Henry Fung

henryhfung21@gmail.com • 202-655-6630 • https://github.com/hfung4

Highlights

Python: pandas, NumPy, SciPy, scikit-learn, Matplotlib, Seaborn, spaCy, Gensim, Keras
Machine Learning: Multiple/Logistic Regression, Decision Trees, Random Forest, Boosting Algorithms, PCA, SVM, KNN, Neural Networks, CNN
Causal Inference: Randomized Control Trial (RCT), A/B Testing, Regression Discontinuity (RD), Instrumental Variables, Panel Data Analysis, Qualitative Research Design
Other: SQL, R, Tableau, Git, STATA, MATLAB, C/C++, LaTeX

Education

Johns Hopkins University, School of Advanced International Studies (SAIS)
MA., International Economics and Quantitative Methods *2018-2020*

McGill University
BA., Economics, Minor in Mathematics, summa cum laude *2015-2018*

Carleton University
MSc., Mechanical Engineering *2008-2010*
- Thesis: Modelling, Simulation, and Control of a Bipedal Walking Robot

BEng., Aerospace Engineering *2003-2008*
- Project: Modelling Human Abdomen Palpation for a Patient Simulator

Projects

Predicting the Intent to Leave of Employees in Large Organizations: Developed classifiers using various statistical learning methods on multi-year survey data (4 million observations) to predict employee's intent to leave and to infer management practices that may induce adverse selection out of large organizations.

Software Requirements Classification with Pre-trained Word Embeddings: Developed a CNN model with pre-trained word embeddings to evaluate its potential application on software requirements classification—a problem that is characterized by small datasets with feature-poor samples.

Regression Discontinuity Study on the Impact of a New Employee Training Program: Evaluated the causal impact of a new training program for government employees with a quasi- experiment (RD). I identified the running variable, established the estimating equations, and tested the identification assumptions. The results of this study informed managers' decision on the continuation of the training program.

Experience

Surgo Ventures, Inc. **Washington, DC**
Research Scientist *9/2020 – Present*

- Support research activities and focus on applying machine learning approaches to problems in global development and public health.

Johns Hopkins University **Washington, DC**
Research Data Analyst *9/2018 – 06/2020*

- Performed data wrangling, cleaning, exploratory data analysis, and feature selection (in Python) on six large public employee survey datasets (4 million observations across 2000 agencies).
- Specified linear regression models to establish associations between 30 management practices and employee attitudes to verify my hypotheses on the effects of supportive management practices on motivation.

Logapps, LLC **Falls Church, VA**
Data Scientist *5/2019 – 9/2019*

- Developed and compared the performance of SVM, Random Forest and CNN with various text vectorization methods (tfidf, pre-trained word embeddings) on software requirements classification. This development was part of a NSF funded project that focuses on automating requirement analysis and software sizing.

- Processed software requirement texts using Python (spaCy and Gensim) and identified semantically similar requirements using pre-trained word embeddings.

Government of Canada **Montreal, Canada**
Data Analytics Specialist *5/2018 – 8/2018*

- Wrangled and cleaned government employee data (in SQL and Python) and developed a panel regression model on 3 million observations to investigate the effects of geographical location, employee experience, and management practices on employee performance.
- Performed a quasi-experiment (Regression Discontinuity) to estimate the impact of a new training program on employee performance and provided actionable recommendations (using data visualizations) to non-technical staffs in the program implementation team.

Textron Inc. **Montreal, Canada**
Avionics Systems Engineer *6/2015 – 8/2016*

- Successfully developed simulated models (in C++) for the Airbus A350 Global Aircraft Position System (GAP) and Flight Control Unit (FCU) based on Airbus technical documents.
- Worked collaboratively with test pilots on the design and implementation of an automated validation process for flight simulators that resulted in the reduction of the validation phase of the project by 80 hours.

CAE Inc. **Montreal, Canada**
Autopilot Software Specialist/Project Engineer *5/2011 – 5/2015*

- Responsible for the design requirement reviews, software development (in C) and testing of simulated autopilot models for Boeing flight simulators (B747 and B787).
- Served as a customer-facing project engineer in five different countries (Norway, Denmark, India, China, and Australia) to support the qualification efforts of simulators and secure customer acceptance.