# Nitesh V. Chawla, PhD

**Fellow AAAI, AAAS, ACM, IEEE**
Frank M. Freimann Professor of Computer Science and Engineering
Lucy Family Director for Data & AI Academic Strategy
nchawla@nd.edu
http://niteshchawla.nd.edu


## Higher Education

B.E., Computer Science, University of Pune, India, 1997.
M.S., Computer Science and Engineering, University of South Florida, 1999.
Ph.D., Computer Science and Engineering, University of South Florida, 2002.

## Professional Appointments

- Lucy Family Director for Data & AI Academic Strategy, 2025 -
- Inaugural Director, Notre Dame Strategic Initiative on Data, AI and Computing, 2025 -
- Founding Co-Editor-In-Chief, *ACM AI Letters,* 2025 -
- Founding Director, Lucy Family Institute for Data & Society, University of Notre Dame, 2020 – 2026
- Frank M. Freimann Professor of Computer Science and Engineering, University of Notre Dame, 2015 –
- Concurrent Professor of Information Technology Analytics & Operations, University of Notre Dame
- Concurrent Professor of Applied and Computational Mathematics and Statistics, University of Notre Dame, 2016 –
- Co-founder, Heddle Health, 2025 -
- Co-founder, Pahtia Technologies, 2024 –
- Co-founder, Intrepid Phoenix, 2021 - 2024
- Co-founder, Aunalytics, Inc., 2012 –
- Director, Data Inference Analytics and Learning Lab, University of Notre Dame, 2007 –
- Interim Director of Applied Data Science, Indiana Biosciences Research, 2016 – 2018
- Frank M. Freimann Collegiate Associate Professor, Computer Science and Engineering, University of Notre Dame, 2011 – 2015
- Assistant Professor, Computer Science and Engineering, University of Notre Dame, 2007 – 2011
- Founding Director, ND-GAIN Index, University of Notre Dame, 2013 – 2016
- Research Assistant Professor, University of Notre Dame, 2004 – 2006
- Senior Risk Modeling Manager, Retail Risk Analytics, Canadian Imperial Bank of Commerce, 2002 – 2004
- Research Assistant, University of South Florida, 1997 – 2002

## Distinctions, Honors, Awards

### International and National

1. Inaugural IJCAI-JAIR Classic Paper Award for SMOTE: Synthetic Minority Over-sampling Technique, 2026
2. Member, National Academy of Artificial Intelligence, 2026
3. AAAS Fellow, 2024
4. AAAI Fellow, 2024
5. ACM Fellow, 2022
6. IEEE Fellow, 2021
7. AAIA Fellow, 2021
8. Heterogeneous Graph Neural Network Paper selected by the International Congress of Basic Science, as a recipient of the 2024 Frontiers of Science Award in Theoretical Computer and Information Sciences.
9. Research Featured in several distinguished outlets including Time Magazine, Wall Street Journal, Chicago Tribune, Harvard Business Review, Wired, InStyle, Elle Australia, Science Daily, Phys.Org, Washington Post, RunnersWorld, Mashable, Fast Company, NPR
10. 1st Source Bank Commercialization Award, 2017
11. IEEE CIS Outstanding Early Career Award, 2015
12. Featured as an innovator in the healthcare industry by PSFK's Future of Health report, 2014
13. Research featured in the UN Global Pulse Blogs, 2014
14. IBM Big Data Award, 2013
15. IBM Watson Faculty Award, 2012
16. Academic Advisory Council and Leadership Conclave for IIT, Gandhinagar, 2012 --
17. Distinguished Alumnus Award, University of Pune, India, 2011
18. National Academy of Engineering New Faculty Fellowship, 2005
19. Winner on a NeurIPS 2004 Classification Challenge Problem on Evaluating Predictive Uncertainty (supported by PASCAL Network of Excellence)
20. Outstanding Dissertation Award, University of South Florida, 2003
21. National Academy of Sciences/National Research Council Post-doctoral Fellowship Award, 2002 (Declined for another opportunity)

### University and Community

1. Notre Dame All Faculty Team, 2024. "*At every home game, the provost honors a distinguished member of the faculty as part of the All-Faculty Team, chosen from across the disciplines for their transformative contributions to Notre Dame and beyond*."
2. Frank M. Freimann Endowed Professor of Computer Science and Engineering, 2015 --
3. Rodney F. Ganey Community Based Research Award, 2014
4. Michiana 40 under 40 Honor, 2013
5. Outstanding Undergraduate Teaching Award

## Research

1. Best Paper Candidate, ACM WWW, 2019
2. Best Application Paper, IEEE Data Science and Advanced Applications, 2017
3. Best Paper Award, IEEE Computational Intelligence Magazine, 2016
4. Best Paper Nominee, ACM Conference on Web Search and Data Mining, 2015
5. Best Paper Runner Up, 4th International Conference Learning Analytics and Knowledge, 2014
6. Best Paper Nominee, 2nd International Conference on Big Data and Analytics in Healthcare, 2014
7. Best Paper Runner Up, 121st ASEE Annual Conference and Exposition, 2014
8. Best of ACM/IEEE ASONAM'2012 Papers
9. Best of ACM/IEEE ASONAM'2011 Papers
10. Advisor of student group that won the DREAM6 Challenge, 2011, organized by IBM DREAM (Dialogue for Reverse Engineering Assessments and Methods) poses fundamental questions about systems biology, and invites participants to propose solutions
11. Paper titled, "Exploring and Exploiting Disease Interactions from Multi-relational Gene and Phenotype Networks," selected for the 2012 IMIA Yearbook of Medical Informatics, 2012 as best of medical informatics papers published in 2011
12. Best Student Paper, ACM SIGKDD Workshop on Sensor Networks, 2009
13. Best of ACM/IEEE ASONAM'2009 Papers
14. NET Institute Award, 2008
15. Best of IEEE ICDM'2007 Papers
16. Best Student Paper, NAACSOS, 2006
17. Best Student Paper Nominee, ACM/IEEE Supercomputing, 2006
18. Notable mention at NIPS 2003 Workshop Feature Selection Challenge (Invited to make a presentation and write a book chapter)

## Teaching

1. National Academy of Engineering New Faculty Fellowship
2. Students' chosen speaker at the Department of Computer Science and Engineering, University of Notre Dame, Graduation Ceremony, 2013
3. Students' chosen speaker at the Department of Computer Science and Engineering, University of Notre Dame, Graduation Ceremony, 2012
4. Outstanding Undergraduate Teaching Award, Department of Computer Science and Engineering, University of Notre Dame, 2011
5. Outstanding Undergraduate Teaching Award, Department of Computer Science and Engineering, University of Notre Dame, May 2008

## Students' Honors

1. PhD student, Chuxu Zhang, selected for New Faculty Highlight, AAAI, 2023
2. PhD student, Yuxiao Dong, received the Rising Star Award, ACM SIGKDD, 2022
3. PhD student Mandana Saebi received the ACM Grace Hopper Outstanding Research Award, 2020.
4. PhD student, Yuxiao Dong, received the Runner-Up for the Outstanding Dissertation Award at *ACM KDD*, 2017
5. PhD student Aastha Nigam received Second Prize for Best Student Research at *ACM Grace Hopper Computing*, 2017
6. PhD student, Pamela Bilo Thomas, recipient of the IBRI Fellowship, 2017
7. PhD student, Keith Feldman, recipient of the CRA-E Fellowship, 2016, 2017
8. PhD student, Yuxiao Dong, selected for the CSE Research Excellence Award, 2016
9. PhD student, Jian Xu, selected for the Outstanding Research Poster Award – Faculty Vote, 2016
10. PhD student, Keith Feldman, selected as the CRA-E Fellow, 2016
11. PhD student, Reid Johnson, selected for the Data Science for Social Good Fellowship, 2015
12. PhD student, Yang Yang, recipient of the CSE Research Excellence Award, 2015
13. PhD student, Everaldo Aguiar, selected for the Data Science for Social Good Fellowship, 2014
14. PhD student, Dipanwita Dasgupta, selected for the Trinity Health System Analytics Fellowship, 2014 (inaugural fellow)
15. PhD student, Keith Feldman, recipient of Outstanding TA Award, 2013
16. Undergraduate student Robert Thompson, recipient of the Steiner Award, 2012
17. PhD student Karsten Steinhaeuser, recipient of the CRC Award for Computational Science and Visualization, 2010
18. PhD student Ryan Lichtenwalter, recipient of the CRC Award for Computational Science and Visualization, 2012
19. Undergraduate student Jake Lussier, recipient of the NSF Graduate Student Fellowship, 2011
20. Undergraduate student Jim Notwell, recipient of the NSF Graduate Student Fellowship, 2010
21. PhD student Karsten Steinhaeuser, selected for presentation to the NSF-sponsored doctoral consortium on computational sustainability, 2010
22. PhD student Darcy Davis, recipient of fellowship for SIAM Data Mining Dissertation Forum, 2010
23. Undergraduate student Jake Lussier, awarded the Barry M. Goldwater Scholarship, 2010
24. PhD student Karsten Steinhaeuser, recipient of an ORNL Significant Event Award (SEA) for "Science Support for a Climate Change War Game and Follow-Up Support to the U.S. Department of Defense," 2009
25. PhD student David Cieslak, recipient of CSE Research Excellence Award, 2008

## Grants and Sponsored Programs

### External Research Grants / Gifts

*Total value of over $83M.*

1. Health Data Asset, *CICP,* $1M, 2026, PI.
2. Proactive 311: Bloomberg Mayoral Challenge project, *Bloomberg,* $1M, 2026, PI.
3. Notre Dame Planning Grant for AI in Higher Education, *Lilly Endowment Inc,* $199,999, 2026. PI.
4. Regional AI Model Adaptation and GenAI Outcomes Analysis*, MasterCard,* $150,000, 2026, PI.
5. Efficient MedLM: Trustworthy and Efficient Healthcare LLM, *Optum,* $150,000, 2025 – 2026, Co-PI.
6. Combatting Online Abuse Against Athletes: A Comprehensive Framework, *Fédération Internationale de l'Automobile (FIA),* 2025 – 2028, $427,999, PI.
7. Collaborative Research: III: Small: Data-driven and AI-powered Innovations for Tackling the Nexus of Poverty-driven Food Insecurity and the Opioid Crisis, *NSF,* 2025 – 2028, $360,000, Co-PI.
8. Explorations: Data Crossings - Bridging to Data Science Careers through Experiential Learning and Training, *NSF,* $999,523, Co-PI
9. SCC-IRG Track 1: A New AI-driven Paradigm to Promote Community Resilience for Teenagers and Young Adults in Preventing Opioid Misuse and Addiction**,** *National Science Foundation***,** $1.5M, 2024– 2028, Co-PI.
10. A Safety-Aware Ecosystem of Interconnected and Reputable sUAS, *National Aeronautics and Space Administration,* $5,297,873, 2023 – 2025. Co-PI.
11. Proto-OKN Theme 1: Exploiting Federal Data and Beyond: A Multi-modal Knowledge Network for Comprehensive Wildlife Management under Climate Change, *NSF,* $1.5M, 2023 – 2026, Co-PI.
12. AnalytiXIN, *Lilly Endowment Inc.,* 2021 – 2025, $6M, PI.
13. Center for Computer Assisted Synthesis (Phase II), *National Science Foundation,* 2022 – 2027, $20M, Co-PI (AI Lead and Executive Committee for the Center).
14. EAGER: A New Explainable Multi-objective Learning Framework for Personalized Dietary Recommendations against Opioid Misuse and Addiction, *National Science Foundation*, $300,000. 2023-2024. Co-PI.
15. Socio Behavioral Science for Public Health, *Saint Joseph County Department of Health,* $345,010, December 2022 - 24. PI
16. Portfolio Synthesis, Intervention Profiles and Decision Support: Analyzing Global Affairs Canada's Agriculture and Food Systems Portfolio Alignment with Ceres2030 High-Impact Interventions*, Global Affairs Canada*, $191,701, November 2022, PI.

17. Health Equity Data Project, *Accenture*, $2.25M, 2022, PI.
18. Institutional Storage for the University of Notre Dame, *National Science Foundation,* 2022 – 2024, $500K, Co-PI
19. SCC-CIVIC-PG Track B: Community-Based Research meets Systems Approach: Closing the Loop on Child Lead Poisoning, *National Science Foundation, 2022 – 2024, $50K,* PI.
20. Safe Deployment of Small Unmanned Aerial Systems through On-Board Monitoring and Assessment, *NASA,* 2021 – 2024, $749,850, Co-PI.
21. SCC-IRG Track 1: Food Information Networks (FINs):Building data-driven supports for increasing access and healthy food choices in low-income neighborhoods, *Department of Agriculture,* $1.7M, 2020 – 2024, Co-PI.
22. Applied Analytics and Emerging Technology Lab, *Lily Endowment,* 2020 – 2025, $5M*,* PI.
23. HDR: DSC: Interdisciplinary Traineeship for Socially Responsible and Engaged Data Scientists (iTREDS), *National Science Foundation,* 2019 – 2023, $1.2M, PI
24. Center for Computer Assisted Synthesis (Phase I), *National Science Foundation,* 2019 – 2021, $1.8M, Co-PI (Machine Learning Lead).
25. Modeling and Personalizing Consumer Engagement, *Conde Nast,* 2019 – 2020, $150K, PI
26. Scalable biological agent detection network, *Physical Sciences Corp (STTR from Army Research Office)),* $260,000, 2018 – 2020, PI.
27. A Comprehensive Approach to Modeling Job Performance via Unobtrusive, Continuous, Multimodal Sensing, IARPA, $10M, 2017 – 2020, Co-PI (Machine Learning Lead).
28. Applied Data Science Core, *Indiana Biosciences Research Institute,* $199,000, PI, 2017 – 2019, PI.
29. Knowledge Management Paradigm, *Boeing,* $493,600, 2017 – 2019, PI.
30. Evaluation of an integrated community based pre-natal care coordination platform for at risk families, *INDIANA CTSI,* $25,000, 2017, PI.
31. Cognitive systems research, *IBM Research***,** $240,000, 2015 – 2018, PI.
32. EAGER: Developing scalable benchmark mini-apps for graph engine comparison, *NSF,* 2016 – 2018, $299,869, Co-PI.
33. II-New: Infrastructure for Supporting Biomedical Application Algorithms, Runtime Development and Resource Management, *NSF,* 2016 – 2019, $500,000, Co-PI.
34. Social and Cognitive Networks Research Center, *Army Research Labs,* 2009 – 2019, $2M, PI.
35. BigData: F: DKM: Addressing the two V's of Veracity and Variety in Big Data, *NSF,* 2014 – 2019, $1M, PI.
36. Coastal SEES Collaborative Research: Changes in Ship-borne Introductions of Invasive Species in Coupled Natural-human Systems: Infrastructure, Global Trade, Climate and Policy, *NSF,* 2014 – 2019, $1.7M. Co-PI.

37. NetHealth: Modeling the Co-Evolution of Social Networks and Health Behaviors, *NIH,* 2014 – 2019, $2.9M, Co-PI.
38. ND-GAIN Urban Adaptation Assessment, *Kresge Foundation,* 2015 – 2017, $483,000, Co-PI
39. Learning Analytics, *PwC,* 2014 – 2017, $300,000, PI
40. Student Travel Support for the 2015 SIAM Data Mining Conference, 2015, *NSF,* $35,424*,* PI.
41. Fraud Analytics, *Windhaven,* 2014 – 2016, $228,000, PI.
42. Community Health Analytics Program, *Everyday Health*, 2014, $200,000, PI.
43. eSeniorCare, *Memorial Hospital of South Bend*, 2013, $86,222, PI.
44. Diabetics Management: A personalized approach, *INDIANA CTSI*, 2013, $25,000, PI.
45. Web-based Mutli-player Social Environment to Reduce Middle School Student Obesity, *United Way Foundation*, 2013 – 2014, $120,000, PI.
46. Ethnography, Oral History, and Network Analysis of Trader Behaviors and Institutions in India, Kenya, and South Sudan, *National Science Foundation (NSF)* 2013, $78,845, Co-PI.
47. ND-Global Adaptation Index, 2013 – 2017, $1,960,000, Co-PI.
48. MRI: Acquisition of a Data Analytics Cluster for Computational Social Science, *National Science Foundation (NSF),* 2012, $451,839, Co-PI.
49. Ensemble-Based Modeling of Large Graphs and Its Applications to Social Networks, 2012 – 2016, *DARPA,* $2.34M, Co-PI.
50. Geroscience Pilot Proposal: Translational Biomedical Informatics in Aging and Aging Associated Disease, 2012 – 2013, Sub-contract from *Buck Institute for Aging* (*NIH)*, $15,000, PI.
51. Conference on Intelligent Data Understanding, 2012, *NASA,* $18,000, PI.
52. Scaling semantic networks, 2012 – 2013, *Battelle Memorial Institute,* $150,000, PI.
53. CDI-Type II: Building and Studying a Virtual Organization for Adaptation to Climate Change, 2010 – 2014, *National Science Foundation,* $1,560,000, PI.
54. Modeling Extremes and Uncertainty in Climate Data Sciences, 2010, *Oak Ridge National Labs,* $15,000, PI.
55. Incremental Learning from Unbalanced Data in Non-stationary Environments, *National Science Foundation,* 2009 – 2012, $164,948, PI.
56. Advanced Analytics: Charting the Path from Data to Knowledge to Insight*, Boeing*, 2009 – 2012, $672,000, PI.
57. Longitudinal Analysis and Modeling of Large-Scale Social Networks Based on Cell Phone Records, *National Science Foundation,* 2008 – 2011, $711,770, Co-PI.
58. Development of a Research Infrastructure for the Multithreaded Computing Community Using the Cray Eldorado, *National Science Foundation,* 2007 – 2012, $500,000, Co-PI.

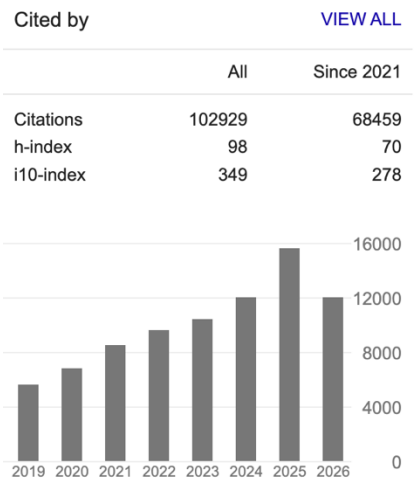
59. Troubleshooting the NWICG Grid, *NWICG*, $25,000, 2007 – 2008, PI.
60. Troubleshooting distributed systems with machine learning, *National Science Foundation*, 2007 – 2008, $30,000, PI.
61. Face recognition from video, *National Institute of Justice*, 2006, $362,476, Co-PI.
62. TeamTrak: An experimental testbed for collaborative wireless environments, *Department of Defense*, 2006, $200,000, Co-PI.

## Internal Funding

- The Interplay of Poverty-driven Food Insecurity and the Opioid Crisis: Examining Their Impact and Developing Integrated Interventions, Co-PI, $299,908, 2025-2028
- CAR.L.L.: A LLM agent for physiologic inversion, *BELS Initiative,* Co-PI, $399,000, 2025 – 2027.
- The Connected World: Global Economic Commerce, Climate Change, and the Spread of Invasive Species, *ND Environmental Change Initiative SRI*, $575,000, PI.
- Notre Dame Collaboratory for the Study of Adaptation to Climate Change, ND SRI, $100,000, Co-PI.

## Publications

*Source: Google Scholar. Sep 2026*




**Books**

1. Yu, T., Chawla, N. V., Simoff, S., (Editors) "Computational Intelligent Data Analysis for Sustainable Development," CRC Press, 2013
2. Vatsavai, R., Omitomu, O. Gama, J., Chawla, N. V., Gaber, M., Ganguly, A., (Editors) "Knowledge Discovery from Sensor Data," Springer, 2010.

**Journal Articles**

1. Li, S., Lalor, J. P., Ahmad, F., Abbasi, A., & Chawla, N. V. (2026). "Modeling Edge-Rich Graphs using Neural Networks," *IEEE Transactions on Knowledge and Data Engineering*.
2. Chawla, N. V. et al. (2026). "Organic Chemistry as a Catalyst for AI Innovation: Challenges, Methods, and Emerging Paradigms. *Chemical Reviews,* 126 (13): 7587–7635.
3. Dong, K., Guo, Z., & Chawla, N. V. (2026). "CORE: Data Augmentation for Link Prediction via Information Bottleneck." *ACM Transactions on Knowledge Discovery from Data*, *20*(3), 1-28.
4. Wang, S., Dong, Y., Zhang, B., Chen, Z., Fu, X., He, Y., Shen, C., Zhang, C., Chawla, N.V. and Li, J., (2026). "Safety in graph machine learning: Threats and safeguards," *IEEE Transactions on Knowledge and Data Engineering*.
5. Sokol, A., Moniz, N., & Chawla, N. V. (2026). "Conformalized selective regression," *Discover Data*, *4*(1), 14.
6. Han, D., Moniz, N., & Chawla, N. V. (2026). "Generation of Loss Functions from Matrix-Based Binary Classification Metrics" *ACM Transactions on Knowledge Discovery from Data*.
7. Zhou, Yujun, et al (2025). "Benchmarking large language models on safety risks in scientific laboratories." *Nature Machine Intelligence:* 1-12.
8. García-Martínez, A., Serván-Mori, E., Márquez-González, H., Schnur, J., Orozco, E., Torres, P. A. C., & Chawla, N. V. (2025). "Embedding social determinants in mHealth for pediatric oncology: co-designing a patient-centred tool for febrile neutropenia in resource-limited settings," International *Journal for Equity in Health*.
9. Germino, J., Brinkmann, B., Chawla, N. V., Cui, J., Fortune, E., Moniz, N., & Faust, L. (2025). "Mixture of checkpoint experts for explainable seizure detection using wearable devices," *Scientific reports*, *15*(1), 43976.
10. Dong, K., Mao, H., Guo, Z. and Chawla, N. V. (2025), "Universal Link Predictor By In-Context Learning on Graphs," *Transactions of Machine Learning Research.*
11. Guo, Z., Zhao, T., Liu, Y., Dong, K., Shiao, W., Ju, M., Shah, N., and Chawla, N. V., (2025) **"**Node Duplication Improves Cold-start Link Prediction," *Transactions of Machine Learning Research.*
12. Khvatskii, G., Lee, Y. S., Angst, C., Gibbs, M., Landers, R., & Chawla, N. V. (2025). "Do multimodal large language models understand welding?" *Information Fusion*, *120*, 103121.
13. Ma, Y., Tian, Y., Moniz, N., & Chawla, N. V. (2025). "Class-imbalanced learning on graphs: A survey," *ACM Computing Surveys*, *57*(8), 1-16.
14. Fogarty, B., García-Martínez, A., Chawla, N. V., & Serván-Mori, E. (2025). "Social and economic predictors of under-five stunting in Mexico: a comprehensive approach through the XGB model," *Journal of Global Health*, *15*, 04065.
15. Tian, Y., Pei, S., Zhang, X., Zhang, C., & Chawla, N. V. (2025). "Knowledge distillation on graphs: A survey." *ACM Computing Surveys*, *57*(8), 1-16.

