Saiyang Na

+1 (650) 334-9184 | Arlington, TX | sxn3892@mavs.uta.edu | github.com/nasyxx | nasy.me |
orcid.org/0000-0002-2662-5063

Experienced Computer Science Ph.D. candidate with expertise in deep learning. Created multiple open-source libraries with JAX and Python, focusing on deep learning and LLM.

Education

University of Texas at Arlington — Arlington, TX
Ph.D in Computer Science, supervised by Dr. Junzhou Huang — *Aug 2021 — present*

New Jersey Institute of Technology — Newark, NJ
Master in Computer Science, supervised by Dr. Xinyue Ye — *Aug 2019 — May 2021*

Central University of Finance and Economics — Beijing, China
Bachelor of Economics, major in Science of Investment — *Aug 2014 — May 2018*

Publications

- **Saiyang Na** et al., (2024), "Segment Any Cell: A SAM-based Auto-prompting Fine-tuning Framework for Nuclei Segmentation", *TNNLS* (In Revision).
- Bing Song, Kaiwen Wang, **Saiyang Na** et al., (2024), "Cmai: Predicting Antigen-Antibody Interactions from Massive Sequencing Data", *Nature Cancer* (To Appear).
- Bing Song, Kaiwen Wang, **Saiyang Na** et al., (2024), "An Artificial Intelligence Model for Profiling the Landscape of Antigen-binding Affinities of Massive BCR Sequencing Data", *bioRxiv*.
- Feng Jiang, Yuzhi Guo, Hehuan Ma, **Saiyang Na** et. al, "GTE: a graph learning framework for prediction of T-cell receptors and epitopes binding specificity", *Briefings in Bioinformatics*.
- Lu Zhang, **Saiyang Na** et al., (2023), "Multimodal deep fusion in hyperbolic space for mild cognitive impairment study", *The 26th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, **Oral**.
- Xinyue Ye, Jiaxin Du, Xi Gong, **Saiyang Na** et al., (2021), Geospatial and semantic mapping platform for massive COVID-19 scientific publication search, *Journal of Geovisualization and Spatial Analysis*.

Experience

Retrosynthesis — 2024
- Achieved state-of-the-art performance in multi-step retrosynthesis on the USPTO dataset by combining GNN and LLM, building upon the LocalRetro framework.
- Participated in the Standard Industries Chemical Innovation Challenge on HeroX, a retrosynthesis competition, and advanced to the semi-finals, placing in the top 10.

Cell Segmentation — 2023 — present
- Developed auto-prompt generation pipeline for SAM achieving SOTA performance in cell segmentation.
- Created live demo at https://segment-any-cell.com.
- Paper under review at *TNNLS*

Protein-Protein Interaction — 2022 — present
- Developed contrastive learning framework for TCR-pMHC binding prediction with SOTA accuracy.
- Built end-to-end protein sequence-to-structure pipeline integrating RoseTTAFold.
- Published in top journals including *Nature Cancer* (2nd round review) and *Briefings in Bioinformatics*
- Open-source implementations: GTE (https://github.com/uta-smile/GTE) and Cmai (https://github.com/ice4prince/Cmai)

Deep Learning in Hyperbolic Space — 2023
- Applied hyperbolic graph neural networks to MCI classification, achieving 92.30% accuracy.
- Published in MICCAI 2023 (Oral presentation). Source code: https://github.com/nasyxx/MDF-HS.
- Developed JAX-based hyperbolic neural networks library: https://github.com/nasyxx/jaxrie.

Zebrafish Segmentation — 2023
- Performed zebrafish juvenile heart segmentation using nnNet.
- Code: https://github.com/nasyxx/zebrafish_seg

Research on Heterogeneous Graph Neural Networks 2022

- Analysis of the evolution process of Heterogeneous Graphs in deep learning, on the graph-related tasks, such as node classification tasks.

uuUNet in Medical Image Segmentation 2021

- Implemented 2D/3D UNet architectures with PyTorch and TensorFlow
- Developed efficient data pipeline reducing processing time
- Created TPU-optimized data augmentation techniques for medical images

Cord Search 2020

- Source: https://github.com/nasyxx/cord_search
- Related paper:
  - Xinyue Ye et al., "Geospatial and semantic mapping platform for massive COVID-19 scientific publication search", *Journal of Geovisualization and Spatial Analysis*
  - https://doi.org/10.1007/s41651-021-00073-y
- Build a web app to search for covid-19 related papers.

Open-Source Projects

Owner, nadl (https://github.com/nasyxx/nadl) 2023 — present

- A deep learning framework based on JAX and Equinox, which includes custom model implementations, JAX and GPU-compatible dataloaders, an Equinox trainer, and various JAX-based helper utilities.

Owner, jaxrie (https://github.com/nasyxx/jaxrie) 2023

- A JAX hyperbolic neural networks library

Owner, naipyext (https://github.com/nasyxx/naipyext) 2019 — present

- An IPython Extensions, which includes a better trace exception and auto performance process and CPU timer.

Teaching and Researching Assistance

- Research Assistance of Dr. Xinyue Ye, 2019 to 2020
- Research Assistance of Dr. Junzhou Huang, start from 2025
- CSE1310, Introduction to Computers & Programming, Fall 2021, Spring 2022, Fall 2022
- CSE5311, Design and Analysis of Algorithms, Fall 2023, Fall 2024
- CSE5324, Software Engineering, Summer 2022
- CSE5334, Data Mining, Summer 2024
- CSE6392, Special Topics in Deep Learning, Spring 2024

Skills

- **Expert**: Python, JAX with Equinox, PyTorch, NumPy
- **Proficient**: TensorFlow with Keras, Lisp (Emacs Lisp), LLM
- **Familiar**: Haskell, JavaScript, C/C++

Internship

Data Analyst 2017 — 2018

Cihon Technology Co., Ltd, Beijing *Beijing, China*

- Our team mapped the route and found the coincident points. We analyzed the right path, corrected the real-time direction to match the track and the bus's designated route, and realized the bus's real-time position.

Data Analyst 2016 — 2017

Power Xene Digital Technology *Beijing, China*

- Participated in Establishing real time advertisement/commercial bidding (RTB) model.
- Built target people labeling system and made classification with logistic regression.
- The model was well applied into company's practices.