# PRANAV DULEPET

+1(925) 997-0461 ⋄ San Ramon, CA

[ps.dulepet@gmail.com](mailto:ps.dulepet@gmail.com) ⋄ [linkedin.com/in/pranavdulepet](https://linkedin.com/in/pranavdulepet) ⋄ [pranavdulepet.github.io](https://pranavdulepet.github.io)

## EDUCATION

**M.S.E. Computer Science — Human Language Technology**, *Johns Hopkins University* — Aug 2026
*Relevant Coursework*: Natural Language Processing, Human Language Technology, AI Safety/Alignment/Governance

**B.S. Computer Science — ML**, *University of Maryland, College Park* — May 2025
*Honors*: Dean's List; QUEST Consulting Honors Program
*Relevant Coursework*: Systems for Machine Learning, Long-Context LLMs, LLM Privacy & Security, Deep Learning, Machine Learning, Artificial Intelligence, Algorithms, Data Structures

## SKILLS

**Languages/Technologies**  Python, Java, Swift, JavaScript, Git, AWS, closed & open source LLMs
**Libraries/Frameworks**  Triton, TensorFlow, PyTorch, Keras, FastAPI, Pandas, MongoDB, Firebase, React

## EXPERIENCE

**AI Research Intern**, *DataTecnica* — Feb 2026 - Present
- Building AI Agents for biomedical research
- Creating a novel multi-turn benchmark for the biomedical domain
- Distilling "aggregation with thinking" into an open model that generalizes to new tools

**Research Assistant**, *Johns Hopkins CLSP (PI: Benjamin Van Durme)* — Aug 2025 - Present
- Applying confidence calibration to LLM-based search agents

**Software Engineer Intern (AI/ML)**, *Apple* — May 2025 - Aug 2025
- AI/LLM infrastructure and API work for Apple Intelligence, all work is under NDA
- Built prototypes and presented to various teams, and senior leadership

**Technology Policy Researcher**, *Paragon Policy Fellowship* — Sep 2024 - May 2025
- Designed a streamlined AI model approval process for the Santa Clara County Government, reducing redundant labor by consolidating three separate interviews into a centralized questionnaire
- Developed and implemented an AI Usage Guidelines document to assist clients in accurately and comprehensively submitting GenAI applications. [View here, AI Use Case Guidelines for Mitigating AI Risk](#)

**Software Developer Intern (AI/ML)**, *Amazon* — Jun 2024 - Aug 2024
- Developed end-to-end recommendation pipeline using customer-Alexa interaction data (Alexa+ feature)
- Built data pre-processing framework with PySpark for over 60TB of interaction data
- Used Claude 3 Sonnet through AWS Bedrock to generate structured outputs from customer interaction data with an acceptance rate of 94%

**Undergraduate Researcher**, *UMD PIRL (PI: Ramani Duraiswami)* — Jan 2023 - Jan 2025
- Helped develop a factorable attention mechanism reducing transformers' complexity to $O(N)$, inspired by fast multipole and Gauss transform methods. [View here, FAST](#)
- Ensured this still captures complete data relationships, avoiding data loss often seen with similar methods
- Contributed to [The Prompt Report](#), particularly prompt engineering definitions and multimodal sections

**Machine Learning Intern**, *Capital One* — Jan 2023 - May 2023
- Implemented [NMSLIB similarity search frameworks](#) on financial graph embeddings as part of the Enterprise Graph Services Team to detect transaction fraud
- Applied to samples of up to 5 million in size with high-dimensional outputting >90 recall (success rate)

## PROJECTS

**agora.** *Large Language Models, LangChain, Python, SwiftUI, Swift, AWS, MongoDB, Rest APIs*
Developed iOS app and agentic LLM pipeline to provide personalized and affordable meals for students. Adapted Stable Diffusion to generate visuals. Integrated Amazon Fresh and Kroger API to buy ingredients. [(website link)](#)

**College RO** *Swift, SwiftUI, Python, Node.js, Rest APIs, MongoDB, AWS, Google/Firebase Analytics*
Launched CollegeRO on the App Store helping college students find research opportunities, reaching a peak of 1.5k app units. [(app link)](#)