

Carla E. Brodley

May, 2014

30 Gloucester St, Unit 3
Boston, MA 02155
cell phone: 617-913-7658

work phone: 617-627-3652
email: brodley@cs.tufts.edu
<http://www.cs.tufts.edu/brodley/>

Education:

B.A. 1985 McGill University, Mathematics and Computer Science
M.S. 1991 University of Massachusetts at Amherst, Computer Science
Ph.D. 1994 University of Massachusetts at Amherst, Computer Science

Professional Experience:

05/84 – 08/84 Summer Research Assistant, Image Analysis Lab, Dana Farber Cancer Inst.
09/85 – 09/86 Research Associate, Jensen Associates, Boston, MA
10/86 – 02/88 Systems Analyst, The Wyatt Company, Wellesley, MA
02/88 – 06/88 Consultant, Unisys Corporation, Cambridge, MA
09/88 – 08/94 RA, Department of Computer Science, Univ. of Massachusetts at Amherst
12/94 – 08/00 Assistant Professor, Electrical and Computer Engineering, Purdue University
08/00 – 08/05 Associate Professor, Electrical and Computer Engineering, Purdue University
08/04 – 10/04 Visiting Professor, Department of Computer Science, Tufts University
10/04 – Present Professor, Department of Computer Science, Tufts University
07/05 – 06/06 Acting Chair, Department of Computer Science, Tufts University
09/10 – 08/13 Chair, Department of Computer Science, Tufts University
10/11 – Present Professor, Clinical and Translational Science Institute, Tufts Medical Center

Honors:

1. NSF-CAREER Award, 1998-2002
2. Ruth and Joel Spira outstanding teacher award, Purdue University, 1998
3. Member of the 2004-2005 Defense Science Study Group, DARPA
4. Elected Member, AAAI Executive Council, 2008-2011
5. Alumni Award: Outstanding Educator, Department of Computer Science, the University of Massachusetts at Amherst, 2010
6. Elected Member, International Machine Learning Society Board, 2011-present
7. AAAI Fellow, 2014

Summary of Scholarship:

- Google Scholar Statistics on May 30, 2014: h-index = 44, i10-index = 74, total citations = 8310
- Publications: 26 journal articles, 47 highly-refereed conference papers, 6 magazine articles, 3 book chapters, 25 symposia and workshop papers

Research Grants and Contracts Received:

1. *Jet Propulsion Laboratory*, \$9,978, Principal Investigator: Carla Brodley, "Development of Pattern Recognition and Machine Learning Techniques to Detect and Characterize Structure in Images of Scientific Interest," 6/17/96-10/30/96.
2. *IBM*, \$48,212, Principal Investigator: Carla Brodley, "Knowledge Discovery in Databases," 6/15/96-6/14/98.
3. *SGI*, \$50,000, Principal Investigator: Carla Brodley, "Purdue-SGI Collaboration in Data Mining," 1/7/97-12/15/97.
4. *NSA*, Role: Co-PI, (my portion: \$116,476), "A Proposal on Autonomous Security Agents and Vulnerability Database," 1/7/97-9/30/99, (PI: Gene Spafford, Computer Science, Purdue University).
5. *Showalter Trust*, \$49,877, Co-Principal Investigators: Carla Brodley and Avi Kak, "Content-Based Image Retrieval for Large-Scale Medical Databases," 7/1/97-6/30/98.
6. *NSF*, \$526,417, Principal Investigator: Carla Brodley, "Content Based Image Retrieval for Medical Databases," 9/1/97-8/30/00, (Co-PIs: Avi Kak, ECE, Purdue University and Alex Aisen, Indiana University Medical School).
7. *NASA*, \$81,113, Principal Investigator: Carla Brodley, "Machine Learning to Improve Land Cover Classifications from Multisensor and Multitemporal Data," 3/1/98-2/28/01, (Co-PI's: with separate proposals and separate funding: Mark Friedl, Geography, Boston University and Ruth DeFries, Geography, University of Maryland).
8. *NSF-CAREER*, \$236,729, Principal Investigator: Carla Brodley, "A Foundation for Applied Machine Learning," 9/1/98-8/31/02.
9. *NIH*, \$1,067,940, Principal Investigator: Carla Brodley, "Content-Based Image Retrieval for Medical Databases," 8/1/98-7/31/01, (Co-PIs: Avi Kak, ECE, Purdue University, Alex Aisen, Indiana University Medical School and Lynn Broderick, Radiology, University of Wisconsin).
10. *Schlumberger*, \$30,000, Principal Investigator: Carla Brodley, "Automatic Detection of Denial of Service Attacks," 9/01/01-9/01/02.
11. *NSF*, \$20,000, Principal Investigator: Carla Brodley, "NSF Student Scholarship Program for the International Conference on Machine Learning," 2001.
12. *NASA*, \$750,872, Principal Investigator: Carla Brodley, "Machine Learning and Data Mining for Intelligent Data Understanding of High Dimensional Earth Science Data," 03/01/01-02/28/04 (Subcontract to Prof. M. Friedl, Department of Geography, Boston University).
13. *AFRL*, \$60,000, Principal Investigator: Carla Brodley, "Automatic Detection of Anomalous Behavior in Networks," 10/1/02-9/30/03 (Co-PI: Catherine Rosenberg).
14. *CISCO*, \$40,000, Co-PI's: Carla Brodley and Catherine Rosenberg, "Lightweight Software-only Anomaly Technique to Identify Pervasive Network Problems," 9/1/02-8/31/03.
15. *AFRL*, \$109,000, Principal Investigator: Carla Brodley "User Re-authentication via Mouse Behavior," 8/1/03-7/31/05.

