Jennifer Neville
Microsoft Research Redmond
Departments of Computer Science and Statistics, Purdue University
West Lafayette, IN 47907-2107
Phone: (765) 496-9387 · Email: prof.jen.neville@gmail.com

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

2024–present:	*Partner Research Manager*, Microsoft Research Redmond
2022–2024:	*Senior Principal Research Manager*, Microsoft Research Redmond
2021–2022:	*Senior Principal Researcher*, Microsoft Research Redmond
2020–present:	*Samuel Conte Professor of Computer Science*, Purdue University Joint appointment: Computer Science (75%) and Statistics (25%)
2020	*Visiting Researcher*, Microsoft Research/Microsoft MSAI
2019–2020:	*Professor*, Purdue University Joint appointment: Computer Science (75%) and Statistics (25%)
2014–2019:	*Miller Family Chair*, Purdue University
2013–2019:	*Associate Professor*, Purdue University Joint appointment: Computer Science (75%) and Statistics (25%)
2013	*Visiting Scientist*, Simons Institute for Theoretical Computing, UC Berkeley
2006–2013:	*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

Program Committee Chair, 37th AAAI Conference on Artificial Intelligence, 2023

Distinguished Lecture, University of California San Diego, Computer Science Department, 2020

Purdue University Faculty Scholar, 2020

Keynote Address, 23rd Pacific-Asia Conference on Knowledge Discovery and Data Mining, 2019

Keynote Address, 2nd IEEE Data Science Workshop, 2019.

Keynote Address, 2nd Machine Learning in Science and Engineering Forum, Georgia Tech 2019.

Keynote Address, IUPUI Data Science Summit, 2019.

Invited Plenary Panelist: *Future of Artificial Intelligence*,
33rd AAAI Conference on Artificial Intelligence, 2019

Program Committee Chair, 18th SIAM International Conference on Data Mining, 2019

Distinguished Lecture, Virginia Tech University, Computer Science Department, 2019

ACM Doctoral Dissertation Award Committee, 2019-2021

Keynote Address, 28th Keck Annual Research Conference, 2018

Invited Plenary Panelist: *Societal Impact of Data Science and Artificial Intelligence*,
ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2018

AAAI Executive Council, Elected Councilor, 2015-2018

Keynote Address, 27th International Conference on Inductive Logic Programming, 2017

Keynote Address, 3rd ACM International Conference on the Theory of Information Retrieval, 2017

Keynote Address, 6th International Conference on Complex Networks and Their Applications, 2017

Invited Plenary Panelist *Is Deep Learning the New 42?*,
ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2016

Distinguished Lecture, Max Planck Institute of Informatics, 2016

Program Committee Chair, 9th ACM International Conference on Web Search and Data Mining, 2016

Keynote Address, 2nd European Network Intelligence Conference, 2015

Purdue College of Science Research Award, 2014

Keynote Address, 27th International Conference of the Florida Artificial Intelligence Research Society, 2014

Purdue College of Science Graduate Mentoring Award, 2013

Outstanding Achievement by a Young Alum, School of CS, University of Massachusetts Amherst, 2013

NSF Career Award, 2012

Purdue College of Science Team Award, 2012

Purdue Seed for Success Award, 2011-2013

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[1]

Journal Articles

1. Generating Post-Hoc Explanations for Skip-gram-based Node Embeddings by Identifying Important Nodes with Bridgeness
   H. Park, and J. Neville.
   Neural Networks, 2023.

2. Ensemble learning for relational data.
   H. Eldardiry, J. Neville, and R. Rossi.
   Journal of Machine Learning Research, 21 (1), 1780-1816.

3. Scalable and exact sampling method for probabilistic generative graph models.
   S. Moreno[†], J. Pfeiffer III[†], and J. Neville.
   Data Mining and Knowledge Discovery, Volume 32, Issue 6, pp 1561-1596, 2018.

4. Designing Size Consistent Statistics for Accurate Anomaly Detection in Dynamic Networks.
   T. La Fond[†], J. Neville, and B. Gallagher.
   ACM Transactions on Knowledge Discovery from Data, 12:14, 2018.

5. Tied Kronecker Product Graph Models to Capture Variance in Network Populations.
   S. Moreno, J. Neville, and S. Kirshner
   ACM Transactions on Knowledge Discovery from Data, 12:3, 2018.

6. Graphlet Decomposition: Framework, Algorithms, and Applications.
   N. Ahmed[†], J. Neville, R. Rossi[†], N. Duffield, T. Willke
   Knowledge and Information Systems, Volume 50, Issue 3, pp 689-722, 2017.

7. Network Sampling: From Static to Streaming Graphs.
   N. Ahmed[†], J. Neville, and R. Kompella
   Transactions on Knowledge Discovery and Data Mining, Vol. 8, Issue 2, 2014.

8. Transforming Graph Data for Statistical Relational Learning.
   R. Rossi[†], L. McDowell, D. Aha and J. Neville
   Journal of Artificial Intelligence Research, Vol. 45, 363-441, 2012.

9. Correcting Evaluation Bias of Relational Classifiers with Network Cross Validation.
   J. Neville, B. Gallagher, T. Eliassi-Rad, and T. Wang
   Knowledge and Information Systems, 30-1, 31-55, 2012.

10. 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.

11. Guided Data Repair
    M. Yakout[†], A. Elmagarmid, J. Neville, M. Ouzzani, and I. Ilyas
    Proceedings of the VLDB Endowment, 2011.

12. Prediction models for long-term Internet prefix availability
    R. Khosla[†], S. Fahmy, Y. C. Hu, and J. Neville
    Computer Networks, 2010.

13. A Bias-Variance Decomposition for Collective Inference Models.
    J. Neville and D. Jensen.
    Machine Learning Journal, 73:1, pages 87-106, 2008.

---

[1]My authorship policy is to list students authors ahead of me, regardless of contribution. Graduate students are denoted with [†]; undergraduates are denoted with [‡].

14. Dependency Networks for Relational Data.
    J. Neville and D. Jensen.
    Journal of Machine Learning Research, 8(Mar):653–692, 2007.

15. 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

16. Node Similarities under Random Projections: Limits and Pathological Cases.
    Tvrtko Tadic, Cassiano Becker, J. Neville.
    Proceedings of the 13th International Conference on Learning Representations (ICLR), 2025.

17. Symbolic prompt program search: A structure-aware approach to efficient compile-time prompt optimization.
    T. Schnabel and J. Neville.
    Findings of the Association for Computational Linguistics (EMNLP), 2024.

18. TnT-LLM: Text mining at scale with large language models.
    M. Wan, T. Safavi, S. Jauhar, Y. Kim, S. Counts, J. Neville, S. Suri, C. Shah, R. White, L. Yang, R. Andersen, G. Buscher, D. Joshi, N. Rangan.
    Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 2024.

19. Automatic Pair Construction for Contrastive Post-training.
    C. Xu, C. Rosset, E. Chau, L. Corro, S. Mahajan, J. McAuley, J. Neville, A. Awadallah, N. Rao.
    Findings of the Association for Computational Linguistics (NAACL), 2024.

20. On Overcoming Miscalibrated Conversational Priors in LLM-based Chatbots.
    C. Herlihy, J. Neville, T. Schnabel, A. Swaminathan.
    Proceedings of the 40th Conference on Uncertainty in Artificial Intelligence (UAI), 2024.

21. MS MARCO Web Search: a Large-scale Information-rich Web Dataset with Millions of Real Click Labels
    Q. Chen, X. Geng, C. Rosset, C. Buractaon, J. Lu, T. Shen, K. Zhou, C. Xiong, Y. Gong, P. Bennett, N. Craswell, X. Xie, F. Yang, B. Tower, N. Rao, A. Dong, W. Jiang, Z. Liu, M. Li, C. Liu, Z. Li, R. Majumder, J. Neville, A. Oakley, K. Risvik, H. Simhadri, M. Varma, Y. Wang, L. Yang, M. Yang, C. Zhang.
    Companion Proceedings of the ACM on Web Conference, 2024.

22. Interpretable User Satisfaction Estimation for Conversational Systems with Large Language Models.
    Y. Lin, J. Neville, J. Stokes, L. Yang, T. Safavi, M. Wan, S. Counts, S. Suri, R. Andersen, X. Xu, D. Gupta, S. Jauhar, X. Song, G. Buscher, S. Tiwary, B. Hecht, and J. Teevan.
    Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (ACL), 2024.

23. DYANE: DYnamic Attributed Node rolEs Generative Model.
    G. Zeno and J. Neville.
    Proceedings of the 32nd ACM International Conference on Information and Knowledge Management (CIKM), 2023.

