linkedin.com/in/taicai
github.com/padtai
hackerrank.com/padtai

— TAI CAI —

padtai.ca
Canadian Citizen with H-1B

Updated 05/2019
tcai@mail.harvard.edu
(857) 998-8619

SUMMARY

PhD training	Bayesian statistics (causal inference, hierarchical models, model averaging, nonparametric theory) Classical statistics (mixed models, statistical learning, experimentation, asymptotic theory)
Current interests	Artificial neural networks, statistical properties of machine learning methods, Bayesian methods for machine learning, ProjectEuler+
Frequently used	Python, Tensorflow, Keras, R, SQL, Spark, Jupyter Notebook, Amazon S3, Amazon EC2, LaTeX
Occasionally used	Java, C++, HTML, CSS, Javascript

EXPERIENCE

Senior Data Scientist, BitSight — *Boston MA* — 02/2019–Present
Data Scientist, BitSight — *Boston MA* — 10/2017–02/2019

- Oversee all data science requirements for third-party risk management projects through cross-functional collaboration in designing metrics, building prototypes, deploying machine learning models, and communicating results to stakeholders
- Design observational studies to evaluate the association and causality of relationships pertaining to efficacy of products, impact of extraneous events, and influence of internal interventions
- Supervise one data science intern working on sales/marketing analytics and external data validation
- Lead reading groups covering advanced topics in forecasting methods, prediction intervals, and model evaluation metrics
- Program internal Python scripts and modules for data storage, collection, cleaning, analysis, and visualization

Technical Advisor, Insight Data Science — *Boston MA* — 05/2019–Present

- Mentoring data science projects for five PhD graduates/postdoctoral researchers by providing feedback on project ideation, data considerations, modeling techniques, and communicating results
- Organize mock interviews based on the type of role that interests the fellow
- Run workshops on probability, statistics, Python coding, and business cases

Data Science Fellow, Insight Data Science — *Boston MA* — 05/2017–09/2017

- Generated idea for predicting supply/demand at Hubway bike-sharing stations in Boston, despite a lack of labeled data
- Consolidated and cleaned multiple data sources to tally labeled information for 200 bike stations over 11 million time points
- Predicted and visualized real-time supply/demand for each bike station using various machine learning and statistical models

PhD Student, Research Assistant, Harvard University — *Cambridge MA* — 08/2012–09/2017

- Conducted original statistical research on applied problems in genetics, health care policy, and end-of-life care
- Developed novel R code for handling datasets with complexities such as sampling bias, misclassified outcomes, correlated outcomes, hierarchical structures, and confounding
- Taught graduate-level labs with topics ranging from introductory statistics to seminar topics in Bayesian nonparametrics, decision theory, and sequential methods
- Selected to tutor fellow PhD students for the biostatistics written qualifying exam

Full-time Co-op Work Semesters, University of Waterloo — *Waterloo ON* — 05/2008–12/2011

- Completed six semesters of full-time work at Munich Re, Manulife, Hewitt Associates, and Logitech
- Applied statistical and actuarial methods to problems in biometric research, enterprise risk management, segregated fund valuation, group benefits pricing, pension administration, and database analysis

EDUCATION

PhD, Biostatistics, Harvard University — *Cambridge MA* — 2017

- Thesis: Statistical Methods for the Analysis of Observational Data with Multiple Correlated Outcomes
- Advisors: Tianxi Cai, Francesca Dominici

Bachelor of Mathematics, University of Waterloo — *Waterloo ON* — 2012

- Honours Actuarial Science/Finance Option, Honours Statistics, Co-operative Program

Deep Learning Specialization, deeplearning.ai — *Coursera* — 2018
Associate of the Society of Actuaries, Society of Actuaries — *Schaumburg IL* — 2012