

ANUJ KARPATNE

Contact Details

+1 (540) 231-6420
karpatne@vt.edu
people.cs.vt.edu/karpatne

Mailing Address

114 McBryde Hall (0106)
Dept. of CS, Virginia Tech.
Blacksburg, VA 24061

WORK EXPERIENCE

Assistant Professor <i>Department of Computer Science, Virginia Tech</i>	2018 – present
PostDoctoral Associate <i>Kumar Research Group, UMN</i>	2017 – 2018
Research Assistant <i>Expeditions in Computing Research Group, UMN</i>	2011 – 2017

EDUCATION

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

RESEARCH INTERESTS

Data mining and machine learning; Combining scientific knowledge (physics) with data science; Theory-guided data science; Physics-guided machine learning.

FUNDING

- **NSF-OAC-1940247**, “Collaborative Research: Biology-guided neural networks for discovering phenotypic traits,” (VT PI, \$422,000; Duration: 10/01/2019-09/30/2021; collaborative project with Battelle, Tulane University, Drexel University, and University of Washington Seattle).
- **NSF-IIS-2026710**, “EAGER: Collaborative Research:III: Exploring Physics Guided Machine Learning for Accelerating Sensing and Physical Sciences,” (VT PI, \$54,452; Duration: 05/01/2020-04/30/2021; collaborative project with Ohio State University, SUNY Binghamton, and University of Massachusetts Lowell).

TEACHING

- Instructor** for “CS 4824 / ECE 4424: Machine Learning”, Virginia Tech, Spring 2020.
- Instructor** for “CS(STAT) 5525: Data Analytics I”, Virginia Tech, Fall 2019.
- Instructor** for “CS(STAT) 5525: Data Analytics I”, Virginia Tech, Spring 2019.
- Instructor** for “CS 6804: Machine Learning Meets Physics”, Virginia Tech, Fall 2018.
- Instructor** for Summer School on “Intelligent Systems for Geosciences (IS-GEO)”, UT Austin, 2017.

HONORS AND AWARDS

Named the **Inaugural Research Fellow** by the Intelligent Systems for Geosciences (IS-GEO), sponsored by Petrobras, for 2019.

Recipient of the **Doctoral Disseration 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.

ADVISING AND MENTORING ROLES

Advising the following Ph.D. students:

- *Arka Daw*, Dept. of Computer Science, Virginia Tech (Fall 2018 – present).
- *Md Abdullah Al Maruf*, Dept. of Computer Science, Virginia Tech (Summer 2019 – present).
- *Jie Bu*, Dept. of Computer Science, Virginia Tech (Summer 2019– present).
- *Mohannad Elhamod*, Dept. of Computer Science, Virginia Tech (Fall 2019 – present).
- *Snehal More*, Dept. of Forest Resources and Environmental Conservation, Virginia Tech (Fall 2019 – present).

Advising the following M.S. students:

- *Jie Bu*, Dept. of Computer Science, Virginia Tech (Summer 2019 – present).
- *Ioannis Papakis*, Dept. of Computer Science, Virginia Tech (Fall 2019 – present).
- *Arya Shahadi*, Dept. of Computer Science, Virginia Tech (Fall 2019 – present).
- *Zheng Li*, Dept. of Computer Science, Virginia Tech (Spring 2020 – present).
- *Gengrui Wei*, Dept. of Computer Science, Virginia Tech (Spring 2020 – present).
- *Sandhya Bhaskar*, Dept. of Electrical and Computer Engineering, Virginia Tech (Spring 2020 – present).
- *Prathamesh Kalyan Mandke*, Dept. of Electrical and Computer Engineering, Virginia Tech (Spring 2020 – present).

Mentoring the following graduate students:

- *Nikhil Muralidhar*, Ph.D. in Computer Science (advised by Prof. Naren Ramakrishnan), Virginia Tech (Fall 2018 – present).
- *Mansoorah Ahmadian*, Ph.D. in Computer Science (advised by Prof. Yang Cao), Virginia Tech (Fall 2018 – present).
- *Yun Dong*, Ph.D. in Electrical and Computer Engineering (advised by Prof. Elena Lind), Virginia Tech (Spring 2019 – present).

