

ANUJ KARPATNE

Contact Details

+1 (612) 222-8228
karpa009@umn.edu
www-users.cs.umn.edu/~anuj/

Postal Address

200 Union Street SE
Minneapolis, MN 55455

EDUCATION

- PhD, Computer Science** 2011 – 2017
University of Minnesota (UMN), Twin Cities
Thesis: “Predictive Learning with Heterogeneity in Populations”
Advisor: Vipin Kumar
- Integrated M.Tech, Mathematics and Computing** 2006 – 2011
Indian Institute of Technology Delhi (IITD)

RESEARCH EXPERIENCE

- PostDoctoral Associate** Oct 2017 – present
Kumar Research Group, UMN
- Research Assistant** 2011 – 2017
Expeditions in Computing Research Group, UMN
- Research Intern** Summer 2013
Smarter Planet Group, IBM T.J. Watson Research Center

RESEARCH INTERESTS

Data mining and machine learning; Combining scientific knowledge (physics) with data science; Spatio-temporal data mining; Climate and Earth science problems.

HONORS AND AWARDS

- Recipient of the **Doctoral Dissertation Fellowship** by the University of Minnesota for 2015.
- Recipient of University of Minnesota Informatics Institute (**UMII Graduate Fellowship**) for 2015.
- Recipient of **Student Travel Awards** at SIAM International Conference on Data Mining (SDM) 2014 and 2015, IEEE International Conference on Data Mining (ICDM) 2015, Conference on Intelligent Data Understanding (CIDU) 2012, and Climate Informatics Workshop 2013, 2014, and 2015.
- Recipient of two consecutive **Director’s Merit Awards** at IIT Delhi.

PUBLICATIONS

BOOK

- [B1] P. Tan, M. Steinbach, **A. Karpatne**, and V. Kumar “Introduction to Data Mining,” *Pearson Addison–Wesley (Second Edition)*, 2017 (available in print by January, 2018).

JOURNAL ARTICLES

- [J10] **A. Karpatne**, G. Atluri, J. Faghmous, M. Steinbach, A. Banerjee, A. Ganguly, S. Shekhar, N. Samatova, and V. Kumar, “Theory-guided Data Science: A New Paradigm for Scientific Discovery from Data,” *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 29(10), 2318–2331, 2017.
- [J9] *G. Atluri, ***A. Karpatne**, and V. Kumar, “Spatio-temporal Data Mining: A Survey of Problems and Methods,” *ACM Computing Surveys*, 2017 (accepted; *arXiv: 1711.04710*) (* equal contribution).

- [J8] **A. Karpatne**, I. Ebert-Uphoff, S. Ravela, H. A. Babaie, and V. Kumar, “Machine Learning for the Geosciences: Challenges and Research Opportunities,” *IEEE TKDE*, 2017 (*in review; arXiv: 1711.04708*).
- [J7] *A. Khandelwal, ***A. Karpatne**, *M.E. Marlier, J. Kim, D. P. Lettenmaier, and V. Kumar, “An Approach for Global Monitoring of Surface Water Extent Variations Using MODIS Data,” *Remote Sensing of Environment, Elsevier*, 2017 (* equal contribution).
- [J6] **A. Karpatne**, Z. Jiang, R. R. Vatsavai, S. Shekhar, and V. Kumar, “Monitoring Land Cover Changes: A Machine Learning Perspective,” *IEEE Geoscience and Remote Sensing Magazine*, 4(2), 8–21, 2016.
- [J5] **A. Karpatne** and S. Liess, “A Guide to Earth Science Data: Summary and Research Challenges,” *IEEE Computing in Science & Engineering*, 17(6), 14–18, 2015.
- [J4] F. Schrodtt, J. Kattge, H. Shan, F. Fazayeli, J. Joswig, A. Banerjee, M. Reichstein, G. Bónisch, S. Díaz, J. Dickie, A. Gillison, **A. Karpatne**, S. Lavorel, P.W. Leadley, C. Wirth, I. Wright, S.J. Wright, and P.B. Reich, “BHPMF - A Hierarchical Bayesian Approach to Gap-filling and Trait Prediction for Macroecology and Functional Biogeography,” *Global Ecology and Biogeography*, 24(12), 1510–1521, 2015.
- [J3] R. Khemchandani, **A. Karpatne**, and S. Chandra, “Twin Support Vector Regression for the Simultaneous Learning of a Function and its Derivatives,” *International Journal of Machine Learning and Cybernetics*, 4(1), 51–63, 2013.
- [J2] R. Khemchandani, **A. Karpatne**, and S. Chandra, “Proximal Support Tensor Machines,” *International Journal of Machine Learning and Cybernetics*, 4(6), 703–712, 2013.
- [J1] R. Khemchandani, **A. Karpatne**, and S. Chandra, “Generalized Eigenvalue Proximal Support Vector Regressor,” *Expert Systems with Applications*, 38(10), 13136–13142, 2011.

