

LEO XU

+1(236) 838-8957

◇ shenxiaoze.xu@mail.utoronto.ca

◇ github.com/leo-oxu

EDUCATION

Bachelor of Applied Science & Engineering, University of Toronto

Sep 2021 - Apr 2026

Major in Engineering Science – Machine Intelligence + PEY Co-op

CGPA: 3.78/4.0

Relevant courses: Machine Intelligence, Software and Neural Networks(TensorFlow, PyTorch); Artificial Intelligence (Reinforcement Learning, Bayesian Inference), Economic Analysis & Decision Making(Accounting)

SKILLS

Programming & Tools: Python, R, PostgreSQL, MATLAB, C, JavaScript, Linux, Bash, Airflow, Git, GitHub, Docker, LaTeX

Data & ML Libraries: NumPy, Pandas, Matplotlib, Seaborn, Scikit-learn, PyTorch, TensorFlow, JAX

LLM & Systems: Exploring RAG pipelines and multi-agent LLM workflows using LangChain and LangGraph; currently learning Dagster, MLflow, Prometheus, and Grafana for MLOps

Other Software: Microsoft 365

PROFESSIONAL EXPERIENCES

Technical Trainee, City of Toronto Transportation Data & Analytics Unit

Sep 2024 - Present

- Built and automated ETL pipelines with Python (pandas, NumPy) and Apache Airflow, processing around 200 large JSON files weekly from Toronto Police Service collision reports.
- Cleaned and normalized structured and nested JSONB data, and ingested it into a PostgreSQL database to support analysis and reporting workflows.
- Parsed a complex internal PDF specification to extract code-to-meaning mappings, and built a structured reference table to decode raw categorical fields (e.g., weather, traffic control, road surface) during ingestion.
- Conducted performance evaluation of three radar and one video traffic detector at Don Mills Rd & Overlea Blvd using Jupyter, Seaborn, Matplotlib, and psycopg, analyzing detection accuracy across travel directions, vehicle types and weather conditions.
- Designed reusable SQL queries and exploratory notebooks to support data quality checks and facilitate internal review.

PROJECTS

Developer, LLM-Based Travel Recommendation System

March 2025 - Ongoing

- Designing a modular LLM pipeline with Dagster, FastAPI, Docker, and Kubernetes to serve grounded, queryable travel recommendations.
- Integrating Retrieval-Augmented Generation (RAG) for context-aware responses and building multi-agent workflows using LangChain and LangGraph.
- Tracking experiments via MLflow, monitoring system performance with Prometheus + Grafana, and focusing on end-to-end MLOps best practices.
- Project emphasizes open-source data, reproducibility, and scalable infrastructure design for applied LLM systems.

Researcher, Image Super-Resolution with Custom GAN, UofT

Jan 2024 - April 2024

- Built a custom GAN in PyTorch to upscale blurred 512×512 face images to 1024×1024 using the FFHQ dataset in a team of 3.
- Designed Generator with PixelShuffle and residual upsampling; trained with BCE loss and manual early stopping every 30 epochs.
- Evaluated Generator output using PSNR on a validation set; implemented full data pipeline with custom transforms and GPU loaders.