Faith **Lee**

Data Scientist · Statistician

Quebec, QC, Canada

faithlee.ca | faithghlee | faith-lee-6257b193/ | Faith Lee

Skills

Programming	Python, R, SAS, Java, HTML, CSS, LaTeX
Version control	Git
Languages	English (bilingual), Mandarin Chinese (bilingual), French (intermediate advance)
Actuarial Exams	Probability (P), Financial Mathematics (FM)

Work Experience

Department of Mathematics and Statistics, Laval University

Quebec, QC

Teaching Assistant

Sep 2020 - Present

- Teaching assistant to introductory statistics courses
- Duties as a teaching assistant include holding consultation hours, marking assignments, tests and exams, answering Q&A forums and invigilating exams

MindBridge Analytics Inc.

Ottawa, Ontario

Data Scientist

Jun 2018 - Mar 2020

- Researched into possible analyses areas based on client data and industry (insurance, bank, loans) that could be prototyped based on end-user objective. Analyses areas include the use of ordinary least squares regression (linear regression), K-means, generalized linear model (GLM)
- Cleaned / scrubbed large datasets using Unix commands in the command line (100 million rows)
- Created data visualizations to communicate with clients about their data (through the use of heatmaps, lattice plots, forest plots, etc)
- Transformed prototyped code in Python to production level code for integration into the backend
- Researched into computer vision techniques such as image smoothing, use of convolutional neural networks for the purposes of extracting information and removing noise from invoice images (Tesseract, Azure Optical Character Recognition (OCR) API, OpenCV, Tensorflow)
- Aided in decisions through the use of data-driven approaches
- Technologies used: Python, R, Azure Databricks, Azure OCR, CircleCI, JIRA, Docker

Institute for Clinical and Evaluative Sciences

Toronto, Ontario

Analyst (intern)

May 2017 - Dec 2017

- Derived insights about population health outcomes through the use of provincial administrative databases
- Merged databases, Cleaned, formatted and visualized data for analysis and presentation
- Use of statistical modelling to derive insights such as multi-state model to examine adherence in cervical cancer screening in Ontario and negative binomial GEE(generalized estimating equations) to understand survey uptake
- Engaged in manuscript writing for the purposes of publications in journals, validated results from statistical models and created visualizations for publications
- Technologies used: SAS, R

Department of Statistics and Actuarial Science, University of Waterloo

Waterloo, Ontario

Teaching Assistant

Sep 2016 - Apr 2018

- Teaching assistant to introductory statistics courses
- Duties as a teaching assistant include holding consultation hours, marking assignments, tests and exams, answering Q&A forums and invigilating exams

Education

Laval University

Quebec, QC

Doctorat en statistique

Sep 2020 - Present

- Research on missing data in the longitudinal context

Thompson Rivers University

Remote

Non-degree seeking student

May 2020-Sep 2020

- COMP 2231: Data Structures and Algorithms

University of Waterloo *Waterloo, Ontario*

Master of Mathematics in Biostatistics *Sep 2016 - Apr 2018*

- Research Essay: Attenuation from intermittent measurement of time-dependent co-variates in multi-state models

Memorial University of Newfoundland *St. John's, Newfoundland*

Bachelor of Science (Honours) in Statistics *Sep 2013 - Aug 2016*

- Thesis: Comparison of logistic and survival regression models

Ngee Ann Polytechnic *Singapore*

Diploma in Biomedical Science *Apr 2010 - May 2013*

- Research: Characterization of NDM-1 and IMP genes in antibiotic resistant bacteria

Professional Development

Online Learning/ MOOC

Object Oriented Programming in Java Specialization	2020
Introduction à la programmation orientée objet (en Java)	2020
Bayesian Methods for Machine Learning on Coursera	2019
Convolutional Neural Networks on Coursera	2018
Improving Deep Neural Networks: Hyperparameter tuning, Regularization and Optimization on Coursera	2018
Neural Networks and Deep Learning on Coursera	2018
Machine Learning on Coursera	2017
Practical Machine Learning on Coursera	2017
R Programming on Coursera	2017

Certifications

Diplôme d'études en langue française (DELF) B2	2019
Diplôme d'études en langue française (DELF) B1	2019
1Z0-051 Oracle Database 11g: SQL Fundamentals I	2017
SAS Certified Statistical Business Analyst	2016
SAS Certified Advanced Programmer	2015
SAS Certified Based Programmer	2015

Publications

Immigration and adherence to cervical cancer screening: a province-wide longitudinal matched cohort study using multistate transitional models — Journal of Obstetrics and Gynaecology Canada

Lee F., Sutradhar, R. and Paszat, L.

Higher risk of gastric cancer among immigrants to Ontario: a population-based matched cohort study with over 2 million individuals — Gastric Cancer

Sutradhar, R. Asidianya N., **Lee F.**, et. al.

Use of patient reported outcomes in regional cancer centres: a retrospective study — CMAJ

Barbera, L., **Lee., F** and Sutradhar, R.

Scholarships & Awards

University of Waterloo

International Masters Student Award	Fall 2016, Winter 2017, Winter 2018
Statistics and Actuarial Science Chair's Award	Winter 2017, Winter 2018

Memorial University of Newfoundland

Faculty of Science Dean's List	2013-2014, 2015-2016
MUN Faculty of Science Summer Undergraduate Research Award	2015, 2016
The Williams Science Scholarship	2014-2015
Dr. B.K. Kim Scholarship in Statistics	2014-2015
International Undergraduate Scholarship	2013-2014

Conferences

Science Atlantic Mathematics, Statistics, and Computer Science Conference, 2017 *Fredericton, New Brunswick*

Presenter *Oct 2017*

- Topic presented: Multi-state modelling for examining adherence to cervical cancer based on research with Institute for Clinical and Evaluative Sciences

Science Atlantic Mathematics, Statistics, and Computer Science Conference, 2015 *Wolfville, Nova Scotia*

Presenter - 3rd in Best Oral Presentation under Statistics *Oct 2015*

- Topic presented: Comparison of the logistic regression and Cox regression model based on summer research with Memorial University