ lists\ of\ size\ k\ with\ exactly\ i\ changes\ is\ C_{k-1}^{i}.$

$This\ is\ because\ we\ chose\ i\ positions\ out\ of\ k-1\ to\ be\ points\ of\ change.$

$Note\ that\ we\ can\ flip\ -1\ with\ 1\ as\ so\ we\ should\ count\ m\ twice.$

$$So\ we\ have\ that\ m = \sum_{i=3}^{k-1} C_{k-1}^{i} \Rightarrow |\mathcal{H}_D| = 2^k - 2\sum_{i=3}^{k-1} C_{k-1}^{i}$$

$$We\ have\ that\ the\ shattering\ coefficient\ \tau_{\mathcal{H}}(m) = \begin{cases} 2^m, & m \leq 3 \\ 2^m - 2\displaystyle\sum_{i=3}^{m-1} C_{m-1}^{i}, & m \geq 4 \end{cases}$$

$We\ are\ intersted\ in\ the\ case\ where\ m \geq 4$

$$Recall\ that\ \sum_{k=0}^{n} C_n^k = 2^n$$

$We\ have\ that$

$$\tau_{\mathcal{H}}(m) = 2\left(2^{m-1} - \sum_{i=3}^{m-1} C_{m-1}^{i}\right) = 2\left(\sum_{i=0}^{m-1} C_{m-1}^{i} - \sum_{i=3}^{m-1} C_{m-1}^{i}\right) = 2\sum_{i=0}^{2} C_{m-1}^{i} =$$

$$= 2\left(1 + (m-1) + \frac{m^2 - 3m + 2}{2}\right) = m^2 - m + 1$$

$$From\ the\ Sauer's\ Lemma\ we\ have\ the\ upper\ bound\ \sum_{i=0}^{3} C_m^i$$

$$\sum_{i=0}^{3} C_m^i = 1 + m + \frac{(m-1)m}{2} + \frac{(m-2)(m-1)m}{6} = \frac{m^3 + 5m + 6}{6}$$

$$We\ want\ to\ prove\ that\ \tau_{\mathcal{H}}(m) < \sum_{i=0}^{3} C_m^i \Leftrightarrow\ m^2 - m + 1 < \frac{m^3 + 5m + 6}{6}$$

$$\Leftrightarrow 0 < m^3 - 6m^2 + 11m - 6$$

$Let\ f: \mathbb{R} \to \mathbb{R}, f(x) = x^3 - 6x^2 + 11x - 6$

$We\ start\ searching\ for\ it's\ solutions\ between\ divisors\ of - 6$

$Note\ that\ f(1) = 0, f(2) = 0\ and\ f(3) = 0\ and\ that\ f\ is\ a\ polinomyal\ of\ degree\ 3$

$Because\ f(4) = 6\ than\ f(x) > 0, \forall x > 3$

$$So\ we\ have\ that\ 0 < m^3 - 6m^2 + 11m - 6\ \Leftrightarrow \tau_{\mathcal{H}}(m) < \sum_{i=0}^{3} C_m^i, m \geq 4$$

Exercise 5.

$$Let\ C = \left\{\left(x, x+1, x+\frac{3}{2}\right) \middle| x \in \mathbb{R}\right\}$$

$$For\ labels\ (1,1,1)\ we\ can\ consider\ \Theta = x - 3\ \Rightarrow\ \Theta + 2 < x < x + 1 < x + \frac{3}{2}$$

$$For\ labels\ (0,1,1)\ we\ can\ consider\ \Theta = x - \frac{3}{2}\ \Rightarrow\ \Theta + 1 < x < \Theta + 2\ and\ \Theta + 2 < x + 1 < x + \frac{3}{2}$$

$$For\ labels\ (1,0,1)\ we\ can\ consider\ \Theta = x - \frac{1}{4}\ \Rightarrow\ x < \Theta\ and\ \Theta < x + 1 < \Theta + 1\ and\ \Theta + 2 < x + \frac{3}{2}$$

$$For\ labels\ (1,1,0)\ we\ can\ consider\ \Theta = x\ \Rightarrow\ x = \Theta\ and\ x + 1 = \Theta + 1\ and\ \Theta + 1 < x + \frac{3}{2} < \Theta + 2$$

$$For\ labels\ (0,0,1)\ we\ can\ consider\ \Theta = x + \frac{3}{2}\ \Rightarrow\ x < x + 1 < \Theta\ and\ \Theta = x + \frac{3}{2}$$

$$For\ labels\ (0,1,0)\ we\ can\ consider\ \Theta = x + \frac{1}{2}\ \Rightarrow\ x < \Theta < x + 1 < \Theta + 1 < \Theta + 2$$

$$For\ labels\ (1,0,0)\ we\ can\ consider\ \Theta = x - \frac{1}{4}\ \Rightarrow\ \Theta < x < \Theta + 1,\ \ \Theta + 1 < x + 1 < x + \frac{3}{2} < \Theta + 2$$

$$For\ labels\ (0,0,0)\ we\ can\ consider\ \Theta = x + 2\ \Rightarrow\ x < x + 1 < x + \frac{3}{2} < \Theta$$

$As\ there\ is\ a\ \Theta \in \mathbb{R}\ that\ satisfy\ ant\ of\ this\ cases\ we\ end\ up\ with\ |\mathcal{H}_C| = 2^3 = 8$ [1]

