# Kunal Pai

408-620-2339 | pai.kunal05@gmail.com | linkedin.com/in/kunpai | github.com/kunpai | kunpai.space

## Education

**M.S., Computer Science**, University of California, Davis (GPA: **4.0/4.0**) Expected: June 2026

**B.S., Computer Science & Engineering**, University of California, Davis (GPA: **3.8/4.0**) June 2023

## Relevant Skills

**Languages**: Python, C++, C, JavaScript, Java

**ML/AI**: TensorFlow, PyTorch, scikit-learn, LLMs, Prompt Engineering, Ollama, Hugging Face, Multi-agent Systems

**Web/Data**: React, Next.js, Django, Flask, MongoDB, pandas, NumPy, Matplotlib

**Tools**: Git, Docker, Unix/Linux, gem5, Jupyter, LLVM, Clang

## Work Experience

**Graduate Student Researcher, *DavSec Lab @ UC Davis*** Apr 2025 – Present

- Built an automated pipeline for C-to-Rust transpilation using LLMs, with 5 prompt variations, targeting secure systems migration
- Identified Halstead vocabulary as the strongest metric for predicting translation difficulty
- Validated lightweight semantic augmentations (e.g., filename context) that improved functional accuracy by 5%
- Benchmarked state-of-the-art LLMs across 746 C/C++ programs, achieving 70.2% functional accuracy with best prompt design

**Graduate Student Researcher, *DECAL Lab @ UC Davis*** Sept 2022 – Dec 2024

- Developed a 4,500-sample dataset for pairwise code-documentation alignment from 200 open-source Python projects, enabling future research in software maintenance
- Engineered a pipeline for measuring calibration and correctness of large language models for code repair, using Defects4J
- Collaborated in validating efficacy of semantic augmentation of language model prompts for code summarization using precision and recall metrics like ROUGE and METEOR

## Publications (Selected)

**CoDocBench: A Dataset for Code-Documentation Alignment in Software Maintenance**, Pai, K., Devanbu, P. & Ahmed, T., *Mining Software Repositories (MSR) 2025: Data and Tool Showcase Track*

**Calibration and Correctness of Language Models for Code**, Spiess, C., Gros, D., Pai, K., et. al., *International Conference on Software Engineering (ICSE) 2025*

**Automatic Semantic Augmentation of Language Model Prompts (for Code Summarization)**, Ahmed, T., Pai, K., Devanbu, P. & Barr, E. T., *International Conference on Software Engineering (ICSE) 2024*

## Project Experience

**HASHIRU: Hierarchical, Resource-Aware Multi-Agent Framework** March 2025 – Present

*Entrepreneurial Research Venture* *Python, LLMs, Multi-Agent Systems*

- Designed and deployed a multi-agent architecture enabling dynamic, LLM-driven collaboration across diverse tasks
- Implemented task decomposition with intelligent agent delegation based on resource cost models and task specialization
- Engineered autonomous generation of tools and APIs for task execution
- Developed a robust evaluation framework for agent performance across complex, multi-step tasks

**SuperNOVA: Superconducting Graph Accelerator in gem5** Jan 2024 – Present

*Research Project, DArchR Lab @ UC Davis* *C++, Python, gem5, Hardware Simulation*

- Modeling a superconducting graph accelerator in gem5, targeting sparse workloads under cryogenic conditions
- Integrating energy and latency estimates from superconducting literature, and RTL simulations, to improve realism
- Mentoring undergraduate researchers on modeling, benchmarking, and research writing; one co-authored a ModSim 2024 poster

**gem5 Vision** Jan 2023 – Jun 2023

*Resource Discovery Framework* *Next.js, Python, MongoDB, JSON Schema*

- Accelerated resource discovery 20× for 1,200+ gem5 artifacts by optimizing search and categorization logic
- Implemented categorization and semantic versioning across 20+ resource types to streamline retrieval
- Integrated local and remote JSON schemas with MongoDB, improving accessibility for 500+ users