Kris Sankaran

Department of Statistics
University of Wisconsin - Madison
1300 University Avenue
Room 7225C
Madison, WI
53706

Email: ksankaran@wisc.edu
Homepage: http://krisrs1128.github.io/LSLab/

Education

Stanford University

PhD Statistics, 2013 - 2018

Thesis: *Discovery and Visualization of Latent Structure with Applications to the Microbiome*

MS Statistics, 2012 - 2013

BS Mathematics, 2009 - 2013

Professional Experience

Discovery Fellow
July 2022 - Present
Wisconsin Institute for Discovery
University of Wisconsin - Madison

Assistant Professor
August 2020 - Present
Department of Statistics
University of Wisconsin - Madison

Postdoctoral Fellow
August 2018 - April 2020
Département d'Informatique et de Recherche Opérationnelle
Université de Montréal

Senior Machine Learning Researcher
April 2018 - August 2018
Obsidian Security

Data Science for Social Good Fellow
Summer 2016
Center for Data Science and Public Policy
University of Chicago

Quantitative Research Intern
Summer 2012
The Climate Corporation

Areas of Expertise

Statistical methods for multi-omics

Interactive data visualization

Probabilistic latent variable modeling

Analysis of earth systems imagery

Lab members are underlined.

Current Preprints

Xu, Tinghui and Kris Sankaran. Interactive Visualization of Spatial Omics Neighborhoods [version 1; peer review: awaiting peer review]. F1000Research 11:799. 2022 https://doi.org/10.12688/f1000research.122113.1.

Xu, Zhuoyan, and Kris Sankaran. Spatial Transcriptomics Dimensionality Reduction using Wavelet Bases. arXiv:2202.00180. Submitted (May 2022).

Peer-Reviewed Publications

Sankaran, Kris. Bootstrap Confidence Regions for Learned Feature Embeddings. Journal of Computational and Graphical Statistics, 2023. https://doi.org/10.1080/10618600.2023.2197478

Sankaran, Kris and Susan P. Holmes. Generative Models: An Interdisciplinary Perspective. Annual Review of Statistics and its Application. Volume 10, Issue 1, 2023. https://doi.org/10.1146/annurev-statistics-033121-110134

Tian, Weiyushi, Anthony Ortiz, Tenzing C. Sherpa, Finu Shrestha, Mir Matin, Rahul Dodhia, Juan M. Lavista Ferres, Kris Sankaran. Estimating Glacial Lake Trends using Historically Guided Segmentation Models. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. Volume 15, 2023. https://doi.org/10.1109/JSTARS.2022.3215722/

Chevrette, Marc, Thomas, Chris, Hurley, Amanda, Rosario-Meléndez, Natalia, Sankaran, Kris, Tu, Yixing, Hall, Austin, Magesh, Shruthi, and Handelsman, Jo. Microbiome composition modulates secondary metabolism in a multispecies bacterial community. Proceedings of the National Academy of Sciences. Volume 119, Issue 42, 2022. https://doi.org/10.1073/pnas.2212930119

Zheng, Minxing, Xinran Miao, and Kris Sankaran. Interactive Visualization and Representation Analysis Applied to Glacier Segmentation. ISPRS International Journal of Geo-Information. 11(8):415. 2022. https://doi.org/10.3390/ijgi11080415

Sankaran, Kris. "Societal Impacts" in *Tackling Climate Change with Machine Learning.* Collection with Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., ... Luccioni, A. ACM Computing Surveys. Volume 55, Issue 2, 2023.

Fukuyama, Julia, Kris Sankaran, Laura Symul, Multiscale analysis of count data through topic alignment, Biostatistics, 2022. https://doi.org/10.1093/biostatistics/kxac018

Cheong, So-Min, Sankaran, Kris, and Bastani, Hamsa. Artificial intelligence for climate change adaptation. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2022; e1459.

Lê Cao, Kim Anh, Aedin Culhane, Elana Fertig, Vincent Carey, Ricard Argelaguet, Susan Holmes, Michael Love, Guo-Cheng Yuan, Al Abadi, Casey Greene, Ayshwarya Subramanian, Amrit Singh, Alexis Coullomb, Vera Pancaldi, Arshi Arora, Joshua Sodicoff, Joshua Welch, Emily Davis-Marcisak, Matthew Ritchie, Yue You, Genevieve Stein-O'Brien, Kris Sankaran, Ruben Dries, Atul Deshpande. Central themes in single cell multiomics identified through community-wide hackathons. Genome biology, 22(1), 1-21, 2021. https://doi.org/10.1186/s13059-021-02468-y

Cheng, Yuwei, Jiannan Zhu, Mengxin Jiang, Jie Fu, Changsong Pang, Peidong Wang, Kris Sankaran, Olawale Onabola, Yimin Liu, Dianbo Liu, Yoshua Bengio. FloW: A Dataset and Benchmark for Floating Waste Detection in Inland Waters. Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021.