16. *NSF-SGER*, \$50,658, Principal Investigator: Carla Brodley, “Behavioral Authentication of Server Flows,” 9/1/03-8/31/04.
17. *NSF-NRT*, \$819,000 Role: Co-PI (my portion: \$200,000), “Collaborative research: Testing and Benchmarking Methodologies for Future Networking Security Mechanisms,” 8/20/03-8/19/06.
18. *NSF-SGER*, \$61,791, Principal Investigator: Carla Brodley, “SGER: Mining for Planets,” 8/1/05-7/31/06.
19. *AFRL*, \$86,353, Principal Investigator: Carla Brodley “Automatic Detection of Covert Channels in Networks,” 6/6/05-8/31/06.
20. *DARPA*, \$298,042, Principal Investigator: Carla Brodley, “Detecting Natural versus Unnatural DNA,” 7/15/06-7/14/07 (Co-PI’s Lenore Cowen, Donna Slonim, Computer Science, Tufts University and Jonathan Eisen, Genomics, UC-Davis).
21. *NIH*, \$12,484,287, Role: Co-PI (my portion \$572,614), “Microsensor Arrays for Saliva Diagnostics,” 9/1/06-8/31/11, (PI: David Walt, Chemistry, Tufts University, also with BU Dental and Medical Schools).
22. *NSF*, \$212,000, Principal Investigator: Carla Brodley, “SEI: Collaborative Research: Discovering Unexpected Planets and Other Astronomical Oddities,” 8/01/07-7/31/10 (companion proposal with separate funding, Charles Alcock, CFA, Harvard University).
23. *NASA*, \$120,000, Role: Co-PI (my portion \$52,000) “Discovering Anomalies in Astrophysics,” 8/01/07-7/31/09 (PI: Charles Alcock, CFA, Harvard University).
24. *NSF*, \$875,824, Principal Investigators: Carla Brodley and Roni Khardon, “III-CXT-Medium: Interdisciplinary Machine Learning Research and Education,” 09/15/08-8/31/13.
25. *NIH*, \$1,289,608, Role: Faculty (my portion: 1/2 of an RA, one summer month of salary), “Translational Bioinformatics for Developmental Genomics,” 12/01/08-11/30/12, (PI: Donna Slonim, Computer Science, Tufts University; Co-PI: Diana Bianci, Medical School, Tufts University).
26. *NIH*, Role: Faculty, (my portion: \$125,000), “Clinical and Translational Science Award (CTSA),” 5/1/08-4/31/11, (PI: Harry Selker, Tufts Medical Center).
27. *AHRQ*, \$1,135,531 Role: Co-PI, (my portion: 20% academic year support and 1 research scientist), “Semi-Automated Abstract Screening for Comparative Effectiveness Reviews,” 9/1/09-8/31/12, (PI: Tom Trikalinos, Tufts Medical Center),
28. *Draper Lab*, \$131,301, Principal Investigator: Carla Brodley, “User-Guided Clustering: A Synergy of Human Interaction, Machine Learning and Visualization,” 7/1/12-6/30/13.
29. *Multiple Sclerosis Society*, \$34,000, Principal Investigator: Carla Brodley, “Machine Learning for Outcome Prediction in Multiple Sclerosis,” 3/25/13-3/01/14.
30. Under Review:
 - *AHRQ*, Role: Co-PI, (PI: Tanuja Chitnis, Harvard Medical School), “Use of Machine-learning Techniques to Predict Disease Course in Multiple Sclerosis,” to be submitted July 2014.
 - *NIH*, Role: Co-PI (PI: Thomas Thesen, NYU Neuroscience and Epilepsy Center), “Quantitative Imaging Characteristics of Focal Cortical Dysplasia,” submitted November 2013.

- *NSF*, Role: Co-PI (PI: Matthias Scheutz, Computer Science, Tufts), “RI: Small: Online Learning Algorithms for Automatic System Modeling, Fault Detection and Recovery in Autonomous Systems,” submitted January 2014.

Professional Leadership Roles:

- Co-Chair, (with R. Kohavi), The KDD-CUP, 2000.
- Program Co-Chair and General Co-Chair, (with Andrea Danyluk) the Eighteenth International Conference on Machine Learning (ICML), Williamstown, MA, June 28 - July 2, 2001.
- General Chair, the Twenty-First International Conference on Machine Learning (ICML), Banff, Alberta, July 4-8, 2004.
- Co-Chair, (with Kathleen Fisher) Committee on the Status of Women in Computing Research (CRA-W), 2008 - 2011.
- Program Co-Chair and General Co-Chair, (with Peter Stone) AAI-2014, Quebec City, Canada, July 2014.

National/International Boards:

- Committee on the Status of Women in Computing Research (CRA-W), 2002 - 2013.
- AAI Council, 2008-2011 (elected).
- Computing Research Association Board of Directors, 2008 - 2012 (AAI Liason).
- International Machine Learning Society Board, 2011 - present (elected).
- Information Science and Technology Board, DARPA, 2013 - present.

Purdue University Leadership Roles:

- Elected as ECE representative to SOE Dean Search Committee, 2000
- Elected to Head of ECE Search Committee, 2001

Tufts University Leadership Roles:

- Acting Chair, Department of Computer Science 2005-2006
- Chair, Department of Computer Science, 2010-2013
- Member, Dean Search committee, Arts and Sciences 2004-2005
- Member, Tufts President Search Committee, 2010
- Member, Tufts Effectiveness in Administrative Management Project, 2013 (one of two faculty members appointed to this committee)

Professional Service Activities:

Associate Editorships: *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1999 - 2002; *Journal of Artificial Intelligence Research*, 2002 - 2005; *Machine Learning*, 2004 - 2011; *Journal of Machine Learning Research*, 2008 - 2011.

Journal Editorial Boards: *Data Mining and Knowledge Discovery*, 2008-Present; *Journal of Machine Learning Research*, 2000 - 2008, 2011 - present; *Machine Learning*, 1997 - 2001, 2011 - present. *Journal of Artificial Intelligence Research*, 1999 - 2001.

