Epione team / ERC G-Statistics Job offer

Research Engineer for Geomstats

Context

In the context of the ERC Advanced grant G-Statistics, a research engineer position (18 months) is offered to contribute to the development and maintenance of the geomstats Python package.

There is an increasing interest in leveraging differential geometry in the machine learning community. Indeed, in many applications, the data or the model parameters belong to a manifold, i.e. a space that is defined by a set of constraints or invariance properties, and that only locally resembles the usual Euclidean space. It is now common to account for this geometric structure and to define an adequate Riemannian metric to measure distances on the manifold. Statistics and machine learning methods thus need to be generalized to manifold-valued data, both for mathematical coherence and statistical consistency, but also with the aim to encode prior knowledge into geometric constraints and consequently improve learning. This research field, termed Geometric Statistics, has been prolific in the past decades, with applications to many fields, such as directional statistics, medical imaging, computer vision, robotics, and others.

The ERC G-Statistics aims at grounding the mathematical foundations of geometric statistics and demonstrating their impact on selected applications. A strong attention is given to the implementation of novel generic algorithms. In particular, we are actively involved in the development of Geomstats, an open-source Python package that provides tools to perform statistics and machine learning on Riemannian manifolds. The package is designed both for mathematicians and for applied scientists for whom most of the mathematical difficulties are hidden under high-level functions, while providing the suitable material to learn Riemannian geometry. With this tool, we ambition to democratize the use of geometric statistics in data science.

Links: https://github.com/geomstats/geomstats
https://geomstats.github.io/
https://gstats.inria.fr/

Host group/location

The successful candidate will be part of both the G-Stats group and the Epione team at Inria, the French national research institute for digital science and technology that conducts world-class research. She/he will find an exciting working environment in Sophia-Antipolis on the French Riviera, with the ability to collaborate daily with international researchers on technical implementations and practical applications. Through our international collaborations, she/he will have the opportunity to travel, e.g. to our collaborators' labs in the Silicon Valley.

Assignment

The successful candidate will work closely with the researchers of the team and our collaborators in the USA, to pursue the development of the package. She/he will be expected to contribute directly to the codebase by refactoring the sampling and visualization modules to make them work with more geometries, implementing new statistical learning algorithms and implementing new geometries.

Research scientists of the G-Statistics team will bring the mathematical expertise needed for the new geometries and algorithms while the successful candidate will support their personal developments, with the goal to identify what can be generic and help bringing the code to a level that can be included in geomstats. Every code contribution will come with the corresponding documentation and unit-tests and will be reviewed by peers on Github.

She/he will also oversee the deployment of unit-tests to ensure the compatibility of geomstats on GPUs. Furthermore, the successful candidate will be expected to interact with the international community of geomstats contributors and users via Github, by performing code reviews, responding to issues, and addressing bug reports. Finally, a few hackathons may be organized to gather the contributors around a coding challenge.

Main activities

- Contribute to the development of geomstats
- Support the developments of the G-Statistics team members
- Test the library on GPUs
- Respond to issues and address bug reports
- Organize hackathons

Skills

Educational level is expected to be at the Master / Engineering degree. Applicants should have a strong background in mathematics and experience in scientific programming with Python. Knowledge of the standard numerical libraries (Numpy, Scipy, Matplotlib, Scikit-learn) is required. Experience with GPUs and the related libraries (PyTorch, Tensorflow, Jax) is a plus. Applicants should also have good communication skills to interact with the researchers of the team (fluent English speaking and writing is required), be willing to learn some mathematics and be eager to solve complex numerical problems.

How to apply

Applicants should send a CV and links to previous coding projects to Xavier Pennec and Nicolas Guigui (see addresses below), and apply through the Inria procedure.

Starting date

January or February 2022

Contacts

Xavier Pennec (xavier.pennec@inria.fr) & Nicolas Guigui (nicolas.guigui@inria.fr)