

Jennifer Neville
Departments of Computer Science and Statistics
Purdue University
West Lafayette, IN 47907-2107
Phone: (765) 496-9387 · Email: neville@cs.purdue.edu

Research Interests

Artificial intelligence, machine learning, knowledge discovery and data mining, relational learning, social network analysis, link analysis, computational social science.

Education

Ph.D., University of Massachusetts Amherst, Computer Science, 2006

Dissertation: *Statistical Models and Analysis Techniques for Learning in Relational Data*

Advisor: David Jensen

M.S., University of Massachusetts Amherst, Computer Science, 2004

B.S. *summa cum laude*, University of Massachusetts Amherst, Computer Science, 2000

Professional Experience

2006–present: *Assistant Professor*, Purdue University
Joint appointment: Computer Science (75%) and Statistics (25%)

2000–2006: *Research Assistant*, University of Massachusetts Amherst

2000–2000: *Research Intern*, AT&T Shannon Laboratory

Awards and Honors

NSF Career Award, 2012

Purdue College of Science Team Award, 2011

Purdue Seed for Success Award, 2011

Purdue College of Science Interdisciplinary Award, 2009, 2010

ICDM Best Research Paper Award Runner-Up, 2009

IEEE Intelligent Systems Top Ten to Watch, 2008

Microsoft New Faculty Fellowship Finalist, 2007

DARPA Computer Science Study Panel Member, 2007

Nominated for ACM Doctoral Dissertation Award, University of Massachusetts, 2006

KDD Cup First Place Open Task, 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2003

Young Investigator Award, DARPA IPTO Cognitive Systems Conference, 2003

AT&T Labs Graduate Fellowship, 2000–2006

National Science Foundation Graduate Research Fellowship, 2000–2003

National Physical Science Consortium Fellowship, 2000 (*declined*)

Bell Labs Graduate Research Fellowship, 2000 (*declined*)

Publications

Journal Articles

1. Gender demographics trends and changes in U.S. CS departments.
D. Baumann, S. Hambrusch, and J. Neville.
Communications of the ACM, 54:11, 38-42, 2011.
2. Guided Data Repair
M. Yakout, A. Elmagarmid, J. Neville, M. Ouzzani, and I. Ilyas
Proceedings of the VLDB Endowment, 2011.
3. Correcting Evaluation Bias of Relational Classifiers with Network Cross Validation.
J. Neville, B. Gallagher, T. Eliassi-Rad, and T. Wang
Knowledge and Information Systems, 2011.
4. Prediction models for long-term Internet prefix availability
R. Khosla, S. Fahmy, Y. C. Hu, and J. Neville
Computer Networks, 2010.
5. A Bias-Variance Decomposition for Collective Inference Models.
J. Neville and D. Jensen.
Machine Learning Journal, 73:1, pages 87-106, 2008.
6. Dependency Networks for Relational Data.
J. Neville and D. Jensen.
Journal of Machine Learning Research, 8(Mar):653–692, 2007.
7. Exploiting Relational Structure to Understand Publication Patterns in High-Energy Physics.
A. McGovern, L. Friedland, M. Hay, B. Gallagher, A. Fast, J. Neville and D. Jensen.
SIGKDD Explorations, Volume 5, Issue 2, pages 165-172, 2003.

Conference Papers

8. Network Sampling Designs for Relational Classification
N. Ahmed, J. Neville, and R. Kompella
Proceedings of the 6th International AAAI Conference on Weblogs and Social Media (ICWSM), 4 pages, to appear.
9. Time-Evolving Relational Classification and Ensemble Methods
R. Rossi and J. Neville
Proceedings of the 16th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD), 12 pages, to appear.
10. Structured Comparative Analysis of Systems Logs to Diagnose Performance Problems
K. Nagaraj, C. Killian, and J. Neville.
Proceedings of the 9th USENIX Symposium on Networked Systems Design and Implementation (NSDI), to appear. (Acceptance rate: 18%)
11. Understanding Propagation Error and Its Effect on Collective Classification
R. Xiang and J. Neville
Proceedings of the 11th IEEE International Conference on Data Mining, to appear. (Acceptance rate (full paper): 12%)
12. Correcting Bias in Statistical Tests for Network Classifier Evaluation
T. Wang, J. Neville, B. Gallagher, and T. Eliassi-Rad
Proceedings of the 21st European Conference on Machine Learning, 16 pages, 2011, (Acceptance rate: 20%)

