

James A. Brofos

CONTACT INFORMATION

24 Hillhouse Avenue
New Haven, CT 06511

E-mail: james@brofos.org
Website: www.brofos.org

RESEARCH INTERESTS

I am interested in designing intelligent systems that leverage new computational methods in statistical inference. I want to develop algorithms that generalize, comprehend uncertainty, and scale to problems of practical significance. I have worked in scalable estimation for graphical models, variational inference, robustness to errors-in-variables, and Bayesian optimization. My current research in Bayesian learning is motivated by applications in engineering, cybersecurity, and the humanities.

EDUCATION

Yale University Connecticut, United States
Ph.D., Statistics & Data Science (August 2018 - Present)
Research Specialization: Bayesian machine learning

Dartmouth College New Hampshire, United States
B.A., Mathematics & Computer Science (September 2011 - June 2015)
Relevant Coursework: Machine Learning · Deep Learning · Parallel Computing · Advanced Statistics · Mathematical Optimization · Advanced Econometrics · Theoretical Machine Learning · Communication Protocols and Complexity
Cumulative GPA: 3.63 / 4.00

Undergraduate Honors Theses

Statistical Estimation of Ising Graphical Models (High Honors) Communication Complexity of Distributed Statistical Algorithms (Honors)

Advisor: Eugene Demidenko

Advisor: Peter Doyle

Developed a scalable and parallel algorithm for estimating the structure of Ising models. Proved bounds on the reconstruction error of the maximum likelihood estimator.

Determined new probabilistic bounds for the communication complexity of distributed statistical estimation problems using information theory.

INDUSTRY EXPERIENCE

The MITRE Corporation Bedford, Massachusetts
Senior Data Scientist (June 2015 - Present)

Deployed Bayesian optimization as a service for personnel to tune hyperparameters of machine learning systems in Python, R, and MATLAB. Developed a distributed infrastructure using TensorFlow and MPI to enable scalable sampling of Bayesian neural network posterior distributions. Developed a decision-making system that identified over \$300M in anomalous payments by healthcare insurance providers. Constructed a likelihood ratio test to enable efficient verification and accreditation of a digital radar target simulator for the US military.

ACADEMIC EXPERIENCE

Carnegie Mellon Computer Science Department Pittsburgh, Pennsylvania
Summer Machine Learning Researcher (June 2014 - August 2014)
Wrote parallel C code using MPI for large-scale analysis of Markov random fields.

Dartmouth Department of Classics Hanover, New Hampshire
Machine Learning Research Assistant (March 2014 - June 2014)

Applied machine learning for anomaly detection with language-based features for authorship attribution. Developed metrics to analyze similarity between texts in sound, content, and style.

Dartmouth Department of Mathematics Hanover, New Hampshire
NSF-Supported Research Assistant (June 2013 - September 2013)

Implemented an adaptive quadrature schema that enabled systems described by Laplace and Helmholtz equations to be solved more efficiently and accurately via boundary integral equations.

Dartmouth Department of Physics Hanover, New Hampshire
Cluster System Administrator (February 2013 - May 2013)

Built and configured a high-performance Linux environment in CentOS to meet the computational requirements of studying Bose-Einstein condensates for the Department of Physics.

TEACHING
EXPERIENCE

Dartmouth Academic Skills Center Hanover, New Hampshire
Teaching Assistant (June 2012 - May 2014)

Courses: Mathematical Finance · Advanced Topics in Statistics · Statistical Analysis and Machine Learning · Discrete Probability · Econometrics · Data Visualization

PUBLICATIONS

Journal Papers

- [1] Joseph Dexter, Theodore Katz, Nilesh Tripuraneni, Tathagata Dasgupta, Ajay Kannan, James Brofos, Jorge Bonilla Lopez, Lea Schroeder, Adriana Casarez, Maxim Rabinovich, Ayelet Haimson Lushkov, and Pramit Chaudhuri. Quantitative criticism of literary relationships. *Proceedings of the National Academy of Sciences*, 114(16), 2017.
- [2] James Brofos. Man vs. machine: The rivalry in chess. *Dartmouth Undergraduate Journal of Science*, 16(2), 2014.

Conference Proceedings

- [1] James Brofos, Rui Shu, and Frank Zhang. The optimistic method for model estimation. In *15th International Symposium on Intelligent Data Analysis (IDA)*. Springer International Publishing, 2016.
- [2] James Brofos and Rui Shu. Parallelization of minimum probability flow on binary Markov random fields. In *14th International Conference on Machine Learning and Applications (ICMLA)*. IEEE, 2015.

