

Hanfu Liu

☎ (647)975-5980

✉ liuhanfu1@outlook.com

🌐 [linkedin.com/in/hanfuliu](https://www.linkedin.com/in/hanfuliu)

🚀 Skip the Read - See What I Have Built at hanfuliu.com/demos 🚀

QUALIFICATIONS

AI & Machine Learning:	• LLM Integration, LLM Application Development, TensorFlow, Scikit-learn, RAG Pipelines, Prompt Engineering, ML Models
Languages & Frameworks:	• Python (PySpark, Pandas, FastAPI, Flask), SQL (PostgreSQL, SQL Server), JavaScript/TypeScript (Next.js, React), C/C++, R
Data Engineering:	• Apache Airflow, Databricks, PySpark, Data Lake Architecture, ETL/ELT Pipelines, Power BI
Cloud & DevOps:	• AWS (EC2, RDS, Lambda, S3, Glue, Amplify), Azure (Databricks, Data Factory, DevOps), Docker, Nginx, CI/CD Pipelines

PROFESSIONAL EXPERIENCE

Data Engineer

Dec 2023 - Present
Oxford Properties Group
Toronto, Canada

- **ITSM AI Automation Platform (Next.js + FastAPI + AWS)**: Architected enterprise AI solution with 3 integrated modules: real-time incident streaming with SSE, LLM-powered troubleshooting assistant, and intelligent ticket classification. Built with OpenAI GPT, Next.js, TypeScript, and FastAPI. Deployed on AWS (EC2, RDS, Amplify) with JWT authentication and rate limiting.
- **Enterprise Data Engineering (PySpark + Airflow + Databricks)**: Designed scalable ELT pipelines processing 400K+ daily records across 20+ ServiceNow tables into data lake and SQL Server. Implemented Airflow orchestration with incremental watermarks, parallel ingestion, and LLM-powered anomaly detection.
- **Budget Management Platform (FastAPI + React)**: Built enterprise solution managing \$10M+ annual budget, replacing spreadsheet workflows with intuitive web application. Delivered accessible, user-friendly interface with real-time dashboards, reducing budget processing time by 40% and enhancing audit capabilities.
- **Reverse Engineering & Documentation**: Developed LLM-powered pipeline using GPT-4 to reverse-engineer and document 1,000+ legacy SQL/ETL jobs. Generated runbooks and technical specifications, saving \$2.4M vs manual documentation.
- **Technical Leadership**: Established CI/CD pipelines using Azure DevOps, managed on-prem deployments, and mentored junior data engineer on software engineering.

Data Engineer (Contractor)

Aug 2022 - Aug 2023
Branchy Solution
Toronto, Canada

- **Cloud Migration & AWS Architecture**: Led zero-downtime cloud migrations for multiple clients, architecting highly available AWS solutions with optimized resource allocation. Achieved 100% uptime for critical systems while significantly reducing infrastructure costs.
- **ML Job Matching Pipeline (AWS Serverless)**: Architected end-to-end serverless recommendation system using Lambda, Glue, and DynamoDB. Automated resume ingestion, job post aggregation, and ML-powered candidate matching—reducing manual screening effort by 90%.
- **Client Engagement**: Delivered technical presentations on architecture designs and ROI analyses to stakeholders, resulting in contract renewals and expanded partnerships.

Data Analyst (Intern)

Mar 2022 - Jun 2022
SAXO Financial Technology
Shanghai, China

- **ML Bond Yield Prediction**: Built ensemble models (Random Forest + Neural Networks) achieving 20% accuracy improvement for bond yield forecasting.
- **Data Quality Automation**: Developed validation framework improving data consistency from 43% to 92% across Risk Technology systems.

EDUCATION

Bachelor of Math

University of Waterloo
2018-2022

- Double major in Statistics and Combinatorics & Optimization
- Coursework: Machine Learning, Data Structures, Algorithms, Database Systems, Data Analysis

PROJECTS & COMPETITIONS

Protein Location Prediction

Kaggle Competition

- Led team of 4 building ensemble ML models (Random Forest, SVM, Neural Networks) for protein localization. **Top 3 among 38 teams** (graduate & undergraduate).