

David Liu

COMPUTER SCIENCE · STATISTICS

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

Skills

Languages Python, Java, C++, Scala, SQL, R, SAS, JavaScript, HTML
Technologies MapReduce, Spark, Flink, Azure, AWS, Databricks, Flask, Docker, Elasticsearch, Tableau, Power BI
Others Fellow of the Society of Actuary (Quantitative Finance track)

Experience

Amazon

Software Development Engineer

TORONTO, ON

May 2021 - Present

Huawei Technologies

Computer Vision Engineer Intern

TORONTO, ON

Sep 2020 - Apr 2021

- Developed hand gesture-based command capabilities in smartphones and smart TVs
 - Refactored development Python codes into production Java codes to run optimally on low-CPU Android devices
 - Introduced hand landmark detector that drastically increased tracking stability and pose classification accuracy
 - Replaced time-consuming bitmap operations with matrix-based operations, decreasing runtime by up to 80%

Munich Re

Data Scientist Intern

TORONTO, ON

May 2020 - Aug 2020

- Developed a BERT-based text similarity matching model in PyTorch to streamline reinsurance underwriting
 - Explored state-of-the-art biomedical word embeddings and Siamese network objective function
 - Deployed the final model as a web API using Flask and containerized using Docker
- Enhanced existing medical concept extraction tool using a novel multi-hop question-answering approach, resulting in 3 times the relevant information extracted. Integrated into production Airflow pipeline
- Refactored several model development and explanatory notebooks into production Python packages

Capital One

Data Scientist

TORONTO, ON

Sep 2018 - Aug 2019

- Built supervised-learning models on customer transactions to identify profitable customer segments
 - Presented model findings and proposed A/B testing strategies to the senior leadership team using Tableau
- Revamped analytical table ETL process using Jenkins and SQL, winning company-level initiative award
- Supervised co-op students on building a model input validation pipeline in R

Milliman

Actuarial Consultant

CHICAGO, IL

Sep 2016 - Aug 2018

- Developed actuarial appraisal models for several multi-billion dollar M&A transactions of life insurance companies

Projects

Code Review Dynamics and Software Quality in OpenStack

Mar 2021 - Apr 2021

Mined Git 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 and Supervised Learning Framework

Mar 2020 - Apr 2020

Spark framework in Scala that guarantees optimal workload distribution in parallelized data deduplication.

Multi-Domain Source Adaptation of Image Classification

Nov 2019 - Dec 2019

PyTorch model that classifies images from an unseen target domain using training images from multiple source domains.

Real-Time Twitter Streaming

Nov 2019 - Dec 2019

Flink application 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 and Actuarial Science, minor in Computer Science

Sep 2011 - Apr 2016