MLT Homework 3

Ana Borovac
Jonas Haslbeck
Bas Haver

October 1, 2018

Question 1

Show the following monotonicity property of VC-dimension: For every two hypothesis classes if $\mathcal{H}' \subset \mathcal{H}$ then $VCdim(\mathcal{H}') \leq VCdim(\mathcal{H})$.

Solution

Let assume:

$$C' \subseteq \mathcal{H}' : \text{ VCdim}(\mathcal{H}') = |C'|$$

We also know:

$$C' \subseteq \mathcal{H}' \subset \mathcal{H} \quad \Longrightarrow \quad C' \subset \mathcal{H}$$

From that we can conclude:

$$\text{VCdim}(\mathcal{H}) \geq |C'| = \text{VCdim}(\mathcal{H}') \quad \Longrightarrow \quad \text{VCdim}(\mathcal{H}') \leq \text{VCdim}(\mathcal{H})$$

Question 2

Given some finite domain set, $\mathcal{X}$, and a number $k \leq |\mathcal{X}|$, figure out the VC-dimension of each of the following classes (and prove your claims).

Subquestion 2.1

$\mathcal{H}^{\mathcal{X}}_{=k} = \{h \in \{0,1\}^{\mathcal{X}} : |\{x : h(x) = 1\}| = k\}$. *That is, the set of all functions that assign the value 1 to exactly k elements of $\mathcal{X}$.*

Solution

Claim: $\text{VCdim}(\mathcal{H}^{\mathcal{X}}_{=k}) = \min\{k, |\mathcal{X}| - k\}$.
Proof: Let C denote a subset of $\mathcal{X}$ such that $|C| = \min\{k, |\mathcal{X}| - k\}$.

In the "worst" case we want to label all elements of C with 0. In order to do that we need at least k elements in a set $\mathcal{X} \setminus C$. From there we can conclude that the biggest subset of $\mathcal{X}$ which can be shattered by $\mathcal{H}^{\mathcal{X}}_{=k}$ is of size $|\mathcal{X}| - k$.

On the other hand if we want to label all elements of C with 1, the size of C can not be bigger than k.

If we combine both explanations above, we get:

$$\text{VCdim}(\mathcal{H}^{\mathcal{X}}_{=k}) = \min\{k, |\mathcal{X}| - k\}$$

Subquestion 2.2

$\mathcal{H}_{at-most-k} = \{h \in \{0,1\}^{\mathcal{X}} : |\{x : h(x) = 1\}| \le k \text{ } or \text{ } |\{x : h(x) = 0\}| \le k\}$.

Solution

Claim: $\text{VCdim}(\mathcal{H}_{at-most-k}) = \min\{2k, |\mathcal{X}|\}$
Proof: Let's analyse the case that we do not want to happen. We do not want to have a subset C where one of labelings contains $> k$ elements that are labeled 0 and $> k$ elements labeled 1. So, subsets bigger that $2k$ can not be shattered by $\mathcal{H}_{at-most-k}$. From this and the fact that VC-dimension can not be bigger than $|\mathcal{X}|$ we conclude:

$$\text{VCdim}(\mathcal{H}_{at-most-k}) = \min\{2k, |\mathcal{X}|\}$$

Question 3

Let $\mathcal{X}$ be the Boolean hypercube $\{0,1\}^n$. For a set $\mathcal{I} \subseteq \{1, 2, \ldots, n\}$ we define a parity function h_I as follows. On a binary vector $\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in \{0,1\}^n$,

$$h_I(\boldsymbol{x}) = (\sum_{i \in I} x_i) mod\ 2.$$

(That is, h_I computes parity of bits in I.) What is the VC-dimension of the class of all such parity functions, $\mathcal{H}_{n-parity} = \{h_I : I \subseteq \{1, 2, \ldots, n\}\}$?

Solution

By both proving that the VC-dimension of the parity functions is both bounded from above and from below by n, we will get to the conclusion that the VC-dimension is equal to n.

