

JOB OFFER

Software developer Deep learning tools for brain image analysis

Keywords: Python, deep learning, reproducibility, MLOps, neuroimaging, image analysis

The topic: ClinicaDL – Open-source deep learning software for reproducible neuroimaging processing

ClinicaDL is an end-to-end deep learning framework for deep learning users working on neuroimaging data that aims to prevent common pitfalls that we identified and described in our previous study [Wen et al., 2020]. These pitfalls are the difficult use of neuroimaging data sets by users with little expertise, data leakage during training and testing, and insufficient reproducibility. ClinicaDL [Thibeau-Sutre et al., 2021] includes a set of tools to prepare data for deep learning tasks (such as quality check, label definition, generation of synthetic data), architecture search, network training, as well as result inference, model evaluation and interpretation. In addition, it implements a set of technical solutions to avoid the main methodological issues causing data leakage found in the literature. ClinicaDL allows its users to work with a great diversity of neuroimaging data sets as it interacts with a neuroimaging standard, the Brain Imaging Data Structure [Gorgolewski et al., 2016]. This structure ease data processing by Clinica (www.clinica.run) [Routier et al., 2021], the companion project of ClinicaDL, that prepares images and stores them in a folder hierarchy called CAPS (Clinica Processed Structure). ClinicaDL takes this folder as input and outputs a Model Analysis and Processing Structure (MAPS), which contains all the information necessary for the reproduction of deep learning experiments.

ClinicaDL is convenient for research experiments but is not scalable enough and lacks tools for model deployment. The main aspect of the project is the integration of deep learning tools widely adopted by the community, such as MLflow and Pytorch Lightning, to improve model management and experiment tracking. To this end, the candidate will have to adapt and integrate these tools to ClinicaDL to make it more accessible and flexible by converging towards community standards.

- ClinicaDL: <https://github.com/aramis-lab/clinicadl> | <https://clinicadl.readthedocs.io>
- Clinica: <https://github.com/aramis-lab/clinica> | www.clinica.run
- Thibeau-Sutre, E., Diaz, M., Hassanaly, R., Routier, A., Dormont, D., Colliot, O., Burgos, N. ClinicaDL: An Open-Source Deep Learning Software for Reproducible Neuroimaging Processing, (2021). <https://hal.archives-ouvertes.fr/hal-03351976>
- Wen, J., Thibeau-Sutre, E., Diaz-Melo, M., ..., D., Durrleman, S., Burgos, N., Colliot, O. (2020). Convolutional neural networks for classification of Alzheimer’s disease: Overview and reproducible evaluation. *Medical Image Analysis*, 63: 101694. <https://doi.org/10.1016/j.media.2020.101694>
- Routier, A., Burgos, N., Díaz, M., ..., Colliot, O. (2021). Clinica: An Open-Source Software Platform for Reproducible Clinical Neuroscience Studies. *Frontiers in Neuroinformatics*, 15, 689675. <https://doi.org/10.3389/fninf.2021.689675>
- Gorgolewski, K. J., Auer, T., Calhoun, V. D., ... Poldrack, R. A. (2016). The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific Data*, 3(1), 160044. <https://doi.org/10.1038/sdata.2016.44>

Your mission

You will be in charge of the:

- integration of standard deep learning community tools (e.g. MLflow, Pytorch Lightning, TorchIO),
- development of new features (data augmentation, refactoring of TSV tools, configuration file maker),
- software maintenance,
- user support and animation of the community,
- contribution to training and dissemination with the other engineers of the team.

In addition, you will be presenting the software at international scientific conferences and other events (organized for instance by Inria, ICM, CNRS).

A vibrant scientific, technological, clinical and ethical environment

You will work within the ARAMIS Lab (www.aramislab.fr) at the Paris Brain Institute (<http://www.icm-institute.org>), one of the world top research institutes for neurosciences. The institute is ideally located at the heart of the Pitié-Salpêtrière hospital, downtown Paris.

The ARAMIS Lab, which is also part of Inria (the French National Institute for Research in Digital Science and Technology), is dedicated to the development of new computational approaches for the analysis of large neuroimaging and clinical data sets. You will interact locally with the PhD students, postdoctoral fellows and engineers of the ARAMIS Lab.

Your profile

- PhD degree or Master
- Strong programming skills in Python
- Good understanding of the software development process and tools (Git, continuous integration, tests)
- Basic knowledge of deep learning is mandatory and experience with the associated tools (e.g. PyTorch) would be a strong plus
- Knowledge of digital image processing is mandatory and experience with medical imaging would be a strong plus
- Excellent relational and communication skills to interact with users and lab members
- Good writing skills (documentation, website, scientific articles)

Salary: depending on experience

Type of contract: fixed-term contract

Starting date: as soon as possible

Ready to take up the challenge?

Send your CV to ninon.burgos@icm-institute.org.