Algorithmic Learning Theory - Exercise sheet 4

Christopher Brix 343982 | Arne Nix 331423 | Sascha Welten 333544

May 9, 2018

Exercise 1

Let $\mathcal{X} = \{0,1\}^n$, I $\subseteq \{1,2,...,n\}$, and $h_I = (\sum_{i\in I})$ mod 2.

We show VCdim($\mathcal{H}_{n-parity}$) $\geq$ n.
We define the set S_n as the set of all n unit vectors $e_j = (e_{j,1},...,e_{j,n})$. If e_j was labeled positively, we add i to the set I if $e_{j,i} = 1$. In the end, we obtain a set of indices in combination with h_I which fulfills the labeling. With this setup, we can construct all necessary hypotheses. So, $\mathcal{H}_{n-parity}$ shatters a set of at least size n.
We show VCdim($\mathcal{H}_{n-parity}$) $\leq$ n.
We know that it holds: VCdim($\mathcal{H}_{n-parity}$) $\leq log_2(|\mathcal{H}_{n-parity}|)$.
Since we only have 2^n parity functions (2^n subsets of $\{1,2,...,n\}$) it holds:
VCdim($\mathcal{H}_{n-parity}$) $\leq log_2(2^n) = n$

Exercise 2

a.

Chose $k \geq d \geq 3$ such that there is a set S with size k that can be shattered by $\bigcup_{i=1}^r \mathcal{H}_i$. This k exists, because $\max_i VCdim(\mathcal{H}_i) \geq 3$, so $VCdim(\bigcup_{i=1}^r \mathcal{H}_i) \geq 3$ as well.

By Sauer's lemma, we know (assuming $d \geq m$):

$$\begin{aligned}
\tau_{\bigcup_{i=1}^r \mathcal{H}_i}(m) &\leq \sum_{i=1}^r \tau_{H_i}(m) \\
&\overset{(1)}{\leq} r \cdot \tau_{H_j}(m) \\
&\overset{(2)}{\leq} r \cdot \sum_{i=0}^d \binom{m}{i} \\
&\overset{(3)}{\leq} r \cdot \sum_{i=0}^d \binom{d}{i} \\
&= r \cdot 2^d \\
&\overset{(4)}{\leq} r \cdot k^d
\end{aligned}$$

(1) Chose $j = \text{argmax}_{j\in[r]} \left\{\tau_{\mathcal{H}_j}(m)\right\}$
(2) $VCdim(\bigcup_{i=1}^r \mathcal{H}_i) < d \Rightarrow VCdim(H_j) < d$, apply Sauer's lemma
(3) $d \geq m$ (Assumption!)
(4) $k \geq 3$

Because S can be shattered, $\tau_{\bigcup_{i=1}^r \mathcal{H}_i}(k) \geq 2^k$, so we know that $2^k \leq rk^d$

b.

Proof that $VCdim(\mathcal{H}_1 \cup \mathcal{H}_2) \leq 2d + 1$:

- Let $VCdim(\mathcal{H}_1 \cup \mathcal{H}_2) \geq 2d + 2$ with $VCdim(\mathcal{H}_1) = d_1$ and $VCdim(\mathcal{H}_2) = d_2$ and $d = max\{d_1, d_2\}$

- W.l.o.g. let $\mathcal{H}_1$ shatter $X = \{x_1, \dots, x_{d_1}\}$ and $\mathcal{H}_2$ shatter $Y = \{y_1, \dots, y_{d_2}\}$
- Let $Z = X \bigcup Y \bigcup \{z_1, z_2\} \subseteq \mathcal{X}$ be a set that is shattered by $\mathcal{H}_u := \mathcal{H}_1 \cup \mathcal{H}_2$
- Let h_I be the hypothesis that accounts for $I \subseteq Z$
- Any function $h_{I'}$ for I' with $I' \backslash Z \neq \emptyset$, we can find an equivalent function $h_{I' \cap Z}$ which accounts for the same subsets of Z
- $\mathcal{H}_1$ must contain h_{I_x} for all $I_x \subseteq X$
- $\mathcal{H}_2$ must contain h_{I_y} for all $I_y \subseteq Y$
- The h_{I_z} for all $I_z \in (\mathcal{P}(Z) \backslash \mathcal{P}(X)) \backslash \mathcal{P}(Y)$ are distributed among $\mathcal{H}_1$ and $\mathcal{H}_2$
- Claim: one of these holds: $\mathcal{H}_1$ shatters $X \cup \{z_1\}$ or $X \cup \{z_2\}$ or $\mathcal{H}_2$ shatters $Y \cup \{z_1\}$ or $Y \cup \{z_2\}$
  Proof:
  - For simplicity (w.l.o.g) we assume that $\mathcal{H}_1$ does not shatter $X \cup \{z_1\}$
  - Then there is one $I = \{x_{i_1}, \dots, x_{i_n}, z_1\} \subseteq X \cup \{z_1\}$
    s.t. $h_J \notin \mathcal{H}_1$ for $J = I \cup I'$ for all $I' \in \mathcal{P}(Y \cup \{z_2\})$
    i.e. exclude all hypotheses that could somehow account for I
  - Otherwise any of the h_J could account for I
  - Then $h_J \in \mathcal{H}_2$ for all J.
  - This implies that $\mathcal{H}_2$ shatters $Y \cup \{z_2\}$ since the h_J's can account for all $I' \in \mathcal{P}(Y \cup \{z_2\})$
- Consequently we have $VCdim(\mathcal{H}_1) > d_1$ or $VCdim(\mathcal{H}_2) > d_2$
  $\Rightarrow$ Contradiction

