Seminar class 4

Exercise 1 (exercise 6.6 in the book)

. Consider the class $\mathcal{H}^d_{mcon}$ of monotone Boolean conjunctions over $\{0,1\}^d$. Monotonicity here means that the conjunctions do not contain negations. As in $\mathcal{H}^d_{con}$, the empty conjunction is interpreted as the all-positive hypothesis. We augment $\mathcal{H}^d_{mcon}$ with the all-negative hypothesis h^-. Show that $\text{VCdim}(\mathcal{H}^d_{mcon}) = d$.

Exercise 2 (exercise 6.9 in the book)

Let $\mathcal{H}$ be the class of signed intervals, that is, $\mathcal{H} = \{h_{a,b,s} : a \leq b, s \in \{-1,1\}\}$ where

$$h_{a,b,s}(x) = \begin{cases} s & \text{if } x \in [a,b] \\ -s & \text{if } x \notin [a,b] \end{cases}$$

Calculate $\text{VCdim}(\mathcal{H})$.

Exercise 3 (exercise 3.13 in the book "Foundations of Machine Learning", 2nd edition)

3.13 VC-dimension of union of k intervals. What is the VC-dimension of subsets of the real line formed by the union of k intervals?

Exercise 4 (exercise 3.15 in the book "Foundations of Machine Learning", 2nd edition)

VC-dimension of subsets. What is the VC-dimension of the set of subsets I_α of the real line parameterized by a single parameter α: $I_\alpha = [\alpha, \alpha+1] \cup [\alpha+2, +\infty)$?

Exercise 5 (exercise 8.1 in the book)

Let $\mathcal{H}$ be the class of intervals on the line (formally equivalent to axis aligned rectangles in dimension $n = 1$). Propose an implementation of the $\text{ERM}_{\mathcal{H}}$ learning rule (in the agnostic case) that given a training set of size m, runs in time $O(m^2)$.
Hint: Use dynamic programming.

19. 04. 2019

Seminar 4

1) $H^d_{mcon} = \{$ all $h : \{0,1\}^d \to \{0,1\}$, $h(x_1 \ldots x_d) = \wedge\, l(x_i)$
$l(x_i) \in \{x_i, 1\}\}$
$\cup \{h^-\}$

$H^d_{con} = \{ h : \{0,1\}^d \to \{0,1\}$, $h(x_1 \ldots x_d) = \wedge\, l(x_i)$
$l(x_i) \in \{x_i, \bar{x}_i, 1\}\}$

sem 3 $\quad$ VCdim$(H^d_{con}) = d$, $|H^d_{con}| = 3^d + 1$

$d = 2 \Rightarrow x_1 \wedge 1 = x_1$
$x_1 \wedge x_2 = x_1 \wedge x_2$
$1 \wedge x_2 = x_2$
$1 \wedge 1 = 1 \quad = h^+$

$|H^d_{mcon}| = 2^d + 1$

$h^-(x_1 \ldots x_d) = 0$

$h^+(x_1 \ldots x_d) = 1$

VCdim$(H^d_{mcon}) = d$

a) VCdim$(H) \leq \lfloor \log_2(|H|) \rfloor$

VCdim$(H^d_{con}) \leq \lfloor \log_2(2^d + 1) \rfloor = d$

b) $H^d_{mcon} \subseteq H^d_{con} \Rightarrow$ VCdim$(H^d_{mcon}) = d$

Trebuie să găsim o mulț. A, $|A| = d$ care este saturată de H^d_{mcon}.

$$\left| H^d_{mcon\,|A} \right| = 2^{|A|}$$

În H^d_{mcon} $\exists$ $h_1 \dots h_{2^d}$ care generează toate etichetările posibile pe $\{a_1 \dots a_d\} = A$, $a_i \in \{0,1\}^d$

$h_1(a_1) \dots h_1(a_d) = (0, \dots 0)$

$h_2(a_1) \dots h_2(a_d) = (1 \dots 0)$

$\vdots$

$h_{2^d}(a_1) \dots h_{2^d}(a_d) = (1 \dots 1)$

$\Leftrightarrow$ $\forall B \subseteq A$ $\exists$ h_B a.î. $h_B(a) = 1$, $a \in B$

$h_B(a) = 0$, $a \notin B$

Pt H^d_{con} luam A:

$A = \{ e_i \mid i = 1 \dots d \}$

$$A = \{ (1 \dots 0), (0, 1 \dots 0), \vdots, (0 \dots 1) \}$$

Scanned with CamScanner

$\forall B \subseteq A \ \exists\, h_B \ , \ h_B(e) = \begin{cases} 1, & e \in B \\ 0, & e \notin B \end{cases}$

$$h_B = \bigwedge_{e_i \notin B} \bar{x}_i$$

$B = \{e_1, e_2, e_3\}$

$h_B = \wedge \bar{x}_4 \wedge \bar{x}_5 -$

$h_B(e_1) = h_B(1 \ldots 0) = 1$

Nu pot avea $h_B = \wedge \bar{x}_i$ deoarece nu aparține lui H^d_{mcon}.

