# Allen Wang

908-834-7299 | aw578@cornell.edu | github.com/aw578 | linkedin.com/in/aw578

## Education

**Cornell University, NY** – *M.Eng, B.A in Computer Science* August 2021 - May 2025

- Cumulative GPA: 3.8

**Coursework:** Machine Learning | Distributed Systems | Parallel Computing | Operating Systems | Databases | Algorithms

## Work Experience

**Software Engineer, Roblox** August 2025 - Present

**Software Engineer, Legal Information Institute** January 2025 - May 2025

- Built an end-to-end search pipeline incorporating sparse / dense embeddings and LLM reranking, boosting search accuracy from **70%** to **96%**.
- Collaborated with lawyers to generate a 20,000-question benchmark for accurate search quality evaluations using existing state laws.
- Reviewed and implemented multiple techniques to improve retrieval accuracy, including hypothetical document testing, late chunking, contextual embedding, and prompt engineering.
- Performed user research to identify layman needs and behaviors, then implemented result summarization to improve understanding.

**Software Engineering Intern, Capital One** June 2024 - August 2024

- Designed and implemented a machine learning model, combining similarity search and SQL query generation to enable natural language searches for customer transaction and search data.
- Researched, implemented, and evaluated prompting techniques including chain-of-thought reasoning, constitutional AI, and self-consistency to improve answer accuracy for complex queries by **40%**.
- Finetuned similarity search model using a custom training process to train the model on targeted document attributes, reducing model convergence time from **1 day to 30 minutes** and improving document recall by **83%**.
- Built a synthetic dataset of over **10,000 queries** to establish a company-wide benchmark for evaluating and improving model performance, along with a generation framework to build queries from user-defined templates.

**Teaching Assistant (Computer Systems), Cornell University** January 2023 - January 2024

- Enabled students to copy code locally by migrating assignments to self-hosted servers, improving the student experience, eliminating **180 regrade requests and saving $8,000 per year**.
- Led development of a Docker container management system to generate and monitor student containers, eliminating **30 hours** of manual setup and management time per assignment.
- Held weekly office hours to assist students with course materials and assignments.

## Extracurricular Activities

**Educational Robot Developer, Cornell Cup Robotics** September 2022 - May 2024

- Simplified installation process by Dockerizing client application, reducing onboarding time from **5 hours to <1**.
- Added integration tests with physical robots by rearchitecting the Flask backend to eliminate threading issues.
- Added remote control functionality using a controller support scheme, enabling live demos for project sponsors.

## Projects

**Distributed Key-Value Store** February 2024 - May 2024

- Built a distributed key-value database with sharding, dynamic load balancing and atomic multi-key updates.
- Reimplemented the Paxos consensus protocol to reliably replicate state for shards, achieving a maximum wait time of less than **1 second** under constant dropped messages and repartitions.

**Bookmark Sorter**

- Developed a Chrome extension to automatically manage and sort bookmarks using JQuery.
- Built a machine learning model to sort URLs into folders using text embeddings and hierarchical clustering, then deployed it as an API for the Chrome extension on Firebase.