16. Carvalho, T., Moniz, N., Antunes, L., & Chawla, N. V. (2025). Differentially-private data synthetisation for efficient re-identification risk control. *Machine Learning*, *114*(7), 164.
17. Germino, J., Brinkmann, B., Chawla, N. V., Worrell, G., & Faust, L. (2025). Mixture of checkpoint experts for explainable seizure detection using wearable devices. *Scientific Reports*, *15*, 43976.
18. Germino, J., Moniz, N., and Chawla, N. V. (2024) "FairMOE: Counterfactually-Fair Mixture of Experts with Levels of Interpretability," *Machine Learning*: 1-21.
19. Dablain, D., Bellinger, C., Krawczyk, B., Aha, D. W., & Chawla, N. (2024). "Understanding imbalanced data: XAI & interpretable ML Framework," *Machine Learning*, 1-19.
20. Dablain, D., Jacobson, K., Bellinger, C., Roberts, M., and Chawla, N. V. (2023). "Understanding CNN fragility when learning with imbalanced data," *Machine Learning*, https://doi.org/10.1007/s10994-023-06326-9
21. Dablain, D., Krawczyk, B., and Chawla, N. V. (2024) "Towards a holistic view of bias in machine learning: bridging algorithmic fairness and imbalanced learning," *Discover Data* 2.1 (2024): 4.
22. Ma, Y., Islam, M., Cleland-Huang, J., and Chawla, N. V. (2023). "Detecting Anomalies in Small Unmanned Aerial Systems via Graphical Normalizing Flows," *IEEE Intelligent Systems*, doi: 10.1109/MIS.2023.3252810.
23. Andrés, J., et al. (2023). "Environment and shipping drive environmental DNA beta-diversity among commercial ports," *Molecular Ecology*, 00, 1– 14. doi:10.1111/mec.16888.
24. Saebi, M., Nan, B., Herr, J., Wahlers, J., Guo,. Z., Zuranski, A., Kogej, T., Norrby, P., Doyle, A., Chawla, N.V., and Wiest, O (2023). "On the use of real-world datasets for reaction yield prediction," *Chemical Sciences*, doi: 10.1039/D2SC06041H.
25. Schnur, J. & Chawla, N. V. (2022). "Information Fusion via symbolic regression: A tutorial in the context of human health," *Information Fusion*. doi: 10.1016/j.inffus.2022.11.030.
26. Saebi, M., Kreig, S., Zhang, C., Jiang, M., Kajdanowicz, T., & Chawla, N. V. (2022). "Heterogeneous relational reasoning in knowledge graphs with reinforcement learning," *Information Fusion*, *88*, 12-21.
27. Bielak, P., Kajdanowicz, T., & Chawla, N. V. (2022). "Graph Barlow Twins: A self supervised representation learning framework for graphs," *Knowledge-Based Systems*, 109631.
28. Bielak, P., Kajdanowicz, T., and Chawla, N. V. (2022). "AttrE2vec: Unsupervised attributed edge representation learning," *Information Sciences,* 592, 82 – 96.
29. Avilés-Robles, M., Schnur, J. J., Dorantes-Acosta, E., Márquez-González, H., Ocampo-Ramírez, L. A., & Chawla, N. V. (2022). "Predictors of Septic Shock or Bacteremia in Children Experiencing Febrile Neutropenia Post-Chemotherapy," *Journal of the Pediatric Infectious Diseases Society*.
30. Krieg, S. J., Schnur, J. J., Miranda, M. L., Pfrender, M. E., & Chawla, N. V. (2022). "Symptomatic, Presymptomatic, and Asymptomatic Transmission of SARS-CoV-2 in a

University Student Population, August–November 2020," *Public Health Reports*, 00333549221110300.

31. Wu, X., Granda, P., Huang, C., and Chawla, N. V. (2022). "Representation Learning on Variable Length and Incomplete Wearable-Sensory Time Series," *ACM Transactions on Intelligent Systems and Technology.*
32. Wang, D., Zhao, Yu, W., Chawla, N. V., and Jiang, M. (2022), "Deep Multimodal Complementarity Learning," *IEEE Transactions on Neural Networks and Learning Systems.*
33. Guo, Z., Tao, J., Chen, S., Chawla, N. V., Wang, C. (2022), "SD^ 2: Slicing and Dicing Scholarly Data for Interactive Evaluation of Academic Performance," *IEEE Transactions on Visualization and Computer Graphics.*
34. Krieg, S., Smith, C., Chatterjee, R., and Chawla. N. V. (2022), "Predicting terrorist attacks in the United States using localized terrorist data," *Plos One.*
35. Marshall, J., Jimenez-Pazmino, P., Metoyer, R., & Chawla, N. V. (2022). "A survey on healthy food decision influences through technological innovations," *ACM Transactions on Computing for Healthcare (HEALTH)*, *3*(2), 1-27.
36. Krieg, S., Avendano, C., Grantham-Brown, E., Asburn, A. L., Schnur, J., Miranda, M L, and Chawla, N. V (2022), "Data-driven testing program improves detection of COVID-19 cases and reduces community transmission," *npj Digital Medicine.*
37. Dablain, D., Krawczyk, B., and Chawla, N. V. (2022), "DeepSMOTE: Fusing deep learning and SMOTE for imbalanced data," *IEEE Transactions on Neural Networks and Learning systems.*
38. Amon, M.J., Mattingly, S., Necaise, A., Mark, G., Chawla, N. V., Dey, A., and D'Mello, S (2022), Flexibility versus routineness in multimodal health indicators: A sensor-based longitudinal in situ study of information workers, *ACM Transactions on Computing for Health.*
39. Tian, Y., Zhang, C., Metoyer, R., & Chawla, N. V. (2022). "Recipe Recommendation with Hierarchical Graph Attention Network," *Frontiers in Big Data*, 123.
40. Faust, L., Feldman, K., Lin, S., Mattingly, S., D'Mello, S., & Chawla, N. V. (2021). "Examining Response to Negative Life Events Through Fitness Tracker Data," *Frontiers in digital health*, *3*, 37.
41. Mirjafari, S., Bagherinezhad, H., Nepal, S., Martinez, G. J., Saha, K., Obuchi, M., ... & Campbell, A. T. (2021). "Predicting Job Performance Using Mobile Sensing," *IEEE Pervasive Computing*, *20*(4), 43-51.
42. Martinez, G. J., Mattingly, S. M., Robles-Granda, P., Saha, K., Sirigiri, A., Young, J., ... & Striegel, A. (2021). "Predicting Participant Compliance With Fitness Tracker Wearing and Ecological Momentary Assessment Protocols in Information Workers: Observational Study," *JMIR mHealth and uHealth*, *9*(11), e22218.
43. Bielak, P., Tagowski, K., Falkiewicz, M., Kajdanowicz, T., and Chawla, N. V. (2021), "FILDNE: A Framework for Incremental Learning of Dynamic Networks Embeddings," *Knowledge-Based Systems.*

44. Wang, D., Zeng, Q., Chawla, N. V., & Jiang, M. (2021). "Modeling Complementarity in Behavior Data with Multi-Type Itemset Embedding, "*ACM Transactions on Intelligent Systems and Technology (TIST)*, *12*(4), 1-25.
45. Syed, M., Wang, D., Jiang, M., Conway, O., Juneja, V., Subramanian, S., & Chawla, N. V. (2021). "Unified Representation of Twitter and Online News Using Graph and Entities," *Frontiers in big Data*, *4*.
46. Wang, D., Zhang, Z., Ma, Y., Zhao, T., Jiang, T., Chawla, N., & Jiang, M. (2021). "Modeling Co-evolution of Attributed and Structural Information in Graph Sequence," *IEEE Transactions on Knowledge and Data Engineering*.
47. Feldman, K., Rohan, A. J., & Chawla, N. V. (2021). "Discrete Heart Rate Values or Continuous Streams? Representation, Variability, and Meaningful Use of Vital Sign Data," *CIN: Computers, Informatics, Nursing*, *39*(11), 793-803.
48. Zhang, C., Yao, H., Yu, L., Huang, C., Song, D., Chen, H., ... & Chawla, N. V. (2021). "Inductive Contextual Relation Learning for Personalization," *ACM Transactions on Information Systems (TOIS)*, *39*(3), 1-22.
49. Faust, L., Feldman, K., Lin, S., Mattingly, S., D'Mello, S., & Chawla, N. V. (2021). "Examining Response to Negative Life Events through Fitness Tracker Data," *Frontiers in Digital Health*, *3*, 37.
50. Robles-Granda, P., Lin, S., Wu, X., et al. (2021). "Jointly predicting job performance, personality, cognitive ability, affect, and well-being," *IEEE Computational Intelligence Magazine*, *16*(2), 46-61.
51. Krieg, S. J., Schnur, J. J., Marshall, J. D., Schoenbauer, M. M., & Chawla, N. V., (2020), "Pandemic Pulse: Unraveling and Modeling Social Signals during the COVID-19 Pandemic," *ACM Digital Government: Research and Practice*, *2*(2), 1-9.
52. Saebi, M., Xu, J., Curasi, S. R., Grey, E. K., Chawla, N. V., & Lodge, D. M., (2020). "Network analysis of ballast-mediated species transfer reveals important introduction and dispersal patterns in the Arctic," *Scientific reports*.
53. Kundu, S., Kajdanowicz, T., Kazienko, P. and Chawla, N., (2020), "Fuzzy Relative Willingness: Modeling Influence of Exogenous Factors in Driving Information Propagation Through a Social Network," *IEEE Access*, *8*, pp.186653-186662.
54. Saebi, M., Ciampaglia, G.L., Kaplan, L.M. and Chawla, N.V., (2020), "HONEM: learning embedding for higher order networks," *Big Data*, *8*(4), pp.255-269.
55. Saebi, M., Xu, J., Grey, E.K., Lodge, D.M., Corbett, J.J. and Chawla, N., (2020), "Higher-order patterns of aquatic species spread through the global shipping network," *Plos one*, *15*(7), p.e0220353.
56. Jiang, T., Zeng, Q., Zhao, T., Qin, B., Liu, T., Chawla, N., & Jiang, M., (2020), "Biomedical knowledge graphs construction from conditional statements," *IEEE/ACM transactions on computational biology and bioinformatics*.
57. Barbieri, D., Chawla, N., Zaccagni, L., Grgurinović, T., Šarac, J., Čoklo, M., & Missoni, S. (2020). Predicting cardiovascular risk in athletes: resampling improves classification

performance. *International Journal of Environmental Research and Public Health*, *17*(21), 7923.

58. Saebi, M., Xu, J., Kaplan, L. M., Ribeiro, B., & Chawla, N. V. (2020). Efficient modeling of higher-order dependencies in networks: from algorithm to application for anomaly detection. *EPJ Data Science*, *9*(1), 15.
59. Jiang, T., Zeng, Q., Zhao, T., Qin, B., Liu, T., Chawla, N. V., and Jiang, M. (2020), "Biomedical Knowledge Graph Construction from Conditional Statements," *IEEE / ACM Transactions on Computational Biology and Bioinformatics.*
60. Faust, L., Feldman, K., Mattingly, S., Hachen, D., and Chawla, N. V. (2020), "Deviations from normal bedtimes are associated with short-term increases in resting heart rate," *Nature Digital Medicine,* 3(2), 1 – 9.
61. Feldman, K., Solymos, G.M.B., de Albuquerque, M.P. & Chawla, N. V. (2019) *"*Unraveling Complexity about Childhood Obesity and Nutritional Interventions: Modeling Interactions Among Psychological Factors." *Nature Sci Rep* **9,** 18807.
62. Faust, L., Feldman, K., & Chawla, N. V. (2019). "Examining the weekend effect across ICU performance metrics." *Critical Care*, *23*(1), 207.
63. Lin, S., Faust, L., Robles-Granda, P., Kajdanowicz, T., & Chawla, N. V. (2019). "Social network structure is predictive of health and wellness." *PloS one*, *14*(6), e0217264.
64. Wang, S., Minku, L. L., Chawla, N., & Yao, X. (2019). "Learning from data streams and class imbalance." *Connection Science*, *31*, 103-104.
65. Faust, L., Wang, C., Hachen, D., Lizardo, O., and Chawla, N. V. (2019). "Physical Activity Trend eXtraction: A Framework for Extracting Moderate-Vigorous Physical Activity Trends From Wearable Fitness Tracker Data," *Journal of Medical Informatics Research (JMIR).*
66. Nigam, A., Johnson, R., Wang, D., and Chawla, N. V. (2019), "Characterizing Online Health and Wellness Information Consumption: A Study," *Information Fusion,* 46: 33 – 43.
67. Gonya, J., Harrison, T., Feldman, K., Stein, M., & Chawla, N. (2019). "Nursing networks in the NICU and their association with maternal stress: A pilot study." *Journal of Nursing Management*, doi: 10.1111/jonm.12679.
68. Wang, S., Minku, L., Chawla, N. V., and Yao, X. (2019), "Learning in the presence of class imbalance and concept drift," *Neurocomputing.*
69. Yang, Y., Chawla, N. V., & Uzzi, B. (2019). "A network's gender composition and communication pattern predict women's leadership success." *Proceedings of the National Academy of Sciences (PNAS)*, 201721438.
70. Tao, J., Imre, M., Wang, C., Chawla, N. V., Guo, H., Sever G., and Kim, S. (2019) "Exploring Time-Varying Multivariate Volume Data Using Matrix of Isosurface Similarity Maps," *IEEE Transactions on Visualization and Computer Graphics*, *25*(1), 1236-1245.
71. Thomas, P. B., Robertson, D. H., & Chawla, N. V. (2018). "Predicting onset of complications from diabetes: a graph-based approach," *Applied network science*, *3*(1), 48.
72. Gonya, J., Feldman, K., Brown, K., Stein, M., Keim, S., Boone, K., Chawla, N. V.& Butter, E. (2018). "Human interaction in the NICU and its association with outcomes on the Brief

Infant-Toddler Social and Emotional Assessment (BITSEA)." *Early human development*, *127*, 6-14.

73. Tao, J., Wang, C., Chawla, N. V., Shi, L., & Kim, S. H. (2018). "Semantic flow graph: A framework for discovering object relationships in flow fields." *IEEE transactions on visualization and computer graphics*, *24*(12), 3200-3213.
74. Li, S., Juang, Y., Chawla, N. V. and Zhu, Z. (2018) "Multi-Label Learning from Crowds," *IEEE Transactions on Knowledge and Data Engineering.*
75. Nagrecha, S., Johnson, R. A., & Chawla, N. V. (2018), "FraudBuster: Reducing Fraud in an Auto Insurance Market." *Big data*, *6*(1), 3-12.
76. Fernandez, A., Garcia, S., Herrera, F., and Chawla, N. V. (2018), "SMOTE for Learning from Imbalanced Data: Progress and Challenges, Marking the 15-year Anniversary," *Journal of Artificial Intelligence Research (JAIR),* 61: 863 – 905.
77. Huang, H., Dong, Y., Tang, J., Yang, H., Chawla, N.V. and Fu, X., (2018). "Will Triadic Closure Strengthen Ties in Social Networks?" *ACM Transactions on Knowledge Discovery from Data (TKDD)*, *12*(3), p.30.
78. Feldman, K., Johnson, R. A., & Chawla, N. V. (2018). "The State of Data in Healthcare: Path Towards Standardization," *Journal of Healthcare Informatics Research*, 1-24.
79. Feldman, K., Kotoulas, S., & Chawla, N. V. (2018). « TIQS: Targeted Iterative Question Selection for Health Intervention," *Journal of Healthcare Informatics Research*, 1-23.
80. Nigam, A., Dambanemuya, H.K., Joshi, M. and Chawla, N.V. (2017). "Harvesting Social Signals to Inform Peace Processes Implementation and Monitoring." *Big data*, *5*(4), pp.337-355.
81. Mursalin, M., Zhang, Y., Chen, Y., & Chawla, N. V. (2017). "Automated Epileptic Seizure Detection using Improved Correlation-based Feature Selection with Random Forest Classifier." *Neurocomputing*, *241*, 204-214.
82. Huang, C., Wang, D., & Chawla, N. (2017). "Scalable Uncertainty-Aware Truth Discovery in Big Data Social Sensing Applications for Cyber-Physical Systems." *IEEE Transactions on Big Data*.
83. Wang, S., Song, J., Yang, Y., Chawla, N. V., Ma, J., & Wang, H. (2017). "Rs12970134 near MC4R is associated with appetite and beverage intake in overweight and obese children: A family-based association study in Chinese population." *PloS one*, *12*(5), e0177983.
84. Wang, S., Song, J., Yang, Y., Zhang, Y., Chawla, N. V., Ma, J., & Wang, H. (2017). "Interaction between obesity and the Hypoxia Inducible Factor 3 Alpha Subunit rs3826795 polymorphism in relation with plasma alanine aminotransferase." *BMC medical genetics*, *18*(1), 80.
85. González, P., Castaño, A., Chawla, N.V. and Coz, J.J.D., (2017). "A Review on Quantification Learning." *ACM Computing Surveys (CSUR)*, *50*(5), p.74.
86. Yan, Z., Liu, J., Yang, L. T., & Chawla, N. (2017). "Big Data Fusion in Internet of Things," *IEEE Access.*

87. Bahulkar, A., Szymanski, B.K., Chawla, N., Lizardo, O. and Chan, K. (2017). "Influence of Personal Preferences on Link Dynamics in Social Networks. *Complexity*.

88. Fernández, A., del Río, S., Chawla, N. V., & Herrera, F. (2017). An insight into imbalanced Big Data classification: outcomes and challenges," *Complex & Intelligent Systems*, 1-16.

89. González, P., Díez, J., Chawla, N., & del Coz, J. J. (2017). "Why is quantification an interesting learning problem?" *Progress in Artificial Intelligence*, *6*(1), 53-58.

90. Dong, Y., Chawla, N.V., Tang, J., Yang, Y. and Yang, Y., 2017. "User Modeling on Demographic Attributes in Big Mobile Social Networks." *ACM Transactions on Information Systems (TOIS)*, *35*(4), p.35.

91. Tao, J**.,** Wang, C., Chawla, N., Shi, L. and Kim, S.H. (2017). "Semantic Flow Graph: A Framework for Discovering Object Relationships in Flow Fields." *IEEE Transactions on Visualization and Computer Graphics*.

92. Dasgupta, D., Chaudhry, B., Koh, E. and Chawla, N.V., (2016). "A Survey of Tablet Applications for Promoting Successful Aging in Older Adults. *IEEE Access*, *4*, pp.9005-9017.

93. Nagrecha, S. and Chawla, N. V. (2016) "Quantifying decision making for data science: from data acquisition to modeling," *EPJ Data Science,* 5:27, DOI: 10.1140/epjds/s13688-016-0089-x.

94. Wang, S., Song, J., Shang, X., Chawla, N. V., Yang, Y., Meng, X., Wang, H., and Ma. (2016) "Physical activity and sedentary behavior can modulate the effect of PNPLA3 variant on childhood NAFLD: a case-control study in a Chinese population." *BMC Medical Genetics,* 17 (1), 90.

95. Feldman, K., Stiglic, G., Dasgupta, D., Kircheff, M., Obradovic, Z., and Chawla, N. V. (2016), "Insights into population health management from disease diagnoses networks," *Nature Scientific Reports,* 6: 30465.

96. Kusimba, S., Yang, Y., and Chawla, N. (2016). "Heartholds of mobile money in western Kenya," *Economic Anthropology*, *3*(2), 266-279.

97. Dong, Y., Johnson, R. A., & Chawla, N. V. (2016). "Can Scientific impact Be predicted?" *IEEE Transactions on Big Data*, 2(1), 18-30.

98. Xu, Jian, Wickramarathene, T., and Chawla, N. V. (2016), "Representing higher-order dependencies in networks," *Science Advances, 2*(5), e1600028.

99. Siwo, G., Rider, A., Tan, A., Pinapati, R., Emrich, S., Chawla, N., V. and Ferdig, M. (2016). "Prediction of fine-tuned promoter activity from DNA sequence," *F1000Research*, *5*.

100. Feldman K., Darcy, D., and Chawla, N. V., (2015). "Scaling and contextualizing personalized healthcare: A case study of disease prediction algorithm integration," *Journal of Biomedical Informatics (JBI)*, 57: 377 – 385.

101. Feldman, K. and Chawla, N. V., (2015) "Does medical school training relate to practice? Evidence from Big Data," *Big Data Journal,* 3(2): 103-113. doi:10.1089/big.2014.0060. 2015. Featured on the cover page of journal.

102. Dong, Y., Tang, J., Chawla, N. V., Lou. T., Yang, Y., Wang, B., (2015) "Inferring social status and rich club effects in enterprise communication networks," *PLoS ONE* 10(3): e0119446. doi:10.1371/journal.pone.0119446.

103. Yang, Y., Lichtenwalter, R., Chawla, N. V. (2015) "Evaluating Link Prediction Methods," *Knowledge and Information Systems,* 45(3), 751—782, 2015.

104. Kusimba, S. B., Yang, Y. and Chawla, N. V., (2015), "Family networks of mobile money in Kenya," *Information Technologies and International Development,* 11(3), 1 – 21.

105. Yang, Y, Dong, Y, Chawla, N. V., (2014) "Predicting node centrality with the Node Prominence Profile, " *Scientific Reports (A Nature Publication), 14:7236* (Impact Factor: 5.228)

106. Zhou, Z., Chawla, N. V., Jin, Y., Williams, G., (2014) "Big Data opportunities and challenges: Discussions from data analytics perspectives," *IEEE Computational Intelligence Magazine*, 9(4): 62 – 74. (Impact Factor: 2.076)

107. Aguiar, E.**,** Ambrose, A., Chawla, N. V., Goodrich, V., and Brockman, J. (2014), "Engagement vs Performance: Using Electronic Portfolios to predict first semester engineering student persistence," *Journal of Learning Analytics,* 1 (3): 7 – 33.

