Yuanyao (Ann) Xiao

San Jose, CA | annxyy25@gmail.com

Education

New York University
M.S. in Computer Science — Sept 2023 – May 2025

University of Wisconsin – Madison
B.S. in Computer Science & Data Science (Dean's List) — Sept 2019 – May 2023

Technical Skills

Languages: Python, C, C++, SQL, Java
Frameworks/Tools: PyTorch, TensorFlow, CUDA, OpenMP
Additional: Rust, R, JavaScript, Spark, Hadoop, Spring Boot, Redis, Vue.js

Professional Experiences

Cancer–Dementia Association Analysis (CDC NHIS Data)

- Built a **Python pipeline** to process **120k+ CDC patient records (2019–2022)** for dementia risk analysis.
- Applied tree-based ML models to identify key predictors such as **age, depression, and diabetes**.
- Improved performance under severe class imbalance using **SMOTE + undersampling**, raising precision from **0.07** → **0.30** and recall to **0.47**.
- Produced demographic insights supporting **reproducible healthcare analytics workflows**.

Student Assistant Platform (App)

- Built a **full-stack platform** serving thousands of UW–Madison students with **Vue.js, Spring Boot, and Redis**.
- Implemented authentication, caching, and performance optimizations, scaling to support **3,000+ active students** with reduced latency.
- Delivered reliable day-to-day services for student operations across multiple campus workflows.

Distributed Transaction Management System

- Implemented a **replicated database system** with transaction lifecycle management (begin, read, write, commit, abort).
- Designed concurrency control with **cycle detection** to ensure consistency under conflicting transactions.
- Simulated **site failures and automated recovery**, validating fault tolerance and reliability across replicas.
- Built a **CLI and automated test runner** to execute and verify transaction workloads at scale.

System Programming

- Built a **Unix shell** with I/O redirection, multi-pipe support, and job control (C, fork/exec/waitpid).
- Implemented a **multithreaded run-length encoder** with thread pool, mutexes, and condition variables.
- Developed a **FAT32-like filesystem** with recovery tools for deleted files and lost metadata.

Web A/B Testing Platform for Conversion Optimization

- Developed a **Flask-based web app** with multiple e-commerce flows (homepage, browse, donate, signup).
- Deployed two homepage variants to evaluate the impact of **UI changes** on user engagement and conversion metrics.
- Implemented an event logging pipeline (visits, clicks, signups) and applied **statistical significance testing** (chi-square) to measure conversion-rate differences.

Geographic Loan Data Analysis

- Processed and analyzed **mortgage loan records** with demographic and financial attributes using **Python and SQLite**.
- Designed an object-oriented data model with a **Binary Search Tree index** to support efficient queries on loan interest rates and applicant profiles.
- Generated and visualized **demographic and financial insights** (age, race distributions, loan-to-value ratios).