13. Relational Active Learning for Joint Collective Classification Models
A. Kuwadekar and J. Neville
Proceedings of the 28th International Conference on Machine Learning, 8 pages, 2011. (Acceptance rate: 25%)
14. Across-Model Collective Ensemble Classification
H. Eldardiry and J. Neville
Proceedings of the 25th Conference on Artificial Intelligence, 7 pages, 2011. (Acceptance rate: 25%)
15. Methods to Determine Node Centrality and Clustering in Graphs with Uncertain Structure
J. Pfeiffer III and J. Neville
Proceedings of the 5th International AAAI Conference on Weblogs and Social Media, 4 pages, 2011.
16. Relational Learning with One Network: An Asymptotic Analysis
R. Xiang, J. Neville
Proceedings of the 14th International Conference on Artificial Intelligence and Statistics (AISTAT), 11 pages, 2011. (Oral presentation, acceptance rate: 8%)
17. ERACER: A Database Approach for Statistical Inference and Data Cleaning
C. Mayfield, J. Neville, and S. Prabhakar
Proceedings of the 2010 ACM SIGMOD Conference, 12 pages, 2010. (Acceptance rate: 19%)
18. Predicting Prex Availability in the Internet
R. Khosla, S. Fahmy, C. Hu, and J. Neville
Proceedings of the 29th IEEE Conference on Computer Communications (INFOCOM) Mini-Conference, 5 pages, 2010. (Acceptance rate: 24%)
19. Randomization tests for distinguishing social influence and homophily effects
T. LaFond and J. Neville
Proceedings of the 19th International World Wide Web Conference (WWW), 10 pages, 2010. (Acceptance rate: 14%)
20. Modeling Relationship Strength in Online Social Networks
R. Xiang, J. Neville, and M. Rogati
Proceedings of the 19th International World Wide Web Conference (WWW), 10 pages, 2010. (Acceptance rate: 14%)
21. Using Transactional Information to Predict Link Strength in Online Social Networks
I. Kahanda and J. Neville
Proceedings of the the 3rd Int'l AAAI Conference on Weblogs and Social Media. 8 pages 2009.
22. Evaluating Statistical Tests for Within-Network Classifiers of Relational Data.
J. Neville, B. Gallagher, and T. Eliassi-Rad. **Best Paper Award Runner-Up**
Proceedings of the 9th IEEE International Conference on Data Mining, 10 pages, 2009. (Acceptance rate (full paper): 9%)
23. Temporal-Relational Classifiers for Prediction in Evolving Domains.
U. Sharan and J. Neville
Proceedings of the 8th IEEE International Conference on Data Mining, 10 pages, 2008. (Acceptance rate (full paper): 10%)
24. A Shrinkage Approach for Modeling Non-Stationary Relational Autocorrelation.
P. Angin and J. Neville
Proceedings of the 8th IEEE International Conference on Data Mining, 6 pages, 2008. (Acceptance rate (short paper): 20%)
25. Pseudolikelihood EM for Within-Network Relational Learning.
R. Xiang and J. Neville

