

Pautas Memoria Practica Clustering 2019



- Práctica de Clustering:
 - -Algoritmo k-medias (sólo inicialización aleatoria + distancia euclídea, con estimación previa del parámetro k con chain-map): este algoritmo es obligatorio.
 - -Algoritmo secuencial : este algoritmo es obligatorio.
 - -algoritmo mainshift(utilizar una librería externa) : este algoritmo es opcional y sólo para mejorar nota.
 - La memoria es básicamente un informe sobre los resultados obtenidos con los algoritmos de clustering aplicados a los data sets disponibles en la asignatura: UCI (Iris,wine, cancer, isolet) y MNIST.
- Apartados de la memoria:
- 1.- Implementación: breve descripción del desarrollo de la práctica; incluir el listado de los programas
 - 2.- Resultados: descripción de los resultados obtenidos [índice de compactación %C] con cada algoritmo para todos los data sets, con especial énfasis en la optimización del índice de compactación %C en relación a los parámetros de diseño de cada algoritmo: en el caso del k-medias para la optimización de k es **muy recomendable** aplicar el chain-map y estimar a ojo el valor ótimo de k ([RAF 2014-Clustering Chainmap Libro.pdf](#)) en el caso del secuencial se debe optimizar con un bucle for(variando los valores de sus parámetros); para el mainshift , experimentar con la ventana de búsqueda de la librería externa. [Indices Evaluación Clustering 2015.pdf](#)
 - 3.- Conclusiones: breve comentario final sobre los resultados obtenidos en esta práctica.
 - Entrega de la memoria en formato pdf mediante correo directo al profesor[dario.maravall@upm.es] Esta práctica se puede realizar y entregar individualmente o en pareja.
 - [RAF2018-clustering-transparencias.pdf](#)