

Tarea 2: Aplicaciones de la Matemática en Ingeniería

25 de Octubre de 2021

1. La siguiente tabla muestra un conjunto de datos que contiene 6 observaciones, 1 covariable y 1 variable respuesta categórica:

x	1	5	10	20	40	80
y	a	b	b	b	b	a

- Muestre que este conjunto de datos no se puede clasificar de manera perfecta con un clasificador lineal.
 - Pruebe que al expandir el espacio de la covariable a través del mapeo $\phi : \mathbb{R} \rightarrow \mathbb{R}^2$ dado por $\phi(x) = (x, \sqrt{x})$, si se puede hallar un clasificador lineal perfecto.
 - Plantee el problema de optimización que permite hallar el clasificador lineal perfecto de margen máximo. Escriba explícitamente la función objetivo y todas las restricciones involucradas.
 - Identifique los vectores soporte en el problema de clasificación del inciso **c.**
2. En un problema de clasificación binario se utilizó la regla de Bayes con 3 umbrales de decisión diferentes, obteniendo las matrices de confusión de más abajo:

	actual	
predicted	0	1
0	22	3
1	22	15

	actual	
predicted	0	1
0	32	6
1	12	12

	actual	
predicted	0	1
0	43	11
1	1	7

- Para cada uno de estos umbrales, calcule la sensibilidad y la especificidad.
 - Haga un bosquejo de la curva ROC correspondiente y obtenga el área bajo la curva.
 - ¿Qué se puede decir sobre el desempeño del clasificador?
3. Considere el conjunto de datos del archivo `datos_T2.txt` (disponible en AULA), el cual contiene 200 observaciones, 2 covariables (X_1 y X_2) y 1 variable respuesta categórica (X_3). Aplique el método de K vecinos más cercanos para abordar el problema de clasificación. Específicamente, use validación cruzada para evaluar el desempeño del método en términos de K (considere al menos 3 valores distintos de K). Explique claramente el procedimiento usado y sus conclusiones.