Dougal J. Sutherland

Contact Information

Gatsby Computational Neuroscience Unit
University College London
25 Howland Street
London, UK W1T 4JG

email: dougal@gmail.com
web: www.gatsby.ucl.ac.uk/~dougals/
office: +44 20 3108 8120

Research interests

Machine learning, particularly kernel methods and their integration with deep learning. Problems including generative models, two-sample testing, density estimation, and distribution regression and classification. Active and human-in-the-loop learning. Nonparametric statistics, statistical learning theory.

Academic Positions and Education

2016 – **Research Associate**, *Gatsby Computational Neuroscience Unit, University College London*.
Postdoctoral position with Arthur Gretton.

2011 – 2016 **Ph.D., Computer Science**, *Carnegie Mellon University*.
Thesis Title: Scalable, Flexible, and Active Learning on Distributions.
Committee: Jeff Schneider (chair), Barnabás Póczos, Maria-Florina Balcan, Arthur Gretton.
Included an M.S. obtained in 2015. Unofficial GPA: 3.96 / 4.

2007 – 2011 **B.A., Computer Science**, *Swarthmore College*, with high honors.
Minors in Linguistics (with high honors) and Mathematics & Statistics. GPA: 3.93 / 4.
Thesis Title: Integrating Human Knowledge into a Relational Learning System.

Honors and Awards

2014 – 2016 Sandia Campus Executive Program fellowship. (Renewed in 2015.)

2013 National Science Foundation Graduate Research Fellowship Program: Honorable Mention.

2011 Ivy Award for "the senior man outstanding in leadership, scholarship, and contributions to the college community" by Swarthmore faculty vote.

2011 Elected Phi Beta Kappa.

2011 Drew Pearson Prize for excellence in journalism.

Research and Academic Experience

2013 – 2016 **XDATA workshops**, *DARPA*.
Analyzed various datasets with teams from across academia and industry as a testbed for development of open-source data-analytic software libraries. Developed a Python library, `skl-groups`, for machine learning on distributions. Participated in development of a financial analysis application in use at a federal agency. Led a small team of CMU participants and managed collaborations.

2011 – 2016 **Ph.D. research**, *Carnegie Mellon University*.
Research in machine learning with Jeff Schneider. Particular focus on machine learning on samples from distributions and on active learning problems. Additional unpublished empirical work related to learning on distributions in the analysis of financial anomalies, fusion reactor behavior, web browsing traffic, shipping behavior, terrorist activities, and Twitter language use.

2011 **Linguistics senior honors study**, *Swarthmore College*.
Studied the phonotactics of Chaha, including computational approaches, for Colleen Fitzgerald.

2010 **REU in machine learning**, *University of Oklahoma*.
Improved and added human interaction to a relational concept learning system with Andrew Fagg.

2009 **Howard Hughes Medical Institute fellowship**, *Swarthmore College*.
Worked on natural language processing and medical information extraction with Rich Wicentowski.

2009 **Directed independent study project**, *Pitzer in Nepal*.
Examined the interaction of language use and pedagogical techniques in rural Nepali schools.

Publications

Below, ** denotes equal contribution.

Preprints

Li Wenliang**, Dougal J. Sutherland**, Heiko Strathmann, and Arthur Gretton. "Learning deep kernels for exponential family densities." 2018. arXiv: 1811.08357.

Journal and low-acceptance-rate conference papers

Michael Arbel**, Dougal J. Sutherland**, Mikołaj Bińkowski, and Arthur Gretton. "On gradient regularizers for MMD GANs." *Neural Information Processing Systems (NeurIPS)*. 2018. arXiv: 1805.11565.

Mikołaj Bińkowski**, Dougal J. Sutherland**, Michael Arbel, and Arthur Gretton. "Demystifying MMD GANs." *International Conference on Learning Representations (ICLR)*. 2018. arXiv: 1801.01401.

Dougal J. Sutherland**, Heiko Strathmann**, Michael Arbel, and Arthur Gretton. "Efficient and principled score estimation with Nyström kernel exponential families." *Artificial Intelligence and Statistics (AISTATS)*. 2018. arXiv: 1705.08360. Selected for oral presentation.

Ho Chung Leon Law**, Dougal J. Sutherland**, Dino Sejdinovic, and Seth Flaxman. "Bayesian Approaches to Distribution Regression." *Artificial Intelligence and Statistics (AISTATS)*. 2018. arXiv: 1705.04293.