PROFESSIONAL SERVICE

Co-Editor-in-Chief (EiC) of the ACM Special Interest Group in Artificial Intelligence (SIGAI) quarterly newsletter, “AI Matters.”

Review Editorial Board Member for “Data-driven Climate Sciences” section in *Frontiers in Big Data*.

Workshop Co-chair for KDD 2019.

Co-organizer of session on “How AI and Knowledge Centers are Changing Societal Views of Critical

Earth Resources” at *American Association for the Advancement of Science (AAAS) Annual Meeting*, 2019.

Co-organizer of workshop on “Fragile Earth: Theory Guided Data Science to Enhance Scientific Discovery (FEED)” at *KDD*, 2018.

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

Program Committee Member for the following workshops and conferences:

- *KDD (Research Track)*, 2020.
- *International Joint Conference on Artificial Intelligence (IJCAI)*, 2020.
- *SIAM International Conference on Data Mining (SDM)*, 2020.
- *Association for the Advancement of Artificial Intelligence (AAAI) Conference (AI for Social Impact Track)*, 2020.
- *AAAI Workshop on Knowledge Discovery from Unstructured Data in Financial Services*, 2020.
- *NeurIPS*, 2019.
- *ICML*, 2019.
- *KDD (Research Track)*, 2019.
- *SDM*, 2019.
- *AAAI*, 2019.
- *Workshop on “A new paradigm in lake and reservoir research and management through global monitoring, modeling, and engaging and empowering people networks,”*, 2018.
- *KDD (Research Track)*, 2018.
- *SDM Workshop on Mining Big Data in Climate and Environment*, 2017.
- *IJCAI*, 2013.

Reviewer for the following conferences and journal proceedings:

- *Springer Neural Processing Letters 2020, Nature Scientific Reports 2020, Physical Review Letters 2020, Artificial Intelligence Review 2020, Medical Physics 2020, IEEE Transactions on Knowledge and Data Engineering 2018, IEEE Transactions on Pattern Analysis and Machine Intelligence 2018, Elsevier: Information Sciences 2013, Springer: Neural Computing and Applications 2014, Springer: Data Mining and Knowledge Discovery 2014, Wiley Ecosphere 2018, BMC Bioinformatics 2015, IEEE Geoscience and Remote Sensing Letters 2016, IEEE Transactions on Geoscience and Remote Sensing 2017, Springer: Remote Sensing of Environment 2018, PLoS One 2018.*

PROFESSIONAL AFFILIATIONS

Member, Society for Industrial and Applied Mathematics (SIAM).

Member, Association for Computing Machinery (ACM).

Member, American Association for the Advancement of Science (AAAS).

Member, American Geophysical Union (AGU).

Member, Interdisciplinary Graduate Education Program (IGEP) in Remote Sensing, Virginia Tech.

Member, Virginia Water Resources Research Center, Virginia Tech.

INVITED TALKS

[T13] “Physics-guided Machine Learning: Advances in an Emerging Paradigm Combining Scientific Knowledge with Machine Learning for Geoscience Applications,” **Seminar Talk in the Dept. of Atmospheric, Oceanic, and Planetary Physics at the University of Oxford**, Jan 6, 2020.

[T12] “Physics-guided Machine Learning: Advances in an Emerging Paradigm Combining Scientific Knowledge with Machine Learning,” **Invited Talk at the Macromolecules Innovation Institute (MII) workshop**, November 4, 2019.

[T11] “Theory-guided Data Science: Foundations of an Emerging Paradigm Combining Physics and Machine Learning,” **Keynote Talk at DARPA Physics of AI (PAI) Review Meeting**, Ann Arbor, MI, October 2, 2019.

