Instructions

The project is written in Java 8 in IntelliJ IDEA on macOS Sierra operating system. The following java library dependencies are included in the project: opencsv, smile-core, weka, poi, poi-ooxml.

GADriver, PSODriver and KMeansDriver are classes used to run experiments on GA-based, PSO-based and K-Means clustering algorthms respectively. The project contains thus 3 jar files corresponding to each of those classes, which should be run from the project root. Clustering solutions in the form of array are saved into txt file (Passed arguments are included in file's name).

GADriver and PSODriver take four arguments as an input: argument 1 is a name of configuration (e.g. mgaC13, CONF5); argument 2 is a name of dataset (e.g. GLASS, DIM064); argument 3 is a seed to start experiments from (e.g. 0); argument 4 is a number of experiments (e.g. 30).  Examples:

- java -jar out/artifacts/GADriver/GADriver.jar mgaC13 GLASS 0 30

- java -jar out/artifacts/GADriver/GADriver.jar mgaC26 COMPOUND 0 30

- java -jar out/artifacts/PSODriver/PSODriver.jar CONF5 DIM064 0 15

- java -jar out/artifacts/PSODriver/PSODriver.jar CONF5 DIM064 15 15

KMeansDriver has the same parameters except name of configuration. If arguments are passed, experiments are run to determine optimal number of clusters in the dataset using Elbow Method for K-Means Clustering. Otherwise, that is if no arguments passed, experiments are run on K-Means using a pre-defined number of clusters, obtained from experimental results utilizing Elbow Method for K-Means clustering. Example:

- java -jar out/artifacts/KMeansDriver/KMeansDriver.jar PATHBASED 0 30

- java -jar out/artifacts/KMeansDriver/KMeansDriver.jar

To evaluate produced clustering solutions and ResultsGenerator is used, which generates performance report in the form of excel file from given txt file where clustering solutions produced by algorithms are stored.