- Proceedings of the 8th IEEE International Conference on Data Mining*, 6 pages, 2008. (Acceptance rate (short paper): 20%)
26. Database support for probabilistic attributes and tuples.
S. Singh, C. Mayfield, R. Shah, S. Prabhakar, S. Hambrusch, J. Neville, R. Cheng.
The 24th International Conference on Data Engineering, 9 pages, 2008. (Acceptance rate: 19%)
 27. Bias/Variance Analysis for Relational Domains.
J. Neville and D. Jensen.
The 17th International Conference on Inductive Logic Programming, *Lecture Notes in Artificial Intelligence 4894*, pages 27-28, 2007. (Acceptance rate: 34%)
 28. Leveraging Relational Autocorrelation with Latent Group Models.
J. Neville and D. Jensen.
Proceedings of the 5th IEEE International Conference on Data Mining, pages 322-329, 2005. (Acceptance rate: 11%)
 29. Using Relational Knowledge Discovery to Prevent Securities Fraud.
J. Neville, O. Simsek, D. Jensen, J. Komoroske, K. Palmer and H. Goldberg.
Proceedings of the 11th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pages 449-458, 2005. (Acceptance rate: 19%)
 30. Dependency Networks for Relational Data.
J. Neville and D. Jensen.
Proceedings of the 4th IEEE International Conference on Data Mining, pages 170-177, 2004. (Acceptance rate: 9%)
 31. Why Collective Inference Improves Relational Classification.
D. Jensen, J. Neville and B. Gallagher.
Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pages 593-598, 2004. (Acceptance rate: 25%)
 32. Simple Estimators for Relational Bayesian Classifiers.
J. Neville, D. Jensen and B. Gallagher.
Proceedings of the 3rd IEEE International Conference on Data Mining, pages 609-612, 2003. (Acceptance rate: 23%)
 33. Learning Relational Probability Trees.
J. Neville, D. Jensen, L. Friedland and M. Hay.
Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pages 625-630, 2003. (Acceptance rate: 27%)
 34. Avoiding Bias When Aggregating Relational Data with Degree Disparity.
D. Jensen, J. Neville and M. Hay.
Proceedings of the 20th International Conference on Machine Learning, pages 274-281, 2003. (Acceptance rate: 32%)
 35. Autocorrelation and Linkage Cause Bias in Evaluation of Relational Learners.
D. Jensen and J. Neville.
Proceedings of the 12th International Conference on Inductive Logic Programming, pages 101-116, 2002. (Acceptance rate: 54%)
 36. Linkage and Autocorrelation Cause Feature Selection Bias in Relational Learning.
D. Jensen and J. Neville.
Proceedings of the 19th International Conference on Machine Learning, pages 259-266, 2002. (Acceptance rate: 33%)

Referred Workshop Papers

37. Role-Dynamics: Fast Mining of Large Dynamic Networks
R. Rossi, Brian Gallagher, J. Neville, and Keith Henderson
Proceedings of the 1st International Workshop on Large Scale Network Analysis, WWW, 9 pages, to appear.
38. Understanding Propagation Error and Its Effect on Collective Classification
R. Xiang and J. Neville
Proceedings of the 9th Workshop on Mining and Learning with Graphs, KDD, 8 pages, 2011.
39. Modeling the Variance of Network Populations with Mixed Kronecker Product Graph Models
S. Moreno and J. Neville and S. Kirshner and S.V.N. Vishwanathan. **Most Promising Paper Award**
Proceedings of the Workshop on Analyzing Networks and Learning with Graphs, 24th Annual Conference on Neural Information Processing Systems, 8 pages, 2010.
40. Reconsidering the Foundations of Network Sampling
N. Ahmed, J. Neville, and R. Kompella
Proceedings of the 2nd Workshop on Information in Networks, 5 pages, 2010.
41. Time-Based Sampling of Social Network Activity Graphs
N. Ahmed, F. Berchmans, J. Neville, and R. Kompella
Proceedings of the 8th Workshop on Mining and Learning with Graphs, KDD, 8 pages, 2010.
42. Multi-Network Fusion for Collective Inference
H. Eldardiry and J. Neville
Proceedings of the 8th Workshop on Mining and Learning with Graphs, KDD, 8 pages, 2010.
43. Probabilistic Paths and Centrality in Time
J. Pfeiffer III and J. Neville
Proceedings of the 4th SNA-KDD Workshop, KDD, 8 pages, 2010.
44. Combining Semi-supervised Learning and Relational Resampling for Active Learning in Network Domains
A. Kuwadekar and J. Neville. **Best Paper Award**
Proceedings of the Budgeted Learning Workshop, ICML, 8 pages, 2010.
45. Modeling the Evolution of Discussion Topics and Communication to Improve Relational Classification
R. Rossi and J. Neville
Proceedings of the 1st Workshop on Social Media Analytics, KDD, 8 pages, 2010.
46. Ranking for Data Repairs
M. Yakout, A. Elmagarmid, and J. Neville
Proceedings of the 4th International Workshop on Ranking in Databases, ICDE 2010, 6 pages, 2010.
47. Modeling Relationship Strength in Online Social Networks.
R. Xiang and J. Neville
Proceedings of the Workshop on Analyzing Networks and Learning with Graphs, 23rd Annual Conference on Neural Information Processing Systems, 8 pages, 2009.
48. An Investigation of the Distributional Characteristics of Generative Graph Models.
S. Moreno and J. Neville.
Proceedings of the the 1st Workshop on Information in Networks, 5 pages, 2009.
49. A Shrinkage Approach for Modeling Non-Stationary Relational Autocorrelation.
P. Angin and J. Neville.
Proceedings of the 2nd Social Network Analysis Workshop, 14th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 6 pages, 2008.

