COLUMBIA | ALLIANCE
COLUMBIA UNIVERSITY | ÉCOLE POLYTECHNIQUE
SCIENCES PO | PARIS 1 PANTHÉON-SORBONNE

20 Years Ans

ALLIANCE RESEARCH INTERNSHIP PROGRAM

COLUMBIA UNIVERSITY

ACADEMIC YEAR 2024-2025

DEADLINE TO APPLY:

13 DECEMBER 2024

COLUMBIA | ALLIANCE
COLUMBIA UNIVERSITY | ÉCOLE POLYTECHNIQUE
SCIENCES PO | PARIS 1 PANTHÉON-SORBONNE

20 Years Ans

Created in 2002, the Alliance Program is an innovative joint venture between Columbia University, the École Polytechnique, Sciences Po, and Paris 1 Panthéon-Sorbonne University. Every year, Columbia University offers several student internships in scientific disciplines open to École Polytechnique students. The process for applying to these internships is outlined below.

I. Internship Description

- Students work with a faculty member, who acts as an academic advisor and supervises their research project.
  Internships will start in March/April 2025. The duration, objectives and tasks of the internship will be discussed with the supervisor at the host center or department.

- If a stipend is offered, it will be specified in the internship offer.

- Students are responsible for finding housing.

- *All students are required to apply for a J1 Visa to conduct an internship in the United States*.

II. Applications requirements

- Applicants must include: a CV, a cover letter (1 page), and a letter of recommendation.

- Students must send their application to the Alliance Program: alliance@columbia.edu

- All materials must be submitted in English.

DEADLINE – 13 December 2024
All applications must be sent to alliance@columbia.edu

COLUMBIA | ALLIANCE
COLUMBIA UNIVERSITY | ÉCOLE POLYTECHNIQUE
SCIENCES PO | PARIS 1 PANTHÉON-SORBONNE

20 Years Ans

IRVING INSTITUTE FOR CANCER DYNAMICS (IICD)

1. Faculty Sponsor: Professor Simon Tavaré (Director, Irving Institute for Cancer Dynamics [IICD]; Professor of Statistics and of Biological Sciences)

2. Number of interns: One (1) or Two (2)

3. Type of support offered:

   - ✓ Stipend: $3,250 per month over a four-month period
   - ✓ Access to campus services and facilities
   - ✓ Immigration and visa assistance/sponsor

4. Internship Title:

Stochastic modeling and inference for cell population dynamics

5. Description:

Stochastic models of cell dynamics are often used to identify the relative role of different biological parameters on those dynamics. In some case the stochastic models are simple enough that explicit distributions for observable features of the system may be found, for example as solutions to partial differential equations. On the other hand, models do not have to be very complicated before this is impossible. The usual next step is to resort to simulation-based methods, which can be exploited for prospective questions of the form "if the parameters are like this, the observables would look like this", or the (usually much harder) retrospective question of the form "if the observables look like this, the parameters would be like this". The retrospective questions correspond to statistical inference, and usually suffer from uncomputable likelihoods. Such inference questions require likelihood-free approaches such as Approximate Bayesian Computation.

We have a number of collaborations which involve problems of this sort. One involves using the Drosophila ovary as a model for investigating both germline stem cells and epithelial stem cells (also known as follicle stem cells (FSCs)). These stem cells work together to produce eggs that develop from egg chambers that bud off a structure called the germarium. Our collaborators use an elegant live-cell imaging system to observe FSCs undergoing back-and-forth movements in the germarium and moving anteriorly into the region where a population of nondividing somatic escort cells reside. Put simply, FSCs either divide or move into the FC layer where they are no longer followed. What is observed is the number of labeled FSCs and whether or not any FSCs have become FCs. The aim is to estimate the rates at which divisions and emigrations occur, and to extend the system to allow for division and

death in the FSC compartment, or to allow for more layers. The simplest of these settings is one in which explicit likelihoods may be found, the second one in which the ABC methods will be needed. We have a considerable amount of control and mutant data on which to perform the inference.

Reference: Reilein A et al. (2018) Live imaging of stem cells in the germarium of the Drosophila ovary using a reusable gas-permeable imaging chamber. Nat Protoc. 13: 2601–2614.

6. Skills:

Background in mathematics, statistics, computer science or biological sciences with a strong quantitative component. Experience in R. Prior experience in cancer research is useful, but not required.

7. Additional Information:

The IICD is an interdisciplinary institute located on the Morningside Heights campus of Columbia University and focused on the interplay between the mathematical sciences and cancer research, collaborating across disciplinary boundaries to develop tools and methods that can improve our understanding of cancer biology, origins, treatment and prevention. Our website, at cancerdynamics.columbia.edu, gives an overview of our research teams and our current projects. Knowledge of statistics, basic Markov chain theory, machine learning, PDEs, stochastic simulation and skills in R (or similar) preferred. Introductory-level biology useful but not required.