$h_B = \bigwedge_{e_i \in B} x_i$ - nu merge pt $B = \{e_1, e_2\}$

$A = \{ 1 - e_i \mid i \in \{1 \ldots d\} \} = \{ (0\ 1 \ldots 1)$
$(1\ 0\ 1 \quad 1)$
$\vdots$
$(1\ 1\ 1\ .\ 0) \}$

$\begin{array}{c} 1 \\ 1 \\ 1 \\ \vdots \\ 0 \end{array}$ pt $B = 3$

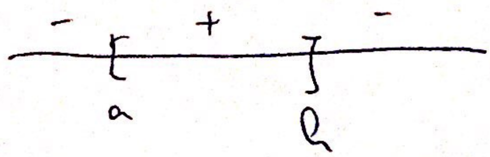
$$\Rightarrow h_B = \bigwedge_{1-e_i \notin B} x_i$$

$B = \{1-e_1, 1-e_2, 1-e_3\}$

$h_B = x_4 \wedge x_5 - \wedge x_d$

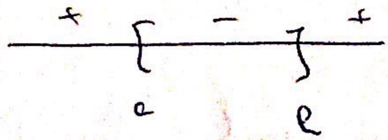
2) $H^{sign}_{int} = \{ h_{a,b,s} : a \le b , s \in \{-1, 1\} \}$

$$h_{a,b,s}(x) = \begin{cases} s, & x \in [a,b] \\ -s, & x \notin [a,b] \end{cases}$$

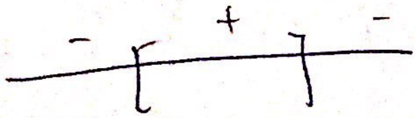


$s = 1$

$s = -1$



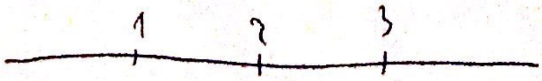
$H_{int} = \{ h_{a,b} : a \le b \}$, $h_{a,b} = \begin{cases} 1, & x \in [a,b] \\ 0, & x \notin [a,b] \end{cases}$



$VCdim(H_{int}) = 2$

$VCdim(H^{sign}_{int}) \ge 3$

$A = \{1, 2, 3\}$



+	+	+	$h_{1,3,1}$
+	+	−	$h_{1,2,1}$
+	−	+	$h_{2,2,-1}$
−	+	+	$h_{2,3,1}$
+	−	−	$h_{1,1,1}$

VC dim (H sign int) < 4

$\forall A = \{a_1, a_2, a_3 - a_4\}$

$|H_{\text{sign int}, A}| \neq 2^{|A|} = 2^4 = 16$

Fie $A = \{a_1, a_2, a_3, a_4\}$

$a_1 \leq a_2 \leq a_3 \leq a_4$

$(+1, -1, +1, -1)$ nu poate fi generată de o fc. $h \in H_{\text{sign int}}$

Pp că $\exists\, h_{a,b,s}$ care gen. $(+1, -1, +1, -1)$

$$\left.\begin{array}{l} h_{a,b,s}(a_1) = 1 \\ h_{a,b,s}(a_3) = 1 \end{array}\right\} \Rightarrow h_{a,b,s}(a_2) = 1 \text{ - contradicție}$$

$h_{a,b,s}(a_2) = -1 \Rightarrow h_{a,b,s}(a_4) = 1$ - contradicție.

3) $H_{k\text{-int}} = \{h_{a_1, b_1, a_2, b_2 - a_k, b_k} : \mathbb{R} \to \{0,1\}\}$

$$h_{a_1, b_1 - a_k, b_k}(x) = \begin{cases} 1, & \exists\, j \text{ a.î. } x \in [a_j, b_j] \\ 0, & \text{altfel} \end{cases}$$

$- [\ + \ ] - [\ + \ ] - [\ + \ ] -$

$k = 1 \Rightarrow H_{int} \rightarrow VCdim() = 2$

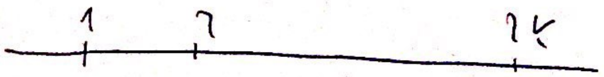
$k = 2 \rightarrow H_{2-int} \rightarrow VCdim() = 4$

$VCdim(H_{k-int}) = 2k$

$A = \{1, 2, 3, 4 - 2k\}$

$\forall B \subseteq A \ \exists h_B$ a.î. $h_B(c) = \begin{cases} 1, & c \in B \\ 0, & c \notin B \end{cases}$

Numărăm câte schimbări de semn avem



\+ + + 0

[+ + +] - [+ +] 2

[+ +] - - [+ +] - [+ +]

[+] - [+] - [+] - [+] - [+] - [+]
\- [+] - [+] - [+] - [+] - 2k - 1

Cu k intervale sunt să generez etichete cu 2k - 1 schimbări de semn care este nr. maxim.

