linkedin.com/in/taicai

github.com/caikitlearn

hackerrank.com/caikitlearn

– TAI CAI, PHD –

caikitlearn.com

Canadian Citizen | US Permanent Resident

Updated 03/2022

tcai@mail.harvard.edu

(857) 998-8619

SUMMARY

Profile	· Quantitative scientist working at the intersection of applied statistics, machine learning, research science, software engineering, and product management
Areas of Expertise	· Cross-functional collaboration to productionize/engineer data-driven products and features · Bayesian statistics (causal inference, hierarchical models, model selection, nonparametric methods) · Classical statistics (statistical learning, experimental design, mixed models, time series) · Deep learning (artificial/recurrent/convolutional neural networks)
Frequently Used	· Python, Tensorflow, Keras, LightGBM, R, SQL, Spark, jupyter, Vim, LaTeX
Sometimes Used	· Java, C++, HTML, CSS, Javascript

EXPERIENCE

Research Data Scientist, Infrastructure Quantitative Engineering   *Meta*, *New York, NY* 02/2021–Present

· Collaborate as the core scientific point of contact in a 100 person team to build, improve, and migrate users to internal infrastructure tools used by engineers and scientists at Meta
· Serve as technical lead for members of the data team through planning, 1-on-1 meetings, and mentorship
· Oversee all components of the data team, including prioritization, resource allocation, production modeling, dashboards, ad hoc requests, data pipelines, and stakeholder management
· Provide direction to the team by meeting with other core members in engineering, product, design, and UX research

Applied Scientist, Global Intelligence   *Uber*, *San Francisco, CA* 01/2020–02/2021

· Estimated point and interval values of competitive intelligence metrics used in executive decision-making
· Engineered an ensemble of predictive models using internal and third-party data to identify rare user groups for multi-million dollar experiments; improved legacy model performance by a factor of 2
· Identified and corrected multiple errors and inefficiencies found in the team repository, ranging from production code to deliverables used by stakeholders

Senior Data Scientist, Data and Research   *BitSight*, *Boston, MA* 10/2017–01/2020

· Led all data science projects for third-party cybersecurity risk management, including BitSight's first machine learning based product, multiple analytical tools, and observational studies to evaluate association and causality
· Supervised Master's-level intern working on 1) validation of external financial data and 2) sales/marketing analytics
· Organized reading groups covering advanced topics in forecasting, prediction intervals, and model evaluation

Technical Advisor, Data Science Fellow   *Insight Data Science*, *Boston, MA* 05/2017–09/2019

· As Technical Advisor: mentored data science projects for PhD graduates and postdoctoral fellows by providing weekly 1-on-1 feedback on project ideation/viability, data considerations, modeling techniques, and communication skills
· 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 at 200 bike-sharing stations

EDUCATION

PhD, Biostatistics   *Harvard University*, *Cambridge, MA* 08/2012–09/2017

· Conducted and published statistical research on applied problems in genetics, health care policy, and end-of-life care
· Developed novel methods and R code for handling sampling bias, misclassified outcomes, correlated outcomes, intractable cross-validation, hierarchical structures, and confounding
· Taught graduate-level classes in Bayesian statistics, statistical methods, and introductory statistics/epidemiology

BMath, Statistics and Actuarial Science   *University of Waterloo*, *Waterloo, ON, Canada* 09/2007–06/2012

· Graduated with Distinction - Dean's Honours List (Highest Honors)
· Completed six work semesters at Logitech, Hewitt Associates (now Blackstone), Manulife Financial, and Munich Re
· René Descartes National Scholar