Janik, Adrianna, Jonathan Dodd, Georgiana Ifrim, Kris Sankaran, Kathleen Curran. Interpretability of a deep learning model in the application of cardiac MRI segmentation with an ACDC challenge dataset. Accepted to SPIE Medical Imaging Conference (Image Processing Track). February 2021.

Baraka, Shimaa, Benjamin Akera, Bibek Aryal, Tenzing Sherpa, Finu Shresta, Anthony Ortiz, Kris Sankaran, Juan Lavista Ferris, Mir Matin, and Yoshua Bengio. Climate Change AI Workshop at NeurIPS 2020 [Spotlight Talk].

Gonzalez Villareal, Jessenia, Debjani Bhowmick, and Kris Sankaran. Segmentation of water bodies in Peru using knowledge-transfer based convolutional neural networks. Workshop on ML for the Developing World at NeurIPS 2019.

Goytom, Israel, Qin Wang, Tianxiang Yu, Kunjie Dai, Kris Sankaran, Xinfei Zhou, Dongdong Lin. Nanoscale Microscopy Images Colorization Using Neural Networks. Microstructures Analysis. December 2019.

Goytom, Israel and Kris Sankaran. Forecasting Maxima in Climate Time Series. 9th International Workshop on Climate Informatics. October 2019.

Yuan, Tianle, Hua Song, David Hall, Victor Schmidt, Kris Sankaran and Yoshua Bengio. Artificial intelligence based cloud distributor (AI-CD): probing clouds with generative adversarial networks. 9th International Workshop on Climate Informatics. October 2019.

Janik, Adrianna and Kris Sankaran. Discovering Concepts in Learned Representations using Statistical Inference and Interactive Visualization. Workshop on Explainable AI/ML (XAI) for Accountability, Fairness, and Transparency at KDD 2019. https://github.com/adrijanik/adrijanik.github.io/.

Razzak, Muhammed and Kris Sankaran. Interactive Segmentation for Disaster Relief Mapping [Oral presentation]. AI for Good Workshop at IJCAI 2019.

Janik, Adrianna, Kris Sankaran, and Anthony Ortiz. Interpreting Black-Box Semantic Segmentation Models in Remote Sensing Applications. In *Machine Learning Methods in Visualisation for Big Data* edited by Archambault, Daniel and Nabney, Ian and Peltonen, Jaakko (2019). DOI 10.2312/mlvis.20191158.

Benabbou, Loubna, Zouheir Malki, Kris Sankaran, and Hicham Bouzeckri. Machine Learning based Predictive Maintenance for Renewable Energy: The Case of Power Plants in Morocco. Climate Change AI workshop at ICML 2019.

Schmidt, Victor, Alexandra Luccioni, Surya Mukkavilli, Narmada Balasooriya, Kris Sankaran, Jennifer Chayes, and Yoshua Bengio. (2019). Visualizing the consequences of climate change using cycle-consistent adversarial networks. AI for Social Good Workshop at ICLR 2019. arXiv preprint arXiv:1905.03709.

Min, Yan, Xiaoguang Ma, Kris Sankaran, Yuan Ru, Lijin Chen, Mike Baiocchi, and Shankuan Zhu (2019). Sex-specific association between gut microbiome and fat distribution. Nature communications, 10(1), 2408. https://doi.org/10.1145/3485128.

Sankaran, Kris. Remembrances of States Past [Oral presentation]. 2nd Workshop on Visualization for AI Explainability at IEEE VIS 2019.

Sankaran, Kris, and Susan P. Holmes. Multitable Methods for Microbiome Data Integration. Frontiers in Genetics 10, 2019. https://doi.org/10.3389/fgene.2019.00627

Huang, Chin-Wei, Kris Sankaran, Eeshan Dhekane, Alexandre Lacoste, Aaron Courville. Hierarchical Importance Weighted Autoencoders. Proceedings of the 36th International Conference on Machine Learning, PMLR 97:2869-2878, 2019.