Conference/Workshop Organizer: National Academy of Engineering's sixth Annual Symposium on Frontiers of Engineering, Irvine, CA September, 14-16, 2000; VISSEC/DMSEC Workshop held in conjunction with CCS, Washington DC, 2005 (Co-Chair); Snowbird, 2012; Meaningful Use of Complex Medical Data, 2013, 2014 (Co-Chair).

Tutorial Chair: The Eleventh International Conference on Knowledge Discovery and Data Mining, 2005.

Senior Program Committee Member: IEEE Conference on Computer Vision and Pattern Recognition, 1999; International Conference on Machine Learning, 2005, 2008, 2009, 2010, 2011; European Conference on Machine Learning, 2005, 2012, 2013; National Conference on Artificial Intelligence (AAAI), 2007, 2010, 2013; International Conference on Knowledge Discovery and Databases, 2006, 2009, 2010, 2012, 2013; International Health Informatics Symposium, 2011; ICHI 2013.

Conference Program Committees: International Conference on Machine Learning, 1996, 1997, 1998, 2000, 2002, 2003, 2007, 2008; National Conference on Artificial Intelligence, 1996, 1998, 1999, 2000, 2008; International Conference on Knowledge Discovery and Data Mining, 1997, 1998, 2003, 2004, 2007, 2008; European Conference on Machine Learning, 2007, 2010; Annual ACM SIGIR Conference 2008; KDD 2012 Workshop on Health Informatics.

Diversity Outreach:

- Member of the Committee on the Status of Women in Computing Research (CRA-W), 2002-2013. Activities included: Editor of "Expanding the Research Pipeline," *Computing Research News*, 2002-2005; Fundraising Co-Chair (with Lori Clarke and Carla Ellis), 2005-2008; Co-Chair (with Kathleen Fisher), 2008-2011.
- Speaker/Panelist/Mentor at: CRA-W Academic Careers Workshops, 1996, 1999, 2005; CRA-W Graduate Cohort Workshop 2004, 2006, 2011, 2013; Grace Hopper Conference, 2012.
- Faculty Advisor, Women in Computer Science, Tufts University, 2012-present.

Courses Taught:

Undergraduate:

1. Introduction to Computer Security, Fall 2002, Fall 2003, Spring 2005, Spring 2006
2. Information Extraction, Retrieval and Security, Fall 2000
3. Discrete Math for Computer Engineers, Spring 1995, Fall 2001

4. Introduction to Artificial Intelligence, Spring 1996, Spring 1997

Undergraduate/Graduate:

1. Introduction to Machine Learning, Fall 2007, Fall 2009, Fall 2011, Fall 2012, Fall 2013, Spring 2014
2. Research in Machine Learning, Spring 2012

Graduate:

1. Graduate-Level Machine Learning and Data Mining, Fall 1995, Spring 1997, Spring 1999, Spring 2001, Spring 2003, Fall 2004, Fall 2006
2. Graduate-Level Artificial Intelligence, Fall 1996, Fall 1997, Fall 1998, Fall 1999
3. Interdisciplinary Research in Computer Science, Chemistry and Biology, Fall 2006, Fall 2007, Spring 2008 (with David Walt, Chemistry, Tufts University)
4. Machine Learning for Predictive Medicine – co-listed with the medical school, Spring 2011 (with Chris Schmid, Sackler School of Graduate Biomedical Sciences, Tufts University).

Ph.D. Thesis Supervision Completed:

1. Craig Codrington, (co-chair Prof. M. Tenorio) May 1997, “Image Segmentation: A Competitive Approach,” currently a researcher in the Physics Department, Purdue University.
2. Terran Lane, August, 2000, “Machine Learning Techniques for the Computer Security Domain of Anomaly Detection,” previously a tenured Associate Professor, University of New Mexico, now at Google, Cambridge.
3. Jennifer Dy, May 2001, “Feature Selection for Unsupervised Learning Applied to Content-Based Image Retrieval,” currently a tenured Associate Professor, Northeastern University.
4. Xiaoli Fern Zhang, July 2005, “High Dimensional Data Clustering and Correlation Pattern Analysis,” currently a tenured Associate Professor, Oregon State.
5. James Early, July 2005, “Behavioral Authentication for Computer Security,” currently an Assistant Professor, SUNY, Oswego.
6. Serdar Cabuk, December 2006, “Network Covert Channels: Design, Analysis, Detection, and Elimination,” currently Director, IT Risk and Assurance at EY, London, UK.
7. Maja Pusara, (co-chair Gene Spafford) June 2007, “An Examination of User Behavior for Re-Authentication,” currently a consultant at Ab Initio.
8. D. Sculley, August 2008, “Advances in Online Learning-Based Spam Filtering,” currently a Software Engineer at Google, Pittsburgh (leading a group of 30 researchers at Google).
9. Rachel Lomasky, February 2010, “Active Acquisition of Informative Training Data,” currently an Analytics Director at Opera Solutions.
10. Umma Rebbapragada, July 2010, “Strategic Targeting of Outliers for Expert Review,” currently a Research Scientist, Jet Propulsion Laboratory.

11. Byron Wallace, June 2012, “Machine Learning in Health Informatics: Making Better use of Domain Experts,” currently an Assistant Professor, School of Information, University of Texas at Austin. Selected as *the Runner-Up for the 2013 ACM SIGKDD Doctoral Dissertation Award*.
12. Jingjing Liu, May 2014, “Clustering with Domain Knowledge,” currently on the job market.

Ph.D. Thesis Students Currently Being Supervised:

Bilal Ahmed (expected May 2015), Yijun Zhao, Eric Murray, Moses Huang, Yuqi Xian.