PEER-REVIEWED CONFERENCE PAPERS

- [C9] **A. Karpatne**, W. Watkins, J. Read, and V. Kumar, “Physics-guided Neural Networks (PGNN): An Application in Lake Temperature Modeling,” *SIAM International Conference on Data Mining (SDM)*, 2018 (*in review; arXiv: 1710.11431*).
- [C8] X. Jia, Y. Hu, A. Khandelwal, **A. Karpatne**, and V. Kumar, “Joint Sparse Auto-encoder: A Semi-supervised Spatio-temporal Approach in Mapping Large-scale Croplands,” *IEEE International Conference on Big Data*, 2017.
- [C7] S. Agrawal, G. Atluri, **A. Karpatne**, S. Chatterjee, S. Liess, and V. Kumar, “Triples: A New Class of Relationships in Time Series Data,” *ACM International Conference on Knowledge Discovery and Data Mining (KDD)*, 697–706, 2017.
- [C6] X. Jia, X. Chen, **A. Karpatne**, and Vipin Kumar, “Identifying Dynamic Changes with Noisy Labels in Spatial-temporal Data: A Study on Large-scale Water Monitoring Application,” *IEEE International Conference on Big Data*, 1328–1333, 2016.
- [C5] **A. Karpatne** and V. Kumar, “Adaptive Heterogeneous Ensemble Learning Using the Context of Test Instances,” *IEEE International Conference on Data Mining (ICDM)*, 787–792, 2015.
- [C4] **A. Karpatne**, A. Khandelwal, and V. Kumar, “Ensemble learning methods for binary classification with multi-modality within the classes,” *SDM*, (82) 730–738, 2015.
- [C3] **A. Karpatne**, A. Khandelwal, S. Boriah, and V. Kumar, “Predictive learning in the presence of heterogeneity and limited training data,” *SDM*, (29) 253–261, 2014.
- [C2] **A. Karpatne**, M. Blank, M. Lau, S. Boriah, K. Steinhaeuser, M. Steinbach, and V. Kumar, “Importance of vegetation type in forest cover estimation,” *NASA Conference on Intelligent Data Understanding (CIDU)*, 71–78, 2012.
- [C1] *X. Chen, ***A. Karpatne**, *Y. Chamber, V. Mithal, M. Lau, K. Steinhaeuser, S. Boriah, M. Steinbach, V. Kumar, C.S. Potter, S.A. Klooster, T. Abraham, J.D. Stanley, and J.C. Castilla-Rubio, “A new data mining framework for forest fire mapping,” *CIDU*, 104–111, 2012 (* equal contribution).

BOOK CHAPTERS

[BC2] **A. Karpatne**, A. Khandelwal, X. Chen, V. Mithal, J. Faghmous, and V. Kumar, “Global monitoring of inland water dynamics: State-of-the-art, challenges, and opportunities,” In *Computational Sustainability*, J. Lässig, K. Kersting, and K. Morik (Eds.), Springer, 121–147, 2016.

[BC1] **A. Karpatne**, J. Faghmous, J. Kawale, L. Styles, M. Blank, V. Mithal, X. Chen, A. Khandelwal, S. Boriah, K. Steinhaeuser, M. Steinbach, and V. Kumar, “Earth science applications of sensor data,” In *Managing and Mining Sensor Data*, C. Aggarwal (Ed.), Springer, 505–530, 2013.

