

Dynamic multi-domain translation for high-throughput sequencing data

Lead supervisor: Paul Villoutreix (CENTURI, LIS)
Co-supervisor: Thierry Artières (ECOLE CENTRALE MARSEILLE)

Project description

Multi-domain translation is a problem of major interest in biology. When studying a biological system such as a developing embryo, many acquisition techniques are available. Each of them brings out unique features of the system, however, they are often incompatible and cannot be performed at the same time. To overcome this challenge we need to develop multi-domain integration techniques. Current approaches rely either on the tools of optimal transport, or multiple autoencoders, however, they are not designed to address temporal data. With this project, we propose to take advantage of multi-domain dynamical data in high-dimensional spaces to infer a dynamical coupling between sequencing data acquisition techniques (such as sc-RNASeq) and microscopy data. This will include theoretical work and computational experiments on artificial and real data. The results of the project are expected to have large impact in the machine learning community and be of wide applicability in real world biological problems. The scientific environment for this project is ideal as it combines expertise in interdisciplinary approaches of machine learning applied to biological data, and expertise in theoretical machine learning.

Keywords

Optimal Transport, Autoencoders, Single Cell Transcriptomics

Expected profile

We are looking for a PhD in machine learning, computer science, applied mathematics with strong interest in machine learning and its applications to biology. The postdoc will take place in Paul Villoutreix's interdisciplinary team (Learning meaningful representation of life <http://bioml.lis-lab.fr/>) and the Machine Learning team of the Computer Science lab in Marseille (<https://qarma.lis-lab.fr>)

References

Towards a general framework for spatio-temporal transcriptomics
Julie Pinol, Thierry Artières, Paul Villoutreix, NeurIPS, LMRL workshop, 2020

Gene expression cartography
Nitzan, Mor, et al., Nature, 2019

Multi-domain translation by learning uncoupled autoencoders
Karren D Yang, Caroline Uhler, Arxiv, 2019

