

James A. Brofos

Pittsburgh, Pennsylvania | +1 (978) 330-0765 | james@brofos.org | www.brofos.org

EDUCATION

Yale University, New Haven, CT

Ph.D. in Statistics and Data Science

May 2022

M.A. in Statistics

May 2020

Dissertation: Essays on Numerical Integration in Riemannian Manifold Hamiltonian Monte Carlo

Awards: National Science Foundation Graduate Research Fellowship (2018)

Relevant Coursework: Applied Data Mining & Machine Learning, Linear Models, Topics in Functional Analysis, Statistical Signal Processing, Theory of Statistics, Harmonic Analysis

GPA: Honors in 11 / 12 courses (approx. 3.95 on 4.00 scale)

Dartmouth College, Hanover, NH

B.A. in Mathematics (with High Honors) and Computer Science (with Honors)

June 2015

Bachelor's Theses:

- Mathematics: Statistical Estimation of Ising Graphical Models
- Computer Science: Communication Complexity of Distributed Statistical Algorithms

Awards: Dartmouth Institute for Writing & Rhetoric's Dickerson Essay Prize (2012), Dartmouth Presidential Scholar (2013), Neukom Institute Prize for Undergraduate Research (2013, 2014, 2015)

Relevant Coursework: Deep Learning, Parallel Computing, Advanced Statistics, Advanced Econometrics, Mathematical Optimization, Theoretical Machine Learning

GPA: 3.63 / 4.00

SKILLS AND QUALIFICATIONS

- *Programming Skills:* Python (9 Years), MATLAB (9 Years), LaTeX (8 Years), Cluster and Distributed Computing (4 Years), R (3 Years), HTML & CSS (2 Years), C (1 Year)
- *Language Skills:* Swedish (Fluent), German (Basic Reading)
- *Security Clearance:* DOD - Top Secret

RESEARCH AND WORK EXPERIENCE

Yale University, New Haven, CT

Ph.D. Researcher

August 2018 – Present

- Developed and executed a multi-year research agenda studying efficient methods of geometric numerical integration within the context of Markov chain Monte Carlo using Python (NumPy and SciPy), and large-scale distributed computing (Linux cluster and Slurm resource manager).
- Collaborated with researchers at Yale University, York University, and École Polytechnique to develop a theory of adaptive methods of Markov chain Monte Carlo leveraging methods of deep generative modeling (normalizing flows) with applications to Bayesian inference in physical systems.
- Collaborated with faculty at Dartmouth College and University of Texas to develop word embedding (Word2Vec) and TF-IDF models for the identification of influence groups from classical Greco-Roman prose and poetry.

The MITRE Corporation, Bedford, MA

Senior Data Scientist (Quarter-Time)

September 2018 – Present

- Worked directly with the Senior Advisor to the CEO to assess (using Pandas and Matplotlib) the correlation between non-pharmaceutical interventions and COVID-19 propagation, and briefed MITRE's CEO and senior leadership on findings.
- Developed prototypes and advised senior technical staff on the use of the bootstrap method in MATLAB to generate approximate confidence intervals in a COVID-19 survival analysis study.
- Developed prototypes and advised project leadership on causal inference using directed acyclic graphs in application to the Veteran Benefits Administration's claim analysis program.
- Advised project leadership engaged in aircraft ground safety assessments for the FAA on the applications of Bayesian linear regression using R and Stan.

