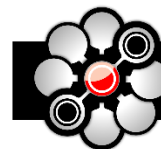




ANR-11-INBS-0010



METABOHUB

Metabolome-IDF

Data Scientist/ Biostatistician

About MetaboHUB

MetaboHUB (MTH) is the national French metabolomics and fluxomics infrastructure. Launched in 2013, MTH is a leading international infrastructure serving more than 700 scientists worldwide. MetaboHUB gathers 5 regional facilities including more than 80 permanent staffs, 15 NMRs, 43 MS, robotic and computational platforms. MTH aims at pushing forward the field to develop metabolomics and fluxomics from single cell to population. Your contribution will serve a broad range of researchers in the fields of biotechnologies, Human health and nutrition and plant science. Joining MTH, you will be involved in cutting edge research within a highly skilled and motivated consortium.

About MTH-Metabolome-IDF

The Metabolome-IDF platform has been specialized for more than 15 years in metabolomics, lipidomics, glycomics by mass spectrometry and data science for biomarker discovery in health. The data science team (Odisce) develops innovative methods and software libraries in applied mathematics and statistics for the processing (signal processing), integrative statistical analysis (machine learning) and annotation (cheminformatics) of large-scale data within large cohorts. The data and algorithms are publicly available to the community, notably through a suite of software tools available on R/Bioconductor and Galaxy/Workflow4Metabolomics.

The mission

Data integration is a major opportunity in the omics sciences to obtain more robust predictive results that are better interpretable. Numerous statistical methods have been described for building integrative predictive models and selecting variables of interest (using clustering, latent variables, kernel mixing, model fusion, etc.) [1].

The multi-infrastructures ProMetIS project [2], aims to develop high-throughput data generation and integration workflows for the deep phenotypic characterization of genetic mutations. A comprehensive multi-platform, multi-tissue and multi-omics dataset has been generated [3].

Key responsibilities

The data scientist will develop statistical integration methods based on model fusion or stacking, and compare them with alternative approaches on the ProMetIS data (e.g. Canonical Correlation Analysis, Matrix Factorization, Similarity Matrices) to identify the most effective in terms of prediction and interpretation, depending on the level of integration considered (inter-platform, inter-tissue, inter-omics). Methods development and evaluation will be implemented in a FAIR computational workflow to guarantee the reproducibility of the results.

[1] Pierre-Jean, *et al.* (2019) Clustering and variable selection evaluation of 13 unsupervised methods for multi-omics data integration. *Briefings in Bioinformatics*, 6:2011-2030.

[2] <https://github.com/IFB-ElisirFr/ProMetIS>

[3] Imbert *et al.* (2021) ProMetIS, deep phenotyping of mouse models by combined proteomics and metabolomics analysis. *Scientific Data*, 8, 311.



Profile

Biostatistician, Data Scientist
• Master in applied mathematics or statistics



Skills

Data analysis
Machine learning
Programming



Informations

CDD (Research engineer)
12 months contract, full-time position
Starting from: January 2023

How to apply?

The application should contain the following attachments:

1. motivation letter
2. CV including contact details of two references
3. relevant diplomas or university certificates

Contact

Etienne Thévenot
Data sciences team (Odisce)
UMR « Médicaments et Technologies pour la Santé » (SPI/LI-MS)
CEA, Centre de Saclay
F-91191 Gif sur Yvette, France
etienne.thevenot@cea.fr

More informations

www.metabohub.fr

<https://odisce.github.io>