First note that the Boolean hypercube is a finite set and therefore, as is stated in section 6.3.4 in the book,

$$\text{VC-dim}(\mathcal{H}_{n-parity}) \leq \log_2(|\mathcal{H}_{n-parity}|) = \log_2(2^n) = n.$$

So we now have $\text{VCdim}(\mathcal{H}_{n-parity}) \leq n$.
We also know that in $\{0,1\}^n$ there are n different unit vectors e_i which have a one at the i-th spot and a zero everywhere else. For any collection of these unit vectors, excluding the one containing all the unit vectors, we can choose I to be corresponding with one of the unit vectors (for example $I = 1$ when the unit vector e_1 is in the collection), to obtain 1 and choose I corresponding to a unit vector which is not in the collection to obtain 0. Now when all unit vectors are in our collection, we can choose an $I = \{1\}$ to obtain a 1 and $I = \{1, 2\}$ to obtain a zero. So for every collection of unit vectors we can obtain bot a 0 and a 1. Since there are n of these unit vectors we have $\text{VC-dim}(\mathcal{H}_{n-parity}) \geq n$ and so $\text{VC-dim}(\mathcal{H}_{n-parity}) = n$

Question 4

VC-dimension of Boolean conjunctions: *Let $\mathcal{H}^d_{con}$ be the class of Boolean conjunctions over the variables $x_1, \ldots, x_d$ ($d \geq 2$). We already know that this class is finite and thus (agnostic) PAC learnable. In this question we calculate VCdim($\mathcal{H}^d_{con}$).*

1. Show that $|\mathcal{H}^d_{con}| \leq 3^d + 1$.
   Since having a literal twice in a product give's 1 for every x_i, it does not contribute anything when multiplying this with other literals. So first consider the possibilities where there is no literal used twice. Then for any $i \in [d]$ we have the possibilities of using a literal giving x_i, one giving $\bar{x}_i$, and of course the possibility of not using any literal of x_i in your product. Having these three options for every $i \in [d]$ results in a maximal amount of 3^d functions. Combined with the possibility of having a literal twice only, so not multiplied by any other product (of course multiplying it with twice any literal can be done, still not contributing anything to our final value), will result in a label 1 and is in the class as well. This concludes the task and gives us $|\mathcal{H}^d_{con}| \leq 3^d + 1$.

2. Conclude that $\text{VCdim}(\mathcal{H}) \leq d \log 3$.
   Now since the all-positive function does only give us 1 as a result, it can not be in our set determining the VC-dimension. So the problem is equivalent to finding teh VC-dimension of a set of size 3^d, which is finite. So the VC-dimension is smaller than the log over the size, and thus we can conclude $\text{VCdim}(\mathcal{H}) \leq d \log 3$ immediately.

3. Show that $\mathcal{H}^d_{con}$ shatters the set of unit vectors $\{\mathbf{e}_i : i \leq d\}$.
   For every collection of unit vectors, we can choose an i such that e_i is in the collection. Then $f(\mathbf{x}) = x_i$ gives our value 1 and $f(\mathbf{x}) = \bar{x}_i$ gives us

0. Since this is available for all combinations of the unit vectors, we have that the set of unit vectors can be shattered.

4. **(**)** Show that $\mathrm{VCdim}(\mathcal{H}^d_{con}) \le d$.
Hint: Assume by contradiction that there exists a set $C = \{c_1, \dots, c_{d+1}\}$ that is shattered by $\mathcal{H}^d_{con}$. Let $h_1, \dots, h_{d+1}$ by hypotheses in $\mathcal{H}^d_{con}$ that satisfy
$$\forall i, j \in [d+1], h_i(c_j) = \begin{cases} 0 & i = j \\ 1 & \text{otherwise} \end{cases}$$
For each $i \in [d+1]$, h_i (or more accurately, the conjunction that corresponds to h_i) contains some literal l_i which is false on c_i and true on c_j for each $j \neq i$. Use the Pigeonhole principle to show that there must be a pair $i < j \le d+1$ such that l_i and l_j use the same x_k and use that fact to derive a contradiction to the requirements from the conjunctions h_i, h_j.

5. 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 $\mathrm{VCdim}(\mathcal{H}^d_{mcon}) = d$.

Solution