Jie Jeff Xu

Computer Science graduate student interested in all things data and security

jeffxu2018@gmail.com | jjx003.pages.dev/about | github.com/jjx003 | (562) 618-7588

Education

Georgia Institute of Technology — Atlanta, GA
M.S. Computer Science, Machine Learning Specialization — *2022 – 2025*

University of California, San Diego — La Jolla, CA
B.S. Computer Science — *2018 – 2022*

Research Experience

Georgia Institute of Technology, Data to Insights Lab — Atlanta, GA
Ph.D. Student Researcher — *Aug 2022 – Dec 2025*

- Developed a model-agnostic visual-concept explanation system for identifying performance gaps in object detection models.
- Created customized image compression and search system supporting multi-vector and spatial search using concept indices, product quantization, and R-trees.
- Developed keyframe-based video querying system with fuzzy search and user-defined functions.
- Extended the keyframe system with LLM-based natural language query translation via structured output generation.
- Utilized active learning techniques to help users iteratively rerank keyframe query results.

Work Experience

Cribl — San Francisco, CA
Part-time Software Engineer — *June 2025 – Current*

- Developing adaptive MFA solutions for non-SSO customers to secure thousands of customers.
- Diagnosed and fixed various issues in the OpenFGA roles, Auth0 account and tenant metadata, and login validation.

Dolby Laboratories — Remote
Research Intern — *July 2024 – Aug 2024*

- Collaborated with hardware security and financial teams to organize device reporting data.
- Devised internal tool to increase data observability with visualizations and undercounting reports.

CrowdStrike — Remote
Software Engineering Intern (Log Platform Team) — *June – Aug 2021 & 2022*

- Shipped CPU/RAM scaling microservice for thousands of instances, reducing bottlenecks and alerts.
- Reduced technical debt by absolving long-standing race conditions and locking issues in S3 storage.
- Implemented PagerDuty API scraper to forward incident-data and metadata to Splunk and Grafana.

Splunk — Remote
Software Engineering Intern (Team Argus) — *June – Sept 2020*

- Shipped an automated and scalable health-probe in Golang to concurrently check the reporting-health of various Splunk stacks and forward statuses to Prometheus.

Projects

Resource-Efficient LLM Post-Training — Oct 2025 – Dec 2025

- Implemented an accelerated post-training workflow for Qwen-4B, leveraging Unsloths' optimizations like LoRA and quantization to run RLHF on a single 24GB consumer GPU.
- Implemented a semantic data filtering pipeline eliminate redundant training data for better downstream model performance.
- Experimented and compared various RL algorithms like GRPO, DrGRPO, and DAPO.

LLM Quantization Experiments — Oct 2025 – Nov 2025

- Implemented uniform quantization, non-uniform quantization, and power-of-two quantization.
- Applied quantization methods to compare post-training quantization (PTQ) against quantization-aware-training (QAT).

LLM Based Video Clipper — July 2024 – Sept 2024

- Devised a method to find interesting clips from long-form videos by combining OpenAI's Whisper transcriber and Google's Gemini Flash 2.0.
- Created a video pipeline to stitch generated clips together with scene transitions and text subtitles.
- Pitched methodology and video streaming ideas to the startup Overlap for collaboration.

TECHNICAL SKILLS

Languages: Golang, Python, C++, C, Lua, TypeScript, Java, JavaScript, ARM Assembly, SQL
Tools & Frameworks: Redis, Node.js, React, PyTorch, WanDB, CUDA, Docker, Git, Splunk, Prometheus, FFmpeg, LangChain, DSPY

PUBLICATIONS

VCR: Interpretable and interactive debugging of object detection models with visual concepts. **First Author**.
Information Systems (2025): 102652. [Elsevier]

Demonstration of VCR: A Tabular Data Slicing Approach to Understanding Object Detection Model Performance. **First Author**.
PVLDB, 17(12): 4453 - 4456, 2024. [VLDB]

Rethinking Similarity Search: Embracing Smarter Mechanisms over Smarter Data. Third Author.
Preprint 2023. [arXiv]

Hot Pixels: Frequency, Power, and Temperature Attacks on GPUs and Arm SoCs. Third Author.
USENIX Security, pp. 6275–6292, Aug 2023. [USENIX] [arXiv] [CVE]