Journal Articles:

1. Draper, B., Brodley, C. E., and Utgoff, P. E., “Goal-directed classification using linear machine decision trees,” *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 16 (9), pp. 888-893, 1994.
2. Brodley, C. E. and Utgoff, P. E., “Multivariate decision trees,” *Machine Learning*, vol. 19, pp. 45-77, 1995.
3. Brodley, C. E., “Automatic bias selection for classifier construction,” *Machine Learning*, vol. 20, pp. 63-94, 1995.
4. Brodley, C. E. and Smyth, P., “Applying classification algorithms in practice,” *Statistics and Computing*, vol. 7, pp. 45-56, 1997.
5. Friedl, M. and Brodley, C. E., “Decision tree classification of land cover from remotely sensed data,” *Remote Sensing of Environment*, 61 (3), pp. 399-409, 1997.
6. Friedl, M., Brodley, C. E., and Strahler, A., “Maximizing land cover classification accuracies produced by decision trees at continental to global scales,” *IEEE Transactions on Geoscience and Remote Sensing*, 37 (2), pp. 969-977, 1999.
7. Shyu, C., Brodley, C. E., Kak, A., Kosaka, A., Aisen, A. and Broderick, L., “ASSERT, A physician-in-the-loop content-based image retrieval system for HRCT image databases,” *Computer Vision and Image Understanding*, vol. 74, pp. 111-132, 1999.
8. Brodley, C. E. and Friedl, M., “Identifying mislabeled training data,” *Journal of Artificial Intelligence Research*, vol. 11, pp. 131-167, 1999.
9. Lane, T. and Brodley, C. E., “Temporal sequence learning and data reduction for anomaly detection,” *ACM Transactions on Computer Security*, 2 (3), pp 295–331, 1999.
10. Bradford, J. P. and Brodley, C. E., “The effect of instance-space partition on significance,” *Machine Learning*, 42, pp 269–286, 2001.
11. Stough, T. M. and Brodley, C. E., “Focusing attention on objects of interest using multiple matched filters,” *IEEE Transactions on Image Processing*, 10 (3), pp. 419-426, 2001.
12. Shyu, C. R., Kak, A., Brodley, C. E., and Broderick, L., “Using human perceptual categories for content-based retrieval from a medical image database,” *Computer Vision and Image Understanding*, vol. 88, pp. 119-151, 2002.
13. MacArthur, S. Brodley C. E., and Broderick, L., “Interactive content-based image retrieval using relevance feedback,” *Computer Vision and Image Understanding*, vol. 88, pp. 55-75, 2002.

14. Dy, J., Brodley, C. E., Kak, A., Broderick, L. S., and Aisen, A. M., "Unsupervised feature selection applied to content-based image retrieval of lung images," *IEEE Transactions on Pattern Recognition and Machine Intelligence*, vol. 25, pp. 373-378, 2003.
15. Aisen, A. M. Broderick, L. S., Winer-Muram, H., Brodley, C. E., Kak, A. C., Pavlopoulou C., Dy, J., and Marchiori, A., "Automated storage and retrieval of medical images to assist diagnosis: Implementation and preliminary assessment," *Radiology*, 228 (1), pp. 265-270, 2003.
16. Lane, T. and Brodley, C. E., "An empirical study of two approaches to sequence learning for anomaly detection," *Machine Learning*, 51(1), pp. 73-107, 2003.
17. Dy, J. and Brodley, C. E., "Feature selection for unsupervised learning," *Journal of Machine Learning Research*, vol. 5, pp. 845-889, 2004.
18. Ozdoganoglu, H., Jalote, A., Vijaykumar, T. N., Brodley, C. E., and Kuperman, B. A., "Smash-Guard: A hardware solution to prevent security attacks on the function return address," *IEEE Transactions on Computers*, 55 (10), pp. 1271-1285, 2006.
19. Rebbapragada, U., Protopapas, P., Brodley, C. E., and Alcock, C. "Finding Anomalies in Periodic Time Series," *Machine Learning*, 74 (3), pp. 281-313, 2009.
20. Scabuk, C., Brodley, C. E. and Shields, C., "IP Covert Channel Detection," *ACM Transactions on Information and Systems Security (TISSEC)*, 12 (4), pp. 1-29, 2009.
21. Preston, D., Protopapas, P., and Brodley, C. E., "Discovering arbitrary events types in time series," *Statistical Analysis and Data Mining*, 2 (5), pp. 396-411, 2010.
22. Wallace, B., Trikalinos, T., Lau, J., Brodley, C., Schmid, C. "Semi-automated screening of biomedical citations for systematic reviews," *BMC Bioinformatics*, 11 (55), 2010.
23. Whitaker, R. D., Pember, S., Wallace, B. C., Brodley, C. E., Walt, D. R., "Single cell time resolved quorum responses reveal dependence on cell density and configuration," *Journal of Biological Chemistry*, vol. 286, pp. 21623-21632, 2011.
24. Noto, K., Slomin, D., and Brodley C. E., "FRaC: A feature-modeling approach for semi-supervised and unsupervised anomaly detection," *Data Mining and Knowledge Discovery*, 25 (1), pp. 109-133, 2011.
25. Wallace, B., Small, K., Brodley, C. E., Lau, J., Schmid, C., Bertram, Li., Lill, C., Cohen, J. and Trikalinos, T. "Towards modernizing the systematic review pipeline: Efficient updating via data mining," *Genetics in Medicine*, vol. 14, pp 663-669, 2012.
26. Liu, J., Healy, B., Chitnis, T., Brodley, C. E., "Finding homogeneous groups of multiple sclerosis patients by removing physician subjectivity via constraint-based clustering," to appear *AI and Medicine*.

Articles under Review:

1. Zhao, Y., Healy, B., Rotstein, D., Brodley, C. E., and Chitnis, T. "Predicting early clinical course in multiple sclerosis using machine learning," submitted to *Neurology*.
2. Ahmed, B., Thesen, T., Barah, G., Carlson, C., Kuzniecky, R., Doyle, W., Blackmon, K., Devinsky, O., and Brodley, C. E., "Machine learning methods for automated detection of MRI-elusive lesions: A surface-based MRI morphometric approach," submitted to *Epilepsia*.

