1. Assume that $H_n$ is a class of binary classifiers with instance base $X_n = \{-1, 1\}^n$, for $n \in \mathbb{N}$. Show that $\log |H_n|$ is polynomial in $n$, if and only if $\text{VCdim}(H_n)$ is polynomial in $n$. (You may use Sauer’s Lemma.)

2. (a) Assume that $H_2$ is obtained from $H_1$ by adding $k$ new hypotheses: $H_2 = H_1 \cup \{ h_1, \ldots, h_k \}$. Prove or disprove that $\text{VCdim}(H_2) \leq \text{VCdim}(H_1) + k$. Hint: start by considering $k = 1$.

(b) Assume $H = H_1 \cup H_2$. Prove or disprove that $\text{VCdim}(H) \leq \text{VCdim}(H_1) + \text{VCdim}(H_2)$.

3. Let $H_n$ be the concept class of $n$-dimensional axis-parallel rectangles ("boxes"). In other words, $H_n$ consists of functions $h$ such that for some constants $a_1, b_1, \ldots, a_n, b_n$ we have $h(x_1, \ldots, x_n) = 1$ if and only if $a_i \leq x_i \leq b_i$ for all $i$. Show that $\text{VCdim}(H_n) = 2^n$.

4. Given a sequence $a_1 \leq \cdots \leq a_k$ where $a_i \in \mathbb{R}$ for $1 \leq i \leq k$, define further $a_0 = -\infty$ and $a_{k+1} = \infty$. The sequence $a$ defines two binary classifiers with instance space $\mathbb{R}$:
   
   - $f_a(x) = 1$ iff $a_i < x \leq a_{i+1}$ where $i$ is odd,
   - $\tilde{f}_a(x) = 1$ iff $a_i < x \leq a_{i+1}$ where $i$ is even.

   A function $h: \mathbb{R} \to \{-1, 1\}$ that is either $f_a$ or $\tilde{f}_a$ for some $a \in \mathbb{R}^k$ is called an interval classifier with $k$ splits. Let $H_k$ be the class of interval classifiers with $k$ splits.

   (a) What is the VC dimension of $H_k$?

   (b) Give an efficient empirical risk minimisation algorithm for the hypothesis class $H_k$. Hint: use dynamic programming.

5. On pages 168–169 of the lecture notes we saw how the Rademacher complexity can be estimated by empirical risk minimisation, assuming that $H$ is closed under complementation and we know how to do empirical risk minimisation. Show that if $H$ is not closed under complementation, we can similarly estimate the Rademacher complexity, assuming we know how to do both empirical risk minimisation and empirical risk maximisation.