24. Hindsight Learning for MDPs with Exogenous Inputs.
    S. Sinclair, F. Frujeri, C. Cheng, L. Marshall, H. Barbalho, J. Li, J. Neville, I. Menache, A. Swaminathan.
    Proceedings of the 40th International Conference on Machine Learning (ICML), 2023.

25. Stationary Algorithmic Balancing Over Dynamic Email Re-Ranking Problem
J. Liu[†] and J. Neville.
Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 2023.

26. Workplace Recommendation with Temporal Network Objectives
K. Tomlinson, J. Neville, L. Yang, M. Wang, C. Lu
Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 2023.

27. Expressive and Efficient Representation Learning for Ranking Links in Temporal Graphs
S. Suresh[†] M. Shrivastava, A. Mukherjee, J. Neville, and P. Li.
Proceedings of the 32nd ACM Web Conference, 2023.

28. Adversarial Graph Augmentation to Improve Graph Contrastive Learning
S. Suresh[†] P. Li, C. Hao, and J. Neville.
Advances in Neural Information Processing Systems (NeurIPS), 2021.

29. A Collective Learning Framework to Boost GNN Expressiveness for Node Classification.
M. Hang[†] J. Neville, and B. Ribeiro.
Proceedings of the 38th International Conference on Machine Learning (ICML), 2021.

30. Breaking the Limit of Graph Neural Networks by Improving the Assortativity of Graphs with Local Mixing Patterns
S. Suresh[†] V. Budde, J. Neville, P. Li, J. Ma.
Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 2021.

31. Towards Decentralized Social Reinforcement Learning via Ego-Network Extrapolation
M. Goindani[†] and J. Neville.
Proceedings of the 20th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS), 2021.

32. Dymond: Dynamic motif-nodes network generative model
G. Zeno[†] T. La Fond, and J. Neville.
Proceedings of the 30th International Web Conference, 2021.

33. A Hybrid Model for Learning Embeddings and Logical Rules Simultaneously from Knowledge Graphs
S. Suresh[†] and J. Neville.
Proceedings of the 20th IEEE International Conference on Data Mining (ICDM), 6 pages, 2020. (Acceptance rate: 20%)

34. Online Bayesian Sparse Learning with Spike and Slab Priors
S. Fang, S. Zhe, K. Lee, K. Zhang, and J. Neville.
Proceedings of the 20th IEEE International Conference on Data Mining (ICDM), 8 pages, 2020. (Acceptance rate: 10%)

35. MERL: Multi-View Edge Representation Learning in Social Networks
Y. Lai[†], and J. Neville.
Proceedings of the 29th ACM International Conference on Information and Knowledge Management (CIKM), 10 pages, 2020. (Acceptance rate: 21%)

36. Role Equivalence Attention for Label Propagation in Graph Neural Networks
H. Park[†] and J. Neville
Proceedings of the 24th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD), 12 pages, 2020. (Acceptance rate: 21%)

37. Cluster-Based Social Reinforcement Learning
    M. Goindani[†] and J. Neville.
    Proceedings of the 19th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS), 3 pages, 2020. (Acceptance rate: 40%)

38. ReadNet: A Hierarchical Transformer Framework for Web Article Readability Analysis
    C. Meng[†], M. Chen, J. Mao, and J. Neville
    Proceedings of the 42nd European Conference on Information Retrieval (ECIR), 15 pages, 2020. (Acceptance rate: 23%)

39. Exploiting Interaction Links for Node Classification with Deep Graph Neural Networks
    H. Park[†] and J. Neville
    Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI), 8 pages, 2019. (Acceptance rate: 18%)

40. Learning How to Intervene in True News Diffusion to Combat Fake News Spread.
    M. Goindani[†] and J. Neville.
    Proceedings of the 35th Conference on Uncertainty in Artificial Intelligence (UAI), 10 pages, 2019. (Acceptance rate: 26%)

41. HATS: A Hierarchical Sequence-Attention Framework for Inductive Set-of-Sets Embeddings.
    C. Meng[†], J. Yang[†], B. Ribeiro, and J. Neville.
    Proceedings of the 25th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 10 pages, 2019. (Acceptance rate: 9%)

42. A Stein-Papangelou Goodness-of-Fit Test for Point Processes.
    J. Yang[†], V. Rao and J. Neville.
    Proceedings of the 23nd International Conference on Artificial Intelligence and Statistics (AISTAT), 10 pages, 2019. (Acceptance rate: 32%)

43. TransConv: Relationship Embedding in Social Networks.
    Y. Lai[†], J. Neville, and D. Goldwasser.
    Proceedings of the 33rd Conference on Artificial Intelligence (AAAI), 9 pages, 2019. (Acceptance rate: 16%)

44. Multi-level hypothesis testing for populations of heterogeneous networks
    G. Gomes[†], J. Neville, and V. Rao.
    Proceedings of the 18th IEEE International Conference on Data Mining (ICDM), 6 pages, 2018. (Acceptance rate: 20%)

45. The Indian Buffet Hawkes Process to Model Evolving Latent Influences.
    X. Tan[†], V. Rao and J. Neville.
    Proceedings of the 34th Conference on Uncertainty in Artificial Intelligence (UAI), 10 pages, 2018. (Acceptance rate: 31%)

46. Exploring Student Check-In Behavior for Improved Point-of-Interest Prediction.
    M. Hang[†], I. Pytlarz, and J. Neville.
    Proceedings of the 24th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 10 pages, 2018. (Acceptance rate: 22%)

47. Goodness-of-fit Testing for Discrete Distributions via Stein Discrepancy.
    J. Yang[†], Q. Liu, V. Rao and J. Neville.
    Proceedings of the 35th International Conference on Machine Learning (ICML), 9 pages, 2018. (Acceptance rate: 25%)

48. Nested CRP with Hawkes-Gaussian Processes.
    X. Tan[†], V. Rao and J. Neville.
    Proceedings of the 22nd International Conference on Artificial Intelligence and Statistics (AISTAT), 10 pages, 2018. (Acceptance rate: 33%)

49. Subgraph Pattern Neural Networks for High-Order Graph Evolution Prediction.
    C. Meng[†], C. Mouli[†], B. Ribeiro, and J. Neville.
    Proceedings of the 32nd Conference on Artificial Intelligence (AAAI), 9 pages, 2018.
    (Acceptance rate: 24%)

50. Decoupling Homophily and Reciprocity with Latent Space Network Models
    J. Yang[†], V. Rao and J. Neville
    Proceedings of the 33rd Conference on Uncertainty in Artificial Intelligence (UAI), 10 pages, 2017.
    (Acceptance rate: 31%)

51. Unified Representation and Lifted Sampling for Generative Models of Social Networks
    P. Robles[†], S. Moreno, and J. Neville
    Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI), 8 pages, 2017.
    (Acceptance rate: 25%)

52. Should We Be Confident in Peer Effects Estimated From Partial Crawls of Social Networks?
    J. Yang[†], B. Ribeiro and J. Neville
    Proceedings of the 11th International AAAI Conference on Weblogs and Social Media (ICWSM), 4 pages, 2017. Acceptance rate: 19%)

53. Deep Collective Inference
    J. Moore[†] and J. Neville
    Proceedings of the 31st Conference on Artificial Intelligence (AAAI), 9 pages, 2017.
    (Acceptance rate: 21%)

54. Sampling of Attributed Networks from Hierarchical Generative Models
    P. Robles[†], S. Moreno, and J. Neville
    Proceedings of the 22nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 10 pages, 2016. (Acceptance rate: 9%)

55. Efficient Graphlet Counting for Large Networks
    N. Ahmed[†], J. Neville, R. Rossi[†], and N. Duffield.
    Proceedings of the 15th IEEE International Conference on Data Mining (ICDM), 10 pages, 2015.
    (Acceptance rate: 8%)

56. Overcoming Relational Learning Biases to Accurately Predict Preferences in Large Scale Networks
    J. Pfeiffer III[†], J. Neville, and P. Bennett
    Proceedings of the 24th International World Wide Web Conference (WWW), 11 pages, 2015.
    (Acceptance rate: 14%)

57. Incorporating Assortativity and Degree Dependence into Scalable Network Models
    S. Mussmann[‡], J. Moore[‡], J. Pfeiffer III[†], and J. Neville
    Proceedings of the 29th Conference on Artificial Intelligence (AAAI), 9 pages, 2015.
    (Acceptance rate: 12%)

58. Composite Likelihood Data Augmentation for Within-Network Statistical Relational Learning
    J. Pfeiffer III[†], J. Neville, and P. Bennett
    Proceedings of the 14th IEEE International Conference on Data Mining (ICDM), 10 pages, 2014.
    (Acceptance rate: 10%)

