2018-01040 - Engineer

Contract type : Public service fixed-term contract
Level of qualifications required: Graduate degree or equivalent
Fonction : Temporary scientific engineer

About the research centre or Inria department

Located at the heart of the main national research and higher education cluster, member of the Université Paris Saclay, a major actor in the French Investments for the Future Programme (Idex, LabEx, IRT, Equipe) and partner of the main establishments present on the plateau, the centre is particularly active in three major areas: data and knowledge, safety, security and reliability; modelling, simulation and optimisation (with priority given to energy).

The 450 researchers and engineers from Inria and its partners who work in the research centre's 31 teams, the 100 research support staff members, the high-level equipment at their disposal (image walls, high-performance computing clusters, sensor networks), and the privileged relationships with prestigious industrial partners, all make Inria Saclay Île-de-France a key research centre in the local landscape and one that is oriented towards Europe and the world.

Context

What kind of relationships can one draw between the brain and the mind? This cognitive neuroscience question has motivated the acquisition and analysis of many brain imaging datasets that assess such relationships. These datasets typically associate brain responses with a description of the stimuli assumed to involve given cognitive functions under investigation. This information is currently gathered in repositories such as NeuroVault http://neurovault.org. Analyzing these resources is thus a way to map cognitive labels with brain patterns associated with them. When considered together, the contrasts in these images form a manifold that embeds cognitive terms into a common observation space in a complex way.

Our aim is to learn the statistical relationships between brain images and the associated labels. The difficulty is that both brain images and labels are noisy, but both have a latent structure that can be leveraged: Indeed, as a resource contributed on a voluntary basis by the community, NeuroVault contains images with an incomplete field of view, as well as atlas images that do not represent any brain activity. On the other hand, the spatial structure common to all well-registered images can be leveraged to obtain low-dimensional representations of the signal, which facilitates further computations.

Similarly, manually entered labels or tags are often inconsistent or redundant, but they do not appear in isolation. Topic models can be used to de-noise and simplify the information carried by the labels.

Assignment

In this work, we propose to rely on the automatic NeuroVault data downloader built as part of the Nilearn library http://nilearn.github.io to perform a large-scale analysis of the statistical associations between images and cognitive labels. The study will develop upon the seminal work of [5] and will consist in the following steps:

- Extract usable images from NeuroVault and exclude outlier images
- Perform a curation of the labels of the associated with these images
- Apply unsupervised dimension reduction of the images
- Study the structure underlying the tags topics, hierarchical models
- Learn a mapping from labels to images, using a so-called encoding model.
- Learn reverse mappings from images to the labels, using a decoding model.
- Study similarities between brain images on the hand, labels on the other hand

Main activities

Importantly, we will consider adaptive strategy such as transfer learning (taming the heterogeneity related to data provenance to improve generalization), and differential dataset weighting. A further validation will be carried out on independent data by using the Individual brain Charting dataset, set up at NeuroSpin, that contains images of many tasks observed in a small group of participants. Further validations on OpenfMRI datasets will be carried out if time permits.

Skills

The successful candidate will be interested in applications of machine learning and in the understanding of human cognition. Knowledge of scientific computing in Python (Numpy, Scipy) is encouraged. All the work will be done in Python based on the Nilearn library http://nilearn.github.io.

Benefits package

- Subsidised catering service
- Partially-reimbursed public transport
- Social security
- Paid leave
- Flexible working hours
- Sports facilities

Remuneration

Salary in regards of diplomas and professional experiences.