



Contribution ID: 36

Type: **not specified**

## Fully automated clustering by accurate non-parametric density estimation

*Thursday, December 10, 2015 3:40 PM (20 minutes)*

The use of the right density estimates in the framework of density-based cluster analysis is one of the keys to reveal the properties of a given dataset. We developed a new unsupervised and adaptive density estimator [1], that is able to reconstruct the point local density in an accurate way, even in highly inhomogeneous datasets. By combining the new density estimator with a generalized version of a recently developed clustering approach [2], we can automatically recognize sets of data points organized in clusters, regardless of the dataset characteristics (i.e. space dimensionality, shape of the clusters, distance metrics). We demonstrate the power of the algorithm performing cluster analyses on biological systems to disentangle complexity patterns. In particular, interesting results have been obtained by analysing rRNA sequences from human GUT microbiota.

[1] M. d'Errico, A. Laio, E. Facco and A. Rodriguez, "An accurate and unsupervised density estimator for highly inhomogeneous datasets" (In preparation).

[2] A. Rodriguez, A. Laio, "Clustering by fast search and find density peaks", Science, 2014, 344, 1492-1496

**Primary author:** Dr D'ERRICO, Maria (SISSA)

**Co-authors:** Prof. LAIO, Alessandro (SISSA); Dr RODRIGUEZ, Alex (SISSA); Ms FACCO, Elena (SISSA)

**Presenter:** Dr D'ERRICO, Maria (SISSA)

**Session Classification:** Session 5