59. A Scalable Method for Accurate Sampling from Kronecker Models
    S. Moreno, J. Pfeiffer III[†], and J. Neville
    Proceedings of the 14th IEEE International Conference on Data Mining (ICDM), 10 pages, 2014.
    (Acceptance rate: 10%)

60. Active Exploration in Networks: Using Probabilistic Relationships for Learning and Inference
    J. Pfeiffer III[†], J. Neville, and P. Bennett
    Proceedings of the 23rd ACM International Conference on Information and Knowledge Management (CIKM), 10 pages, 2014. (Acceptance rate: 21%)

61. Graph Sample and Hold: A Framework for Big-Graph Analytics
N. Ahmed[†], N. Duffield, J. Neville, and R. Kompella
Proceedings of the 20th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 10 pages, 2014. (Acceptance rate: 15%)

62. Attributed Graph Models: Modeling network structure with correlated attributes
J. Pfeiffer III[†], S. Moreno[†], T. La Fond[†], J. Neville, and B. Gallagher
Proceedings of the 23rd International World Wide Web Conference (WWW), 11 pages, 2014. (Acceptance rate: 13%)

63. Network Hypothesis Testing Using Mixed Kronecker Product Graph Models
S. Moreno[†] and J. Neville
Proceedings of the 13th IEEE International Conference on Data Mining (ICDM), 6 pages, 2013. (Acceptance rate: 19%)

64. Learning Mixed Kronecker Product Graph Models with Simulated Method of Moments
S. Moreno[†], J. Neville and S. Kirshner
Proceedings of the 19th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 9 pages, 2013. (Acceptance rate: 17%)

65. Collective Inference for Network Data with Copula Latent Markov Networks
R. Xiang[†] and J. Neville
Proceedings of the 6th ACM International Conference on Web Search and Data Mining (WSDM), 10 pages, 2013. (Acceptance rate: 19%)

66. Modeling Dynamic Behavior in Large Evolving Graphs
R. Rossi[†], B. Gallagher, J. Neville and K. Henderson
Proceedings of the 6th ACM International Conference on Web Search and Data Mining (WSDM), 10 pages, 2013. (Acceptance rate: 19%)

67. An Analysis of How Ensembles of Collective Classifiers Improve Predictions in Graphs
H. Eldardiry[†] and J. Neville
Proceedings of the 21st ACM International Conference on Information and Knowledge Management (CIKM), 10 pages, 2012. (Acceptance rate: 13%)

68. Fast Generation of Large Scale Social Networks While Incorporating Transitive Closures
J. Pfeiffer III[†], T. La Fond[†], S. Moreno[†], and J. Neville
Proceedings of the 4th ASE/IEEE International Conference on Social Computing (SocialCom), 12 pages, 2012. (Acceptance rate: 10%)

69. The Impact of Communication Structure and Interpersonal Dependencies on Distributed Teams
T. La Fond[†], D. Roberts[‡], J. Neville, J. Tyler, and S. Connaughton
Proceedings of the 4th ASE/IEEE International Conference on Social Computing (SocialCom), 8 pages, 2012. (Acceptance rate: 17%)

70. 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, 2012. (Acceptance rate: 26%)

71. 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, 2012. (Acceptance rate: 36%)

72. 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), 2012. (Acceptance rate: 18%)

73. Understanding Propagation Error and Its Effect on Collective Classification
R. Xiang[†] and J. Neville
Proceedings of the 11th IEEE International Conference on Data Mining (ICDM), 10 pages, 2011.
(Acceptance rate, full paper: 12%)

74. 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 (ECML), 16 pages, 2011.
(Acceptance rate: 20%)

75. Relational Active Learning for Joint Collective Classification Models
A. Kuwadekar[‡] and J. Neville
Proceedings of the 28th International Conference on Machine Learning (ICML), 8 pages, 2011.
(Acceptance rate: 25%)

76. Across-Model Collective Ensemble Classification
H. Eldardiry[†] and J. Neville
Proceedings of the 25th Conference on Artificial Intelligence (AAAI), 7 pages, 2011.
(Acceptance rate: 25%)

77. 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 (ICWSM), 4
pages, 2011.

78. Relational Learning with One Network: An Asymptotic Analysis
R. Xiang[†] and J. Neville
Proceedings of the 14th International Conference on Artificial Intelligence and Statistics (AISTAT), 11
pages, 2011. (Acceptance rate, oral presentation: 8%)

79. ERACER: A Database Approach for Statistical Inference and Data Cleaning
C. Mayfield[†], J. Neville, and S. Prabhakar
Proceedings of the 2010 ACM SIGMOD Conference (SIGMOD), 12 pages, 2010.
(Acceptance rate: 19%)

80. Predicting Prefix 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%)

81. 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%)

82. 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%)

83. Tied Kronecker Product Graph Models to Capture Variance in Network Populations
S. Moreno[†], S. Kirshner, J. Neville and S.V.N. Vishwanathan
Proceedings of the 48th Annual Allerton Conference on Communications, Control and Computing, 8
pages, 2010.

84. Using Transactional Information to Predict Link Strength in Online Social Networks
I. Kahanda[†] and J. Neville
Proceedings of the 3rd International AAAI Conference on Weblogs and Social Media (ICWSM), 8
pages, 2009. (Acceptance rate: 17%)

85. 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 (ICDM), 10 pages, 2009.
    (Acceptance rate, full paper: 9%)

86. Temporal-Relational Classifiers for Prediction in Evolving Domains.
    U. Sharan[†] and J. Neville
    Proceedings of the 8th IEEE International Conference on Data Mining (ICDM), 10 pages, 2008.
    (Acceptance rate, full paper: 10%)

87. A Shrinkage Approach for Modeling Non-Stationary Relational Autocorrelation.
    P. Angin[†] and J. Neville
    Proceedings of the 8th IEEE International Conference on Data Mining (ICDM), 6 pages, 2008.
    (Acceptance rate, short paper: 20%)

88. Pseudolikelihood EM for Within-Network Relational Learning.
    R. Xiang[†] and J. Neville
    Proceedings of the 8th IEEE International Conference on Data Mining (ICDM), 6 pages, 2008.
    (Acceptance rate, short paper: 20%)

89. 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 (ICDE), 9 pages, 2008.
    (Acceptance rate: 19%)

90. 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* (ILP), pages 27-28, 2007. (Acceptance rate: 34%)

91. Leveraging Relational Autocorrelation with Latent Group Models.
    J. Neville and D. Jensen.
    Proceedings of the 5th IEEE International Conference on Data Mining (ICDM), pages 322-329, 2005.
    (Acceptance rate: 11%)

92. 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 (KDD), pages 449-458, 2005. (Acceptance rate: 19%)

93. Dependency Networks for Relational Data.
    J. Neville and D. Jensen.
    Proceedings of the 4th IEEE International Conference on Data Mining (ICDM), pages 170-177, 2004.
    (Acceptance rate: 9%)

94. 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 (KDD), pages 593-598, 2004. (Acceptance rate: 25%)

95. Simple Estimators for Relational Bayesian Classifiers.
    J. Neville, D. Jensen and B. Gallagher.
    Proceedings of the 3rd IEEE International Conference on Data Mining (ICDM), pages 609-612, 2003.
    (Acceptance rate: 23%)

96. 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 (KDD), pages 625-630, 2003. (Acceptance rate: 27%)

97. 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 (ICML), pages 274-281, 2003. (Acceptance rate: 32%)

98. 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 (ILP), pages 101-116, 2002. (Acceptance rate: 54%)

99. Linkage and Autocorrelation Cause Feature Selection Bias in Relational Learning.
D. Jensen and J. Neville.
Proceedings of the 19th International Conference on Machine Learning (ICML), pages 259-266, 2002. (Acceptance rate: 33%)

Referred Workshop Papers[2]

100. Creating Generalizable Downstream Graph Models with Random Projections.
A. Amirov, C. Quirk, J. Neville.
Proceedings of the Workshop on Machine Learning on Graphs, (WSDM), 2023.

101. Lightweight Compositional Embeddings for Incremental Streaming Recommendation
M. Hang[†] T. Schnabel, L. Yang, and J. Neville.
Proceedings of the Workshop on Machine Learning on Graphs, (WSDM), 2022.

102. Social Reinforcement Learning
M. Goindani[†] and J. Neville
Proceedings of the AAAI Symposium on Challenges and Opportunities for Multi-Agent Reinforcement Learning, (COMARL), 4 pages, 2020.

