

## Práctica 4

### Parte Obligatoria

El archivo excel `4000datosSimuladosEnergia.xlsx` contiene datos de 4000 clientes ficticios de una empresa comercializadora de energía eléctrica y gas. Cada contrato tiene 10 atributos:

1. Antigüedad del contrato
2. Superficie del domicilio
3. Potencia eléctrica contratada
4. Consumo de electricidad en punta
5. Consumo de electricidad en llano
6. Consumo de electricidad en valle
7. Consumo de gas medio
8. Generación fotovoltaica
9. Bonificaciones consumo
10. Programa puntos

Se desea visualizar los 4000 contratos en un mapa, de forma que se pueda identificar visualmente si hay grupos de contratos con las mismas propiedades. Como ayuda, se proporciona una columna adicional, llamada

11. Perfil de contrato

Esta columna es la solución del problema. Solo se empleará para colorear los puntos del mapa (es decir, si algún método produjese un mapa correcto, éste debería consistir en figuras aisladas de diferente color).

Ninguno de los algoritmos de generación de mapas debería hacer uso de esta variable salvo para colorear el resultado final.

Tomando como base el notebook `manifold-learning.ipynb` visto en clase de teoría (se incluye una copia en este enunciado, puedes cortar y pegar de este notebook lo que necesites), crea un notebook llamado `comercializadora.ipynb` donde se le apliquen al problema de la comercializadora los métodos indicados a continuación.

Lista de métodos:

1. PCA
2. LLE LTSA
3. MDS
4. UMAP

En el algoritmo LLE, prueba a cambiar el número de vecinos hasta que obtengas la representación más informativa. Discute qué método produce la solución mejor.

**Parte opcional 1:**

Aplica los siguientes métodos:

1. Sammon
2. Isomap
3. t-SNE

Comenta si alguno de estos métodos mejora a los cuatro probados anteriormente.

**Parte opcional 2:**

Aplica el método Kernel PCA con los kernels sigmoide y rbf (ajusta los parámetros que necesites) y discute los resultados.