$Lets\ verify\ if\ there\ is\ possible\ to\ map\ the\ labels\ (1,0,1,0)\ to\ x_1, x_2, x_3, x_4 \in \mathbb{R},$

$with\ x_1 \le x_2 \le x_3 \le x_4$

$If\ x_1 < \Theta\ than\ x_1 will\ have\ label\ 0\ \Rightarrow x_1 \ge \Theta$

$If\ x_1 \ge \Theta + 2\ than\ x_2\ will\ have\ label\ 1 \Rightarrow x_1 < \Theta + 2$

$If\ \Theta + 1 < x_1 < \Theta + 2\ than\ x_1 will\ have\ label\ 0\ \Rightarrow x_1 \le \Theta + 1$

$We\ have\ \Theta \le x_1 \le \Theta + 1$

$If\ x_2 \le \Theta + 1\ than\ x_2\ will\ have\ label\ 1\ \Rightarrow x_2 > \Theta + 1$

$If\ x_2 \ge \Theta + 2\ than\ x_2\ will\ have\ label\ 1\ \Rightarrow x_2 < \Theta + 2$

$We\ have\ \Theta + 1 < x_2 < \Theta + 2$

$If\ x_3 < \Theta + 2\ than\ x_3\ will\ have\ label\ 0\ \Rightarrow x_3 \ge \Theta + 2\ \Rightarrow x_4 \ge \Theta + 2 \Rightarrow x_4\ will\ have\ label\ 1$

$As\ we\ covered\ all\ cases\ and\ still\ ended\ up\ with\ a\ contradiction: x_4\ with\ label\ 1$

$That\ means\ the\ initial\ asumption\ is\ false\ \Rightarrow VCdim(\mathcal{H}) < 4$ [2]

$[1], [2]\ \Rightarrow\ VCdim(\mathcal{H}) = 3$

Exercise 6.

$Let\ L = \{(c_1, b_1), \dots, (c_l, b_l)\}, x \in \mathbb{R}^n, b_i, b \in \{0,1\}\ and\ c_i\ a\ conjunction\ of\ one\ literal\ over\ x_1, \dots, x_n$

$Note\ c_i \in \{x_j, \overline{x_j}\}, i \in \overline{1,k}, j \in \overline{1,n}$

$We\ define\ c: \mathbb{R}^l \to \{0,1\}^n, c(a) = \left(c_1(a_{k_1}), \dots, c_l(a_{k_l})\right), k_i \in \overline{1,n}\ \forall\ i \in \overline{1,l}$

$w \in \mathbb{R}^l\ with\ w_i = 2^{l+1-i} * (b_i * 2 - 1)$

$f_L(x) = 1_{[\langle x,w \rangle + b > 0]}$

$We\ want\ to\ prove\ that\ f_L(c(a)) = L(a)$

$Let\ L(a) = b_j, that\ means\ c_i(a_{k_i}) = 0, \forall\ i \in \overline{1,j}$

$Let's\ note\ h_L(x) = \langle x, w \rangle + b$

$So\ h_L(c(a)) = w_j c_j\left(a_{k_j}\right) + \cdots + w_l c_l(a_{k_l}) + b$

$If\ b_j = 1 \Rightarrow h_L(c(a)) = 2^{l+1-j} + w_{j+1} c_{j+1}\left(a_{k_{j+1}}\right) + \cdots + w_l c_l(a_{k_l}) + b$

$Let's\ note\ S = w_{j+1} c_{j+1}\left(a_{k_{j+1}}\right) + \cdots + w_l c_l(a_{k_l}) + b$

$S\ it's\ at\ least\ S_{min} = -\sum_{i=0}^{l-j} 2^i = -(2^{l+1-j} - 1)$

$\Rightarrow h_L\big(c(a)\big) > 0 \Rightarrow f_L\big(c(a)\big) = 1 = b_j$

$If\ b_j = 0 \Rightarrow h_L\big(c(a)\big) = -2^{l+1-j} + w_{j+1}c_{j+1}\left(a_{k_{j+1}}\right) + \cdots + w_l c_l\big(a_{k_l}\big) + b \ldots < 0$

$S\ it's\ at\ most\ S_{max} = \sum_{i=0}^{l-j} 2^i = 2^{l+1-j} - 1$

$\Rightarrow h_L\big(c(a)\big) < 0\ \Rightarrow f_L\big(c(a)\big) = 0 = b_j$

$We\ proved\ that\ f_L(c(a)) = L(a) = b_j\ for\ an\ arbitrary\ value\ b_j\ \Rightarrow\ f_L(c(a)) = L(a)$

$Now\ we\ can\ identify\ \mathcal{H}_{1-decision\ list}\ with$

$\mathcal{H}S^n = \left\{h_a \colon \mathbb{R}^n \to \{0,1\} \middle| h_a(x) = 1_{[a_1x_1+\cdots+a_nx_n+a_{n+1}>0]}, a \in \mathbb{R}^{n+1}\right\}$

$We\ have\ from\ lecture\ 7\ that\ VCdim(\mathcal{H}S^n) = n+1\ \Rightarrow VCdim(\mathcal{H}_{1-decision\ list}) = n+1$

$So\ there\ is\ \alpha = 1, \beta = 0, \Upsilon = 1, \delta = 2\ as\ so:$

$\alpha n + \beta \leq VCdim(\mathcal{H}_{1-decision\ list}) \leq \Upsilon n + \delta$