Siyi Wei

Fresh Graduate
With Bachelor Degree

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

Skills:

Programming:
Advanced:
Python/R/MATLAB/C#/SQL
TensorFlow/Linux (Ubuntu)/Git
Keras/Pytorch/Scikit-Learn
Experienced:
Java/C/Jenkins/Docker/CUDA
Apache Hadoop

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

About me:

I am an active student who focus on statistical learning related questions and applications. I am also passionate on discovering the cutting-edge challenges in interdisciplinary area between statistics and computer science. Driven by my passion, I keep 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.

Education

Candidate Honours Bachelor of Science at University of Toronto Scarborough
May 2015 - Present
- 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   **Relevant Skills: C#/Ruby/QA/Git/Docker/Jenkins/SQL**
- Composed QA web story using Jenkins, Ruby and selenium. Helping team analyze customer behaviors through Google analytics. Wrote SQL scripts for test automation and data extraction.
- 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   **Relevant Skills: MATLAB/Python/NumPy/SkLearn**
- CSCC11: Introduction to Machine Learning and Data Mining
- CSCA08: Introduction to Programming
- MATA22/24: Linear Algebra I&II   STAB23: Statistics for Social Science

Project Experience

Personal Website and Blogs (Personal Project)
September 2019 – Present.   **Relevant Skills: Programming/Writing Skills**

Speech Recognition for mental health data (Research Project)
January 2020 – Present   **Relevant Skills: Ubuntu/Shell Commands**
- Contribute to open source code by fixing the bugs of COVFEFE on GitHub. Successfully transfer audio to text for mental health related conversations.

Nominal GDP prediction (Competition for Business Data Science Challenge 2020)
January 2020 – Present   **Relevant Skills: Python/LSTM/Time Series**
- Predict Nominal GDP for different industry sectors using LSTM and ARIMA.

Anti-plagiarism Statistical Analysis (Research Project)
January 2020 – Present   **Relevant Skills: R/Cluster Analysis/PCA**
- Analyze the pattern of collected MOSS (anti-plagiarism program) data using R. Visualize the data pattern through R shiny.

Automation for ARIMA (Academic Project for times-series analysis in R)
August – December 2019   **Relevant Skills: R/Time Series**
- Automate the program for external resources selection in GDP deflators prediction. The improved model decreases the MAPE from 11.77% to 9.37%.