108. Wickramarathne, T. L., Premaratne, K., Murthi, M., Chawla, N. V., (2014) "Convergence analysis of iterated belief revision in complex fusion environments," *IEEE Journal of Selected Topics in Signal Processing,* 8(4): 598 – 612.

109. Licthenwalter, R. and Chawla, N. V. (2014), "Vertex collocation profiles: theory, computation, and results," *SpringerPlus,* 3:116.

110. Rider, A., Milenkovic, T., Siwo, G., Emrich, S., Ferdig, M., Chawla, N. V. (2014) "Network characteristics are important for systems biology," *Network Science,* 2(2): 139-161.

111. Rider, A., Siwo, G., Emrich, S., Ferdig, M., Chawla, N. V. (2014), "Supervised Learning Approach to the Ensemble Clustering of Genes," *International Journal of Data Mining and Bioinformatics,* 9(2): 199-219.

112. Meyer, S., et al., (2013) "Inferring gene expression from ribosomal prompter sequences, a crowdsourcing approach," *Genome Research,* 23(11), 1928 –1937.

113. Chawla, N. V. and Davis, D. A., (2013) "Bringing Big Data to Personalized Healthcare: A Patient-Centered Framework," *Journal of General Internal Medicine,* 28 (3), 660 – 665.

114. Regola, N. and Chawla, N. V. (2013) "Storing and Using Health Data in Amazon Virtual Private Cloud," *Journal of Medical Internet Research,* 15(3):e63.

115. Hoens, T. R., Blanton, M., Steele A., Chawla, N. V., (2013) "Reliable Medical Recommendation Systems with Patient Privacy," *ACM Transactions on Intelligent Systems and Technology (ACM TIST),* 4 (4), 31 pages.

116. Wang, C., Lizardo, O., Hachen, D., Toroczkai, T., Chawla, N. V., (2013) "Weighted reciprocity in human communication networks," *Network Science,* 1 (1), pp. 31 – 48.

117. Ercsey-Ravasz, M., Lichtenwalter, R., Chawla, N. V., Toroczkai, Z., (2012) "Range-limited centrality measures in complex networks," *Physical Review E* 85.6 (2012): 066103.

118. Davis, D. A., Lichtenwalter, R., Chawla, N. V. (2012), "Supervised methods for multi-relational link prediction," S*ocial Network Analysis and Mining (SNAM*): 1-15.

119. Hoens, T. R., Polikar, R., Chawla, N. V. (2012), "Learning from streaming data with concept drift and imbalance," *Progress in Artificial Intelligence* 1.1 (2012): 89-101.

120. Steinhaeuser, K., Ganguly, A., Chawla, N. V. (2011), "Multivariate and multiscale dependence in the global climate system revealed through complex networks," *Climate Dynamics*, *doi:10.1007/s00382-011-1135-9.*

121. Davis, D. A. and Chawla, N. V. (2011), "Exploring and Exploiting Disease Interactions from Multi-relational Gene and Phenotype Networks," *PLOS One, 6(7), e22670.*

122. Torres, J., Raeder, T., Rodriguez, R., Chawla, N. V., Herrera, F. (2011), "A Unifying View on Dataset Shift in Classification," *Pattern Recognition,* DOI: 10.1016/j.patcog.2011.06.019.

123. Raeder, T., Lizardo, O., Hachen D. and Chawla, N. V. (2011), "Predictors of short-term decay of cell phone contacts in a large scale communication network," S*ocial Networks*, *33*(4), 245-257.

124. Lichtenwalter, R. and Chawla, N. V. (2011), "LPmade: Link Prediction Made Easy," *Journal of Machine Learning Research.*

125. Cieslak, D., Hoens, T. R., Chawla, N. V., Kegelmeyer, P. (2011) "Hellinger Distance Decision Trees and Robust and Skew Insensitive," *Data Mining and Knowledge Discovery. 24*(1), 136-158.

126. Steinhaeuser, K., Chawla, N. V., Ganguly, A. (2010), "Complex networks as a unified framework for descriptive analysis and predictive modeling in climate sciences," *Statistical Analysis and Data Mining Journal*, DOI: 10.1002/sam.10100. 2nd most cited paper published in the journal in 2011-2012.

127. Rider, A., Siwo, G., Chawla, N. V., Ferdig, M., Emrich, S., (2010) "A statistical approach to finding overlooked genetic associations," *BMC Bioinformatics*, 11(1), 526—624. Highly Accessed,

128. Raeder, T. and Chawla, N. V. (2010). "Market Basket Analysis with Networks," *Social Networks Analysis and Modeling (SNAM) Journal.* DOI: 10.1007/s13278-010-0003-7.

129. Steinhaeuser, K. and Chawla, N. V. (2010) "Identifying and Evaluating Community Structure in Complex Networks." *Pattern Recognition Letters,* 31(5), 413 – 421.

130. Davis, D. A., Chawla, N. V., Christakis, N., Barabasi, A.-L. (2010). "Time to CARE: A recommendation system for prospective health care," *Data Mining and Knowledge Discovery Journal,* 20(3)*,* 388 – 415.

131. Gray, J., Davis, D. A., Pursley, D.M., Smallcomb, J, Geva, A., and Chawla, N. V. (2010). "Network Analysis of Team Structure in the Neonatal Intensive Care Unit," *Pediatrics,* 125 (6), 1460 – 1467.

132. Beretta, L., Cappiello, F., Santaniello, A., Chawla, N. V., Allanore, Y., Mazzanoi, A. Bertolotti, F., Scorza, R., (2010) "Development of a five-years mortality model in systemic

sclerosis patients by different analytical approaches," *Clinical and Experimental Rheumatology*, *28*(2), S18.

133. Steinhaeuser, K., Chawla, N. V., and Ganguly, A. (2010). "An Exploration of Climate Data Using Complex Networks," *ACM SIGKDD Explorations,* 12(1), 25 – 32.

134. Raeder, T. and Chawla, N. V. (2009). "Model Monitor (M^2): Evaluating, Comparing, and Monitoring Models," *Journal of Machine Learning Research,* 10, 1387-1390.

135. Cieslak, D. and Chawla, N. V. (2009). "A Framework for Monitoring Classifiers' Performance: When and Why Failure Occurs?" *Knowledge and Information Systems (KAIS)*, 18(1), 83-108.

136. Tang, Y., Zhang, Y., Chawla, N. V. (2009). "SVMs Modeling for Highly Imbalanced Classification", *IEEE Transactions on Cybernetics,* 39(1): 281-288.

137. Taft, L., Evans, S., Egger, M., Chawla, N. V., Mitchell, J., Thornton, S., Bray, B., Varner, M. (2009). "Countering Imbalanced Datasets to Improve Adverse Drug Event Predictive Models in Labor and Delivery", *Journal of Biomedical Informatics,* 42(2): 356 – 364.

138. Vatsavai R., Omitaomu O., Gama J., Chawla N, Gaber M. M., Ganguly A., "Knowledge Discovery from Sensor Data," *ACM SIGKDD Explorations*, 10 (2): 68-73, December 2008.

139. Chawla, N. V., Cieslak, D., Hall, L., Joshi, A. (2008). "Automatically countering imbalance and its empirical relationship to cost" *Data Mining and Knowledge Discovery Journal,* 17(2), 225-252.

140. Liao, J. Q., Cieslak, D. A., Striegel, A. D., Chawla, N. V. (2008). "Using selective, short-term memory to improve resilience against DDoS exhaustion attacks", *Security and Communication Networks*, 1(4), 287 – 299.

141. Pawling, A., Chawla, N. V., Madey, G. (2007). "Anomaly detection in mobile communication networks," *Computational and Mathematical Organizational Theory (CMOT),* 2007: 3(4): 407-422.

142. Liu, Y., Chawla, N. V., Shriberg, E., Stolcke, A., Peskin, B., Harper, M. (2006). "A Study in Machine Learning from Imbalanced Data for Sentence Boundary Detection in Speech Recognition," *Computer Speech and Language,* 20(4), 468 – 494.

143. Chawla, N. V., Karakoulas, G. (2005). "Learning from labeled and unlabeled data: An empirical study across techniques and domains," *Journal of Artificial Intelligence Research,* 23, 331-366.

144. Chawla, N. V. "Book review for Discovering knowledge in data: An introduction to data mining," *Journal of Briefings in Bioinformatics,* June 2005 (Invited).

145. Radivojac, P., Chawla, N. V., Dunker, K., Obradovic, Z. (2004). "Classification and knowledge discovery in protein databases," *Journal of Biomedical Informatics*, 37(4), 224 – 239.

146. Chawla, N. V., Hall, L. O., Bowyer, K. W., Kegelmeyer, W. P., "Learning ensembles from bites: A scalable and accurate approach," *Journal of Machine Learning Research*, **5** (2004), 421 – 451.

147. Chawla, N. V., Japkowicz, N., Kolcz, A. (2004). "Learning from Imbalanced Datasets," *ACM SIGKDD Explorations*, 6(1), 1 – 6.

148. Chawla, N. V., Moore, T. E., Hall, L. O., Bowyer, K.W., Kegelmeyer, W.P., Springer C., "Distributed Learning with Bagging like Performance," *Pattern Recognition Letters,* Vol. 24 (1-3) (2003) pages: 455-471.

149. Chawla, N. V., Bowyer, K. W., Hall, L. O., Kegelmeyer, W.P. (2002) "SMOTE: Synthetic Minority Over-sampling Technique," *Journal of Artificial Intelligence Research (JAIR),* 16, 321 – 357.

150. Eschrich, S., Chawla, N. V, Hall, L. O., "Learning to predict in complex biological domains," *Journal of System Simulation*, vol. 14 (11), 2002, pages 1464-1471.

**Conference Proceedings**

1. Bystroński, M., Han, D., Chawla, N. V., & Kajdanowicz, T. J. "Continuous Context Sampling Allows Extending Diversity Boundaries of Large Language Models," In *Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics (ACL),* 2026.
2. Bao, Han, et al. "PolicyLLM: Towards Excellent Comprehension of Public Policy for Large Language Models." In *Findings of the Association for Computational Linguistics (ACL), 2026*.
3. Li, P., Huang, X., Hua, T., & Chawla, N. V. "CrochetBench: Can Vision-Language Models Move from Describing to Doing in Crochet Domain?." *Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics (ACL),* 2026.
4. Pan, D., Murugesan, K., Hua, T., Moniz, N., & Chawla, N. V. "Context attribution with multi-armed bandit optimization," In *Findings of the Association for Computational Linguistics (ACL),* 2026.
5. Huang, X., Ma, Y., Gurajapu, A., Schleinitz, J., Guo, Z., Reisman, S. E., & Chawla, N. V. "Chemhgnn: A hierarchical hypergraph neural network for reaction virtual screening and discovery," *ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD),* 2026.
6. Nogueira, B., Montiel, G., Jiang, M. Chawla, N. V., Moniz, N., "SPECTRA: Spectral Domain-Aware Graph Generation for Imbalanced Molecular Property Regression," *ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD),* 2026.
7. Zhou, et al. "Alignment Risks from Capability: Seeking RL Training," *International Conference on Machine Learning (ICML),* 2026.
8. Li, P., Tang, X., Chen, S., Cheng, Y., Metoyer, R., Hua, T., & Chawla, N. V. "Adaptive Testing for LLM Evaluation: A Psychometric Alternative to Static Benchmarks," *International Conference on Machine Learning (ICML),* 2026.
9. Liu, G., Chen, J., Zhu, Y., Sun, M., Luo, T., Chawla, N. V., & Jiang, M. "Graph diffusion transformers are in-context molecular designers," *International Conference on Learning Representations (ICLR),* 2026.

10. Huang, Y. et al., "TrustGen: A Platform of Dynamic Benchmarking on the Trustworthiness of Generative Foundation Models," *International Conference on Learning Representations (ICLR),* 2026.
11. Li, P., Huang, X., Hua, T., & Chawla, N. V. "CrochetBench: Can Vision-Language Models Move from Describing to Doing in Crochet Domain?" *64th Annual Meeting of the Association for Computational Linguistics (ACL),* 2026.
12. Pan, D., Murugesan, K., Moniz, N., & Chawla, N. "Context Attribution with Multi-Armed Bandit Optimization," *64th Annual Meeting of the Association for Computational Linguistics (ACL),* 2026.
13. Nogueira, B., Geyer, W., Anderson, A. A., Li, T. J. J., Kim, D., Moniz, N., & Chawla, N. V. "From Verification Burden to Trusted Collaboration: Design Goals for LLM-Assisted Literature Reviews," *31st International Conference on Intelligent User Interfaces*, 2026.
14. Tian, Y., Zhang, C., Kou, Z., Liu, Z., Zhang, X., & Chawla, N. V. "Adaptive and Context-rich Generative Self-supervised Learning on Graphs," In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2026.
15. Le, H. K., Dey, S., Galindo, M. M., Lopez, V., Hua, T., Chawla, N., & Hoang, L. T. "Towards Generalist Large Language Models for Molecular Property Prediction: Distilling Knowledge from Specialist Models." *AAAI Conference on Artificial Intelligence,* 2026.
16. Huang, Y. et al. "ChemOrch: Empowering LLMs with Chemical Intelligence via Groundbreaking Synthetic Instructions." *Advances in Neural Information Processing Systems (NeurIPS)* 38, 2026.
17. Wang, Z., Zhang, Z., Ma, T., Chawla, N. V., Zhang, C., & Ye, Y. "Beyond Message Passing: Neural Graph Pattern Machine," *42nd International Conference on Machine Learning (ICML),* 2025.
18. Wang, Z., Zhang, Z., Ma, T., Chawla, N. V., Zhang, C., & Ye, Y. "Towards Graph Foundation Models: Learning Generalities," 42nd *International Conference on Machine Learning (ICML),* 2025.
19. Sokol, A., Daly, E., Hind, M., Piorkowski, D., Zhang, X., Moniz, N., & Chawla, N. V. (2026). "BenchmarkCards: Standardized Documentation for Large Language Model Benchmarks," *Advances in Neural Information Processing Systems (NeurIPS)*, *38*, 2025.
20. Chen, S., Metoyer, R., Le, K., Acunin, A., Molnar, I., Ambrose, A., Chawla, N. V., & Metoyer, R. "Bridging the AI Adoption Gap: Designing an Interactive Pedagogical Agent for Higher Education Instructors." In *International Conference on Artificial Intelligence in Education* (pp. 171-178). Cham: Springer Nature Switzerland, 2025.
21. Zhang, Z., Wang, Z., Ma, T., Taneja, V.S., Nelson, S., Le, N.H.L., Murugesan, K., Ju, M., Chawla, N.V., Zhang, C. and Ye, Y., "Mopi-hfrs: A multi-objective personalized health-aware food recommendation system with llm-enhanced interpretation." *31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 1* (pp. 2860-2871), 2025.
22. Tian, Y., Han, Y., Chen, X., Wang, W., & Chawla, N. V., "Beyond answers: Transferring reasoning capabilities to smaller llms using multi-teacher knowledge distillation." In

*Proceedings of the Eighteenth ACM International Conference on Web Search and Data Mining* (pp. 251-260), 2025.

23. Wang, X., Yang, T., Rohr, J., Scheffers, B., Chawla, N., & Zhang, X. "WildlifeLookup: A Chatbot Facilitating Wildlife Management with Accessible Data and Insights." In *Proceedings of the Eighteenth ACM International Conference on Web Search and Data Mining* (pp. 1064-1067), 2025.
24. Guo, T., Ma, C., Chen, X., Nan, B., Guo, K., Pei, S., Wiest, O., Chawla, N.V. and Zhang, X.. "ReactionTeam: Teaming Experts for Divergent Thinking Beyond Typical Reaction Patterns," In *IEEE International Conference on Big Data (BigData)* (pp. 520-529). IEEE, 2025.
25. Sokol, A., Sisk, M. L., Alvarez, J. E., & Chawla, N. V. "Ventana a la Verdad (Window to the Truth): A Chatbot Application for Navigating The Colombian Truth Commission's Archives."In *Proceedings of the Eighteenth ACM International Conference on Web Search and Data Mining* (pp. 1052-1055), 2025.
26. Zhang, Z., Li, Y., Le, N.H.L., Wang, Z., Ma, T., Galassi, V., Murugesan, K., Moniz, N., Geyer, W., Chawla, N. V., Zhang, C., & Ye, Y. "NGQA: A Nutritional Graph Question Answering Benchmark for Personalized Health-aware Nutritional Reasoning." In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (ACL)* (pp. 5934-5966), 2025.
27. Ye, J., Wang, Y., Huang, Y., Chen, D., Zhang, Q., Moniz, N., Gao, T., Geyer, W., Huang, C., Chen, P.-Y., Chawla, N. V., & Zhang, X. "Justice or Prejudice? Quantifying Biases in LLM-as-a-Judge." In *Proceedings of the Thirteenth International Conference on Learning Representations (ICLR)*, 2025.
28. Pan, D., Moniz, N., & Chawla, N. V. "Fast Explanations via Policy Gradient-Optimized Explainer." In *Proceedings of the 34th International Joint Conference on Artificial Intelligence (IJCAI)* (pp. 475-483), 2025.
29. Karr, J. A., Smith, E., Hauenstein, M., Curto, G., & Chawla, N. V. "What is Behind Homelessness Bias? Using LLMs and NLP to Mitigate Homelessness by Acting on Social Stigma." In *Proceedings of the 34th International Joint Conference on Artificial Intelligence (IJCAI)*, 2025.
30. Guo, K., Shen, Y., Gonzalez-Montiel, G. A., Huang, Y., Zhou, Y., Surve, M., Guo, Z., Das, P., Chawla, N. V., Wiest, O., & Zhang, X. "Artificial Intelligence in Spectroscopy: Advancing Chemistry from Prediction To Generation and Beyond." In *Proceedings of the 34th International Joint Conference on Artificial Intelligence (IJCAI)* (pp. 10445-10454), 2025.
31. Wang, Z., Zhang, C., Li, J., Chawla, N. V., & Ye, Y. "Graph Foundation Models: Challenges, Methods, and Open Questions." In *31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 2* (pp. 6184-6194), 2025.
32. Guo, K., Liu, Z., Guo, Z., Nan, B., Isayev, O., Chawla, N. V., Wiest, O., & Zhang, X. "Proto-Yield: An Uncertainty-Aware Prototype Network for Yield Prediction in Real-world Chemical Reactions." In *Proceedings of the 34th ACM International Conference on Information and Knowledge Management (CIKM)* (pp. 791-801), 2025.

33. Gao, C., Huang, Y., Wang, X., Wu, S., Chawla, N. V., & Zhang, X. "Think it Image by Image: Multi-Image Moral Reasoning of Large Vision-Language Models." In *Proceedings of the 34th ACM International Conference on Information and Knowledge Management (CIKM)* (pp. 4722-4727), 2025.
34. Huang, Y., Chen, C., Cheng, L., Kailkhura, B., Chawla, N. V., & Zhang, X. "Socially Responsible and Trustworthy Generative Foundation Models: Principles, Challenges, and Practices." In *Proceedings of the 34th ACM International Conference on Information and Knowledge Management (CIKM)* (pp. 6825-6828), 2025.
35. Dong, K., Jonnalagedda, P., Gao, X., Acharya, A., Kissa, M., Flores, M., Chawla, N. V., & Das, K. "Transaction Categorization with Relational Deep Learning in QuickBooks." In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases (ECML/PKDD)* (pp. 143-160), 2025.
36. Sokol, A., Zhang, X., & Chawla, N. V. "WildInsight: a Chatbot for Wildlife Conservation Research." In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases (ECML/PKDD)* (pp. 511-515), 2025.
37. Khvatskii, G., Garcia Martinez, A., Pan, D., Belcher, M., Medrano Loera, G., Pineda Pérez, D., Ferrari Muñoz-Ledo, J. E., Márquez-González, H., Moniz, N., & Chawla, N. V. "Leveraging Artificial Intelligence to Bridge Gaps in Pediatric Oncology Care for Marginalized Spanish-Speaking Communities." In *Proceedings of the 34th International Joint Conference on Artificial Intelligence (IJCAI)* (pp. 9763-9771), 2025.
38. He, X., Tian, Y., Sun, Y., Chawla, N. V., Laurent, T., LeCun, Y., Bresson, X., & Hooi, B., "G-Retriever: Retrieval-Augmented Generation for Textual Graph Understanding and Question Answering," In Proceedings of the *Thirty-eighth Annual Conference on Neural Information Processing Systems (NeurIPS),* 2024.
39. Dong, K., Guo, Z., & Chawla, N. V., "Pure Message Passing Can Estimate Common Neighbor for Link Prediction," In Proceedings of the *Thirty-eighth Annual Conference on Neural Information Processing Systems (NeurIPS),* 2024.
40. Wang, Z., Zhang, Z., Chawla, N.V., Zhang, C., & Ye, Y., "Computation Tree: A Transferable Pattern Towards Graph Foundation Models," In Proceedings of the *Thirty-eighth Annual Conference on Neural Information Processing Systems (NeurIPS),* 2024.
41. Guo, K., Nan, B., Zhou, Y., Guo, T., Guo, Z., Surve, M., Liang, Z., Chawla, N. V., Wiest, O., & Zhang, Z., "Can LLMs Solve Molecule Puzzles? A Multimodal Benchmark for Molecular Structure Elucidation," In Proceedings of the *Thirty-eighth Annual Conference on Neural Information Processing Systems Datasets and Benchmark Track (NeurIPS),* 2024.
42. Schnur, J. J., Garcia-Martínez, A., Soga, P., Badillo-Urquiola, K., Botello, A. J., Calderon Raisbeck, A., ... & Chawla, N. V. "SaludConectaMX: Lessons Learned from Deploying a Cooperative Mobile Health System for Pediatric Cancer Care in Mexico," In *Companion Publication of the 2024 Conference on Computer-Supported Cooperative Work and Social Computing (CSCW)*, 2024.
43. Li, P., Huang, X., Tian, Y., & Chawla, N. V., "ChefFusion: Multimodal Foundation Model Integrating Recipe and Food Image Generation," In Proceedings of the *33rd ACM International Conference on Information and Knowledge Management (CIKM)*, 2024.