[T10] “Theory-guided Data Science: How Can Machine Learning and Physical Knowledge Come Together to Accelerate Scientific Discovery,” **Invited Talk at Oak Ridge National Lab (ORNL) AI Workshop**, Oak Ridge, TN, September 20, 2019.

[T9] “Physics-guided Data Science: Challenges and Opportunities in Combining Machine Learning with Physical Knowledge in Geosciences,” **Invited Talk at the VT Office of GIS and Remote Sensing (OGIS) Research Symposium**, April 26, 2019.

[T8] Lightning Talk Representing the NSF Expeditions project: “Understanding Climate Change: A Data-driven Approach” at the **NSF Expeditions in Computing PI Meeting: 10 Years of Transforming Science and Society**, Washington D.C., Dec 10, 2018.

[T7] “Theory-guided Data Science: A New Paradigm for Scientific Discovery from Data,” **Invited Talk at UCLA IPAM Workshop on HPC for Computationally and Data-Intensive Problems**, November 9, 2018.

[T6] “Theory-guided Data Science: A New Paradigm for Scientific Discovery from Data,” **Invited Talk at IS-GEO Seminar for Energy Industry (with support from Petrobras) at Texas Advanced Computing Center (TACC)**, September 20, 2018.

[T5] “Theory-guided Data Science: A New Paradigm for Scientific Discovery Combining Physics with Machine Learning,” **Invited CISL Seminar Talk at National Center for Atmospheric Research (NCAR)**, May 3, 2018.

[T4] “Theory-guided Data Science: A New Paradigm for Scientific Discovery from Data,” **Invited Talk at Oak Ridge National Laboratory (ORNL)**, March 6, 2018.

[T3] “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.

[T2] “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.

[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.

PANEL DISCUSSIONS

[PD4] “Data Mining Challenges and Opportunities for Earth Science,” **Panel Discussion at ACM SIGKDD 2019 Earth Day Session**, Anchorage, AK, Aug 5, 2019.

[PD3] “Physics guided machine learning: a new paradigm for modeling dynamical systems, **Panel Discussion at American Geophysical Union (AGU) Annual Meeting**, Washington D.C., Dec 10, 2018.

[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.

PUBLICATIONS

BOOK

[B1] P. Tan, M. Steinbach, **A. Karpatne**, and V. Kumar “Introduction to Data Mining (2nd Ed.),” *Pearson Addison-Wesley*, ISBN-13: 978-0133128901, 2018.

JOURNAL ARTICLES

[J12] P. C. Hanson, A. B. Stillman, X. Jia, **A. Karpatne**, H. A. Dugan, C. C. Carey, J. Stachelek, N. K. Ward, Y. Zhang, J. S. Read, and V. Kumar, “Predicting lake surface water phosphorus dynamics using theory-guided data science,” *International Journal on Ecological Modelling and Systems Ecology* by Elsevier, (Accepted) 2020.

[J11] J. S. Read, X. Jia, J. Willard, A. P. Appling, J. A. Zwart, S. K. Oliver, **A. Karpatne**, G. J. A. Hansen, P. C. Hanson, W. Watkins, M. Steinbach, and V. Kumar, “Process-guided deep learning predictions of lake water temperature,” *Water Resources Research*, 55, 91739190, 2019.

[J10] **A. Karpatne**, I. Ebert-Uphoff, S. Ravela, H. A. Babaie, and V. Kumar, “Machine Learning for the Geosciences: Challenges and Research Opportunities,” *IEEE Transactions on Knowledge and Data Engineering*, 2018.

[J9] *G. Atluri, ***A. Karpatne**, and V. Kumar, “Spatio-temporal Data Mining: A Survey of Data Types, Problems, and Methods,” *ACM Computing Surveys*, 51(4), 83:1–83:41, 2018 (* equal contribution).

[J8] **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.

[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*, 202: 113–128, 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

[C16] A. Daw, R. Q. Thomas, C. C. Carey, J. S. Read, A. P. Appling, and **A. Karpatne**, “Physics-Guided Architecture (PGA) of Neural Networks for Quantifying Uncertainty in Lake Temperature Modeling,” *SDM*, 532–540, 2020.

