

## Tarea 4 (parte 1)

Entregar a más tardar el martes 8PM

- Supongamos que para un problema de clasificación binaria, se construye un clasificador donde se permite, además de regresar como predicción 0 y 1, también abstenerse.

El costo de predecir 1 si la verdadera categoría es 0, es 1 peso. El costo de predecir 0 si la verdadera categoría es 1, también es 1 peso. El costo de abstenerse es  $\theta$ , una constante dada de antemano:  $0 < \theta < \frac{1}{2}$ .

Calcula el clasificador Bayesiano óptimo en función de  $\theta$  y  $P(Y = 1|X = x)$ .

- Considera un problema de clasificación binaria con predictores  $X$ . Supongamos que  $P(Y = 1) = P(Y = 0)$  y que  $P(X|Y = i)$  sigue una distribución Poisson con parámetro  $\lambda_i$ .

Derive el clasificador Bayesiano óptimo si el costo de un falso positivo es dos veces el costo de un falso negativo.

- Supongamos que  $X, Y$  sean v.a. discretas:

	X=0	X=1	X=2
Y=0	0.1	0.3	0.25
Y=1	0.25	0.05	0.05

Si  $L(0, 1) = L(1, 0)$ , calcula el clasificador Bayesiano óptimo de  $Y$  usando  $X$ .

Si  $L(0, 1) = 2L(1, 0)$  calcula el clasificador Bayesiano óptimo y su error (promedio) correspondiente.

- (no entregar) Calcula el clasificador Bayesiano óptimo para una función de costo simétrico,  $Y|X = x \sim \mathcal{N}(\mu, \sigma_y^2)$ ,  $y \in \{0, 1\}$  y  $P(Y = 1) = P(Y = 0)$ .
- (no entregar) Para un problema de clasificación binaria y  $x \in \mathcal{R}^2$ , dibuja un conjunto de datos con tres observaciones donde el clasificador 1-NN tiene un error empírica (sobre el conjunto de entrenamiento) que no sea cero.