Exercise 3

a.

In addition to the all-negative hypothesis, 3^d hypotheses are possible by either choosing for every literal either the positive or negative interpretation, or omitting it from the hypothesis. Hypthoeses, where one literal appears multiple times in the same form are logically equivalent to any hypothesis where this literal appears only once.
$\Rightarrow$ $|\mathcal{H}^d_{con}| \leq 3^d + 1$

We now want to show that $VCdim(\mathcal{H}^d_{con}) \leq d \cdot \log_2 3$
Proof by contradiction: Assume $VCdim(\mathcal{H}^d_{con}) \geq d \cdot \log_2 3 + 1$
$\Rightarrow \mathcal{H}^d_{con}$ shatters X with $|X| = \lceil d \cdot log_2 3 + 1 \rceil$
$\Rightarrow |\mathcal{H}^d_{con}| \geq 2^{|X|} \geq 2^{d \cdot log_2 3 + 1} = 2^{log_2 3^d} \cdot 2 = 3^d \cdot 2 > 3^d + 1 \geq |\mathcal{H}^d_{con}|$
$\Rightarrow$ contradiction

b.

Let $S_d = \{e_i | i \leq d\}$ with unit vector $e_i = (e_{i,1}, ..., e_{i,j}, ..., e_{i,d})$ and $e_{i,j} = 1$. So, $|S_d| = d$. We construct a family of $h_i \in \mathcal{H}^d_{con}$ which shatters the set S_d. The base structure of h_i is the conjunction where all literals are negated. According to the labeling, we omit a negative literal x_j if e_i was labeled positively (Omit the literal where the unit vector has $e_{i,j} = 1$). Consequently, we obtain 2^d different conjunctions/hypotheses and this implies that $\mathcal{H}^d_{con}$ shatters S_d.

c.

Assume $VCdim(\mathcal{H}_{con}^d) \geq d+1)$. Then there is a set $C = (c_1, \ldots, c_{d+1})$ of boolean functions with size $d+1$ and $c_i \neq c_j \forall i,j \in [d+1] \wedge i \neq j$ that can be shattered by $\mathcal{H}_{con}^d$.

Set $C_i = C \setminus \{c_i\}$ for all $i \in [d+1]$. Because C can be shattered, for every C_i, there is a hypothesis $h_i \in \mathcal{H}_{con}^d$ as defined in the hint (it accepts exactly the $c_j \in C_i$). l_i is defined as in the hint as well, i.e. it's a literal that's false on c_i and true on all other c_j with $j \neq i$.

Because there are only d variables, but $d+1$ such literals, at least one variable has to be used at least twice. Let l_a, l_b be two literals that both use the same variable.

Case 1: Both literals are positive or both are negative. Then both literals are identical because they use the same variable. Then $h_a = h_b$ and therefore $S_a = S_b$. Hence, $c_a = c_b$. This is a contradiction, because C was chosen to contain pairwise different elements.

Case 2: One literal is positive, the other negative. W.l.o.g., let l_a be the positive, l_b the negative one. From h_a we know that l_a is true for all c_i with $i \neq a$. Because l_b is the negation of l_a, l_b is false for all c_i with $i \neq b$. But this implies that all c_i with $i \neq b$ are not in h_b. This is a contradiction to the choice of h_b.

