

# Sean Tai

916-397-1266 | stai2404@gmail.com | <https://www.linkedin.com/in/seanytai>

## EDUCATION

### University of California, Berkeley

Expected May 2027

*B.A. Computer Science and Data Science, Emphasis in Applied Math & Modeling*

GPA: 3.75

- Coursework: Deep Neural Networks, Machine Learning, Agentic AI, Data Structures, Algorithms, Data Science, Optimization, Probability, Data Mining, Numerical Linear Algebra, Discrete Math, Internet, Poker Theory
- Activities: Generative AI at Berkeley, YC AI Startup School, SIG Technology Discovery Fellow

## EXPERIENCE

### Amazon Web Services

May 2025 - August 2025

*Software Development Engineer Intern | AWS Security Hub*

Boston, MA

- Automated cloud infrastructure deployments in pre-production, to generate 24/7 security data, with AWS CDK.
- Cut manual overhead, slashing deployment time from >1 hr to secs, with AWS Lambda and CloudFormation.
- Built multi-account testing pipelines producing 500+ logs/day by simulating customer behaviors in TypeScript.
- Ensured security compliance and cost-efficiency by designing TTL cleanup for vulnerable AWS resources in Java.
- Improved reliability with event-driven scheduling, deploying infrastructure in dependency-ordered batches.

### Block (formerly Square)

February 2025 - May 2025

*Machine Learning Research Engineer*

Oakland, CA

- Developed context management strategies and MCP Server evals for an open-source autonomous coding agent.
- Achieved a 15% F1-score increase over RAG baselines by building a sequential multi-agent LLM context approach.
- Reduced latency through parallelized hierarchical multi-agent RAG experimentation using LangGraph.
- Created benchmarking metrics (popularity, security, quality) to evaluate MCP servers for agent-driven workflows.

### Digs

July 2024 - October 2024

*Machine Learning Engineer Intern*

Vancouver, WA

- Pitched strategic GraphRAG improvements to executives and AI leadership to boost query accuracy and speed.
- Designed ETL pipeline to annotate 5k images using JSON ground-truth files, for LLM training on HuggingFace.
- Rendered knowledge graphs with 12,000 entity relationships through GraphRAG, Azure, and SQL database.
- Optimized dataset generator, cutting processing time by 30%, with batch processing and robust error handling.

### Wonderfil

June 2024 - August 2024

*Software Engineer Intern*

Berkeley, CA

- Developed a Python backend for real-time device health monitoring and automated anomaly alerts.
- Tracked 50 voltage readouts/sec for 4 wires on a Raspberry Pi prototype, by using a multithreading solution.
- Ensured 100% presence verification across 12 devices by developing two APIs and queue-based voltage-averaging.

## PROJECTS

### Vulnerability Detection for Vibe Coders | Anthropic x PearVC Hackathon

- **Top 10** out of 250+ hackers; Built an MCP Server to connect with Claude Code CLI to find risks in a repository.
- Performed static analysis with 100% secret retrieval by building a tool-calling, multi-agentic hierarchical system.

### All-Inclusive Health App | Berkeley x SkyDeck AI Hackathon

- **Finalist** for AI for Good award (\$25k) out of 1,200+ hackers by building an AI app synced with watch data.
- Processed 3M data in 3.5s by building categorization algorithms for LLM health risk diagnosis via Bedrock API.

### Faulty GitHub Commit Prediction | Python, Pytorch, Numpy, Seaborn, Matplotlib

- Detected faulty commits in GitHub repositories by designing a multi-layer neural network in Pytorch.
- Optimized dataloader, achieving 18000+ rows/sec throughput, by preprocessing with OHE, imputation, etc.

## SKILLS

**Programming Languages:** Python, Java, TypeScript, SQL, Javascript, LaTeX

**Libraries:** Pandas, Pytorch, Scikit-learn, Hugging Face, LangChain, Seaborn, matplotlib, NumPy, regex

**Tools:** Git, AWS, Jupyter, Azure, Jira, DuckDB SQL, Google CoLab, Balena, Docker, MySQL

**Concepts:** Backend development, CI/CD Pipelines, Infrastructure as Code (IaC), Agile, DevOps, ML/AI, GenAI