

Agrupamiento

Fabián Villena

Enero 2024

Se le pide detectar la existencia de grupos de objetos en un conjunto de datos de características de semillas.

En particular, trabajaremos con un conjunto de datos que contiene múltiples atributos físicos de semillas y que se cree que tiene 3 variedades, pero no se tienen las clases de cada una de las semillas.

Los resultados esperados son que apliquen una técnica de agrupamiento y sugieran el grupo al cual pertenece cada una de las semillas.

<https://github.com/fvillena/biocompu/blob/2023/data/seeds.csv>

Preguntas

Responda las siguientes preguntas en un *Jupyter Notebook* con código desarrollado en el lenguaje de programación Python.

1. Verifique y preprocese si es necesario las características.
2. Realice agrupamiento con los algoritmos *k-Means* y Agrupamiento Jerárquico utilizando distintas cantidades de grupos.
3. Utilice el método del codo para seleccionar una buena cantidad de grupos.
4. Explore los grupos sugeridos por los métodos ajustados con la cantidad de grupos encontrada por el método del codo.