Kyle Lai

San Jose, CA • kylelai@gmail.com • www.kylelai.me • github.com/KylePythons • linkedin.com/in/kylelai-cs

EDUCATION

San José State University
M.S. in Computer Science — Expected GPA: 4.0 / 4.0 | Expected Graduation: May 2028

University of California, Merced
B.S. in Computer Science and Engineering — GPA: 3.83 / 4.0 | Aug 2022 – May 2026

TECHNICAL SKILLS

Programming Languages: Python, C++, C, JavaScript, Swift, HTML, MIPS Assembly, RISC-V
Technologies: React, Flask, FastAPI, SQL, AWS (Bedrock, DynamoDB), Docker, SwiftUI, SpriteKit, WordPress

EXPERIENCE

App Development Intern – Curiosities (ed-tech startup) Jun 2026 – Sep 2026
- Developed animated UI features for ACE, Curiosities' curiosity-driven learning app for kids ages 5–12, transforming static sprites into dynamic character animations
- Implemented resource abuse detection workflows, saving hundreds of dollars in operational costs
- Expanded internal dashboard analytics to provide deeper session tracking and insights

Software Engineering Capstone – E. & J. Gallo Winery Jan 2026 – May 2026
- Engineered a checklist and inspection tool that replaced Gallo's paper-based sanitation audits, backed by a FastAPI/DynamoDB backend supporting 4 configurable inspection types (checklist, visual score, pH level, ATP swab) across facilities and production lines
- Designed an AI assistant on AWS Bedrock that translates manager questions into a constrained, read-only query plan, pulls live submission data, and summarizes it in plain business language
- Wrote Cypress end-to-end tests covering checklist template creation and inspection submission flows

INDEPENDENT RESEARCH

"Visualizing Emergent Concepts in Transformer Hidden States" — full paper Jan 2026 – May 2026
- Analyzed how a language model's internal representation of a prompt changes as it passes through each layer, using dimensionality-reduction techniques (PCA, t-SNE, UMAP) to project those high-dimensional vectors into 3D and visualize how prompts with similar meaning cluster together with depth
- Curated a 104-prompt evaluation set across 4 semantic categories; found UMAP produced the most separable clusters while PCA lost roughly 50% of variance

PROJECTS

RISC-V CPU Simulator Apr 2026 – May 2026
- Developed single-cycle and multi-cycle CPU simulators in Python implementing the RISC-V instruction set architecture, executing programs from raw 32-bit machine code
- Implemented the fetch-decode-execute-memory-writeback datapath, including instruction decoding, a 32-register file, and data hazard detection that inserts stall cycles to resolve read-after-write dependencies

Ambulance Placement Optimization Web App Nov 2025
- Created a web application that recommends optimal ambulance placements using city data (population density, traffic conditions, hospital routes)
- Integrated Google's Gemini LLM to dynamically analyze inputs and generate adaptable, real-time placement recommendations

Drawing Web App Apr 2025 – May 2025
- Shipped a full-stack drawing game with a Flask/SQLAlchemy backend: hashed-password authentication, a lobby system with unique join codes and capacity limits, and an admin dashboard for managing games and players
- Implemented a resolution-independent canvas engine, storing strokes as relative coordinates, supporting pen, eraser, and undo using React, JavaScript, and CSS

RELEVANT COURSES

Database Systems, Full Stack Web Development, Algorithm Design and Analysis, Data Structures, Digital Image Processing, Computer Architecture, Software Engineering, Natural Language Processing