ATAC-Guided Robust Graph Learning for Reliable Single-Cell Multi-omic Analysis

Mentors: Jingfeng Wang, Zhichao Hou

Required Prerequisites: Basic Python programming; commitment to consistent weekly hours; willingness to learn

Preferred Prerequisites: Linear algebra (matrices, vectors); prior coursework or self-study in machine learning; familiarity with command line / Git; introductory biology or interest in genomics.

Project Summary:

Much of modern biology comes down to comparing cells, which cells are present in a tissue, how they differ, what changes in disease. Single-cell technologies now measure gene expression and chromatin accessibility in the same individual cell, giving two complementary views of cell identity. Turning these measurements into biological insight is largely a computational problem: most analysis pipelines build a “cell map" from gene expression alone, and the resulting maps contain hidden errors from technical artifacts that can quietly mislead biological interpretation.

You'll help develop a computational method that uses chromatin accessibility as an independent check on those maps. Along the way, you'll work with real paired datasets, learn to identify cell types from marker genes, and see firsthand how the choices made at the computational stage shape the biological conclusions one can draw.

The project suits undergraduates curious about how cells work and interested in approaching biological questions through code. Some Python experience helps; linear algebra is a plus but not required.