Dougal J. Sutherland, Hsiao-Yu Tung, Heiko Strathmann, Soumyajit De, Aaditya Ramdas, Alex Smola, and Arthur Gretton. "Generative Models and Model Criticism via Optimized Maximum Mean Discrepancy." *International Conference on Learning Representations (ICLR)*. 2017. arXiv: 1611.04488.

Michelle Ntampaka, Hy Trac, Dougal J. Sutherland, Sebastian Fromenteau, Barnabás Póczos, and Jeff Schneider. "Dynamical Mass Measurements of Contaminated Galaxy Clusters Using Machine Learning." *The Astrophysical Journal* 831.2 (2016), p. 135. arXiv: 1509.05409.

Dougal J. Sutherland**, Junier B. Oliva**, Barnabás Póczos, and Jeff Schneider. "Linear-time Learning on Distributions with Approximate Kernel Embeddings." *AAAI Conference on Artificial Intelligence (AAAI)*. 2016. arXiv: 1509.07553.

Dougal J. Sutherland and Jeff Schneider. "On the Error of Random Fourier Features." *Uncertainty in Artificial Intelligence (UAI)*. 2015. arXiv: 1506.02785.

Yifei Ma**, Dougal J. Sutherland**, Roman Garnett, and Jeff Schneider. "Active Pointillistic Pattern Search." *Artificial Intelligence and Statistics (AISTATS)*. 2015.

Michelle Ntampaka, Hy Trac, Dougal J. Sutherland, Nicholas Battaglia, Barnabás Póczos, and Jeff Schneider. "A Machine Learning Approach for Dynamical Mass Measurements of Galaxy Clusters." *The Astrophysical Journal* 803.2 (2015), p. 50. arXiv: 1410.0686.

Dougal J. Sutherland, Barnabás Póczos, and Jeff Schneider. "Active learning and search on low-rank matrices." *Knowledge Discovery and Data Mining (KDD)*. 2013. Selected for oral presentation.

Barnabás Póczos, Liang Xiong, Dougal J. Sutherland, and Jeff Schneider. "Nonparametric kernel estimators for image classification." *Computer Vision and Pattern Recognition (CVPR)*. 2012.

Andrew Stromme, Dougal J. Sutherland, Alexander Burka, Benjamin Lipton, Nicholas Felt, Rebecca Roelofs, Daniel-Elia Feist-Alexandrov, Steve Dini, and Allen Welkie. "Managing User Requests with the Grand Unified Task System (GUTS)." *Large Installation System Administration (LISA)*. 2012.

Peer-reviewed workshop and high-acceptance-rate conference contributions

Ho Chung Leon Law**, Dougal J. Sutherland**, Dino Sejdinovic, and Seth Flaxman. "Bayesian Approaches to Distribution Regression." *Learning on Distributions, Functions, Graphs and Groups (NeurIPS workshop)*. 2017. Selected for oral presentation.

Jay Jin, Kyle Miller, Dougal J. Sutherland, Simon Labov, Karl Nelson, and Artur Dubrawski. "List Mode Regression for Low Count Detection." *IEEE Nuclear Science Symposium (IEEE NSS/MIC)*. 2016.

Dougal J. Sutherland**, Junier B. Oliva**, Barnabás Póczos, and Jeff Schneider. "Linear-time Learning on Distributions with Approximate Kernel Embeddings." *Feature Extraction: Modern Questions and Challenges (NeurIPS workshop)*. 2015.

Yifei Ma**, Dougal J. Sutherland**, Roman Garnett, and Jeff Schneider. "Active Pointillistic Pattern Search." *Bayesian Optimization (NeurIPS workshop)*. 2014.

Technical reports

Dougal J. Sutherland. "Fixing an error in Caponnetto and de Vito (2007)." 2017. arXiv: 1702.02982.

Seth Flaxman, Dougal J. Sutherland, Yu-Xiang Wang, and Yee Whye Teh. "Understanding the 2016 US Presidential Election using ecological inference and distribution regression with census microdata." 2016. arXiv: 1611.03787.

Junier B. Oliva**, Dougal J. Sutherland**, Barnabás Póczos, and Jeff Schneider. "Deep Mean Maps." 2015. arXiv: 1511.04150.