50. A Resampling Technique for Relational Data Graphs.
H. Eldardiry and J. Neville.
Proceedings of the 2nd Social Network Analysis Workshop, 14th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 6 pages, 2008.
51. Pseudolikelihood EM for Within-Network Relational Learning.
R. Xiang and J. Neville
Proceedings of the 2nd Social Network Analysis Workshop, 14th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 8 pages, 2008.
52. Exploiting Time-Varying Relationships in Statistical Relational Models.
U. Sharan and J. Neville.
Proceedings of the 1st Social Network Analysis KDD Workshop, 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 7 pages, 2007.
53. Bias/Variance Analysis for Network Data.
J. Neville and D. Jensen.
Proceedings of the Workshop on Statistical Relational Learning, 23rd International Conference on Machine Learning, 8 pages, 2006.
54. Structure Learning for Statistical Relational Models.
J. Neville.
Proceedings of the 20th National Conference on Artificial Intelligence (Doctoral Consortium), pages 1656-1657, 2005.
55. Autocorrelation and Relational Learning: Challenges and Opportunities.
J. Neville, O. Simsek and D. Jensen.
Proceedings of the Workshop on Statistical Relational Learning, 21st International Conference on Machine Learning, 8 pages, 2004.
56. Collective Classification with Relational Dependency Networks. J. Neville and D. Jensen.
Proceedings of the 2nd Multi-Relational Data Mining Workshop, 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pages 77-91, 2003.
57. Statistical Relational Learning: Four Claims and a Survey.
J. Neville, M. Rattigan and D. Jensen.
Proceedings of the Workshop on Learning Statistical Models from Relational Data, 18th International Joint Conference on Artificial Intelligence, 5 pages, 2003.
58. Clustering Relational Data Using Attribute and Link Information.
J. Neville, M. Adler and D. Jensen.
Proceedings of the Text Mining and Link Analysis Workshop, 18th International Joint Conference on Artificial Intelligence, 6 pages, 2003.
59. Schemas and Models.
D. Jensen and J. Neville.
Proceedings of the Multi-Relational Data Mining Workshop, 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 15 pages, 2002.
60. Supporting Relational Knowledge Discovery: Lessons in Architecture and Algorithm Design.
J. Neville and D. Jensen.
Proceedings of the Data Mining Lessons Learned Workshop, 19th International Conference on Machine Learning, pages 57-64, 2002.
61. Correlation and Sampling in Relational Data Mining.
D. Jensen and J. Neville.
Proceedings of the 33rd Symposium on the Interface of Computing Science and Statistics, 14 pages, 2001.

62. Iterative Classification in Relational Data.
J. Neville and D. Jensen.
Proceedings of the Workshop on Learning Statistical Models from Relational Data, 17th National Conference on Artificial Intelligence, pages 42-49, 2000.

Book Chapters

63. Relational Dependency Networks.
J. Neville and D. Jensen.
Introduction to Statistical Relational Learning, L. Getoor and B. Taskar, editors, pages 239-268, 2007.

Invited Papers

64. Tied Kronecker Product Graph Models to Capture Variance in Network Populations
S. Moreno and S. Kirshner and J. Neville and S.V.N. Vishwanathan
Proceedings of the 48th Annual Allerton Conference on Communications, Control and Computing, 8 pages, 2010.
65. Data Mining in Social Networks.
D. Jensen and J. Neville.
National Academy of Sciences Symposium on Dynamic Social Network Analysis, 13 pages, 2002.

