

Hank Sha

San Jose, CA | (510) 320-1116 | hankssha@gmail.com | linkedin.com/in/hank-sha | hanksha.com | Github

EDUCATION

University of California, Santa Barbara
B.S. Statistics and Data Science

Goleta, CA
Sep 2022 – June 2026

EXPERIENCE

Bouken Capital Remote
Venture Scout *Sep 2025 – Apr 2026*

- Sourced 10+ pre-seed to Series A companies into the fund's pipeline by mapping AI founder networks and inbound channels, converting the strongest into 3 intro and first calls with founders.
- Authored 10+ investment memos covering market sizing, competitive landscape, and valuation for pre-seed through Series A opportunities, giving partners a screening basis on 20+ companies without partner-level time spent upstream.
- Built market maps across 4 categories by decomposing pitch decks, product docs, and public traction data into comparable diligence frameworks, surfacing category structure rather than one-off company opinions.
- Recommended passing on the majority of companies reviewed, defending each thesis against partner push-back on market timing, defensibility, and founder-market fit.

UC Riverside Remote
Machine Learning Research Assistant *June 2025 – Mar 2026*

- Ran 12+ controlled experiments across 10K+ topological graphs, isolating which traversal strategies hold up for LLM graph reasoning and scaling a single-graph prototype into an unattended benchmark suite.

ABVC BioPharma Remote
Clinical Data Scientist Intern *Aug 2025 – Nov 2025*

- Pulled my own model before deployment after cross-fold variance exposed an underpowered dataset and a data-leakage defect in my pipeline, redirecting team resources rather than shipping a result that would not replicate.

GitRoll San Francisco, CA
AI/ML Engineer Intern *Sep 2024 – Apr 2025*

- Shipped a production RAG pipeline serving 45K+ users at a developer-analytics startup, owning the core scoring metric end-to-end, direct operating context for evaluating dev tools and AI infrastructure companies.
- Cut LLM unit costs 13% by benchmarking model families on cost-versus-latency and implementing prompt compression, the same inference-economics analysis that determines whether an AI startup's gross margin holds at scale.

PROJECTS

Xeno CLI — open-source auto coding workflow 2026 – present
Python · LangGraph · Docker · Tree-sitter

- Orchestrated 6 specialized agents with local/API routing built to keep proprietary codebases off third-party APIs, building inside the agent-infrastructure category I want to invest in.

Biotech Intelligence Engine — data platform for biotech-catalyst trading 2026 – present
Python · FastAPI · Celery · TypeScript · Docker

- Built a 137K-entry unified feed across 5 sources to test whether clinical-trial catalysts are systematically mispriced, operating a thesis end-to-end from hypothesis to shipped product.

WRITING

Medium — AI Didn't Raise the Bar. It Removed the Ladder 2026
Market-structure analysis of AI collapsing entry-level engineering.

SKILLS TOOLS

Investing Investment memos, market mapping, deal sourcing, technical due diligence, pre-seed–Series A valuation, competitive landscape analysis, Crunchbase, MS Excel

Technical: Python, SQL, TypeScript, LLM systems (RAG, agent orchestration, evals, inference cost optimization), PyTorch, Hugging Face, Postgres, vector databases, GCP, Docker, CI/CD