

Curtis Li

📧 sz7li | 🏠 shao-curtis-li | 🌐 sz7li.github.io | ✉️ sz7li@uwaterloo.ca | 📞 +01 6472726690 | 🇨🇦 Canadian Citizen

SKILLS

Languages and Libraries Python, R, C, C++, SQL, Pandas, NumPy, Matplotlib, Keras, PySpark, Hadoop
Tools and Technologies Jupyter Notebook, Git, PostgreSQL, Docker, Bash

EDUCATION

University of Waterloo - Bachelor of Mathematics, Honours, Co-op **December 2022**
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
- Volunteered on 4-person Faculty Orientation Committee, led the planning and execution of Math Orientation Week in 2021 and 2022 for 250+ Orientation Leaders and 4000+ incoming students

WORK EXPERIENCE

HeadSpin | Engineering Intern | Data Science Team (Remote) Palo Alto, CA
Python, NumPy, Matplotlib, PostgreSQL, OpenCV, MVC *May - Aug 2022*

- Researched, developed, and deployed two analytics features that extends the core platform's existing image matching functionality to support additional use cases for HeadSpin's customers
- Curated targeted video dataset to define measureable success criterias and conducted iterative R&D on image matching from Computer Vision literature
- Shipped features to production using the MVC pattern in Python and published detailed guides in the documentation
- Improved existing image matching feature by introducing robustness to rotation, scale, and background changes

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

- Identified and implemented improvements for the large-scale data de-duplication application, accelerated development cycles for the team and decreased processing times for users
- Initiated test coverage for new Spark features by creating Pytest Continuous Integration and Unit Testing modules
- Automated dependency tracking for Spark jobs, enabling the re-configuration of modules without cluster downtime
- 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
Python, NumPy, Pandas, Matplotlib, spaCy, React, Flask, PostgreSQL, NLP *Jan - Apr / Sep - Dec 2020*

- 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 Toronto, ON
SR&ED, Word, Excel, VBA Scripting *May - Aug 2019*

- Assisted major Canadian technology firms in securing tax credits through the government's SR&ED program
- Compiled documentation from client projects to articulate SR&ED eligibility and drafted technical sections of T661 forms

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

- Improved security features for the inventory management Web App by upgrading authentication and user management workflows, reducing potential vulnerabilities of accessing customer data
- Implemented token-based authentication using JSON Web Tokens and built REST APIs that incorporated LDAP and Microsoft Azure AD credentials for user management