- Advised a team of junior and senior technical staff members on statistical assumption checking and methodological approach within an observational study of the effectiveness of monoclonal antibodies against COVID-19.
- Senior Data Scientist (Full-Time)* January 2017 – September 2018
- Conceived and executed a research program to analyze machine learning methods (Bayesian neural networks in TensorFlow & Keras) for improving the robustness of anti-virus software against adversarial malware.
 - Devised and implemented a statistical hypothesis test in MATLAB to enable efficient evaluation of the USMC Ground / Air Task Oriented Radar (AN/TPS-80) to perform its Ground Weapon Location Mission during developmental testing.
- Data Scientist* June 2015 – January 2017
- Developed an automated analysis system in Python using Pandas that uncovered over \$300M in anomalous payments by healthcare insurance providers.
- Dartmouth College**, Hanover, NH
- Machine Learning Research Assistant* March 2014 – June 2014
- Applied machine learning for anomaly detection with language-based features for authorship attribution in ancient Greco-Roman literature.
- Mathematics Research Assistant* June 2013 – September 2013
- Implemented an adaptive quadrature approach in MATLAB that enabled systems described by Laplace and Helmholtz equations to be solved more efficiently and accurately via boundary integral equations.
- Cluster System Administrator* February 2013 – May 2013
- Constructed and configured a high-performance Linux environment in CentOS to meet the computational requirements of studying Bose-Einstein condensates for the Department of Physics.
- Carnegie Mellon University**, Pittsburgh, PA
- Machine Learning Research Assistant* June 2014 – August 2014
- Implemented algorithms for the large-scale analysis of Markov random fields using C and MPI.

TEACHING EXPERIENCE

- Yale University**, New Haven, CT
- Yale Teaching Fellowship Program* September 2019 – May 2019
- Teaching Assistant for Probability and Statistics (Fall 2019) and Introductory Statistics (Spring 2019).
 - Responsible for evaluating homework, exams, and course projects, and for clarifying key concepts in probability and statistical inference in weekly tutorial sessions.
- The MITRE Corporation**, Bedford, MA
- MITRE Institute Instructor* June 2018 – June 2020
- Designed and presented courses to technical staff on Bayesian optimization and Hamiltonian Monte Carlo.
- Dartmouth College**, Hanover, NH
- Academic Skills Center Grader & Tutor* June 2012 – May 2014
- Graded or provided tutoring services for the following courses: Mathematical Finance, Advanced Topics in Statistics, Statistical Analysis and Machine Learning, Discrete Probability, Econometrics, and Data Visualization.

PEER-REVIEWED PUBLICATIONS

1. James A. Brofos, Marylou Gabri e, Marcus A. Brubaker, Roy R. Lederman. *Adaptation of the Independent Metropolis-Hastings Sampler with Normalizing Flow Proposals*. In International Conference on Artificial Intelligence and Statistics (AISTATS), 2022.
2. James A. Brofos and Roy R. Lederman. *Evaluating the Implicit Midpoint Integrator for Riemannian Hamiltonian Monte Carlo*. In International Conference on Machine Learning (ICML), 2021.
3. Patrick J. Burns, James A. Brofos, Kyle Li, Pramit Chaudhuri, and Joseph P. Dexter. *Profiling of Intertextuality in Latin Literature Using Word Embeddings*. In Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL), 2021.
4. Joseph Dexter, Theodore Katz, Nilesh Tripuraneni, Tathagata Dasgupta, Ajay Kannan, James A. Brofos, Jorge Bonilla Lopez, Lea Schroeder, Adriana Casarez, Maxim Rabinovich, Ayelet Haimson Lushkov, and Pramit

- Chaudhuri. *Quantitative Criticism of Literary Relationships*. In Proceedings of the National Academy of Sciences (PNAS), 2017.
5. James A. Brofos, Rui Shu, and Frank Zhang. *The Optimistic Method for Model Estimation*. In 15th International Symposium on Intelligent Data Analysis (IDA), 2016.
 6. James A. Brofos and Rui Shu. *Parallelization of Minimum Probability Flow on Binary Markov Random Fields*. In 14th International Conference on Machine Learning and Applications (ICMLA), 2015. Received award for best poster.

WORKSHOP PUBLICATIONS

7. James A. Brofos, Marcus A. Brubaker, Roy R. Lederman. *Manifold Density Estimation via Generalized Dequantization*. In International Conference on Machine Learning (ICML) Workshop on Invertible Neural Networks, Normalizing Flows, and Explicit Likelihood Models, 2021.
8. James A. Brofos, Rui Shu, Roy R. Lederman. *A Bias-Variance Decomposition for Bayesian Deep Learning*. In Neural Information Processing Systems (NeurIPS) Workshop on Bayesian Deep Learning, 2019.
9. James A. 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.
10. Rui Shu, James A. 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.
11. James A. 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.