Technical Reports

66. Network Sampling via Edge-based Node Selection with Graph Induction.
N. Ahmed, J. Neville, R. Kompella.
Purdue University, CSD TR 11-0016, 2011.
67. Correcting Bias in Statistical Tests for Network Classifier Evaluation.
J. Neville, T. Wang, B. Gallagher, and T. Eliasi-Rad.
Purdue University, CSD TR 10-012, 2010.
68. Spectral Clustering with Links and Attributes.
J. Neville, M. Adler and D. Jensen.
University of Massachusetts Amherst, Technical Report 04-42, 2004.
69. Randomization Tests for Relational Learning.
D. Jensen, J. Neville and M. Rattigan.
University of Massachusetts Amherst, Technical Report 03-05, 2003.

Invited Presentations

- Workshop on Machine Learning: Theory and Computation (**invited speaker**), *Institute for Mathematics and Its Applications (IMA)*, Minnesota, MN, 2012 (upcoming).
- Workshop on Network Links: Connecting Social, Communication & Biological Network Analysis (**invited speaker**), *Institute for Mathematics and Its Applications (IMA)*, Minnesota, MN, 2012.
- How to learn from a single network: Statistical relational learning for social network domains, Computer Science Colloquia, *Duke University*, Durham, NC, 2011.
- How to learn from one sample? Statistical relational learning for single network domains, AI Seminar, *University of Texas Austin*, Austin, TX, 2011.
- Modeling online social networks to understand and predict user behavior, IROM Seminar, *McCombs School of Business*, Austin, TX, 2011.
- Understanding the Effects of Collective Classification on Learning and Inference (**invited keynote talk**), *Workshop on Collective Learning and Inference for Structured Data*, ECML, 2011.

Mining Social Network Activity to Understand and Predict User Behavior (**invited keynote talk**),
Workshop on Enriching Information Retrieval, SIGIR, 2011.

Modeling Complex Social Networks, *Center for Science of Information*, NSF Science & Technology Center
 Summer School, 2011.

Modeling Complex Social Networks: Challenges and Opportunities for Statistical Learning and Inference,
Machine Learning Summer School @ Purdue, 2011.

Statistical Relational Learning in Single Network Domains,
 PRiML Seminar, *University of Pennsylvania*, Philadelphia, PA, 2011.

Modeling and Mining Social Networks,
 Fantastic Lectures in Computer Science Series, *Bryn Mawr College*, Bryn Mawr, PA, 2011.

Hypothesis testing methods for social network mining,
 AI Seminar, *Information Sciences Institute*, Marina Del Ray, CA, 2010.

Hypothesis testing methods for social network mining,
 Neyman Seminar, *University of California Berkeley*, Berkeley, CA, 2010.

Hypothesis testing methods for social network mining,
 IS Research Seminar, *New York University Stern School of Business*, New York, NY, 2010.

Capturing the Natural Variability of Real Networks with Kronecker Product Graph Models, Sanida National
 Laboratory, Livermore, CA, 2010.

Evaluation Strategies for Network Classification Models (**invited keynote talk**), Workshop on Mining
 and Learning from Graphs, 2010.

Evaluation Strategies for Network Classification Models, University of Maryland College Park, College
 Park, MD, 2010.

Modeling Relationship Strength in Online Social Networks, IUPUI, Indianapolis, IN, 2010.

Modeling Relationship Strength in Online Social Networks, DAIS Seminar, University of Illinois Urbana-
 Champaign, Urbana, IL, 2010.

Social Network Mining (tutorial), with Foster Provost, International Conference on Knowledge Discovery
 and Data Mining (KDD-08), Henderson, NV, 2008.

Social Network Mining (invited tutorial), with Foster Provost, National Conference on Artificial Intelligence
 (AAAI-08), Chicago, IL, 2008.

Exploiting Temporal Variations in Relational Domains, Lawrence Livermore National Laboratory, Liver-
 more, CA, 2008.

Exploiting Temporal Variations in Relational Domains, University of Maryland College Park, College Park,
 MD, 2008.

Statistical Models for Learning and Inference in Complex Relational Domains.
 National Security Agency, Fort Meade, MD, 2007.

Data Mining in Networks.
 MA108: Mathematics as a Profession and a Discipline, Purdue University, West Lafayette, Indiana,
 2006.

