Daniel Zhao

danielzhao02@gmail.com | https://astradzhao.github.io/ | https://www.linkedin.com/in/danzhao02/

Education

University of California, San Diego | San Diego, CA **SEP 2024 - JUN 2026**

M.S. in **Computer Science** **GPA:** 4.00

Relevant Coursework: Deep Learning, Machine Learning, NLP, ML Systems, Artificial Intelligence, Recommender Systems, OS

University of California, Berkeley | Berkeley, CA **AUG 2020 - MAY 2024**

B.A. in **Computer Science** & **B.A.** in **Applied Math** **GPA:** 3.90

Professional Experience

Machine Learning Engineer Intern | **Qualcomm** **JUN 2025 -**

Teaching Assistant | **CSE 234: Data Systems for ML @ UCSD** **DEC 2024 -**

- Construct 3 brand-new programming assignments for students mat-mul optimization, backpropagation using PyTorch and Triton
- Host office hours, grade projects and exams, maintain Ruby & Jekyll course website, assign relevant papers to students for reading

Research Intern | **Hao AI Lab @ UCSD** **APR 2024 -**

- Develop cutting-edge LLM monitoring system, training **SAEs** for new sequence-level early exit strategies in Python & PyTorch
- Configure new CLLM models to be compatible with **qLoRA fine-tuning** on Jacobi trajectories, decreasing memory usage over 3x

Machine Learning Engineer Intern | **Accretional** **JUN 2024 - SEP 2024**

- Design and implement **RAG vector** embedding model to ground LLM with respect to cloud infrastructure
- Create command-line tool in Golang for indexing and vectorization of files, **semantic & lexical search** of keywords
- Fully construct VSCode application in Svelte and TypeScript to help users easily implement and deploy APIs in less than 5 seconds

Software Engineer Intern | **Hinge Health** **MAY 2023 - AUG 2023**

- Overhauled and integrated a tedious (100+ hour) manual user data reset workflow process
- Designed 5+ endpoints and event consumers in TypeScript and Ruby on Rails, restructuring and building new GraphQL queries
- Presented solutions for over 10 UI/UX features in React, linked repos with RabbitMQ event emitting, deployed in Kubernetes

Full Stack Intern | **Cisco Meraki** **MAY 2022 - AUG 2022**

- Spearheaded a full stack project to create uplink ping stat graphs, helping **500,000 customers** visualize wireless trends over time
- Collected ping stats data from over 4 million wireless internet nodes, using Scala scripts to build a grabber and Jenkins to test
- Conducted experiments with global LT & SQL queries across 600 shards to mitigate the size of rapidly-growing tables

Software Engineer Intern*, *Student Tutor | **Breakout Mentors** **JUN 2021 - AUG 2021**

Projects & Research

CoT Reasoning with Sparse Autoencoders, paper under **ICLR 2025** workshop review **JAN 2025**

MusicInterp, interpretability of music using **sparse autoencoders** **JAN 2025 -**

Learner, a full stack **Flask / React** app that prompts & fine-tunes LLMs to help with learning languages **SEP 2024 -**

Convolutional Neural Net, AlexNet, UNet implementation **MAR 2024 - APR 2024**

- Engineered CNN, AlexNet, UNet and variants, including construction of layers, activations, autodiff in NumPy, PyTorch

Insite, **a React, NestJS & TypeScript app that connects high schoolers to professionals** **JUN 2023 - SEP 2023**

- Utilized AWS S3 buckets for file storage, assembled 40+ REST API endpoints for tasks like search, connections, and profile pages
- Built custom Docker containers and PSQL databases, integrating them with CRA frontend & backend repos with networks

Regulus, **a Golang secure file sharing system** **MAR 2023 - MAY 2023**

- Employed over 5 cryptographic tools and libraries like HashKDF, PBKDF, UUID, HMACs for authentication & access revocation

NumC, **a C++ version of numpy matmul**, optimized 40x using loop unrolling, SimD, OpenMP, caches **OCT 2021 - DEC 2021**

Skills

Python, SQL, JavaScript, TypeScript, React, PyTorch, TensorFlow, Java, Golang, Kubernetes, AWS, Ruby on Rails, Scala, Docker, C, C#, C++, RabbitMQ, GraphQL, HTML, CSS, Figma, Jenkins, PHP, Unity, Git / GitLab / GitHub, Unix, Jira, JupyterLab