Owen Feng

510-755-4409 | ofeng@ucsb.edu | owenfeng.me | linkedin.com/in/owenf3ng/ | github.com/ofeng1

Education

University of California, Santa Barbara — Santa Barbara, CA
B.S., Statistics and Data Science — *August 2022 – June 2026*
- GPA: 3.80
- Organizations: UCSB Data Science Collaborative, Chinese Student Union

Experience

Software Engineering Intern, Data Operations | 🌐 — Remote
Vannevar Labs — *October 2024 – April 2025*
- Implementing social media crawling/scraping engine infrastructure, indexing more than 1M documents per day
- Handled data indexing and deduplication in S3 and Postgres to massively improve downstream service scalability
- Generated FAISS index on data, allowing for semantic grouping and querying of data for AI-related workflows
- Collaborated on testing, ensuring robust software performance by writing high-coverage tests, and improving on testing protocols to improve software reliability

Software Engineering Intern | 🌐 — San Francisco, CA
Sonnet AI (Y Combinator S22) — *December 2023 – July 2024*
- Major early contributor to Sonnet, a platform to automatically extract semantic insights from unstructured conversations. Worked with a team of 5 engineers to scale active usership to thousands of teams.
- Helped reimplement frontend component serving and UI/UX flows, leading to 50-70% faster render speed on web assets and approximately 20% less dropoff during onboarding. See linked website (globe icon) for a work sample.
- Implemented fault-tolerant, at least once task queue to handle long-running audio processing models, reducing processing error rates to less than 1%.

Vice President of Operations — Santa Barbara, CA
UCSB Data Science Collaborative — *May 2023 – Present*
- Interfaced with enterprise clients to set up consultation projects and fundraising opportunities. Raised over 5,000 dollars during the 2023-2024 school year.
- Managed day-to-day operations of an organization totalling over 300 members. Coordinated teaching material, project assignments, and team updates regularly.
- Led a team of 4 to build both a lightweight and performant model for predicting heart disease at a campus-wide project showcase for UCSB Data Science professors.

Projects

Predictive Modeling of Heart Disease | *scikit, Python* — May 2024
- Looking at heart disease prediction from an interpretable feature selection perspective. Used dataset with over a dozen features like Old Peak and RBP to understand impact on cardiovascular health.
- Trained Random Forest models to achieve over 94% accuracy on test set while maintaining fast performance.
- Employed techniques like ROC/AUC curve visualization, pairwise feature correlation, and hyperparameter tuning to optimize model performance.

YouTube Query Analyzer | *Postgres, Whisper, GPT, Typescript, pgvector* — March 2024
- Developed a video question answer tool to help users achieve high fidelity understanding of video content quickly.
- Implemented hybrid embedding + full-text similarity search, a ranking mechanism to eliminate useless context from initial page of results, and a planning mechanism to gain more control over the search process.
- Implemented low-latency frontend chat components for streaming model responses according to RESTful streaming paradigm.

Technical Skills

Languages: Java, Python, C, C++, R, SQL, JavaScript, Typescript, NodeJS, React, Ruby
Frameworks: Next.js, Flask, React, AWS, Docker, Kubernetes, Terraform
Developer Tools: Git, Jenkins, Github CI, Webpack
Overall Concepts: Distributed Systems, Databases, Data Visualization, Networking, Full-Stack Web Development, Cloud Infrastructure, Agile, Test-Driven Development, REST API, Data Structures, Version Control, Collaboration