Leveraging Autocorrelation with Latent Group Models.
 Auton Lab, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, 2005.

Leveraging Autocorrelation with Latent Group Models.

Dagstuhl Seminar on Probabilistic, Logical and Relational Learning: Towards a Synthesis. Schloss Dagstuhl, Wadern, Germany, 2005.

Knowledge Discovery with Relational Dependency Networks.

Weekly Computer Science Colloquium, Williams College, Williamstown, MA, 2004.

Dependency Networks for Relational Data.

The Boeing Company, Phantom Works, Mathematics & Computing Technology Unit, Seattle, WA, 2004.

Probability Estimation Trees for Relational Data.

Computer Science Department, Williams College, Williamstown, MA, 2004.

Collective Classification with Relational Dependency Networks (*poster*).

DARPA IPTO Cognitive Systems Conference, Washington, DC, 2003.

Knowledge Discovery in Networks.

Talent Advancement Program Seminar, Computer Science Department, University of Massachusetts, Amherst, MA, 2003.

Clustering Relational Data (*poster*).

Grace Hopper Celebration of Women in Computing, Vancouver, BC, 2002.

Data Mining in Networks.

International Sunbelt Social Network Conference XXII, New Orleans, LA, 2002.

Sponsored Research¹

Career: Machine Learning Methods and Statistical Analysis Tools for Single Network Domains

NSF/CISE/IIS, Primary Investigator

Amount: \$496,638 (100% of total), 01/01/12 - 12/31/16

Toward Intrusion Tolerant Clouds

DARPA/I2O, co-Primary Investigator (subcontract from Johns Hopkins)

\$367,028 (9% of total), 11/01/11 - 10/31/15

Modeling Tools to Support Advanced Analysis of Multi-Source Network Data

IARPA/KDD, co-Primary Investigator (subcontract from SAIC)

\$563,950 (5% of total), 11/01/10 - 01/31/15

Towards Better Modeling of Communication Activity Dynamics in Large-Scale Online Social Networks

NSF/CISE/NETSE, Primary Investigator

\$248,226 (50% of total), 09/01/10 - 08/31/13

Emerging Frontiers of Science of Information

NSF/Science & Technology Center, Senior Personnel

approx. \$200,000 (0.8% of total), 08/01/10 - 07/31/15

Algorithms for Sampling Similar Graphs Using Subgraph Signatures

NSF/CISE/IIS, co-Primary Investigator

\$164,846 (33% of total), 09/01/09 - 08/31/11

Machine Learning Techniques to Model the Impact of Relational Communication on Distributed Team Effectiveness

¹Reported amounts reflect funds 100% under my responsibility.

NSF/SES/IOS, Primary Investigator
\$205,311 (50% of total), 09/01/08 - 08/31/11

MAASCOM: Modeling, Analysis, and Algorithms for Stochastic Control of Multi-Scale Networks

ARO/MURI, co-Primary Investigator (subcontract from Ohio State)
\$250,000 (5% of total), 5/29/08–10/28/08

Fusion and Analysis of Multi-Source Relational Data

DARPA/ISO, Primary Investigator
\$499,877 (100% of total), 06/23/08–06/22/10

Learning Compositional Simulation Models

IARPA/Proactive Intelligence, co-Primary Investigator (subcontract from UMass)
\$122,217 (33% of total), 04/01/07–02/09/09

Mining Transaction Streams to Infer Semantic Relations

Microsoft, Primary Investigator
\$50,000 (100% of total), 06/01/07–06/01/99

Statistical Models and Algorithms to Improve Decision-Making in Relational Domains

DARPA/ISO, Primary Investigator
\$100,000 (100% of total), 04/01/07–12/31/07

Service

Area chair

International Conference on Machine Learning (ICML): 2011
ACM conference on Web Search and Data Mining (WSDM): 2012

Program committees

ACM International Conference on Knowledge Discovery and Data Mining (KDD): 2008-2012
European Conference on Machine Learning (ECML/PKDD): 2007, 2008, 2012
IEEE International Conference on Data Mining (ICDM): 2007, 2009-2012
International Conference on Artificial Intelligence and Statistics (AISTATS): 2009
International Conference on Inductive Logic Programming (ILP): 2007
International Conference on Machine Learning (ICML): 2006, 2008-2012
International Joint Conference on Artificial Intelligence (IJCAI): 2009
National Conference on Artificial Intelligence (AAAI): 2006-2008, 2012
SIAM Conference on Data Mining (SDM), 2006

