Inria Saclay, MIND team http://team.inria.fr/mind

NeuroSpin, bat 145, CEA Saclay

Developer position: Machine learning for Brain Neuroimaging

Main topic: open-source software for neuroimaging data analysis
Keywords: population imaging, visualization, machine learning
Research team: MIND (Inria Saclay and CEA)
Contact: Bertrand Thirion, bertrand.thirion@inria.fr and Demian Wassermann, demian.wassermann@inria.fr
Start and duration of contract: 01/02/2024, for 2 years
Salary: Depending on experience: 28 to 36 kE/year --free from charges
Application: Interested candidate should send CV and motivation letter

Context

Human neuroscience uses magnetic resonance imaging (MRI) to understand the structure and function of the brain and to characterize certain neurological and psychiatric disorders. Large imaging cohorts have recently been established, including a thousand (Human Connectome Project, Abide, ADNI, Imagen, EU-AIMS, 1000brains, ABCD) to a hundred thousand individuals (Enigma consortium, UK Biobank). Such cohorts are necessary to study the impact of many brain pathologies (psychiatric diseases, addictions, neurodegenerative diseases) or risk factors in epidemiological studies (UK Biobank). The corresponding data are generally publicly available. In addition to these large studies, smaller datasets are acquired, and more and more frequently made public (https://openneuro.org), in the context of cognitive neuroscience. The analysis of data from all these studies requires medical image processing tools, but also, increasingly, statistical analysis and learning tools. The brain imaging community has developed a standard, the **Brain Imaging Data Structure** (BIDS)(1)to organize data and facilitate large-scale statistical analyses.
In this framework, MIND produces many contributions to statistical learning in neuroimaging, with a particular interest in supervised learning, simulation-based inference, and covariance model estimation. A part of these contributions are disseminated via the **Nilearn** library (http://nilearn.github.io)(2). Nilearn is a key open-source library in the neuroscience ecosystem, that relies on the scientific Python stack (Numpy, Scikit-learn, Matplotlib). It is very successful ($50k$ downloads/month on pypi). Nilearn is contributed by many people from several countries, see https://github.com/nilearn/nilearn/graphs/contributors. It follows the best practices in terms of software development (exhaustive automated testing, CI, complete API documentation plus narrative documentation, API homogeneity, reasonable dependencies, open discussions on technical choices, etc.) The development is managed by a **coredev team** with 9 members that meets monthly. The developer community is very active, as it provides feedback on public channels like Neurostars, opens issues and pull requests on the Github interface.
Finally, MIND is investing a lot of resources in **clinical collaborations**. Specifically, MIND is engaged in a collaborative initiative with the Assistance Publique - Hopitaux de Paris (AP-HP), Institut Pasteur, Sainte Anne, Stanford University and Neurospin, to address clinical scenarios such as brain tumor surgeries, analysis of stroke-induced lesions (3; 4), understand the relationship between brain structure and cognition, or the use of ultra-high field MRI.

Mission

The job proposal includes two main, equally important, missions.
The first mission is to **set up robust pre-processing pipelines** for data coming from our collaborating institutions such that we can transfer our scientific advancements to clinical and pre-clinical settings. The users of these pipelines will be the members of the MIND team as well as close collaborators. These pipelines will be concerned mainly with two different image modalities, functional and diffusion MRI. Data from

all these modalities has to first undergo artifact correction, denoising, and standardization. The pipelines addressing previously described use cases will build on tools such as Fmriprep, Nilearn, Dipy, ANTs and FreeSurfer. Beyond basic use of these powerful tools, we need to add steps for quality control of the output or special adaptations to the pathological cases in our collaboration, such as brain tumors.
A second, equally important objective is to **ensure the development of Nilearn**. In relation to the other goal, a planed development is an improvement of the connection with other tools from the ecosystem, in particular BIDS standards: leveraging these standards will facilitate the analysis of BIDS-organized datasets. In general, we would like to automate more the main analytical steps in Nilearn to improve user's experience. We also want to upgrade interactive visualization tools with the integration of NiiVue. Finally, depending on the developer's skills on optimization, we aim at improving performance of some Nilearn functionality based on profiling/memory management analysis, use of GPU computing etc.

