244 E. Pearson ST.
Chicago IL 60611

yangyang.li@northwestern.edu
+1-507-606-8822

Yangyang Li

yangyangli.top | github.com/cauliyang | linkedin.com/in/yangyang-liz5

Self-introduction

I am a third-year Ph.D. candidate with a focus on deep learning, highly efficient algorithms, and research software development to tackle complex biological problems. My area of expertise is machine learning-based, data-driven domains. Armed with a passion for applying this knowledge to scientific issues, I am eager to contribute and further my knowledge in a fast-paced professional environment.

Research Experience and Project

Northwestern University — Chicago, US
Ph.D. in Bioinformatics — *June 2022 – June 2025*

- Formulated a deep generative model tailored for sequencing data simulation
- Introduced a graph algorithm to identify non-linear transcripts in long-read data, achieving a 20x speedup
- Crafted a web application for graph algorithm visualization
- Designed a Python interface for a C-based command-line tool, gaining 20% performance boosts

University of Minnesota — Minneapolis, US
Ph.D. in Bioinformatics and Computational Biology — *Sep. 2020 – June 2022*

- Developed a transformer-based deep learning model to predict causality between gene fusion and structural variation
- Pioneered an algorithm to discern non-linear structure variations in transcriptomes
- Conducted a comprehensive assessment of the effectiveness of leading tools for the detection of alternative splicing variants
- Courses (Grade): Advanced Machine Learning (A), Introduction to Data Mining (A), Adv. Algs. & Data (B)

China Agricultural University — Beijing, CN
Master in Crop Bioinformatics — *Sep. 2018 – June 2020*

- Identified pivotal features in 1,400 maize genomics datasets to enhance agronomic traits
- Undertook a study to map the relationship between genetic variations and maize ear attributes in 450 natural populations

Education

Northwestern University — Chicago, US
Ph.D in Bioinformatics. GPA: 3.7 — *June 2022 – June 2025*

University of Minnesota — Minneapolis, US
Ph.D. in Bioinformatics and Computational Biology. GPA: 3.68 — *Sep. 2020 – June 2022*

China Agricultural University — Beijing, CN
Master in Crop Bioinformatics. GPA 3.14 — *Sep. 2018 – June 2020*

Northeast Agricultural University — Harbin, CN
Bachelor of Arts in Agricultural Engineering. GPA 3.04 — *Sep. 2014 – June 2018*

Technical Skills

Languages and Frameworks: C/C++, Python, Rust, Pytorch, Jax, Candle, GGML
Development Stack: Neovim, Git, Numpy, Pandas, Matplotlib, Docker, GitHub Action, CMake, HTML, Gcc, Clang, Linux, LaTeX
Specializations: Algorithm Development, Concurrency Programming, Data Analysis and Visualization, Natural Language Processing

Publications

Fry, J., **Li, Yangyang**, & Yang, R. (2022, 09). ScanExitronLR: characterization and quantification of exitron splicing events in long-read RNA-seq data. *Bioinformatics*. doi: 10.1093/bioinformatics/btac626

Li, Yangyang, & Yang, R. (2023). Pxblat: An ergonomic and efficient python binding library for blat. *bioRxiv*. doi: 10.1101/2023.08.02.551686