Journal editorial boards

Journal of Artificial Intelligence Research, 2010-present
Machine Learning Journal, 2011-present

Journal reviewing

Journal of Machine Learning Research
Machine Learning Journal
Transactions on Knowledge Discovery
Data Mining and Knowledge Discovery Journal

Proposal reviewing

NSF Information & Intelligent Systems Panel, 2005, 2008, 2010, 2011

NASA Earth Science Technology Office Proposals, 2005

Conference/workshop organization

Tutorial Chair: ACM International Conference on Knowledge Discovery and Data Mining (KDD): 2012

Tutorial Chair: International Conference on Machine Learning (ICML): 2009

Treasurer: International Machine Learning Society (IMLS): 2009–present

Organizer: 9th Workshop on Mining and Learning with Graphs (MLG): 2011

Organizer: Machine Learning Summer School, Purdue: 2011

Outreach

CS197: *Topics in Computer Sciences (Honors)*, 2008, 2010

CS291: *Sophomore Development Seminar*, 2008

CS591C: *Research Seminar for Graduate Students*, 2006, 2007, 2008, 2010

MA108: *Mathematics as a Profession and a Discipline*, 2006

STAT VIGRE Seminar: *Exploring Statistical Sciences Research*, 2006, 2008, 2010

College of Science Woman for Purdue event, 2010

College of Science Alumni Board Meeting, 2010

CSWN events with women faculty, 2008, 2010

Purdue President and Provost, Tour of the College of Science, 2010

STAR lunches with incoming freshman, 2009

Undergrad Research Bonanza, 2009

Women in CS Career Day: *Purdue CS outreach event for high school girls*, 2008, 2009, 2010

Teaching/Students

Course Development

Development of graduate data mining course, with Chris Clifton (2007)

Development of undergraduate artificial intelligence course (2008)

Development of joint MS program for CS/Stat, with Sergey Kirshner (2009-10)

Development of undergrad machine learning course for Machine Intelligence track (2011)

Teaching

Data Mining and Machine Learning, Undergraduate Course, CS — Spring 2012

Data Structures and Algorithms, Undergraduate Course, CS — Spring 2011

Data Mining, Graduate Course, CS/STAT (cross-listed) — Fall 2010, Spring 2009, Fall 2007

Artificial Intelligence, Undergraduate Course, CS — Fall 2008

Statistical Network Analysis, Graduate Seminar, STAT/CS (cross-listed) — Spring 2008, Spring 2010

Introduction to Statistical Relational Learning, Graduate Seminar, CS — Spring 2007

Current PhD Students

Rongjing Xiang (CS), Passed Prelim (Summer 2011) *expected graduation January 2012*

Nesreen Ahmed (CS), Passed Qual 2 (Fall 2010)

Timothy La Fond (CS), Passed Qual 2 (Fall 2010)
Sebastian Moreno (CS), Passed Qual 2 (Summer 2011)
Joel Pfeiffer (CS), Passed Qual 2 (Fall 2011)
Ryan Rossi (CS), Passed Qual 1
Hyokun Yun (Stat), co-advised with SVN Vishwanathan, Post-Qual (Spring 2011)

Current Undergraduate Students

Jordan Bates
Christopher Cole
Daniel Roberts

Graduated Students

Hoda Eldardiry
Degree: PhD, February 2012
Thesis: Ensemble Classification Techniques for Relational Domains
Current location: Palo Alto Research Center

Chris Mayfield (co-advised with Sunil Prabhakar)
Degree: PhD, Aug 2011
Thesis: Statistical Inference and Data Cleaning in Relational Database Systems
Current location: James Madison University

Ankit Kuwadekar
Degree: BSc, May 2010
Research project: Active Learning for Relational Domains
Current location: Amazon

Umang Sharan
Degree: MSc, May 2008
Thesis: Temporal-Relational Classifiers for Prediction in Evolving Domains
Current location: YouTube (Google)