

Curtis Li

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

SUMMARY

Languages and Libraries Python, R, C, C++, SQL, Pandas, NumPy, Matplotlib, PySpark, spaCy, NLTK, OpenCV
Tools and Technologies Git, PostgreSQL, Docker, Bash, Hadoop, LaTeX, AWS, GCP
Skills Data Analysis, Data Visualization, NLP, Software Development, Machine Learning

EDUCATION

University of Waterloo | Bachelor of Mathematics Jun 2023
Statistics and Computational Mathematics, Computing Minor cGPA: 82.18/100

- Fall 2022 Research Assistant at 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

Engineering Intern | HeadSpin | Data Science Team Toronto, ON
Python, NumPy, Pandas, Matplotlib, PostgreSQL, OpenCV, MVC May - Aug 2022

- Collaborated with stakeholders to translate business requirements for new software feature into actionable success metrics.
- Created hypotheses, conducted research within CV literature, and presented visualizations of data from iterative experiments. Communicated cost-benefit analyses to Senior Leadership and guided decision-making on the final approach.
- Shipped 2 new software features to production using Python and OpenCV. **Improved existing feature by 23% in F1 score**, extending the product's support for additional customer use cases.

Data Science Co-op | ThinkData Works | Data Labs Team Toronto, ON
Python, NumPy, Pandas, Matplotlib, PySpark, Hadoop, PostgreSQL, Docker, NLP May - Aug 2021

- Automated processes and implemented improvements for the data de-duplication application using Python, PySpark, and SQL. Accelerated productivity for the team and decreased data computation time for customers.
- Analyzed data from benchmarking the performance of multi-threaded sparse matrix multiplication. Gained insights into optimizations leading to **improvements in computation speed by 22% over 6 Million+ records**.

Data Scientist | Canadian Imperial Bank of Commerce | Data Studio / Enterprise Innovation Toronto, ON
Python, NumPy, Pandas, Matplotlib, spaCy, React, Flask, PostgreSQL, NLP Jan - Apr / Sep - Dec 2020

- Built and delivered an end-to-end NLP application using React, Flask, and PostgreSQL. **Reduced manual document analysis time by 66%**, driving higher productivity for CIBC Capital Markets stakeholders.
- Investigated performance of bank branches and understanding of CIBC branding by conducting social media analytics. Performed text clustering, visualized topic and sentiment trends, and communicated insights to Senior Management.
- Built and delivered an internal search engine Web App for CIBC contact center agents to mitigate high call volumes during onset of COVID. Reduced processing times for customer queries and increased efficiency by developing semantic search capability using Sentence-BERT embeddings.

Summer Student | KPMG | Tax Incentives Practice Team Toronto, ON
SR&ED, Technical Writing, Microsoft Office, VBA Scripting Aug - Sep 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.

Technical Security Analyst Co-op | DaVinci Retail | Engineering Team Toronto, ON
Java, Javascript, Spring, AngularJS, PostgreSQL Sep - Dec 2018

- Improved security features for the inventory management Web App by upgrading authentication and user management workflows. Reduced vulnerabilities in 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.
- Applied NLP and timeseries techniques on **2 Million+** text records with PySpark and EMR to identify trending keywords over time. Queried data in AWS Athena and created dashboards to visualize findings in Tableau.