103. Dynamic Network Modeling from Motif-Activity
G. Zeno[†], T. La Fond, and J. Neville
Proceedings of the 4th International Workshop on Mining Actionable Insights from Social Networks, WWW, 8 pages, 2020.

104. Joint Embedding Models for Textual and Social Analysis
C. Li[†], Y. Lai[†], D. Goldwasser, and J. Neville
Proceedings of the 1st Workshop on Deep Structured Prediction, (ICML), 5 pages, 2017.

105. Stochastic Gradient Descent for Relational Logistic Regression via Partial Network Crawls
J. Yang[†], B. Ribeiro and J. Neville
Proceedings of the 7th International Workshop on Statistical Relational AI, UAI, 7 pages, 2017.

106. Deep Dynamic Relational Classifiers: Exploiting Dynamic Neighborhoods in Complex Networks
H. Park[†], J. Moore[†], and J. Neville
Proceedings of the Mining Actionable Insights from Social Networks Workshop, WSDM, 7 pages, 2017.

107. Online Spike-and-slab Inference with Stochastic Expectation Propagation
S. Zhe[†], K. Lee, K. Zhang, and J. Neville
Proceedings of the 2016 Workshop on Advances in Approximate Bayesian Inference, NIPS, 2016.

108. Generating Local Explanations of Network Anomalies via Score Decomposition
T. La Fond[†], J. Neville, and B. Gallagher
Proceedings of the ODD 4.0: Outlier Definition, Detection, and Description on Demand, KDD, 6 pages, 2016.

---

[2]Note: In data mining and machine learning, workshops are a key venue to present and discuss preliminary work at the top conferences. Workshop papers are peer-reviewed and acceptance rates are typically in the 25-75% range.

109. Investigating the Impact of Graph Structure and Attribute Correlation on Collective Classification Performance
G. Zeno[†] and J. Neville
Proceedings of the 13th Workshop on Mining and Learning with Graphs, KDD, 8 pages, 2016.

110. Combining Gradient Boosting Machines with Collective Inference to Predict Continuous Values
I. Alodah[†] and J. Neville
Proceedings of the 6th International Workshop on Statistical Relational AI, IJCAI, 7 pages, 2016.

111. Better Together: Combining Language and Social Interactions into a Shared Representation
Y. Lai[†], C. Li[†], D. Goldwasser, and J. Neville
Proceedings of the TextGraphs 2016, (NAACL), 5 pages, 2016.

112. Analyzing the Transferability of Collective Inference Models Across Networks
R. Niu[‡], S. Moreno and J. Neville
Proceedings of the International Workshop on Information Analysis and Data Mining Over Social Network, (ICDM), 9 pages, 2015.

113. Using Bayesian Network Representations for Effective Sampling from Generative Network Models
P. Robles[†], S. Moreno, and J. Neville
Proceedings of the 5th International Workshop on Statistical Relational AI, UAI, 6 pages, 2015.

114. Assortativity in Chung Lu Random Graph Models
S. Mussmann[‡], J. Moore[‡], J. Pfeiffer III[†], and J. Neville
Proceedings of the 8th SNA-KDD Workshop, KDD, KDD, 8 pages, 2014.

115. Anomaly Detection in Networks with Changing Trends
T. La Fond[†], J. Neville, and B. Gallagher
Proceedings of the Outlier Detection & Description under Data Diversity Workshop, KDD, 10 pages, 2014.

116. Block Kronecker Product Graph Models
S. Moreno[†], P. Robles[†], and J. Neville
Proceedings of the 11th Workshop on Mining and Learning with Graphs, KDD, 6 pages, 2013.

117. Combining Active Sampling with Parameter Estimation and Prediction in Single Networks
J. Pfeiffer III[†], J. Neville, and P. Bennett
Proceedings of the Structured Learning: Inferring Graphs from Structured and Unstructured Inputs Workshop, ICML, 6 pages, 2013.

118. Space-Efficient Sampling from Social Activity Streams
N. Ahmed[†], J. Neville, and R. Kompella
Space-Efficient Sampling from Social Activity Streams. *Proceedings of the 1st International Workshop on Big Data, Streams and Heterogeneous Source Mining, KDD*, 8 pages, 2012.

119. Using Latent Communication Styles to Predict Individual Characteristics
J. Bates[‡], J. Neville, and J. Tyler
Proceedings of the 3rd Workshop on Social Media Analytics, KDD, 8 pages, 2012.

120. Active Sampling of Networks
J. Pfeiffer III[†], J. Neville, and P. Bennett
Proceedings of the 10th Workshop on Mining and Learning with Graphs, ICML, 8 pages, 2012.

121. On the Mismatch Between Learning and Inference for Single Network Domains
R. Xiang[†] and J. Neville
Proceedings of the *Workshop on Inferning: Interactions between Inference and Learning, ICML*, 6 pages, 2012.

122. 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, 2012.

123. 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.

124. Modeling the Variance of Network Populations with Mixed Kronecker Product Graph Models
S. Moreno[†], J. Neville, 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.

125. 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.

126. 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.

127. 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.

128. Probabilistic Paths and Centrality in Time
J. Pfeiffer III[†] and J. Neville
Proceedings of the 4th SNA-KDD Workshop, KDD, 8 pages, 2010.

129. 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.

130. 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.

131. 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.

132. 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.

133. An Investigation of the Distributional Characteristics of Generative Graph Models.
S. Moreno[†] and J. Neville.
Proceedings of the 1st Workshop on Information in Networks, 5 pages, 2009.

134. 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.

135. 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.

136. 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.

137. 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.

138. 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.

139. Structure Learning for Statistical Relational Models.
J. Neville.
Proceedings of the 20th National Conference on Artificial Intelligence (Doctoral Consortium), pages 1656-1657, 2005.

140. 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.

141. 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.

142. 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.

143. 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.

144. 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.

145. 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.

146. 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.

147. 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

148. 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

149. 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

150. Identifying User Survival Types via Clustering of Censored Social Network Data.
C. Mouli†, A. Naik‡, B. Ribeiro, and J. Neville.
ArXiv e-prints, 1703.03401, 2017.

151. Learning the Latent State Space of Time-Varying Graphs.
N. Ahmed†, C. Cole‡, and J. Neville.
ArXiv e-prints, 1403.3707, 2014.

152. Network Sampling via Edge-based Node Selection with Graph Induction.
N. Ahmed†, J. Neville, R. Kompella.
Purdue University, CSD TR 11-0016, 2011.

153. 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.

154. Spectral Clustering with Links and Attributes.
J. Neville, M. Adler and D. Jensen.
University of Massachusetts Amherst, Technical Report 04-42, 2004.

155. Randomization Tests for Relational Learning.
D. Jensen, J. Neville and M. Rattigan.
University of Massachusetts Amherst, Technical Report 03-05, 2003.

Invited Presentations

Towards Relational AI: The good, the bad, and the ugly of learning over networks (**distinguished lecture**), University of California San Diego, Computer Science Department, 2020.

Learning Over Sets, Subgraphs, and Streams (**invited speaker**), Microsoft Research, 2020.

Learning Over Sets, Subgraphs, and Streams (**invited speaker**), From Local to Global Information Workshop, NSF Center for Science of Information, 2020.

Natural Intelligence vs. Artificial Intelligence (**invited panelist**), Purdue Engineering Distinguished Lecture Panel Discussion w/James DiCarlo, 2019.

Analyzing Behavioral Traces to Improve Student Outcomes (**invited speaker**), Microsoft Faculty Summit, Microsoft Research, Bellevue, Washington, 2019.

Towards Relational AI: The good, the bad, and the ugly of learning over networks (**invited keynote**), IEEE Data Science Workshop, Minneapolis, Minnesota, 2019.

Towards Relational AI: The good, the bad, and the ugly of learning over networks (**invited keynote**), MLSE Data Science Forum, Georgia Tech Institute for Data Engineering and Science, 2019.

Deep Learning for Relational Networks (**invited keynote**), IUPUI Data Science Summit, IUPUI, Indianapolis, Indiana, 2019.

Towards Relational AI: The good, the bad, and the ugly of learning over networks (**invited keynote**), Pacific-Asia Conference on Knowledge Discovery and Data Mining, Macau, China, 2019.

Statistical methods for prediction and anomaly detection in dynamic networks (**invited speaker**), Foundations of Data Science Workshop, Purdue University, 2019.

Can we trust AI? (**invited panelist**), Purdue 150th Celebration Ideas Festival Panel Discussion w/Vint Cerf, 2019.

Towards Relational AI: The good, the bad, and the ugly of learning over networks (**distinguished lecture**), Virginia Tech University, Computer Science Department, 2019.