Sankaran, Kris, and Susan P. Holmes. Latent Variable Modeling for the Microbiome. Biostatistics, June 3, 2018. https://doi.org/10.1093/biostatistics/kxy018

Sankaran, Kris, and Susan P. Holmes. Interactive Visualization of Hierarchically Structured Data. Journal of Computational and Graphical Statistics, October 18, 2017. https://doi.org/10.1080/10618600.2017.1392866

Gonzalez Villareal, Jessenia, Kris Sankaran, Victor Ayma, and Cesar Beltran. Application of Semantic Segmentation with Few Labels in the Detection of Water Bodies from Perusat-1 Satellite's Images. 2020 IEEE Latin American GRSS & ISPRS Remote Sensing Conference (LAGIRS). IEEE, 2020.

Bastidas, Alexei, Matthew Beale, Yoshua Bengio, Anna Bethke, Pablo Fonseca, Jason Jo, Dale Kunce, Sean McPherson, Vincent Michalski, Anthony Ortiz, Kris Sankaran, and Hanlin Tang. Foundational mapping of Uganda to assist American Red Cross disaster response to floods and pandemics. AI for Social Good Workshop at NeurIPS 2018.

Fukuyama, Julia, Laurie Rumker, Kris Sankaran, Pratheepa Jeganathan, Les Dethlefsen, David A. Relman, and Susan P. Holmes. Multidomain analyses of a longitudinal human microbiome intestinal cleanout perturbation experiment. PLOS Computational Biology 13, no. 8 (2017): e1005706.

Callahan, Ben J., Kris Sankaran, Julia A. Fukuyama, Paul J. McMurdie, and Susan P. Holmes. Bioconductor workflow for microbiome data analysis: from raw reads to community analyses. F1000Research 5 (2016).

Rhee, Soo-Yon, Kris Sankaran, Vici Varghese, Mark Winters, Christopher B. Hurt, Joseph J. Eron, Neil Parkin, Susan P. Holmes, Mark Holodniy, and Robert W. Shafer. HIV-1 Protease, Reverse Transcriptase, and Integrase Variation. Journal of Virology (2016): JVI-00495.

Sankaran, Siva, and Kris Sankaran. Improving Online Course Performance Through Customization: An Empirical Study Using Business Analytics. International Journal of Business Analytics (IJBAN) 3, no. 4 (2016): 1-20.

Sankaran, Kris, and Susan P. Holmes. structSSI: Simultaneous and Selective Inference for Grouped or Hierarchically Structured Data. Journal of Statistical Software 59, no. 13. 2014. https://doi.org/10.18637/jss.v059.i13

Supervision

Postdoctoral Trainees: Margaret Thairu (Joint with Jo Handelsman)

PhD students: Xinran Miao, Zhuoyan Xu (Joint with Yin Li), Hanying Jiang, Shuchen Yan.

MS students: Minxing Zheng, Tinghui Xu, Margaret Turner, Kaiyan Ma.

Undergraduate students: Yixing Tu, Weiyushi Tian, Zhihao Lyu (VISP), Zhuoxin Shi, Harshita Narayanan (WISCERS), Yuliang Peng (Honors Advisor), Zhi Zheng.

Thesis Committee Membership: Siyu Zhang, Emily Diaz Vallejo.

Interns supervised (Mila): Anthony Ortiz, Md. Rifat Arefin, Pablo Fonseca, Muhammed Razzak, Israel Goytom, Adrianna Janik, Yasser Salah Eddine Bouchareb, Sara Ebrahim, Mustafa Alghali, Jessenia Gonzalez Villareal, Shimaa Baraka, Debjani Bhowmick, Joseph Baafi, James Assiene, Benjamin Akera, and Olawale Onabola.

Interns co-supervised (Mila): The biasly AI team, Wisdom d'Almeida.

Talks

Community

Generative Models for Microbiome Mediation Analysis. CMStatistics (12/2022)

Visualization and Simulation in Microbial Community Analysis. Indiana University Bloomington Statistics Department Colloquium (10/2022)

Revisiting Iterative Data Structuration. University of Pittsburgh Department of Statistics Seminar (04/2022)

Multiscale Topic Visualization in the Microbiome. ML + X Talks, UW Madison Data Science Hub (11/2021)

Multiscale Analysis of Count Data through Topic Alignment. McMaster University Mathematics and Statistics Department Seminar (11/2021)

Latent Structure Laboratory. UW Madison Statistics Department Seminar (09/2021)

Discovery and Visualization. UW Madison Statistics Department Lightning Talks (03/2021)

Visualizing Mapping Models. UW Madison Data Science Bazaar (02/2021)

Climate Change AI Team Intro. TEDx Climate Countdown (10/2020)

