**Clustering**

## **Dataset**

Choose any dataset you like at openml.org. The selected data set must contain at least 100 objects. Remember to vectorize the attributes of your data set (move from categories to numbers), fill in the blanks (if any), and normalize.

## **Algorithms**

**Option 1 (100/100)**

Implement several algorithms (with sklearn):

* DBSCAN
* k-Means
* Hierarchical algorithm

**Option 2 (120/100)**

Implement several algorithms (without sklearn):

* DBSCAN
* k-Means
* Hierarchical algorithm

## **Analysis (obligatory for option 1 and option 2)**

For each partition you make (for each clustering algorithm) evaluate it with two internal metrics: Calinski-Harabasz and Davies-Bouldin

**Soft deadline: Monday, February 5**