PEER-REVIEWED WORKSHOP PROCEEDINGS

[W7] **A. Karpatne** and V. Kumar, “Learning Physics-based Models in Hydrology under the Framework of Generative Adversarial Networks”, *American Geophysical Union (AGU) Fall Meeting*, 2017.

[W6] **A. Karpatne**, W. Watkins, J. Read, and V. Kumar, “Physics-guided Learning of Neural Networks: An Application in Lake Temperature Modeling”, *NIPS Workshop on Deep Learning for Physical Sciences*, 2017.

[W5] **A. Karpatne**, H. Babaie, S. Ravela, V. Kumar, and I. Ebert-Uphoff, “Machine Learning for the Geosciences—Opportunities, Challenges, and Implications for the ML process”, *SDM Workshop on Mining Big Data in Climate and Environment*, 2017.

[W4] S. Gopal, **A. Karpatne**, R. R. Vatsavai, and V. Kumar, “Modeling the Food-Energy-Water Nexus in Critical Biodiverse Landscapes,” *ACM KDD Workshop on Data Science for Food, Energy and Water*, 2016.

[W3] **A. Karpatne**, A. Khandelwal, R. Anderson, M. Blank, S. Boriah, and V. Kumar, “Group-specific local learning for global lake monitoring”, *The Fourth International Workshop on Climate Informatics*, 2014.

[W2] **A. Karpatne**, J. Faghmous, M. Blank, R. Anderson, S. Boriah, S. Liess, and V. Kumar, “Understanding the Influence of Sea Surface Temperatures on Terrestrial Ecosystem Disturbances”, *The Third International Workshop on Climate Informatics*, 2013.

[W1] **A. Karpatne**, M. Blank, J. Middleton, S. Boriah, K. Steinhaeuser, M. Steinbach, S. Chatterjee, and V. Kumar, “Understanding relationships between fire activity and sea surface temperature anomalies”, *AGU Fall Meeting*, 2012.

INVITED TALKS AND PANEL DISCUSSIONS

[PD2] “Theory-guided Data Science: A New Paradigm for Scientific Discovery,” **Panel Discussion at International Conference on Scientific and Statistical Database Management**, June 29, 2017.

[PD1] “Understanding and Narrowing Gaps Between Data Science and Mechanistic Theories in Physical Sciences,” **Panel Discussion at SDM Workshop on Mining Big Data in Climate and Environment**, April 29, 2017.

[T4] “How Can Physics Inform Deep Learning Methods in Earth System Science?: Recent Progress and Future Prospects,” **Invited Keynote Talk at ICDM Workshop on Data Mining in Earth System Science**, November 18, 2017.

[T3] “Theory-guided Data Science: A New Paradigm for Scientific Discovery in the Era of Big Data,” **Invited Talk at American Institute of Chemical Engineers (AIChE) Annual Meeting**, October 30, 2017.

[T2] “Theory-guided Data Science: A New Paradigm for Advancing Scientific Discovery in the Big Data Era,” **Invited Seminar Talk at University of Notre Dame**, February 9, 2017.

[T1] “Global Monitoring of Inland Surface Water Dynamics Using Remote Sensing Data,” **Invited Talk at 96th American Meteorological Society Annual Meeting**, January 11–14, 2016.

PATENTS

[P2] **A. Karpatne** and V. Kumar, “Multi-Modal Data and Class Confusion: Application in Water

Monitoring,” US Patent 15/403,708, issued January, 2017.

[P1] A. Hamarpur, **A. Karpatne**, H. Li, X. Liu, R. Lougee, B. Qian, and S. Xing, “Characterizing relationships among space-time events,” US Patent 20,160,034,323, issued February, 2016.

