

Rudy Jin

(510) 610-6881 | rudyjin2@gmail.com | linkedin.com/in/rudy-jin-3932041a7 | Canadian Citizen

EDUCATION

University of Illinois, Urbana-Champaign

Master of Computer Science

August 2024 - December 2025

GPA: 4.0/4.0

University of Toronto, St. George

Bachelor of Applied Science, Computer Engineering

September 2019 - May 2024

Major GPA: 3.77/4.0

PROFESSIONAL EXPERIENCE

Taiwan Semiconductor Manufacturing Company (TSMC)

Full-Stack Software Engineer Intern (BSID - Business System Integration Division)

Hsinchu City, Taiwan

June 2025 - August 2025

- Developed a Sentry-based webpage activity monitoring system with **React** frontend and **Spring** backend, automating dynamic DSN key configuration across **5+** production environments and supporting **800+** daily users.
- Designed the backend system to automate custom product configuration orders; implemented **RESTful APIs** utilizing **TypeScript** that interfaced with **MariaDB** database through **SQL**, handling **300+** customized order requests each day.
- Improved the customer order process by refining server backend validation logic to replace generic exceptions with case specific error messages in **Java**; implemented thread-safe contextual logging for error tracing using **SLF4J** and **MDC**.
- Deployed applications to **Kubernetes** clusters utilizing **GitOps** workflow and **CI/CD** pipelines on Microsoft Azure.

Hong Kong University of Science and Technology

Machine Learning Engineer Intern

Remote

May 2023 - August 2023

- Conducted dataset collection on **CIFAR-10** and custom datasets; performed data pre-processing and augmentation.
- Enhanced text-to-image generation quality by benchmarking and fine-tuning **Stable Diffusion** Models (1.5, 2.1, XL).
- Applied in-context learning on **LLMs** and created prompt adapters, enhancing input prompt elaboration, style customization, and handling of negative prompts; engineered a genetic algorithm to determine the best prompt selection.

Intel

Software Engineer Intern (PSG – Programmable Solutions Group)

San Jose, California

May 2022 - April 2023

- Implemented low-level system algorithms in **C++** to optimize clock alignment logic and resource scheduling tasks, improving overall application performance and reducing end-to-end system latency by **6%**.
- Developed high-performance **C++** memory API modules in firmware systems, optimizing buffer management and synchronization overhead to improve I/O concurrency and throughput by **4%**.
- Created a system-wide validation module for debugging purposes with real-time observability into the I/O subsystem.
- Automated the pipeline of read/write efficiency simulation for different memory controllers and benchmarks in **Bash** and **Python**.

PROJECTS & RESEARCH

Distributed System Development

January 2025 - May 2025

- Constructed decentralized causal-consistent event ordering algorithms in **Go** across distributed nodes with error-handling and failure process recovery; achieved 100% accuracy with lower latency than naive solution (top-ranked in class).
- Implemented **Raft** consensus for leader election, log replication, and entry commit to ensure total ordering execution.

Android World Map Application

August 2024 - December 2024

- Configured a database system on cloud using **FireBase** to manage user log in credentials and application preferences.
- Incorporated dynamic UI customization, API for real-time weather update, and **LLM**-based weather insights in **Java**.
- Created comprehensive instrumented test cases for modules and functionalities, achieving over **96%** statement coverage.

Carrot: Deep Learning for Cloud CDN Cache Optimization

Research Thesis with Professor Mark Jeffrey

Toronto, ON, Canada

July 2023 - April 2024

- Hard-coded **C++** solution based on Belady's Min Algorithm as a benchmark for theoretical optimal performance.
- Designed a custom container class within CacheLib, improving data structures for efficient memory allocation and retrieval.
- Integrated ML models into the caching engine with **Python** wrapper, facilitating execution by the **C++** application.
- Adapted **LSTM** with architectural advancements and applied **Network Compressions** techniques, achieving lower time cost by **3%** and higher hit rate by **4%** compared to conventional LRU replacement with Microsoft and Twitter datasets.