Website: padtai.ca
LinkedIn: linkedin.com/in/taicai
GitHub: github.com/padtai

**Tianyi (Tai) Cai**
*Brookline MA*
Canadian Citizen with H-1B

Updated 07/2018
tai@mail.harvard.edu
(857) 998-8619

## Summary

**Graduate-level training in:** Bayesian statistics (multilevel models, causal inference, model averaging, nonparametric methods), classical statistics (hypothesis testing, mixed models, regularization, asymptotic theory)

**Comfortable using:** Python (scikit-learn, NumPy, pandas, matplotlib, SQLAlchemy, Flask, Beautiful Soup), R (ggplot2, Markdown, Shiny), Jupyter Notebook, Amazon Web Services, LaTeX/LyX, Microsoft Excel, HTML, CSS, Javascript, git, tmux

**Some exposure to:** Keras, TensorFlow, Hadoop, Spark, Scala, Java, SAS, Stata

## Experience

**Data Scientist**, BitSight Technologies, *Cambridge MA*  ·  10/2017-Present
· Presented data-driven solutions for projects involving new products, engineering support, marketing and sales analytics, insurance modeling, and customer success management
· Designed studies to evaluate the association and causality of relationships pertaining to efficacy of products, impact of extraneous events, and influence of internal interventions
· Programmed scripting tools for data collection, cleaning, analysis, visualization, and storage
· Led reading groups covering forecasting methods, prediction intervals, and model evaluation metrics
· Supported 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 (as an alumni fellow) for subsequent cohorts of fellows

**PhD Student, Research Assistant**, Harvard University, *Cambridge MA*  ·  08/2012-09/2017
· Developed R code for handling complex datasets with issues in sampling bias, misclassified outcomes, correlated outcomes, hierarchical structures, and confounding
· Conducted original statistical research on problems in mental disorders, health care policy, and end-of-life care
· Taught graduate-level courses 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, Biometric Research, Munich Re, *Toronto ON*  ·  09/2011-12/2011
· Actuarial Analyst, Enterprise Risk Management, Munich Re, *Toronto ON*  ·  01/2011-04/2011
· Actuarial Analyst, Segregated Funds, Manulife, *Waterloo ON*  ·  05/2010-08/2010
· Actuarial Analyst, Group Benefits, 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*  ·  08/2012-09/2017
· Thesis: Statistical Methods for the Analysis of Observational Data with Multiple Correlated Outcomes

**Bachelor of Mathematics**, University of Waterloo, *Waterloo ON*  ·  09/2007-06/2012
· Honours Actuarial Science/Finance Option, Honours Statistics, Co-operative Program

**Associate of the Society of Actuaries**, Society of Actuaries, *Schaumburg IL*  ·  01/2009-03/2012