3. Brodley, C. E., Liu, J., Khardon, R., Preston, D. Sulla-Menashe, D., and Friedl, M., "Refining class structure via class-level constraint-based clustering," submitted to *Machine Learning*.
4. Priesol, A. J., Cao, M., Brodley, C. E., Lewis, R. F., "Clinical vestibular testing assessed with machine learning algorithms," submitted to *Science Translational Medicine*

Magazine Articles:

1. Brodley, C. E., Lane, T., and Stough, T., "Knowledge discovery and data mining," *American Scientist*, 87 (1), pp. 54-61, 1999, (Invited article).
2. Kohavi, R., Brodley, C. E., Frasca, B., Mason, L., and Zheng, Z., "KDD-Cup 2000: Organizers' report," *SIGKDD Explorations*, 2 (2) pp. 86-98, 2000.
3. Bajcsy, R., et al. (one of 24 authors), "Cyber Defense Technology Networking and Evaluation," *Communications of the ACM*, 47 (3), pp. 58-61, 2004.
4. Kuperman, B., Brodley, C. E., Ozdoganoglu, H., Vijaykumar, T.N. and Jalote, A., "Detection and prevention of stack buffer overflow attacks," *Communications of the ACM*, 48 (11), pp. 51-56, 2005.
5. Pollack, M.E., et al., (one of 26 authors) "Computing: Report leaps geographical barriers but stumbles over gender," Correspondance, *Nature*, vol. 44, 2006.
6. Brodley, C. E., Rebbapragada, U., Small, K., and Wallace, B., "Challenges and opportunities in applied machine learning," *AI Magazine*, 33 (1), pp. 11-24, 2012.

Book Chapters:

1. Brodley, C. E. and Utgoff, P. E., "Dynamic recursive model class selection for classifier construction," In Cheeseman and Oldford (Eds.), *Selecting Models from Data*, pp. 329-338, Springer-Verlag, 1994. (A preliminary version appeared in *The Proceedings of the Fourth International AI and Statistics Workshop*.)
2. Brodley, C. E. and Clouse, J., "Machine Learning," In the *Encyclopedia of Computer Science and Technology*, vol. 41 (A. Kent and J. G. Williams, eds.), Marcel Dekker, New York, 1999, pp. 137-146.
3. Early, J. P. and Brodley, C. E. "Behavioral features for network anomaly detection," in M. Maloof, (Ed.) *Machine Learning and Data Mining for Computer Security: Methods and Applications*, Springer, 2005, pp. 107-124.

Conference Proceedings and Presentations: Full Paper Reviewed:

1. Utgoff, P. E. and Brodley, C. E., "An incremental method for finding multivariate splits in decision trees," *The Seventh Conference on Machine Learning*, June, 1990, Austin, TX, pp. 58-65.
2. Brodley, C. E., "Addressing the selective superiority problem: Automatic algorithm/model class selection," *The Tenth International Conference on Machine Learning*, June, 1993, Amherst, MA, pp. 17-24. (Note: one of 5% accepted for plenary session presentation.)

3. Brodley, C. E., "Automatic selection of split criterion during tree growing based on node location," *The Twelfth International Conference on Machine Learning*, July, 1995, Tahoe City, CA, pp. 73-80.
4. Brodley, C. E. and Friedl, M. A., "Identifying and eliminating mislabeled training instances," *The Thirteenth National Conference on Artificial Intelligence*. August 4-8, 1996, Portland, OR, pp. 799-805.
5. Stough, T. and Brodley, C. E., "Image feature reduction through spoiling: Its application to multiple matched filters for focus of attention," *The Third International Conference on Knowledge Discovery and Data Mining*, August, 1997, Newport Beach, CA, pp. 255-258.
6. Lane, T. and Brodley, C. E., "An application of machine learning to anomaly detection," *The Twentieth Annual National Information Systems Security Conference*, October, 1997, Washington, DC, pp. 366-380.
7. Moss, J. E., Utgoff, P., Cavozos, J., Precup, D., Stefanovic, D., Brodley, C. E., and Scheeff, D., "Learning to schedule straight-line code," *Neural Information Processing Systems*, December 1997, Denver, CO, pp. 929-935.
8. Bradford, J. P., Kunz, C., Kohavi, R., Brunk, C., and Brodley, C. E., "Pruning decision trees with misclassification costs," *The Tenth European Conference on Machine Learning*, April, 1998, Chemnitz, Germany, pp. 131-136.
9. Lane, T. and Brodley, C. E., "Approaches to online learning and concept drift for user identification in computer security," *The Fourth International Conference on Knowledge Discovery and Data Mining*, August, 1998, New York, NY, pp. 259-263.
10. Lane, T. and Brodley, C. E., "Temporal sequence learning and data reduction for anomaly detection" *The Fifth ACM Conference on Computer and Communications Security*, November, 1998, San Francisco, CA, pp. 150-158.
11. Dy, J. G., Brodley, C. E., Kak, A. Shyu, C., and Broderick, L. S., "The customized-queries approach to CBIR," *Storage and Retrieval for Image and Video Databases VII, IS&T/SPIE Electronic Imaging 1999*, vol. 3656, January, 1999, San Jose, CA, pp. 22-32.
12. Dy, J. G., Brodley, C. E., Kak, A. Shyu, C., and Broderick, L. S., "The customized-queries approach to CBIR using EM," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 2, June, 1999, Fort Collins, CO, pp. 400-406.
13. Zhou, Y. and Brodley, C. E., "A hybrid lazy-eager approach to reducing the computation and memory requirements of local parametric learning algorithms," *The Sixteenth International Conference on Machine Learning*, June, 1999, Bled, Slovenia, pp. 503-512.
14. Brodley, C. E., Kak, A. C., Dy, J. G., Shyu, C. R., Aisen, A., and Broderick, L., "Content-based retrieval from medical image databases: A synergy of human interaction, machine learning and computer vision," *The Sixteenth National Conference on Artificial Intelligence*, July, 1999, Orlando, FL, pp. 760-767.
15. Kapadia, N., Fortes, J. and Brodley, C. E., "Predictive application-performance modeling in a computational grid environment," *the Eighth IEEE International Symposium on High Performance Distributed Computing*, August, 1999, Redondo Beach, CA, pp. 47-54.

