# Xiran "Derek" Wang

+1 919-579-1665 | xiran.wang@duke.edu | linkedin.com/in/xiran-derek-wang | US Citizen

## Education

**Duke University** *Aug. 2024 – May 2028*
*Bachelor of Science in **Physics** and **Mathematics*** Durham, NC
*Minor in **Statistical Science***
*GPA: **4.00/4.00***

**Relevant Coursework** *Anticipated up to Summer 2026*

*Graduate:* Real Analysis (anticipated), Abstract Algebra, Mathematical Modeling (anticipated), Predictive Modeling and Statistical Learning, Minicourse on Introductory Differential Geometry (anticipated)

*Undergraduate:* Mathematics of Machine Learning, Quantum Mechanics, Probability Theory, Classical Mechanics (anticipated), Optics & Modern Physics, Differential Equations, Linear Algebra, Multivariable Calculus, Experimental Physics Laboratory

## Research Experience

**Undergraduate Algorithms Researcher @ Liberata** *Aug. 2025 – Present*
*Liberata, associated with Duke University* Durham, NC

- Develop **graph theoretic metrics** and **algorithms** for interaction modeling, spectral analysis, and metrics computation for a new academic publishing system based on continuous authorship shares
- Implement algorithms based on scalable matrix methods in an open-source **Python library**

**Undergraduate Machine Learning Researcher @ You Lab** *Jan. 2025 – Present*
*Duke University* Durham, NC

- Implement **variational autoencoder**- and **neural network**-based models to predict bacterial growth dynamics and understand antibiotic resistance
- Deploy and analyze domain-informed ML architectures to model **chaotic ODE systems**, specifically targeting improving sloppy parameter estimation
- Formulate **mechanistic models** based on biophysical principles to generate synthetic training data

## Projects and Manuscripts

**A Foundation Model for Microbial Growth Dynamics** *Mar. 2025 – Oct. 2025*
*Duke University* Durham, NC

- Assisted in the development of a variational autoencoder foundation model trained on ∼370,000 microbial growth curves, which learns 8-dimensional latent embeddings of growth dynamics
- Analyzed predictions of absolute abundances in microbial consortia with latent embeddings of relative abundances, which outperformed using raw inputs across training regimes
- Ran reproducibility validation tests on the project code base

**Associated manuscript (submitted for publication)**

"A foundation model for microbial growth dynamics," *Nature Methods* (submitted, 2025),
With Z. Holmes, I. Shyti, A. Hoffman, K. Duncker, *et al.*
Preprint on *bioRxiv.* doi: *10.64898/2025.12.01.691707.*

**A Graph Model to Detect Communities of Interest in North Carolina** *Nov. 2025*
*2025 Intercollegiate Math Modeling Challenge* Durham, NC

- Developed a graph model for North Carolinian communities based on ACS data and employed clustering algorithms to identify communities of interest among census block groups

- Introduced an interpretable mathematical framework for quantifying similarity between communities and evaluating electoral district maps with a political misalignment metric

**Associated manuscript (submitted to competition)**

"A Graph Model to Detect Communities of Interest and Inform Districting in North Carolina," (2025), With E. Fazal and J. Warrier.
Achieved **Finalist Team** standing.

**Variational Autoencoders for ODE Parameter Estimation** *Feb. 2025 – Aug. 2025*
*Duke University* Durham, NC

- Developed and benchmarked baseline variational autoencoder models in estimating parameters for microbial growth curves under antibiotic stress, trained on synthetic simulation data
- Integrated the model training and evaluation process into a modular Python pipeline to enhance reproducibility

## Professional Experience

**AI Development Intern @ Lenovo** *June 2025 – Aug. 2025*
*Lenovo, Cloud Service Providers* Morrisville, NC

- Assembled a **high-performance computing** and storage **AI cluster testbed** ground-up, installing and troubleshooting server hardware and software
- Investigated optimization strategies for **congestion control algorithms** in distributed AI cluster networks, targeting high-throughput training for LLM workloads
- Wrote and deployed **Python** automation scripts via **Ansible** for validation, benchmarking, and reproducibility

**Co-Director @ HackDuke** *Sept. 2024 – Present*
*Duke University* Durham, NC

- Lead the **Design Team** in creating graphics, merchandise, and the annual website for Duke's premier annual "Code for Good" hackathon (**300+ participants**, **55 projects** in 2025)
- Collaborate with fellow co-directors in managing logistics, budgeting, recruitment, and outreach

## Teaching Experience

**Grader for Math 466: Mathematics of Machine Learning** *Spring 2026*
*Duke University Math Department* Durham, NC

**Grader for Math 221: Linear Algebra (proof-based)** *Fall 2025*
*Duke University Math Department* Durham, NC

## Activities and Involvements

**Research Committee Lead** *Oct. 2025 – Present*
*Duke University, Statistical Science Majors Union* Durham, NC

- Run an undergraduate reading group on theory/methods in statistics, with a focus on foundational high dimensional data and machine learning techniques
- Organize a mini-seminar series on modern statistics research, targeting undergraduate audiences

**Freelance Graphic Designer and Photographer** *Sep. 2022 – Present*
*Duke University & Cary Academy* Cary, NC

- Design graphics for t-shirts, posters, websites, and social media with Photoshop and Figma; recreational nature/event photography with a Nikon D3200
- Worked with the Cary Academy Athletics Department, Communications Department, and varsity teams
- Staff photographer for Duke Chronicle, volunteer photographer for the Sarah P. Duke Gardens, designer and event photographer for the Duke Math Union

## Honors

**Finalist Team** *Nov. 2025*
*2025 Intercollegiate Math Modeling Challenge*

**Renssalaer (RPI) Medal with $160,000 Merit Scholarship** *May 2024*
*Cary Academy and Renssalaer Polytechnic Institute*

## Technical Skills

**Programming Languages**: Python, C++, R, Java, TypeScript, HTML/CSS
**Machine Learning and Data**: PyTorch, TensorFlow, scikit-learn, SciPy, pandas, numpy
**Research Tools**: LaTeX, Git, Linux, SLURM, high performance computing
**Spoken Languages**: English (bilingual), Mandarin Chinese (bilingual), French