



ASA GEORGIA CHAPTER

Winter Lecture Series

4:00 PM, Tuesday, December 8, 2020

Zoom link will be posted on <https://www.amstatgeorgia.org/events>

Dr. Bin Yu

Departments of Statistics and EECS, UC Berkeley

Veridical Data Science

Veridical data science extracts reliable and reproducible information from data, with an enriched technical language to communicate and evaluate empirical evidence in the context of human decisions and domain knowledge. Building and expanding on principles of statistics, machine learning, and the sciences, we propose the predictability, computability, and stability (PCS) framework for veridical data science. Our framework is comprised of both a workflow and documentation and aims to provide responsible, reliable, reproducible, and transparent results across the entire data science life cycle. Moreover, we propose the PDR desiderata for interpretable machine learning as part of veridical data science (with PDR standing for predictive accuracy, predictive accuracy and relevancy to a human audience and a particular domain problem). The PCS framework will be illustrated through the development of iterative random forests (iRF) for extracting predictable and stable non-linear interactions in genomics studies. Finally, a general DNN interpretation method based on contextual decomposition (CD) will be discussed with applications to sentiment analysis and cosmological parameter estimation.

About the Speaker

Bin Yu is The Class of 1936 Second Chair in the College of Letters and Science, and Chancellor's Distinguished Professor, Departments of Statistics and of Electrical Engineering & Computer Sciences, University of California at Berkeley and a former chair of Statistics at UC Berkeley. She heads the Yu Group at Berkeley, which consists of 15-20 students and postdocs from Statistics and EECS. She was formally trained as a statistician, but her research interests and achievements extend beyond the realm of statistics. Her work has leveraged new computational developments to solve important scientific problems by combining novel statistical machine learning approaches with the domain expertise of her many collaborators in neuroscience, genomics and precision medicine. She also develops relevant theory to provide insight and guide practice. She is a member of the U.S. National Academy of Sciences and a fellow of the American Academy of Arts and Sciences. She was a Guggenheim Fellow in 2006, and the Tukey Memorial Lecturer of the Bernoulli Society in 2012. She was President of IMS (Institute of Mathematical Statistics) in 2013-2014 and the Rietz Lecturer of IMS in 2016. She received the E. L. Scott Award from COPSS (Committee of Presidents of Statistical Societies) in 2018. She was the Breiman Lecturer at NeurIPS 2019.



For more information, please contact us at:

Phone: 706.542.5232 E-Mail: stat@uga.edu

Parking is available in the North Campus Parking Deck.

For a UGA Campus map, please see: <http://aviary.camplan.uga.edu/CampusMap/Default.aspx>