$\Rightarrow VCdim(H_{k-int}) \geq 2k.$

$A = \{a_1 \ldots a_{2k+1}\}$

[+] - [+]   [+]   k+1 intervale.

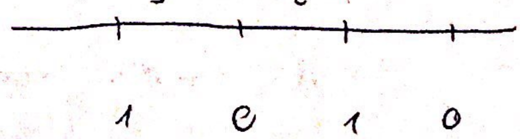
4) $i_\alpha = [\alpha, \alpha+1] \cup [\alpha+2, +\infty)$

$$H = \left\{ h_\alpha : \mathbb{R} \to \{0,1\},\ h_\alpha(x) = \begin{cases} 1, & x \in [\alpha, \alpha+1] \cup [\alpha+2, +\infty) \\ 0, & \text{altfel} \end{cases},\ \alpha \in \mathbb{R} \right\}.$$

VCdim(H) = ?

[+	+]	$\alpha = 0$
+	-	$\alpha = -0.5$
-	-	$\alpha = 20$
-	+	$\alpha = 1$

[+] - [+ -



$\Rightarrow 2 \le \text{VCdim}(H) \le 3$

$A = \{a_1 \le a_2 \le a_3\}$

$+ + - \Rightarrow d_1 \le a_1 \le a_2 \le d+1$

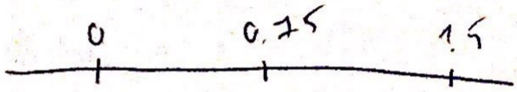
$+ - + \Rightarrow d_2 \le a_1 \le d_2 + 1$

$d_2 + 1 < a_2 < d_2 + 2$

$d_2 + 2 \le a_3$

$- + +$

$A = \{0, 0.75, 1.5\}$



$+ + + \Rightarrow d_1 = -100$

$+ + - \Rightarrow d_2 = 0$

$+ - + \Rightarrow d_3 = -0.5$ sau -0.4

$- + + \Rightarrow d_4 = 0.70$

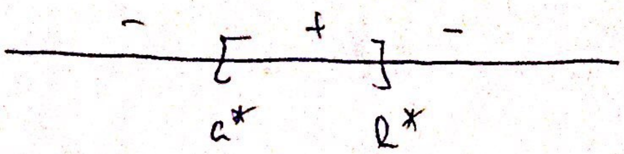
$+ - - \Rightarrow d_5 = -0.3$

$- + - \Rightarrow d_6 = 0.20$

$- - + \Rightarrow d_7 = 1.4$

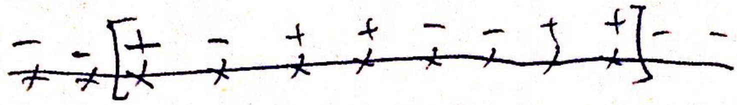
$- - - \Rightarrow d_8 = 100$

5)



Hint

$$h_{a^*,b^*} \text{ cu } \min_{h \in H} L_S(h)$$



agnostic case:

- distribuție pe $\mathbb{R} \times \{0,1\}$. $\nexists$ fc. de etichetare
- $\exists$ fc de etichetare, dar $f \notin H$

$$S = \{ (x_i, y_i),\ x_i \in \mathbb{R},\ y_i \in \{0,1\} \}$$

O implementare ~~...~~ mașină a ERM_H este:

- pas 1: construim toate intervalele posibile cu pct x_i din S.

  $\frac{n(n+1)}{2}$ intervale.

- pas 2: pt. fiecare interval în parte calculez $loss([x_i, x_j]) =$

  $$= \frac{\#\ +\ \text{din afară} + \#\ -\ \text{dinăuntru}}{n}$$

- output $[x_i^*, x_j^*]$ cu ERM min

$O(n^3)$.

$$l_{\theta} = \frac{\#\, \text{elem. cu } +}{n}$$, dam la tat –

$$\begin{array}{ccc} x_1 & & x_n \\ \begin{bmatrix} l_{11} & & & \\ & l_{22} & & \\ & & \ddots & \\ & & & l_{nn} \end{bmatrix} \end{array}$$

$$l_{ii} = \begin{cases} l_{\theta} - \frac{1}{n} \\ l_{\theta} + \frac{1}{n} \end{cases}$$

$$l_{ij+1} = \begin{cases} l_{ij} - \frac{1}{n}, & x_{j+1} = 1 \\ \quad + \frac{1}{n}, & \end{cases}$$

$\Rightarrow O(n^2)$.