16. Dy, J. and Brodley, C. E., "Feature subset selection and order identification for unsupervised learning," *The Seventeenth International Conference on Machine Learning*, June, 2000, Stanford University, pp. 247-254.
17. Lane, T. and Brodley, C. E., "Data reduction techniques for instance-based learning of human/computer interface data," *The Seventeenth International Conference on Machine Learning*, June, 2000, Stanford University, pp. 519-526.
18. Dy, J. and Brodley, C. E., "Visualization and interactive feature selection for unsupervised data," *The ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, August, 2000, Boston, MA, pp. 360-364.
19. Fern, X. Z. and Brodley, C. E., "Boosting lazy decision trees," *The Twentieth International Conference on Machine Learning*, August 2003, Washington D.C., pp. 178-185.
20. Fern, X. Z. and Brodley, C. E., "Random Projection for high dimensional data clustering: A cluster ensemble approach," *The Twentieth International Conference on Machine Learning*, August 2003, Washington D.C., pp. 186-193.
21. Early, J., Brodley, C. E. and Rosenberg, C., "Behavioral Authentication of Server Flows," in the *The Nineteenth Annual Computer Security Applications Conference*, December 2003, Las Vegas, pp. 49-55.
22. Fern, X. Z. and Brodley, C. E., "Solving cluster ensemble problems by bipartite graph partitioning," *The Twentieth-First International Conference on Machine Learning*, July, 2004, Banff, Alberta.
23. Scabuk, C., Brodley, C. E., and Shields, T. C., "IP covert timing channels: An initial exploration," *The Eleventh ACM Conference on Computer and Communications Security*, October, 2004, Washington D.C., pp. 176-187.
24. Fern, X. Z., Brodley, C. E., and Friedl, M. A., "Correlation clustering for learning mixtures of canonical correlation models," *SIAM International Conference on Data Mining (SDM)*, April 2005, Newport Beach, CA, pp. 439-448.
25. Hasan, J., Jalote, A., Vijaykumar, T. N., and Brodley, C. E., "Heat stroke: Power-density-based denial of service in SMT," *Proceedings of the 11th International Symposium on High-Performance Computer Architecture*, December 2005, Goa, India, pp. 166-177.
26. Sculley, D. and Brodley, C. E., "Compression and machine learning: A new perspective on feature space vectors," *Proceedings of DCC06, IEEE Data Compression Conference*, March 2006, Brandeis University, pp. 332-342.
27. Sculley, D., Wachman, G. and Brodley, C. E. "Spam classifiers and inexact string matching features," *TREC 2006: The Fifteenth Text Retrieval Conference Proceedings*, November, 2006.
28. Jacob, N. and Brodley, C. E., "Offloading IDS computation to the GPU," *The Twentysecond Annual Computer Security Applications Conference*, December 2006, Washington D.C., pp. 371-380.
29. Rebbapragada, U. and Brodley, C. E., "Improving classifier accuracy by assigning confidence labels to training data," *The Eighteenth European Conference on Machine Learning*, September 2007, Warsaw, Poland, pp. 708-715.

30. Lomasky, R., Brodley, C. E., Aerneke, M., Walt, D., and Friedl, M. "Active Class Selection," *Eighteenth European Conference on Machine Learning*, September 2007, Warsaw, Poland, pp. 640-647.
31. Wei, X., Cowen, L., Brodley, C. E., Brady, A., Sculley, D. and Slonim, D. "A distance-based method for detecting horizontal gene transfer in whole genomes," *Proceedings of the 4th International Symposium on Bioinformatics Research and Applications (ISBRA 2008)*, Springer Lecture Notes in Computer Science, 2008.
32. Rebbapragada, U., Mandrake, L., Wagstaff, K., Gleeson, D., Castao, R., Chien, S., and Brodley, C. E., "Improving onboard analysis of Hyperion images by filtering mislabeled training data examples," *Proceedings of the 2009 IEEE Aerospace Conference*.
33. Preston, D., Protopapas, P. and Brodley, C. E. "Event Discovery in Time Series," *SIAM International Conference on Data Mining*, April 2009, Sparks, Nevada.
34. Wallace, B., Small, K., Brodley, C. E., and Trikalinos, T., "Active learning for biomedical citation screening," *The Sixteenth International Conference on Knowledge Discovery and Data Mining*, August 2010, Washington D.C., pp. 173-182.
35. Preston, P., Brodley, C. E., Khardon, R., Sulla-Menashe, D., and Friedl, M., "Redefining class definitions using constraint-based clustering: An application to remote sensing of the Earth's surface," *The Sixteenth International Conference on Knowledge Discovery and Data Mining*, August 2010, Washington D.C.
36. Wallace, B., Small, K., Brodley, C. E., and Trikalinos, T. "Modeling annotation time to reduce workload in comparative effectiveness reviews," *ACM International Health Informatics Symposium*, November 2010, Arlington, VA, pp. 28-35.
37. Small, K., Wallace, B., Brodley, C. E., and Trikalinos, T., "The constrained weight-space SVM: Learning with Labeled Features," *The Twenty-eighth International Conference on Machine Learning*, June 2011, Bellevue, WA, pp. 754-763.
38. Wallace, B., Small, K., Brodley, C. E. and Trikalinos, T., "Who should label what? Instance allocation in multiple expert active learning," *Siam International Conference on Data Mining (SDM)*, April 2011, Phoenix, AZ, pp. 176-187.
39. Wallace, B., Small, K., Brodley, C. E., and Trikalinos, T. "Class imbalance, redux," *IEEE 11th International Conference on Data Mining (ICDM)*, December 2011, Vancouver, B.C., pp. 754-763.
40. Wallace, B., Small, K., Brodley, C. E., Lau, J. and Trikalinos, T., "Deploying an interactive machine learning system in an evidence-based practice center," *In Proc. of the Second ACM International Health Informatics Symposium (IHI)*, January 2012, Miami, FL, pp. 819-826.
41. Brown, E., Liu, J., Brodley, C. E., and Chang, R., "Dis-Function: Learning distance functions interactively," *IEEE Conference on Visual Analytics Science and Technology (VAST)*, October 2012, Seattle, WA, pp. 83-92.
42. Rebbapragada, U., Brodley, C. E., Sulla-Menashe, D., and Friedl, M., "Active label correction," *IEEE 12th International Conference on Data Mining (ICDM)*, December 2012, Brussels, pp. 1080-1035.

