Tarea n°2

El objetivo de esta tarea es investigar el comportamiento de la dimensión de Vapnik-Chernonenkis V. Para una clase $\mathcal{A}$ de subconjuntos de $\mathbb{R}^d$ y elementos $x_1, \ldots, x_n \in \mathbb{R}^d$, denotamos

$$\mathcal{A}(x_1^n) = \{b = (b_1, \ldots, b_n) \in \{0,1\}^n : \exists A \in \mathcal{A} : b_i = \mathbb{1}_{x_i \in A}, i = 1, \ldots, n\}.$$

El coeficiente de *añicos* (eg : *shatter* coefficient) es

$$\mathcal{S}_{\mathcal{A}}(n) = \max_{x_1, \ldots, x_n \in \mathbb{R}^d} |\mathcal{A}(x_1^n)|$$

Finalmente, sea $V(\mathcal{A})$ el mayor n tal que $\mathcal{S}_{\mathcal{A}}(n) = 2^n$. El numero $V(\mathcal{A})$ se llama la dimensión de Vapnik-Chervonenkis (VC) de $\mathcal{A}$. Si por cada n, $\mathcal{S}_{\mathcal{A}}(n) = 2^n$ pongamos $V(\mathcal{A}) = \infty$. (¡Cuidado! Este definición cambia de una unidad de la del curso.)

P.1 Sean $\mathcal{A}$ y $\mathcal{B}$ dos clases de subconjuntos de $\mathbb{R}^d$ y sean $n, m \geq 1$ números naturales. Mostrar que

1. $\mathcal{S}_{\mathcal{A}}(n+m) \leq \mathcal{S}_{\mathcal{A}}(n)\mathcal{S}_{\mathcal{A}}(m)$.
2. Si $\mathcal{C} = \mathcal{A} \cup \mathcal{B}$, entonces $\mathcal{S}_{\mathcal{C}}(n) \leq \mathcal{S}_{\mathcal{A}}(n) + \mathcal{S}_{\mathcal{B}}(n)$.
3. Si $\mathcal{C} = \{C = A^c : A \in \mathcal{A}\}$, entonces $\mathcal{S}_{\mathcal{C}}(n) = \mathcal{S}_{\mathcal{A}}(n)$.
4. Si $\mathcal{C} = \{C = A \cap B : A \in \mathcal{A} \text{ and } B \in \mathcal{B}\}$, entonces $\mathcal{S}_{\mathcal{C}}(n) \leq \mathcal{S}_{\mathcal{A}}(n)\mathcal{S}_{\mathcal{B}}(n)$
5. Si $\mathcal{C} = \{C = A \cup B : A \in \mathcal{A} \text{ and } B \in \mathcal{B}\}$, entonces $\mathcal{S}_{\mathcal{C}}(n) \leq \mathcal{S}_{\mathcal{A}}(n)\mathcal{S}_{\mathcal{B}}(n)$
6. Si $\mathcal{C} = \{C = A \times B : A \in \mathcal{A} \text{ and } B \in \mathcal{B}\}$, entonces $\mathcal{S}_{\mathcal{C}}(n) \leq \mathcal{S}_{\mathcal{A}}(n)\mathcal{S}_{\mathcal{B}}(n)$

P.2 Justificar que la dimensión VC es bien definida.

Digamos que un conjunto $B \subset \{0,1\}^n$ *rompe* un conjunto $S = \{s_1, \ldots, s_m\} \subset \{1, 2, \ldots, n\}$ si la restricción de B a los componentes $s_1, \ldots, s_m$ es el hypercubo binario completo de dimensión m. Eso significa que

$$\{(b_{s_1}, \ldots, b_{s_m}) : b = (b_1, \ldots, b_n) \in B\} = \{0,1\}^m.$$

P.3.a Sea $B_0 \in \{0,1\}^n$, $S = \{s_1, \ldots, s_m\}$ y

$$\Psi_1 : \begin{array}{rcl} \mathcal{P}(\{0,1\}^n) & \longrightarrow & \mathcal{P}(\{0,1\}^n) \\ B & \longmapsto & \Psi_1(B) \end{array}$$

una transformación de cada $b = (b_1, \ldots, b_n) \in B$ en $\bar{b}$ tal que :

— Por los b tal que $b_1 = 1$, $\bar{b} = (0, b_2, \ldots, b_n)$ si $(0, b_2, \ldots, b_n) \notin B$ y $\bar{b} = b$ si no.
— Por los b tal que $b_1 = 0$, $\bar{b} = b$.

Denotamos $B_1 = \Psi_1(B_0)$. Probar que $|B_0| = |B_1|$ y que B_1 rompe $S \implies B_0$ rompe S.

