Sample Problems 3

To be discussed on 24.05.2024

Sample Problem 3.1: Growth Function and VC Dimension of Unions

Let $\mathcal{H}, \mathcal{H}' \subset \{\pm 1\}^{\mathcal{X}}$ be two hypothesis classes. Bound the growth function of $\mathcal{H} \cup \mathcal{H}'$ in therms of the growth functions of $\mathcal{H}, \mathcal{H}'$. From there, derive a bound on the VC dimension of $\mathcal{H} \cup \mathcal{H}'$.

Sample Problem 3.2: Growth Function and VC Dimension for Decision Stumps

We generalise the decision stumps of $\mathbb{R}$ to *multi-dimensional decision stumps* $\mathcal{H}_p \subseteq \{\pm 1\}^{\mathbb{R}^p}$. Let $x^{(i)}$ denotes the i-th coordinate of $x \in \mathbb{R}^p$.

1. State the corresponding Hypothesis class.
2. Compute an upper bound on the growth function of the form $\tau_{\mathcal{H}_p}(m) \leq$?
3. Show that the bound is tight for $p = 1$.
4. What does this imply on the VC dimension of $\mathcal{H}_p$?

Sample Problem 3.3: Graph dimension of multi-class decision stumps

In this problem, we look at 3-class classification over $\mathbb{R}$ using a generalisation of decision stumps. Define a function $h_{t_1,t_2,a,b,c} : \mathbb{R} \to \{0, 1, 2\}$ as

$$h_{t_1,t_2,a,b,c}(x) = a \cdot \mathbf{1}\{x < t_1\} + b \cdot \mathbf{1}\{t_1 \leq x < t_2\} + c \cdot \mathbf{1}\{x \geq t_2\},$$

parametrised by threshold $t_1, t_2 \in \mathbb{R}$ and integers $a, b, c \in \{0, 1, 2\}$ such that a, b, c are distinct.

Define the hypothesis class $\mathcal{H}_{3ds} \in \{0, 1, 2\}^{\mathbb{R}}$ as the set of all above predictors, $\mathcal{H}_{3ds} = \{h_{t_1,t_2,a,b,c} \;:\; t_1, t_2 \in \mathbb{R},\; a, b, c \in \{0, 1, 2\} \text{ distinct}\}$.

1. Let growth function be defined as in the lecture, $\tau_{\mathcal{H}}(m) = \max\limits_{C:|C|=m} \left|\mathcal{H}_{|C}\right|$. Compute the growth function for the above hypothesis class $\mathcal{H}_{3ds}$.

2. Graph dimension is a generalisation of VC-dimension for multiclass classifiers and is defined as follows. Let $\mathcal{H} \subseteq \mathcal{Y}^{\mathcal{X}}$ be a hypothesis class. We say that $\mathcal{H}$ G-shatters a finite set $C \subseteq \mathcal{X}$ if there exists a function $f : C \to \mathcal{Y}$ such that for every $S \subseteq C$, there is a $h \in \mathcal{H}$ such that $h(x) = f(x)$ for every $x \in S$ and $h(x) \neq f(x)$ for every $x \in C \setminus S$.

   Compute the Graph dimension of the hypothesis class $\mathcal{H}_{3ds}$.