Dougal J. Sutherland, Liang Xiong, Barnabás Póczos, and Jeff Schneider. "Kernels on Sample Sets via Nonparametric Divergence Estimates." 2012. arXiv: 1202.0302.

Junier B. Oliva, Dougal J. Sutherland, and Yifei Ma. "Finding Representative Objects with Sparse Modeling." CMU 10-725 Optimization course project. 2012. Best poster award.

Matthew Bodenhamer, Thomas Palmer, Dougal J. Sutherland, and Andrew H. Fagg. "Grounding Conceptual Knowledge with Spatio-Temporal Multi-Dimensional Relational Framework Trees." 2012.

Invited Talks

Sep 2018 *Kernel Distances for Better Deep Generative Models*.
Advances in Kernel Methods (workshop at the Gaussian Process Summer School, GPSS).

Jun 2018 *Better GANs by using the MMD*.
Facebook AI Research New York.

Jun 2018 *Efficiently Estimating Densities and Scores with Kernel Exponential Families*.
Gatsby Tri-Center Meeting.

Jun 2018 *Better GANs by using the MMD*.
Machine Learning reading group, Google New York.

Jun 2018 *Better GANs by using the MMD*.
Machine Learning reading group, Columbia University.

May 2018 *Advances in GANs based on the MMD*.
Machine Learning Seminar, University of Sheffield.

Dec 2017 *Efficient and principled score estimation with kernel exponential families*.
Approximating high dimensional functions (workshop at the Alan Turing Institute).

Dec 2017 *Efficient and principled score estimation with kernel exponential families*.
Computational Statistics and Machine Learning seminar, University College London.

Aug 2017 *Evaluating and Training Implicit Generative Models with Two-Sample Tests*.
Implicit Models (workshop at the International Conference on Machine Learning, ICML).

Apr 2017 *Two-Sample Tests, Integral Probability Metrics, and GAN Objectives*.
Theory of Generative Adversarial Networks (workshop at Data Analysis, Learning, and Inference, DALI).

Feb 2017 *Generative Models and Model Criticism via Optimized Maximum Mean Discrepancy*.
Computational Statistics and Machine Learning seminar, Oxford University.

Relevant Teaching Experience

Guest lectures.

- "New Kernel Distances for Better Deep Generative Models." December 2018.
  Advanced Topics in Machine Learning, University College London.
- "What Is Machine Learning?" April 2016.
  Capstone Course, Jackson Institute for Global Affairs, Yale University.
- "What Is Machine Learning?" December 2014.
  Capstone Course, Jackson Institute for Global Affairs, Yale University.

Spring 2014 **Teaching Assistant**, *15-853 Algorithms in the Real World*, Carnegie Mellon University.
Ph.D.-level course on algorithms with real-world applications. (Guy Blelloch and Anupam Gupta)

Fall 2013 **Teaching Assistant**, *10-701 Machine Learning*, Carnegie Mellon University.
Introductory Ph.D.-level course in machine learning. (Alex Smola and Geoff Gordon)

Summer 2011 **Teaching Assistant and Residential Mentor**, *Summer Science Program*, `ssp.org`.
Intense five-week program for high schoolers from around the world, who learned programming, vector calculus, and astronomy to determine an asteroid's orbit from their own observations.

2009 – 2011 **Editor-in-Chief**, *The Daily Gazette*, Swarthmore.
Supervised staff in writing and editing news stories, as well as managing all newspaper operations.

Service

2015 – **Program committee or equivalent**, *NeurIPS, ICML, ICLR, AISTATS, AAAI*.
NeurIPS 2018: top 216 (of 3,045) reviewers. ICML 2018: Outstanding Reviewer.

2014 – **Reviewer**, *JMLR, IEEE TSP, IEEE T-PAMI, MLJ, COLT, SoCG, IJCAI, ECML-PKDD*.

2017 **Session chair**, *ICML*.

2016 – **External seminar organizer**, *Gatsby*.

2013 **Immigration Course organizer**, *CMU*.

2015 – **Top 50 annual contributor**, *Cross Validated*, `stats.stackexchange.com`.

2018 – **Core member**, `conda-forge.org`, scientific software packaging ecosystem.

Other

Programming Expert: Python scientific/deep learning ecosystem. Experienced: C/C++, web languages.

Languages Native English; practical Nepali; coursework in Chinese, Arabic, ASL, and Latin.

Citizenship U.S.

Last update: 12 December, 2018.