44. Huang, X., Surve, M., Liu, Y., Luo, T., Wiest, O., Wang, X., & Chawla, N.V., "Application of Large Language Models in Chemistry Reaction Data Extraction and Cleaning," In Proceedings of the *33rd ACM International Conference on Information and Knowledge Management (CIKM)*, 2024.
45. Dablain, D. A., & Chawla, N. V. "Data Augmentation's Effect on Machine Learning Models when Learning with Imbalanced Data," In Proceedings of the *2024 IEEE 11th International Conference on Data Science and Advanced Analytics*, 2024.
46. Ruz, G. A., & Chawla, N. V. "SMOTE for gene regulatory network sampling," In Proceedings of the *2024 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, 2024.
47. Yuan, X., Tian, Y., Zhang, C., Ye, Y., Chawla, N. V., & Zhang, C. "Graph Cross Supervised Learning via Generalized Knowledge," In Proceedings of the *30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, 2024.
48. Han, D., Moniz, N., & Chawla, N. V. "AnyLoss: Transforming Classification Metrics into Loss Functions." In Proceedings of the *30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, 2024.
49. Ren, X., Tang, J., Yin, D., Chawla, N., & Huang, C. "A survey of large language models for graphs," In Proceedings of the *30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, 2024.
50. Zhang, Z., Wang, Z., Hou, S., Hall, E., Bachman, L., Galassi, V., White, J., Chawla, N.V., Zhang, C. and Ye, Y., 2024. "Diet-ODIN: A Novel Framework for Opioid Misuse Detection with Interpretable Dietary Patterns." In Proceedings of the *30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, 2024.
51. Dablain, D. & Chawla, N. V. "The Hidden Influence of Latent Feature Magnitude When Learning with Imbalanced Data," In Proceedings of the *International Conference on Pattern Recognition and AI (ICPRAI),* 2024.
52. Krieg, S. J., Chawla, N. V., & Feldman, K. "Representing Outcome-driven Higher-order Dependencies in Graphs of Disease Trajectories." In *2024 IEEE 12th International Conference on Healthcare Informatics (ICHI)*, 2024. IEEE.
53. Wu, L., Tian, Y., Lin, H., Huang, Y., Li, S., Chawla, N. V., & Li, S. Z. "Learning to Predict Mutation Effects of Protein-Protein Interactions by Microenvironment-aware Hierarchical Prompt Learning," In Proceedings of the *41st International Conference on Machine Learning (ICML),* 2024.
54. Wan, G., Tian, Y., Huang, W., Chawla, N. V., & Ye, M. "S3GCL: Spectral, Swift, Spatial Graph Contrastive Learning," In Proceedings of the *41st International Conference on Machine Learning (ICML),* 2024.
55. Ma, Y., Yan, N., Li, J., Mortazavi, M. and Chawla, N.V., "HetGPT: Harnessing the Power of Prompt Tuning in Pre-Trained Heterogeneous Graph Neural Networks," In *Proceedings of the ACM on Web Conference,* 2024.
56. Liu, Z., He, X., Tian, Y., & Chawla, N. V. "Can we soft prompt LLMs for graph learning tasks?" In *Proceedings of the ACM on Web Conference,* 2024.

57. Ma, Y., Huang, X., Nan, B., Moniz, N., Zhang, X., Wiest, O., & Chawla, N. V. "Are we making much progress? Revisiting chemical reaction yield prediction from an imbalanced regression perspective." In *Proceedings of the ACM on Web Conference,* 2024.
58. Ma, C., Guo, T., Yang, Q., Chen, X., Gao, X., Liang, S., Chawla, N. and Zhang, X., "A Property-Guided Diffusion Model For Generating Molecular Graphs," In Proceedings of the *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 2365-2369), 2024.
59. Tian, Y., Song, H., Wang, Z., Wang, H., Hu, Z., Wang, F., Chawla, N.V. and Xu, P., March. "Graph neural prompting with large language models." In Proceedings of the *AAAI Conference on Artificial Intelligence* (Vol. 38, No. 17, pp. 19080-19088), 2024.
60. Chaudhry, B. M., Islam, M. U., & Chawla, N. V. "Longitudinal Evaluation of Casual Puzzle Tablet Games by Older Adults." In Proceedings of the *ACM Designing Interactive Systems Conference*, 2024.
61. Guo, T., Nan B., Liang Z., Zhichun Guo, Chawla N. V., Wiest O, and Zhang X. "What Can Large Language Models Do in Chemistry? A Comprehensive Benchmark on Eight Tasks." In Proceedings of the *37th Conference on Neural Information Processing Systems* (NeurIPS 2023), 2023.
62. Germino, J., Moniz, N., Chawla, N.V.: Fairness-Aware Mixture of Experts With Interpretability Budgets. *26th International Conference on Discovery Science, (DS 2023) pp 341–355*, 2023.
63. Akoglu, L., Chawla, N.V., Kumar, S., Nagrecha, S., Das, M., Naware, V.M., Faruquie, T.A. " KDD Workshop on Machine Learning in Finance." *29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining ( (KDD '23),* 2023.
64. Cheng, D.Z., Patel, D., Pang, L., Mehta, S., Xie, K., H, E., Liu, W., Chawla, N.V., Bailey, J. " Foundations and Applications in Large-scale AI Models: Pre-training, Fine-tuning, and Prompt-based Learning." *29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '23). pp 5853–5854,* 2023
65. Guo, Z., Shiao, W., Zhang, S., Liu, Y., Chawla, N.V., Shah, N., Zhao, T. "Linkless link prediction via relational distillation." *40th International Conference on Machine Learning, (PMLR 202:12012-12033),* 2023.
66. Zhang, Q., Pei, S., Yan, X., Zhang, C., Chawla, N.V., Zhang, X. "Cross-Domain Few-Shot Graph Classification with a Reinforced Task Coordinator." *AAAI Conference on Artificial Intelligence, 37(4), 4893-4901. 37, 4893–4901*, 2023.
67. Guo, Z.M., Zhang, C., Yu, F., Tian, Y., Zhang, C., Chawla, N.V. "Boosting graph neural networks via adaptive knowledge distillation." *AAAI Conference on Artificial Intelligence, 37(6), 7793-7801. 37, 7793–7801*, 2023.
68. Tian, Y., Dong, K., Zhang, C., Zhang, C., Chawla, N.V. "Heterogeneous graph masked autoencoders." *AAAI Conference on Artificial Intelligence, 37(8), 9997-10005. 37, 9997–10005,* 2023.

69. Krieg, S.J., Burgis, W.C., Soga, P.M., Chawla, N.V. "Deep Ensembles for Graphs with Higher-order Dependencies." *The Eleventh International Conference on Learning Representations (ICLR 2023),* 2023.
70. Dablain, D. A., Bellinger, C., Krawczyk, B., & Chawla, N. V. "Efficient augmentation for imbalanced deep learning." *IEEE 39th International Conference on Data Engineering (ICDE)* (pp. 1433-1446). 2023.
71. Dong, K., Tian, Y., Guo, Z., Yang, Y., & Chawla, N. V. "FakeEdge: Alleviate Dataset Shift in Link Prediction," *Learning on Graphs Conference* (pp. 56-1). *PMLR,* 2023.
72. Tian, Y., Zhang, C., Guo, Y, Zhang, X. and Chawla, N. V., "Learning MLPs on Graphs: A Unified View of Effectiveness, Robustness, and Efficiency," *International Conference on Learning Representation (ICLR)*, 2023.
73. Zhang, C., Tian, Y., Ju, M., Liu, Z., Ye, Y., Chawla, N. V., and Zhang, C. "Chasing All Round Graph Representation Robustness: Model, Training, and Optimization," *International Conference on Learning Representation (ICLR)*, 2023.
74. Zhang, Q., Pei, S., Yang, Q., Zhang, C., Chawla, N. V., and Zhang, X., "Cross-domain Few-shot Graph Classification with a Reinforced Task Coordinator," *AAAI,* 2023.
75. Guo, Z., Zhang, C., Fan, Y., Tian, Y., Zhang, C., and Chawla, N. V. "Boosting Graph Neural Networks via Adaptive Knowledge Distillation," *AAAI,* 2023.
76. Tian, Y., Dong, K., Zhang, C., and Chawla, N. V., "Heterogeneous Graph Masked Autoencoders," *AAAI,* 2023.
77. Qian, Y., Zhang, Y., Chawla, N. V., Ye, Y., and Zhang, C., "Malicious Repositories Detection with Adversarial Heterogeneous Graph Contrastive Learning," *31st ACM International Conference on Information and Knowledge Management (ACM CIKM),* 2022.
78. Zhang, C., Ding, K., Li, J., Zhang, X., Ye, Y., Chawla, N. V., Liu, H., "Few Shot Graph Learning: A Survey," *The 31st International Joint Conference on Artificial Intelligence (IJCAI) – Survey Track,* 2022.
79. Tian, Y., Zhang, C., Guo, Z., Huang, C., Metoyer, R., and Chawla, N. V., "RecipeRec: A Heterogeneous Graph Learning Model for Recipe Recommendation," *The 31st International Joint Conference on Artificial Intelligence (IJCAI),* 2022.
80. Tian, Y., Zhang, C., Guo, Z., M, Y. Metoyer, R., and Chawla, N. V., "Multi-modal Recipe Representation Learning with Graph Neural Networks,**"** *The 31st International Joint Conference on Artificial Intelligence (IJCAI),* 2022. (
81. Ma, Y., Gerard, P., Guo, Z., Tian, Y., Chawla, N. V., "Hierarchical Spatio-Temporal Graph Neural Networks for Pandemic Forecasting," *31st ACM International Conference on Information and Knowledge Management (ACM CIKM)*, 2022.
82. Cleland-Huang, J., Chawla, N., Cohen, M., Al Islam, M. N., Sinha, U., Spirkovska, L., ... & Chowdhury, M. T, "Towards Real-Time Safety Analysis of Small Unmanned Aerial Systems in the National Airspace." In *AIAA AVIATION 2022 Forum* (p. 3540).
83. Yuan, Z., Guo, Z., Chawla, N. V., Yang, T., "Compositional Training for end-to-end deep AUC maximization," *International Conference on Learning Representations (ICLR),* 2022.

84. Lin, S., Wu, X., & Chawla, N. V. "motif2vec: Semantic-aware Representation Learning for Wearables' Time Series Data." *IEEE 8th International Conference on Data Science and Advanced Analytics (DSAA)* 2021.
85. Tian, Y., Zhang, C., Metoyer, R., & Chawla, N. V. "Recipe representation learning with networks." In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*, 2021.
86. Guo, Z., Zhang, C., Yu, W., Herr, J., Wiest, O., Jiang, M., & Chawla, N. V. "Few-Shot Graph Learning for Molecular Property Prediction," *Proceedings of the Web Conference* 2021
87. Tiang, Y., Zhang, C., Metoyer, R., and Chawla, N. V., "Recipe Representation with Networks," *ACM Conference on Information and Knowledge Management (CIKM),* 2021.
88. Chaudhry, B., Dasgupta, D., Mohamed, M., & Chawla, N. "Teaching Tablet Technology to Older Adults." *International Conference on Human-Computer Interaction*, 2021.
89. Dasgupta, D., Chaudhry, B., & Chawla, N. "A Qualitative Usability Evaluation of Tablets and Accessibility Settings by Older Adults." *International Conference on Human-Computer Interaction*, 2021.
90. Dong, K., Lu, K., Xia, X., Cieslak, D., and Chawla, N. V., "An Optimized NL2SQL System for Enterprise Data Mart," *Joint European Conference on Machine Learning and Knowledge Discovery in Databases (ECML/PKDD)*, 2021.
91. Lin, S., Faust, L., & Chawla, N. V. "Lan: Learning to Augment Noise Tolerance for Self-report Survey Labels." *IEEE International Conference on Pervasive Computing and Communications (PerCom)*, 2021.
92. Wang, D., Zhao, T., Chawla, N. V., Jiang, M., "Dynamic Attributed Graph Prediction with Conditional Normalizing Flows," *IEEE International Conference on Data Mining (ICDM),* 2021.
93. Lin, S., Faust, L., D'Mello, S., Martinez, G., & Chawla, N. V. "MBead: Semi-supervised Multilabel Behaviour Anomaly Detection on Multivariate Temporal Sensory Data." *IEEE International Conference on Big Data (Big Data)*, 2020.
94. Syed, M., Wang, D., Jiang, M., Conway, O., Juneja, V., Subramanian, S., & Chawla, N. V. (2020, December). Overcoming Data Sparsity in Predicting User Characteristics from Behavior through Graph Embeddings. In *2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)* (pp. 32-36). IEEE.
95. Krieg, S. J., Robertson, D. H., Pradhan, M. P., & Chawla, N. V. (2020, November). Higher-order Networks of Diabetes Comorbidities: Disease Trajectories that Matter. In *2020 IEEE International Conference on Healthcare Informatics (ICHI)* (pp. 1-11). IEEE.
96. Schnur, J., Karl, R., Garcia-Martinez, A., Jiang, M., and Chawla, N.V., "Imputing Growth Snapshot Similarity in Early Childhood Development: A Tensor Decomposition Approach," *2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Seoul, Korea (South), 2020.

97. Krieg, S.J., Kogge, P.M. and Chawla, N.V., "GrowHON: A Scalable Algorithm for Growing Higher-order Networks of Sequences.," *International Conference on Complex Networks and Their Applications*, 2020.

98. Zhang C, Yu L, Saebi M, Jiang M, Chawla N., "Few-Shot Multi-Hop Relation Reasoning over Knowledge Bases*," In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP),* 2020.

99. Eliassi-Rad, T., Chawla, N., Colizza, V., Gardner, L., Salathé, M., Scarpino, S. and Wu, J.T., "Fighting a Pandemic: Convergence of Expertise, Data Science and Policy," In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2020.

100. Wang, D., Jiang, M., Syed, M., Conway, O., Juneja, V., Subramanian, S. and Chawla, N.V., "Calendar Graph Neural Networks for Modeling Time Structures in Spatiotemporal User Behaviors." In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 2020.

101. Tang, P., Jiang, M., Xia, B.N., Pitera, J.W., Welser, J. and Chawla, N.V., "Multi-label patent categorization with non-local attention-based graph convolutional network," In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020.

102. Wu X, Mattingly S, Mirjafari S, Huang C, Chawla NV., "Personalized Imputation on Wearable-Sensory Time Series via Knowledge Transfer," *In Proceedings of the 29th ACM International Conference on Information & Knowledge Management,* 2020.

103. Guo Z, Yu W, Zhang C, Jiang M, Chawla NV., "GraSeq: Graph and Sequence Fusion Learning for Molecular Property Prediction," *In Proceedings of the 29th ACM International Conference on Information & Knowledge Management,* 2020.

104. Ford, T., Jiminez, P., Metoyer, R. and Chawla, N. V., "Identifying Bridge Users: the Knowledge Transfer Agents in Enterprise Collaboration Systems" In *Proceedings of the 52nd Hawaii International Conference on System Sciences (HICSS)*, 2020.

105. Wu, X., Huang, C., Zhang, C., and Chawla, N. V. "Hierarchically Structured Transformer Networks for Fine-Grained Spatial Event Forecasting", *The Web Conference / World Wide Web Conference (WWW)*, 2020.

106. Wu, X., Cetintas, S., Kong, D., Miao, L., Yang, J., and Chawla, N. V., "Learning from Cross-Modal Behavior Dynamics with Graph-Regularized Neural Contextual Bandit," *The Web Conference / World Wide Web Conference (WWW)*, 2020.

107. Zhang, C., Yao, H., Huang, C., Jiang, M., Li, Z., and Chawla, N. V., "Few-shot Knowledge Graph Completion," *AAAI,* 2020.

108. Yao, H., Zhang, C., Wei, Y., Jiang, M., Wang, S., Huang, J., Chawla, N. V. and Li, Z., "Graph Few-shot Learning via Knowledge Transfer," *AAAI,* 2020.

109. Lin, S., Wu, X., Martinez, G., and Chawla, N. V., "Filling Missing Values on Wearable-sensory Time Series Data," *SIAM Conference on Data Mining (SDM),* 2020. (**Acc. Rate: 19.3%).**

110. Talkad Sukumar, P., Martinez, G., Grover, T., Mark, G., D'Mello, S., Chawla, N. V., Mattingly, S., and Striegel, A., "Characterizing Exploratory Behaviors on a Personal Visualization Interface Using Interaction Logs," *EuroVis,* 2020.

111. Huang, C., Wu, X., Zhang, X., Lin, S., and Chawla, N. V., "Deep Prototypical Networks for Imbalanced Time Series Classification under Data Scarcity," *Proceedings of the 28th ACM International Conference on Information and Knowledge Management* (CIKM), 2019.

112. Huang, C., Shi, B., Zhang, X., Wu, X., and Chawla, N. V, "Similarity-Aware Network Embedding with Self-Paced Learning." In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management* (*CIKM*), 2019.

113. Jiang, T., Zhao, T., Qin, B., Liu, T., Chawla, N., & Jiang, M. "Multi-Input Multi-Output Sequence Labeling for Joint Extraction of Fact and Condition Tuples from Scientific Text.," In *Proceedings of Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP),* 2019.

114. Zhang, C., Song, D., Huang, C., Swami, A., and Chawla, N. V., "HetGNN: Heterogeneous Graph Neural Network," *24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 2019.

115. Jiang, T., Zhao, T., Qin, B., Liu, T., Chawla, N. V., Jiang, M., "The Role of "Condition": A Novel Scientific Knowledge Graph Representation and Construction Model," *24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 2019.

116. Wang, D., Jiang, T., Chawla, N. V., Jiang, M., "TUBE: Embedding Behavior Outcomes for Predicting Success," *24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 2019.

117. Huang, C., Wu, X., Zhang, X., Zhang, C., Zhao, J., Yin, D., and Chawla, N. V., "Online Purchase Prediction via Multi-Scale Modeling of Behavior Dynamics," *24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 2019.

118. Faust, L., Jiménez-Pazmino, P., Holland, J. K., Lizardo, O., Hachen, D., & Chawla, N. V. "What 30 Days Tells Us About 3 Years: Identifying Early Signs of User Abandonment and Non-Adherence." In *Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare* (pp. 216-224). ACM, 2019.

119. Mattingly, S., *et al.,* "The Tesserae Project: Large-Scale, Longitudinal, In Situ, Multimodal Sensing of Information Workers," *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, 2019.

120. Mirjafari, S., *et al.*. "Differentiating Higher and Lower Job Performers in the Workplace Using Mobile Sensing." *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (ACM IMWUT)*, *3*(2), p.37, 2019.

121. Nwanganga, F. and Chawla, N. V., "Using Structural Similarity to Predict Future Workload Behavior in the Cloud," *IEEE Cloud, 2019.*

122. Markley, C., Feldman, K., and Chawla, N. V., "Outside the Hospital Walls: Associations of Value Based Care Metrics and Community Health Factors," *IEEE EMBS International Conference on Biomedical &Health Informatics (BHI),* 2019.

123. Chaudhry, B. M., Faust, L., & Chawla, N. V. "Development and Evaluation of a Web Application for Prenatal Care Coordinators in the United States," *International Conference on Design Science Research in Information Systems and Technology* (pp. 140-156). Springer, Cham, 2019.

124. Saha, K., *et al.,* "Imputing Missing Social Media Data Stream in Multisensor Studies of Human Behavior," *8th International Conference on Affective Computing and Intelligent Interaction (ACII)*, 2019.

125. Syed, M., Anggara, T., Lanski, A., Duan, X., Ambrose, A., and Chawla, N. V., "Integrated Closed-loop Learning Analytics Scheme in a First Year Experience Course," *9th International Conference on Learning Analytics and Knowledge (LAK),* 2019.

126. Syed, M., Chetlur, M., Afzal, S., Ambrose, A., and Chawla, N. V., "Implicit and explicit emotions in MOOCs," *Educational Data Mining (EDM),* 2019.

127. Huang, C., Zhang, C., Zhao, J., Wu, X., Yin, D., & Chawla, N. V., "MiST: A Multiview and Multimodal Spatial-Temporal Learning Framework for Citywide Abnormal Event Forecasting," *Proceedings of the Web Conference (WWW),* 2019. *Selected for Best Papers Track.*

128. Wu, X., Shi, B., Dong, Y., Huang, C., & Chawla, N. V. "Neural Tensor Factorization for Temporal Interaction Learning." *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining* (*WSDM)*, 2019

129. Zhang, C., Swami, A., & Chawla, N. V. "SHNE: Representation Learning for Semantic-Associated Heterogeneous Networks." *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining* (*WSDM)*, 2019

130. Zhang, C., Song, D., Chen, Y., Feng, X., Lumezanu, C., Cheng, W., Ni, J., Zong, B., Chen, H. and Chawla, N. V., "A Deep Neural Network for Unsupervised Anomaly Detection and Diagnosis in Multivariate Time Series Data", *Proceedings of AAAI International Conference on Artificial Intelligence (AAAI)*, 2019.

131. Nwanganga, F., Chawla, N. V., & Madey, G., "Statistical Analysis and Modeling of Heterogeneous Workloads on Amazon's Public Cloud Infrastructure." In *Proceedings of the 52nd Hawaii International Conference on System Sciences (HICSS)*, 2018.

132. Wu, X., Shi, B., Dong, Y., Huang, C., Faust, L., and Chawla, N. V. "RESTFul: Resolution-Aware Forecasting of Behavioral Time Series Data," *ACM Conference on Information and Knowledge Management (CIKM),* 2018.

133. Huang, C., Zhang, J., Zheng, Y., and Chawla, N. V. "DeepCrime: Attentive Hierarchical Recurrent Networks for Crime Prediction," *ACM Conference on Information and Knowledge Management (CIKM),* 2018.

134. Zhang, C., Yu, L., Zhang, X., and Chawla, N. V. "Task Guided and semantic-aware Ranking for academic author-paper correlation inference,' *IJCAI,* 2018.

135. Wang, D., Jiang, M., Zeng, Q., Eberhart, Z., and Chawla, N. V. "Multi-Type itemset embedding for learning behavior success," *23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 2018.

136. Zhang, C., Huang, C., Yu, L., Zhang, X., and Chawla, N. V., "CAMEL: Content Aware and Meta-path Augmented Metric Learning for author identification," *International Conference on World Wide Web (WWW),* 2018.

137. Moniz, N., Ribeiro, R., Cerqueira, V. and Chawla, N., V. "SMOTEBoost for Regression: Improving the Prediction of Extreme Values." *IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA)*, 2018.