Essential machine learning hacks that everyone should know about (**invited keynote**), *Women in Data Science*, Purdue, West Lafayette, Indiana, 2019.

The role of machine learning in AI (**invited panelist**), Plenary Panel, 33rd AAAI Conference on Artificial Intelligence, 2019.

Statistical methods for prediction and anomaly detection in dynamic networks,
Microsoft Research, Redmond, WA, 2018.

Deep Learning for Relational Networks (**invited keynote**), 5th *International Workshop on High Performance Big Graph Data Management, Analysis, and Mining*, IEEE BigData 2018.

Deep Learning for Relational Networks (**invited speaker**), Purdue Workshop on Deep Learning, West Lafayette, Indiana, 2018.

Machine learning for network data (**invited panelist**), 4th *Annual Financial Stability Conference*, Department of the Treasury, Washington DC, 2018.

Using machine learning to exploit complex data in biomedical domains (**invited keynote**), Keck Annual Conference, Houston, TX, 2018.

Societal Impact of Data Science and Artificial Intelligence (**invited panelist**), Plenary Panel, ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2018.

Networks, Crowds & Confusion (**invited panelist**), Purdue Engineering Distinguished Lecture Panel Discussion w/Eva Tardos, 2018.

Mining social network interactions to understand and predict user behavior (**invited speaker**),
6th *Purdue Symposium on Psychological Sciences*–Big Data for Psychological Sciences, West Lafayette, Indiana, 2018.

Network Machine Learning (**invited speaker**),
Women in Data Science, Purdue, West Lafayette, Indiana, 2018.

Learning from single networks—the impact of network structure on relational learning (**invited keynote**),
27th *International Conference on Inductive Logic Programming*, France, 2017.

Learning in networks: How to exploit relationships to improve predictions (**invited keynote**),
3rd *ACM International Conference on the Theory of Information Retrieval*, Netherlands, 2017.

The impact of network structure on relational machine learning (**invited keynote**),
6th *International Conference on Complex Networks and Their Applications*, France, 2017.

Learning in networks: How to exploit relationships to improve predictions,
Ecole Polytechnique Federale de Lausanne (EPFL), Computer Science Department, 2017.

Exploiting User Relationships to Accurately Predict Preferences in Large Scale Networks (**invited speaker**), *First Data Institute Conference*, San Francisco, CA, 2017.

Semi-supervised learning for node classification in networks (**invited speaker**), *Workshop on Machine Learning in Network Science*, NetSci, Indianapolis, IN, 2017.

Lifted and Constrained Sampling of Attributed Graphs with Generative Network Models (**invited speaker**), *Workshop on Statistical Inference of Network Models*, NetSci, Indianapolis, IN, 2017.

Learning in networks: How to exploit relationships to improve predictions (**invited speaker**), *Workshop on Diversity in AI*, AAAI, San Francisco, CA, 2017.

AI-Easy vs AI-Hard (**invited speaker**), Dawn or Doom Syposium, Purdue University, West Lafayette, IN, 2016.

Is Deep Learning the New 42? (**invited panelist**), Plenary Panel, ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016.

Statistical methods for modeling network distributions (**invited keynote talk**), *Workshop on Mining and Learning from Graphs*, KDD, San Francisco, CA, 2016.

Statistical methods for modeling network distributions, *MIT Lincoln Laboratory*, Boston, MA, 2016.

AI-Easy vs AI-Hard: Machine learning and its impact on the development of AI systems (**invited speaker**), Purdue Student Pugwash Midwest Regional Conference, West Lafayette, IN, 2016.

Network machine learning: How to exploit relationships to improve node-level predictions (**invited speaker**), Institute of Science and Technology Austria, Young Scientist Symposium, Klosterneuburg, Austria, 2016.

Statistical methods for modeling network distributions (**invited speaker**), Mathematical Biosciences Institute Workshop on Generalized Network Structures and Dynamics, Columbus, Ohio, 2016.

Learning in networks: How to exploit relationships to improve predictions (**distinguished lecture**), Max Planck Institute for Informatics, Saarbrucken, Germany, 2016.

Exploiting User Relationships to Accurately Predict Preferences in Large Scale Networks (**invited speaker**), Netflix Workshop on Personalization, Recommendations, and Search, Los Gatos, CA, 2016.

Learning How to Transfer Collective Classification Models Across Networks (**invited speaker**), Santa Fe Institute Workshop: Inference on Networks, Santa Fe, New Mexico, 2015.

Collective Classification in Large-Scale Networks (**invited tutorial**), *Data Science and Advanced Analytics* (DSAA-15), Paris, France, 2015.

Sampling Attributed Networks From Generative Graph Models (**invited speaker**) Workshop on Information in Networks (WIN), New York, NY, 2015.

Network machine learning: How to use friends to your advantage algorithmically, 2nd European Network Intelligence Conference (**invited keynote**), Karlskrona, Sweden, 2016.

Machine learning methods for accurate estimation and prediction in partially-labeled complex networks, Skytree, San Jose, CA, 2015.

Network Sampling: Methods and Applications (**tutorial**), with Mohammad Al Hasan and Nesreen Ahmed, *SIAM International Conference on Data Mining* (SDM-15), Vancouver, Canada, 2015.

An introduction to big data: Opportunities for CS/AG collaborations, Purdue Big Data in Agriculture Seminar Series, West Lafayette, IN, 2015.

Predicting user behavior in networks: The impact of structure on machine learning methods, Indiana University, Bloomington, IN, 2014.

Predicting user behavior in networks: The impact of structure on machine learning methods. Data, Society, and Inference Seminar, *Stanford University*, Palo Alto, CA, 2014.

Predictive modeling for online social networks: Machine learning methods for networks (**invited speaker**), Westwood Colloquium, Purdue University, , West Lafayette, IN, 2014.

Are we too smart for our own good? How large-scale machine learning systems can vastly exceed human level decision-making abilities (**invited speaker**), Dawn or Doom: The New Technology Explosion, Purdue University, West Lafayette, IN, 2014.

How to Exploit Network Properties to Improve Learning in Relational Domains, IBM Thomas J Watson Research Center, Yorktown Heights, NY, 2014.

Prediction in complex networks: The impact of structure on learning and inference (**invited keynote**), 27th International Conference of the Florida Artificial Intelligence Research Society, Pensacola, FL, 2014.

How to learn from a single network to support classification and hypothesis testing in graphs, iCeNSA Seminar, *Notre Dame*, Notre Dame, IN, 2014.

Strengthening Computer Science, with Sunil Prabhakar, *Purdue President's Council*, Mollenkopf/Keyes Weekend, Naples, FL, 2014.

Network Sampling: Methods and Applications (**tutorial**), with Mohammad Al Hasan and Nesreen Ahmed, *International Conference on Data Mining* (ICDM-13), Dallas, TX, 2013.

Prediction in complex networks: The impact of structure on learning and inference (**invited keynote**), Women in Machine Learning Workshop, NIPS 2013, Lake Tahoe, CA, 2013.

Supporting hypothesis testing over graphs (**invited speaker**), Workshop on the Frontiers of network analysis: Methods, models, and applications, NIPS 2013, Lake Tahoe, CA, 2013.

Prediction in complex networks: The impact of structure on learning and inference, RAIN Seminar, *Stanford University*, Palo Alto, CA, 2013.

The impact of network structure on relational learning and inference, Simons Institute, *University California Berkeley*, Berkeley, CA, 2013.

Purdue Moves: Growth in Computer Science, with Sunil Prabhakar, *President's Forum*, Purdue University, November 2013.

How to learn from a single network to support classification and hypothesis testing in graphs, *SRI International*, Menlo Park, CA, 2013.

How to learn from a single network to support classification and hypothesis testing in graphs, *Facebook*, Menlo Park, CA, 2013.

Prediction in complex networks: The impact of structure on learning and inference, *PARC*, Palo Alto, CA, 2013.

Prediction in complex networks: The impact of structure on learning and inference, Neyman Seminar, *University of California Berkeley*, Berkeley, CA, 2013.

Machine learning methods for diagnosis, maintenance, and repair of data, AMP Lab, *University California Berkeley*, Berkeley, CA, 2013.

Prediction in complex networks: The impact of structure on learning and inference, *Google*, Mountain View, CA, 2013.

How to learn from a single network to support classification and hypothesis testing in graphs,
  Sandia National Laboratory, Livermore, CA, 2013.

Network Sampling: Methods and Applications (**tutorial**), with Mohammad Al Hasan and Nesreen Ahmed[†],
  International Conference on Knowledge Discovery and Data Mining (KDD-13), Chicago, IL, 2013.

