Jack Yang

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

EDUCATION

New York University — May 2025
Bachelor's Degree in Computer Science — New York, NY
- Relevant Courses: Software Engineering, Operating Systems, Parallel Programming, Algorithms, Data Structures

EXPERIENCE

Kura Labs x Google — Aug 2025 – Oct 2025
Software Engineering Fellow — New York, NY
- Led 5 engineers to develop full-stack application using FastAPI, Uvicorn, MCP servers, and LangGraph workflow
- Engineered a multi-agent distributed system containerized with Docker and deployed on GCP Compute Engine
- Achieved 100% test coverage utilizing Pytest across unit, integration, system, and end-to-end tests, with logging
- Integrated 3 FastMCP and Node.js services, enabling agent tools like SQL querying on a PostgreSQL database

Pulp — Apr 2025 – Jun 2025
Software Engineering Intern — New York, NY
- Reduced resume analysis runtime by 80% by enabling asyncronous processing for 16 AWS Lambda functions
- Built mobile-responsive website with RAG chatbot using Next.js, React, and TypeScript, hosted on AWS Amplify
- Created REST endpoints for automated ETL pipeline using Pinecone SDK to manage vector database for LLMs
- Optimized database API usage with 90% fewer calls, enhancing scalability and maintaining budget across SDLC

Angi — Jun 2024 – Aug 2024
Software Engineering Intern — New York, NY
- Increased Google Search click-through rate by 12% by modifying JSON-LD schema for SEO and Google indexing
- Enabled Contentful workflow for 2,800 employees by integrating 3 microservices across distributed architecture
- Achieved 2 s updates by building REST API in Spring Boot with Redis caching to fetch from GraphQL data layer
- Authored 10+ unit tests with JUnit and deployed 5 production releases to Kubernetes via Jenkins CI/CD pipeline

People Inc. — Jun 2023 – Aug 2023
Software Engineering Intern — New York, NY
- Reduced max page latency by 39% by lazy loading videos and thumbnails served via Akamai CDN and AWS S3
- Improved avg. page load time at scale to 850 ms for 40M+ visitors, validated through A/B testing and Grafana

NYU FAMS Robotics Lab — Jul 2022 – Apr 2023
Software Engineering Intern — New York, NY
- Led 20+ member Agile team to build waste sorter using 3 NVIDIA robot arms, open-source CNNs, and OpenCV
- Developed low-latency motion control algorithms and embedded Linux applications via CUDA, C++, and Python

PROJECTS

Corpus Optimization Library | Python, PyTorch, Hugging Face, Scikit-Learn, NumPy — GitHub
- Created pruning algorithms to optimize machine learning training using k-means, cosine similarity, and rankings
- Boosted precision of SLM (2.7B) from 35% to 72% on Q&A after fine-tuning using 4:1 dataset compression ratio

SKILLS

Languages: Python, Java, C, C++, JavaScript, TypeScript, SQL, GraphQL, Bash, Kotlin, HTML, CSS, XML
Frameworks & Databases: PostgreSQL, MongoDB, Redis, Spring Boot, FastAPI, Node.js, Next.js, React
Infrastructure & Tools: AWS, GCP, Azure, Docker, Kubernetes, Jenkins, Linux, Git, JUnit, Jira