

## INTERNSHIP PROPOSAL

### Master 2 or Engineer (COMPUTER SCIENCE or BIONFORMATICS)

**Title:** Neural network-based model for the identification of Bladder Cancer subtypes

**Keywords:** Neural networks, Graph neural networks, bioinformatics, cancer prediction

**Supervisors:**

- Farida Zehraoui, Lecturer, AROB@S Team, IBISC laboratory, Paris-Saclay University, Univ. Evry
- François Radvanyi, Research Director, Molecular Oncology Team, Institut Curie
- Constance Creux, PhD student, AROB@S team, IBISC laboratory, Paris-Saclay University, Univ. Evry
- Fariza Tahiri, Professor, AROB@S Team, IBISC laboratory, Paris-Saclay University, University of Evry

**Contact:** farida.zehraoui@univ-evry.fr

**Duration of internship:** 6 months

**Location:** IBISC Laboratory, IBGBI, University of Evry, 23 Boulevard de France, 91000 Evry.

**Description:**

Precision medicine, also called stratified medicine, consists of mainly using the genomic characteristics of patients for personalized care.

“Omics” technologies: genomics (DNA sequencing), transcriptomics (microarrays), proteomics, have considerably modified the scale of data and make it possible to generate massive quantities of genomic data on patients. These data can cover all the mechanisms involved in the variations that occur in the cellular networks that influence the functioning of organ systems in humans. They can be used for diagnosis, prognosis, prediction of personalized patient treatment, etc.

Artificial intelligence, particularly machine learning, has become a promising tool over the past decade to support precision medicine in oncology. Deep learning, which is a subfield of machine learning, will play a major role in improving the accuracy of cancer susceptibility, recurrence, and survival predictions.

As cancer is a heterogeneous disease, several subtypes can be identified. Treatments and diagnostics must be tailored to each subtype. In this project, we will focus on the prediction of subtypes of a common cancer in humans, bladder cancer.

The aim of this internship, which is part of a collaboration between the IBISC laboratory and the Curie Institute, is to develop a multi-source and multiscale machine learning based on graph neural networks, in order to identify cancer subtypes and potentially discover new ones, using heterogeneous data sources representing different types of omics, images and clinical data associated with patients.

### Bibliography

1. V. Bourgeais, F. Zeharoui, B. Hanczar. GraphGONet: a self-explaining graph-based neural network encapsulating the Gene Ontology for phenotype prediction on gene expression. *Bioinformatics*, 2022.
2. V. Bourgeais, F. Zeharoui, B. Hanczar. Deep GONet: Self-explainable deep neural network based on Gene Ontology for phenotype prediction from gene expression data. *BMC bioinformatics* 22, 455, 2021.
3. Graw, S., Chappell, K., Washam, C.L., Gies, A.J., Bird, J.T., Robeson, M.S., & Byrum, S.D. (2020). Multi-omics data integration considerations and study design for biological systems and disease. *Molecular omics*.

4. B. Hanczar, F. Zehraoui, T. Issa, M. Arles. Biological interpretation of deep neural network for phenotype prediction based on gene expression. *BMC bioinformatics* 21 (1), 1-18, 2020.
5. L. Platon, F. Zehraoui, F.Tahi. Localized Multiple Sources Self-Organizing Map. *International Conference on Neural Information Processing, ICONIP (3) 2018* : 648-659.
6. L. Platon, F. Zehraoui, A. Bendahmane, F. Tahi. IRSOM, a reliable identifier of ncRNAs based on supervised Self-Organizing Maps with rejection. *Bioinformatics*, 34(17), i620-i628, 2018.
7. L. Platon, F. Zehraoui, F.Tahi. Self-Organizing Maps with supervised layer. In *Proceedings of the 12th International Workshop on Self-Organizing Maps and Learning Vector Quantization, Clustering and Data Visualization (WSOM+)*. 2017: 161-168.
8. Boucheham A, Somnard V, Zehraoui F, Boualem A, Batouche M, Bendahmane A, Israeli D, Tahi F. IpiRId: Integrative approach for piRNA prediction using genomic and epigenomic data. *PLoS One* 2017, 16;12(6):e0179787.