Supporting Statistical Hypothesis Testing over Graphs, (**invited session**)
  Statistical Inference for Networks, *2013 Joint Statistical Meetings* (JSM), Montreal, QC, Canada, 2013.

Supporting Statistical Hypothesis Testing over Graphs, (**invited speaker**)
  International Conference on Network Science (NetSci), Copenhagen, Denmark, 2013.

Mining Social Network Activity to Understand and Predict User Behavior,
  Smith College, Northampton, MA, 2013.

Prediction in complex networks: The impact of structure on learning and inference, (**invited speaker**)
  Graph Exploitation Symposium, MIT Lincoln Laboratory, MA, 2013.

Prediction in complex networks: The impact of structure on learning and inference,
  Statistics Department Colloquia, *Carnegie Melon University*, Pittsburgh, PA, 2013.

Active Exploration in Networks,
  CSoI Big Data Workshop, Honolulu, HI, 2013.

Prediction in complex networks: The impact of structure on learning and inference,
  Computer Science Colloquia, *University of Alberta*, Edmonton, AB, Canada, 2013.

Prediction in complex networks: The impact of structure on learning and inference,
  Microsoft Research, Redmond, WA, 2012.

How to learn from a single network: Statistical relational learning for social network domains,
  Computer Science Colloquia, *University of Maryland College Park*, College Park, MD, 2012.

Measurement and methodology for mining mobile, cloud-based, social systems, (**invited speaker**)
  NSF Workshop on Social Networks and Mobility in the Cloud, Arlington, VA, 2012.

Prediction in complex networks: The impact of structure on learning and inference,
  Computer Science Colloquia, *Ohio State University*, Columbus, OH, 2012.

How to learn from a single network: Statistical relational learning for social network domains,
  Statistics Colloquia, *Columbia University*, New York, NY, 2012.

Prediction in complex networks: The impact of structure on learning and inference, (**invited speaker**)
  Workshop on Information in Networks (WIN), New York, NY, 2012.

Mining Social Network Activity to Understand and Predict User Behavior (**invited speaker**),
  Data Sciences Summer Institute, University of Illinois Urbana-Champaign, Urbana, IL, 2012.

How to learn from a single network: Relational learning for social networks, (**invited speaker**)
  Workshop on Machine Learning: Theory and Computation, *Institute for Mathematics and Its Applications* (IMA), Minnesota, MN, 2012.

Supporting Statistical Hypothesis Testing Over Graphs, (**invited speaker**)
  Workshop on Network Links: Connecting Social, Communication & Biological Network Analysis, *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, Athens, Greece, 2011.

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

Modeling Complex Social Networks: Challenges and Opportunities for Statistical Learning and Inference,
(**invited speaker**), *Machine Learning Summer School*, Purdue, West Lafayette, IN, 2011.

Modeling Complex Social Networks: Challenges and Opportunities for Statistical Learning and Inference,
(**invited speaker**), *Science of Information Summer School*, Purdue, West Lafayette, IN, 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,
Sandia National Laboratory, Livermore, CA, 2010.

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

Predictive Modeling with Social Networks (**invited tutorial**), with Foster Provost,
International Conference on Weblogs and Social Media (ICWSM-09), San Jose, CA, 2009.

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, Livermore, 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.

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.

Service

Outreach

Machine learning and its impact on the development of AI systems, Cyber Technology Conference & Policy Summit, Indianapolis, IN, 2019.

Session organizer: Fairness and Bias in Machine Learning, SciFoo'18, Google, 2018.

A few things about machine learning, Anvil Software Stir lecture, West Lafayette, Indiana, 2017.

Machine Learning Over Social Networks, Iridescent Learning, Ask the Experts Lecture and Interview, 2017.
`https://www.youtube.com/watch?v=UOvtr6LOrzU&t=78s`
`https://www.youtube.com/watch?v=iKUXaaguO-I&t=45s`

AI-easy vs. AI-hard?, Purdue Dawn or Doom Symposium, 2016.
`https://www.c-span.org/video/?415869-2/jennifer-neville-discusses-artificial-intelligence`

Machine learning and its impact on the development of AI systems, Purdue Student Pugwash Midwest Regional Conference, 2016.

Panelist: Techpoint Indianapolis, Dawn or Doom Panel, 2015.

An introduction to big data: Opportunities for CS/AG collaborations, Purdue Big Data in Agriculture Seminar, 2015.

Are we too smart for our own good?, Purdue Dawn or Doom Symposium, 2014.
`https://www.youtube.com/watch?v=GpltbAljmbg`

Strengthening Computer Science, with Sunil Prabhakar, *Purdue President's Council*, Naples, FL, 2014.

University

150th Ideas Festival Working Group, 2018-2019

Faculty Search Committee OIE Facilitator, 2014-2018

Diversity and Inclusion Symposium, Panel Moderator, 2014

Conference for Pre-Tenure Women, Mentor, 2014

Presentations:
  CoE EFC on Security and Privacy, 2019
  CoE Engineering Advisory Panel, 2017
  Center for The Environment Big Data Panel, 2017
  Westwood Colloquium, 2014
  President's Council Weekend, 2014
  President's Forum, 2013

College

Integrative Data Science Initiative Steering Committee, 2018-2019

Integrated Imaging Cluster Hiring Committee, 2012-2014

CS Realignment Task Force, 2009-2010

Presentations:
  College of Science Woman for Purdue event, 2010, 2014
  College of Science Alumni Board Meeting, 2010, 2104
  Purdue President and Provost, Tour of the College of Science, 2010

Other:
  Purdue Day of Giving Interview, 2014
  Purdue Insights Article, 2011, 2014

Department

Computer Science

 AI Search Committee Chair, 2019-2020

 Head Search Committee, 2018-2019

 Hiring Committee, 2007-2008, 2012-2016, 2017-2019

 New Data Science Degree Committee, 2016-present

 Hiring Committee, 2007-2008, 2012-2016

 Big Data Cluster Hiring Committee, 2012-2016

 Advisory Committee, 2014-present

 Colloquium Chair, 2008-2010

Course Development
  Development of introductory data science course for Data Science major (2017-2018)
  Development of undergrad machine learning course for Machine Intelligence track (2011)
  Development of joint MS program for CS/Stat (2009-10)

Graduate Curriculum Committee, 2011-2013

Strategic Planning Committee, 2009-2010

Undergraduate Curriculum Committee, 2006-2007, 2010-2011

Qual 2 Committees[3] (31 students):
    2006: Rimma Nehme
    2007: Jing Li, Yinian Qi
    2008: Suleyman Cetintas
    2009: Mohamed Yakout
    2010: Nan Ding, Yi Fang, Yao Zhu, Chandima Hewanadungodage, Youhan Fang
    2011: Pei He, Bin Shen, Fang Yu Rao, Feng Yan, Ahmet Bugdayci, Jiaqi Ge, Xi Tan, Syed Abaas
    2012: Ruby Tahboub, Romilla Pradhan, Jihwan Lee
    2014: Ning Zhang, Xuejiao Kang, Hao Peng
    2015: Shandian Zhe, Shawn Merrill, Zhiwei Zhang
    2016: Cosmo Xiang, Ashish Ghoshal, Kristen Johnson, Jin Di

Prelim Committees (that do not include PhD committee):
    Feng Yan
    Jiaqi Qi
    Qifan Wang
    Lei Cen
    Bahar Saberidokht
    Amr Ebiad
    Tanay Saha

PhD Committees (44 students):
    Computer Science:
    Ferit Akova, Vasil Denchev, Nicole Eikmeier, Juan Esquivel, Jiaqi Ge, Ashish Ghoshal, Chandima Hewanadungodage, Andrew Hoblitzell, Kristen Johnson, Ishita Khan, Chang Li, I-Ta Lee, Jihwan Lee, Koray Manchuan, Mohamed Mohamed, Abbas Naqvi, Rimma Nehme, Mehmet Nergiz, Gen Nishida, Keewan Park Hao Peng, Rahul Potharaju, Yinian Qi, Mahmudur Rahman, Sarvjeet Singh, Christine Task, Leonardo Teixeira, Mohamed Yakout, Pinar Yardanag, Dan Zhang, Xiao Zhang

    Aerospace Engineering: Alden Black

    Biology: Biaobin Jiang, Muyi Liu

    Electrical and Computer Engineering:
    Suhas Javagal, Subrata Mitra, Nawanol Theera-Ampornpunt, Tsung-Tai Yeh

    Industrial Engineering: Arora Viplove

    Sociology: Min Yim

    Statistics: Emery Goosens, Ryan Murphy, Hyokun Yun

    External: Shuangfei Fan, Virginia Tech

Presentations:
    CS197: *Topics in Computer Sciences (Honors)*, 2008, 2010, 2012
    CS291: *Sophomore Development Seminar*, 2008
    CS591C: *Research Seminar for Graduate Students*, 2006, 2007, 2008, 2010
    Women in CS Career Day: *Purdue CS outreach event for high school girls*, 2008, 2009, 2010
    Undergrad Research Bonanza, 2009
    Undergrad Track Fair, 2014
    Corporate Partners Program, 2009
    Grad Orientation, 2009, 2012
    Corporate presentations 2014-2016: John Deere, General Electric, Northrup Grumman, Lockheed Martin, Infosys, AFRL, Tsukuba

---

[3]Note that Purdue CS eliminated Qual 2 exams in 2016.

