

Xin (Allen) Wang

805-259-5932 | xin.wang.1@vanderbilt.edu | [Website](#)

EDUCATION

University of California, Santa Barbara

Goleta, CA

B.S. Statistics and Data Science, Overall GPA: 3.88/4.0, Major GPA: 3.92/4.0

Aug. 2019 – Jun 2023

- Graduation with College Honors

Vanderbilt University

Nashville, TN

M.S Computer Science, GPA: 4.0/4.0

Aug. 2023 – May 2025

- Graduate Fellowship

RESEARCH EXPERIENCE

Network and Data Science (NDS) Lab | Vanderbilt University

Aug. 2023 – Present

- Research areas: *Data-driven ML, AI for Biochemistry, Generative Model*
- Developed [WelQrate](#): Defining the Gold Standard in Small Molecule Drug Discovery
- [BioML](#) Challenge: designed binding domains for a Chimeric Antigen Receptor that will bind CD20
- A Data-centric augmentation framework for covariate shift in small drug datasets via Diffusion Model
- Learning Molecules as Cellular Complexes

Geometric Intelligence Lab | UC, Santa Barbara

Jan. 2023 – Sep. 2023

- ECE594N: Contributed to the *Geomstats* package, brain decoding via manifold learning
- Debugged CAN on *TopoModelX* and created a training tutorial for the molecule graphs

Caves Lab | UC, Santa Barbara | Poster, Code

Nov. 2022 – Jun. 2023

- Developed a program for identifying and measuring visual acuity features from 2D bee images

Math Directed Reading Program | UC, Santa Barbara | Poster

Jan. 2022 – May. 2022

- Conducted a study on Universal Approximation Theory and its applications
- Gave an oral presentation of learning outcomes and won the People Choice's Award

EMPLOYMENT

Undergraduate Teaching Assistant

Sep. 2022 – Jun. 2023

UC, Santa Barbara, PSTAT Department

Goleta, CA

- Assisted teaching in Statistical Machine Learning, Regression Analysis, and Big Data Analytics courses
- Held lab/office hours to help students with homework and coding problems

Algorithm Engineer Intern

Jun. 2022 – Aug. 2022

PING AN TECH, Intelligent City Group

Shenzhen, China

- Mined and analyzed monthly macro-data of cities and trained an LSTM-based model
- Cleaned the Policy Database used for the Policy Recommendation System

Algorithm Intern

Jun. 2021 – Aug. 2021

BOSERA FUND, Index and Quantitative Department

Shenzhen, China

- Collected and modeled the quantitative data on daily stock transactions

TECHNICAL SKILLS

Programming Languages: Python, R, C/C++, SAS, SQL

Libraries/Softwares: PyTorch, PyTorch-Geometric, PyMOL, AutoDock, AmberTools,

Relevant Coursework: Computational Structural Biochemistry, Graph ML, Deep Learning, Geometric ML, Advanced ML, Representation Learning, Stochastic Process, Statistical Computing, Real Analysis

SERVICE

- **Subreviewer:** CIKM 2024
- **PC Member:** WSDM2024-MLoG, GTA3-2024

PUBLICATIONS

1. WelQrate: Defining the Gold Standard in Small Molecule Drug Discovery (coming soon)
2. Wang Yu, Nedim Lipka, Ruiyi Zhang, Alexa Siu, Yuying Zhao, Bo Ni, Xin Wang, Ryan Rossi, and Tyler Derr. "Topology-aware Retrieval Augmentation for Text Generation." In Proceedings of the 33rd ACM International Conference on Information and Knowledge Management (CIKM '24), 2442–2452. Boise, ID, USA: Association for Computing Machinery, 2024. <https://doi.org/10.1145/3627673.3679746>.