Hank Sha

No visa sponsorship required ✉ hankssha@gmail.com ☎ (510) 320-1116   LinkedIn   Personal Site

WORK EXPERIENCE

AI/ML Engineer Intern   San Francisco, CA
GitRoll   *September 2024 – May 2025*

- Led design and deployment of an LLM-based uniqueness grading system, translating ambiguous contribution quality into measurable signals for 1.4M+ users.
- Built evaluation and benchmarking frameworks in absence of ground-truth labels; incorporated structured senior engineer feedback, improving performance by 20%.
- Analyzed cost–quality tradeoffs across foundation models (ChatGPT, Gemini, Claude), reducing inference costs by 13% and informing infrastructure decisions.

Research Assistant   Riverside, CA (Remote)
UC Riverside   *June 2025 - Present*

- Translated recent ML research into reproducible PyTorch experiments, building structured evaluation pipelines for robustness under distribution shifts.
- Proposed strategic tooling tradeoffs under time constraints, optimizing research velocity over custom engineering.
- Implemented CUDA optimizations, achieving 400% runtime reduction (6h vs. 23–24h projected) for experiment.

Clinical Data Scientist Intern   Fremont, CA (Remote)
ABVC BioPharma   *August 2025 - November 2025*

- Defined analytical scope under ambiguous project direction, researching FDA guidelines and trial workflows to structure modeling deliverables.
- Developed interpretable predictive models (83% AUC-ROC, 100% recall) to assess trial risk and subgroup behavior, supporting operational decisions.
- Presented structured findings directly to executives, reinforcing data-driven trial optimization initiatives.

Venture Scout   (Remote)
Bouken Capital   *Sept 2025 - Present*

- Perform TAM/SAM/SOM modeling and competitive landscape analysis to assess scalability and differentiation.
- Engage founders to evaluate product strategy, deployment approach, and commercialization pathways.
- Synthesize market and technical insights to inform partner-level investment discussions.

PROJECTS & COMPETITIONS

Project Nexus - 1st Place (Civic Engagement & Policy Track), UCSB Datathon 2026   Link

- Architected and deployed an LLM-powered RAG system and chatbot, owning embedding, retrieval, and orchestration design.
- Integrated an AI agent for entity extraction and relationship graph generation from unstructured email data; 1st Place among 30+ teams.

Biotech Intelligence Engine   Repo

- Built an asynchronous data pipeline (Python, LangChain, Celery) to ingest and structure clinical and regulatory data.
- Implemented **LLM-based extraction workflows** to generate structured outputs for downstream analytical use.
- Built full-stack system with **TypeScript frontend** and modular backend architecture.

EDUCATION

University of California, Santa Barbara   **Goleta, CA**
B.S. Statistics and Data Science   *June 2026*

SKILLS

Language: Python, SQL, TypeScript, R
AI & ML: PyTorch, LLM systems, RAG pipelines, model evaluation, scikit-learn, LangChain
Systems & Data: REST APIs, asynchronous processing (Celery), data pipelines, MongoDB, Pinecone