

Aspiring machine learning engineer with strong foundations in computer science and software development

EDUCATION

University of California San Diego <i>Master of Science in Computer Science</i>	Sept 2024 – Dec 2025 3.9/4.0 GPA
University of Virginia <i>Bachelor of Science in Computer Science</i>	Aug 2020 – May 2024 3.8/4.0 GPA
Thomas Jefferson High School for Science and Technology	Aug 2016 – May 2020

COURSEWORK

Data Structures and Algorithms, Operating Systems, Machine Learning, Computer Architecture, Databases, Networks, Distributed Systems, Hardware Accelerators, Robotics, Cybersecurity, LLM and Agent Infrastructure,

EXPERIENCE

EleutherAI <i>Visiting Researcher</i>	Aug 2025 – Current
As part of EleutherAI's SOAR Researcher program, I worked on the automated construction and evaluation of attribution graphs for large language models to improve interpretability and explainability	
Pytorch <i>Open Source Software Engineer</i>	May 2024 – Sept 2024
- Engaged in the development of TorchAO, an architecture optimization library for AI model inference, by opening issues, performing code reviews, and updating documentation - Implemented Activation-aware Weight Quantization (AWQ) which is used by 1000+ models on Hugging Face to decrease model size by 60% while maintaining accuracy on several benchmarks	
Capital One <i>Software Engineering Intern</i>	June 2023 – Aug 2023
- Built and deployed a scalable full-stack cloud application using React, GraphQL, and AWS DynamoDB to create a reddit-style discussion page for all Capital One datasets - Optimized local development build times by decoupling our service and saved 100+ hours of development time	
University of Virginia <i>Teaching Assistant</i>	Aug 2022 – Dec 2022
Led 100+ students in laboratory sessions and office hours by conducting code reviews and peer mentoring	

PROJECTS

LLM Security	Jan 2025 – July 2025
- Uncovered a vulnerability in OpenAI's Deep Research tool which allowed for the discovery of exposed API keys - Applied GCG attacks onto DeepSeek distilled reasoning models showing that test-time inference doesn't inherently improve adversarial defenses - Reproduced emergent misalignment on Gemini-Flash-1.5 which elicited harmful behavior 2.5% of the time - Utilized Kubernetes and VLLM to run large scale evaluations of LLMs across various datasets and attacks, showcasing that reasoning models do not inherently improve adversarial robustness	
TritonTube	Mar 2025 – Jun 2025
- Created a fault tolerant, scalable web server written in Go for serving video content - Utilized RAFT for storing metadata, consistent hashing for scalable storage, and MPEG-DASH for video streaming	
Slider	Mar 2022 – Mar 2023
- Co-developed and published an award winning puzzle game called Slider which has over 10,000 unique players - Developed animation acceleration based on input queue size to maintain responsive gameplay using Unity Engine	

SKILLS

Languages: Java, Go, Python, C/C++, CUDA, Triton, Bash, SQL, C#, JavaScript, HTML, CSS

Other: Github, Docker, AWS, JIRA, Weights and Biases, Kubernetes, Claude, PyTorch, MPI, NCCL, React, Angular, REST, GraphQL, Tailwind, Next.js, prompt-engineering, Copilot, coding agents, Unity