P.3.b Denotamos $\Psi_2, \ldots, \Psi_n$ las mismas transformaciones análogas por cada coordenada $2, \ldots, n$. Sea $B_n = \Psi_n \circ \cdots \circ \Psi_1(B_0)$ y $V \in \mathbb{N}^*$. Supongamos que para cada $S \subset \{1, \ldots, n\}$ con $|S| = m > V$, B_0 no rompe S. Mostrar que si $v = (v_1, \ldots, v_n) \in \{0,1\}^n$ es un vector fijo tal que $v \in B_n$, entonces

$$T_v = \{b \in \{0,1\}^n : b_i = 0 \text{ si } v_i = 0\}$$

es un subconjunto de B_n. Deducir que v no contiene más que V unos.

P.3.c Sea $T = \cup_{v \in \mathcal{V}} T_v$, donde $\mathcal{V}$ es el conjunto de vectores que contienen no más que V unos. Mostrar que $B_n \subset T$ y calcular el valor de $|T|$ en función de V y n.

P.3.d Utilizando las preguntas anteriores mostrar el **Lema (de Sauer)** :
Sea $\mathcal{A}$ una clase tal que $V(\mathcal{A}) < \infty$, entonces para cada n,

$$\mathcal{S}_{\mathcal{A}}(n) \leq \sum_{i=0}^{V} \binom{n}{i}.$$

P.4 Sea $\mathcal{A}$ con $V = V(\mathcal{A}) < \infty$. Mostrar que $\mathcal{S}_{\mathcal{A}}(n) \leq (n+1)^V$. Mostrar que en el caso particular $n \geq V$, $\mathcal{S}_{\mathcal{A}}(n) \leq \left(\frac{ne}{V}\right)^V$. (Así si $V < \infty$ el shatter coefficient crece de manera polinomial...)

P.5 Sean $\mathcal{A}$ y $\mathcal{B}$ dos clases de dimensión VC, V_1 y V_2 respectivamente. Mostrar que la dimensión VC de $\mathcal{A} \cup \mathcal{B}$ es menor or igual a $V_1 + V_2 + 1$.

Sea ρ la distancia sobre $\{0,1\}^n$ dada por $b, c \in \{0,1\}^n$,

$$\rho(b,c) = \sqrt{\frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{b_i \neq c_i}}.$$

En los siguientes preguntas supongamos que $\mathcal{A}$ tiene una dimensión de Vapnik-Chervonenkis **finita** V.

P.6.a Fijo $x_1, \ldots, x_n$, sea $B = \mathcal{A}(x_1^n)$ y $0 \leq r \leq 1$. Sea $B_r \subset \{0,1\}^n$ los centros de las bolas de una cubierta mínima de B por bolas de radio r. (Justificar que siempre existe). Mostrar que existe un conjunto $C_r \subset B$ tal que, para cada $b, c \in C_r$, $\rho(b,c) > r$ y tal que $|B_r| \leq |C_r|$.

P.6.b Denotamos $c^{(1)}, \ldots, c^{(M)}$ los elementos del conjunto C_r definido en la pregunta anterior. Sean, para cada $c^{(i)}, c^{(j)}$

$$A_{i,j} = \{m = 1, \ldots, n : c_m^{(i)} \neq c_m^{(j)}\}.$$

Ademas, sean $K \in \mathbb{N}^*$ fijo y $Y_1, \ldots, Y_K$ variables aleatorias independientes y uniformes sobre $\{1, \ldots, n\}$. Mostrar que, $\forall k \leq K$,

$$\mathbb{P}\left(Y_k \in A_{i,j}\right) \geq r^2.$$

P.6.c Probar que

$$\mathbb{P}\left(\forall i, j \leq M : i \neq j, \text{ al menos un } Y_k \text{ pertenece a } A_{i,j}\right) \geq 1 - M^2 e^{-Kr^2}.$$

P.6.d Tomar $K = \lceil 2\log M/r^2 \rceil + 1$. Mostrar que existen unos indices $y_1, \ldots, y_K \in \{1, \ldots, n\}$ todos diferentes tal que todos los elementos de $(C_r)_{|y}$ construidos por restricción de los elementos de C_r a las coordenadas $y_1, \ldots, y_K$ son distintos. Mostrar que $|(C_r)_{|y}| = M$.

P.6.e Justificar que C_r no rompe cualquier conjunto de tamaño mas grande que V. Deducir que

$$M \leq \left(\frac{eK}{V}\right)^V \qquad \text{por } K \geq V.$$

P.6.f Utilizando la pregunta anterior mostrar que

$$\log M \leq \frac{V}{1 - 1/e} \log \frac{4e}{r^2}$$

(Pista : Pensar que para cada $x > 0, \log x \leq x/e$. Estudiar los casos $\log M \geq V$ y $\log M \leq V$)

P.6.g Concluir que

Teorema Sea $\mathcal{A}$ una clase de subconjuntos de $\mathbb{R}^d$ de dimensión de Vapnik-Chervonenkis finita V. Entonces, por cada $x_1, \ldots, x_n \in \mathbb{R}^d$ y $0 \leq r \leq 1$,

$$H(r, \mathcal{A}(x_1^n), \rho) \leq \frac{V}{1 - 1/e} \log \frac{4e}{r^2}$$