

Curtis Li

🌐 sz7li | 🌐 shao-curtis-li | 🌐 sz7li.github.io | ✉️ crtsz7li@gmail.com | 📞 647-272-6690 | 🇨🇦 Canadian Citizen

SKILLS

Languages and Libraries Python, R, C, C++, SQL, Pandas, NumPy, Matplotlib, PySpark
Tools and Technologies Git, PostgreSQL, Docker, Bash, LaTeX

EDUCATION

University of Waterloo - Bachelor of Mathematics, Honours, Co-op **Jun 2023**
Statistics and Computational Mathematics, Computing Minor *cGPA: 82.18/100*

- Fall 2022 Research Assistant in the Vision and Image Processing Lab, research on Neural Radiance Fields.
- Served on 4-person Orientation Committee and organized Math O-Week 22' for 200+ Leaders and 2000+ new students.

WORK EXPERIENCE

HeadSpin | Engineering Intern | Data Science Team | Manager evaluation: **Outstanding (7/7)** Toronto, ON
Python, NumPy, Matplotlib, PostgreSQL, OpenCV, MVC *May - Aug 2022*

- Researched, developed, and deployed two image analytics features that extends the core platform's image matching functionality to support additional use cases for clients.
- Curated a targeted video dataset, conducted R&D from CV literature, and shipped features to production in Python. **Improved** existing image matching feature by **23%** in F1 score under key scenarios.

ThinkData Works | Data Science Co-op | Data Labs Team | Manager evaluation: **Excellent (6/7)** Toronto, ON
Python, NumPy, Pandas, Matplotlib, PySpark, Hadoop, PostgreSQL, Docker, NLP *May - Aug 2021*

- Identified and implemented improvements for the Spark-based large-scale data de-duplication application. Accelerated development cycles for the team and decreased data computation time for customers.
- Benchmarked performance of multi-threaded sparse matrix multiplication; monitored and gathered data on CPU/memory usage in key use cases, **improving de-duplication computation speed by 22%**.

Canadian Imperial Bank of Commerce | Data Scientist | Data Studio / Enterprise Innovation Team Toronto, ON
Manager Evaluation: **Outstanding (7/7)** *Jan - Apr / Sep - Dec 2020*

Python, NumPy, Pandas, Matplotlib, spaCy, React, Flask, PostgreSQL, NLP

- Developed features for a document summarization Web App using React, Flask, and PostgreSQL, **reducing manual document analysis time** for CIBC Capital Markets stakeholders by 66%.
- Used Google Places API to extract reviews data from CIBC bank branches and performed unsupervised text clustering; created visualizations of topic and sentiment trends in Google Data Studio and presented to C-level management team.
- Developed semantic search functionality in Web App using Sentence-BERT, reducing COVID response-related processing times for customers at CIBC Contact Center.

KPMG | Summer Student | Tax Incentives Practice Team | Manager Evaluation: **Excellent (6/7)** Toronto, ON
SR&ED, Word, Excel, VBA Scripting *May - Aug 2019*

- Assisted Canadian firms in securing tax credits through the government's SR&ED program by compiling technical documentation from client IT projects to identify and articulate SR&ED eligible activities.

DaVinci Assortment Solutions | Technical Security Analyst Co-op | Engineering Team Toronto, ON
Manager Evaluation: **Very Good (5/7)** *Sep - Dec 2018*

Java, Javascript, AngularJS, Spring, PostgreSQL

- Improved security features for the inventory management Web App by upgrading authentication and user management workflows, reducing vulnerabilities related to accessing customer data.

PROJECTS

UWaterloo Subreddit Data Analysis | DEMO | GitHub
Python, SQL, Tableau, PySpark, AWS Glue, EMR, Athena, Lambda

- Extracted data from Reddit's API using Python in AWS Lambda. Transformed and stored data in S3 with Glue.
- Used PySpark, EMR to perform NLP and data analysis on **2 Million+** records. Queried results in AWS Athena and created dashboards to visualize trending keywords in Tableau.