Therefore, no such set C can exist. Because not set of size $d+1$ can be shattered, not set with size $\geq d+1$ can be shattered, so $VCdim(\mathcal{H}_{con}^d) \leq d$.

d.

$VCdim(\mathcal{H}_{mcon}^d) \leq d$: The all positive hypothesis is contained in $\mathcal{H}_{con}^d$. The all negative hypothesis can be generated by the conjunction $x_1 \wedge \neg x_1 \in \mathcal{H}_{con}^d$. All other hypotheses in $\mathcal{H}_{mcon}^d$ are conjunctions, and as such part of $\mathcal{H}_{con}^d$. Therefore, $\mathcal{H}_{mcon}^d \subseteq \mathcal{H}_{con}^d$. By reducing the set of available hypotheses, the VC dimension can not be increased. If it is possible to shatter a set of size m using hypotheses in $\mathcal{H}_{mcon}^d$, this set could be shattered by $\mathcal{H}_{con}^d$ as well. In c) we proved that $VCdim(\mathcal{H}_{con}^d) \leq d$, so $VCdim(\mathcal{H}_{mcon}^d) \leq d$ as well.

$VCdim(\mathcal{H}_{mcon}^d) \geq d$: Chose $S = \{s_1, \ldots, s_d\}$ with $s_{ij} = 0 \Leftrightarrow i = j$ (ie. element s_i has 1s for every variable except x_j, there it has a 0).

- The subset $S' = \emptyset \subseteq S$ can be generated using the all negative hypothesis.
- The subset $S' = S$ can be generated using the all positive hypothesis.
- For every other subset $S' \subseteq S$, let $I = \{i \in [d] | s_i \in S'\}$. To find a hypothesis h that only accepts elements in S', one can build the conjunction of all variables x_j with $j \notin I$. Due to the choice of S and h, every element not in S' has a negative value for one of the variables used in h. Every element of S' has only positive values for the variables in h. Therefore, h accepts exactly those elements that are in S'. Because h can be build for every $S' \subseteq S$, S can be shattered. Hence, $VCdim(\mathcal{H}_{mcon}^d) \geq d$.

By combining both proofs, it is shown that $VCdim(\mathcal{H}_{mcon}^d) = d$.

Exercise 4

Let $h : \{0,1\}^n \to \{0,1\}$ be a symmetric function and $\mathcal{H}$ the set of all symmetric functions.

First, we show that the VCdim($\mathcal{H}$) $\geq$ n+1.
Therefore, we take a set $S \subset \{0,1\}^n$ of maximal size such that there are no $s_1, s_2 \in S$ with the same number of 1's, i.e. $\{(0,0,...,0),(1,0,...,0),(1,1,0,...,0),...,(1,1,...,1)\}$. Consequently, $|S| = n+1$. Now we show that we can shatter the set S. We define the family of symmetric functions $h_{1,...,i}$ with $h_i(x) = 1$ iff x has exactly i ones (e.g. $h_1((1,0,0)) = 1$, $h_2((1,0,0)) = 0$). Since the value of the function h_i depends on the number of ones, it is symmetric. With this set of functions, we are able to shatter a set of size n+1 by combining h_i and $h_{i'}$ (i $\neq$ i') to a single function $h_{i,i'}$ which is also symmetric. Meaning that

$$h_{i_1,\ldots,i_n}(x) = \begin{cases} 1 \text{ if number(ones)} = i_1 \\ 1 \text{ if number(ones)} = i_2 \\ \ldots \\ 1 \text{ if number(ones)} = i_n \\ 0 \text{ else} \end{cases}$$

With these symmetric functions, we are able to shatter S by setting the i's according to the labeling.
$\implies$ VCdim($\mathcal{H}$) $\geq$ n+1.

Now, we show that VCdim($\mathcal{H}$) $\leq$ n+1:
Assume there is a set $|S'| = (n+1)+1$ which can be shattered by $\mathcal{H}$. Then, there are $x_1, x_2 \in S'$ with the same number of 1's due to the finite instance space and the pigeonhole principle. This means that for any $h \in \mathcal{H}$ it always holds $h(x_1) = h(x_2)$ due to the definition of symmetric functions. Consequently, there is no h^* with $h^*(x_1) \neq h^*(x_2)$. In terms of labeling, there is no different labeling of x_1 and x_2 possible.
$\implies$ specific labelings are not possible $\implies$ no shattering of S' $\implies$ VCdim($\mathcal{H}$) $\leq$ n+1