138. Wu, X., Dong, Y., Shi, B. Swami, A., Chawla, N. V., "Who will attend this event together? Event attendance prediction via deep LSTM Networks." *SIAM Conference on Data Mining (SDM)*, 2018.

139. Nigam, A., Shin, K., Bahulkar, A., Hooi, B., Hachen, D., Szymanski, B., Faloutsos, C. and Chawla, N. V. "ONE-M: Modeling the Co-evolution of Opinions and Network Connections." In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases (ECML/PKDD)*, 2018.

140. Faust, L., Hachen, D., Lizardo, O., and Chawla, N. V., "Quantifying subjective well-being trends in weekend activity," *IEEE International Conference on Health Informatics,* 2018.

141. Feldman, K., Duarte, M., Carrasco, W., and Chawla, N. V. "Leveraging Health and Wellness Platforms to Understand Childhood Obesity: A Usability Pilot of FitSpace," *IEEE International Conference on Biomedical and Health Informatics,* 2018.

142. Chaudhry, B., Faust, K. and Chawla, N. V., "Towards an integrated platform for community-based maternity health workers," *EAI International Conference on Pervasive Computing Technologies for Healthcare,* 2018.

143. Zhi, Q., Lin, S., He, S., Metoyer, R., and Chawla, N. V., "VisPod: Content-Based Audio Visual Navigation," *Proceedings of the 23rd International Conference on Intelligent User Interfaces,* 2018.

144. Wu, X., Dong, Y., Tao, J., Huang, C., Chawla, N. V., "Reliable fake review detection via modeling temporal and behavioral patterns," *IEEE International Conference on Big Data,* 2017.

145. Zhang, C., Yu, L., Zhang, X., Chawla, N. V., "ImWalkMF: Joint Matrix Factorization and Implicit Walk Integrative Learning for Recommendation," *IEEE International Conference on Big Data,* 2018.

146. Tang, P., Pitera, J., Zubarev, D., and Chawla, N. V. "Materials Science Literature-Patent Relevance Search: A Heterogeneous Network Analysis Approach", In *IEEE International Conference on Data Science and Advanced Analytics*, 2017. Best Application Paper Award.

147. Nagrecha, S., Thomas, P. B., Feldman, K., & Chawla, N. V. "Predicting Chronic Heart Failure Using Diagnoses Graphs." In *International Cross-Domain Conference for Machine Learning and Knowledge Extraction*, Springer, 2017.

148. Dong, Y., Johnson, R. A., Xu, J., & Chawla, N. V. "Structural Diversity and Homophily: A Study Across More Than One Hundred Big Networks." In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, ACM, 2017.

149. Dong, Y., Chawla, N. V., & Swami, A. "metapath2vec: Scalable Representation Learning for Heterogeneous Networks." In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 2017.

150. Nwanganga, F., Saebi, M., Madey, G. and Chawla, N.V. "A Minimum-Cost Flow Model for Workload Optimization on Cloud Infrastructure." In *IEEE Cloud Computing (CLOUD),* 2017.

151. Xu, J. and Chawla, N. V. "Mining Features Associated with Effective Tweets." In *Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017 (ASONAM),* 2017.

152. Xu, J., Tao, J., Chawla, N. V., & Wang, C. "Visual Analytics of Higher-order Dependencies in Sensor Data." In *IEEE/ACM Internet-of-Things Design and Implementation (IoTDI),IEEE/ACM Second International Conference,* 2017.

153. Wu, X., Dong, Y., Huang, C., Xu, J., Wang, D. and Chawla, N.V. "UAPD: Predicting Urban Anomalies from Spatial-Temporal Data." In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases (EMCL/PKDD),* 2017.

154. Nagrecha, S., Dillon, J. Z., & Chawla, N. V. "MOOC Dropout Prediction: Lessons Learned from Making Pipelines Interpretable." In *Proceedings of the 26th International Conference on World Wide Web* (*WWW*), 2017.

155. Tao, J., Xu, J., Wang, C. and Chawla, N.V. "HoNVis: Visualizing and exploring higher-order networks." In *Pacific Visualization Symposium (PacificVis),* 2017.

156. Afzal, S., Sengupta, B., Syed, M., Chawla, N., Ambrose, G.A. and Chetlur, M. "The ABC of MOOCs: Affect and its inter-play with behavior and cognition." *Seventh International Conference on Affective Computing and Intelligent Interaction (ACII),* 2017.

157. Nigam, A., Aguinaga, A., and Chawla, N. V., "Connecting the dots to infer followers' topical interest on Twitter," *International Conference on Behavioral, Economic, and Social Modeling,* 2016.

158. Bahulkar, A., Szymanski, B., Lizardo, O., Dong, Y., Yang, Y., and Chawla, N. V., "Analysis of link formation, persistence and dissolution in NetSense data," *IEEE/ACM International Conference on Advances in Social Network Analysis and Mining (ASONAM)*, 2016.

159. Pal, S., Dong, Y., Thapa, B., Chawla, N. V., Swami, A., and Ramanathan, R., "Deep learning for network analysis: problems, approaches, and challenges," *MILCOM,* 2016.

160. Dasgupta, D. and Chawla, N. V., "Medcare: Leveraging medication similarity for disease prediction," *IEEE International Conference on Data Science and Advanced Analytics (DSAA),* 2016.

161. Feldman, K. Hazecamp, N., and Chawla, N. V., "Mining the clinical narrative: all text are not equal," *IEEE International Conference on Health Informatics (ICHI),* 2016.

162. Dasgupta, D. Reeves, K., Chaudhry, B., Duarte, M., and Chawla, N. V., "eSeniorCare: Technology for promoting well-being of older adults in independent living facilities," *IEEE International Conference on Health Informatics (ICHI),* 2016**.**

163. Dasgupta, D., Reeves, K., Chaudhry, B. and Chawla, N. V., "Design and evaluation of a medication adherence application with communication for seniors in independent living communities," *AMIA Annual Symposium,* 2016.

164. Chaudhry, B., Greeves, K., and Chawla, N. V., "Successful Aging for Low-Income Older Adults: Towards Design Principles," *10th EAI International Conference on Pervasive Computing Technologies for Healthcare,* 2016.

165. Nigam, A., and Chawla, N. V., Link "Prediction in a Semi-bipartite Network for Recommendation." In *Conference on Intelligent Information and Database Systems* (pp. 127-135). Springer Berlin Heidelberg, 2016.

166. Dong, Y., Johnson, R., and Chawla, N. V., "Collaboration Signatures Reveal Scientific Impact," *ACM/IEEE International Conference on Advances in Social Network Analysis and Modeling (ASONAM),* 2015.

167. Nagrecha, S., Chawla, N. V., Bunke, H., "Recurrent Subgraph Prediction", *ACM/IEEE International Conference on Advances in Social Network Analysis and Modeling (ASONAM)*, 2015.

168. Nagrecha, S., Aguiar, E. and Chawla, N. V., "Predicting Online Video Engagement Using Clickstreams," *IEEE International Conference on Data Science and Advanced Analytics (IEEE DSAA),* 2015.

169. Dong, Y., Pinelli, F., Gkoufas, Y., Nabi, Z., Calabrese, F., and Chawla, N.V., "Inferring Unusual Crowd Events from Mobile Phone Call Data Records*," European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD)*, 2015. Huang, C., Wang, D. and Chawla, N. V., "Towards Time-sensitive Truth Discovery in Social Sensing Applications," 12th *IEEE International Conference on Mobile Ad-hoc and Sensor Systems* (IEEE MASS), 2015.

170. Dong, Y., Johnson, R., and Chawla, N. V., "Will This Paper Increase your H-index?" *European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD)*, 2015.

171. Dong, Y., Chawla, N. V., Tang, J., and Yang, Y. "The Evolution of Social Relationships and Strategies Across the Lifespan," *European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD)*, 2015.

172. Dong, Y., Zhang, J., Tang, J., Chawla, N. V., Wang, B., "CoupledLP: Link Prediction in Coupled Networks," *21st ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD),* 2015.

173. Huang, C., Wang, D. and Chawla, N. V., "On spatial-temporal truth finding in social sensing," *12th IEEE Conference on Sensing, Communication, and Networking (SECON*), 2015.

174. Nwanganga, F., Aguiar, E., Ambrose, A., Goodrich, V., and Chawla, N. V., "Qualitatively exploring electronic portfolios: a text mining approach to measuring student emotion as an early warning indicator," *5th ACM International Conference on Learning Analytics and Knowledge*, 2015.

175. Johnson, R., Raeder, T., and Chawla, N. V., "Optimizing Classifiers for Hypothetical Scenarios," *19th Pacific Asia Conference on Knowledge Discovery and Data Mining (PAKDD)*, 2015.

176. Dong, Y., Johnson, R., and Chawla N. V.,"Will This Paper Increase Your h-index? Scientific Impact Prediction," *8th ACM International Conference on Web Search and Data Mining (WSDM),* 2015. Best Paper Nominee.

177. Feldman, K. and Chawla, N. V., "Admission Duration Model for Infant Treatment (ADMIT)," *IEEE International Conference on Bioinformatics and Biomedicine*, 2014.

178. Xu, J., Wickramarathne, T. L., Chawla, N. V., Grey, E., Steinhaeuser, K., Keller, R. P., Drake, J. M., Lodge, D. M., "Improving Management of Aquatic Invasions by Integrating Shipping Network, Ecological and Environmental Data: Data Mining for Social Good," *20th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD),* 2014

179. Dong, Y., Yang, Y., Tang, J., and Yang, Y., Chawla, N. V., "Inferring User Demographics and Social Strategies in Mobile Social Networks," *20th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD),* 2014.

180. Goodrich, V., Aguiar, E., Ambrose, G. A., McWilliams, L. H., Brockman, J., Chawla, N. V., "Integration of ePortfolios in a First Year Engineering Course for Measuring Student Engagement," 121st ASEE Annual Conference and Exposition (ASEE'14). **Best paper runner-up.**

181. Feldman, K. and Chawla, N. V., "Scaling Personalized Healthcare with Big Data," *2nd International Conference on Big Data Analytics and Healthcare*, 2014. **Best paper nominee.**

182. Dasgupta, D. and Chawla, N. V., "Disease and Medication Networks: An Insight into Disease-Drug Interactions," *2nd International Conference on Big Data Analytics and Healthcare*, 2014.

183. Dal Pozzolo, A., Johnson, R. A., Caelen, O., Waaterschoot, S., Chawla, N. V., and Bontempi, G., "HDDT to avoid instance propagation in unbalanced and evolving data streams," *IEEE International Joint Conference on Neural Networks (IEEE IJCNN),* 2014.

184. Dasgupta, D., Feldman, K., Waghray, D., Mikels-Carrasco, W., Willaert, P., Raybold, D., and Chawla, N. V., "Integrated Digital Care Framework for Successful Aging," *IEEE International Conference on Biomedical and Health Informatics,* 2014.

185. Aguiar, E., Chawla, N. V., Brockman, J., Ambrose, G. A., Goodrich, V., "Engagement vs Performance: Using Electronic Portfolios to Predict First Semester Engineering Student Retention," *Fourth International Conference on Learning Analytics and Knowledge (LAK),* 2014.

186. Pandit, S., Yang, Y., Chawla, N.V., and Uzzi, B., "Red Black Network: Temporal and Topological Analysis of Two Intertwined Social Networks," *Military Communications Conference (MILCOM),* 2013.

187. Rider, A. and Chawla, N. V., "An Ensemble Topic Model for Sharing Healthcare Data and Predicting Disease Risk," *ACM International Conference on Bioinformatics, Computational Biology and Biomedical Informatics (ACM BCB),* 2013.

188. Rider, A., Johnson, R. A., Davis, D. A., Hoens, T. R., and Chawla, N. V., "Classifier Evaluation with Missing Negative Class Labels," *Twelfth International Symposium on Intelligent Data Analysis (IDA),* 2013.

189. Yang, Y., Chawla, N. V., Basu, P., Prabhala, B., and LaPorta, T., "Link Prediction in Human Mobility Networks," *ACM/IEEE ASONAM,* 2013.

190. Dong, Y., Tang, J., Lou, T., Wu, B. and Chawla, N. V., "How long will she call me? Distribution, Social Theory and Duration Prediction," *Twenty Third European Conference on Machine Learning and Practice of Knowledge Discovery in Data Bases (ECML/PKDD),* 2013.

191. Yang, Y., Chawla, N. V., Adali, S., "Prominence in Networks: A co-evolving process," *IEEE Network Science Workshop,* 2013.

192. Thompson, R. and Chawla, N. V., "Addressing Challenges in Prescription Management," *24th Annual Conference of the Production and Operations Management Society,* 2013.

193. Dong, Y., Tang, J., Wu, S., Tian, J., Chawla, N. V. Rao, J., Cao, H. "Link Prediction and Recommendation Across Heterogeneous Social Networks," *IEEE International Conference on Data Mining (ICDM),* 2012.

194. Yang, Y., Chawla, N. V., Sun, Y., Han, J. "Predicting Links in Multi-Relational and Heterogeneous Networks," *IEEE International Conference on Data Mining (ICDM),* 2012.

195. Johnson, R. A., Chawla, N. V., Hellmann, J. J., "Species Distribution Modeling and Prediction: A class imbalance problem," *IEEE Conference on Intelligent Data Understanding,* 2012.

196. Lichtenwalter, R. and Chawla, N. V., "Link Prediction: Fair and Effective Evaluation," *IEEE/ACM International Conference on Social Networks Analysis and Mining,* 2012.

197. Hoens, T. R. and Chawla, N. V., "Learning in non-stationary environments with class imbalance," *ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD),* 2012.

198. Regola, N., Cieslak, D. and Chawla, N. V., "Constraints of magnetic versus flash disk capabilities in big data analysis," *Architectures and Systems for Big Data,* 2012.

199. Sun, Y., Han, J., Aggarwal, C., and Chawla, N. V., "When will it happen? Relationship prediction in heterogeneous information networks," *ACM International Conference on Web Search and Data Mining (WSDM),* 2012.

200. Lichtenwalter, R. and Chawla, N. V., "Vertex Collocation Profiles: Subgraph Counting for Link Analysis and Prediction," *International Conference on World Wide Web (WWW),* 2012,

201. Hoens, T. R., Qen, Q., Chawla, N. V., Zhou Z., "Building Decision Trees for Multi-class Imbalanced Data Learning," *PAKDD,* 2012.

202. Hoens, T. R., Chawla, N. V., Polikar, R., "Heuristic Updatable Weighted Random Subspaces for Nonstationary Environments," *IEEE International Conference on Data Mining (ICDM),* 2011.

203. Mooney, S., Li, B., Davis, D. A., Chawla, N. V., Morris D., Anderson J., "The link between Recent Parkinson's Disease Diagnoses and Bipolar Disorder Diagnoses in an Aging Population in the United States," *The Translational Bioinformatics Conference,* 2011.

204. Lussier, J. and Chawla, N.V., "Network Effects of Tweeting," *14th International Conference of Discovery Science,* 2011.

205. Pandit, S., Yang, Y., Kawadia, V., Srinivasan, S., Chawla, N. V., "Detecting communities in time-evolving proximity networks," *IEEE Network Science Workshop,* 2011.

206. Steinhaeuser, K., Chawla, N. V., Ganguly, A., "Comparing predictive power in climate data: Clustering Matters," *Symposium on Spatial and Temporal Databases,* 2011.

207. Lichtenwalter, R. N. and Chawla, N. V., "DisNet: A framework for distributed graph computation," *International Conference on Advances in Social Networks Analysis and Mining (ASONAM),* 2011.

208. Davis, D. A., Lichtenwalter, R. N., Chawla, N. V., "Multi-relational link prediction in heterogeneous information networks," *International Conference on Advances in Social Networks Analysis and Mining (ASONAM),* 2011.

209. Yang, Y., Sun, Y., Pandit, S., Chawla, N. V., Han, J., "Is objective function the silver bullet? A case study of community detection algorithms on social networks," *International Conference on Advances in Social Networks Analysis and Mining (ASONAM),* 2011. (short paper**)**

210. Pelan, A., Steinhaeuser, K. Chawla, N. V., Pitts, D., Ganguly, A. R., "Empirical comparison of correlation measures and pruning levels in complex networks for representing the global climate system," *IEEE Symposium on Computational Intelligence and Data Mining,* 2011.

211. Raeder, T., Hoens, T. R., Chawla, N. V., "Consequences of Variability in Classifier Performance Estimates," *IEEE International Conference on Data Mining (ICDM),* 2010.

212. Rider, A., Siwo, G., Emrich, S., Ferdig, M., Chawla, N. V., "A Supervised Learning Approach to the Unsupervised Clustering of Genes," *IEEE International Conference on Bioinformatics and Biomedicine (BIBM),* 2010.

213. Lussier, J., Raeder, T., Chawla, N. V., "Digging the Dirt on User Generated Content Consumption," *16th International Conference on Management of Data (COMAD)*, 2010.

214. Hoens, T. R., Blanton, M., Chawla, N. V., "Reliable Medical Recommendation Systems with Patient Privacy, *ACM International Health Informatics Symposium,* November 2010.

215. Steinhaeuser, K., Chawla, N. V., Ganguly, A. R., "Complex Networks in Climate Science: Progress, Opportunities," *NASA Conference on Intelligent Data Understanding (CIDU),* October 2010.

216. Ditzler, G., Chawla, N. V., Polikar, R. "An Incremental Learning Algorithm for Nonstationary Environments and Class Imbalance," *Int. Conference on Pattern Recognition (ICPR),* 2010.

217. Lichtenwalter, R., Lussier, J., Chawla, N. V., "New Perspectives and Methods in Link Prediction," *ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD),* 2010, pages: 243 – 252.

218. Davis, D. A. and Chawla, N. V., "Exploring Disease Interactions Using Combined Gene and Phenotype Networks," *International Conference on Intelligent Systems for Molecular Biology (ISMB),* 2010.

219. Hoens, T. R., Blanton, M., Chawla, N. V., "A Private and Reliable Recommendation System for Social Networks," *IEEE International Conference on Information Privacy, Security, Risk and Trust (PASSAT),* 2010.

220. Liao, Q., Striegel, A., Chawla, N. V., "Visualizing Dynamics and Similarity in Enterprise Networks," *International Symposium on Visualization for Cyber-Security,* 2010.

221. Hoens, T. R. and Chawla, N. V., "Generating Diverse Ensembles to Counter the Problem of Class Imbalance," *" Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD)*, 2010.

222. Raeder, T., Blanton, M., Chawla, N. V., Frikken, K**.,** "Privacy-Privacy-Preserving Network Aggregation," *" Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD)* 2010**.**

223. Liu, W., Chawla, S., Cieslak, D., Chawla, N. V., "A robust decision tree algorithm for imbalanced data sets," *SIAM Conference on Data Mining (SDM),* (**Acc: 23.36%**), 2010.

224. Lussier, J., Raeder, T., Chawla, N. V., "User Generated Content Consumption and Social Networking in Knowledge-Sharing OSN", *International Conference on Modeling of Social Behavior and Prediction (SBP),* 2010.

225. Lichtenwalter, R., Lichtenwalter, K., Chawla, N. V., "Applying Learning Algorithms to Music Generation," *International Conference on Artificial Intelligence,* December 2009.

226. Marcos, F., Lamanna, C., Chawla, N. V., Izaguirre, J., "Determination of Specificity Residues in Two Component Systems using Graphlets," *International Conference on Bioinformatics & Computational Biology,* August 2009. **(Acc: 27%)**

227. Raeder, T. and Chawla, N. V., "Modeling the Product Space as a Social Network," *ACM/IEEE Conference on Advances in Social Network Analysis and Modeling (ASONAM),* pages: 143 – 152, July 2009. **In the Best of ASONAM, nominated for a journal publication.**

228. Lichtenwalter, R. and Chawla, N. V., "Learning to classify data streams with imbalanced class distributions," *Proceedings of PAKDD Workshops,* 2009.

229. Cieslak, D. and Chawla, N. V., "Start Globally, Optimize Locally, Predict Globally: Improving Performance on Unbalanced Data." *IEEE International Conference on Data Mining (ICDM),* pages: 143 – 152, December 2008.

230. Moretti, C., Steinhaeuser, K., Thain, D., Chawla, N. V., "Scaling Up Classifiers to Cloud Computers*." IEEE International Conference on Data Mining (ICDM),* pages: 472 – 481, December 2008.

231. Davis, D. A., Chawla, N. V., Blumm, N., Christakis, N. and Barabasi, A-L., “"Predicting individual disease risk based on medical history" *ACM Conference on Information and Knowledge Management (CIKM)*, pages: 769 – 778, October 2008.

232. Cieslak, D., Chawla, N. V., Thain, D., “"Troubleshooting Thousands of Jobs on Production Grids Using Data Mining Techniques" *IEEE/ACM International Conference on Grid Computing (GRID)*, pages: 217 – 224, July 2008.

233. Cieslak, D. and Chawla, N. V.,” "Learning Decision Trees for Unbalanced Data" *European Conference on Machine Learning (ECML),* pages: 241 – 256, September 2008.

234. Cieslak, D. and Chawla, N. V., "Analyzing Classifier Performance on Imbalanced Datasets when Training and Testing Distributions Differ*" Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD)*, pages: 519 – 526, May 2008.

235. Steinhaeuser, K. and Chawla, N. V., “Scalable Learning with Thread-Level Parallelism.” *Midwest Artificial Intelligence and Cognitive Science Conference (MAICS)*, August 2008, 8 pages.

236. Cieslak, D. and Chawla, N. V., “Detecting fracture points in classifier performance,” *IEEE International Conference on Data Mining (ICDM),* October 2007.

237. Thomas, V., Chawla, N. V., Bowyer, K. W., Flynn, P., “Learning to predict gender from irises,” *IEEE Biometrics: Theory, Application and Systems (BTAS)*, pages: 1 – 5, September 2007.

238. Madey, G., Barabasi, A.-L., Chawla, N. V., et al., “Enhancing Situational Awareness: Application of DDDAS Concepts to Emergency and Disaster Management,” *International Conference on Computational Science,* pages: 1090 – 1097.

239. Chawla, N. V. and Bowyer, K. W., “Actively Exploring Creation of Face Spaces for Improved Face Recognition,” *Advancement of Artificial Intelligence (AAAI),* pages: 809 – 814, 2007.

240. Chapple, M., Chawla, N. V., Striegel, A. “Authentication Anomaly Detection: A case study on VPN.” *Proceedings of ACM MineNet*, pages: 17 – 22, 2007.

