Kunal Pai

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

Education

University of California, Davis — Davis, CA
M.S. in Computer Science — *Sep 2023 – Present*

University of California, Davis — Davis, CA
B.S. in Computer Science and Engineering — *Sep 2019 – Jun 2023*

Publications

Spiess, C., Gros, D., Pai, K., Pradel, M., Rabin, MRI, Alipour, A., Jha, S., Devanbu, P., & Ahmed, T. (2025).
Calibration and Correctness of Language Models for Code — *ICSE 2025.*

Pai, K., Nand, A. & Lowe-Power, J. (2024).
Potential and Limitation of High-Frequency Cores and Caches — *ModSim 2024.*

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

Shah, P., Pai, K., Patel, H., & Ali, A. (2023).
gem5 Vision — *ISCA 2023: gem5 Workshop.*

Pai, K., Qiu, Z., & Lowe-Power, J. (2023).
Validating Hardware and SimPoints with gem5: A RISC-V Board Case Study — *ISCA 2023: gem5 Workshop.*

Research Experience

Graduate Student Researcher — Jun 2023 – Present
DArchR Lab @ University of California, Davis — *Davis, CA*

- Working with Prof. Jason Lowe-Power on computer architecture research.
- Leading a project to research cryogenic computing and superconducting electronics, including developing models for them in gem5.
- Collaborating with a team of 3 researchers to validate hardware and sampling techniques with gem5.
- Mentoring 5 undergraduate students in the Davis Computer Architecture Lab.

Technical Skills

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

Tools & Technologies: Git, Docker, MongoDB, Django, React, Next.js, TensorFlow, PyTorch, gem5

Projects

Automated Frameworks of Semantic Augmentation to Improve MWP Solving — Apr 2024 - Jun 2024
Machine Learning Project — *Python, NLP, Prompt Engineering*

- Developed a framework leveraging digit-level semantics to improve language model performance in solving mathematical word problems (MWPs) by incorporating context-aware prompt generation techniques.
- Utilized BLEU scores and Levenshtein distance in a demonstration selection algorithm and a prompt selection model to identify the most similar equations for one-shot examples, optimizing prompt accuracy.
- Fine-tuned TinyLlama and prompted PaLM2 with prompt-enhanced GSM8K questions, resulting in improved accuracy on the SVAMP dataset and uncovering insights into the varying context requirements of different mathematical operations.
- Completed as part of ECS 271 : Machine Learning & Discovery at UC Davis.

Effects of Toxicity on Disengagement in Open Source Projects — Jan 2024 - Mar 2024
Software Engineering Project — *Python, GitHub mining, scikit-learn*

- Analyzed factors contributing to developer disengagement in open-source projects, focusing on the impact of toxic communications on contributor retention and project sustainability.

• Utilized sentiment analysis to quantify toxicity within open-source communities and examined its correlation with developer engagement across various programming communities and corporate/non-profit projects.

• Proposed community-driven policies and practical implications for project maintainers, aiming to mitigate negative interactions and enhance collaborative practices in open-source environments.

• Completed as part of ECS 260 : Software Engineering at UC Davis.

Behavior of Spectre in Different Branch Predictors in gem5 Oct 2023 - Dec 2023
Computer Architecture Project *C++, gem5*

• Investigated susceptibility of x86-based in-order and out-of-order processors to Spectre V1 attacks using gem5 v23.0, revealing correlations between Spectre effectiveness and branch predictor characteristics.

• Suggested features for a Spectre-resistant branch predictor, emphasizing extended training periods and mitigation of biased branches to improve processor security against speculative execution attacks.

• Completed as part of ECS 235A : Computer & Info Security at UC Davis.

gem5 Vision Jan 2023 – Jun 2023
Framework *Next.js, Python, MongoDB, JSON Schema*

• Implemented advanced search functionality to enhance resource discovery within the gem5 ecosystem.

• Introduced semantic versioning and built upon a comprehensive resource categorization system for easy access to newer resource versions under specific categories.

• Expanded gem5's database support by integrating local and remote JSON files and MongoDB, improving efficiency and user accessibility.

UNIfy - Course Assistant Jan 2022 – Jan 2022
Hackathon Project *Python, JavaScript, Discord API*

• Utilized the UC Davis Schedule Builder API to extract class timings and professors.

• Formulated a class-based hierarchized dictionary to maintain schedules of over 100 server members in five Discord servers.

• Extracted data from APIs of Rate My Professor and Google Calendar to add additional features to the bot.

Work Experience

Teaching Assistant Sep 2023 – Dec 2023
University of California, Davis *Davis, CA*

• Assisted 180 students in understanding course material and assignments for ECS 132: Probability & Statistical Modeling for Computer Science.

• Conducted weekly office hours and provided feedback on student assignments.

Software Developer Intern & Tech Lead Jan 2022 – Jun 2022
humanID *Davis, CA*

• Coordinated project task completion with global teams, delivering 10 completed projects.

• Co-created documentation on the implementation of a Discord bot that combats spam and fake users.

• Algorithmically hierarchized 100 users and their associated permissions in a Django-based web application.

Technical Product Marketing Intern July 2021 – Sep 2021
SiTime Corporation *Santa Clara, CA*

• Co-authored a product requirements document for a newly-released timing chip in the semiconductor industry.

• Conducted a market survey on optical transceivers and studied potential applications of MEMS timing chips.

• Presented strategy to hierarchize distributors and maximize stakeholders' earned profits by $250,000.