COLUMBIA | ALLIANCE
COLUMBIA UNIVERSITY | ÉCOLE POLYTECHNIQUE
SCIENCES PO | PARIS 1 PANTHÉON-SORBONNE

20 Years Ans

IRVING INSTITUTE FOR CANCER DYNAMICS (IICD)

1. Faculty Sponsor: Elham Azizi (Herbert & Florence Irving Assistant Professor of Cancer Data Research, Irving Institute for Cancer Dynamics, Assistant Professor of Biomedical Engineering, Affiliated Faculty of Computer Science, Affiliated Member of Data Science Institute)

2. Number of interns: One (1)

3. Type of support offered:

   - ✓ Stipend: $3,250 per month over a four-month period
   - ✓ Access to campus services and facilities
   - ✓ Immigration and visa assistance/sponsor

4. Internship Title:

Exploring generative models for single-cell resolution genomic data analysis

5. Description:

We are interested in exploring zero-shot alternatives to existing conditional generative models for analyzing single-cell RNA sequencing data. This work is motivated by the shortcomings of foundation models for single-cell RNA data and whether we can extend the state-of-the-art methods to the zero-shot setting effectively. This project will involve an extensive literature review of existing methods and a systematic exploration of methods for the task.

6. Skills:

Probabilistic Machine Learning, Deep Learning, Bayesian Statistics, Computational Biology, Software Engineering, Programming, Single-cell transcriptomics analysis.

IRVING INSTITUTE FOR CANCER DYNAMICS (IICD)

1. Faculty Sponsor: Sanja Vickovic (Assistant Professor of Biomedical Engineering in the Herbert and Florence Irving Institute for Cancer Dynamics)

2. Number of interns: Two (2)

3. Type of support offered:

   - ✓ Stipend: $3,250 per month over a four-month period
   - ✓ Access to campus services and facilities
   - ✓ Immigration and visa assistance/sponsor

4. Internship Title:

Partial graph matching for in vivo imaging

5. Description:

Project 1: The ability to generate paired functional (eg, neural spiking) and spatial transcriptomic (ST) data across the same tissue, both in vivo during behavior and in situ following tissue processing, would represent a significant improvement in our ability to understand how neuronal cell state and architecture give rise to complex brain and behavior states[1,2]. It is now conceivable to co-register volumetric immunofluorescent (IF) data measuring transcriptional activity with in vivo functional imaging data collected using a two photon (2P) microscope. We hypothesize that it is possible to align functional and transcriptomic data at nearly single cell resolution and that the resulting single time point transcriptomic profiles can identify specific features that will enable an improved prediction and understanding of spontaneous and evoked neuronal activity in health and disease states. One of the major barriers to such an approach is the difficulty aligning the sectioning plane from which IF images are gathered post mortem with one of the principle planes along which in vivo images were acquired. Successful approaches to this problem have proven laborious and impractical to scale. Thus there is a need for a more flexible computational approach to address this issue. To facilitate this work, it is necessary to develop and validate a tool to align the 2D fluorescent images acquired during a typical ST workflow with the volumetric data that was acquired from the brain of awake, behaving mice. Graph neural networks, and in particular subgraph matching algorithms, offer an appealing framework to solve this problem. The goal of this summer project is to develop and test a general purpose toolbox for 1) loading volumetric two photon imaging data and subsetting serial 2D slices at any angle with respect to the existing planes;

2) graph embedding of a designated target section and serial query sections from from 2D and 3D data, respectively; and 3) development of the framework for an iterative subgraph matching approach to localize any 2D section within a reference volume.

6. Skills:

Machine learning, GNNs, image processing, tensorflow

COLUMBIA SURGERY – PEDIATRIC/CONGENITAL CARDIAC SURGERY

1. Faculty Sponsor: David Kalfa MD, PhD (Director, Pediatric Heart Valve Center, Surgical Director, Initiative for Pediatric Cardiac Innovation Director, Kalfa research lab)

2. Number of interns: Two (2)

3. Type of support offered:

- ✓ Stipend: $3,000 per month
- ✓ Access to campus services and facilities
- ✓ Immigration and visa assistance/sponsor

4. Internship Title:

   Transnational research applied to the care of congenital heart defect

5. Description:

   Device development
   Material development
   Lining valve development
   AI
   Computational modeling

6. Skills:

   Biomedical Engineering
   Mechanical
   Polymer science
   Computational model

7. Additional Information:

   Multidisciplinary team