

Examen rápido No. 1

Reconocimiento de Patrones (2023-2)

Julio Weissman Vilanova

Nombre: _____

Problema 1

La desigualdad de Hoeffding es un medio para caracterizar el error fuera de muestra de una hipótesis de aprendizaje a partir de una cota superior:

$$P [|E_{out}(h^*) - E_{in}(h^*)| > \varepsilon] \leq 2|\mathcal{H}|e^{-2\varepsilon^2 M}, \quad \varepsilon > 0.$$

Si establecemos que $\varepsilon = 0.05$ y queremos que la cota superior $2|\mathcal{H}|e^{-2\varepsilon^2 M}$ sea a lo más 0.03, ¿Cuál sería el número mínimo de datos de entrenamiento M dependiendo de la cardinalidad del modelo de aprendizaje?

- Para el caso $|\mathcal{H}| = 1$:
- Para el caso $|\mathcal{H}| = 1$:
- Para el caso $|\mathcal{H}| = 1$:

Problema 2

Consideremos el modelo de aprendizaje que vamos a llamar *2-intervalos*. En este modelo \mathcal{H} vamos a considerar que:

$$h : \mathbb{R} \rightarrow \{0, 1\},$$

donde $h(x) = +1$ si el punto $x \in \mathbb{R}$ se encuentra dentro de alguno de dos intervalos (i.e. $[a, b]$ y $[c, d]$) preestablecidos. En caso contrario, $h(x) = -1$.

- ¿Cual es el breakpoint de \mathcal{H} ?
- ¿Que valor tiene $d_{VC}(\mathcal{H})$?

Ahora consideremos el caso genérico *M-intervalos*, donde \mathcal{H} está definido como el conjunto de funciones $h : \mathbb{R} \rightarrow \{0, 1\}$ tales que $h(x) = +1$ si x se encuentra en alguno de los M intervalos establecidos y $h(x) = -1$ en caso contrario.

- ¿Que valor tiene $d_{VC}(\mathcal{H})$?