Junyu Li

junyu1928@gmail.com ♦ linkedin.com/in/junyu1928 ♦ github.com/Junyu06

EDUCATION

HOFSTRA UNIVERSITY Hempstead, NY

Dual Degree in Computer Science

Master (GPA 4.0 / 4.0) Jan 2025 – Dec 2025

Concentration in Artificial Intelligence

Bachelor of Science (GPA 3.8 / 4.0) Sep 2022 – May 2025

Concentration in Networking and Cybersecurity

Queensborough Community College | AS in Computer Science (GPA: 3.8 / 4.0) Sep 2019 – Jun 2022

EXPERIENCE

SafeClick – Phishing Detection System | Winner: 1st Place (Hackathon) Apr 2025 – Present

Lead Software Engineer

- Engineered a multi-stage detection pipeline (URL structure analysis → sandbox evidence → LLM-assisted evidence interpretation) to address high false-negative rates in zero-day phishing detection.
- Implemented code-enforced risk scoring and schema-constrained validation (Pydantic) to override LLM hallucinations and maintain deterministic security logic.
- Architected a domain reputation caching and scan-job deduplication system, significantly reducing inference latency and API costs for concurrent requests.
- Led a 5-engineer team through the full system integration, evolving the project from a winning hackathon entry toward a production-oriented security service.

Belle SPA Inc – Huntington, NY

Applied AI Engineer (Part Time / Internal Tools) May 2023 – Present

- Architected and deployed an on-prem LLM system to replace cloud-based SaaS tools, significantly reducing recurring API costs and ensuring data privacy for sensitive internal workflows.
- Engineered a fault-tolerant Python pipeline with strict schema validation (Pydantic), eliminating brittle free-form outputs and stabilizing daily operations that rely on structured data.
- Optimized on-prem model deployment (GGUF/Ollama) for low-latency, predictable inference, enabling reliable internal use independent of external network conditions or service outages.
- Developed long-context handling and self-repair recovery mechanisms, maintaining production-grade stability across the full system lifecycle from architecture to iteration.

PROJECT

Long-Document LLM Pipeline for Financial Research Sep 2025 – Dec 2025

Software Engineer

- Designed and implemented a MapReduce-style pipeline with semantic chunking to enable stable, bounded-context processing of 20k-word documents for researchers, ensuring consistent outputs under strict context limits.
- Integrated schema-constrained generation (Pydantic) to enforce structured outputs, treating LLMs as unreliable components to prevent silent data corruption from malformed responses.
- Developed idempotent, resumable pipeline stages with self-repair logic, allowing long-running summarization jobs to recover from partial failures without full reprocessing.
- Engineered fast-fail validation and explicit failure boundaries, ensuring high-fidelity data extraction that maintains integrity for downstream system consumption.

SKILLS

Languages: Python, TypeScript, SQL, C++

AI/ML: LLM Orchestration, Pydantic, Semantic Chunking, Prompt Engineering, GGUF, MapReduce, Idempotent pipelines, System Design

Infrastructure: Docker, MongoDB, GCP, Local LLM Deployment, CI/CD, RESTful APIs

HONORS AND ACHIEVEMENTS

1st Place & Most Secure Project, Hofstra-Pensar Hackathon (2025)

3rd Place, Business Competition: Foundations of Leadership & Innovation (Pitched to VCs) (2024)

Provost's Scholars (2024) & Dean's List (2022-2023), Hofstra University