Internship Opportunity: Integrative Analysis of Clinical and Multi-Omics Data for Blood Cancer Treatment Responses

Location: Gustave Roussy (Villejuif, France) and CentraleSupélec (Paris-Saclay)
Duration: 6 months
Start Date: March 2025

About the Project

This internship, a collaboration between **Gustave Roussy**, Europe's top cancer research center, and **CentraleSupélec**, a leading institution in applied mathematics and artificial intelligence, seeks to push the boundaries of personalized medicine. The project focuses on **predicting treatment responses in patients with subtypes of myeloid leukemia** called myelodysplastic neoplasms, undergoing Erythropoietin (EPO) and Azacitidine (AZA) therapy. Using advanced AI and machine learning techniques, the goal is to integrate clinical and multi-omics data to uncover the factors that predict treatment outcomes, to ultimately optimize therapies.

Supervision and Collaborative Environment

This internship offers a unique interdisciplinary experience with expert supervision from **CentraleSupélec's AI specialists** and **Gustave Roussy's computational biologists**. At Gustave Roussy, you'll access extensive clinical datasets and collaborate with clinicians to understand the healthcare context. CentraleSupélec will support your technical development in AI, machine learning, and data integration.

Regular joint meetings will ensure a dynamic exchange of knowledge across fields, enriching your understanding of AI applications in oncology and personalized medicine.

What You'll Do

In this role, you will have the chance to apply your mathematical and AI expertise to a real-world healthcare challenge. You'll work on creating robust data representations and survival models that will help predict patient responses, gaining hands-on experience with practical applications of AI in oncology.

Your main responsibilities will include:

1. **Data Exploration and Model Strategy Development**
   - Perform exploratory analysis on clinical and molecular datasets, uncovering meaningful patterns.
   - Develop strategies to integrate complex datasets and construct predictive models.
2. **Data Integration and Survival Modeling**
   - Use dimensionality reduction, machine learning, and deep learning techniques to merge clinical and omics data.
   - Build and validate survival analysis models, leveraging AI tools to create accurate patient outcome predictions.
3. **Explainability and Key Factor Analysis**
   - Apply explainable AI (XAI) techniques, such as SHAP or LIME, to interpret model outputs and identify the most impactful clinical and molecular features.

- ○ Highlight key predictors for personalized treatment planning.
- ○ Participate in the biological interpretations of the results through literature search and collaborative discussions

What We're Looking For

We are seeking an enthusiastic candidate with a strong foundation in applied mathematics and artificial intelligence, ready to bring their skills to bear on impactful clinical research.

Ideal qualifications:

- **Educational Background:** Currently pursuing or recently completed a Master's in Applied Mathematics, Data Science, AI, or a related field.
- **Technical Skills:** Proficiency in Python, with experience in machine learning, data manipulation, and data visualization.
- **Analytical and Problem-Solving Skills:** Strong interest in applying mathematical models to complex real-world problems.
- **Mathematical and Statistical Expertise:** Background in survival analysis, dimensionality reduction, or similar statistical modeling is optional.
- **Communication Skills:** Ability to collaborate with interdisciplinary teams and present findings clearly.

What You'll Gain

This internship provides the opportunity to:

- Apply advanced mathematical and AI techniques to clinical data analysis and survival modeling.
- Work within a unique interdisciplinary team at two leading institutions in cancer research and applied mathematics.
- Make a meaningful contribution to personalized cancer treatment strategies.

Application Instructions

To apply, please send us your resume, a cover letter, and a brief description of any relevant projects or experience in AI or data science, as well contact information (name and email) of 1 or 2 references. We're excited to see how your skills in applied mathematics can drive innovation in personalized oncology.

Hakim Benkirane: hakim.benkirane@centralesupelec.fr

Elsa Bernard: elsa.bernard@gustaveroussy.fr

[1] Bernard, E., Tuechler, H., Greenberg, P. L., Hasserjian, R. P., Arango Ossa, J. E., Nannya, Y., ... & Papaemmanuil, E. (2022). Molecular international prognostic scoring system for myelodysplastic syndromes. *NEJM evidence*, *1*(7), EVIDoa2200008.

[2] Cantini, L., Zakeri, P., Hernandez, C., Naldi, A., Thieffry, D., Remy, E., & Baudot, A. (2021). Benchmarking joint multi-omics dimensionality reduction approaches for the study of cancer. *Nature communications*, *12*(1), 124.

[3] Harrell, F.E. (2001). Introduction to Survival Analysis. In: Regression Modeling Strategies. Springer Series in Statistics. Springer, New York, NY. https://doi.org/10.1007/978-1-4757-3462-1_16.