Qi (Savvy) Liu

Toronto, ON | (647) 848-2699 | liu1693@mcmaster.ca | linkedin.com/in/qiliu5 | github.com/SavvyLiu

Education

McMaster University | Bachelor of Engineering, Software Engineering (Co-op) Graduating April 2026

- **Cumulative GPA 3.8/4.0,** Dean's Honor List
- **Relevant Coursework:** Object Oriented Programming, Data Structures and Algorithms, Linear Optimization, Dynamic Systems and Control, Software Testing and Requirements, Probability and Statistics, Computer Architecture

Professional Experience

Data Engineering Intern | Travelers Canada | Toronto, Ontario May 2024 – August 2024

- Built ETL pipelines using Python and Pandas, extracted data from Oracle and MS SQL Server databases to automate data aggregation and validation of company-wide financial data, achieving an **80% speedup** over previous methods
- Conducted extensive end-to-end and unit testing using Python and VBA to ensure consistent results between baseline manual test results and pipeline output, shrinking **discrepancy rates below 1%**
- Integrated pipeline into daily QA workflows as a **key component of the company's tooling** for data validation

Projects

Data Integration Team Lead | McMaster Solar Car | McMaster University September 2024 – Present

- **Leading a 4-member team** within a multidisciplinary project of over 100 members, managing SQL databases and developing Python-based data pipelines to optimize vehicle speed, steering, and battery management in real time
- Directing collaboration with software sub-teams to stream, store, and analyze live data from vehicle DAQ systems using ZeroMQ sockets and CAN Bus integration, providing the drivers with a seamless **response in under 5ms**
- Migrated local databases to cloud-based InfluxDB servers, **improving data access speeds by 65%**

Histology Slide Segmentation Model | Natale Lab | Queens University February 2025 - Present

- Partnering with a research team specializing in fetal development, applying machine learning techniques to aid manual histology slide processing, reducing time per slide from ~30 minutes to ~5 minutes, (**600% speedup**)
- Created an end-to-end pipeline to process placental slides and train a U-Net segmentation model, achieving **98% accuracy** in blood space identification, validated using the Jaccard Index to compare against ground truth
- Exploring techniques to consolidate results of different segmentation methods to further enhance accuracy

Character Recommendation Algorithm | Personal Project | Source Code January 2025

- Collected data from Riot Games API using the Requests library, **storing over 50,000 records** in a SQLite database
- Developed and trained a multiclass classification model via Python and scikit-learn, implementing a random forest classifier with hyperparameter tuning to reliably predict character preferences based on player gameplay patterns
- Experimented with and implemented a content-based filtering approach using embeddings and TensorFlow to train deep learning networks, resulting in a **15% improvement in top-5 accuracy** compared to baseline models

Dialect Translation Model | Hack the Valley 9 | Hackathon Winner | Source Code October 2024

- Innovated an application providing accurate translations of dialects unsupported by conventional translation services
- Curated and preprocessed datasets by leveraging Google's API to extract YouTube comments under creators based in relevant dialects and integrating university-provided public data, using Python with PySpark for efficient processing
- Trained a seq2seq ML model on Databricks using TensorFlow, impressing judges and **winning multiple categories**

Technical Skills

Languages: Python, SQL, R, JavaScript, Java, C, C++, MATLAB, VBA

Libraries and Frameworks: Pandas, NumPy, TensorFlow, PySpark, scikit-learn, Requests, Streamlit, Flask

Developer Tools and Practices: Linux, Git, VS Code, Excel, PowerBI, WSL, Databricks, Jira, Agile, Maven, Junit