



TAREA 1

Entrega el 10 de marzo

Tarea 1. Clasificador Bayesiano.

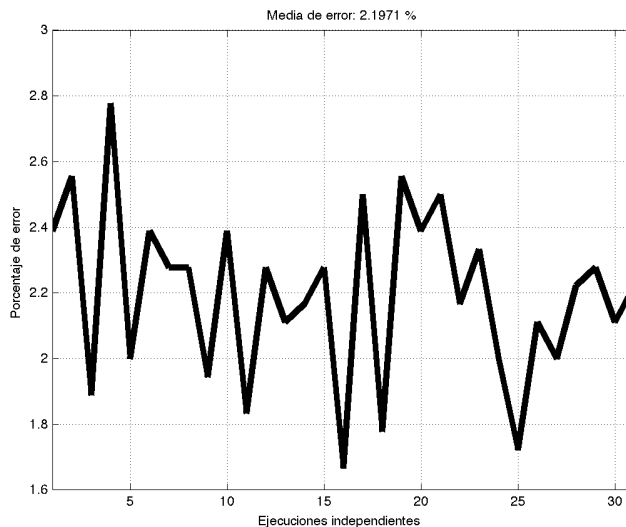
- Implementar las siguientes funciones en Matlab:
 - $[\mu_{wi}, S_{wi}, P_{wi}] = \text{trainBayes}(X_{tr}, Y_{tr})$, donde X_{tr} es el conjunto de entrenamiento representado por una matriz $d \times n$ (d dimensiones y n patrones), Y_{tr} es un vector $1 \times n$ con la clase conocida c para cada patrón de entrenamiento, μ_{wi} es una matriz $d \times c$ con los valores de las medias para cada clase c , S_{wi} es una hipermatriz $d \times d \times c$ con las matrices de covarianza $d \times d$ para cada clase c , y P_{wi} es un vector $1 \times c$ con las probabilidades de cada clase.
 - $[Y_{pred}, err] = \text{classifyBayes}(X_{tt}, Y_{tt}, \mu_{wi}, S_{wi}, P_{wi})$, donde X_{tt} es el conjunto de prueba representado por una matriz $d \times n$ (d dimensiones y n patrones), Y_{tt} es un vector $1 \times n$ con la clase verdadera de cada patrón de entrenamiento. Implementar la regla de decisión MAP. Y_{pred} es la clase asignada (o predicha) por el clasificador para cada patrón desconocido de entrenamiento, y err es el porcentaje de error del conjunto de prueba definido como:

$$err = \frac{100}{N} \sum_{i=1}^N Q[y_i, r_i]$$

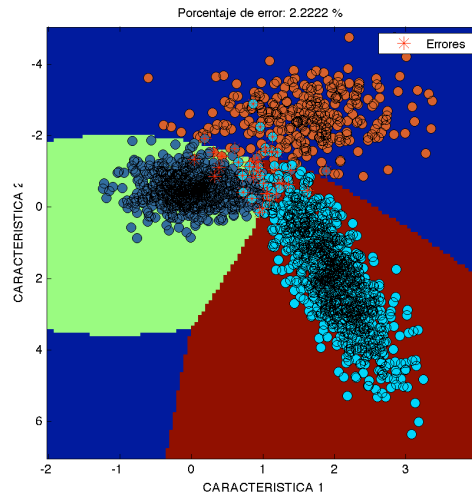
donde N es el número total de patrones de prueba y $Q[\cdot]$ es la discrepancia entre la clase verdadera, y_i , y la clase asignada, r_i , del patrón i , definida como:

$$Q[y_i, r_i] = \begin{cases} 0 & \text{si } r_i = y_i \\ 1 & \text{si } r_i \neq y_i \end{cases}$$

- Se proporcionan cinco conjuntos de datos simulados los cuales contienen dos variables: X , vectores de patrones bidimensionales ($2 \times n$), e Y , etiqueta de clase para cada patrón ($1 \times n$). El conjunto original se dividirá aleatoriamente en patrones de entrenamiento (70%) y patrones de prueba (30%). Puede hacerse uso de la función “`crossvalind`” para dividir el conjunto original. Se harán 31 ejecuciones independientes y en cada una se medirá el porcentaje de error. Se deberá mostrar el gráfico de error versus número de ejecuciones, como se muestra a continuación:



Adicionalmente, se deberá mostrar la partición del espacio de característica generado en alguna de las ejecuciones independientes (por ejemplo, la última ejecución), donde se muestren los patrones de prueba sobre el espacio particionado y marcando aquellos clasificados erróneamente, como se muestra a continuación:



- La práctica se ejecutará desde un archivo principal, para clasificar los conjuntos de datos proporcionados. El código debe ir debidamente comentado, no olvidando colocar un encabezado que identifique al autor de la tarea.
- Entregar reporte de la práctica en archivo PDF con formato libre, con los elementos mencionados en la Clase00_RP.pdf.