Peer-Reviewed Workshop Contributions

- [1] James Brofos, Michael Downs, and Rui Shu. Detecting evasive malware with loss-calibrated Bayesian neural networks. In *International Conference on Machine Learning (ICML) Workshop on Machine Learning for Safety-Critical Applications in Engineering*, 2018.
- [2] James Brofos, Michael Downs, and Rui Shu. Deep Bayesian defenses against adversarial malware. In *Applied Machine Learning Days (AML D)*, 2018.
- [3] Rui Shu, James Brofos, Frank Zhang, Hung Hai Bui, Mohammad Ghavamzadeh, and Mykel Kochenderfer. Stochastic video prediction with conditional density estimation. In *European Conference on Computer Vision (ECCV) Workshop on Action and Anticipation for Visual Learning*, 2016.

- [4] James Brofos, Rui Shu, Michael Downs, and Matthew Jin. Leveraging deep neural networks as kernels for survival analysis. In Neural Information Processing Systems (NIPS) Workshop on Machine Learning in Healthcare, 2015.

Technical Reports

- [1] James Brofos, Frank Zhang, and Rui Shu. A statistical approach to the circular error probable, 2017.
- [2] James Brofos, Frank Zhang, and Rui Shu. Batch Implicit Gradient Descent, 2017.

TALKS AND TUTORIALS

- [1] Bayesian Optimization. The MITRE Corporation, 2018.
- [2] Introduction to Bayesian optimization. The MITRE Corporation, 2016.
- [3] Tutorial on optimistic estimation of statistical models. The MITRE Corporation, 2016.
- [4] Adaptive neural network representations for parallel and scalable Bayesian optimization. Dartmouth Computer Science Research Symposium, 2015. With Rui Shu and Matthew Jin.

HONORS AND AWARDS

- [1] National Science Foundation Graduate Research Fellowship (2018)
- [2] Honorary Dartmouth James B. Reynolds Scholarship (2018)
- [3] IEEE International Conference on Machine Learning and Applications Best Poster Award (2015)
- [4] Sigma Xi Honor Society – Associate Member (2015)
- [5] Neukom Institute’s Undergraduate Research Prize (2013, 2014, 2015)
- [6] John G. Kemeny Computing Prize – Honorable Mention (2015)
- [7] Duke University Data Science Competition – Honorable Mention (2014)
- [8] Quora Machine Learning Competition – Fourth Place (2013)
- [9] Neukom Institute’s Saying Science Prize (2012, 2013)
- [10] Dartmouth Digital Arts Exhibition – Honorable Mention (2013)
- [11] Dartmouth College Occom Scholar (2012)
- [12] Dartmouth Institute for Writing & Rhetoric’s Dickerson Essay Prize (2011)
- [13] German Consulate, Language and Literature Commendation (2011)

GRANTS

The MITRE Corporation

- [1] *Research Program: Applications of Bayesian Reinforcement Learning to Cyber-security* \$92,000 (2017)
- [2] *Research Program: Statistics for the Circular Error Probable* \$8,500 (2016)

Dartmouth College

- [1] *Neukom Institute: Bounding the Cumulative Regret for Bayesian Optimization Algorithms* \$1,000 (2014)
- [2] *Kaminski Family Fund: Distributed Statistical Estimation of Ising Graphical Models* \$4,300 (2014)
- [3] *Presidential Scholarship: Applications of Deep Learning to Cross-Language Machine Translation* \$1,500 (2013)

- [4] *Neukom Institute*: Machine Learning for Political Analysis \$1,000 (2013)
[5] *Neukom Institute*: Ensembles for Stock Return Prediction \$1,000 (2013)

REVIEWING AND
SERVICE

Journal Reviewing

Model Assisted Statistics and Applications (2014)

OPEN-SOURCE
PROJECTS
Available on GitHub

Thor

Implements an API and user interface, hosted via a Flask server, for facilitating Bayesian optimization of machine learning systems. Includes state-of-the-art advances in hyperparameter search such as portfolios of acquisition functions, distributed optimization, and low-discrepancy pseudo-random restarts for kernel parameter tuning. Thor possesses client-side interfaces for Python, R, and MATLAB.

Odin

Implements an event-driven live trading and backtesting infrastructure in Python. Supports a modular design that allows retail traders to leverage low-level control over signal generation, portfolio management, equity rebalancing, and data streams. Odin integrates closely with a dedicated Postgres equities database and incorporates modern portfolio performance metrics.

SKILLS

Programming Proficiency: Python · MATLAB · R · Swift · Objective-C · Java · C ·
MPI · HTML & CSS
Operating Systems: Linux (Ubuntu) · macOS · iOS