10 Final lab

You can use external libraries for linear algebra operations but you are expected to write your own algorithms.

10.1 Exercise 1

Implement your own versions of:

- density peaks clustering (without the "halo" feature)
- Normalized Mutual Information (NMI)
- F-score

Using the `Aggregation` dataset available at `https://cs.joensuu.fi/sipu/datasets/` under Shape sets:

1. Use **your** version of density peaks at several values of the `dc` parameter and choose the one that maximises the NMI with the ground truth (third column).
2. Compute the NMI with the ground truth F-score for this assignation.
3. Perform k-means with k equal to the true number of clusters, compute the NMI witht the ground truth and the F-score.

Notes

Hints for point 1. As a rule of thumb, `dc` should be chosen in such a way that, on average, between 1% and 2% of the data points are included in the neighbourhood. Use the exponential kernel.