Ilya Liu

+1 (734) 961-9541 | guyibin01@gmail.com | 616 E. Washington St., Ann Arbor, MI 48104

EDUCATION

Georgia Institute of Technology   2025/08 - Present
B.S. in Computer Science
Rensselaer Polytechnic Institute   2023/09 - 2025/05
B.S. in Computer Science, GPA: 3.9/4.0, Dean's Honor List
Coursework: Data Structures, Computer Org., Principles of Software, Intro to Algo, Foundation of Computer Science

WORKING EXPERIENCE

Beijing Mingguang Zhenduo Data Technology Enterprise   2025/05 - Present
Software Engineer Intern
- Refined the retrieval workflow of a biomedical assistant system to help clinicians surface the most relevant documents from 10M+ records, improving result accuracy by 18% and cutting query time by 40%.
- Modernized a clinical dashboard by migrating legacy WPF views to **Vue.js**, resolving data inconsistencies and reducing average page load latency by over 40% during clinical pilot usage.
- Refactored a domain-specific biomedical search engine by optimizing **FAISS** and **Elasticsearch** indices and redesigning ingestion pipelines, improving top-k retrieval accuracy and ensuring consistent results across large-scale clinical datasets.
- Constructed proactive monitoring and telemetry pipelines with **Prometheus** and **Grafana**, resolving memory leaks via profiling and refactoring to ensure stable model API performance in clinical production.
- Productionized real-time biomedical QA services for clinician decision-making via **KServe** on Kubernetes, ensuring >99.5% uptime and uninterrupted, low-latency access during critical hospital pilot use.

Moyu Corporation   2025/05 - Present
Software Engineer Intern
- Rolled out a behavior-driven recommendation system for a North American resale platform by aligning browsing and purchase patterns with lightweight retrieval logic, elevating homepage relevance and boosting user retention.
- Integrated real-time recommendation results into the **React**-based user portal by dynamically rendering 1.6M + scored listings, enabling personalized homepage feeds and boosting average user dwell time by 18%.
- Optimized a **two-tower retrieval model** to handle burst traffic and noisy catalog queries, leveraging async embedding and lightweight encoders to cut inference latency by 3× and maintain real-time ranking stability in production.
- Built modular data pipelines to refresh user/item features and embeddings, enabling real-time inference and daily updates via unified **PostgreSQL** integration, improving recommendation relevance and model responsiveness.

Lightspeed Technology Enterprise   2024/05 - 2024/08
Software Engineer Intern
- Revamped an internal configuration tool for PCB workflows using **React** and **Zustand**, replacing a legacy UI with a more intuitive layout and real-time updates, reducing operator errors and misconfigurations.
- Refactored brittle configuration endpoints into modular, versioned **Fastify** APIs, decoupling backend logic from UI rendering and enabling seamless, scalable UI updates across agile release cycles.
- Spearheaded a flexible data layer with **MongoDB** to centralize form data and layout metadata in PCB tools, enabling rapid iteration, safe rollbacks, and schema-free updates that significantly reduced UI deployment errors and backend reliance.
- Enforced field-level data validation and role-based access safeguards in MongoDB-backed forms, preventing malformed inputs and unauthorized edits that previously caused UI crashes and misrouted PCB submissions.

SKILLS

Language: Java, C/C++, Python, Javascript, TypeScript C#, HTML/CSS, SQL, Bash, Go, Ruby, Scala
Tech: Linux, Shell, Springboot, Pytest, Django, Spark, MongoDB, Node.js, Kubernetes, AWS, Redis, Pytorch, JavaFX