

## Práctica: Análisis Discriminante

- (LDA univariado) Implemente una función **univariateLDA** que reciba los siguientes parámetros:
  - $X$  vector de valores.
  - $Y$  vector de clases para cada valor  $X$ .
  - $k$  cantidad de clasesy devuelva *otra* función que sea una predictor. Es decir, debe devolver otra función que permita, dado un valor  $x$ , obtener su clase.
- Genere 1000 datos  $z_i \sim U(-1, 2)$ , y asigne la clase 0 si el  $z_i$  correspondiente es menor que cero y la clase 1 en caso contrario. Defina  $x_i = 0.5 + 5 * z_i + \epsilon_i$  donde  $\epsilon_i$  sigue una distribución normal estándar, será el valor observado. Obtenga una clasificación LDA usando la función anterior de un conjunto de training elegido de 700 elementos y en base a eso clasifique los 300 elementos restantes, obteniendo métricas de error. Compare con algún esquema de regresión logística.
- Proceda como en el primer ejercicio, pero para el caso multivariado, llamándolo **multivariateLDA**.
- Considere el archivo **iris.data**, que tiene información sobre la longitud del sépalo, su ancho, la longitud del pétalo y su ancho, todo en centímetros. Finalmente, la última columna tiene información sobre la clase a la que pertenece. Implemente un mecanismo de clasificación por LDA para estos datos utilizando todas las variables.
- Efectúe una clasificación por LDA de los datos **abalone.txt**, discriminando entre adulto e infante, utilizando aquellas variables que considere más pertinentes. Elija el mejor modelo, usando criterio, sentido común y alguno de los métodos propuestos (K-fold, por ejemplo), para determinarlo. Se recomienda separar inicialmente un 20%-30% de los datos para poder hacer la evaluación final a través de una matriz de confusión.
- Compare, para el ejercicio anterior, el “mejor” esquema de clasificación que usted considere entre regresión logística, vecinos más cercanos y análisis discriminante. Comente sobre la métrica o las métricas de comparación que se utilizaron.