43. Liu, J., Brodley, C. E., Healy, B., Chitnis, T. “Removing confounding factors via constraint-based clustering: An application to finding homogeneous groups of MS patients,” *Data Mining for Healthcare Workshop at ICHI* (included in ICHI conference proceedings), September 2013, Philadelphia, PA (invited for submission to the special issue of *AI and Medicine* on the best of ICHI).
44. Noto, K., Brodley, C. E., Majidi, S., Bianchi, D. and Slonim, D. “Characterizing systematic anomalies in eXpression data,” *RECOMB*, Pittsburgh, PA, April 2014.
45. Zhao, Y., Brodley, C. E., Chitnis, T., and Healy, B. “Addressing human subjectivity via transfer learning: An application to predicting disease course in multiple sclerosis patients,” *in the Proceedings of the 2014 SIAM International Conference on Data Mining (SDM)*, Philadelphia, PA, April 2014.
46. Moran, K., Wallace, B., and Brodley, C. E., “Discovering better AAAI keywords via clustering with crowd-sourced constraints,” accepted to appear at AAAI 2014.
47. Ahmed, B., Thesen, T., Blackmon, K., Devinsky, O., Kuzniaky, R., and Brodley, C. E., “Hierarchical conditional random fields for unsupervised outlier detection: An application to detecting epileptogenic cortical malformations,” accepted to appear at ICML 2014.

Conference Proceedings and Presentations: Abstract Reviewed:

1. Sandholm, M., Sandholm, T., Brodley, C. E., and Vidovic, A., “Linear and logistic regression, symbolic induction methods, and neural networks in morbidity diagnosis and morbidity prediction in equine gastrointestinal colic,” *The Second Annual SEPSIS/SIRS: Reducing Mortality to Patients & Suppliers*, February 1996, Washington D.C.
2. Brodley, C. E., Friedl, M. A., and Strahler, A., “New approaches to classification in remote sensing: Using hybrid decision tree and ensemble classifiers to map land cover,” *Proceedings of the International Geoscience and Remote Sensing Symposium*, vol. I, May 27-31, 1996, Lincoln, NE, pp. 532-534.
3. Brodley, C. E. and Friedl, M. A., “Improving automated land cover mapping by identifying and eliminating misclassified instances,” *Proceedings of the International Geoscience and Remote Sensing Symposium*, vol. II, May 27-31, 1996, Lincoln, NE, pp. 1382-1384.
4. Aisen, A. M. Broderick, L. S., Winer-Muram, H. Brodley, C. E., Kak, A. C., and Pavlopoulou C., “A Content-based Image Retrieval System for HRCT Images of the Lung: Implementation and Initial Validation,” *Radiology*, 2001.
5. Rebbapragada, U., Lomasky, R., Brodley, C. E., Friedl, M. A., “Generating high-quality training data for automated land-cover mapping,” *The International Geoscience and Remote Sensing Symposium*, IGARSS 2008.
6. Preston, D., Protopapas, P., and Brodley, C. E., “Event discovery in astronomical time series,” *Astronomical Data Analysis Software and Systems (ADASS)* conference, 2008, Quebec City.
7. Rebbapragada, U., Protopapas, P., Brodley, C. E., and Alcock, C. “Anomaly detection in catalogs of periodic variable stars,” *Astronomical Data Analysis Software and Systems (ADASS)* conference, 2008, Quebec City.

- Ahmed, B., Brodley, C. E., Carlson, C., Kuzniecky, R., Devinsky, O., French, J., and Thesen, T., "Detection of structural brain abnormalities in single patients: Focal Epilepsy," *International Conference on Basic and Clinical Multimodal Imaging*, Geneva, September, 2013.

Symposia and Workshop Papers:

