Wu, Renjie (Roger)

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

ABOUT ME

Computational Biologist with a robust background in machine learning and biology. Strong mathematical basis, 4+ years of Python development experience and 2+ years of hands-on experience developing and applying ML models for applications. Skilled at working in collaborative, interdisciplinary environments. Seeking to leverage domain knowledge, research expertise and problem-solving skills to make real-world impacts in healthcare innovation with AI.

EDUCATION

Harvard University, School of Public Health   Boston, MA
M.S. in Computational Biology   08/2024 – 05/2026 (Expected)
· **GPA**: 3.8/4.0; **Courses**: Data Science, Computational Biology, High Performance Computing, Software Systems
· **Awards**: Harvard Chan Fellowship (2024)

Peking University, School of Life Sciences   Beijing, China
B.Eng. in Bioinformatics   09/2020 – 07/2024
· Major **GPA**: 3.8/4.0; **Courses:** Data Structure & Algorithms, Machine Learning, Statistical Inference, Stochastic Process
· **Awards**: Award for Contribution in Student Organizations (2022), Qin Wanshun – Jin Yunhui Scholarship (2021)

SKILLS

· **Domains:** Data Science, Machine Learning, Software Engineering, Statistical Analysis, Computational Genomics
· **Programming Languages:** Proficient in Python, R | Experienced in Bash, C/C++, HTML, JavaScript, SQL
· **Libraries/Tools:** PyTorch, Scikit-learn, Scipy, Scanpy, Flask, React, Git, Docker, Linux, AWS

EXPERIENCES

Protein–Ligand Interaction Network Classification Model Optimization   Boston, MA
Research Assistant in Dr. Giulia Menichetti's Lab, Harvard Medical School   03/2025 – 05/2025
· Collected, cleaned, and processed protein-ligand interaction datasets, employing network-based negative sampling methods to mitigate label distribution bias.
· Encoded proteins and ligands with pre-trained large language models, yielding significant improvements in the interaction classifier's transductive and inductive testing performance (AUROC = 0.920; AUPRC = 0.938).

Web App Engineering of LLM-Powered Knowledge Management System 🎧 ▶️   Boston, MA
Course Project in Architecting & Engineering Software System, MIT   11/2024 – 12/2024
· Developed independently a full-stack web application for end-to-end deployment using Flask (backend) and React (frontend) for a knowledge management system with a user-friendly interface.
· Integrated Large Language Model (LLM) API and Graph RAG techniques to enable effective information retrieval, enhancing the visualization of knowledge and theoretically improving retrieval accuracy by up to 3 times compared with the non-graph-based approaches.

Single-Cell Deep Learning Model for Predicting Gene Expression Dynamics 🎧 📄   Beijing, China
Undergraduate Researcher in Dr. Jinzhuo Wang's Lab, Peking University   12/2023 – 05/2024
· Developed a machine learning model to predict gene expression changes over time using single-cell RNA velocity data.
· Built scalable data preprocessing pipelines for feature transformation, selection, and dimensionality reduction.
· Designed and implemented a hybrid deep learning architecture combining Graph Neural Networks and Transformer in Pytorch framework, optimized by regularization, hyper-parameter tuning, and ablation studies, achieving a 35% prediction accuracy improvement compared with the linear model.
· Applied the trained model to infer gene regulatory networks, reaching an AUC of 0.78, surpassing prior studies

Multi-omics Deep Learning Model for Predicting Cancer Subtypes 📄   Beijing, China
Research Assistant in Dr. Jinzhuo Wang's Lab, Peking University   03/2023 – 06/2023
· Collaborated with computational specialists, contributing to a publication in *Briefings in Bioinformatics* as a co-author.
· Integrated biological insights to guide both architecture design and biological validation of a deep learning model framework for predicting cancer subtypes and markers from multi-model genomic datasets (~9,000 cells).