DATA SCIENCE

TRAINING PROGRAM

Clustering Lab
27.11.2019

Exercise

1. Implement K-means algorithm (or use ready sklearn implementation) in Python and generate 2 datasets with 1000 samples each using **make_circles** and **make_moons** to test your clustering implementation.
2. Find an optimal number of clusters using Silhouette method using sklearn.
3. ⭐Additional task: read [DBScan paper](#) (Section 4.2) and find for make_moons dataset for minPt=4, optimal Eps using nearest neighbour approach. Apply DBScan with those parameters using sklearn implementation.