# Renjie Wu

Boston, MA renjiewu@hsph.harvard.edu +1 (857) 209-1097 linkedin.com/in/renjiewu ren-jie-wu.github.io

## ABOUT ME

Computational biology graduate student focused on machine learning for biomedical research. Strong training in statistical inference and representation learning, solid mathematical foundations and programming skills, and hands-on experience implementing and evaluating deep-learning models. Emphasizes rigorous experimental design, real-world validity and clear interpretation, with an interest in turning complex biological data signals into actionable insights.

## EDUCATION

**Harvard School of Public Health, Department of Biostatistics**, Boston, MA 08/2024 – 05/2026
*M.S. in Computational Biology & Quantitative Genetics*

- GPA: 3.9/4.0; Relevant Courses: Data Science, Advanced Computational Biology, Software Engineering.
- Awards: Harvard Chan Fellowship (2024 & 2025).

**Peking University, School of Life Sciences**, Beijing, China 09/2020 – 07/2024
*B.Eng. in Bioinformatics*

- Major GPA: 3.8/4.0; Relevant Courses: Data Structures & Algorithms, Machine Learning, Statistical Inference.
- Awards: Award for Student Organizations (2022); Qin Wanshun–Jin Yunhui Scholarship (2021).

## SKILLS

**CompBio:** scRNA/scATAC analysis, integration, GRN inference; Protein/DNA language model
**DL:** CNN, GNN, Transformer, VAE, deep RL; PyTorch, scikit-learn.
**Programming & Tools:** Python, R, Bash, C/C++ (basic), SQL (basic); Git, Linux, Docker (basic), AWS (basic).
**Soft Skills:** Clear written/verbal communication, attention to detail, critical thinking
**Languages:** English (Proficient), Mandarin Chinese (Native).

## EXPERIENCE

**Multi-condition Single-cell Data Integration via Optimal Transport** Boston, MA
*Graduate Researcher, Dr. Luca Pinello's Lab, Mass General Hospital* 11/2025 – Present

- Developed a multi-condition extension of single-cell embedding method (SIMBA) to integrate scRNA/scATAC across samples/conditions while preserving condition effects and cell-state structure.
- Implemented a joint objective combining graph embedding construction (ELBO loss), cross-context gene embedding alignment, and OT-based cross-batch cell distribution alignment.
- Built a training and evaluation pipeline in PyTorch for config sweep and benchmarking across methods.

**Protein–Ligand Interaction Modeling and Performance Evaluation** Boston, MA
*Research Intern, Dr. Giulia Menichetti's Lab, Brigham and Women's Hospital* 02/2025 – 05/2025

- Cleaned and curated protein–ligand interaction datasets; implemented bias-aware negative sampling to address label imbalance.
- Built and evaluated predictive models under multiple experimental settings (transductive/inductive splits); assessed performance with AUROC/AUPRC (AUROC=0.920, AUPRC=0.938).

**Single-Cell Deep Learning Model for Predicting Gene Expression Dynamics** Beijing, China
*Research Intern, Dr. Jinzhuo Wang Lab, Peking University* 12/2023 – 05/2024

- Built scalable preprocessing pipelines and a hybrid GNN + Transformer model in PyTorch to predict expression dynamics using scRNA velocity data
- Optimized the model by regularization, hyper-parameter tuning; achieved 35% improvement over linear baselines.
- Applied the trained model to infer gene regulatory networks (AUC=0.78) and to perform downstream analyses (trajectory, enrichment) as biological validation.

**Multi-omics Modeling for Cancer Subtype Prediction** Beijing, China
*Research Intern, Dr. Jinzhuo Wang's Lab, Peking University* 03/2023 – 06/2023

- Supported development, statistical validation and biological interpretation on multi-omics dataset modeling (~9,000 cells); contributed to a co-authored publication in *Briefings in Bioinformatics*.