GRANT WRITING

Major contributor in the preparation of the following research grants:

- NSF Expeditions in Computing. (Invited for Reverse Site Visit) “ <i>Water in the 21st Century: A Data-guided Approach</i> ”	2017–2022 \$10,000,000
- NSF Innovations at the Nexus of Food, Energy and Water Systems (INFEWS). (Funded) “ <i>Innovations for Sustainable Food, Energy, and Water Supplies in Intensively Cultivated Regions: Integrating Technologies, Data, and Human Behavior</i> ”	2017–2022 \$2,500,000
- DARPA World Modelers Program. (Funded) “ <i>Model Integration Through Knowledge-Rich Data and Process Composition</i> ”	2017–2021
- NSF/NIH Smart and Connected Health (SCH) Program. (Funded) “ <i>Group-Specific Learning to Personalize Evidence-Based Medicine</i> ”	2016–2019 \$500,000
- NASA. (Funded) “ <i>Integrating Distributed Data Mining Algorithms into NEX</i> ”	2012–2014 \$100,000
- NASA. (Funded) “ <i>Automated Detection of Precursors to Aviation Safety Incidents</i> ”	2012–2017 \$500,000

EDUCATIONAL ACTIVITIES

Instructor for Summer School on “Intelligent Systems for Geosciences (IS-GEO)”, 2017.

- Presented lecture course on “Introduction to Machine Learning.”
- Provided hands-on tutorial on “Monitoring Surface Water using Remote Sensing Data.”

Teaching Assistant for graduate-level “Introduction to Data Mining” course, Fall 2012–2014.

Completed **Preparing Future Faculty Course** at University of Minnesota, Spring 2016.

Mentored the following **undergraduate** students:

- *University of Minnesota*: Robert Leunberger (Summer 2015–Spring 2017); Connor Tarczon (Spring 2016–Spring 2017); Vishnu Arun (Spring 2016–Fall 2016); Kaiwei Ang (Summer 2016–Fall 2016); Yizheng Ding (Summer–Fall 2015); Stryker Thomson (Fall 2015–Spring 2016); Eric Mccaleb (Spring 2015–2016); Reid Anderson (Fall 2012–Spring 2016); Mace Blank (Spring 2012–2016).
- *Ausburg University*: Nyssa Capman (Summer 2017); *Delhi College of Engineering, India*: Anirudh Mittal (Summer 2016); *North Carolina A&T State University*: James Middleton (Summer 2012).

Mentored the following **graduate** students:

- *University of Minnesota*: Saurabh Agarwal; Xiaowei Jia; Ankush Khandelwal; Kshitij Tayal.

PROFESSIONAL SERVICE

Convener for session on “Intelligent Systems for Geosciences: Accelerating Discovery and Building Community” at *AGU Fall Meeting*, 2017.

Program Committee Member for the following conferences:

- *SDM Workshop on Mining Big Data in Climate and Environment*, 2017.
- *International Joint Conference on Artificial Intelligence*, 2013.

Reviewer for the following conferences and journal proceedings:

- *ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, *IEEE International Conference on Data Mining (ICDM)*, *SIAM International Conference on Data Mining*, *Association for the Advancement of Artificial Intelligence (AAAI) Conference*.
- *BMC Bioinformatics*, *IEEE Geoscience and Remote Sensing Letters*, *IEEE Transactions on Geoscience and Remote Sensing*, *Elsevier: Information Sciences*, *Springer: Neural Computing and Applications*, *Springer: Data Mining and Knowledge Discovery*.

REFERENCES

Vipin Kumar

Regents Professor and William Norris Chair in Large Scale Computing
Department of Computer Science and Engineering
University of Minnesota, Twin Cities
email: kumar001@umn.edu

Arindam Banerjee

Professor
Department of Computer Science and Engineering
University of Minnesota, Twin Cities
email: baner029@umn.edu

Yolanda Gil

President-Elect, Association for the Advancement of Artificial Intelligence (AAAI)
Information Sciences Institute
University of Southern California
email: gil@isi.edu

Imme Ebert-Uphoff

Research Faculty
Department of Electrical and Computer Engineering
Colorado State University
email: iebert@colostate.edu

Sai Ravela

Principal Research Scientist
Department of Earth Atmospheric and Planetary Sciences
Massachusetts Institute of Technology
email: ravela@mit.edu

Dennis Lettenmaier

Distinguished Professor
Department of Geography
University of California, Los Angeles
email: dlettenm@ucla.edu