Outreach:
 STAR lunches with incoming freshman, 2009
 CSWN events with women faculty, 2008, 2010

<u>Statistics</u>

Hiring Committee, 2007-2008, 2017-2018

Presentations:
 MA108: *Mathematics as a Profession and a Discipline*, 2006
 STAT VIGRE Seminar: *Exploring Statistical Sciences Research*, 2006, 2008, 2010

Professional

Program committee chair

 National Conference on Artificial Intelligence (AAAI): 2023

 SIAM International Conference on Data Mining (SDM): 2019

 ACM International Conference on Web Search and Data Mining (WSDM): 2016

Associate chair

 Neural Information Processing Systems (NIPS), 2017-2019

 International Joint Conference on Artificial Intelligence (IJCAI), Machine Learning Track: 2015, 2018

 National Conference on Artificial Intelligence (AAAI), Machine Learning Track: 2013, 2019

Area chair

 ACM International Conference on Knowledge Discovery and Data Mining (KDD): 2016-2019

 ACM International Conference on on Web Search and Data Mining (WSDM): 2013, 2014, 2017

 International Conference on Machine Learning (ICML): 2011, 2015, 2018-2019

 International Conference on Uncertainty in Artificial Intelligence (UAI): 2019

 International Joint Conference on Artificial Intelligence (IJCAI), 2016, 2017

 National Conference on Artificial Intelligence (AAAI): 2014

 Neural Information Processing Systems (NIPS): 2014

 IEEE International Conference on Data Mining (ICDM): 2014, 2015

 SIAM conference on Data Mining (SDM): 2013, 2014

Program committees

 ACM International Conference on Knowledge Discovery and Data Mining (KDD): 2008-2012, 2014

 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, 2011

 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

 International World Wide Web Conference (WWW): 2011

 National Conference on Artificial Intelligence (AAAI): 2006-2008, 2012

 Neural Information Processing Systems (NIPS): 2012

 SIAM Conference on Data Mining (SDM), 2006

Journal editorial boards

Journal of Artificial Intelligence Research, 2010-2013

Machine Learning Journal, 2011-present

Data Mining and Knowledge Discovery, 2014-present

Journal reviewing

Journal of Machine Learning Research

Machine Learning Journal

Transactions on Knowledge Discovery

Data Mining and Knowledge Discovery Journal

Proposal reviewing

NSF Expeditions Blue Ribbon Panel, 2015, 2018

NSF Information & Intelligent Systems Panel, 2005, 2008, 2010-2015, 2017

NSF Information & Intelligent Systems CAREER Panel, 2012, 2016

NASA Earth Science Technology Office Proposals, 2005

Conference/workshop organization

Workshop Co-chair:
Relational Representation Learning Workshop, NeurIPS: 2018
NII Shonan Meeting Seminar 113, Meta-Programming for Statistical Machine Learning: 2018
13th Workshop on Mining and Learning with Graphs (MLG): 2017
Unifying Theory and Experiment for Large-Scale Networks, Simons Institute, UC Berkeley: 2014
4th ACM SIGSPATIAL International Workshop on Location-Based Social Networks (LBSN): 2012
Statistical Issues with Modeling of Networks, 8th Int'l Purdue Symposium on Statistics: 2012
9th Workshop on Mining and Learning with Graphs (MLG): 2011
Machine Learning Summer School, Purdue: 2011

Award Committee:
Doctoral Dissertation Award: Association for Computing Machinery (ACM): 2019-2021
Best Papers, ACM Int'l Conf on Knowledge Discovery and Data Mining (KDD): 2015, 2019
Test-of-Time Papers, ACM Int'l Conf on Knowledge Discovery and Data Mining (KDD): 2018
Best Papers (Chair), ACM Int'l Conf on Knowledge Discovery and Data Mining (KDD): 2017
Dissertation Award: ACM Int'l Conf on Knowledge Discovery and Data Mining (KDD): 2011-2013

Diversity events:
Women in KDD Panel, ACM International Conference on Knowledge Discovery and Data Mining (KDD): 2017-2018
Working Group on Gender Diversity, ACM International Conference on Web Search and Data Mining (WSDM): 2013
Women in AI, International Conference on Artificial Intelligence (AAAI): 2016-2017, 2019

Tutorial Chair:
ACM International Conference on Knowledge Discovery and Data Mining (KDD): 2012, 2015
International Conference on Machine Learning (ICML): 2009, 2015

Panelist:
White House Office of Science and Technology Policy, Workshop on AI and Social Good, 2016

Treasurer:
International Machine Learning Society (IMLS): 2009-2013

Professional societies

Association for the Advancement of Artificial Intelligence (AAAI)

Association for Computing Machinery (ACM)

ACM Special Interest Group on Knowledge Discovery and Data Mining (SIGKDD)

Institute of Electrical and Electronics Engineers (IEEE)

International Machine Learning Society (IMLS)

Sigma Xi

Sponsored Research[4]

Learning from Theorem Proving Search via Graph Representations
    NSF/CISE/FMitF, co-Primary Investigator
    $250,000 (33% of total), 08/01/19 - 07/31/22

I^2DS: *Intelligent Interaction Defense System*
    DARPA/I2O, Primary Investigator
    $744,302 (33% of total), 09/01/18 - 08/31/22

Robust Machine Learning
    Integrative Data Science Institute Purdue, Primary Investigator
    $264,989 (100% of total), 06/01/18 - 05/31/20

Transfer Learning Within & Across Networks for Collective Classification
    NSF/CISE/IIS, Primary Investigator
    $495,308 (100% of total), 07/01/16 - 06/30/19

Models, Algorithms, and Software for Spatial-Relational Networks
    NSF/CISE/IIS, co-Primary Investigator
    $300,000 (33% of total), 09/01/15 - 08/31/19

Frontiers of Science of Information
    NSF/Science & Technology Center, Senior Personnel
    approx. $200,000 (0.8% of total), 08/01/15 - 07/31/20

A Heterogeneous Inference Framework for 3D Modeling and Rendering of Sites
    NSF/CISE/CVG, co-Primary Investigator
    $297,782 (50% of total), 07/01/13 - 06/30/16

Parametric Statistical Models to Support Statistical Hypothesis Testing over Graphs
    NSF/CISE/IIS, co-Primary Investigator
    $245,920 (50% of total), 09/01/12 - 08/31/15

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

Sampling and Modeling Dynamic Streaming Networks

---

[4]Reported amounts reflect funds 100% under my responsibility.

CISCO, co-Primary Investigator
$102,076$ (50% of total), 02/07/11 - 02/07/99

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

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

Teaching

Spring 2020-Fall 2023: Research leave

CS471: Introduction to Artificial Intelligence, Fall 2019
Requirement for Machine Intelligence Track
115 students

CS 59000DEI: Data Engineering I and CS 59000FDM: Foundations of Decision Making, Spring 2019
Developed material and delivered lectures for online modules

CS242: Introduction to Data Science, Fall 2018
    Requirement for new Data Science Major
    25 students

CS242: Introduction to Data Science, Spring 2018
    Requirement for new Data Science Major
    10 students

CS573: Data Mining, Spring 2017
    Graduate table course, EPE distance education course
    107 students

CS497: Honors Research, Spring 2017
    Undergraduate honors project
    Student: Benjamin Staiger
    Project: Classifying Relationships Based on Time Series Data

CS473: Web Information Search and Management, Fall 2016
    Requirement for Machine Intelligence Track
    75 students

ST598T: Statistical Network Analysis (cross-listed as CS590), Spring 2016
    Graduate seminar
    12 students

CS39000-DM0: Data Mining & Machine Learning, Fall 2015
    Requirement for Machine Intelligence Track
    55 students

CS573: Data Mining, Spring 2015
    Graduate table course, EPE distance education course
    101 students

CS39000-DM0: Data Mining & Machine Learning, Fall 2014
    Requirement for Machine Intelligence Track
    30 students

