Steven Liu

Phone: +1 (519)721-5934 | Email: bingfanliu36@outlook.com | Web: https://www.linkedin.com/in/steven-bingfan-liu-714195146/

SUMMARY

- 4-year experience using R, Python, SQL for data analysis.
- Machine learning experience using Keras, PyTorch, Scikit-learn, Numpy, Matplotlib, Pandas and Google Colab.
- Computer vision experience using Scikit-image, ChainerCV, Torchvision, PIL.
- Strong statistics background in statistical modeling, quality control, hypothesis testing and non-parametric analysis.

EXPERIENCE

Machine Learning Researcher in Functional Brain Signal Data Analysis | *University of Waterloo* Sep. 2019–Aug. 2020
- Constructed a novel functional machine learning model for functional data classification and feature selection.
- Invented a fast convex optimization algorithm for loss minimization and high dimensional coefficients estimation.
- Applied the model for disease prediction and disease-related brain area allocation using EEG and fMRI.
- Conducted extensive independent research including literature reviews, methodology discussions and presentations.

Teaching Assistant in Statistics | *University of Waterloo* Sep. 2019–Aug. 2020
- Tutored over 700 undergraduate students' probability and statistical theories.
- Assisted over 200 students' hypothesis testing, regression analysis and algorithm implementation using R.

Data Scientist Intern | *UNDP* Jun. 2019–Dec. 2019
- Achieved 83% accuracy in forecasting the severe rainfall using PCA, random forest and regression analysis.
- Participated in the research of poverty prediction for African countries using CNN and satellite night light image data.
- Communicated with data collection team and produced bi-weekly progression report.

Research Assistant in Law Economics | *University of Waterloo* Jan. 2017–Apr. 2017
- Improved the speed of data collection by 400% for Canadian legislative activities by designing Python algorithms.
- Extensive independent investigation and resource checking for legal publications.

EDUCATION

University of Waterloo
- **Master of Mathematics, Statistics |** *GPA: 88/100 (3.89/4)* Sep. 2019–Aug. 2020
- **Bachelor of Arts, Honors Economics Joint Honors Mathematics** | *GPA: 90/100 (3.88/4)* Sep. 2016–Apr. 2019

- **Relevant Courses**: Computer Vision, Machine Learning, Bayesian Statistics, A/B testing, Time Series, GLM.
- **Publication:** "L1-Regularized Functional Support Vector Machine." (Submitted to AAAI 2021).

PROJECTS

Machine Learning: Life Satisfaction Prediction, Kaggle Competition Feb. 2020–Apr. 2020
- Achieved 88.3% AUC score in predicting public life satisfaction level using EU survey data.
- Built a stacking model by blending random forest, L1-GLM, L1-SVM, GBM and implemented a DNN using Keras.

R Package: fdp, Functional Data Preprocessing Package Jan. 2020–Apr. 2020
- Developed a functional data processing package including tools for data smoothing and dimension reduction.
- Created comprehensive unit tests and data simulations for testing package stability.

Computer Vision: Deep Image Prior, Semantic Segmentation, Semi-Supervised Segmentation Dec. 2019–Apr. 2020
- Detected and analyzed prior information of DNN and its applications in denoising, inpainting and super-resolution.
- Performed supervised Semantic Segmentation by constructing a self-designed neural net using PyTorch.
- Implemented a Microsoft adopted semi-supervised Segmentation algorithm using graph cut and clustering.

A/B Testing: Netflix Revenue Optimization Oct. 2019–Dec. 2020
- Minimized Netflix users' average browsing time by 30% through performing a response surface method analysis.
- Optimized the combined level of the design factors in the Netflix.com using a central composite design.

AWARDS

Winner of The American Statistical Association DataFest 2019
- Led a team of 4 predicting and visualizing athlete fatigue level using times series and machine learning on Rugby 7s dataset.

Winner of The American Statistical Association DataFest 2018
- Led a team of 5 analyzing the demand and supply structure of the labor market in the next 5 years on Indeed.com dataset.