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

Tianyi (Tai) Cai

Brookline MA
Canadian Citizen with H-1B

Updated 02/2019
tcai@mail.harvard.edu
(857) 998-8619

Summary

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

Experience

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

- Oversee all data science requirements for third party risk management projects by collaborating with cross-functional teams
- Partner with product, sales, and marketing managers to devise data-driven strategies from unorthodox data structures
- Design studies to evaluate the association and causality of relationships pertaining to efficacy of products, impact of extraneous events, and influence of internal interventions
- Program Python scripts and modules for data storage, collection, cleaning, analysis, and visualization
- Lead reading groups covering advanced topics in forecasting methods, prediction intervals, and model evaluation metrics
- Support data science team members in ad hoc statistical tasks

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 real-time supply/demand for each bike station using various machine learning and statistical models
- Created website to visualize directions and to predict bike availability for user input origins and destinations
- Provided mentorship and feedback for subsequent cohorts of fellows

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

- Conducted original statistical research on applied problems in mental disorders, health care policy, and end-of-life care
- Developed 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

- Actuarial Analyst, Enterprise Risk Management Munich Re — *Toronto ON* — 09/2011–12/2011
- Actuarial Analyst, Biometric Research Munich Re — *Toronto ON* — 01/2011–04/2011
- Actuarial Analyst, Segregated Fund Valuation, Manulife — *Waterloo ON* — 05/2010–08/2010
- Actuarial Analyst, Group Benefits Pricing, Manulife — *Waterloo ON* — 09/2009–12/2009
- Pension Administrator, Hewitt Associates — *Toronto ON* — 01/2009–04/2009
- Database Analyst, Logitech — *Mississauga ON* — 05/2008–08/2008

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

Associate of the Society of Actuaries, Society of Actuaries — *Schaumburg IL* — 2012