

Yimian Liu

i@yimian.xyz | 607-391-5915 | <https://yimian.xyz/>

EDUCATION

Cornell University

Aug 2022 – Dec 2023

Master of Engineering in Computer Engineering, GPA: 3.95/4.00

Ithaca, NY

Relevant coursework: Deep Learning (A+), Computer Vision (A+), Machine Learning (A), Distributed Computing Principles, Computer Networks, Embedded Operating System, Digital Systems with Microcontrollers, UNIX Tools and Scripting

University of Liverpool

Sep 2017 – Jul 2021

Bachelor of Engineering with Honors in Electrical and Electronic Engineering, GPA: 3.97/4.00

Liverpool, UK

TECHNICAL SKILLS

Programming Languages: Python, TypeScript, JavaScript, C, C++, HTML/CSS, Java, Go, Rust, SQL, Bash, Shell

Tools/Frameworks: React, Node.js, FastAPI, PyTorch, TensorFlow, MySQL, Redis, MongoDB, TCP, gRPC, Docker, Kubernetes, Nginx, Kafka, AWS (EC2, RDS, S3, Bedrock), GCP (Cloud Run, BigQuery, Firebase, Vertex AI), Linux, Git, REST, MCP, A2A

Concepts: Agentic AI, GenAI, LLM, RAG, OOP, Web Development, App Development, Web3, DevOps

WORK EXPERIENCE

Guzman Energy

Oct 2024 – Present

Software Engineer

Denver, CO

- Architect and led development of AI-Enabled Guzman Info System (AEGIS) platform, featuring a React TypeScript frontend, GCP-based microservice backend, and BigQuery data warehouse, used daily by 50+ colleagues.
- Delivered 10+ production applications on AEGIS, such as Transmission Sniper, Least Cost Pathfinder, CAISO Bid, streamlining bid submission and trading strategy optimization for traders and analysts.
- Built an **Agentic AI** assistant using multi-agent design, RAG, and MCP tools, enabling non-technical staff to chat with BigQuery in natural language, auto-visualize via VegaChart, and export CSV, reducing ad hoc SQL requests by 60%.
- Enhanced the Agentic AI with closed-loop output reflection, dynamic retrieval of human-curated prior knowledge, long-term memory, reducing hallucinations and improving multi-step task accuracy by 70%.
- Introduced Cursor IDE to the team, hosted weekly AI frontier paper discussions, proposed and drove adoption of OpenAPI spec, unit test, and CI/CD, fostering continuous learning and AI-first engineering culture.

Elseware Capital

Jan 2024 – Sep 2024

Full Stack Developer (Remote)

New York, NY

- Built from scratch a full-stack real estate trading web app integrated with Polymesh, supporting on-chain asset tokenization, portfolio management, KYC, and live blockchain transactions using React, FastAPI, and MySQL.
- Created a responsive frontend for desktop and mobile using React and Less, incorporating Chart.js for trading data visualization, Leaflet for interactive maps, and 2FA login flows to enhance account security.
- Engineered scalable, low-latency RESTful APIs in FastAPI with JWT auth, async processing, and optimized MySQL query/index, reducing response times by 80% in load tests simulating 1,000+ concurrent users.

Ecesis Investment

Jul 2023 – Dec 2023

Frontend Developer Intern (Remote) (Part-time)

New York, NY

- Delivered cross-platform AI scribing mobile app for iOS and Android using React Native, featuring audio recording, playback, real-time transcription, and LLM summarization, collaborating closely with UI/UX and backend teams.
- Reduced real-time transcription latency by 60% via optimizing bi-directional WebSocket streaming; improved app responsiveness by implementing caching and Redux state management, reducing rendering time by 80%.

PROJECTS

Transformer & Reinforcement Learning: Designed a Transformer model for predator control in boids simulation, pre-trained with supervised learning on rule-based data, fine-tuned using PPO reinforcement learning for emergent strategy against boids flocking dynamics, and distilled into a 40% smaller model achieving <10ms in-browser inference via WebAssembly. [\[Link\]](#)

Machine Learning & Deep Learning: Fine-tuned a DenseNet121-based CNN via transfer learning to classify 9 tomato leaf diseases using the 54K-image PlantVillage dataset, achieving 99% test accuracy and 95.1% accuracy on external datasets, with model performance evaluated via ROC curves, precision-recall, and confusion matrices. [\[Link\]](#)

DevOps & Cloud Infrastructure: Orchestrated a high-availability Kubernetes cluster (3 Debian servers) for containerized web services, leveraging Helm for deployment, Nginx Ingress for load balancing, Prometheus/Grafana for real-time monitoring and alerting, and a centralized ELK stack (Elasticsearch, Logstash, Kibana) for log aggregation and visualization. [\[Link\]](#)

Distributed System & Database: Developed a distributed, sharded key/value store in Java using Multi-Paxos, later optimizing it with a Raft-like consensus algorithm for improved throughput; implemented a two-phase commit protocol to ensure atomic, strongly consistent cross-shard transactions in a horizontally scalable system. [\[Link\]](#)