

Dustin Tran

26 Everett Street, RM 110
Cambridge, MA 02138
✉ dtran@g.harvard.edu
dustintran.com

Research interests

Bayesian statistics and computing strategies for massive datasets—with an emphasis on Bayesian nonparametrics, variational inference, and stochastic approximations.

Education

- 2014–present **Ph.D. Statistics**, *Harvard University*, Cambridge, MA.
S.M. Computational Science and Engineering, *Harvard University*, Cambridge, MA.
Advisor: Edoardo Airoldi.
- 2010–2014 **B.A. Mathematics, Statistics**, *University of California, Berkeley*, Berkeley, CA.
Graduated with *Highest Honors*. Advisor: Jim Pitman.

Awards and honors

- 2014 Dorothea Klumpke Roberts Prize in Mathematics
- 2010–2014 Regents' and Chancellor's Scholarship (Top 0.5% of Applicants)
- 2013 Rose Hills Foundation Science & Engineering Grant

Research

- 9/2014–present **Statistical Machine Learning**, *Harvard University*.
Currently studying stochastic gradient methods under a statistical framework with Professor Edoardo Airoldi. Paper under review for ICML 2015.
- 1/2014–5/2014 **Convex Optimization**, *University of California, Berkeley*.
Explored different algorithms to automate model selection in machine learning, reformulated as a constrained optimization problem, with Professor Ben Recht. Led to a research paper with theoretical results and simulations indicating that the proposed method achieves comparable performance to standard estimates which use cross validation.
- 8/2013–12/2013 **Matrix Factorization**, *University of California, Berkeley*.
Examined randomized algorithms for low rank approximations under Professor John Strain. Analyzed their computational efficiency and robustness compared to classical techniques such as SVD and Krylov subspace methods.

Publications

2. **Dustin Tran**, Panos Toulis, and Edoardo Airoldi. Stable and optimal: Implicit stochastic gradient descent with averaging. Submitted, 2015.
1. **Dustin Tran**. Convex Techniques for Model Selection. Preprint arXiv:1411.7596 [math.OC], 2014.

Selected talks

- *Implicit stochastic gradient methods for principled estimation*, Max Planck Institute for Intelligent Systems, Tübingen, Germany. January 2015.
- *Facebook: Tree-like Structure in Social and Information Networks*, Institute for Applied Computational Science Seminars, Cambridge, MA. November 2014.
- *Detecting contagion in financial networks*, Spatial Networks Seminar, Berkeley, CA. December 2013.

Industry

5/2014–present **Data Scientist**, *Earnest*, San Francisco, CA.
Currently working part-time at startup, applying tools for feature learning in transactions data. Built the primary algorithm for loan decision-making, which predicts the risk of default for a loan applicant using ensemble methods. Developed the infrastructure for web reporting, which would be used for internal operations, business development, and marketing.

Teaching

Fall 2014 Teaching Fellow, **Numerical Methods** (Applied Math 205), *Harvard University*.

Spring 2013 Teaching Assistant, **Methods in Calculus, Statistics, Combinatorics** (Math 10B), *University of California, Berkeley*.

Summer 2011 Teaching Assistant, **Numerical Analysis** (Math 128A), *University of California, Berkeley*.

Programming

- Languages: Python (+numpy, +pandas, +sklearn), R, C++, JavaScript (+D3.js), {Ba,z}sh
- Software: Vim, Git, Hadoop, SQL
- Operating Systems: GNU/Linux, BSD