Siyi Wei

Undergraduate Candidates

📍 Toronto, Canada
📞 647-870-8349
✉ siyia.wei@mail.utoronto.ca
🌐 www.wei-siyi.com
⌨ github.com/superp0tat0

About me:

Siyi is an active student who focus on statistical learning related questions and applications. He is passionate on discovering the cutting-edge challenges in interdisciplinary area between statistics and computer science. Driven by his passion, he keeps building relative research and teaching experiences. Beyond university work, he has attended DataFest UofT 2019, AmpHack RBC 2019 and Business Data Science Challenge Statistics Canada 2020 as data scientist.

Skills:

Programming:
Advanced:
Python/R/MATLAB/C#/SQL/
TensorFlow/Linux (Ubuntu)/Git
Experienced:
Java/C/Jenkins/Docker/CUDA

Data Science:
Advanced:
Feature Engineering
Data Modelling/Visualization
Machine Learning algorithms
Statistical Analysis

Education

Candidate Honours Bachelor of Science
May 2015 - Present
University of Toronto Scarborough

- Specialist in Statistics: Statistical Machine Learning and Data Mining
- Minor in Computer Science
- Relevant Courses：
  - CSCC11 Intro to Machine Learning (85)
  - CSC413 Deep Learning (Audit)
  - STAD68 Advanced Statistical Learning (88)
  - STAD57 Time Series (85) STAD80 Big Data (On going)

Work Experience

QA Analyst and Web development, Rakuten Kobo
January – May 2019

- Composed QA web story using Jenkins, Ruby and selenium. Helping team analyze customer behaviors through Google analytics.
- Delivered stories following custom Agile management and report to release manager during daily release meetings.
- Develop New kobo subscription plan web page and API services using C#, report it to product manager.

Teaching Assistant, University of Toronto Scarborough
August 2018 - Now

- CSCC11: Introduction to Machine Learning and Data Mining
- STAB23: Introduction to Statistics for Social Science
- CSCA08: Introduction to Programming
- MATA22/24: Linear Algebra I&II

Project Experience

NLP analysis for mental health data (Research Project)
January 2020 – Present

- QA the most recent BERT NLP model performance for mental health conversation records using COVFEFE and Kaldi
- Contribute to open source code by fixing the bugs of COVFEFE through GitHub.
- Currently investigating how to use CUDA boost Kaldi.

Anti-plagiarism Statistical Analysis (Research Project)
January 2020 – Present

- Analyze the pattern of collected MOSS (anti-plagiarism program) data using R. Test different hypothesis and report the suspicious group of students to instructor.
- Visualize the reported data using Google Charts, and implement the data interaction through Shiny in R.