Lecture 04 Cristiano Panighel 2058351

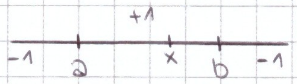
- VC-Dimension of an interval in $\mathbb{R}$

Let's suppose we have an interval $[a,b]$ with $a,b \in \mathbb{R}$, $a<b$ and an hypothesis function $h \in \mathcal{H}$, where $\mathcal{H}$ is the hypothesis space, defined as:

$$h(x) = \begin{cases} +1 & \text{if } a \leq x \leq b \\ -1 & \text{otherwise} \end{cases}$$

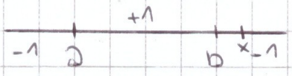
$VC(\mathcal{H}) \geq 1$

Let's start with the trivial case in which we have only 1 point. We can implement all the dichotomies, in fact when x should be $+1$ we can choose an interval $[a,b]$ that simply includes it.



(+1)

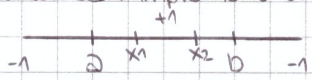
Similar intuition also for the other case, if x should be -1, we can choose an interval $[a,b]$ in which x is outside.



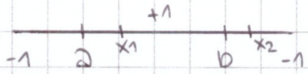
(-1)

$VC(\mathcal{H}) \geq 2$

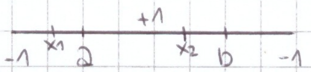
Let's consider 2 points. In this case we have four possible dichotomies and it's easy to see that we can implement them all



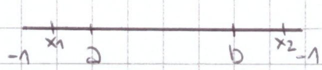
(+1,+1)



(+1,-1)



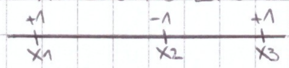
(-1,+1)



(-1,-1)

$VC(\mathcal{H}) \geq 3$

Let's consider 3 points. In this case we have 8 possible dichotomies but we cannot implement them all. In fact, in the case of $(+1,-1,+1)$, there is not such an interval that includes both x_1 and x_3 and exclude x_2 in a fixed interval $[a,b]$ where $x_1<x_2<x_3$



A little more formally, we have $x_1<x_2<x_3$ and since x_1 and x_3 are "marked" with $+1$, then follows that $a \leq x_1 < x_3 \leq b$. At the same time to assign -1 to x_2, we have that $b<x_2$ and this is not possible because $b \geq x_3 > x_2$. It's not possible to find a number which is both greater and less than x_2.

In conclusion, we can say that the VC-Dimension of an interval in $\mathbb{R}$ is 2