

D. Sculley

Department of Computer Science, Tufts University
161 College Ave, Medford, MA
dsculley@cs.tufts.edu
(617) 899-8628

Education

Tufts University, Ph.D Candidate in Computer Science. Research in machine learning, data mining, and information retrieval. Dissertation work on information filtering problems such as spam filtering, with focus on efficient, online machine learning methods for large-scale systems. Other work on digital libraries and online advertising. Dissertation defense date: June 17th, 2008.

Tufts University, M.S. in Computer Science. Research in use of compression algorithms for text classification. Coursework in machine learning, learning theory, algorithms, and computational theory. Diploma received May, 2005.

Harvard University, M.Ed in School Development. Focus on educational leadership and innovative practice. Diploma received June, 2002.

Harvard University, B.A. in Visual and Environmental Studies. Diploma received June, 1997, *magna cum laude*.

Work Experience

Tufts University. (Medford, MA). Research assistant and teaching assistant in machine learning, focusing on problems in spam detection, computational biology, and support vector machines. January, 2005 through present (excluding internships).

Google, Inc. (Pittsburgh, PA). Internship, applying data mining and machine learning methods for problems in online advertising. Fall, 2007.

Yahoo!, Inc. (Sunnyvale, CA.) Research internship, performing data mining and machine learning for optimizing online advertising. Spring, 2006.

Perseus Digital Library. (Medford, MA) Researcher developing algorithms for automatic word sense disambiguation and personalized reading support. Jan. 2005-Jan. 2006.

Webb Schools of California. (Claremont, CA) High school teacher, developed innovative curriculum introducing computer science, logic, and information theory. 2000-'01

The American School in Switzerland. (Lugano, Switz.) High school teacher of history, English. 1999.

Escuela Camp Alegre. (Caracas, Venezuela) High school teacher of math and phys. ed. 1998.

American Community School of Abu Dhabi. (Abu Dhabi, U.A.E.) High school teacher of computer science. 1997-'98.

Publications

- D. Sculley and Bradley M. Pasanek . Meaning and Mining: The Impact of Implicit Assumptions in Data Mining for the Humanities. To appear in: *Literary and Linguistic Computing*. 2008.
- Bradley M. Pasanek and D. Sculley. Mining Millions of Metaphors. To appear in: *Literary and Linguistic Computing*. 2008.
- Xintao Wei, Lenore Cowen, Carla Brodley, Arthur Brady, D. Sculley, and Donna Slonim. A Distance-Based Method for Detecting Horizontal Gene Transfer in Whole Genomes. *ISBRA 2008: International Symposium on Bioinformatics*. May, 2008.
- D. Sculley and Gabriel M. Wachman. Relaxed Online SVMs in the TREC Spam Filtering Track. To Appear in *TREC 2007: The Sixteenth Text Retrieval Conference Proceedings*. 2007.
- D. Sculley. Practical Learning from One Sided Feedback. *KDD 2007: Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2007.
- D. Sculley. Online Active Learning Methods for Fast Label-Efficient Spam Filtering. *CEAS 2007: Proceedings of the Fourth Conference on Email and Anti-Spam*. 2007.
- D. Sculley and Gabriel M. Wachman. Relaxed Online SVMs for Spam Filtering. *SIGIR, 2007: Proceedings of the 30th Annual International ACM SIGIR Conference*. 2007. (Best student paper award, SIGIR 2007.)
- D. Sculley. Rank Aggregation for Similar Items. *Proceedings of 2007 SIAM International Conference on Data Mining*. 2007.
- D. Sculley, Gabriel M. Wachman, and Carla E. Brodley. Spam Classification with On-line Linear Classifiers and Inexact String Matching Features. *TREC 2006: The Fifteenth Text Retrieval Conference Proceedings*. 2006.
- D. Sculley and Carla E. Brodley. Compression and Machine Learning: A New Perspective on Feature Space Vectors. *Data Compression Conference 2006*.
- Gregory Crane, David Bamman, Lisa Cerrato, Alison Jones, David M. Mimno, Adrian Packel, D. Sculley, Gabriel Weaver. Beyond Digital Incunabula: Modeling the Next Generation of Digital Libraries. *ECDL*, 2006.

Invited Talks and Seminars

- University of Waterloo. Waterloo, Ontario, Canada. February, 2008.
- AOL, Inc. Dulles, Virginia. February, 2008.
- Carnegie Mellon University. Pittsburgh, PA. September, 2007.
- MITRE Corporation. Bedford, MA. August, 2007.
- Google, Inc. Mountain View, CA. February, 2006.

Patents

“Systems and Methods for Identifying Unwanted or Harmful Electronic Text.” Inventors: D. Sculley, Gabriel M. Wachman, Carla E. Brodley. US Patent application filed in July, 2006, on behalf of Tufts University.

“Keyword Set and Target Audience Profile Generalization Techniques.” Inventors: Usama M. Fayyad, Pavel Berkin, Andrew Tomkins, Rajesh Girish Parekh, Jignashu Parikh, and D. Sculley. US Patent application filed in May, 2006, on behalf of Yahoo!, Inc.

Awards

Outstanding Graduate Researcher, Tufts University. 2008.

Best Student Paper, SIGIR. 2007.

Magna cum laude. Harvard University, 1997.