241. Chawla, N. V. and Sylvester, J. “Exploiting diversity in ensembles: Improving performance on unbalanced datasets.” *Proceedings of Multiple Classifier systems (MCS),* pages: 397 – 406, 207.

242. Malik, T., Burns, R., Chawla, N. V., “A Black—Box approach to Query Cardinality Estimation.” *ACM Conference on Innovative Data Systems Research (CIDR),* pages: 56 – 67, 2007.

243. Malik, T., Burns, R., Chawla, N. V., Szalay, A., “Estimating Query Result Sizes for Proxy Caching in Scientific Database Federations”, *ACM/IEEE Supercomputing,* pages: 102 – 112, 2006.

244. Pawling, A., Chawla, N. V., Chaudhary, A., “Evaluation of Summarization Schemes for Learning in Streams,” *Proceedings of 10th European Conference on Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD),* pages: 347 – 358, 2006.

245. Pawling, A., Chawla, N. V., Madey, G., "Anomaly Detection in a Mobile Communication Network", *Annual Conference of the North American Association for Computational Social and Organizational Science (NAACSOS),* June 2006.

246. Cieslak, D., Thain, D., Chawla, N. V., "Troubleshooting Distributed Systems via Data Mining", Proceedings of *Hot Topics Sessions: 15th IEEE International Symposium on High Performance Distributed Computing (HPDC-15),* pages: 309 – 312, June 2006.

247. Sylvester, J., Chawla, N. V., "Evolutionary Ensemble Creation and Thinning," *Proceedings of IEEE International Joint Conference on Neural Networks (IJCNN)*, pages: 5148 – 5155, July 2006.

248. Cieslak, D., Chawla, N. V., Striegel, A., "Combating imbalance in network intrusion datasets," *Proceedings of IEEE Conference on Granular Computing (GrC)*, pages: 732 – 737, May 2006.

249. Chawla, N. V. and Bowyer, K. W., "Ensembles in face recognition: Tackling the extremes of high dimensionality, temporality, and variance in data." *Proceedings of IEEE Conference on Systems, Man and Cybernetics (SMC),* pages: 2346 – 2351, October 2005.

250. Chawla, N. V., "Teaching Data Mining by Coalescing Theory and Applications," *Proceedings of Frontiers in Education (FIE),* October 2005.

251. Mack, D., Chawla, N. V., Madey, G., "Activity Mining in Open Source Software," *Proceedings of NAACSOS,* June 2005.

252. Chawla, N. V. and Bowyer, K. W., "Random subspaces and subsampling for face recognition", *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* pages: 582 – 589, June, 2005. .

253. Chawla, N. V. and Bowyer, K. W., "Designing multiple classifier systems for face recognition", *Proceedings of Multiple Classifier Systems (MCS),* pages: 407 – 416, 2005.

254. Chawla, N. V., Lazarevic, A., Hall, L. O., Bowyer, K. W. "SMOTEBoost: Improving the prediction of the minority class in Boosting," *Proceedings of 7th European Conference on Principles and Practices of Knowledge Discovery in Databases (ECML/PKDD),* pages, 107 – 119, September 2003.

255. Chawla, N. V., Hall, L. O., Bowyer, K. W., Moore, T. E., Kegelmeyer, W. P., "Distributed Pasting of Small Votes," *Proceedings of Multiple Classifier Systems (MCS)*, pages: 52 – 61, June 2002.

256. Chawla, N. V., Moore, T. E., Hall, L. O., Bowyer, K.W., Kegelmeyer, W. P., Springer C., "Bagging is a Small-Data-Set Phenomenon," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages: 684 – 689, December 2001.

257. Chawla, N. V., Eschrich, S., Hall, L. O., "Creating Ensembles of Classifiers," *Proceedings of IEEE International Conference on Data Mining (ICDM)*, November 2001.

258. Chawla, N. V., Bowyer, K. W., Hall, L. O., Kegelmeyer, W. P., "SMOTE: Synthetic Minority Over-sampling Technique", *Proceedings of International Conference on Knowledge Based Computer Systems (KBCS)*, Mumbai, India, December 2000.

259. Bowyer, K. W., Hall, L. O., Moore, T. E., Chawla, N. V., Kegelmeyer, W. P., "A Parallel Decision Tree Builder for Mining Very Large Visualization Data sets," *Proceedings of IEEE Conference on Systems, Man, and Cybernetics (SMC)*, October 2000.

260. Hall, L. O., Chawla, N. V., Bowyer, K. W., "Decision Tree Learning on Very Large Data sets", *Proceedings of IEEE Conference on Systems, Man, and Cybernetics (SMC)*, July 1998

**Refereed full papers at Workshops**

1. Faust, L., Purta, R., Hachen, D., Striegel, A., Poellabauer, C., Lizardo, O., & Chawla, N. V. Exploring Compliance: Observations from a Large Scale Fitbit Study. In *Proceedings of the 2nd International Workshop on Social Sensing*, 2017.
2. Golestanian, M., Poellabauer, C. and Chawla, N., "RSSI-Based Pedestrian Localization Using Artificial Neural Networks." In *Proceedings of the 2nd ACM International Workshop on Smart, Autonomous, and Connected Vehicular Systems and Services*, 2017.
3. S. Pandit, Y. Yang, Chawla, N. V., "Maximizing Information Spread Through Influence Structures in Social Networks*," IEEE ICDM Workshop on Data Mining in Networks*, 2012.
4. Johnson, R. A., Yang, Y., Rider, A., Chawla, N. V., "ALIVE: A Multi-Relational Link Prediction Environment for the Healthcare Domain," *PAKDD Medical Data Mining,* 2012.
5. Steinhaeuser, K., Chawla, N. V. and Ganguly, A., "An Exploration of Climate Data Using Complex Networks," *ACM SIGKDD Workshop on Knowledge Discovery from Sensor Data,* 2009.
6. Notwell, J., McRoskey, S., Chawla, N. V. and Poellabauer, C., "Mining in a Mobile Environment," *ACM SIGKDD Workshop on Knowledge Discovery from Sensor Data,* 2009. (**Best Student Paper Award)**
7. Lichtenwalter, R. and Chawla, N. V., "Adaptive Methods for Classification in Arbitrarily Imbalanced and Drifting Data Streams," *PAKDD Workshop Series*, 2009.
8. Davis, D. A. and Chawla, N. V., "CARE For Your Future: Prospective Disease Prediction Using Collaborative Filtering". *Proceedings of the KDD 2008 Workshop on Mining Medical Data*, 2008.
9. Rajan, D., Poellabauer, C., Chawla, N. V., "Resource Access Pattern Mining for Dynamic Energy Management**,"** *Proceedings of the Workshop on Autonomic Computing: A New Challenge for Machine Learning*, *ECML/PKDD,* September 2006.
10. Steinhaueser, K., Chawla, N. V., Kogge, P., "Exploiting Thread-level Parallelism to Build Decision Trees.**"** *Proceedings of the Workshop on Distributed and Parallel Data Mining,*, *ECML/PKDD,* September 2006.
11. Chawla, N. V. and Li, X., "Pricing scheme for benefit scoring," *Proceedings of ACM SIGKDD Workshop on Utility Based Data Mining*.
12. Chawla, N. V. and Cieslak, D., "Evaluating Calibration of Probability Estimation Trees", *Proceedings of the AAAI Workshop on the Evaluation Methods in Machine Learning,* July 2006.

13. Steinhaueser, K., Chawla, N. V., Poellabauer, C., "Towards Learning-based Sensor Management," *Proceedings of the First Workshop on Tackling Computer Systems Problems with Machine Learning Techniques, SIGMETRICS,* June 2006.
14. Pawling, A., Chawla, N. V., Chaudhary, A., "Information Gain computation on data streams", *Proceedings of IEEE ICDM Workshop on Temporal Data Mining*, November 2005.
15. Chawla, N. V., Hall, L. O., and Joshi, A., "Wrapper-based computation and evaluation of sampling methods for imbalanced datasets", *Proceedings of ACM SIGKDD Workshop on Utility-based Data Mining*, August 2005.
16. Sylvester, J. and Chawla, N. V., "Evolutionary Ensembles: Combining learning agents using genetic algorithms," *Proceedings of AAAI Workshop on Multi-Agent Systems*, July 2005.
17. Chawla, N. V., "Many are better than one: Improving probabilistic estimates for decision trees," *Proceedings of PASCAL Challenges Workshop*, Southampton, 2005 (Abstract and Presentation only).
18. Chawla, N. V., Karakoulas, G., Roobaert, D., "Lessons learned from the NIPS Feature Selection Challenge", *Proceedings of NIPS Workshop on Feature Selection*, December 2003.
19. Chawla, N. V., "C4.5 and Imbalanced Data sets: Investigating the effect of sampling method, probabilistic estimate, and decision tree structure", *Proceedings of ICML Workshop on Learning from Imbalanced Data sets II,* August 2003.
20. Eschrich, S., Chawla, N. V., Hall, L. O., "Generalization Methods in Bioinformatics," *Proceedings of SIGKDD workshop on Bioinformatics*, August 2002.
21. Chawla, N. V., Moore, T. E., Hall, L. O., Bowyer, K. W., Kegelmeyer, W. P., Springer, C., "Investigation of bagging-like effects and decision trees versus neural nets in protein secondary structure prediction" *Proceedings of ACM SIGKDD Workshop on Data Mining in Bio-Informatics, KDD,* August 2001.
22. Hall, L. O., Chawla, N. V., Bowyer, K. W., "Combining Decision Trees Learned in Parallel", *Proceedings of KDD Workshop on Distributed Data Mining*, New York, August 1998.

**Book Chapters**

1. Nagrecha, S. and Chawla, N. V., "Cambio Score: quantifying climate-change impacts for MSMEs in developing countries," *Private-sector action in adaptation: Perspectives on the role of micro, small and medium size enterprises, 2018.*
2. Feldman, K., Faust, L., Wu, X., Huang, C. and Chawla, N.V., 2017. Beyond Volume: The Impact of Complex Healthcare Data on the Machine Learning Pipeline. In *Towards Integrative Machine Learning and Knowledge Extraction* (pp. 150-169). Springer.
3. Chawla, N. V., and Yang, Y., "Link Prediction: A Primer," Encyclopedia of Social Network Analysis and Mining, 2014.
4. Regola, N., Cielsak, D., Chawla, N. V., "The Need to Consider Hardware Selection when Designing Big Data Applications Supported by Metadata," Big Data Management, Technologies and Applications, 2013.

5. Rider, A., Chawla, N. V., Emrich, S., "A survey of current integrative network algorithms for systems biology," Systems Biology, 479 – 495.
6. Hoens, T. R., and Chawla, N. V., "Imbalanced Datasets: From Sampling to Classifiers," Imbalanced Learning: Foundations, Algorithms, and Applications, 2012.
7. Raeder, T., Forman, G. and Chawla, N. V., "Learning from Imbalanced Datasets: Evaluation Matters," *Data Mining: Foundations and Intelligent Paradigms*, 2011.
8. Steinhaeuser, K. and Chawla, N. V., "A Network Based Approach to Understanding and Predicting Diseases." *Social Computing, Behavioral Modeling and Prediction*, H. Liu, J.J. Salerno, M.J. Young (Eds.), Springer Verlag, 2009.
9. Steinhaeuser, K. and Chawla, N. V., Community Detection in a Large Real-World Social Network." *Social Computing, Behavioral Modeling and Prediction*, H. Liu, J.J. Salerno, M.J. Young (Eds.), Springer Verlag, 2008.
10. Chawla, N. V. "Many are better than one: Improving Probabilistic Estimates from Decision Trees", *Evaluating Predictive Uncertainty, Textual Entailment and Object Recognition Systems,* Eds. (J. Candela, I. Dagan, B. Magnini, F. D'Alche), Springer, 2005.
11. Chawla, N. V., "Mining for Imbalanced Datasets: An overview," *Data Mining and Knowledge Discovery Handbook: A Complete Guide for Practitioners and Researchers*, Eds. (L. Rokach and O. Maimon), Kluwer, 2005.
12. Roobaert, D., Karakoulas, G., Chawla, N. V., "Information Gain and SVM's," *Feature Extraction, Foundations and Applications*, Eds. (I. Guyon, S. Gunn, M. Nikravesh, and L. Zadeh), Springer, 2004.
13. Hall, L. O., Chawla, N. V., Bowyer, K. W., Kegelmeyer, W. P., "Learning Rules from Distributed Data", *Large-scale Parallel Data Mining*, V. 1759 LNAI, Eds. (M. Zaki and H. Ho), Springer-Verlag, 2000.

## Refereed Abstracts

1. Steinhaeuser, K. and Chawla, N. V., "Discovery of Climate Patterns with Complex Networks," *Network Science (NetSci),* 2009.
2. Steinhaeuser, K. and Chawla, N. V., "Is Modularity the Answer to Evaluating Community Structure in Networks," *Network Science (NetSci),* 2008.
3. Raeder, T. and Chawla, N. V., Modeling Product Space as Network for Causality and Profitability. *Statistical Challenges in Electronic Commerce Research*, 2008.

### Edited Proceedings / Special Issues

1. Chawla & Liu, "Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining," 2022.
2. Senthil, K. et al., "Proceedings of the Workshop on Machine Learning in Finance," *ACM KDD,*2021, 2022, 2023, 2024

3. Budak, C., Soroka, S., Singh, L., Bailey, M., Bode, L., Chawla, N., ... & Traugott, M. (2021). "Modeling Considerations for Quantitative Social Science Research Using Social Media Data," psyarxiv, 2021.
4. AI and Covid-19, Special Issue, *Journal of Artificial Intelligence Research (JAIR),* 2021.
5. Chawla, N. V. and Demeniconi, C., "Proceedings of 2020 SIAM International Conference on Data Mining," 2020.
6. Wang, S., Minku, L., Chawla, N. V., and Yao, X., "Learning in the presence of class imbalance and concept drift," *Neurocomputing,* 2018.
7. Chawla, N. V. and Wang, W., "Proceedings of SIAM International Conference on Data Mining," *SIAM Data Mining Conference,* 2017.
8. Yan, Z., Liu, J., Yang, L & Chawla, N. V, "Big Data Fusion in Internet of Things." *Information Fusion*. 40. 10.1016/j.inffus.2017.04.005, 2018.
9. Wang, S., Minku, L.L., Chawla, N. and Yao, X., Proceedings of the IJCAI 2017 Workshop on Learning in the Presence of Class Imbalance and Concept Drift (LPCICD'17). *arXiv preprint arXiv:1707.09425*, 2016
10. Chawla, N. V. and Wang W., "Proceedings of the SIAM Conference on Data Mining," 2017.
11. Y. Zhang, L. Zhang, E. Oki, Chawla, N. V. and A. Kos, "IEEE Access Special Issue: Big Data Analytics for Smart and Connected Health," in *IEEE Access,* vol. 4, pp. 9906-9909, 2016. doi: 10.1109/ACCESS.2016.2646158,
12. Chawla & Hammer, "IEEE BigDataSE," 2015
13. Chawla et al., "IEEE Symposium on Big Data," 2014.
14. Chawla et al., "Workshop on Data Mining for Medicine and Healthcare at SIAM Data Mining" SIAM Data Mining, 2014.
15. Chawla et al., "Workshop on Data Mining for Medicine and Healthcare at KDD" 2013.
16. Chawla et al., "Workshop on Data Mining for Medicine and Healthcare at KDD", 2012.
17. Stefanowski, J., Japkowicz, N., and Chawla, N. V., "Workshop on Class Imbalance: Past, Present, and Future," 2012.
18. Chawla, N. V. and Srivastava, A., "IEEE Conference on Intelligent Data Understanding," 2012.
19. Chawla, N. V., King, I. and Sperduti, A., "IEEE Symposium on Computational Intelligence and Data Mining," 2011.
20. Chawla, N. V. and Srivastava, A., "NASA Conference on Intelligent Data Understanding," October 2011.
21. Chawla, N. V., Ganguly, A. R., Kumar, V., and Steinhaeuser, V., "Knowledge Discovery from Climate Data," 2010.
22. Chawla, N. V. and Yu, P. S, "NASA Conference on Intelligent Data Understanding," October 2010.

23. Chawla, N. V., Ganguly, A., Kumar, V., Steinbech, M., and Steinhaeuser, K., "International Workshop on Knowledge Discovery from Climate Data", *IEEE International Conference on Data Mining,* December 2009.
24. Chawla, N. V., Japkowicz, N., and Zhou, Zhi-Hua, "Workshop on Data Mining when Classes are Imbalanced and Errors have Cost," *Pacific-Asia Conference on Knowledge Discovery and Databases,* May 2009.
25. Omitomu, O. Vatsavai, R., Gama, J., Chawla, N. V., Gaber, M. and Ganguly, A., "Fourth International Workshop on Knowledge Discovery from Sensor Data," *ACM SIGKDD Conference on Knowledge Discovery and Data Mining,* 2010.
26. Chawla, N.V., Japkowicz, N., and Kolcz, A.,"Proceedings of the Workshop in Learning from Imbalanced Datasets II," *International Conference on Machine Learning,* August 2003.

## Conference Tutorials and Special Lectures

1. Knowledge-enhanced Graph Learning, *AAAI,* 2024.
2. Learning on Graphs, *8th International School on Deep Learning*, University of Bournemouth, UK, January 2023.
3. Toward Graph Minimally-Supervised Learning, *28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, 2022.
4. Graph Minimally-supervised Learning, *Fifteenth ACM International Conference on Web Search and Data Mining (WSDM)*, 2022.
5. Multi-modal Network Representation Learning, *26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining,* August 2020.
6. Learning in the Presence of Class Imbalance and Changing Distributions, *6th International Winter School on Big Data*, Ancona, Italy, 2020.
7. Network Science: Representation Learning and Higher Order Networks, *5th International Winter School on Big Data*, Cambridge, United Kingdom, 2019**.**
8. Beyond Graph Mining: Higher-Order Data Analytics for Temporal Network Data, *26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining,* August 2018.
9. Mining in Distribution Sensitive Environments: Tackling the problem of class imbalance and predictive uncertainties, *PAKDD,* May *2010*.
10. Learning in distribution sensitive environments: Tackling the problem of class imbalance and predictive uncertainties, *ACML*, November, 2009.
11. Mining When Classes are Imbalanced, Rare Events Matter More, and Errors Have Costs Attached, *SIAM Data Mining Conference*, May 2009.

## Patents

1. Disease Diagnoses-Bases Disease Prediction, United States Patent no. 8,504,343, Issued August 6th, 2013.

## Software Artifacts

1. MetaMGNN: Few Shot Learning for Molecular Property prediction, available via http://www.nd.edu/~dial
2. HetGNN: Heterogeneous Graph Neural Networks, available via http://www.nd.edu/~dial.
3. Metapath2vec: Representation Learning on Heterogeneous Graphs, available via http://www.nd.edu/~dial.
4. HON: Higher Order Network Representation, http://www.higherordernetworks.com
5. LPmade: Link Prediction Made Easy, available via http://www.mloss.org and http://www.nd.edu/~dial.

   Number of downloads from mloss.org: over 15,895 (In the Top 25 of most downloaded software).
6. DisNet, Distributed Network Processing, Available via http://www.nd.edu/~dial.
7. CARE: Collaborative Assessment and Recommendation Engine for Prospective Healthcare. Accepted as a Software Demo at the *ACM International Symposium on Health Informatics*, 2010.
8. Model Monitor, available via http://www.mloss.org and http://www.nd.edu/~dial.
9. Research on ensembles and imbalanced data learning implemented in the Avatar toolkit at Sandia National Labs in daily production use at Sandia and LLNL.
10. SMOTE, released as part of the Open-source Machine Learning Toolkit WEKA. SMOTE algorithm has also been independently implemented in R, KNIME, and MATLAB.
11. Hellinger distance decision tree software, Available via http://www.nd.edu/~dial. Also released as part of WEKA.
12. Condor Log Analyzer, http://condorlog.cse.nd.edu/.

## Keynote and Distinguished Speaker

1. "Local Context Is Everything: Foundation Models, Knowledge Graphs, and Tools People Use," *ASONAM,* August 2026.
2. "Everything is Connected: Structure and Trust in Agentic Systems for the Long Tail," *Second Workshop on SciSoc Agents & LLMs: Agentic AI for Scientific and Societal Advances,* South Korea, August 2026.
3. "Graph Machine Learning for Scientific Discovery," *Second SIGKDD Workshop on Frontiers in Graph Machine Learning for the Large Model Era*, South Korea, August 2026.
4. "Fraud and Abuse: Where the Minority Class Always Matters," *First SIGKDD Workshop on AI for Fraud and Abuse*, South Korea, August 2026.

5. "The Token Economy," *Citi Wealth Advisory Leadership Summit,* New York, August 2026.
6. "Everything is Connected: Long Tail and Graphs for Integral Ecology," *Horizons 2026, VinUniversity,* Vietnam, July 2026.
7. "Do LLMs Actually Believe Anything?" *NetSci,* June 2026.
8. "Responsible Inclusive Safe and Empowering AI," *Ibero University, Mexico City,* May 2026.
9. "AI: From Real World to Foundations," *University of Texas, Dallas*, April 2026.
10. "AI for Science and Society," *Georgetown University*, April 2026.
11. "AI for Science and Society," *Strathmore University, Kenya*, March 2026
12. "Decoding AI," *Hesburgh Lecture at Cincinnati, Ohio*, March 2026.
13. "Panelist: AI and Collective Flourishing", *University of Notre Dame*, February 2026.
14. "Panelist: AI and Notre Dame," *Notre Dame Board of Trustees Meeting*, Naples Florida, February 2026.
15. "AI for Science and Society," *Illinois Tech*, February 2026.
16. "Heterogeneous Graph Neural Networks," *University of Porto*, January 2026.
17. "AI for Science and Society," *IEEE Big Data,* December 2025.
18. "Panel Discussion (part of panel with VPRs of IU and Purdue) at CICP Board Meeting," *CICP, Indianapolis*, November 2025.
19. "Wealth Tech, AI and the Industry Panel Discussion," *Wealth Management Symposium*, November 2025.
20. "Fireside Chat on AI and Notre Dame," *VISA HQ*, November 2025.
21. "AI For Good and for Everyone: Sustainable AI," *Pontifical Academy of Social Sciences, Vatican City*, October 2025
22. "AI for Common Good," *ILI Signature Event, University of Notre Dame*, October 2025.
23. "Stepping it Up with AI: Research & Education," *Universidad Los Andes*, August 2025.
24. "Learning from Imbalanced Data," *Ukrainian Catholic University,* August 2025.
25. "LLMs as a Force for Good," *Ukrainian Catholic University,* August 2025.
26. "Intelligent Care: The Role of Data, Informatics, and AI in the next Era of Medicine," *Hospital Infantil Mexico-Federico Gomez, Mexico City, Mexico*, July 2025.
27. "Convergence Catalysts: Data & AI Transforming Science, Health & Society," *VinUniversity,* Vietnam, July 2025.
28. "Building Efficient and Explainable AI Models," *NetSci Conference,* Maastricht Netherlands, June 2025.
29. "Knowledge-centric Machine Learning on Graphs," *Wroclaw University of Science and Technology,* Poland, May 2025.
30. "Toward Responsible AI: Advancing Fairness, Explainability, and Scalable Interpretability," *17th Congreso Mexicano de Inteligencia Artificial (COMIA),* Mexico City, May 2025.

