

David Liu

☎ (+1) 647-674-6834 | ✉ davidsiqiliu@gmail.com | 🏠 david-siqi-liu.github.io | 🌐 [davidsiqiliu](https://www.linkedin.com/in/davidsiqiliu)

Skills

Languages Python, Java, JavaScript, C++, Scala, SQL, R

Certifications AWS Certified Associate Solutions Architect, Fellow of the Society of Actuary (Quantitative Finance)

Experience

Amazon

TORONTO, ON

Machine Learning Engineer

May 2021 - Present

- Maintain and develop core ML infrastructure for a tier-1 service serving **100K+ requests per second**
 - Design and implement ML pipelines for sponsored ads widgets (e.g., highly rated, trending now)
 - Launch A/B tests and analyze results to **increase user clicks and ads revenue**
- Deployed an end-to-end ML pipeline in AWS Sagemaker for **document recommendation**
- Built a full stack, **highly scalable** knowledge sharing web application in AWS from ground up
 - Designed core system components such as edge-optimized REST APIs, automatic failover, load balancing and auto scaling that ensured stable performance during traffic spikes and **fault tolerance** during regional AWS outage
 - Responsible for designing and implementing new features in backend (gRPC server) and frontend (React SPA)
- Developed user monitoring solutions to improve **system observability** and presented at company innovation tech talk
 - Provided metrics on system performance and user engagement to support making data-driven decisions
 - Collaborated with data engineers to build **ETL pipelines** for batch analysis and dashboard display
 - Worked with legal to ensure **security and privacy compliance** with corporate and GDPR policies

Huawei Technologies

TORONTO, ON

Computer Vision Engineer

Sep 2020 - Apr 2021

- Developed hand gesture-based command modules for media control in Android phones and smart TVs
 - Introduced landmark detector using Google MediaPipe that **increased hand tracking stability by 40%**
 - Optimized time-consuming bitmap operations with matrix-based operations, **reducing model runtime by up to 80%**

Munich Re

TORONTO, ON

Data Scientist

May 2020 - Aug 2020

- Implemented a BERT-based **text similarity matching model** in PyTorch to streamline reinsurance underwriting
- Replaced a deterministic medical concept extraction tool with a novel Transformer-based, multi-hop question-answering model, resulting in **3 times the relevant information retrieved**

Capital One

TORONTO, ON

Data Scientist

Sep 2018 - Aug 2019

- Built classification models using credit card transactions to identify profitable customer segments
- Streamlined model monitoring using Databricks and Spark, **reducing reporting time from 2 weeks to 1 day**
- Revamped analytical table ETL process using Jenkins and SQL, winning company-level initiative awards

Projects

Code Review Dynamics and Software Quality in OpenStack

Mar 2021 - Apr 2021

Mined open-source repositories and built models in R to understand code review dynamics and impact on software quality.

Cancer Type Classification using Whole Slide Pathology Images

Mar 2021 - Apr 2021

PyTorch model with permutation-invariant layer (DeepSets) and graph convolution layer (ChebConv).

Distributed Data Deduplication Framework

Mar 2020 - Apr 2020

Spark framework in Scala that orchestrates parallel data deduplication tasks while guaranteeing optimal workload distribution.

Yelp Reviews Sentiment Analysis

Mar 2020 - Apr 2020

Ablation study on several text preprocessing techniques and ML models for their effectiveness in predicting user sentiments.

Real-Time Twitter Streaming

Nov 2019 - Dec 2019

Flink application in Python that streams Twitter data to Elasticsearch and produces real-time analytical dashboards in Kibana.

Education

University of Waterloo

M.Math. in Computer Science

Sep 2019 - Apr 2021

B.Math. in Statistics

Sep 2011 - Apr 2016