

Trabajo Práctico 2: Selección de variables

1. Prepare código en R para los métodos:
 - a) Wrapper greedy backward
 - b) Filter con test no-paramétrico (Kruskal-Wallis)
 - c) RFE

Se puede usar como base el código de ejemplo del wrapper greedy forward proporcionado.

2. En el código proporcionado se crean dos datasets artificiales (datosA y datosB). Analice cada dataset usando proyecciones que den información de las variables (por ejemplo un biplot de una PCA). Aplique los 4 métodos de selección de variable (los 3 desarrollados más el forward) a los dos datasets. Analice los resultados.

3. Prepare un dataset del problema diagonal del práctico 1 con 10 variables, 50 puntos por clase y $C=2$. Agregue 90 variables de ruido uniforme. Aplique los 4 métodos de selección. Repita el experimento 30 veces (repetir todo, crear el dataset y elegir las variables). Calcule para cada método el porcentaje de aciertos en el ranking (cuántas de las 10 variables relevantes están en los primeros 10 lugares del ranking).

4. Opcional (2 puntos): Busque un dataset que considere interesante, muestre un biplot (PCA mas variables originales), seleccione variables sobre el mismo con alguno de los métodos estudiados, analice y explique sus resultados.

Entrega: notebook con todo el código funcionando y con los comentarios correspondientes a cada resultado-