

1154-90-1259

Soledad Villar* (soledad.villar@nyu.edu), 60 5th ave, New York, NY 10011. *Monte Carlo approximation certificates for k-means clustering.*

Efficient algorithms for k-means clustering frequently converge to suboptimal partitions, and given a partition, it is difficult to detect k-means optimality. In this paper, we develop an a posteriori certifier of approximate optimality for k-means clustering. The certifier is a sub-linear Monte Carlo algorithm based on Peng and Wei's semidefinite relaxation of k-means. In particular, solving the relaxation for small random samples of the dataset produces a high-confidence lower bound on the k-means objective, and being sub-linear, our algorithm is faster than k-means++ when the number of data points is large. (Received September 14, 2019)