CS39000-DM0: Data Mining & Machine Learning, Spring 2014
    Requirement for Machine Intelligence Track
    40 students

CS497: Honors Research, Spring 2014
    Undergraduate honors project
    Student: Seong Gyeong Lee
    Project: Sentiment Analysis of Facebook Posts

Fall 2013: Sabbatical

CS39000-DM0: Data Mining & Machine Learning, Spring 2013
    Requirement for Machine Intelligence Track
    40 students

CS497: Honors Research, Spring 2013
    Undergraduate honors project
    Student: Dan Coroian
    Project: Playing Solitaire with Reinforcement Learning

CS573: Data Mining, Fall 2012
    Graduate table course
    66 students

CS497: Honors Research, Fall 2012
    Undergraduate honors project
    Student: Brian Donovan
    Project: Textual Analysis of Facebook Messages to Predict Social Characteristics

CS39000-DM0: Data Mining & Machine Learning, Spring 2012
    Requirement for Machine Intelligence Track
    32 students: 27 CS, 5 Other

CS251: Data Structures and Algorithms, Spring 2011
    Undergraduate requirement
    61 CS students

CS573: Data Mining, Fall 2010
    Graduate table course
    32 students, 32 CS

ST598T: Statistical Network Analysis (cross-listed as CS590), Spring 2010
    Graduate seminar
    19 students: 9 CS, 4 STAT, 5 Eng, 1 IPSC

Fall 2009: Buyout

CS497: Honors Research, Spring 2009
    Undergraduate honors project
    Student: Ankit Kuwadeker
    Project: Active Learning for Relational Classification

CS573: Data Mining (cross-listed as ST598M), Spring 2009
    Graduate table course
    34 students, 21 CS, 3 STAT, 9 Eng, 1 EAS

CS471: Artificial Intelligence, Fall 2008
    Undergraduate elective
    30 CS students

ST598T: Statistical Network Analysis, Spring 2008
    Graduate seminar
    6 students: 4 CS, 2 STAT

CS590D: Data Mining (now CS573, cross-listed as ST598M), Fall 2007
    Graduate table course
    23 students, 16 CS, 2 STAT, 5 Eng

CS590N: Statistical Relational Learning, Spring 2007
    Graduate seminar
    10 students: 6 CS, 1 STAT, 3 Eng

Students

Current PhD Students

1. Jiayi Liu
   Advising since: Fall 2020
   Current status: Post-prelim
   Funding: Teaching assistant

Graduated Students

1. Ying-Chun (Jean) Lin
   Degree: PhD, Dec 2024
   Thesis: Interpretable Frameworks for User-Centered Structured AI Models
   Current location: Amazon

2. Giselle Zeno
   Degree: PhD, July 2023
   Thesis: Dynamic Network Modeling from Temporal Motifs and Attributed Node Activity
   Current location: MIT Lincoln Labs

3. Susheel Suresh
   Degree: PhD, Dec 2022
   Thesis: Models and Representation Learning Mechanisms for Graph Data
   Current location: Microsoft

4. Mengyue Hang
   Degree: PhD, Dec 2021
   Thesis: Graph Representation learning for unsupervised and semi-supervised learning tasks
   Current location: Facebook

5. Yi-yu Lai
   Degree: PhD, May 2021
   Thesis: Relational Representation Learning incorporating Textual Communication for Social Networks
   Current location: Intel

6. Mahak Goindani
   Degree: PhD, Dec 2020
   Thesis: Social Reinforcement Learning
   Current location: Apple

7. Hogun Park
   Degree: PhD, July 2020
   Thesis: Neural Representation Learning for Semi-Supervised Node Classification and Explainability
   Current location: Assistant Professor, College of Computing, Sungkyunkwan University, South Korea

8. Changping Meng
   Degree: PhD, April 2020
   Thesis: Scalable Representation Learning with Invariances
   Current location: ML Software Engineer, Google

9. Guilherme Gomes
   Degree: PhD, Dec 2019
   Thesis: Hypothesis Testing and Community Detection on Networks with Missingness & Block Structure
   Current location: Research Scientist, Bayer Pharma

10. Jiasen Yang
    Degree: PhD, June 2019
    Thesis: Statistical Learning and Model Criticism for Networks and Point Processes
    Current location: Quantitative Associate, Two Sigma

11. Xi Tan
    Degree: PhD, May 2018
    Thesis: Bayesian Nonparamterics to Model Content, User, and Latent Structure in Hawkes Processes
    Current location: Quantitative Associate, Goldman Sachs

12. Elizabeth Tigner
    Degree: BSc, May 2018
    Research project: Analysis of network discussion trends in Twitter using hashtag clusters
    Current location: Engineer, SEAKR Engineering

13. Caleb Belth
    Degree: BSc, May 2018
    Research project: Embedding influence graphs
    Current location: PhD Program, CS, University of Michigan

14. Shandian Zhe
    Degree: PhD, Nov 2017
    Thesis: Scalable Bayesian Nonparametrics and Sparse Learning for Hidden Relationships Discovery
    Current location: Assistant Professor, University of Utah, Computer Science Dept

15. Pablo Robles Granda
    Degree: PhD, Oct 2017
    Thesis: Generating Attribute Networks: Modeling, Learning, and Sampling
    Current location: Research Assistant Professor, Notre Dame University, Computer Science Dept

16. Gouthami Kamalnath
    Degree: BSc, May 2017
    Current location: Software Engineer, Microsoft

17. Shubhika Barjatya
    Degree: BSc (Statistics), May 2017
    Current location: Data Scientist, Wayfair

18. Benjamin Staiger
    Degree: BSc, May 2017
    Research project: Classifying Relationships Based on Time Series Data
    Current location: Software Engineer, Citadel

19. Sait Celebi
    Degree: MSc, May 2017
    Current location: Software Engineer, Google

20. Iman Alodah
    Degree: MSc, Dec 2016
    Current location: TBD

21. John Moore
    Degree: MSc, Dec 2016
    Thesis: Deep Collective Inference
    Current location: Software Engineer, Microsoft

22. Timothy La Fond
    Degree: PhD, July 2016

Thesis: Controlling for Confounding Network Properties in Hypothesis Testing and Anomaly Detection
Current location: Postdoc, Lawrence Livermore National Labs

23. Ransen Niu
Degree: BSc, May 2016
Research project: Transfer learning in networks
Current location: PhD Program, CS, Cornell University

24. Nesreen Ahmed
Degree: PhD, July 2015
Thesis: Scaling Up Network Analysis and Mining: Statistical Sampling, Estimation, and Pattern Discovery
Current location: Senior Staff Research Scientist, Intel Research

25. Stephen Mussmann
Degree: BSc, May 2015
Research project: Assortativity in statistical models of graphs
Current location: PhD Program, CS, Stanford University

26. Joseph Pfeiffer III
Degree: PhD, May 2015
Thesis: Overcoming Uncertainty for Within-Network Relational Machine Learning
Current location: Applied Researcher, Microsoft

27. Sebastian Moreno
Degree: PhD, August 2014
Thesis: Network Hypothesis Testing for Relational Data
Current location: Assistant Professor, Universidad Adolfo Ibanez, Chile

28. Suvidha Kancharla
Degree: MSc, May 2014
Current location: Software Engineer, Microsoft

29. Seong Gyeong Lee
Degree: BSc, May 2014
Research project: Sentiment Analysis of Facebook Messages

30. Karthik Nagaraj
Degree: PhD, October 2013
Thesis: Enabling Richer Insight Into Runtime Execution of Systems
Current location: Software Engineer, Google

31. Dan Coroian
Degree: BSc, May 2013
Research project: An Application of SARSA Learning to Klondike Solitaire
Current location: PhD Program, Computer Science, Duke University

32. Brian Donovan
Degree: BSc, May 2013
Research project: Using Facebook Text to Predict Social Characteristics
Current location: PhD Program, Civil Engineering, UIUC

33. Christopher Cole
Degree: BSc, May 2013
Research project: Analysis of Dynamic Email Graphs
Current location: Software Engineer, Amazon

34. Daniel Roberts
    Degree: BSc, May 2013
    Research project: The Impact of Interpersonal Dependencies on Distributed Teams
    Current location: Software Engineer, EMC Isilon

35. Rongjing Xiang
    Degree: PhD, August 2012
    Thesis: Statistical Relational Learning for Single Network Domains
    Current location: Software Engineer, Google

36. Hoda Eldardiry
    Degree: PhD, February 2012
    Thesis: Ensemble Classification Techniques for Relational Domains
    Current location: Associate Professor, Virginia Tech University, CS

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

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

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