Leon Zhao

zhaoleon03@gmail.com • (302) 388-2989

Education

BS Applied Mathematics & Computer Science, New York University — Sep 2022 - May 2025
Selected Coursework: Game Theory*, Topology*, Statistics*, Computational Geometry*, Deep Learning*, PDE, Numerical Analysis, Probability, Analysis, Operating Systems, Computer Architecture, Algorithms (*Graduate level)

BS Applied Mathematics & Computer Science, University of Delaware (Transferred) — Sep 2021 - May 2022
Selected Coursework: Spectral Graph Theory*, Algebra*, Machine Learning*, Combinatorics*, Discrete Math, Calculus, Linear Algebra, Computer Vision, Bioinformatics, Abstract Algebra (*Graduate level)

Experience

Research Assistant, NYU Center For Data Science — New York, NY — Feb 2025 - Current
- Exploring high-frequency control strategies for robotics and alternative approaches to Vision Transformers

Software + Research Intern (Cascade), Codeium — Mountain View, CA — Sep 2024 - Dec 2024
- Designed and evaluated new algorithm for extracting signal for relevant code from codebases
- Deployed new kubernetes instance to manage new context source for improved context retreival
- Heavily optimized algorithm and service across the entire stack to efficiently manage larger codebases
- Help smooth out the Cascade user experience and squashed bugs

Software Intern (Research Platform), Citadel Securities — Miami, FL — Jun 2024 - Aug 2024
- Redesigned and refactored internal CLI used by the entire firm to manage IO with GCS.
- The CLI serves a thousand researchers and several developer teams across two firms with IO usage peaking at 22 Tb/s
- Reduced cold start latency from seconds to microseconds, alone saving up to $10k in compute annually
- Developed native Python, Java, and C++ clients for the IO library
- Created a new Jenkins server to automate testing and deployment onto custom OS images for our worker orchestrator
- Migrated the entire build and testing to bazel

Software Intern (Core/Databases), Snowflake — San Mateo, CA — Jan 2024 - May 2024
- Explored some of the last remaining compile time constant-fold opportunities in customer SQL queries to cut customers' costs
- Optimized ~8% of the remaining slow queries for large enterprise customers reducing total query time by up to 30 seconds
- Built dashboards to analyze remaining queries' runtime and where optimization opportunities exist

Software Intern (Low Latency), Citadel Securities — Chicago, IL — Jun 2023 - Aug 2023
- Low Latency owns the infrastructure for the firms' high frequency trading system, processing millions of orders per second
- Proposed and implemented new, robust communication protocol for internal order communications
- Wrote high performance, generic C++ code to manage the return order flow information where billions in revenue pass through
- Significantly expanded testing suite of the existing internal order flow code for future changes

Software Intern (Alexa), Amazon — Cambridge, MA — Jun 2022 - Aug 2022
- Automated testing of the wake-word detection models by deploying a custom (internal solution) CI/CD pipeline linked to internal server managing 20+ different Alexa device / OS combinations. Built through AWS and fully connected to researcher dashboard

Teaching

Algorithms TA, New York University — Sep 2023 - Dec 2023
- Created homework and exam problems. Graded homeworks and led office hours

Discrete Math TA, University of Delaware — Sep 2021 - Dec 2021
- Led weekly recitation teaching and reviewing new topics. Graded homework and exams

Algorithms Instructor, AlphaStar — Jun 2021 - Aug 2021
- Designed and taught algorithm topics to 30 students concurrently over zoom leading daily lectures

Honors & Awards

Neo Scholar Finalist — 2023 & 2024

MIT Battlecode Finalist (9th) — 2022

Rank 300 Putnam — 2022

USACO Platinum Qualifier — 2021

AIME x2 — 2021

Apple Swift Student Challenge Winner — 2021

Qualify Junior Olympics (100m dash) — 2020