31. "Building Bridges: AI and Interdisciplinarity," I*nterdisciplinary Graduate Symposium, University of Notre Dame*, April 2025.
32. "Responsible AI," Hesburgh Lecture Series, University of Notre Dame Atlanta Alumni Club, April 2025.
33. "Stepping up with R.I.S.E. AI," *Nigerian Open University Nigeria*, February 2025.
34. "Responsible & Ethical AI," Notre Dame-IBM Lab Tech Ethics Forum, January 2025.
35. "Graphs, Multimodal Data, LLMs, and Science: From Foundation Models to Applications," *IEEE 3th International Symposium on Information and Communication Technology*, Da Nang, Vietnam, December 2024.
36. "Data & AI Convergence to Translation: Advancing Common Good," *Tulane University*, November 2024.
37. "Traversing the Journey of Data and AI: From Convergence to Translation," *ACM CIKM,* 2024.
38. **"**Traversing the AI + X Continuum for Impact," *University of South Florida, AI+X Symposium,* 2024.
39. "Innovating for Social Good," *Telangana Global AI Summit*, *INDIA* September 2024.
40. "Data & AI Convergence to Translation: Advancing Common Good," *9th Workshop on Data Science for Social Good, European Conference on Machine Learning*, September 2024.
41. "Panelist: Generative AI", *ACM KDD*, August 2024.
42. "How is Generative AI being used in Scientific Disciplines?" *Second NSF-Sponsored Workshop on AI-Enabled Scientific Revolution, University of Minnesota,* August 2024.
43. "Learning on Graphs: From Representation to Generative*," Ukrainian Catholic University, Lviv*, August 2024.
44. Hesburgh Lecturer for the University of Notre Dame and Hosted by Notre Dame Alumni Clubs. *Invited to give guest lectures at Milwaukee (WI), Sarasota (FL), Los Angeles (CA), White Plains (NY),* 2023, 2024.
45. "Generative AI Convergence to Translation for Science and Society," *PUC, Chile*, April 2024.
46. "Generative AI Convergence to Translation for Science and Society," *University of Southern California*, March 2024.
47. "AI & Computing Convergence to Translation," *Purdue University,* February 2024.
48. "Computing and AI: Traversing the Journey from Innovation to Convergence to Translation," *University of Michigan*, December 2023.
49. "Charting the AI Journey from Innovation to Translation", *IDS Distinguished Lecture, University of Hong Kong*, December 2023.
50. "Learning on Graphs," *Portuguese Conference on Artificial Intelligence*, September 2023.
51. Panelist. "AI and Science*,*" *ACM KDD*, August 2023.

52. “Innovations in Artificial Intelligence to Improve Healthcare Access and Reduce Disparities,” *Hospital Infantil Mexico de Federico Gomez*, Mexico City, June 2023.
53. “Digital Ethology,” *Social Capital and the City Seminar Series*, University of Notre Dame (England), July 2023.
54. “Interdisciplinarity of Data Science: From Innovation to Translational Impact,” *Wroclaw University of Science and Technology*, May 2023.
55. “Interdisciplinarity of Data Science: From Innovation to Translational Impact,”, *University of Iowa*, April 2023.
56. “Interdisciplinarity of Data Science: From Innovation to Translational Impact,” *Translational Data Analytics Institute, The Ohio State University*, April 2023.
57. “Driving Innovation and Advancing Societal Impact with Data Science,” *Second International Israel Data Science Initiative Conference*, January 2023.
58. “Learning on Graphs and Multimodal Data: From Representation to Minimally Supervised*”, IEEE International Conference on Knowledge and Systems Engineering, Vietnam*, November 2022.
59. “Deep Ensembles for Graphs with Higher Order Dependencies,” *SIAM Conference on Mathematics of Data Science*, October 2022.
60. “From Data to Graphs: Do we trust the representation?”, *Workshop on Trustworthy Learning on Graphs, CIKM*, October 2022.
61. “Food for AI Thought: From Ingredients to Recipes,” *Menu Directions Conference,* October 2022.
62. Panelist, *Deep Learning Practice and Theory for High-Dimensional Sparse and Imbalanced Data, KDD, August 2022.*
63. “Learning from Imbalanced Data: Progress and Challenges,” *Vietnam National University*, June 2022.
64. “Learning on Graphs: From Representation to Minimally Supervised,” *VinAI*, *Hanoi, Vietnam* June 2022.
65. “Interdisciplinarity of Data Science and AI: Driving Innovation and Advancing the Common Good,” *VinUniversity, Hanoi, Vietnam*, June 2022.
66. “Knowledge Graphs: An Interdisciplinary Perspective,” *Knowledge Graph Conference, Cornell Tech,* May 2022.
67. “Learning on Graphs and Multimodal Data: From Representation to Minimally Supervised,” *National Higher School of AI*, Algeria, June 2022.
68. “Interdisciplinarity of Data Science and AI: Driving Innovation and Advancing Societal Impact**,”** *University of Pisa*, Italy, June 2022.
69. “Structure, Content, Behavior and Dynamics in Heterogeneous (Knowledge) Graphs,” *Winter School on Deep Learning*, *Indian Statistical Institute*, March 2022.
70. “Future of Research– Answering Big Questions through Interdisciplinary Initiatives”, *Science Advisory Council, University of Notre Dame,* November 2021.

71. "DeepSMOTE: Deep Learning for Imbalanced Data," *University of Iowa*, October, 2021.
72. "Panelist: Anomaly and Novelty Detection, Explanation and Accommodation Workshop," ACM *KDD,* August, 2021.
73. "Panelist: Workshop on Knowledge Graph," *ACM KDD,* August, 2021.
74. "Structure, Content, Behavior, and Dynamics in Heterogenous (Knowledge) Graphs," *KDD* 2021.
75. "SMOTE: From Shallow to Deep," *S2D Workshop, International Conference on Learning Representation,* May 2021.
76. "A vision for Data and Society," *Indiana CTSI,* March 2021.
77. "Equitable Data Science: Advancing the Common Good," *Vaibhav Summit India*, January 2021.
78. "Panelist: Why aren't numbers just numbers? How data science and algorithms work in our lives," *ThinkND,* September 2020.
79. "Moderator: Fighting a Pandemic: Convergence of Expertise, Data Science and Policy," *KDD,* August 2020.
80. "Equitable Data Science: Advancing the Common Good," *Intuit*, August 2020
81. "Learning from Multi-Modal Data," *Illinois Institute of Technology (IIT),* February 2020.
82. "Learning from Imbalanced Data," *International School on Big Data (BigDat 2020), Ancona, Italy,* January 2020.
83. "Representation Learning on Heterogeneous Graphs: From Shallow to Deep Embedding." *ACM CIKM Workshop on Heterogeneous Information Network and Analysis,* China, November 2019.
84. "What does quantified self tell us about qualified self," *KDD'19 Workshop,* August 2019.
85. "Democratization of Data Science: Why, How and What," *KDD'19 Panel,* August 2019.
86. "Representation Learning on Heterogeneous Networks," *Sun Yat Sen University,* China, June 2019.
87. "What does quantified self tell us about qualified self," *Distinguished Lecture Series, University of California San Diego (UCSD)*, February 2019.
88. "Representing and Learning Higher Order Networks," *International School on Big Data (BigDat 2019), Cambridge University*, *UK,* January 2019.
89. "Future Ready Education of Engineers," *Maharashtra Institute of Technology*, *India* January 2019.
90. "Machine Learning and Big Data for Financial Risk Management," PRMIA, Toronto, Canada, November 2018.
91. "Interdisciplinarity of Data Science: Driving Innovation and Advancing the Common Good," *Dean's Seminar Series, Monash University, Australia*, October 2018.
92. "Representing Higher Order Dependencies in Networks," *Didi,* China, August 2018

93. "Representing Higher Order Dependencies in Networks," *JD.com,* China, August 2018

94. "Engaging in and Developing International Research Collaborations," *Belgrade, Serbia*, August 2018.

95. "What does the quantified self tell us about qualified self*," 4th Annual International Conference on Computational Social Science*, July 2018.

96. "AI, Data, and Design Thinking: The Trifecta for Personalized and Population Healthcare," *AI Innovations for Life Sciences and Health Care Summit*, June 2018.

97. "Data Science: A Trifecta of People, Data, and Technology," *Engineer Week*, Indiana, February 2018.

98. "Human Machine Partnership: Thinking *about* the future," *Mendoza College of Business*, *University of Notre Dame*, February 2018.

99. "Being a Data'ologist: Data will see you know," *Stanford University*, February 2018.

100. "Representation Learning for Heterogeneous Networks," *International Workshop on Heterogeneous Networks, Analysis, and Mining,* February 2018.

101. "Being a Data'ologist: Data will see you know," *AAAI Workshop on Health Intelligence,* February 2018.

102. "Representing, Modeling, and Learning Higher Order Networks," *Iowa State University,* January 2018.

103. "Interdisciplinarity of Data Science: Driving Innovation and Advancing Common Good," *Politechnika Wrocławska,* December 2017.

104. "Representing, Modeling, and Learning Higher Order Networks," *Indiana University,* November 2017.

105. "Representing, Modeling, and Learning Higher Order Networks," *IIT Bombay,* November 2017.

106. "Being a Data'ologist: Data will see you know," *Indiana Life Sciences Summit*, October 2017.

107. "15 Years of SMOTE: Progress, Opportunities, and Challenges," *ECML/PKDD Workshop on Learning from Imbalanced Data*, September 2017.

108. "Representing, Modeling, and Visualizing Higher Order Networks," *KDD Workshop on Machine Learning in Graphs,* August 2017.

109. "Being a Data'ologist: Data will see you know," *Network Medicine Symposium,* June 2017.

110. "Representing, Modeling, and Learning Higher Order Networks," *Machine Learning in Network Science Symposium,* June 2017.

111. "Computational Lens on Social Science: Enabling Collective Learning and Discovery," *Cultural Innovation and Data Science (Fudan University),* June 2017.

112. Being a Data'ologist: Data will see you know," 21st Century Cures: Southeast Conference, May 2017.

113. "Computational and Network Lens on Social Science: Enabling Collective Learning" *Army Research Labs NS CTA,* March 2017.

114. "From complex systems to networks: discovering and modeling the *correct* network", IBM Cognitive Systems Institute, February 2017.

115. "Interdisciplinarity of Data Science: Driving Innovation and Advancing Common Good," *Arizona State University*, February 2017.

116. "Machine learning and design thinking for personalized healthcare," *Johns Hopkins University,* October 2016.

117. "The Trifecta: People, Data, Algorithms", *Bank of Montreal,* October 2016.

118. "Machine learning and design thinking for personalized healthcare," *PARC Forum,* August 2016.

119. "From complex systems to networks: discovering and modeling the higher order network," *SRI International*, August 2016.

120. "Being a Dataologist: Role of Data and Analytics in Healthcare," *Great Lakes Science Bootcamp,* July 2016.

121. "Data Science for Social Good: Innovation and Impact in Climate and Environmental Sciences," *South East Europe Forum on Data Science*, June 2016.

122. "Leveraging Electronic Medical Records for Personalized and Population Healthcare," *Belgrade Bioinformatics Conference,* June 2016.

123. "Big and Small Data: From Personalized to Population Health Management," *Epidemiology Day,* IUPUI, May 2016.

124. "Higher Order Dependencies in Networks," *Network of Networks Workshop, Network Science Conference*, May 2016.

125. "Computational Lens on Large Scale Social Networks," *Facebook,* April 2016

126. "Being a Data'ologist: From Personalized Healthcare to Population Healthcare", *San Antonio Life Sciences Institute,* February 2016.

127. "Being a Data'ologist: From Personalized Healthcare to Population Healthcare", *Texas FreshAIR (Academia, Industry, Roundtable),* February 2016.

128. "A perspective on data and network science," *2016 IEEE Computational Intelligence Society (CIS) Winter School on Big Data in Computational Intelligence,* February 2016.

129. "Data and Network Science at Notre Dame: From Data to Innovation to Impact," *National Chiao Tung University*, Taiwan, October 2015.

130. "Being a Dataologist: From Personalized to Population Healthcare", *Arizona State University,* October 2015.

131. "Panel Discussion: Trustworthy Data Fusion and Analytics," *IEEE TrustCom/ISPA/BigDataSE*, August 2015.

132. "Big Data for the Common Good: Being a Dataologist" *Lilly Corporation Grand Rounds Seminars*, August 2015.

133. "Big Data and Small data for Personalized Healthcare," *Advances in interactive Knowledge Discovery and Data Mining in complex and big data sets,* July 2015.

134. "Big Data and Personalized Healthcare", *IBRI Scientific Advisory Board*, July 2015.

135. "Being a Dataologist: From Data to Networks to Personalized Healthcare*," 9th International Conference on Computer Recognition Systems*, May 2015

136. "Classifier Evaluation Under Changing Scenarios," *Quality Issues, measures of interestingness, and evaluation of data mining models Workshop,* May 2015

137. "Know Thy Audience: The Data and Digital Convergence," *America East Conference (sponsored by Pennsylvania News Media Association)*, April 2015.

138. "Big Data for the Common Good: Being a 'Dataologist,'" *Wheaton College Science Symposium Series*, March 2015

139. "Interviewed on Big Picture Science Radio show on Big Data", *Big Picture Science*, February 2015

140. "Big Data for the Common Good: Being a 'Dataologist'," *IIT Bombay*, India, February 2015

141. "Research, Education, and Service at Notre Dame: The Drive for Common Good," *Discover ND Event*, India, February 2015

142. "Big Data for the Common Good: Being a 'Dataologist,'" *Indiana University, Bloomington,* December 2014.

143. "Modeling Evolution of Social Networks with the Node Prominence Profile," *INFORMS,* November 2014.

144. "Leveraging Electronic Medical Records for Personalized to Population Health," *6th Annual Meeting of Indiana CTSI, Indianapolis,* September 2014.

145. "Bringing Big Data to Personalized Healthcare: A Patient-Centered Framework," *Wolfram Data Summit*, September 2014.

146. "Know Thy Audience: The Data and Network Convergence," *New York Press Association Fall Convention*, September 2014.

147. "Bringing Big Data to Personalized Healthcare: A Patient-Centered Framework," *Joint Statistical Meeting*, August 2014.

148. "Big Data and Computational Intelligence Panel," *IEEE WCCI,* July 2014.

149. "From Population Health to Personalized Health," *eHealth Initiative's Data & Analytics Council*, June 2014.

150. "Riding the Cloud to Big Data Analytics in and for Healthcare", *BioIT World,* May 2014.

151. "Co-evolution in Complex Networks," *Northwestern University*, March 2014.

152. "Know Thy Audience: The Big Data Way," *Indy Big Data Conference*, February 2014.

153. "Big Data for Common Good: The Synergistic Effects of Wellness in Communities", *TedX Talk*, January 2014.

154. "From Population Health Data to Personalized Health: Big Data for Common Good", *Scientia Talk Science Seminar Series*, University of Notre Dame, December 2013.

155. Panelist for "Health: Abundance Intelligence, A Systems Approach," *Global South Summit, A Global Action Platform,* November 2013.

156. "How hospitals can use data analytics tools to improve costs and quality," *FierceHealthIT,* November 2013.

157. "Personalizing Healthcare Byte by Byte," *#CXO Chat*, *IBM Big Data Analytics,* October 2013.

158. Panelist for "Big Data: Real Challenges and Current Status," *IDEAL Conference,* China, October 2013.

159. "Co-evolving Networks: From cell phone to social to disease networks," INFORMS Minneapolis, October 2013.

160. "From Population Health to Patient-Centered Outcomes," *SJRMC Technical Summit,* September 2013.

161. "Co-evolving Networks: From cell phone to social to disease networks," *Networks Over Time Workshop,* June 2013.

162. "Link Dynamics on Network Evolution," *Social Dynamics Workshop,* June 2013.

163. "Connecting the dots for prospective and personalized healthcare," *Michigan State University*, February 2013.

164. "Big Data and Networks in Healthcare," *University of Iowa,* November 2012.

165. "Big Data and Networks in Healthcare," *Temple University,* October 2012.

166. "Connecting Data and Networks in Healthcare: A vision for personalized healthcare," *Keynote Talk at Indiana Global Health Innovation*, October 2012.

167. "Link Prediction: A recipe for networks, imbalanced data, and evaluation," *Purdue University,* September 2012.

168. "Collaboratory for Adaptation to Climate Change," University of Minnesota, August, 2013.

169. "Perspective on Link Prediction," *IIIT Hyderabad*, July 2012.

170. "Connecting the dots for personalized healthcare," *Epic Systems*, June 2012.

171. *"*Imbalanced Data: From Learning to Evaluation," *IBM Research Beijing,* May 2012.

172. "Perspective on Link Prediction," *Tsinghua University,* May 2012.

173. "Personalized Healthcare: From Electronic Health Record to Genetic Associations," *STAT Day, University of Notre Dame,* April 2012.

174. "Personalized Healthcare: From Electronic Health Record to Genetic Associations," *Translational Informatics Conference, IUPUI,* April 2012.

175. "Personalized Healthcare: A perspective," *Production and Operation Management Society Conference*, April 2012.

176. "The age of social networks and social media," *University of Notre Dame Mendoza College of Business,* March 2012.

177. "Connecting the dots for prospective and personalized healthcare," *IBF Innovation Summit,* February 2012.

178. "Perspective on Link Prediction," *University of Michigan*, February 2012.

179. "Connecting the dots for prospective and personalized healthcare," *Wayne State University,* February 2012.

180. "From Phenotype to Genotype: Charting the personalized healthcare," *IU Simon Cancer Center Grand Rounds,* November 2011.

181. "Connecting the dots for prospective and personalized healthcare," *IBM TJ Watson Research, NY*, October 2011.

182. "Computational thinking in healthcare," *Northshore Health System,* October 2011.

183. "Connecting the dots for prospective healthcare," *ACM SIGKDD Panel on Data Mining for Medicine and Healthcare,* August 2011.

184. "Computational thinking for climate data sciences: From description to prediction to adaptation," August 2011.

185. "Networked thinking in communication, social and natural systems," *Ericsson Research India,* August 2011.

186. "Networked thinking in communication, social and natural systems," *IIT Gandhinagar*, August 2011.

187. "Healthcare: A grand challenge for machine learning," *ICML Workshop on Global Challenges,* July 2011.

188. "Leveraging Electronic Healthcare Record (EHR) for Prospective Healthcare," *Community Health Engagement Program,* June 2011.

189. "Process of link formation in networks," *Spreading, the Influencing, and Cascading in Social and Information Networks Workshop, NetSci*, June 2011.

190. "Evaluation Conundrum in Data Mining," *Quality Issues, Measures of Interestingness, and Evaluation in Data Mining Workshop, PAKDD,* May 2011.

191. "Link Prediction: A Perspective," *Social and Cognitive Networks Academic Research Center, RPI*, February 2011.

192. "Scaling classification to clouds," *IEEE ICDM International Workshop on Knowledge Discovery Using Clouds*, December 2010.

193. "Making most of the electronic health records in personalized medicine," *Purdue University*, September 2010.

194. "Prospective Healthcare: A unified view of genetic and phenotypic data," *Department of Biological Sciences, University of Notre Dame*, September 2010.

195. "Learning from Imbalanced Data," *Applied Artificial Intelligence Conference*, Spain, June 2010.

196. "Perspectives in Link Prediction" *Yahoo! Research*, Barcelona, June 2010.

197. "Mining in Distribution-sensitive Environments," *Laboratory of Artificial Intelligence and Decision Support, University of Porto*, Portugal, June 2010.

198. "Networked thinking in communication, social, and natural systems," *Faculty of Economics, University of Porto*, Portugal, June 2010.

199. "CARE for Your Future: Prospective Disease Prediction Based on Individual Disease Histories," *Buck Institute*, CA. March 2010.

200. "Connected: The New Mantra", *Mendoza College of Business*, University of Notre Dame, February 2010.

201. "Connected: The New Mantra" *Indian School of Business*, India, December 2009.

202. "A framework for monitoring classifiers' performance," *Nanjing University***,** China, October 2009.

203. "Learning to Knowledge Discovery to Action in Distribution Sensitive Scenarios," *NASA CIDU Conference*, October 2009.

204. "A Complex Networks Perspective on Global Climate, Commerce and Environment," *NSF Next Generation Data Mining Workshop*, October 2009.

205. "Learning from Imbalanced Data, School of Informatics," *Indiana University, Bloomington*, August 2009.

206. "Learning to Knowledge Discovery to Action in Distribution Sensitive Scenarios," *ASIAS Symposium (sponsored by FAA, MITRE and NASA***),** Washington DC, August 2009.

207. "Learning to Knowledge Discovery to Action in Distribution Sensitive Scenarios," *Oak Ridge National Labs*, July 2009.

208. "Modeling the product space as a network for causation and profitability," *OCBC Bank Executive Talk*, Singapore, May 2009.

209. "Managing the Tipping Point: When, why and how a model may fail," *SAS, Singapore*, May 2009.

210. Keynote talk, "A framework for monitoring classifiers' performance: when and why failure occur" *Quality issues, measures of interestingness and evaluation of data mining models Workshop at PAKDD*, May 2009.

211. "Personalized Disease Prediction from Electronic Health care databases," *Indiana University School of Medicine*, April 2009.

212. "A recommendation system for disease prediction," *University of Notre Dame Workshop on Biocomplexity*, March 2009.

213. "Prospective Health Care," *Beth Israel Hospital, Harvard University*, October 2008.