Job Offer description

More in detail, the following actions will be undertaken:

- Organization of several datasets available to the MIND team using BIDS conventions
- Further automate Nilearn pipelining in BIDS contexts to improve user experience
- Setting up initial Functional and Diffusion MRI pipelines based on existing processing tools
- Extraction of High Level Features from the preprocessed dataset for use in neuroscientifi studies
- Integration of Niivue for interactive visualization in Nilearn
- Improve performance of some Nilearn functionality (profiling/memory management/GPU...).
- Animation of the Nilearn community

Skills and profile

- Love high-quality code and open source
- Worry about users and like to communicate
- Be curious about data (ie like looking at data and understanding it)
- Have an affinity for problem-solving tradeoffs
- Good scientific Python coders
- Enjoy interacting with a community of developers
- Interest in brain imaging and its applications.
- Experience in optimization is a plus.

Working at Inria

Established in 1967, Inria is the only public research body fully dedicated to computational sciences.
Combining computer sciences with mathematics, Inria's 3,500 researchers strive to invent the digital technologies of the future. Educated at leading international universities, they creatively integrate basic research with applied research and dedicate themselves to solving real problems, collaborating with the main players in public and private research in France and abroad and transferring the fruits of their work to innovative companies.
The researchers at Inria published over 4,500 articles in 2022. They are behind over 300 active patents and 120 start-ups. The 220 project teams are distributed in eight research centers located throughout France.

Working with MIND team

Besides permanent researchers, the developer will be in contact with PhD students that do software development as part of their PhD contract, and with the developer team that contributes to many tools of the scientific Python ecosystem (sklearn, joblib, Benchopt). He/she will also be in contact with cognitive and clinical neuroscientists at NeuroSpin.
Mind researchers use English as a common language for their activities (daily interactions, weekly meetings, yearly retreats).

Benefits

- Canteen and cafeteria;
- Sports equipment;
- Partial transport reimbursement

References

[1] K. J. Gorgolewski, T. Auer, V. D. Calhoun, R. C. Craddock, S. Das, E. P. Duff, G. Flandin, S. S. Ghosh, T. Glatard, Y. O. Halchenko, D. A. Handwerker, M. Hanke, D. Keator, X. Li, Z. Michael, C. Maumet, B. N. Nichols, T. E. Nichols, J. Pellman, J.-B. Poline, A. Rokem, G. Schaefer, V. Sochat, W. Triplett, J. A. Turner, G. Varoquaux, and R. A. Poldrack, "The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments," *Scientific Data*, vol. 3, p. 160044, June 2016.

[2] A. Abraham, F. Pedregosa, M. Eickenberg, P. Gervais, A. Mueller, J. Kossaifi, A. Gramfort, B. Thirion, and G. Varoquaux, "Machine learning for neuroimaging with scikit-learn," *Frontiers in Neuroinformatics*, vol. 8, 2014.

[3] J. Benzakoun, M.-A. Deslys, L. Legrand, G. Hmeydia, G. Turc, W. B. Hassen, S. Charron, C. Debacker, O. Naggara, J.-C. Baron, B. Thirion, and C. Oppenheim, "Synthetic FLAIR as a Substitute for FLAIR Sequence in Acute Ischemic Stroke," *Radiology*, p. 211394, Jan. 2022.

[4] J. Benzakoun, S. Charron, G. Turc, W. B. Hassen, L. Legrand, G. Boulouis, O. Naggara, J.-C. Baron, B. Thirion, and C. Oppenheim, "Tissue outcome prediction in hyperacute ischemic stroke: Comparison of machine learning models," *Journal of Cerebral Blood Flow and Metabolism*, p. 0271678X2110243, June 2021.