

COMS 4773 Fall 2026 Problem Set 2

Problem 1. This problem is about learning concept classes on the Boolean n -cube $\{0, 1\}^n$.

- (a) Consider the concept class $\mathcal{C}_a := \{d_{c_1, c_2, c_3} \mid c_1, c_2, c_3 \in \text{CONJ}_n\}$ on domain $\{0, 1\}^n$, where

$$d_{c_1, c_2, c_3}(x) = c_1(x) \vee c_2(x) \vee c_3(x) \quad \forall x \in \{0, 1\}^n.$$

Explain how to use Attribute Elimination to learn $\mathcal{C}_a$ in the online mistake bound learning model. The mistake bound should be $O(n^C)$ and the running time of the algorithm in each round should be $O(n^{C'})$, for some positive constants $C, C' > 0$. Try to make C and C' as small as possible.

Hint: Use the distributive property of Boolean algebra, as well as some input transformations.

- (b) Consider the concept class $\mathcal{C}_b := \{c_{d_1, d_2, d_3} \mid d_1, d_2, d_3 \in \text{DISJ}_n\}$ on domain $\{0, 1\}^n$, where

$$c_{d_1, d_2, d_3}(x) = d_1(x) \wedge d_2(x) \wedge d_3(x) \quad \forall x \in \{0, 1\}^n.$$

Give an algorithm to learn $\mathcal{C} := \mathcal{C}_a \cup \mathcal{C}_b$ in the online mistake bound learning model. The mistake bound should be $O(n^C)$ and the running time of the algorithm in each round should be $O(n^{C'})$, for some positive constants $C, C' > 0$. The C and C' should not be larger than what you could achieve in Part (a).

For each part, an informal but clear description of the algorithm should suffice.

Problem 2. Let m and n denote positive integers. Suppose $x_1, \dots, x_m$ are m unit vectors from $\mathbb{R}^n$ (so $\|x_i\| = 1$ for all $i \in [m]$). For $\gamma \in (0, 1)$, define

$$\mathcal{C}_\gamma := \left\{ c: \mathbb{R}^n \rightarrow \{-1, 1\} \mid \exists w \in \mathbb{R}^n \text{ such that } \|w\| \leq 1 \text{ and } c(x_i) \langle w, x_i \rangle \geq \gamma \text{ for all } i \in [m] \right\}.$$

Prove that the number of different ways of “labeling” the m vectors $x_1, \dots, x_m$ using functions from $\mathcal{C}_\gamma$ is $O(m^{1/\gamma^2})$, i.e.,

$$\left| \left\{ (c(x_1), \dots, c(x_m)) \mid c \in \mathcal{C}_\gamma \right\} \right| \leq O(m^{1/\gamma^2}).$$

(Above, γ is treated as a constant; the asymptotic parameters are m and n , so we are seeking a bound that has no dependence on n .)

Hint: Use the Perceptron algorithm.

Problem 3. Let $K_n = (V, E)$ denote the complete graph on $n \geq 4$ vertices.

- (a) Let $\mathcal{C}_{K_n} := \{c_v \mid v \in V\}$ be the concept class on the domain E , where for each $v \in V$,

$$c_v(e) := \mathbb{1}\{v \in e\} \quad \forall e \in E.$$

Determine the value of $\text{Ldim}(\mathcal{C}_{K_n})$ by exhibiting a mistake tree of the requisite depth and also giving a learning algorithm with the requisite mistake bound.

- (b) If $\mathcal{C}$ is a concept class on the domain $\mathcal{X}$, then its dual class $\mathcal{C}^* := \{\mathbf{e}_x \mid x \in \mathcal{X}\}$ is a concept class on the domain $\mathcal{C}$, where for each $\mathbf{e}_x \in \mathcal{C}^*$ (called the point evaluation function for x),

$$\mathbf{e}_x(c) := c(x) \quad \forall c \in \mathcal{C}.$$

Determine the value of $\text{Ldim}(\mathcal{C}_{K_n}^*)$ (as above).

- (c) Recall the EXPERTS concept class on the domain $\{0, 1\}^N$. Describe the dual class EXPERTS^* , and determine the value of $\text{Ldim}(\text{EXPERTS}^*)$ (as above).
- (d) It turns out that $\text{Ldim}(\mathcal{C}^*) \leq 2^{2^{\text{Ldim}(\mathcal{C})+2}-1} - 2$ for every concept class $\mathcal{C}$.

We say $((x_1, c_1), \dots, (x_\ell, c_\ell)) \in (\mathcal{X} \times \mathcal{C})^\ell$ is a ladder of length ℓ in a concept class $\mathcal{C}$ on domain $\mathcal{X}$ if, for all $i, j \in [\ell]$,

$$\begin{aligned} i < j &\implies c_j(x_i) = 1, \\ i > j &\implies c_j(x_i) = 0. \end{aligned}$$

The length of the longest ladder in a concept class is the same as that for its dual class. Furthermore, the existence of long ladders in a concept class can be related to the existence of deep mistake trees for the concept class.

Fact 1. *For any non-negative integer ℓ and concept class $\mathcal{C}$, there exists a ladder of length ℓ in $\mathcal{C}$ iff there exists a ladder of length ℓ in $\mathcal{C}^*$.*

Fact 2. *For any non-negative integer d and concept class $\mathcal{C}$, if there is a ladder of length $2^{d+1} - 1$ in $\mathcal{C}$, then there is a mistake tree for $\mathcal{C}$ of depth d .*

Fact 3. *For any non-negative integer ℓ and concept class $\mathcal{C}$, if there is a mistake tree for $\mathcal{C}$ of depth $2^\ell - 1$, then there is a ladder in $\mathcal{C}$ of length ℓ .*

(i) Prove Fact 1.

(ii) Prove Fact 2.

Hint: Use induction. For $m := 2^d$, take the middle pair (x_m, c_m) in a ladder of length $2m - 1$, and construct a mistake tree by annotating the root node with x_m and then invoking the inductive hypothesis (twice!) on subsets of $\mathcal{C}$ to get the rest of the tree.

- (iii) Use Fact 3 (without proof), as well as Facts 1 and 2, to prove that for any concept class $\mathcal{C}$,

$$\text{Ldim}(\mathcal{C}^*) \leq 2^{2^{\text{Ldim}(\mathcal{C})+2}-1} - 2.$$