



3IA

Institut interdisciplinaire
d'intelligence artificielle (3IA)

Côte d'Azur

Internship in Statistical/Machine Learning Deep generative models for clustering functional data

Equipe MAASAI, Institut 3IA Côte d'Azur, Inria & Université Côte d'Azur
in collaboration with EDF

Context

Clustering is the task of organizing similar objects into meaningful groups. With the big data phenomenon, modern data are now high dimensional and /or heterogeneous. This provides new challenges and there is a need to develop new clustering methods adapted to such data. One approach for clustering such data relies on data embedding, which consists in embedding the original data into a low-dimensional space in which the clustering can be easily performed. Recently, deep learning models have been successfully used in many domains, and in particular for the task of data embedding.

Subject

The goal of the internship is to develop a clustering approach based on deep learning models for functional data. The main missions are:

- to study the recent development in clustering methods using deep learning
- to develop a model on the basis of an original idea proposed by the internship supervisor,
- to test this model on simulated data and onto data provided by EDF (electricity consumption data).

An article presenting the model will be written during the internship.

Internship conditions

Location: the intern will join the Maasai team of Inria Sophia-Antipolis and Université Côte d'Azur, which is composed of 25 researchers in statistical and machine learning (web: <https://team.inria.fr/maasai/>). The team is part of the Institut 3IA Côte d'Azur.

Duration: 6 months, starting in March 2021

Salary: approx. 550€ / month

How to apply:

The candidates should be Master 2 students in Applied Mathematics or Computer Science, with a strong background in statistical and machine learning. The candidates should have good knowledge of R/Python and TensorFlow. To apply, send a CV + M1/M2 grades + motivation letter by email to charles.bouveyron@inria.fr.

References

- C. Bouveyron, L. Bozzi L., J. Jacques J. and F-X. Jollois (2018). The Functional Latent Block Model for the Co-Clustering of Electricity Consumption Curves, Journal of the Royal Statistical Society, Series C, 67 [4], 897-915.
- DLC2018: IEEE ICDM Workshop on deep learning and clustering, November 17th 2018, Singapore.
- N. Dilokthanakul, P. A. M. Mediano, M. Garnelo, M.C. H. Lee, H. Salimbeni, K. Arulkumaran & M. Shanahan, Deep unsupervised clustering with gaussian mixture variational autoencoders, working paper 2017.