[C15] N. Muralidhar, J. Bu, Z. Cao, L. He, N. Ramakrishnan, D. Tafti, and **A. Karpatne**, “PhyNet: Physics Guided Neural Networks for Particle Drag Force Prediction in Assembly,” *SDM*, 559–567, 2020.

[C14] X. Jia, M. Wang, A. Khandelwal, **A. Karpatne**, and V. Kumar, “Recurrent generative networks for multiresolution satellite data: An application in cropland monitoring,” *IJCAI*, 2628–2634, 2019.

[C13] X. Jia, J. Willard, **A. Karpatne**, J. Read, J. Zwart, M. Steinbach, and V. Kumar, “Physics Guided RNNs for Modeling Dynamical Systems: A Case Study in Simulating Lake Temperature Profiles,” *SDM*, 558–566, 2019.

[C12] X. Jia, G. Nayak, A. Khandelwal, **A. Karpatne**, and V. Kumar, “Classifying Heterogeneous Sequential Data by Cyclic Domain Adaptation: An Application in Land Cover Detection,” *SDM*, 540–548, 2019.

[C11] X. Jia, S. Li, A. Khandelwal, G. Nayak, **A. Karpatne**, and V. Kumar, “Spatial Context-Aware Networks for Mining Temporal Discriminative Period in Land Cover Detection,” *SDM*, 5113–521, 2019.

[C10] N. Muralidhar, M. Islam, M. Marwah, **A. Karpatne**, and Naren Ramakrishnan, “DANN: Incorporating Prior Domain Knowledge into Model Training,” *IEEE Big Data*, 2018.

[C9] **A. Karpatne**, W. Watkins, J. Read, and V. Kumar, “Physics-guided Neural Networks (PGNN): An Application in Lake Temperature Modeling,” *arXiv: 1710.11431*, 2017.

[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*, 1173–1182, 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

- [W11] **A. Karpatne**, “Mapping Surface Water Globally using Remote Sensing Data: A Physics-guided Data Science Approach,” *International Indian Statistical Association (IISA) Conference session on Analysis of Big Remote Sensing Imagery: Tools and Techniques*, 2019.
- [W10] **A. Karpatne**, “Theory-guided Data Science: A New Paradigm for Scientific Discovery from Data,” *SIAM Conference on Computational Science and Engineering (CSE 19)*, 2019.
- [W9] **A. Karpatne**, “Physics-guided Machine Learning: Opportunities in Combining Physical Knowledge with Data Science for Weather and Climate Sciences,” *American Geophysical Union (AGU) Fall Meeting*, 2018.
- [W9] **A. Karpatne**, “Physics-guided AI: Applications of a New Paradigm combining AI with Physics in Geosciences,” *American Geophysical Union (AGU) Fall Meeting*, 2018.
- [W8] **A. Karpatne**, “How can Physics Inform Deep Learning Methods in Scientific Problems?: Recent Progress and Future Prospects,” *Workshop on Physics Informed Machine Learning*, 2018.
- [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. Steinhäuser, M. Steinbach, S. Chatterjee, and V. Kumar, “Understanding relationships between fire activity and sea surface temperature anomalies”, *AGU Fall Meeting*, 2012.

PATENTS

[P5] A. Khandelwal, **A. Karpatne**, and V. Kumar, “Satellite image classification across multiple resolutions and time using ordering constraint among instances,” 16/103,523, filed February 21, 2019.

[P4] V. Kumar, X. Jia, A. Khandelwal, and **A. Karpatne**, “Discovery of shifting patterns in sequence classification,” US Patent 16/371,274, filed October 3, 2019.

[P3] V. Kumar, X. Jia, A. Khandelwal, and **A. Karpatne**, “Predicting land covers from satellite images using temporal and spatial contexts,” US Patent 16/371,269, filed October 3, 2019.

[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.