Invitation to Research. UW Madison Statistics Department Seminar (09/2020)

Measuring Feature Stability. Bernoulli IMS One World Symposium (08/2020)

Formalization of Complex Workflows. Stanford University Data Science Scholars (08/2020)

Navigation and Evaluation in High-Dimensional Data - San Francisco State Mathematics Department Seminar (02/2020)

Navigation and Evaluation in High-Dimensional Data - UC Davis Statistics Department Seminar (02/2020)

Navigation and Evaluation in High-Dimensional Data - UW Madison Statistics Department Seminar (01/2020)

Navigation and Evaluation in High-Dimensional Data - Penn State Statistics Department Seminar (01/2020)

Navigation and Evaluation in High-Dimensional Data - Waterloo Statistics Department Seminar (01/2020)

Latent Structure in the Microbiome - University of Chicago, Department of Statistics Colloquium (01/2018)

Internal

Notes: A Simple Measure of Conditional Independence (09/2022)

Generative Approaches to Microbiome Mediation Analysis (07/2022)

Mediation and the Microbiome (04/2022)

Generative Models and Experimental Design (01/2022)

Recent Multi-Omics Methods and their Ancestors (10/2021)

Multiresolution Analysis of Count Data through Topic Alignment (08/2021)

Software

`mbtransfer`: An R package for modeling microbiome dynamics which provides functions for fitting transfer function models, simulating counterfactual trajectories, and selecting significantly perturbed taxa using mirror statistics.

`alto`: An R package that aligns topics from different LDA models, computes metrics for quantifying the goodness of alignment, and provides visualization functions to explore the alignment and robustness of topics across environments or LDA hyper-parameters.

`NBFvis`: An R package for neighborhood-based featurization and visualization for spatial genomics data.

`waveST`: An R package for wavelet-guided dimensionality reduction for spatial transcriptomics data.

`treelapse`: An R package for interactive visualization of hierarchically structured time series. Suitable for multiscale analysis of longitudinal microbiome data.

`structSSI`: An R package implementing hierarchically structured and grouped multiple hypothesis testing procedures.

`GFLasso`: An R package for the graph-fused lasso.

Funding

Fall Research Competition [$48K, conditional on pending NSF application] for "Differential Interactome Analysis for Multifactorial and Multi-Omics Experiments." December 2022.

AI for Earth Microsoft Azure Compute Grant [$18K] for "Glacier Monitoring for Adaptation to Climate Change." October 2019.

AI for Earth Microsoft Azure Compute Grant [$10K] for "Automated River Plastic Debris Removal." October 2019.

Teaching

Long Courses

STAT 992 (Multi-Omics Data Analysis). University of Wisconsin - Madison. Spring 2023.

STAT 436 (Statistical Data Visualization). University of Wisconsin - Madison. Spring 2023.

STAT 679 (Advanced Statistical Data Visualization). University of Wisconsin - Madison. Fall 2022.

STAT 479 (Statistical Data Visualization). University of Wisconsin - Madison. Spring 2022.

STAT 424 (Statistical Experimental Design). University of Wisconsin - Madison. Fall 2021.

STAT 479 (Statistical Data Visualization). University of Wisconsin - Madison. Spring 2021.

STAT 424 (Statistical Experimental Design). University of Wisconsin - Madison. Fall 2020

IFT6758 (Introduction to Data Science) co-instructor Université de Montréal. Fall 2019.

STATS 390 (Free Statistical Consulting Workshop). Stanford University. Summer 2017

Short Courses

UW Madison MSTP M2 Journal Club. Lecture on Microbiome Analysis.

UW Madison STAT 627. Lecture on Is a PhD right for me? (02/2022)

Third Nepal Winter School in AI instructor. Session on "Foundations of Machine Learning" (12/2021)

Data Science Nigeria AI Bootcamp instructor. Session on "Understanding Machine Learning" (10/2021)

Data Science Nigeria AI Bootcamp instructor. Sessions on "AI for Social Good" and "Statistics and Machine Learning" (11/2021)

McMaster Innovate 1Z03 Lecture on AI, Sustainability, and Climate Change.

Deep Learning Indaba instructor. Session on "Memory and Attention." Nairobi, Kenya. (08/2019)

First Nepal Winter School in AI instructor. Sessions on "Deep Learning Foundations", "Advanced Deep Learning", and "Humanitarian AI." Kathmandu, Nepal (12/2018)

Strategies and Techniques for Analyzing Microbial Population Structures TA Marine Biological Laboratory, Woods Hole. (08/2017)

Other

Madison Teaching and Learning Excellence Fellow - Chi Cohort.