- Brodley, C. E. and Utgoff, P. E., "Dynamic recursive model class selection for classifier construction," *The Fourth International AI and Statistics Workshop*, Ft. Lauderdale, FL, January 1993. (Full version appeared as a book chapter).
- Brodley, C. E. and Rissland, E. L., "Measuring concept changes," *AAAI Spring Symposium Series, Training Issues in Incremental Learning*, Stanford University, Palo Alto, CA, March 23-25, 1993.
- Brodley, C. E. and Smyth, P., "Applying machine learning in practice," *IMLC-95 Workshop: Applying Machine Learning in Practice*, Tahoe City, CA, July 10, 1995.
- Sandholm, T., Brodley, C. E., Vidovic, A., Nyholm, K., and Sandholm, M., "Comparison of regression methods, symbolic induction and methods and neural networks in morbidity diagnosis and mortality prediction in equine gastrointestinal colic," *AAAI 1996 Spring Symposium Series: Artificial Intelligence In Medicine: Applications of Current Technologies*, Stanford University, Palo Alto, CA, March 25-27, 1996.
- Brodley, C. E. and Lane, T., "Creating and exploiting diversity," *AAAI-96 Workshop: Integrating Multiple Learned Models for Improving and Scaling Machine Learning Algorithms*, Portland, OR, August 3, 1996.
- Lane T. and Brodley C. E., "Sequence matching and learning in anomaly detection for computer security," *AAAI-97 Workshop: AI Approaches to Fraud Detection and Risk Management*, Providence, RI, July 31, 1997.
- Shyu, C., Brodley, C. E., Kak, A., Kosaka, A., Aisen, A., and Broderick, L., "Local versus global features for content-based image retrieval", *in the Proceedings of the IEEE Workshop on Content-Based Access of Image/Video Library* held in conjunction with CVPR98, Santa Barbara, CA, June 21, 1998.
- Kapadia, N. H., Brodley, C. E., Fortes, J. A. B., and Lundstrom, M. S., "Resource-usage prediction for demand-based network computing," *Proceedings of the 1998 Workshop on Advances in Parallel and Distributed Systems (APADS)*, West Lafayette, IN, October 20, 1998, pp. 372-377.
- Shyu, C. R., Kak, A., Brodley, C. E., and Broderick, L., "Testing for human perceptual categories in a physician-in-the-loop CBIR system for medical imagery," *in the Proceedings of the IEEE Workshop of Content-Based Access of Image and Video Databases*, Fort Collins, CO, June 22, 1999.
- MacArthur, S. D., Brodley, C. E. and Shyu, C. R., "Relevance feedback decision trees in content-based image retrieval," *in the Proceedings of the IEEE Workshop of Content-Based Access of Image and Video Databases*, Hilton Head, SC, June 13, 2000.
- Marchiori, A., Brodley, C. Broderick, L., Dy, J., Pavlopoulou, C., Kak, A. and Aisen, A. "CBIR for Medical Images – An Evaluation Trial," *in the Proceedings of the IEEE Workshop of Content-Based Access of Image and Video Databases*, Hawaii, 2001.

12. Pavlopoulou, C., Kak, A. and Brodley, C. E., "An interactive framework for boundary delineation for medical CBIR," *in the Proceedings of the IEEE Workshop of Content-Based Access of Image and Video Databases*, Hawaii, 2001.
13. Pavlopoulou, C., Kak, A. and Brodley, C. E., "Applications of Semisupervised and Active Learning to Interactive Contour Delineation," *Proc. ICML-EDDS Workshop on the Continuum from Labeled to Unlabeled Data*, August 2003.
14. Pusara, M. and Brodley, C. E., "User Re-authentication via Mouse Movements," to appear at *the Workshop on Statistical and Machine Learning Techniques in Computer Intrusion Detection*, September 24-26th, 2003.
15. Lomasky, R., Brodley, C. E., Bencic, S., Aernecke, M., and Walt, D., "Guiding class selection for an artificial nose," *NIPS Workshop: Testing of Deployable Learning and Decision Systems* 2007.
16. Wallace, B., Dahabreh, I., Moran, K., Brodley, C. E., and Trikalinos, T. Active literature discovery for scoping evidence reviews: How many needles are there? *KDD 2013 - Workshop on Data Mining for Healthcare (DMH)*, 2013.
17. Ahmed, B., Thesen, T., Blackmon, K., Kuzniecky, R., Carlson, C., Quinn, B., Doyle, W., French, J., Devinsky, O., and Brodley, C. E., "Machine learning for detection of MRI-elusive epileptogenic lesions: A surface-based MRI morphometric approach," *SDM Workshop: Data Mining for Medicine and Healthcare*, Philadelphia, PA, April 2014.

Recent Invited Lectures:

1. "Covert Timing Channels over IP" given at:
 - Trust Seminar Series, University of CA, Berkeley, February 2007
 - Google, Mountain View, February 2007
 - ITA Software, Cambridge MA, February 2007
 - University of Massachusetts, Lowell, February, 2008
 - University of Virginia, Charlotte, April 2008
2. "Generating and Recognizing High-Quality Training Data for Supervised Learning" given at:
 - Stern School of Business, New York University, February 2008
 - CFA, Harvard University, March 2008
 - Northeastern University, March 2008
 - Tufts University Medical School, September 2009
3. "Challenges in the Practical Application of Supervised Learning" given at:
 - AI Seminar, Cornell University, September 2009
 - Intelligence Seminar, Carnegie Mellon University, November 2009
 - School of Informatics and Computing Colloquium, Indiana University, February 2010
 - Computer Science Colloquia, Purdue University, February 2010
 - Image and Video Computing Group, Boston University, March 2010

- CDSP Research Workshop, Communications and Digital Signal Processing Center, Northeastern University, March 2010
 - Computer Science Colloquia, University of Massachusetts at Boston, March 2010
 - Computer Science Colloquia, University of Alberta, Edmonton, Alberta, April 2010
 - Operations Research Center Seminar, MIT, February, 2011
 - Draper Laboratory, November 2011
4. “Removing Confounding Factors via Constraint-Based Clustering: An Application to Finding Homogeneous Groups of Multiple Sclerosis Patients” given at:
 - IMS/ASA Spring Research Conference, June 2012
 - Meaningful Use of Complex Medical Data, August 2012
 - Lincoln Lab, May 2013
 5. “Class Imbalance and Noisy Training Data: Detecting FCD in MRI-Negative Epileptic Patients via Machine Learning,” Meaningful Use of Complex Medical Data, August 2013
 6. “Redefining Class Definitions using Constraint-Based Clustering and its Application to Landcover Classification and the AAAI 2014 Keywords,” given at:
 - AI Colloquia University of Texas at Austin, April 2014
 - Data Science Day, Tufts University, May 2014
 - Google, Cambridge, MA, May 2014
 - New England Machine Learning Day, May 2014