SEBASTIAN ZHAO

📞 (814)824-9751 ✉ sebbyzhao@berkeley.edu 🔗 linkedin.com/in/sebbyzhao ⌨ github.com/sebzhao

Education

University of California, Berkeley — **May 2025**
Bachelor of Arts in Computer Science and Data Science (GPA: 3.98) — *Berkeley, CA*

Relevant Coursework

- Structure & Interpretation of Computer Programs, Data Structures, Discrete Math & Probability Theory, Designing Information Device & Systems I & II, Foundations of Data Science, Artificial Intelligence, Probability for Data Science, Machine Structures, Principles and Techniques of Data Science, Machine Learning, Optimization

Experience

Berkeley Artificial Intelligence Research — **September 2022 – Present**
ML Research Intern — *Berkeley, CA*

- Built package and pipeline to replicate and simulate human studies using large language models.
- Developing bridging-based ranking in recommender systems, building upon Community Notes from Twitter.
- Proved that optimizing for proxy objectives in recommender systems is roughly equivalent to optimizing for the true objective.

Launchpad Project Leader — **January 2023 – May 2023**
Project Leader — *Berkeley, CA*

- Led a project integrating a novel transformer architecture with latent diffusion models for culinary uses.
- Created and deployed website hosting model inference using our custom model with Next.js and FastAPI.
- Previously trained generative adversarial networks to image-to-image style transfer pets into a data-augmented Pokemon domain 50% faster and 50% more memory efficient than the previous state-of-the-art.

Computer Science Mentors Tech — **November 2021 – Present**
Software Developer — *Berkeley, CA*

- Implemented back-end in Django for word-of-the-day feature, decreasing time to take attendance by up to 5000%, enabling ability to track section attendance for 2000+ students.
- Analyzed 4 semesters' worth of student and mentor data using pandas to evaluate diversity and efficacy of DE&I efforts.

Case Western Reserve University — **April 2022 – Present**
ML Research Intern — *Cleveland, OH*

- Publishing a convolutional neural network pruning paper building upon a previous ICLR submission, beating previous state-of-the-art pruning method's adversarial robustness and benign accuracy by up to 6% accuracy.
- Updated 5 previous high-impact structured pruning methods and developed an adversarial attack pipeline for benchmarking 10+ pruning methods against our state-of-the-art grouped kernel pruning method.

Projects

YouTube Recommender System Analysis | *Python, Node.js, Docker, NetworkX, pandas* — **September 2022**

- Scraped over 10,000,000 recommendations and videos using a distributed system, conducted exploratory data analysis and network analysis using networkX.
- Migrated scraping to AWS EC2 for scalability, tested scraper deployment on Berkeley's computing cluster with parallelization and load balancing.

Mindflow | *Python, TypeScript, React, FastAPI, OpenAI, Hume, Pinecone* — **June 2023**

- Created a mental health chatbot with the ability to analyze emotions in speech and take notes that provide context in later interactions.
- Improved chatbot capabilities with chain-of-thought prompting that incorporate's Hume's speech prosody analysis.
- Integrated pinecone to allow for long term note-taking, retrieved as a pre-prompt based on embedding similarity.

Lunchpad | *Python, React, Next.js, FastAPI, Ray* — **May 2023**

- Designed a custom recipe and image generation model to create fancy versions of food pictures with their associated recipe; deployed model using Ray Serve.
- Devised novel method to condition recipe generation on previously sampled food classification and ingredients.

Technical Skills

Languages: Python, JavaScript, Java, C, C++, Go, SQL, Scheme, RISC-V
Libraries/Frameworks: React, Django, FastAPI, PyTorch, Ray, CSS, NumPy, pandas, matplotlib, seaborn
Tools: Docker, Git, AWS