214. "Following data: From learning to knowledge to action", *University of Notre Dame, Center of Research Computing*, November 2007.

215. "The first line of action in health care prognostics," *Center for Complex Networks Research, Northeastern University*, May 2007.

216. "Learning in cost distribution sensitive environments," *NIPS Workshop Series*, December 2006.

217. "Learning classifiers in unbalanced and cost-sensitive environments," *University of Louisville*, Kentucky, October 2006.

218. "Data Mining in Customer Analytics and Beyond", *Stern School, NYU*, November 2005.

219. "Biometrics Initiative at Notre Dame," *National Institute of Justice, Washington DC*, October 2005.

220. "Taking Machine Learning to the Real World," *Max Planck Institute of Biological Cybernetics,* Germany, April 2005.

221. "Biometrics: An overview", *School of Informatics, Indiana University*, April 2005.

222. "Extreme learning: Learning from Massive, Imbalanced, and unlabeled data", *GE Global Research*, March 2004.

223. "Data mining in Business Analytics: Mining for Insight and $$", *Fair Isaac***,** CA, February 2004.

224. "Extreme mining: Mining 'needles' in a hay-stack," *IIT Mumbai***,** India, November 2003.

225. "Learning on extremes–size and imbalance–of data", *TATRC, US Army Medical Research and Material Command*, Fort-Detrick, Maryland, May 2002.

226. "Extreme Data Mining: What to do with a flood of training data?" *INFORMS Conference*, Miami, November 2001.

227. "Learning to visualize large scientific data sets," Presentation to the local industry and federal government representatives. *Intelligent Systems Open House***,** University of South Florida, April 2000.

## Students and Post-Doctoral Scholars

### Current PhD Students

1. Jerrick Gerald
2. Bruce Huang
3. Jonathan Karr
4. Grigorii Khvatsky
5. Khiem Le
6. Georgia Li
7. Phuc Minh
8. Alex Nwaganga
9. Brenda Noguiera
10. Anna Sokol

11. Jennifer Schnur

**Graduated PhD Students** (36 students since 2007)

1. Doheon Han, PhD, "Objective-driven Optimization and Context-aware Synthetic Data Generation for Imbalanced Learning," 2026, South Korea Military.
2. Kaiwen Dong, PhD, "Learning on graphs with robustness, efficiency, and adaptability," 2025, Research Scientist, Intuit.
3. Joseph Germino, PhD, "Foundations and Applications of Fair, Explainable, and Interpretable AI," 2025, Chicago Cubs.
4. Yihong Ma, PhD, "Multimodal learning on graphs: Methods and Applications," 2025, Research Scientist, Meta.
5. Zhichun Guo, PhD, "Empowering Graph Neural Networks for Real-world Tasks," 2024, Senior Research Scientist, TikTok.
6. Yijun Tian, PhD, "Knowledge-centric Machine Learning on Graphs," 2024, Applied Scientist, Amazon
7. Steven Krieg, PhD, "Building a Better Neighborhood: Learning Higher-Order Structure and Capturing Complexity in Graphs," 2023, AI Engineer, Research Innovations Incorporated.
8. Damien Dablain, PhD, "Linear Data Augmentation to Improve Generalization for Imbalanced Data," 2023, Research Scientist, AFOSR.
9. Mandana Saebi, PhD, "Learning from Complex Networks," August 2021, Research Scientist, Apple
10. Daheng Wang, PhD, "Learning Complementarity and Dynamics for Contextual Behavioral Modeling,"June 2021, Research Scientist, Amazon.
11. Suwen Lin, PhD, "Personalizing Temporal Multi-modal Sensory Data Mining," March 2021, Research Scientist, FacebookAI.
12. Jermaine Marshall, PhD, "FOODPOLLO: Driving Reliable Food Recommendations from a massive online food portal," August 2020, Senior Data Scientist, Exponent.
13. Pingjie Tang, PhD, "Multiview Representation Learning," August 2020, Research Scientist, Proctor & Gamble.
14. Munira Syed, PhD, "Learning and Inferring User Characteristics from Online Behavior and Content," July 2020, Data Scientist, Proctor & Gamble.
15. Chuxu Zhang, PhD, "Learning from Heterogeneous Data," April 2020, Associate Professor, University of Connecticut.
16. Xian Wu, PhD, "Deep Learning for Sensory and Behavioral Time Series Analysis," March 2020, Research Scientist, Pinterest.
17. Louis Faust, PhD, "Modeling Physiological and Behavioral Data Streams towards Health Insights," June 2020, Data Scientist, Mayo Clinic.

18. Frederick Nwanganga, PhD, “Optimizing Workload Resource Allocation in the Cloud: A Data-Driven Approach", May 2019, Assistant Professor, Mendoza College of Business, Univesity of Notre Dame
19. Chao Huang, PhD, “Spatial-Temporal Data Inference and Forecasting: Models and Applications,” May 2019, Assistant Professor, University of Hong Kong.
20. Aastha Nigam, PhD., “Beyond Who and What: Data Driven Approaches for User Behavior Modeling,” June 2018, Machine Learning Scientist at LinkedIn.
21. Keith Feldman, PhD., “Beyond Modeling: The Emergent Role of Informatics in Advancing Healthcare Knowledge,” June 2018, Associate Professor, University of Michigan.
22. Dipanwita Dasgupta, PhD., “An Evaluation of Tablet Based Solution Impacting Health and Wellness of Older Adults” 2017, User Experience Researcher. Facebook.
23. Yuxiao Dong, PhD, “Computational Lens on Big Networks” 2017, Associate Professor, Tsinghua University.
24. Saurabh Nagrecha, PhD., “Operationalizing Imbalanced Class Problems in Data Science” 2017, Machine Learning Scientist at Capital One Labs.
25. Jian Xu, PhD., “Representing Big Data as Networks: New Methods and Insights” 2017, Data Scientist at Citadel
26. Reid Johnson, PhD., “Data Science for Imbalanced Data: Methods and Applications”, 2016, Data Scientist at Concur (SAP).
27. Everaldo Aguiar, PhD., “Identifying Students at Risk and Beyond: A Machine Learning Approach,” May 2015, Senior Data Scientist at Concur (SAP).
28. Yang Yang, “Network Dynamics: A Social Influence Perspective”, May 2015, Assistant Professor, Syracuse University.
29. Andrew Rider, PhD, “Data and Network Science for Distributed Heterogeneous Systems,” April 2013, Data Scientist at Postmates.
30. Ryan Lichtenwalter, PhD, “Network Analysis and Link Prediction: Effective and Meaningful Modeling and Evaluation,” March 2012. Research Associate at McGill University.
31. Troy Raeder, PhD, “Evaluating and Maintaining Classification Algorithms,”, March 2012. Data Scientist at Dstillery, Inc.
32. T. Ryan Hoens, PhD, “Living in an Imbalanced World,” March 2012. Research Scientist at Microsoft Corporation.
33. Nathan Regola, PhD, “Infrastructure for Big Data,” April 2012. Senior Engineer at Comcast.
34. Darcy A. Davis, PhD, “Network-centric data mining for medical applications,” Data Scientist at Advocate Health Care.
35. Karsten Steinhaeuser, PhD, “Viewing the world with network lens,”, Research Associate at University of Minnesota.
36. David A. Cieslak, PhD “Finding Problems in, Proposing solutions to, and performing analysis on imbalanced data”, July, 2009. Chief Data Scientist, Aunalytics, Inc.

## Current Post-doctoral Scholar

1. Si Chen, 2024 –
2. Yessica Herrera, 2025 –
3. Deng Pan, 2024 -

## Former Post-doctoral Scholars

1. Angelica Garcia Martinez, 2020 – 2025**,** Research Assistant Professor, University of Notre Dame
2. Priscilla Jiminez, 2017 – 2019, Assistant Professor, Grinnell College.
3. Pablo Robles Granda, 2017 – 2019, Research Associate, UIUC
4. Jun Tao, 2016 – 2018, Assistant Professor, Sun Yat-Sen University, China.
5. Beenish Chaudhry, 2014 – 2017. Associate Professor, Department of Computer Science, University of Louisiana.
6. Reid Johnson, 2016-2017. Data Scientist, Zillow.
7. Thanuka Wickramrathne, 2012 – 2015. Assistant Professor, Department of Electrical and Computer Engineering, University of Massachusetts, Lowell
8. Saurav Pandit, PhD, 2010–2012. Head of Data Science, Simple Machine Marketing Technologies.
9. Ryan Lichtenwalter, PhD, 2012. Data Analyst, Alan Edwards Center for Research on Pain, Canada.
10. David Cieslak, PhD, 2011–2012, Chief Data Scientist, Aunalytics, Inc.

## Masters Theses Directed

1. Yao Shen, "The side effects of small molecule drugs: connecting network and adomistic modeling," April 2011, Research Scientist at Columbia University.

## ESTEEM (Entrepreneurship Program) MS Thesis Directed

1. Conor Hanley, Topic: Wellness Platform, 2014
2. Yuan Gao, Topic: Rare Diseases, 2014
3. Daniel Lewis, Topic: Crowd Sourcing, 2012 (EnFocus Fellow)
4. Jennifer Conrads, Topic: Prospective Health Care, 2009

## Global Health MS Thesis Directed

1. Brianna Greary, Increasing the Safety and Efficacy of Short-Term Medical Mission Trips Via Electronic Referral Systems, 2013

## Undergraduate Researchers

1. Matus Vecera, 2026-
2. Max Ficco, 2025 -
3. Aryan Patel, 2025 -
4. Ben Herbst 2025 -
5. Lina McKimson 2025 -
6. Ethan Sotka 2025 -
7. Izzy Molnar, 2025 -
8. Rishi Singh, 2025 -
9. Blake Bezold, 2024 -
10. Satya Kar, 2024
11. Tanay Nager, 2024
12. N. Tejas, 2023
13. John Kim, 2023
14. Sanchita 2023
15. Kate Schinaman, 2023 –
16. Kyle Crosby, 2023 -
17. William Burgis, 2021 – 2022
18. Rebecca Santa Maria 2022
19. Patrick Gerard, 2021 - 2022
20. Patrick Soga, 2020 – 2022
21. Lukasz Matwiejczyj, 2020-21
22. Charles Kraemer, 2020-21
23. Emma Dudrick, 2020-21
24. Alexia Velazquez, 2020-21
25. Xiangyu Dong, 2020-21
26. Attina Zhang, 2019-21
27. Conor Holohan, 2020
28. Mariana Suarez, 2018-2020
29. Catherine Markley, 2018-2020
30. Christopher Giuffrida’19, Amazon Robotics
31. Hind Zahoor, 2018

32. Lisa Huang, 2018
33. Karthik Pansetty, 2018
34. Emily Koh'16, Call Box
35. Shuyang Li'16, Palantir
36. Weizhi Mao'16, Harvard Law School
37. Daniel Tamaru'15, Apple
38. Kimia Ghazi'15, Lilly Corporation
39. Iheanyi Ekechukwu'14, IBM Watson Innovation Labs
40. Steve Kraska'13, Pariveda Solutions, Inc.
41. Jonathan Koch'13, Deloitte
42. Michael O'Brien'13, PhD student at NCSU
43. Robert Thompson'13, Aunalytics
44. Keith Feldman'12, PhD student at Notre Dame
45. Sam Lopes'12, Groupon
46. John Lium'12, Start-up
47. Charles Torbert'12, Amazon
48. Adrian Cunningham'12, PhD student at UC Riverside
49. Jake Lussier'11, PhD student at Stanford University (NSF Graduate Fellowship)
50. Sean McMillan'11, PhD student at University of Michigan
51. Alex Pelan'11, EPIC Systems (Honors thesis)
52. Brian Dentino'11, Edison Engineer at GE Healthcare
53. James Notwell'10, PhD student at Stanford University (NSF Graduate Fellowship)
54. Sean McRoskey'10, Microsoft
55. Christopher Durr'10, Acquity group
56. Joe Thompson'09, PhD student at Notre Dame
57. Philip Whelan'09, Aunalytics
58. Michael Albrecht'08, PhD student at Notre Dame
59. Vincent Thomas'07, FactSet Research Systems
60. Jared Sylvester'06, PhD student at University of Maryland, College Park
61. Daniel Mack'06, MS Columbia University, PhD Student at Vanderbilt University

## Teaching

1. Nitesh V. Chawla, "Teaching Data Mining by Coalescing Theory and Applications," *Proceedings of Frontiers in Education Conference* (Full Paper), 2005. **Selected for the Faculty Fellowship sponsored by NAE/CASEE.**
2. Outstanding Undergraduate Teacher Award, 2008, 2011
3. Student nominated CSE Graduation speaker for 2012, 2013
4. Courses encourage multi-disciplinary participation, student research, and innovation
5. Healthcare Analytics course has students' participating from different colleges: Engineering, Science, Business, Arts & Letters, as well as Global Health and ESTEEM programs.

## Internal Service

1. Inaugural Director, Data, AI, and Computing Strategic Initiative, 2025 -
2. Founding Director, Lucy Family Institute for Data and Society, 2021 – 2026
3. Executive Committee, University Initiative on Ethics, 2023 - 2025
4. Bioengineering and Life Sciences Initiative, Policy & Guidance Committee, 2024 – 2025
5. Computer Science and Engineering (CSE) Graduate Studies Committee, 2017 –
6. Notre Dame International, South Asia Faculty Advisory Committee, 2021 –
7. Executive Committee, Notre Dame Ethics Initiative, 2024 – 2025
8. Steering Committee, Notre Dame Technology Ethics Center, 2020 – 2023
9. Steering Committee, Notre Dame Health and Wellbeing Initiative, 2018 -- 2021
10. Director, Center for Network and Data Science, University of Notre Dame, 2011 -- 2020
11. Fellow, Pulte Institute for Global Development, University of Notre Dame, 2019 --
12. Fellow, Kellogg Institute for International Studies, University of Notre Dame, 2017 –
13. Fellow, Kroc Institute for Peace Studies, University of Notre Dame, 2015 –
14. Fellow, Leo Institute for Asia and Asian Studies, University of Notre Dame, 2014 –
15. Fellow, Reilly Center for Science, Technology and Values, University of Notre Dame, 2012 -
16. Chair, Structural and Social Determinants of Health, Indiana CTSI, 2020 - 2025
17. Advisory Committee, College of Science, 2024 –
18. Data and Computational Science Task force, 2023 – 2024
19. University Strategic Theme Advisory Committee on Health & Wellbeing, 2022.
20. PI on Moment to See Courage to Act Proposal on "A Call to Action: An Initiative to Solve Health and Healthcare Inequities." 2022.
21. Lucy AOV Faculty Search, 2020 –
    - Participation in faculty search across Sociology, ITAO, Global Affairs, AME, ACMS, English, American Studies. Involved attending candidates' seminars, providing feedback to search process, developing MOUs).
    - Lucy AOV Cluster Search Committee Chair
      The Committee reviewed about 200 applicants, organized over a dozen preliminary zoom interviews, match-make with departments / schools, and then helped host the candidates on-campus visit, 2021 –

22. AnalytiXIN, 2021 –
    Recruitment of managing director, faculty fellows, RFP, establishment of agreements with companies and other programming activities).
23. Co-Chair, Faculty Search, Computer Science and Engineering, 2019 – 2020
24. Executive Steering Committee for ND Technology and Ethics Center, 2019 -
25. Executive Steering Committee for Health and Wellbeing Initiative, 2018 – 2021
26. Task Force, Lucy Institute of Data and Society, 2019 – 2020
27. College of Engineering Council, 2017 - 2021
28. Director of CNDS, 2011 – 2020
29. MS Data Science, Advisory Committee, 2016 –
30. Advising India Engagement for the Office of Internationalization, 2014 --
31. Beijing Global Gateway Committee, 2014 --
32. Chair, Big Data Faculty Search Committee in CSE, 2013 – 2016
33. Dean of Science Review Committee, 2013
34. OIT Organizational Analysis and Design (OAD) Governance Team, 2013
35. University Internationalization Committee, 2013 –
36. CSE Committee on Promotion and Tenure, 2011 –
37. Steering Committee, ECI SRI, 2010 –
38. CSE Honesty Committee, 2010 –
39. CSE Faculty Search Committee, 2010 and 2011
40. Panelist on New Faculty Orientation, 2009
41. I-CTSI committee for Design and Biostatistics, 2009 –
42. Advisory Board Member, Bioinformatics Core, 2009 –
43. Dept. Seminar Coordinator: 2007 –

## External Service

- Leadership Roles
    - Founder and Steering Committee Chair, Responsible, Inclusive, Safe and Empowering (RISE) AI Conference and Summit
    - Dissertation Award Chair, ACM KDD 2023, 2024, 2025, 2026
    - Editor-In-Chief, ACM AI Letters, 2025 –
    - Founding Blue Sky Ideas Track Co-Chair, ACM KDD 2026
    - Editor-In-Chief, Springer Nature Discover Data, 2022 – 2025
    - General Co-Chair, CODS-COMAD Conference, 2024
    - Founding Blue Sky Ideas Chair, SIAM Data Mining 2023
    - Specialty Editor in Chief, Frontiers Big Data, 2020 -
    - Area Chair, AAAI, 2022, 2023, 2024, 2026

- Area Chair, IJCAI, 2021, 2022, 2023, 2024
- Program Co-Chair, ACM KDD 2022
- General Co-Chair, SIAM Conference on Data Mining (SDM), 2020
- General Co-Chair, SIAM Conference on Data Mining (SDM), 2019
- Best Paper Awards Committee, SIAM Data Mining Conference, 2013, IEEE International Conference on Data Mining, 2022.
- Panel Co-Chair, ACM KDD, 2019, ACM KDD 2020, SIAM SDM 2022
- Workshops Co-Chair, ACM KDD, 2018
- Scientific Chair, AI In Medicine, 2020
- Program Co-Chair, SIAM Conference on Data Mining, 2017
- Program Co-Chair, IEEE International Conference on Big Data Science and Engineering, 2015
- Program Co-Chair, IEEE Symposium on Computational Intelligence in Big Data, 2014
- Program Co-Chair, NASA CIDU Conference, 2010, 2011
- Program Co-Chair, CIDM Symposium, 2011, 2013
- Chair, IEEE CIS Big Data Task Force, 2012 – 2014
- Chair, IEEE CIS Data Mining Technical Committee, 2011 – 2013
- Contest Co-Chair, IEEE ICDM, 2013
- Contest Co-Chair, PAKDD, 2014  2015
- Sponsorship Co-Chair, SIAM Data Mining, 2015
- General Co-Chair, IEEE CIDU Conference, 2012
- Vice Chair, IEEE ICDM Conference, 2011
- Tutorials Chair, SIAM Data Mining Conference, 2012
- Industry Track Chair, ACM/IEEE Conference on Analysis of Social Networks and Modeling, 2012, 2013
- Co-Chair, ACM SIGKDD Workshop on Data Mining for Medicine and Healthcare, 2011
- Treasurer, ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 2010.
- Co-Chair, IEEE ICDM Workshop on Knowledge Discovery from Climate Data: Prediction, Extremes, and Impacts, 2009, 2010
- Area Chair, European Conference on Machine Learning (ECML/PKDD), 2009
- Co-Chair, ACM SIGKDD International Workshop on Knowledge Discovery from Sensor Data, 2008, 2009, 2010

- Track Chair, ACM Conference on Information and Knowledge Management (CIKM), 2008
- Publications Chair, IEEE First International Conference on Biometric: Theory, Applications, and Systems, 2007
- Special Activities Chair, IEEE Symposium on Computational Intelligence and Data Mining (CIDM), 2006
- Organizing Committee, Special Session on Ensembles for Extreme Environments, IEEE SMC, 2005
- Co-Chair, Workshop on Learning from Imbalanced Data sets at the 20th International Conference on Machine Learning *(ICML*), August 2000.

- Editorial Board/Associate Editor
  - ACM Transactions on Knowledge and Data Discovery, 2020 -
  - IEEE Transactions on Knowledge and Data Engineering, 2015 –
  - ACM Transactions on Intelligent Systems and Technology, 2021 -
  - IEEE Transactions on Big Data, 2019 -
  - Statistical Analysis and Data Mining, 2013 – 2017
  - IEEE Transactions on Neural Networks and Learning Systems, 2016 – 2020
  - Knowledge and Information Systems, 2016 – 2020
  - Nature Scientific Reports, 2014 – 2020
  - IEEE Transactions on Cybernetics – B, 2005 – 2014
- Guest Editor, Statistical Analysis and Data Mining Journal, 2011, 2012
- Panelist, ACM SIGKDD Workshop on Utility Based Data Mining (UBDM), 2005
- Guest-Editor, SIGKDD Explorations, June 2004
- NIH Study Section Reviewer, 2020, 2021, 2022, 2024
- Invited Participant, Workshop for Strategizing and Transforming Higher Education in India, 2022 and 2023
- Invited Participant, Workshop on Indo US Visioning Workshop: Developing a Diverse AI Workforce, 2022
- NSF Panelist, 2008, 2009, 2010, 2012, 2013, 2014, 2015, 2017, 2018, 2020, 2021, 2022
- Science Foundation of Ireland Proposal Review, 2018, 2020, 2023
- Department of Homeland Security Center of Excellence Review, 2019
- Senior / Program Committee
  - ACM KDD 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020, 2021
  - AAAI 2016, 2019
  - IEEE ICHI 2018
  - ACM WSDM 2012. 2013, 2014, 2015, 2016

- SIAM Data Mining 2007, 2009, 2010, 2011, 2012, 2013 2014, 2015, 2016, 2018, 2022
- ACM CIKM, 2012. 2013
- IEEE ICDM 2011, 2013
- ACM/IEEE ASONAM, 2010, 2011, 2012, 2013, 2014
- ICML 2009, 2010
- WWW 2010
- PAKDD 2012, 2013
- IEEE Big Data, 2014, 2018
- IEEE SocialCom, 2014
- BigComp, 2015
- CODS, 2014, 2015
- ACM IHI 2011, 2012
- ACM BCB 2011
- IHIC, 2013
- ICPR 2010
- ECML/PKDD, 2008, 2010
- ISMB 2006, 2007, 2009
- ACML 2009
- FLAIRS Data Mining 2007, 2008, 2009, 2010
- AAAI 2006, 2007
- AAAI Workshop on Evaluation Methods for Machine Learning 2006, 2007
- ACM SIGKDD Workshop on Utility Based Data Mining 2005, 2006
- IEEE GrC Computing 2006, 2007, 2009