

Shuolin (Leo) Yin

+1 (647) 674-6127 | leo.yin@mail.utoronto.ca
<https://github.com/leoyin1127> | <https://www.linkedin.com/in/shuolinyin/>

EDUCATION

University of Toronto

Bachelor of Applied Science and Engineering in Computer Engineering + PEY Co-op

- Recipient of Faculty Of Applied Science And Engineering Admission Scholarship
- Recipient of Edward S. Rogers Sr. Admission Scholarship

Toronto, ON

Sep 2023 – Apr 2027

WORK EXPERIENCE

The Institute of Automation, Chinese Academy of Sciences (CASIA)

Research Assistant under Prof. Yang Yang, Prof. Jinlin Wu and Sr. PD Zhen Chen

Beijing, China

Feb 2024 – Sep 2024

- Streamlined research on **Multimodal large language models (MLLMs)** through systematic literature review and novel experimental design, developing methodology for **MICCAI 2025 submission**
- Engineered an **end-to-end data pipeline** for benchmark creation using **Python**, incorporating automated quality checks and custom annotation tools, resulting in **75%** reduction in processing time and **40%** improvement in annotation accuracy
- Developed a **Multimodal evaluation framework** using **PyTorch**, benchmarking **SOTA MLLMs** across **multiple dimensions** with metrics for **accuracy**, **robustness**, and **cross-modal consistency**.
- Spearheaded **weekly technical discussions** and authored comprehensive **research documentation**, facilitating knowledge transfer across interdisciplinary teams of **Vertical category experts** and **ML researchers**

VolunTrack Org.

President/Founder

Toronto, ON

Jun 2022 – Sep 2024

- Founded and scaled a **Non-Profit Organization** through a structured **Four-tier management system**, growing the company to **50+ members** and establishing partnerships with **100+ global non-profits**.
- Architected an **ML-powered volunteer matching system** using **TensorFlow** and **scikit-learn**, achieving **85% matching accuracy** and reducing manual matching time by **70%**.
- Developed a **full-stack volunteer platform** with **React Native** and **React.js**, integrating **MLOps practices** and employing **CI/CD pipeline** to ensure consistent and robust application builds, serving **1000+ monthly active users**.
- Implemented **real-time analytics** using **Google Cloud Platform**, **Cloud Firestore**, and **TensorFlow.js**, deploying **ML models** for volunteer engagement prediction and churn analysis across **100+ organizations**.
- Established **AI innovation program**, leading to **30+ successful ML projects** and **10+ conference speaking engagement**.

PROJECT EXPERIENCE

YiXing - AI-Driven Personalized Travel Planner

AI Engineering Lead / Co-Founder

Toronto, ON

Jun 2023 – May 2024

- Architected and deployed a **production-grade LLM system** with **Fine-Tuned GPT-Based Models**, training on **1,000+** curated travel itineraries to achieve **85% user satisfaction rate** based on post-trip feedback.
- Engineered a **RAG pipeline** processing **5000+** travel documents using **LangChain**, **Pinecone vector DB**, and **custom chunking strategies**, improving travel suggestion accuracy by **35%** and reducing waiting time by **90%**
- Developed **serverless APIs** using **AWS Lambda** to handle **LLM integration**, **user data management**, and **travel preference processing** for the React Native frontend.
- Led a team of **6**, implementing **A/B testing** and **performance monitoring** to continuously improve fine-tuned model.

ReassurED - AI-Powered Emergency Care Guidance/Navigator

AI/ML Project Lead

Montreal, QC

Jan 2025

- Architected a **AI-powered healthcare system** using **Deepseek-v3**, **React Native (Expo)** and integrated real-time hospital data via **Firestore**, achieving **90%** alignment with standard **Emergency Severity Index** during a 24 hours hackathon.
- Implemented **Few-Shot Learning** through **medical examples**, **JSON schema validation**, and **Chain-of-Thought reasoning prompt engineering**, reducing incorrect classifications by **25%** compared to zero-shot baseline.
- Developed a **hybrid recommendation system** combining LLM outputs with weighting algorithm for hospital suggestions, integrating **real-time wait times**, **distance metrics**, and **facility capabilities** with **70%** reduce of user deciding time
- Built scalable **FastAPI** backend processing **100+ concurrent requests/second**, enabling real-time integration between the LLM service and hospital data system with **99%** uptime.

IEEE UoTT

Web Team Associate

Toronto, ON

Jun 2024 – Present

- Developed and maintained **full-stack web applications** with a **React.js frontend** and **Django backend**, enhancing IEEE UoTT's main website and event platforms with a **responsive, user-friendly design**.
- Improved performance by **optimizing APIs and caching**, reducing load times and enhancing **user experience**.
- Collaborated within an **Agile team** to build and deploy new features and ensure a **seamless experience**.

SKILLS

Programming Languages: **Python**, JavaScript, C/C++, C#, HTML, Swift, **MATLAB**
Frameworks & Libraries: **TensorFlow**, **LangChain**, **OpenCV**, **PyTorch**, **React**, **Node.js**, **Django**
Cloud & DevOps: **Amazon Web Service**, **Google Cloud Platform**, **Firebase**, **Git**, **CI/CD**
Development Tools: **Vertex AI**, **Github**, **Xcode**, **Unity**, **Fusion 360**, **Blender**, **AutoCAD**