

# Dougal J. Sutherland

---

## Contact Information

Carnegie Mellon University  
Computer Science Department  
5000 Forbes Ave  
Pittsburgh, PA 15213

*cell:* 856 · 336 · 8425  
*office:* 412 · 268 · 3046  
*email:* dsuther1@cs.cmu.edu  
*web:* cs.cmu.edu/~dsuther1

---

## Education

- expected 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.
- 2015 **M.S., Computer Science**, *Carnegie Mellon University*.
- 2011 **B.A., Computer Science**, *Swarthmore College*, with high honors.  
Minors in Linguistics (with high honors) and Mathematics & Statistics. GPA: 3.92.

---

## Honors and Awards

- Sep. 2014 – Sandia Campus Executive Program fellowship. (Renewed in 2015.)
- March 2013 National Science Foundation Graduate Research Fellowship Program: Honorable Mention.
- May 2011 Ivy Award for “the senior man outstanding in leadership, scholarship, and contributions to the college community” by Swarthmore faculty vote.
- May 2011 Elected Phi Beta Kappa.
- May 2011 Drew Pearson Prize for excellence in journalism.

---

## Research and Academic Experience

- 2013 – **XDATA workshops**, *DARPA*.  
Addressed challenge problems about 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, particularly with Casey King and Michael Kane of Phronesis, LLC.
- 2011 – **Ph.D. research**, *Carnegie Mellon University*.  
Research in machine learning with Jeff Schneider; frequent collaboration with Barnabás Póczos. Particular focus on machine learning on samples from distributions and on active learning problems. In addition to work represented by the publications below, I have done unpublished empirical work related to learning on distributions in the analysis of financial anomalies, fusion reactor behavior, web browsing traffic, shipping behavior,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.
- Spring 2011 **Linguistics senior honors study**, *Swarthmore College*.  
Analyzed the phonotactics of Chaha, including computational approaches, for Colleen Fitzgerald.
- Summer 2010 **REU in machine learning**, *University of Oklahoma*.  
Worked on a relational concept learning system with Andrew Fagg.
- Summer 2009 **Howard Hughes Medical Institute fellowship**, *Swarthmore College*.  
Worked on natural language processing and medical information extraction with Rich Wicentowski.

Spring 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.

In submission

**Deep Mean Maps.**

J B Oliva\*, D J Sutherland\*, B Póczos, and J Schneider.  
arXiv:1511.04150.

**Dynamical Mass Measurements of Contaminated Galaxy Clusters Using Machine Learning.**

M Ntampaka, H Trac, D J Sutherland, S Fromenteau, B Póczos, and J Schneider.  
arXiv:1509.05409.

Peer-reviewed conference and journal papers

**Linear-time Learning on Distributions with Approximate Kernel Embeddings.**

D J Sutherland\*, J B Oliva\*, B Póczos, and J Schneider.  
Association for the Advancement of Artificial Intelligence (AAAI), 2016.

**On the Error of Random Fourier Features.**

D J Sutherland and J Schneider.  
Uncertainty in Artificial Intelligence (UAI), 2015.

**Active Pointillistic Pattern Search.**

Y Ma\*, D J Sutherland\*, R Garnett, and J Schneider.  
Artificial Intelligence and Statistics (AISTATS), 2015.

**A Machine Learning Approach for Dynamical Mass Measurements of Galaxy Clusters.**

M Ntampaka, H Trac, D J Sutherland, N Battaglia, B Póczos, and J Schneider.  
The Astrophysical Journal, 2015. 803, 50.

**Active Learning and Search on Low-Rank Matrices.**

D J Sutherland, B Póczos and J Schneider.  
ACM Special Interest Group on Knowledge Discovery and Data Mining (KDD), 2013.

**Nonparametric Kernel Estimators for Image Classification.**

B Póczos, L Xiong, D J Sutherland, and J Schneider.  
Computer Vision and Pattern Recognition (CVPR), 2012.

**Managing User Requests with the Grand Unified Task System (GUTS).**

A Stromme, D J Sutherland, A Burka, B Lipton, N Felt, R Roelofs, D-E Feist, S Dini, A Welkie.  
USENIX Large Installation System Administration (LISA), 2012.

Peer-reviewed workshop papers

**Linear-time Learning on Distributions with Approximate Kernel Embeddings.**

D J Sutherland\*, J B Oliva\*, B Póczos, and J Schneider.  
Feature Extraction workshop at NIPS, 2015.

**Active Pointillistic Pattern Search.**

Y Ma, D J Sutherland, R Garnett, and J Schneider.  
Bayesian Optimization workshop at NIPS, 2014.

## Technical reports

### **Finding Representative Objects with Sparse Modeling.**

J Oliva, D J Sutherland, and Y Ma.

CMU 10-725 Optimization course project, 2012. Best poster award.

### **Kernels on Sample Sets via Nonparametric Divergence Estimates.**

D J Sutherland, L Xiong, B Póczos, and J Schneider.

arXiv:1202.0302, 2012.

### **Grounding conceptual knowledge with spatiotemporal multidimensional relational framework trees.**

M Bodenhamer, T Palmer, D J Sutherland, and A H Fagg.

University of Oklahoma School of Computer Science Tech. Rep. TR-AIR-1138, 2012.

### **Integrating Human Knowledge into a Relational Learning System.**

D J Sutherland.

Swarthmore College B.A. thesis, 2011.

---

## Teaching Experience

- 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 and Residential Assistant**, *The Summer Science Program*, Santa Barbara.  
Assisted rising high school seniors from around the world in an intense five-week residential program. Students determined the orbit of a near-Earth asteroid based on their own observations, as well as learning the necessary computer programming, vector calculus, and astronomy to do so.
- 2009 – 2011 **Editor-in-Chief**, *The Daily Gazette*, Swarthmore, PA.  
Supervised staff in writing and editing news stories, as well as managing all newspaper operations.
- 2008 – 2011 **Lead Web Developer**, *The Daily Gazette*, Swarthmore, PA.  
Led small teams in developing an award-winning newspaper site and a campus announcement site.

---

## Graduate Coursework

CMU	F2013	Deep Learning	B Raj	A
CMU	S2013	Optimizing Compilers for Modern Architectures	T Mowry	A
CMU	F2012	Optimization	G Gordon, R Tibshirani	A+
CMU	F2012	Intermediate Statistics	L Wasserman	A+
CMU	S2012	Graduate Algorithms	M Blum	A–
CMU	S2012	Semantics of Programming Languages	S Brookes	A
CMU	F2011	Machine Learning	E Xing	A+
CMU	F2011	Computational Models of Neural Systems	D Touretzky	A
UPenn	S2010	Software Foundations	B Pierce	A+

---

## Service

- 2014 – **Reviewer**, *NIPS*, *ICML*, *IEEE T-PAMI*.
- 2015 – **Top 50 yearly contributor**, *Cross Validated*, `stats.stackexchange.com`.
- 2013 **Immigration Course organizer**, *CMU*.

---

## Other

Programming	Thorough knowledge of Python scientific stack. Experienced with C/C++, Matlab, and web languages. Have extended scikit-learn, Caffe, Django, LLVM, Postgres, and others.
Software	Standard Unix and Macintosh systems, Git, SVN, L <sup>A</sup> T <sub>E</sub> X. System administration on Debian.
Languages	Practical Nepali; coursework in Chinese, Arabic, ASL, and Latin.
Citizenship	U.S.

*Last update: February 1, 2016.*