

OPPENHEIM LECTURE 2024

SPEAKER

Professor Bin Yu is CDSS Chancellor's Distinguished Professor in Statistics, EECS, and Computational Biology, and Scientific Advisor at the Simons Institute for the Theory of Computing, all at UC Berkeley. Her research focuses on the practice and theory of statistical machine learning, veridical data

science, and solving interdisciplinary data problems in neuroscience, genomics, and precision medicine. She and her team have developed algorithms such as iterative random forests (iRF), stability-driven NMF, and adaptive wavelet distillation (AWD) from deep learning models. She is a member of the National Academy of Sciences and of the American Academy of Arts and Sciences. She was a **Guggenheim Fellow, IMS** President, and delivered the IMS **Rietz and Wald Lectures and Distinguished Achievement** Award and Lecture (formerly Fisher Lecture) of COPSS. She holds an Honorary Doctorate from The University of Lausanne.

CONTACT

Department of Mathematics Faculty of Science National University of Singapore Block S17 10 Lower Kent Ridge Road Singapore 119076 Tel: +65 6516 2737



Professor Bin Yu Statistics, EECS, Center for Comp. Bio. and Simons Institute, UC Berkeley

Veridical Data Science: Bridging the Gap in Education and Research



Friday 29 November 2024, 2.30 pm to 3.30 pm NUS Department of Mathematics Block S17 Level 4 Seminar Room 1 (S17-04-06) 10 Lower Kent Ridge Road, Singapore 119076

Admission is free, please register <u>https://forms.office.com/r/Cbbfzi4JKU</u>. For more information, please visit math.nus.edu.sg > Events > Oppenheim Lectures

Abstract: The rapid advancement of AI relies heavily on the foundation of data science, yet its education significantly lags its demand in practice. A new book 'Veridical Data Science: The Practice of Responsible Data Analysis and Decision Making' by Yu and Barter was published in Oct, 2024 by the MIT Press (free online at www.vdsbook.com). It tackles this gap by promoting Predictability, Computability, and Stability (PCS) as core principles for trustworthy data insights. PCS for veridical data science (VDS) has been developed in the process of solving scientific data science problems. It thoroughly integrates these principles into the Data Science Life Cycle (DSLC), from problem formulation to data cleansing and to result communication, fostering a new standard for responsible data analysis. This talk explores PCS' motivations and compares the VDS book approaches with traditional ones. I will end the talk with a PCS-guided project on prostate cancer detection.

About: The Oppenheim Lectures is a distinguished lecture series in honour of Sir Alexander Oppenheim, first Head of the Department of Mathematics at the Raffles College (a predecessor of NUS), and a number theorist known for the Oppenheim Conjecture.

Activities held in conjunction with the Oppenheim Lecture

29 November 2024 4:15pm : Conversation with Professor Bin Yu (By invitation only) 02 December 2024 : Workshop on Machine Learning and AI: Theory and Practice (Venue: S17-04-06)

Jointly Organised by



Department of Mathematics Faculty of Science

