

Practica 2: selección de subconjuntos de atributos

Ejercicio 1

- Utilizar el conjunto de datos Tmusic normalizando sus atributos.
- Utilizar el método de selección de atributos basada en la correlación (CfsSubsetEval) junto al método de selección hacia adelante con las parametrizaciones por defecto propuestas en Weka 3.8.3.
- Comparar la elección de atributos así obtenida con las que obtuvimos con los métodos de filtro (número total de atributos y atributos seleccionados).
- Estimar, mediante 10-XV, la tasa de error de los siguientes algoritmos utilizando los atributos seleccionados, con este método de selección:
 - J48, NB, IBK3, Regresión Logística, MLP (H10), SVM(lineal)
- Añadir los resultados obtenidos a la tabla resumen de cada algoritmo (con la selección individual) y discutir los resultados.

Ejercicio 2

- Utilizar el método de selección de atributos basada en la correlación (CfsSubsetEval) junto al método de selección hacia adelante. Configurar Weka para forzar la selección de 10 atributos.
- El subconjunto obtenido, ¿contiene el subconjunto obtenido con la configuración por defecto?

Ejercicio 3

- Utilizar el conjunto de datos Tmusic normalizando sus atributos.
- Utilizar el método de envoltorio con un método de selección hacia adelante .
- Estimar la tasa de error mediante 10XV para los algoritmos :
 - J4.8, SVM(lineal)
- Añadir estos resultados a las tablas ya obtenidas para estos mismos algoritmos con todos los métodos de selección ya aplicados. Discutir los resultados para cada tabla.
- Resumir, en menos de 300 palabras, las conclusiones obtenidas.