
PROBLEMA 1.

1.a) Definir las cabeceras de las funciones necesarias para aplicar la técnica de validación *bootstrapping*, indicando qué parámetros de entrada y de salida se proponen, así como el tipo de datos de cada parámetro.

1.b) Implementar dichas funciones

1.c) Escribir un programa realizando llamadas a las funciones anteriores para determinar el error que comete el clasificador de mínima distancia sobre la base de datos IRIS, usando la distancia de Mahalanobis, y la técnica de *bootstrapping* para la validación de los resultados.

Se puede hacer uso de la función *d_mahal.m* incluida en la librería pattern.

PROBLEMA 2.

2.a) Implementar una función que estime la función de densidad *d* en los puntos (*x_i*, *y_i*) mediante ventanas de Parzen en 2D usando gaussianas, sabiendo que se dispone de los datos en dos vectores *x* e *y*.

$$d = \text{Parzen2D}(x, y, \text{sigma}, x_i, y_i);$$

Se supone que el espacio 2D es $0 \leq x \leq 1$, $0 \leq y \leq 1$ y que para la definición de las gaussianas 2D, se cumple que $\sigma_x = \sigma_y = \sigma$.

2.b) Se quiere usar la función anterior para estimar el error que se comete al clasificar usando la técnica de estimación de densidad con ventanas de Parzen de la siguiente forma:

- Reducir la dimensión del espacio de entrada a 2 usando la técnica de Fisher
- Estimar la densidad en los puntos de test usando la función *Parzen2D* y los datos de entrenamiento, con un *sigma* adecuado. En este sentido, proponga un valor numérico para *sigma* y justifíquelo.
- Clasificar cada punto de test asignándole la clase que tenga mayor densidad

En la resolución de este apartado se usará la técnica de la validación simple (100 datos para entrenar y 50 para test) ejecutada y promediada un total de 1000 veces.

La resolución de todos los apartados del examen se realizarán en MATLAB.

BASE DE
DATOS
IRIS
3x150