Algorithmic assessment DAAR 7: kMeans2D.

BM Bui-Xuan

Let $\mathcal{S}$ be a set of n points in a 2D plane.

k-means clustering :

The k-means clustering problem consists of computing a partition of $\mathcal{S}$ into k parties in such a way that the total sum of distances between each element of a party to the barycenter of that party is minimized. Heuristics for k-means clustering aim to give a partition with a small total sum of distances (without being the optimal one).

1 Assessment for DAAR 7

The overall goal is to propose a heuristics to the k-means clustering problem in a 2D plane, with $k = 5$. In the GUI canvas, $\mathcal{S}$ is given by the input file `input.points` and (after parsing,) the variable named `points`.

The evaluation of the assessment will mainly depend on the computing time spent for the 200 input instances of the test (key 'g' in the GUI), and also the clarity of the code.

N.B. : however, in order to obtain more than 10/20, your average score must be lower than 110% the score of the 'bot' mentioned in the GUI. In other words, your score must be lower than 137889.

Constraints :

— The sprint is individual work, no team allowed.

— Send one screenshot of the result from the GUI (after hitting 'g'), as well as the code of function `calculKMeans` to `buixuan@lip6.fr`, 3 emails maximum per student.

— Deadline : 14 November 2021, 23h59, by mail server timestamp. Late report penalty : penalty of $0,1 * 2^k$ points for k minutes late.