1/2 years ML research engineer position in Inria Lyon: *Fighting the reproducibility crisis with Benchopt*

Contact: Mathurin Massias, Inria Lyon (mathurin.massias@gmail.fr)
The applicants are encouraged to contact Badr Moufad (badr.moufad@polytechnique.edu), who previously held a similar position (Apr. 2022 – Dec. 2023)

Location: The candidate will work in the OCKHAM team (https://www.inria.fr/en/ockham), in ENS de Lyon, 46 allée d'Italie, 69007 Lyon, France

Start date & Duration: From Sept/Oct 2024. The contract is for 1 year, with a potential 1 year extension.

Profile: We are looking for a highly motivated Masters graduate with a background in mathematics (optimization, probability and statistics, machine learning). Abilities in computer science are required. Experience with Python collaborative development/open source tools is strongly appreciated. *Depending on the outcome, a PhD may follow this contract.*

Salary: The salary is according to INRIA's grid, and around 2000 euros/month.

Keywords: MLOPS, Benchmarks, Python, open source, Machine Learning, Optimization, scikit-learn

1 Context

Numerical validation is at the core of machine learning research as the main medium to assess the actual impact of new methods, and to confirm the agreement between theory and practice. Yet, the rapid development of the field poses several challenges: researchers are confronted with a profusion of methods to compare, limited transparency and consensus on best practices, as well as tedious re-implementation work. As a result, validation is often very partial, which can lead to wrong conclusions that slow down the progress of research. More globally, this phenomenon is a part of the so-called reproducibility crisis.

Benchopt (https://benchopt.github.io) was created by Moreau et al. [1] to tackle these issues. It is a collaborative framework to *automatize, publish and reproduce* benchmarks in machine learning across programming languages and hardware architectures. Benchopt simplifies benchmarking for the community by providing an off-the-shelf tool for running, sharing and extending experiments.

Since its release in 2022, Benchopt has been used by researchers around the world and is now rapidly gaining traction with tens of existing benchmarks[1], 40k downloads and more than 20 citations. It has been acknowledged by David Donoho, one of the best-known researchers in ML, in his seminal paper "Datascience at the singularity" as an instance of tool allowing for the striking recent progress in data science [2].

2 Missions

The first step of Benchopt's development was focused on benchmarks for optimization algorithms. Now, Benchopt is extending its scope to the generic comparison of methods and models, in diverse fields such as machine learning, statistics or imaging.

In close collaboration with Benchopt developers, the candidate will partake in this endeavor through the following tasks:

- implement new features, tailored towards this new paradigm. For example, properly handling cross-validation in benchmarks is a must-have for Machine Learning benchmarks; while saving and displaying the model's results is a necessity in imaging tasks. Depending on the task at hand, metrics may need to be monitored in a different way, which will require novel development.

- continue to develop benchmarks. Despite the progress made through Benchopt, many domains are still lacking reference benchmarks, leading to slower science. In the last months, there have been several open controversies, e.g. in parameter free optimization[2] where the lack of proper benchmark was at the heart of the issue. Proposing new standard benchmarks in these fields will help the community to adopt better standards.

- provide support to the Benchopt community. Yearly sprints have been organized, with the latest in July 2023 in Paris gathering 40 participants [3]. Between sprints, keeping in touch with the developers of new benchmarks eases onboarding and allows gathering feedback.

- integration with other libraries, such as the Pytorch-based Deepinv library[4] for inverse problems in imaging.

- organize competitions around new benchmarks in the context of reproducibility workshops, conference competitions and Masters courses.

---

[1]https://github.com/orgs/benchopt/repositories?q=benchmark_
[2]https://parameterfree.com/2023/08/30/yet-another-icml-award-fiasco/
[3]https://notes.inria.fr/s/OuVJYf8bM
[4]https://deepinv.github.io/

The candidate will integrate the OCKHAM team, focused on new methods for frugal learning with applications in machine learning and signal processing. They will work in a stimulating research environment, with many potential collaborations, on a fundamental problem for today's ML research. They will get the opportunity to discover a variety of theoretical fields while developing a deep practical understanding through first-hand experience.

References

[1] Thomas Moreau, Mathurin Massias, Alexandre Gramfort, Pierre Ablin, Pierre-Antoine Bannier, Benjamin Charlier, Mathieu Dagréou, Tom Dupre la Tour, Ghislain Durif, Cassio F Dantas, et al. Benchopt: Reproducible, efficient and collaborative optimization benchmarks. *Advances in Neural Information Processing Systems*, 35:25404–25421, 2022.

[2] David Donoho. Data science at the singularity. *arXiv preprint arXiv:2310.00865*, 2023.