Instructor: Rocco A. Servedio

Computer Science 4252: Introduction to Computational Learning Theory
Problem Set #4 Fall 2025

Due 11:59pm Tuesday, November 11, 2025

See the course Web page for instructions on how to submit homework.
Important: To make life easier for the TAs, **please start each problem on a new page.**

Remember to strive for both clarity and concision in your solutions;
solutions which are excessively long may be penalized.

Problem 1

(i) A *parity function* $c_S : \{0,1\}^n \to \{0,1\}$ tests whether the parity of some subset S of the n Boolean variables $x_1, \ldots, x_n$ is odd or even. In other words, if the number of variables in S that have value 1 is odd then $c_S(x) = 1$, and if the number is even then $c_S(x) = 0$. For example, the function $c_S = x_1 \oplus x_3 \oplus x_4$ computes the parity of the subset $S = \{x_1, x_3, x_4\}$, and $c_S(0010) = 1$, $c_S(1010) = 0$.

The class of *parity functions* $\mathcal{P}$ consists of all functions that can be described in this way. Formally,

$$\mathcal{P} = \{c_S \mid c_S = \bigoplus_{x_i \in S} x_i, \text{ where } S \subseteq \{x_1, \ldots, x_n\}\}.$$

Determine the exact value of the VC dimension of the class $\mathcal{P}$ of parity functions. You must define a function $f(n)$; give a set of $f(n)$ examples and show that it is shattered by C; and show that no set of size $f(n)+1$ is shattered by C.

(ii) Now consider the instance space X consisting of all words of at most 7 letters that have a dictionary entry at www.dictionary.com. Let $\mathcal{C}$ be a concept class consisting of 26 concepts, c_a through c_z. A word w in X is an element of c_ℓ if the letter ℓ is present in w. So, for example, the word "book" is an element of c_b, c_o and c_k.

What is the VC dimension of $\mathcal{C}$? Justify your answer.

Problem 2

(i) Let $\mathcal{C}$ be a concept class over $\{0,1\}^n$ for which $|\mathcal{C}| = 2^{n^2}$. True or false: the VC dimension of $\mathcal{C}$ is greater than n. Justify your answer.

(ii) Let X be a finite domain with $|X| = 2k+1$ (in other words X contains an odd number of elements). True or false: There is a concept class $\mathcal{C}$ over X for which $|\mathcal{C}| = \frac{1}{2}2^{|X|}$ (so half of all possible concepts over X belong to $\mathcal{C}$) and $\mathrm{VCDIM}(\mathcal{C}) \le k$. Justify your answer.

Problem 3 A *decision tree* over $\{0,1\}^n$ is a full binary tree (each internal node has two children) in which each internal node is labeled with a variable and each leaf node is labeled with an output bit. The output of a decision tree T on an input $a \in \{0,1\}^n$ is the output bit which is reached by the path through T which is determined by a; this path starts at the root of T and goes right when a satisfies the variable at a node and goes left when a does not.

Let $\mathcal{C}_k$ be the class of all decision trees over $\{0,1\}^n$ with k internal nodes (hence with $k+1$ leaves). Show that for *any* subset S of $\{0,1\}^n$ that contains k strings in $\{0,1\}^n$, the set S is shattered by $\mathcal{C}_k$. What can you infer about the VC dimension of $\mathcal{C}_k$ from this?

Problem 4 Recall that a *halfspace* over the domain $X = \{0,1\}^n$, also known as a linear threshold function, is a function of the form $h : \{0,1\}^n \to \{-1,1\}$, $h(x) = \text{sign}(w_1 x_1 + \cdots + w_n x_n - \theta)$, where $\text{sign}(t) = 1$ if $t \geq 0$ and $\text{sign}(t) = -1$ if $t < 0$.) Let $\mathcal{C}$ be the concept class of all "XORs of two halfspaces" over $X = \{0,1\}^n$; i.e. a concept $c : \{0,1\}^n \to \{-1,1\}$ is in $\mathcal{C}$ if there are two halfspaces $h_1, h_2 : \{0,1\}^n \to \{-1,1\}$ such that

$$c(x) = 1 \quad \text{if and only if} \quad \text{EITHER } (h_1(x) = 1 \text{ and } h_2(x) = -1) \text{ OR } (h_1(x) = -1 \text{ and } h_2(x) = 1)$$

(but if $h_1(x) = h_2(x) = 1$ or $h_1(x) = h_2(x) = 0$, then $c(x) = 0$).

Show that the concept class $\mathcal{C}$ is PAC learnable in $\text{poly}(n, 1/\varepsilon, 1/\delta)$ time. (You may use any results stated in class about the learnability of halfspaces.)