

Shao Zhong Li

 sz7li |  shao-curtis-li |  sz7li.github.io |  sz7li@uwaterloo.ca |  +01.647.272.6690

SKILLS

Languages and Libraries Python, R, C, C++, SQL, Pandas, NumPy, Matplotlib, PySpark, Keras
Tools and Technologies Git, Docker, Bash, Visual Studio

EDUCATION

University of Waterloo - Bachelor of Mathematics, Honours, Co-op **Expected December 2022**
Statistics and Computational Mathematics, Computing Minor *cGPA: 82.15/100*

- Fall 2022 Research Assistant in the Vision and Image Processing Lab, research on Neural Radiance Fields
- Faculty Orientation Committee of 4, led the planning and execution of Math Orientation Week in 2021 and 2022 for 70-200 Orientation Leaders and 2000 incoming students

WORK EXPERIENCE

Headspin Palo Alto, CA (Remote)
Engineering Intern - Data Science Team May - Aug 2022

- Researched, developed, and deployed two analytics features using image matching techniques (template & keypoint)
- Identified customer use cases and curated targeted video dataset for testing and validation
- Conducted extensive R&D, applying existing object detection and image retrieval algorithms from literature
- Presented regular updates to team using plots and animated visualizations with Jupyter Notebooks and Matplotlib
- Wrote production code in Python under the MVC framework and published a user-friendly guide in the documentation
- Shipped features to production, improving existing image match feature by introducing robustness to rotation, scale, and background changes

ThinkData Works Toronto, ON
Data Science Co-op - Data Labs Team May - Aug 2021

- Developed features for the NLP Entity Resolution application using PySpark, Hadoop, PostgreSQL, and Docker
- Updated Pytest Continuous Integration and Unit Testing modules, increasing test coverage for new Spark features
- Automated dependency tracking for Spark jobs, improving testability of modules without cluster downtime
- Benchmarked multi-threaded performance of sparse matrix multiplication; monitored and gathered data on CPU/memory usage under core use cases - improving computation speed by 22%

Canadian Imperial Bank of Commerce Toronto, ON
Data Scientist Co-op - CIBC Data Studio / Enterprise Innovation Team 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
- Conducted experiments on the YouTube 8m Dataset by building and training CNN architectures
- Used Google Places API to extract reviews data from CIBC branches and performed unsupervised text clustering
- Created visualizations of topic, sentiment trends in Google Data Studio and presented findings to prominent stakeholders
- Built and trained sentiment classifiers using Scikit-learn and Keras on SemEval datasets to analyze stock trends
- Developed semantic search functionality in Web App using Sentence-BERT, reducing COVID response-related processing times for customers at CIBC Contact Center
- Incorporated keyphrase extraction using RAKE, EmbedRank, and spaCy dependency parsing to improve search relevance

KPMG Toronto, ON
Summer Student - Tax Incentives Practice Team May - Aug 2019

- Assisted major Canadian technology firms in securing tax credits through the government's SR&ED program
- Compiled project documentation to articulate SR&ED eligibility and drafted technical sections of T661 forms
- Performed data entry in Excel and developed automation VBA scripts for Microsoft Word

DaVinci Assortment Solutions Toronto, ON
Technical Security Analyst Co-op - Engineering Team Sep - Dec 2018

- Improved security features for the inventory management Web App with AngularJS, Spring, and PostgreSQL
- Built REST APIs for user management and implemented token-based authentication using JSON Web Tokens
- Upgraded existing user authentication process to incorporate LDAP and Microsoft Azure AD credentials; improving scalability of new user configuration in steps from linear to constant