

UNIDAD TEMÁTICA 5: Aprendizaje No Supervisado, Clustering

Trabajo de Aplicación 2 – Clustering Jerárquico

EJERCICIO 1

ESCENARIO

Como vimos en la presentación, hay dos enfoques básicos del clustering jerárquico:

- Aglomerativo (“bottom-up”)
- Divisivo (“top-down”)

Deseamos realizar una revisión general de los operadores de clustering jerárquico que ofrece RapidMiner, con el fin de conocer cómo se configuran y cuándo se recomienda usar cada uno

Se debe realizar entonces:

- POSTER indicando los operadores de clustering jerárquico de RapidMiner
- Para cada operador, explicar brevemente de qué tipo es, los parámetros que tiene y los valores que estos parámetros pueden asumir
- Utilizando el dataset “Ripley”, construir un proceso en RapidMiner que cree modelos de clustering por k-means, agglomerative clustering y top-down clustering. Evaluar el rendimiento de cada uno de estos modelos (justificar la elección de los parámetros correspondientes en cada caso). Utilizar vectores de performance (ver TA01 de la clase pasada) para obtener la evaluación de cada uno

Poner los resultados en el POSTER (tabla comparativa)