

Jack Yang

jy3784@nyu.edu | linkedin.com/in/jackyang25 | github.com/jackyang25 | jackyang.fun | U.S. Citizen

EDUCATION

New York University

Bachelor of Arts in Computer Science

May 2025

New York, NY

- Relevant Coursework: Software Engineering, Operating Systems, Parallel Programming, Data Structures

EXPERIENCE

Kura Labs x Google

Machine Learning Intern

Aug 2025 – Present

New York, NY

- Deployed Slack bot with n8n and Docker, encapsulating 5 APIs and 2 agents utilizing PostgreSQL memory
- Designed multi-agent prototypes with OpenAI API, LangGraph, and LangChain, leveraging A2A and MCP

Pulp

Software Engineering Intern

Apr 2025 – Jun 2025

New York, NY

- Reduced service runtime by 80% by parallelizing 16 AWS Lambda functions used for scoring candidates
- Built a mobile-responsive demo page with RAG-based chatbot using Next.js, React, TypeScript, and Groq
- Designed endpoints for ETL pipeline and utilized Pinecone SDK to index vector database used by LLaMA

Angi

Software Engineering Intern

Jun 2024 – Aug 2024

New York, NY

- Increased Google Search click-through rate by 12% and deployed SEO changes through CI/CD pipeline
- Enabled no-code landing page updates for org of 2,800 by integrating 3 microservices with Contentful
- Achieved 2s updates by building REST API in Spring Boot with Redis cache to fetch from GraphQL layer

Dotdash Meredith

Software Engineering Intern

Jun 2023 – Aug 2023

New York, NY

- Improved max page latency with videos by 39%, resulting in increased performance for 200M+ visitors
- Reduced average response time on article pages by 2% to 850 milliseconds by optimizing image loading

Flexible AI-Enabled Mechatronic Systems Lab, NYU

Software Engineering Intern

Jul 2022 – Apr 2023

New York, NY

- Led 20+ member Agile team to prototype waste sorting system using 3 NVIDIA robotic arms and OpenCV
- Developed low-latency motion control algorithms and embedded systems using CUDA, C++, and Python

PROJECTS

High-Fidelity Data Distillation Tool for Language Models

Python, PyTorch, Hugging Face, Scikit-Learn

Mar 2023 – Nov 2023

GitHub

- Created algorithms for corpus size reduction using k-means, cosine similarity, and weighted text scoring
- Boosted precision of SLM (2.7B) from 35% to 72% on Q&A after fine-tuning using 4:1 compression ratio

SKILLS

Languages: Python, SQL, JavaScript, TypeScript, C/C++, Java, Kotlin, HTML/CSS, XML

Libraries & Frameworks: TensorFlow, React, Next.js, FastAPI, Spring Boot, Pandas, JUnit

Services & Tools: AWS, Docker, Linux, PostgreSQL, MongoDB, REST APIs, GraphQL, Git, Jira