

 [linkedin.com/in/taicai](https://www.linkedin.com/in/taicai)


 github.com/caikitlearn

 [hackerrank.com/caikitlearn](https://www.hackerrank.com/caikitlearn)


– TIANYI “TAI” CAI, PhD –

Canadian Citizen with H-1B

caikitlearn.com

Updated 10/2019 

tcai@mail.harvard.edu 

(857) 998-8619 

SUMMARY

Profile	· Experienced Data Scientist and Statistician working at the intersection of applied statistics, machine learning, research science, software engineering, and product management
Areas of Expertise	· Data-driven software products and features put in production through cross-functional collaboration · Applied machine learning (data preparation, classification, regression, clustering, visualizations) · Bayesian statistics (causal inference, hierarchical models, model selection, nonparametric methods) · Classical statistics (experimentation, statistical learning, mixed models, time series)
Current interests	· Statistical properties of machine learning methods, causal inference, ProjectEuler+
Frequently used	· Python, R, SQL, Spark, Jupyter Notebook, Amazon S3, Amazon EC2, Vim, \LaTeX
Used as necessary	· Tensorflow, Keras, Java, C++, HTML, CSS, Javascript

EXPERIENCE

Senior Data Scientist, BitSight	<i>Boston MA</i>	10/2017–Present
· Oversee all data science projects for third-party cybersecurity risk management through designing metrics, researching methods, building prototypes, writing and reviewing production code, and communicating results to stakeholders · Shipped multiple new software products and features (including 1 patent pending), working as the sole data scientist in collaboration with engineering, product, customer success, marketing, and sales teams · Designed observational studies to evaluate the association and causality of relationships pertaining to efficacy of products, impact of extraneous events, and influence of internal interventions · Supervised Master’s-level intern working on 1. validation of external financial data and 2. sales and marketing analytics · Led reading groups covering advanced topics in forecasting methods, prediction intervals, and model evaluation metrics · Promoted from Data Scientist on 02/2019 for exceeding expectations across all major projects		
Technical Advisor, Alumni Mentor, Data Science Fellow, Insight Data Science	<i>Boston MA</i>	05/2017–09/2019
· (As Technical Advisor) Mentored data science projects for 4 PhD graduates and postdoctoral fellows by providing weekly 1-on-1 feedback on project ideation/viability, data considerations, modeling techniques, and communication skills · (As Alumni Mentor) Conducted mock interviews covering computer science, machine learning, statistics, and case studies · (As Data Science Fellow) Consolidated and cleaned multiple data sources to tally millions of labeled data points to predict and visualize real-time supply and demand for Hubway bikes at 200 bike-sharing stations		
PhD Student, Research Assistant, Harvard University	<i>Cambridge MA</i>	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 classes 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	<i>Waterloo ON</i>	05/2008–12/2011
· Completed six semesters (two years) 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	<i>Cambridge MA</i>	2017
· Thesis: Statistical Methods for the Analysis of Observational Data with Multiple Correlated Outcomes		
Bachelor of Mathematics, University of Waterloo	<i>Waterloo ON</i>	2012
· Honours Actuarial Science/Finance Option, Honours Statistics · Graduated With Distinction - Dean’s Honours List		
Deep Learning Specialization, deeplearning.ai	<i>Coursera</i>	2018
Associate of the Society of Actuaries, Society of Actuaries	<i>Schaumburg IL</i>	2012