



**Offer #2024-07548**

## **Post-Doctoral Research Visit F/M Robust and effective stochastic optimization methods for training deep learning models**

**Contract type :** Fixed-term contract

**Level of qualifications required :** PhD or equivalent

**Fonction :** Post-Doctoral Research Visit

### **Context**

The project will be funded by the PRAIRIE 3IA Institute -- ANR-19-P3IA-0001.

Travel expenses are covered within the limits of the scale in force.

### **Assignment**

The main objective of this project is the understanding and development of robust and effective stochastic optimization methods for training deep learning models.

State-of-the-art optimization techniques require careful tuning of several hyperparameters, such as learning rate (schedule), momentum, and weight decay. One aspect of this research project is to investigate and develop adaptive techniques for selecting hyperparameters, in order to reduce the tuning effort. For instance, the stochastic Polyak step size and its recent variants have shown promising results for adaptively setting the learning rate.

One challenge in this regard is that theoretically optimal hyperparameter values depend on quantities that are unknown before training. A core part of the project consists of studying online estimation of unknown quantities for hyperparameter selection and exploring its applicability for modern deep learning problem instances.

The project also aims to develop a theoretical framework for robust optimization methods that are resilient to outliers and heavy-tailed noise in gradient distributions. This is particularly relevant for transformer models, which have dominated the field in recent years but often suffer from training instabilities. We aim to investigate the interactions between model architectures and data domains in deep learning, focusing on disentangling the effects of transformer models and their input data distribution on gradient outliers.

The recruited person will be co-advised with Adrien Taylor and Francis Bach.

### **Main activities**

Main activities :

Conduct theoretical research  
Conduct experiments for empirical verification  
Write scientific articles  
Disseminate the scientific work in appropriate venues.

### **Skills**

Technical skills and level required :

Languages : High-level of professional/academic English

Coding skills : Good level of coding in Python and related deep learning libraries

### **Benefits package**

- Subsidized meals
- Partial reimbursement of public transport costs
- Leave: 7 weeks of annual leave + 10 extra days off due to RTT (statutory reduction in working hours) + possibility of exceptional leave (sick children, moving home, etc.)
- Possibility of teleworking and flexible organization of working hours (after 12 months of employment)
- Professional equipment available (videoconferencing, loan of computer equipment, etc.)

- Social, cultural and sports events and activities
- Access to vocational training
- Social security coverage

## General Information

- **Theme/Domain** : Optimization, machine learning and statistical methods  
Statistics (Big data) (BAP E)
- **Town/city** : Paris
- **Inria Center** : [Centre Inria de Paris](#)
- **Starting date** : 2024-09-01
- **Duration of contract** : 2 years
- **Deadline to apply** : 2024-05-18

## Contacts

- **Inria Team** : [SIERRA](#)
- **Recruiter** :  
Simsekli Umut / [umut.simsekli@inria.fr](mailto:umut.simsekli@inria.fr)

## About Inria

Inria is the French national research institute dedicated to digital science and technology. It employs 2,600 people. Its 200 agile project teams, generally run jointly with academic partners, include more than 3,500 scientists and engineers working to meet the challenges of digital technology, often at the interface with other disciplines. The Institute also employs numerous talents in over forty different professions. 900 research support staff contribute to the preparation and development of scientific and entrepreneurial projects that have a worldwide impact.

**Warning** : you must enter your e-mail address in order to save your application to Inria. Applications must be submitted online on the Inria website. Processing of applications sent from other channels is not guaranteed.

## Instruction to apply

### Defence Security :

This position is likely to be situated in a restricted area (ZRR), as defined in Decree No. 2011-1425 relating to the protection of national scientific and technical potential (PPST). Authorisation to enter an area is granted by the director of the unit, following a favourable Ministerial decision, as defined in the decree of 3 July 2012 relating to the PPST. An unfavourable Ministerial decision in respect of a position situated in a ZRR would result in the cancellation of the appointment.

### Recruitment Policy :

As part of its diversity policy, all Inria positions are accessible to people with disabilities.