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 Chancellor’s Professor in the Departments of Statistics and of Electrical Engineering & Computer Sciences at the University of California at Berkeley and a former chair of Statistics at UC Berkeley. Her research focuses on practice, algorithm, and theory of statistical machine learning and causal inference. Her group is engaged in interdisciplinary research with scientists from genomics, neuroscience, and precision medicine. 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. Moreover, Yu was a founding co-director of the Microsoft Research Asia (MSR) Lab at Peking University and is a member of the scientific advisory board at the UK Alan Turing Institute for data science and AI.
History of the Bradley Lecture

The University of Georgia Department of Statistics and the Statistics Club are proud to host the 23rd Annual Bradley Lecture. The event honors former faculty member Dr. Ralph A. Bradley, who was born on November 28, 1923 in Smith Falls, Ontario, Canada, and who grew up in the village of Wellington. After graduating from Queen’s University in 1944 with an honors degree in mathematics and physics, he served in the Canadian Army from 1944 to 1945, and afterwards earned his Masters of Arts degree in 1946. He received his PhD in 1949 at the University of North Carolina Chapel Hill, and went on to a very distinguished career. He was founder of the Department of Statistics at Florida State University and served as its chair from 1959 to 1978. He joined UGA in 1982.

Dr. Ralph Bradley made many contributions to the field of statistics as a researcher in design of experiments, nonparametric statistics, sequential analysis and multivariate analysis. He also had an exemplary record of service to the profession of statistics as a member of ASA, IMS, ISI, as well as by serving as a president of ASA in 1981.

The Bradley Lecture provides an opportunity for UGA graduate students to interact with the speaker, who is normally an eminent statistician of their choice. After the seminar in the afternoon, the speaker gives an after-dinner presentation and often stays for the next day’s spring picnic to mingle with faculty and students.

We hope you’ll join us for what should be an informative and exciting event!

Schedule of Events

Friday, April 24, 2020

3:30pm – 4:00pm
Arrival
Special Collections Library Room 285

4:00pm – 5:00pm
Lecture
Dr. Bin Yu
Special Collections Library Room 285

5:00pm – 6:00pm
Happy Hour
(Cash Bar & Light Refreshments Will Be Served)
Special Collections Library Room 285

6:00pm – 8:00pm
Dinner and After-Dinner Talk
Dr. Bin Yu
Special Collections Library Room 285

Saturday, April 25, 2020

11:00am – 3:00pm
Bradley Spring Picnic
Lake Herrick Pavilion

After-Dinner Talk

Dr. Bin Yu
Special Collections Library Room 285

How to become a data scientist in 30 minutes

Bradley Spring Picnic

Lake Herrick Pavilion

Saturday, April 25, 2020 | 11:00am – 3:00pm

Enjoy a relaxing afternoon full of food, fun and fellowship!

Mingle with the guest speaker, faculty members, and graduate students.