Service

Reviewer for *BMC Bioinformatics*, *PLOS Computational Biology*, *The Journal of Machine Learning Research*, *Annals of Applied Statistics*, *eLife*, *The R Journal*, *Biometrics*, *Biostatistics*, *Frontiers of Computer Science*, *The Journal of the American Statistical Association*, *Microbiome*, *F1000 Research*.

Reviewer for *Practical Machine Learning for Developing Countries* [ICLR 2021], AAAI 2021, *AI for Social Impact* [AAAI 2020], *Climate Change AI* [ICML 2019, NeurIPS 2019, ICLR 2020, NeurIPS 2020, ICML 2021], *Harvard CRCS Workshop on AI for Social Good* 2020, *AI for Social Good* [NeurIPS 2018, ICLR 2019, ICML 2019], *AI for Good* [IJCAI 2019], *ML for the Developing World* [NeurIPS 2019, NeurIPS 2020], *Montreal AI Symposium* [2019]

Judge for UW Madison Statistics Club Data Challenge. October 2021, October 2022.

Judge for UW Madison Data Science Club Datathon. October 2021.

Steering committee member for Climate Change AI. April 2019 - December 2020.

Project scoper for Data Science for Social Good - Solve. July 2020 - October 2020

Writing Mentor, Deep Learning Indaba. 2020.

Graduate School Application Mentor, Black in AI. 2018 - 2019.

Humanitarian AI reading group organizer. 2018 - 2019.

Workshop organizer for *Computer Vision for Global Challenges* [CVPR 2019], *AI for Social Good* [ICML 2019, ICLR 2019], *Tackling Climate Change with Machine Learning* [ICLR 2020].

Red Judge for IBM Watson AI XPRIZE. May - October 2019.

DataKind SF Chapter Core Volunteer. 2016 - 2017,

Stanford Statistics for Social Good. 2013 - 2017.

STATS 390 (Free Statistical Consulting Workshop) participant. 13 quarters, 2012 - 2017

Recognition

Best reviewer award. AI for Social Good Workshop, NeurIPS 2018.

Jerome H. Friedman Applied Statistics Dissertation Award 2018

Ric Weiland Graduate Fellowship, 2016 - 2018

Boothe Prize for Excellence in Writing (Top 6 freshman papers), 2010

Unpublished Reports

Miao, Xinran and Kris Sankaran. Source Data Selection for Out-of-Domain Generalization. arXiv:2202.02155. Submitted (February 2022).

Sankaran, Kris. Measuring the Stability of Learned Features. 2021. arXiv:2102.10388

Ortiz, Anthony, Kris Sankaran, Olac Fuentes, Christopher Kiekintveld, Pascal Vincent, Yoshua Bengio, Doina Precup. Conditional Networks. (Rejected from ICLR 2021, submission available).

Deudon, Michel, Alfredo Kalaitzis, Israel Goytom, Md. Rifat Arefin, Zichao Lin, Kris Sankaran, Vincent Michalski, Samira E. Kahou, Julien Cornebise, and Yoshua Bengio. (2020). Highres-net: Recursive fusion for multi-frame super-resolution of satellite imagery. arXiv preprint arXiv:2002.06460.

Casavant, Ellen, Les Dethlefsen, Kris Sankaran, Dan Sprockett, Susan P. Holmes, David Relman, and Joshua Elias. (2019). Strategies for understanding dynamic, personalized profiles of host-derived proteins and microbes from human stool. bioRxiv, 551143.

Sankaran, Kris, and Susan P. Holmes. Inference of Dynamic Regimes in the Microbiome. arXiv preprint arXiv:1712.00067 (2017).

Sankaran, Kris, Diego Garcia-Olano, Mobin Javed, Maria Fernanda Alcala-Durand, Adolfo De Unánue, Paul van der Boor, Eric Potash, Roberto Sánchez Avalos, Luis Inaki Alberro Encinas, and Rayid Ghani. Applying Machine Learning Methods to Enhance the Distribution of Social Services in Mexico. arXiv preprint arXiv:1709.05551 (2017).

Sankaran, Kris, Suzanne Tamang, and Ami Bhatt. Opioid Atlas: Mapping Access to Pain Medication. arXiv preprint arXiv:1612.00497 (2016).

Hwang, Jessica, and Kris Sankaran. SparkPoint Bundled Services Analysis. San Francisco, CA: United Way Bay Area (2014).

Last updated: July 9, 2023